INL/CON-11-24299
PREPRINT

Agent-Based Cyber Control Strategy Design for Resilient Control Systems: Concepts, Architecture and Methodologies

5th International Symposium on Resilient Control Systems

Craig G. Rieger
Quanyan Zhu
Tamer Başar

August 2012

The INL is a
U.S. Department of Energy
National Laboratory
operated by
Battelle Energy Alliance

Agent-based Cyber Control Strategy Design for Resilient Control Systems: Concepts, Architecture and Methodologies

Craig Rieger, Quanyan Zhu and Tamer Başar

***Abstract*— The implementation of automated regulatory control has been around since the middle of the last century through analog means. It has allowed engineers to operate the plant more consistently by focusing on overall operations and settings instead of individual monitoring of local instruments (inside and outside of a control room). A similar approach is proposed for cyber security, where current border-protection designs have been inherited from information technology developments that lack consideration of the high-reliability, high consequence nature of industrial control systems. Instead of an independent development, however, an integrated approach is taken to develop a holistic understanding of performance. This performance takes shape inside a multi-agent design, which provides a notional context to model highly decentralized and complex industrial process control systems, the nervous system of critical infrastructure. The resulting strategy will provide a framework for researching solutions to security and unrecognized interdependency concerns with industrial control systems.**

Keywords: resilient control; cyber awareness; human systems; complex networked control systems; data fusion; hierarchical architecture; cyber security, and cyber physical systems.

I. Introduction

The implementation of automated regulatory control has been around since the advent of electro-mechanical analog devices early last century. This advance has reduced the burden on the attending human in its earliest inception, allowing them to operate the plant more consistently by focusing on overall operations and settings instead of local interactions. With the introduction of digital control systems, however, the ability to centralize the display of widely dispersed assets has led to the further minimization of localized readings. While initially further reducing the burden on the human, the ease of adding more sensors leads to an increase in the amount of data but overloads humans with tasks of monitoring the digital control system. Therefore, one could argue that the initial benefits of migrating to digital control have unwittingly increased the burden. In addition, while the ability to incorporate advanced control methods into the formerly analog systems has been cost prohibitive, the digital systems can be configured in software with relative ease. This has led also to further complexity in design [1]–[3], and more for the human operator to comprehend when trying to understand the interactions.

C. Rieger is with Instrumentation Control & Intelligent Systems, Idaho National Laboratory, Idaho Falls, Idaho, USA. E-mail: craig.rieger@inl.gov; Q. Zhu and T. Başar are affiliated with Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1308. W. Main St., Urbana, IL, email: {zhu31, basar1}@illinois.edu.

In similar fashion, the current application of cyber security is a dichotomy in that it possesses a cross-section of the issues expressed for the analog and digital control systems. While a centralized interface is provided, as with a digital system, networking appliances are still dependent upon the human to monitor and make corrections to the individual appliances in response to security events. While individual analysis to recognize and even react to malicious behavior exists, the human is still burdened to further analyze and respond to the local settings. In addition, the monitoring of cyber security using tools such as intrusion detection systems (IDS) has led to the same information overload situation of the digital control system. As a result and in considering resilience [4], [5], a similar need arises for the cyber security of control systems that points to the need for a regulatory design, starting with a mechanism to achieve situational awareness based on several forms of cyber-related performance data. Cyber security can be treated as a disturbance in control theory, given that it can be properly identified with fault detection and diagnosis. However, this paper instead focuses on the modification of the cyber architecture, changing the behavior of cyber assets (e.g., data storage, packet transmission, etc.) when a malicious intrusion is believed to be occurring.

In what follows, perspective on a strategy for cyber control based upon control theory will be made. In Section II, we review related works. In Section III, a discussion of both the active and passive cyber control mechanisms will be given, as paralleled to closed loop and open loop process control, respectively. Section IV evolves this discussion to agent based methodologies that would embed cyber-physical aspects to address control system threats to stability, efficiency and security. Section V provides a perspective on the mechanism and challenge in taking real systems and decomposing them into a multi-agent framework. Section VI summarizes the highlights of the paper.

II. Related Work

The area of resilient control systems (RCS) is arguably a new paradigm that encompasses control design for cyber security, physical security, process efficiency and stability, and process compliance in large-scale, complex systems. In [4], RCS is defined as a control system that maintains state awareness and an accepted level of operational normalcy in response to disturbances, including threats of an unexpected and malicious nature. While one might say that resilient design is primarily dependable computing coupled with fault tolerant control, it has been argued in [4] that dependable

computing views malicious faults as a source of failure, but does not consider the effect of these faults on the underlying physical processes in a large-scale complex system.

Recent literature on RCS has studied many aspects of resilience in control systems. In [6], notional examples are used to discuss fundamental aspects of resilient control systems. It has been pointed out that current research philosophies lack the depth or the focus on the control system application to satisfy many aspects of requirements of resilience, including graceful degradation of hierarchical control while under cyber attack. In [7], a hierarchical viewpoint is used to address security concerns at each level of complex systems. The paper emphasizes a holistic cross-layer philosophy for developing security solutions and provides a game-theoretical approach to model cross-layer security problems in cyber-physical systems. In [5], the authors have proposed a game-theoretic framework to analyze and design, in a quantitative and holistic way, robust and resilient control systems that can be subject to different types of disturbances at different layers of the system. In [8], a hybrid system model is used to address physical layer control design and cyber level security policy making for cyber-physical systems that are subject to cascading effects from cyber attacks and physical disturbances.

Cyber security is an essential component of resilience of control systems. Few works have provided quantitative methods of modeling of device configurations and evaluating trade-offs among defense options. In [12], the authors have made a comprehensive survey on game-theoretic methods for different problems of network security and privacy. It has been pointed out that the quantitive methods discussed in the survey can be integrated with cyber-physical systems for analyzing and design resilient control systems. The literature on device configurations can be found in [9]–[11]. In [10], a cooperative game approach has been used to address the static configuration of security devices, such as intrusion detection systems (IDS) and intrusion prevention systems (IPS), in face of adversarial attacks. In [11], the authors address the dynamic counterpart of the configuration problem. The equilibrium cyber policy can be obtained from a game-theoretic analysis of a dynamic zero-sum Markov game, which has taken into account the tradeoffs between different defense mechanisms. In [9], a network level configuration of security devices has been addressed by considering the interdependence of devices in the network.

III. Cyber resilience strategy

The quantity and diversity of control system's vulnerabilities is related to the systems security. However, we currently have few effective ways of modeling how vulnerability, device configuration, and system attributes affect an adversary; few effective ways of modeling how vulnerability, device configuration, and system attributes affect an adversary currently exist; neither can we easily determine or predict the degree to which the system is immune and resilient to an attack [9]–[11]. Critical infrastructure cyber security design, assessment, and measurement must take into account that some vulnerabilities are inherently less severe than others, that not all devices in the system have the same value to the attacker or the same value to the defender, and that not all vulnerabilities are equally accessible to an adversary [3], [9], [12]. Being able to anticipate the malicious actor's likely attack objectives, strategy, processes, and decisions is clearly valuable. Measures and models to simulate and predict these elements, coupled with methods to evaluate the trade-offs among defense options, would enable organizations to improve their security resource allocations and to balance security with other needs and constraints in the critical infrastructure.

With measures and models of system behavior, a basis is provided for achieving some semblance of state awareness. Given this awareness, some level of control (defensive) action can be taken to minimize the access the malicious actor has to assets, if one is suspected. Unlike the parallel to control theory, where the object is to maintain a physical variable at a set point, cyber control is intended to define something more fundamental than just adjusting defensive posture. That is, cyber control should further clarify whether an intrusion truly exists, what the proficiency of the malicious actor is, and what assets the intruder is after. The mechanisms by which the cyber control system can achieve this include a combination of active and passive mechanisms, with the former paralleling a closed feedback loop and the latter an open loop in control theory. Specific applications of techniques will be presented in the sections that follow for illustration.

A. Active Cyber Control

By its nature, the primary reason for having a control system is to operate in a stable way a facility, whether an oil refinery, chemical plant, or electric transmission system. The controller design, based on control theory, provides changes to the control elements to regulate the process based upon feedback on the state. The control elements may be valves, switches, or any number of devices. In looking at passive control of cyber security, similar processes can be conceptually envisioned. Even within current communications security technology, such as intrusion detection and prevention devices (IDS/IPS), certain characteristics or threat signatures are recognized, and in the case of IPS, reacted to by restricting traffic. However, the approach taken in IDS/IPS design has several limitations. First these systems look at historical patterns, whether signature-based or anomaly-based, which are not necessarily predictive of what may be seen in the future from an intelligent adversary attack. Second, these detection methods are not foolproof, and invariably require unfortunate tradeoffs between false positives and false negatives. Last, applying restrictions on traffic flow as the result of detected threat may limit functionality of a control systems communications for no valid reason, and may even be used by an attacker in a denial of service attack. Even if there is good reason to restrict the traffic, it is not possible to be comprehensive while preventing false positives.

Following the same parallel to control theory, an active cyber security feedback loop would involve a mechanism of representative and reproducible measurement, mathematically based methods to model the information streams in the system, and associated attributes for reconfiguration [9], [11]. Each of these three elements comes with its own challenges in finding a verifiable solution, but also does not imply the solution itself is as mathematically rigorous as those achieved with control theory. As physics based models do not normally apply to cyber prediction, except where affects on process data are delineated between physical and cyber exploitation, stochastic techniques and intelligent system predictions can provide solutions that form the basis state awareness of cyber security [9]–[13].

Considering attacker compromise could occur at the sensors, control devices, or at the industrial control system itself, a cyber control action is needed that addresses all three. That is, state awareness of the corruption or usurpation of data at any point in the active cyber control loop is necessary, as suggested by Fig. 1. Given this awareness, control action can be taken to better understand the threat and correct the effects. In the subsections that follow, the state awareness and cyber feedback control aspects of the active control loop are discussed.

1) Cyber State Awareness: Often there will not be a cyber security measurement which reliably, or provably, indicates the current security of the system with complete accuracy. A direct measurement of system state is often unavailable in control theory, and when this occurs, an observer or state estimator is often used. The state estimator is based on a model of the system and uses a combination of existing data and the model to estimate the current state and determine the desired state. If this concept is applied to cyber security design, a model of the adversarial behavior or threat would be needed. However, the normal first principles models used in control theory would likely not apply in this case because of their divergence from a physics-based design. Current cyber security models can become complex quickly and, in general, are not very predictive and do not necessarily provide a framework for optimizing a control response. Consequently, new security ideas and models (e.g., modified attack graphs) are needed to aid estimation of a control systems security state, anticipate expected adversarial actions, and define appropriate system responses. The mathematical framework for describing adversarial behavior will need to accommodate different thought processes, as well as methods of optimization. These models may find some basis in the biological sciences that are being explored for concepts that might lead to next generation of networks.

2) Cyber Feedback Control: Given a measurement value, whether a direct reading or a modeled observer, various control system designs and security mechanisms can be envisioned to protect a network. Some of these are traffic shaping. They may prevent effective attacker planning and actions by blocking or obfuscating messages. Another example includes some form of unpredictable routing of traffic, not only randomizing what an attacker might see, but also leading to a greater chance of detection. A final example would be change the type or level of encryption, or the selection of security and process sensors, providing for more variation when an increased risk is expected or measured. Multiple methods of control can be anticipated, but suggested research would target mechanisms for modifying network traffic and information streams, and algorithms and models for process sensor selection to aid identification and isolation of attacks. These designs and mechanisms would provide frameworks for design rather than the ultimate designs, and be developed into implementations only as proof of principle.

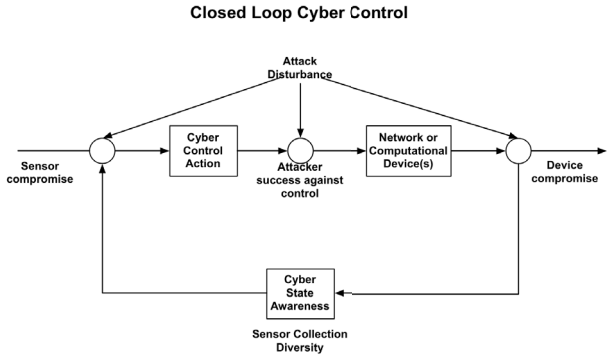

Fig. 1. Active Cyber Control

B. Passive Cyber Control

Unlike active security, passive security includes those items that, by their nature, provide protection to cyber attack without a feedback mechanism. While both active and passive methods are needed for a robust cyber security design, the passive methods are preferred as they do not necessarily require a measurement to be viable. This is especially important in cyber security, where cyber health is ill-characterized, and security measurement poorly understood. Some passive methods of cyber security are available using current technology. Two examples are limiting reachability to software services, and running processes at least required privilege. For both of these examples, the security mechanism and methodology are effectively static in design and not dependent on measurement. One significant weakness in these methods, however, is the dependence they still have on the security of exposed software and system components. Unfortunately, software and system devices have vulnerabilities and, despite best efforts, will continue to be unpredictably vulnerable no matter the level of individual component security assessments.

In considering improvements to the passive protections of next-generation control system networks, it might be suggested that a paradigm shift occur to current methods. Rather than attempting to make control system software more resilient through mitigating identified vulnerabilities with current methods, we should look for low cost trusted mechanisms and methods. These new methods will allow the design and implementation of more secure systems in the

presence of known and unknown software and device vulnerabilities. Unlike the active feedback control loop, corrections will occur without the benefit of state awareness. That is, the nature of the system itself will change to prevent malicious compromise, obfuscating attack pathways and diminishing attacker understanding. Atypical cyber designs that break from traditional, intuitive methods will be required to achieve this. Two resulting passive cyber control strategies (See Fig. 2) to prevent corruption or usurpation of data at any point will be discussed in the next two subsections.

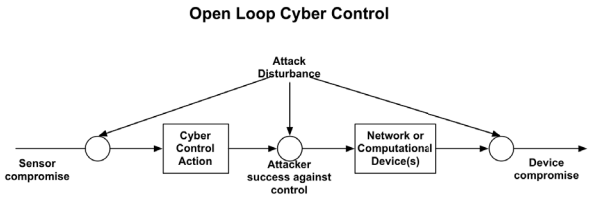

Fig. 2. Passive Cyber Control

1) System Security in the Presence of Insecure Components: There is little empirical evidence to verify whether the security of deployed software improves with the number of vulnerabilities discovered and patched. We know neither the level nor the rate of improvement, neither do we know the conditions under which it happens most quickly. It is sensible to assume that every software component in a control system has vulnerabilities, some of which have not been publicly announced or patched. Further, insecure configurations, including of security devices, may represent system vulnerabilities. Identifying and maintaining secure configurations of components in a changing operational and threat environment is problematic. It is difficult to anticipate or measure the potential security weaknesses in existing or evolving software or induced in the system by insecure configurations. Consequently, there is a need to design in control system security that assures some predictable extent of security even in the presence of vulnerable components. A number of research approaches are available including specification of new security devices (e.g., hardware enforced one way communication) and associated development of techniques (e.g. secure sensor design) to optimize system security design.

2) Randomization of System Attributes: Randomization and automated diversification of various software and system attributes while maintaining control system performance and functionality is another approach for gaining improved control system security. To understand why this might be an effective security mechanism, consider the fact that a malicious actor is often looking for feedback to determine his next move. The attacker is looking for patterns in the control system communications or attributes that can be best used to compare with prior experience or other forms of reference. Considering current research, one randomization technique for improved cyber security against attacks varies library load locations at operating system boot time. A second technique involves automated code rewriting to make exploitation of software vulnerabilities less predictable and, consequently, more difficult for the attacker. New and game changing system randomization techniques are needed to fundamentally dismantle the attacker's inherent advantages of predictably successful system exploration and exploitation. The suggested approach is to apply and vet in the control system environment previously defined techniques for randomization, and to create new randomization concepts, techniques, and tools for aiding control system security [14]. The intent of each of these research activities will be to develop techniques and tools to place attackers in non-deterministic attack states such that their success is highly dependent on the unknown states of particular system attributes. In addition, research will be performed that allow system designers to methodically design in randomization to allow for predictable probabilities of attackers detection and attack failure.

IV. Agent-based methodologies

A. Overview and Integration of Cyber Security

Mechanisms for implementation of cyber intelligence in agent-based designs is important to assure that the interactions are not compromised, causing potentially immeasurable harm to the affected process systems. This agent possesses a mechanism to adjust, within its sphere of influence, to changes within its environment. These changes can include the conditions of the control system components or those outside of the control system, but affect the ability of the control system to fulfill its performance objectives. These performance objectives can be articulated as stability, efficiency, and security (SES) [6]. Stability establishes those characteristics of a control system that assure the system is maintained within the bounds of safe and normal operation. Efficiency provides a term to collect performance characteristics that impact the environmental impact economics of the operation. That is, a control system objective is to minimize waste, maximize product, and minimize the amount of resources consumed to achieve these. And security addresses those aspects of physical and cyber security that, while not recognized as a traditional goal of control system performance, if not addressed, can lead to the undermining of both stability and efficiency objectives.

To understand how security, specifically cyber security, is an important performance parameter for an agent system, an example is required on control system designs, where the dynamics of interchange between one agent and another are already implied. That is, execution (device) layer elements are associated with unit operations, substations, or optimally stabilizable entity. This can be seen from looking at chemical plants, where a collection of separate operations make up an integral unit operation. The unit operation, in this case, defines an area of local optimization. Within the operation, many state and input variables may exist. In a plant made up of many unit operations, the process of determining the stabilizable entities normally results in a minimization of the interactions between individual operations (see Fig. 3). That is, normally only a few state variables will make up the

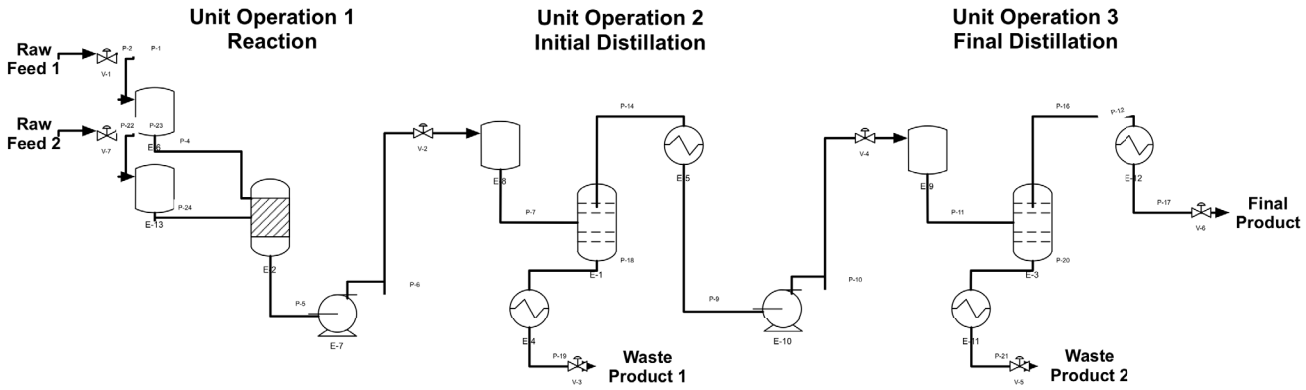

Fig. 3. Series of Unit Operations

interactions between unit operations. For example, the fluid flow of product from one unit operation to the other must remain within a specified range, as the downstream operation is designed to be stabilized for operation within that range. If stability is not achieved, continued plant operation and safety can be impacted.

As cyber attack can affect a control system much like a disturbance, a malicious attack can affect the dynamics and lead to instabilities as with an industrial process control systems disturbance. Therefore, an integrated mechanism is required not only to distinguish that a fault exists, but also the type of fault to ensure an appropriate control action is taken. From a strictly control action standpoint, a recognized cyber disturbance can be corrected by several means, providing one layer of protection at the control loop level. For a sensor compromise, this could include passive cyber-related actions that include methods to recognize and select a known good sensor (or sensor and model), or adjusting the sensory input for the disturbance. However, for an active response, this might include a cyber action that might cut off a communication channel or vary system attributes to attempt a correction that thwarts an attack. Therefore, the attributes of an agent must include a mechanism of anomaly detection for events that affect SES performance indices tied specifically to a corrective action. These actions include both feedback control actions on the associated industrial process, as well as other system actions that are specific to human and malicious aspects that are not well modeled by traditional means.

B. Formation of a Cyber Physical Agent

As information on both the physical and cyber aspects of an industrial process control system can provide synergistic benefit in recognizing faults, a research framework should be established that integrates these perspectives in an interdisciplinary fashion. Consider Fig. 4, which simply depicts integrated selection of sensors based upon cyber-physical threat, decision making to characterize holistically the cyber-physical attributes of the critical infrastructure, and finally independent, but orchestrated regulatory adjustment of the cyber and physical configuration. As the benefit of integrating this data is recognition of impending compromise or degradation of cyber-physical systems, an agent design would need to incorporate a mechanism for shared detection. However, once recognized, the cyber physical control decisions and affected assets controlled are independent.

Considering the cyber threat aspects depicted within Fig. 4, the notional cyber control loops given in Fig. 1 and Fig. 2 might be coupled with state awareness and fusion aspects to form a cyber-physical agent. This agent would perform a sensor selection based upon health and then fuse the relevant information with physical data for the benefit of state awareness and control. The next two subsections will overview a strategy for this solution space.

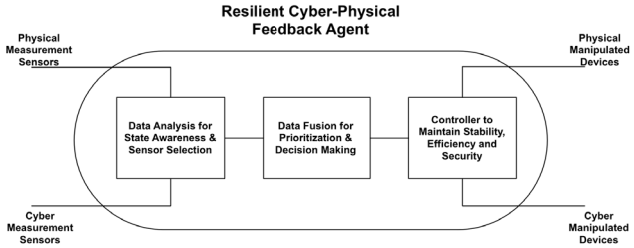

Fig. 4. Cyber-physical Feedback Integration

1) Resilient Cyber Health Mechanism: Recognition and response of the physical sensors to malicious attack is a first layer of protection to a resilient control system. The cyber health of select sensors provides a basis to normalize the data relative to malicious attack in way that is actionable to assure continued awareness [15], [16]. This is in stark contrast to cyber detection mechanisms like signature or anomaly-based IDS, which can give some indication of network and host intrusion, but not to the type of data being compromised. What might be a reasonable approach for business information systems, however, is not well suited to resilient control systems. In developing mechanisms to provide a first layer of protection from malicious intrusion, known secure sensor measurement (KSSM) algorithms, such as those provided in Fig. 5, instead provide a framework for baselining performance and reconfiguring sensors [17]. The KSSM mechanism embeds a method of characterizing a level of confidence in a sensor selection, such as with

encryption, and using this as a basis to contrast network heuristics (message corruption, packet loss, etc) with other sensors for a given network model. While multiple methods can be used for sensor selection, the resulting approach will provide actionable information regarding cyber health of the sensor system. As we will see in the next subsection, the resulting information can then be used to determine the data set that will be used to characterize the state awareness picture.

2) Fusion of Cyber-Physical Data: A necessary attribute of fusion is characterization, prioritization and presentation of the actionable information for a mixed human and automation response. The need to integrate different performance indicators is based upon the ability to ascertain the believability, and the relevance of the data, mechanisms to temporally and spatially orientate the information, and a context in which to normalize the physical and cyber data. This part of the approach can be split into two aspects, one that characterizes the raw cyber physical data to provide a prioritization of response based upon policy, and second, a visualization that integrates the presentation of actionable information for blended tactical response by human operators and the cyber-physical control actions. The raw data fusion aspect of this effort would be to use data-driven, computational intelligence algorithms that have been designed to recognize anomalies that may exist within the cyber or physical realms. In contrast to what was performed in the previous subsection, the anomaly detection design here is to characterize actionable information for control response to cyber and physical devices and not sensor selection [18], [19]. Fig. 6 provides a representation of this mechanism, and highlights the human interaction perspective, considering both the human involvement in establishing the policy governing control system operation and that used for real time tactical decision making. The visualization aspect of this effort is to demonstrate the presentation of physical data with cyber, or threat-related data. The resulting presentation of information would give a full situation awareness that does not overload the human, but provides a target of response.

V. Decomposition of Industrial Process Systems to Multi-Agent Hierarchies

Graph theory provides a technique to interleaf cyber and physical assets of an industrial process control system that have already been decomposed into nodes and edges, such as with an agent-based design [20]. When integrated with SES performance indices, dynamic interactions and the effects on cyber-physical systems can be codified in hierarchical multi-agent dynamical system (HMADS) model. However, the methods of decomposing the critical infrastructure system become the first challenge in establishing a HMADS. And while the prior discussion in this paper has indicated mechanisms for designing an agent, the overall framework for the HMADS hierarch must still be formed. This framework can notionally be based upon three layers, including management, coordination and execution. In the subsections that follow, an overview of this framework is discussed, providing cyber-physical dynamics considerations modeled within each layer.

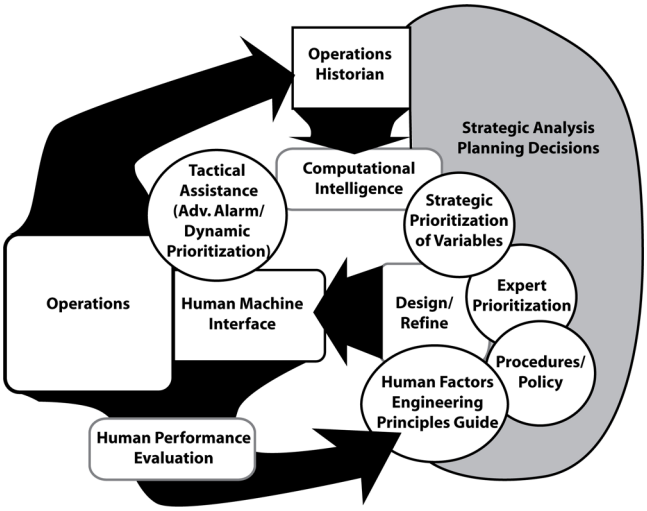

Fig. 6. Fusion Mechanism [18]

A. Decomposing Philosophy to a Multi-agents Hierarchy

While multiple layers can be imagined, for the purposes of illustration three are suggested as suitable to identify distinct and separate functionality [3], [7], [21], defined as follows:

- Upper Layer–Management: This layer provides the overall philosophical goals and priorities for operation. The sources for this design reign from management, regulators, physical constraints of the system, etc.
- Coordination Layer–Coordination: Coordination provides negotiation and potential realignment of resources that best enable meeting the dictated philosophy. Based on renegotiation, the tasking of the execution layer is driven.
- Lowest Layer–Execution: This layer provides direct monitoring of sensors and control of field devices.

There are several factors that influence the philosophies that govern how SES performance goals are set, and these factors require consideration when establishing the policy of the management layer. However, understanding what these factors are and developing a method to decompose them down to constraints of operation are important. In some cases, the constraints are cut and dried. In others, however, an interpretation must be made by those in authority. Below is a list of some factors that should be considered with control system decomposition:

- Regulatory Requirements: Considering primarily governmental agencies that regulate the operation or its products in some fashion [22].
- Desired Performance: Whether a production rate or an efficiency objective, this aspect comes from a desire to maximize profit for the organization using the control system.
- Physics-based Limitations: The physics of the design affect the limits of the operation. While this might

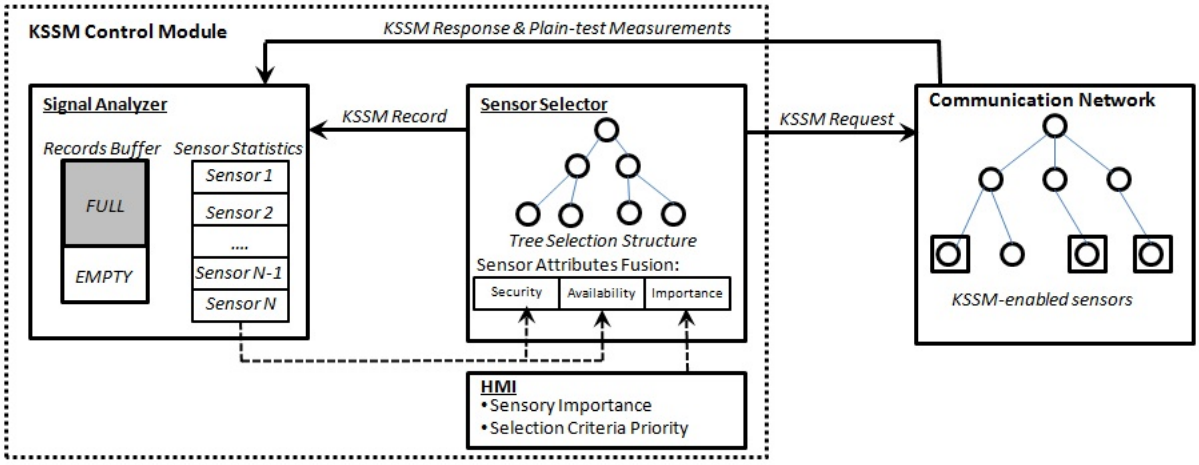

Fig. 5. KSSM Resilient Cyber Health Mechanism [17]

seem obvious, one must collectively include this when considering the tradeoffs of performance [5], [11].

Establishment of a coordination layer requires a mechanism to connect management policy to execution dynamics, establishing a resilience buffer to maintain operational normalcy. In its simplest form, this connection can be considered set points, or even a mathematical relationship that allows flexibility in operation, but also constrains execution to a given set of dynamics. Unlike traditional concepts of set points established to prevent violation of operational limits, however, here the discussion is the development of overall resilience buffers by the nature of the design to achieve optimal performance and prevent loss of critical operation. Below are some of the aspects that can be considered in the coordination layer decomposition:

- The dynamics of the system requires tracking of the optimum path or trajectory to achieve optimum SES, in addition to achieving local SES. Stated another way, the performance of the system remains within its constraints for operation. This implies a performance goal that considers path and endpoints.
- The ability to share, and ultimately negotiate resources is limited by the uniformity of the system. For example, an unmanned air vehicle squadron provides a highly uniform implementation of a HMADS, and therefore, a higher level of resource sharing is theoretically possible within the constraints of the design. The level of uniformity is defined, in this case, as the ability of an agent to provide a necessary functionality in the fulfillment of a need.
- Decisions for shifting of resources can occur at different layers of the HMADS hierarchy, with control action taken at both the middle and lower layers. However, the goal of negotiation is the same regardless of the level, which is to adjust resources to reach optimum performance. The difference lies in the sphere of influence. That is, the coordination layer has responsibility for multiple lower level agents, and as such, will better orchestrate shifts in operation to accommodate the performance goals of the management layer.

B. Stabilizing Hierarchies to Achieve Philosophical Goals

Within the execution layer, the responsibility to operate based upon direction lies. Whereas in a traditional system these directions come from procedures, orders, and judgment, a HMADS directly ties policy to execution. Considering the orchestration of individual agents, methods to achieve some overarching goal have been researched for the last two decades within the mobile robotics community [23].

- Decomposition to minimize, and as a result, simplify agent interactions and complex dynamics
- Development of agent layers in a hierarchical structure
- Optimization of inter-agent interactions with consensus theory to achieve a common objective
- Optimization of intra-agent interactions with applicable control engineering, soft and hard computing, defined by most relevant to situation. The ultimate goal is to stabilize the shared manipulated and controlled variables.

C. Cyber Control Contributions

The HMADS dynamics are multi-factor, and include not only industrial process dynamics, but also human behaviors, both benign and malicious. In developing a complete model of the dynamics, the cyber control system must fulfill its objectives while retaining the SES performance of the industrial process control system. These include methods to integrate network performance requirements, but unlike traditional heuristic designs, these performance parameters will be based upon limits established by the control system design. In addition, some performance indicators, such as latency, will be an overall system parameter that both the cyber and industrial process control elements must achieve within the context of the HMADS.

VI. Conclusions

A perspective for cyber security research can be taken from control theory, and in doing so, an integrated approach will be taken to resilience for industrial process control systems. As the nervous system for critical infrastructure, these

systems to date have depended upon off-the-shelf solutions that have been designed and are suitable for a business system environment. As a first step to developing cyber security research that is targeted to the industrial control system environment, it seems appropriate to take a page from how control theory has been applied to industrial processes. These mechanisms include both open loop and closed loop, or feedback, designs. When integrated as a cyber-physical design, the ability to utilize cyber data for corrective response on the physical system, as well as uncharacterized physical disturbances to correlate cyber exploit, allow for a holistic approach never before possible. In considering research to integrate technologies that address this perspective, a new approach to distributed industrial process control systems is required, which considers both the industrial process control dynamics for SES, as well as the influences of the benign and malicious human. The paradigm of a HMADS offers this notional opportunity.

The resulting HMADS, while not directly replacing humans, is in fact aligning their environment to achieve the desired behaviors. What is currently provided in the form of procedures and policies for benign interactions, resulting from intercommunications of teams and management decisions, are now codified within the design of the HMADS framework. Benign decisions are still made, but occur as the result of interacting with the control system. As a result, a historic understanding of desired interactions with the control system is also developed, and a baseline to recognize malicious behavior. Ultimately, the benefit is a framework for achieving a level of global optimality across multiple facilities and industrial processes, while implementing mechanisms to understand cyber-physical degradation and human performance.

ACKNOWLEDGEMENT

The work of the first author is supported by the U.S. Department of Energy under DOE Idaho Operations Office Contract DE-AC07-05ID14517, performed as part of the Instrumentation, Control, and Intelligent Systems (ICIS) Distinctive Signature of Idaho National Laboratory.

The work of the authors from University of Illinois is partially supported by the AFOSR MURI Grant FA9550-10-1-0573, and also by an NSA Grant through the Information Trust Institute at the University of Illinois.

The authors would like to thank Timothy McJunkin, Miles McQueen from Idaho National Laboratory, and Ondrej Linda, Milos Manic, from University of Idaho, for their comments.

REFERENCES

[1] F. Wang and D. Liu, *Networked Control Systems: Theory and Applications*, Springer-Verlag, London, 2008.

[2] K. J. Aström, P. Albertos, M. Blanke, A. Isidori, W. Schaufelberger and R. Sanz (Eds.) *Control of Complex Systems*, 1st Ed., Springer, 2001.

[3] Q. Zhu and T. Başar, "A hierarchical security architecture for smart grid," In Z. Han, E. Hossain and V. Poor (Eds.), *Smart Grid Communications and Networking*, Cambridge University Press, 2012.

[4] C. G. Rieger, D. I. Gertman, and M. A. McQueen, "Resilient control systems: Next generation design research," 2nd Conference on Human System Interactions, Catania, Italy, pp. 632 636, May 2009.

[5] Q. Zhu and T. Başar, "Robust and resilient control design for cyber-physical systems with an application to power systems," in Proc. of 50th IEEE Conference on Decision and Control and European Control Conference (CDC/ECC), Orlando, Florida, Dec. 12 - 15, 2011.

[6] C. G. Rieger, "Notional examples and benchmark aspects of a resilient control system," 3rd International Symposium on Resilient Control Systems, August, 2010.

[7] Q. Zhu, C. Rieger and T. Başar, "A hierarchical security architecture for cyber-physical systems," in Proc. of the 4th Intl. Symposium on Resilient Control Systems (ISRCS), Boise, ID, Aug. 9 - 11, 2011.

[8] Q. Zhu and T. Başar, "A dynamic game-theoretic approach to resilient control system design for cascading failures," in Proc. of International Conference on High Confidence Networked Systems (HiCoNS) at CPSWeek 2012, in Beijing, China.

[9] Q. Zhu, H. Tembine and T. Başar, "Network security configuration: a nonzero-sum stochastic game approach," in IEEE Proc. of 2010 American Control Conference (ACC), Baltimore, MD, 2010.

[10] Q. Zhu and T. Başar, "Indices of power in optimal IDS default configuration: theory and examples," in Proc. of 2nd Conference on Decision and Game Theory (GameSec 2011), College Park, MD, USA. Nov. 14 - 15, 2011.

[11] Q. Zhu and T. Başar, "Dynamic policy-based IDS configuration," in Proc. of 48th IEEE Conference on Decision and Control (CDC), Shanghai, China, Dec. 2009.

[12] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Başar, J.-P. Hubaux, "Game theory meets network security and privacy," Accepted and to appear in ACM Survey, 2012.

[13] S. B. Shah, K. M. Moudgalya, K. Ramamritham,"Feedback control of Internet applications involving the tracking of dynamic data," IFAC 17th World Congress, pp. 12413-12418, July, 2008.

[14] H. G. Goldman, "Building secure, resilient architectures for cyber mission assurance," MITRE, 2010.

[15] A. Giani, M. McQueen, E. Bitar, P. Khargonekar, K. Poolla, "Smart grid data integrity attacks: Characterizations and countermeasures", SmartGridComm 2011, October, 2011.

[16] A. Giani, M. McQueen, E. Bitar, P. Khargonekar, K. Poolla, "Known secure sensor measurements for critical infrastructure systems: Detecting falsification of system state", 3rd International Workshop on Software Engineering for Resilient Systems, Geneva, Switzerland, Sept 2011.

[17] O. Linda, M. Manic, M. McQueen, Improving Control System Cyber-State Awareness using Known Secure Sensor Measurements, in 7th International Conference on Critical Information Infrastructure Security, in review, 2012.

[18] O. Linda, M. Manic and T. McJunkin, "Anomaly detection for resilient control systems using fuzzy-neural data fusion engine", ISRCS 2011, Boise, ID, August, 2011.

[19] R. Boring et al, "Concept of operations for data fusion visualization," ESREL 2011, Sept 2011.

[20] W. Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," 2005 American Control Conference, pp. 1859-1864, June, 2005.

[21] C. Rehtanz, *Autonomous Systems and Intelligent Agents in Power System Control and Operation*, Springer-Verlag, Berlin, Germany, 2003.

[22] Q. Zhu, M. McQueen, C. Rieger and T. Başar, "Management of control system information security: control system patch management," in Proc. of Workshop on the Foundations of Dependable and Secure Cyber-Physical Systems (FDSCPS-11), CPSWeek 2011, Chicago.

[23] W. Ren and R.W. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control: Theory and Applications*, Springer-Verlag, 2008.