Ensuring Security and Availability through Model-based Cross-Layer Adaptation

Minyoung Kim, Mark-Oliver Stehr, Ashish Gehani, and Carolyn Talcott

SRI International
mkim,stehr,gehani,clt@csl.sri.com

Abstract. Situation- and resource-aware security is essential for the process control systems, composed of networked entities with sensors and actuators, that monitor and control the national critical infrastructure. However, security cannot be addressed at a single layer because of the inherent dependencies and tradeoffs among crosscutting concerns. Techniques applied at one layer to improve security affect security, timing, and power consumption at other layers. This paper argues for an integrated treatment of security across multiple layers of abstraction (application, middleware, operating system including network stack, and hardware). An important step in realizing this integrated treatment of situation- and resource-aware security is first understanding the cross-layer interactions between security policies and then exploiting these interactions to design efficient adaptation strategies (i) to balance security, quality of service, and energy needs, and (ii) to maximize system availability. We propose a novel approach that employs a compositional method within an iterative tuning framework based on lightweight formal methods with dynamic adaptation.

1 Introduction

Physical infrastructure availability relies on the process control systems that can gather, handle, and share real-time data on critical processes from and to networked entities. For example, wireless sensor networks now are being applied in industrial automation to lower system and infrastructure costs, improve process safety, and guarantee regulatory compliance [1]. Harsh environments such as remote areas with potential toxic contamination where mobile ad hoc networks can be the only viable means for communication and information access often necessitate the use of mobile nodes (e.g., surveillance robots with camera and position-changing capability). Optimized control based on continuous observation is an integral part because availability is becoming a fundamental concern in reducing the vulnerability of such systems.

To concretize our approach, we illustrate ideas by using the following motivating scenario. Consider a surveillance system, consisting of a collection of sensors deployed at fixed locations together with mobile nodes, that monitors critical national infrastructure by distributed sensing and actuating. Because of possible jamming attacks and the mobility of nodes, the wireless sensors and mobile nodes need to communicate via opportunistic links that enable the sharing and evaluation of data such as video streams in the presence of unstable connectivity. The challenge here is enabling networked entities to respond to dynamic situations in an informed, timely, and collaborative manner so that the physical infrastructure can safely recover after a cyber-disruption.

The operating scenarios are highly networked, and involve interactions among multiple abstraction layers (application, middleware, OS, hardware) in a distributed real-time environment. Typical wireless sensors and mobile units are limited in communication range, processing power, bandwidth, and residual energy. Often, an emergency situation generates a large volume of communication that must be carefully controlled. Clearly, in such a scenario, the dual goals of ensuring security (with respect to data integrity, confidentiality, authentication, and infrastructure protection) and optimizing resource utilization present a significant challenge. In this paper we focus on integrity and confidentiality for group communication, but we believe that a similar cross-layer treatment of authentication and infrastructure protection would be equally important and possible within the same conceptual framework, e.g., by adopting a situation- and resource-aware posture for authentication and against denial-of-service attacks. Research is needed to develop situation- and resource-aware security solutions that investigate the security implications of existing strategies and integrate them across multiple layers.

We propose the idea of automated verification and configuration of situation- and resource-aware cross-layer security. While existing work has shown the effectiveness of cross-layer adaptation [2], many of these efforts try to address the average-case behavior for energy reduction without verifiable guarantees on their solutions. Our recent work, xTune [3], attempts to provide a comprehensive design methodology, based on formal reasoning that can provide an effective basis for tuning mobile embedded systems under a multitude of constraints. These studies successfully addressed system adaptation and explored the tradeoff with performance, energy, quality of service (QoS), and timeliness for mobile multimedia applications. However, security issues across system abstraction layers in a situation- and resource-aware manner with multidimensional objectives have not been considered until now.

Security goals at each layer can be counterproductive and even harmful. Consider, for instance, the need to protect critical sensor nodes from detection and subsequent subversion by avoiding or reducing their transmissions versus the need of authorized parties to remotely access information. A secure group communication system enables the sharing and evaluation of sensor data. The need for beaconing [1] and rekeying[2] in group membership protocols (at the middleware layer) is in direct conflict with the objective of preventing detection (at the hardware layer). Furthermore, the implementation of security goals is constrained by the available resources. Various solutions ranging from event-

[1] In wireless communication, beaconing refers to the continuous transmission of small packets that advertise the sender's presence.

[2] In cryptography, rekeying refers to the process of changing the encryption key of an ongoing communication to limit the amount of data encrypted with the same key.

driven or on-demand power cycling to reduce transmission power are possible, but the security effects cannot be understood at a single layer. For example, reduced transmission power influences routing and hence requires more reliance on potentially less trustworthy intermediate nodes. In the extreme case, even acknowledgments may have to be avoided due to more opportunities for sniffing and packet loss may need to be compensated for by higher redundancy from the sender (e.g., forward error correction) at the link/network layer. This is why security should be viewed as a multidimensional cross-layer objective for which reasonable tradeoffs must be found in a situation- and resource-aware manner.

This approach posits that cross-layer security opens a large space of feasible solutions exhibiting a range of power, performance, and cost attributes, enabling system designers to optimize and trade off between security, QoS, and resource utilization in response to the operating conditions (i.e., situation and resources). A unified framework is needed to derive, analyze, and validate cross-layer policies and parameters while proving various properties pertaining to security, energy usage, delays, bandwidth, storage, and processing, as the system evolves over time. xTune [3] has demonstrated the feasibility of applying lightweight formal methods to cross-layer adaptation for mobile multimedia with QoS constraints. In this paper, we extend xTune to cope with security issues across layers.

To focus our efforts, we use an existing simulation environment of self-organizing mobile nodes (with sensors and actuators) in wireless networks [4] with a description of our threat model in Section 2. Section 3 surveys security policies for different layers. In Section 4, we explain the extension of the existing cross-layer system tuning framework, xTune, based on a compositional formal approach to accommodate situation- and resource-aware security. A prototype implementation and experimental results are presented in Section 5 and Section 6, respectively. We present related work in cross-layer security in Section 7. Section 8 concludes our paper with future research directions.

2 Threats

We aim to protect physical infrastructure under the following threat model: Rogue sensors or mobile nodes may pretend to be valid entities and, therefore, fraudulent data can be injected to severely compromise the *availability* of infrastructure.

Given this threat model, consider a sensor reading that will trigger a chain of events in response to an emergency (e.g., gas leakage). A video stream must be obtained from the area where gas leakage has been sensed. The video footage then should be delivered via opportunistic network links to the control center. Data may need to be encoded to reduce its volume or to be transmitted in raw data format to save computation time and energy. First, the fraudulent sensor reading will be propagated. This incurs communication overhead in terms of transmission power and bandwidth for data forwarding, which need to be managed. Second, every hop must include a phase of mutual authentication since the node cannot be trusted. Third, the fraudulent nodes should be detected and declared malicious. Finally, the resources of the wireless sensors and mobile nodes need to be provisioned to ensure a certain level of security while avoiding the depletion of residual energy and avoiding congestion. This requires dynamic configuration of individual (seemingly independent) techniques to compose appropriate protections against attack situations while also making optimal use of resources.

3 Security Policies for Different Layers

Specific adaptation policies have been developed within each abstraction layer to enhance a security measure. *Policies* define the individual security techniques available to the system. *Parameters* determine the behavior of a policy. We identify layers, policies, and parameters of interest to effectively demonstrate our concept using the following examples.

Application Layer — In our sample scenario, surveillance is done by sensor readings including video streaming. There is a need to develop techniques that exploit the structure of the data to maximize the efficiency of encryption algorithms. In particular, selective encryption [5] aims to reduce the computational cost of encryption by partial encryption of multimedia content. An example would be encryption of intra-coded blocks, without which the video stream cannot be decoded, to enhance the security of the encoded bitstream.

Middleware Layer — The video data of our sample scenario must be accessed only by authorized consumers. A secure group communication system can be used to enforce this protection. We use a generalization of Secure Spread [6], whose adaptation parameters enable the system to dynamically tailor data transmission to the application requirements and environmental conditions. In the security dimension, this mechanism enables each group to specify the degree of laziness of the key establishment protocol: (i) eager keying will trigger a rekey after every membership change; (ii) key caching will reuse previously cached keys; and (iii) lazy keying will delay rekeying until a message needs to be sent. In the synchrony[3] dimension, groups with different degrees of synchrony can coexist. Taking into consideration the tradeoffs between security and synchrony, we can explore various solutions (e.g., less stringent synchrony semantics with lazy keying protocol) while preserving required security guarantees.

Operating System (OS) Layer — The video sensor stream is received on nodes using an OS. We consider the network stack as part of the OS layer. At the file system level, OS functionality to transparently audit data provenance was prototyped in our recent work, SPADE [7], for software implementation of provenance certification. Continuing this line of research, ciphertext-policy attribute-based encryption (CP-ABE) [8] can provide the desired flexibility in security policy. Using configurable rules, each provenance element will transparently be encrypted at the time of creation with a policy stating which attributes

[3] In a group communication system, the synchrony property ensures that all group members see the same events (group membership changes and incoming messages) and in the same order.

are needed to access it, satisfying the flexible security goal. For example, the provenance record can be verified at runtime to establish mutual authentication to attest the trustworthiness of sensor readings or to be retrospectively checked for forensic purposes.

Hardware Layer — Typical wireless sensors and mobile nodes in our sample scenario are resource constrained. As we explained in Section 1, energy management at the hardware layer has a dramatic impact on the other layers' policies. To save energy or to reduce the risk of an eavesdropping attack, a node can decide to reduce its transmission power [9], which results in residual energy savings at the cost of less coverage (i.e., more hops for end-to-end message delivery). Resource saving on transmission propagates to upper layers (i.e., encryption can consume more energy) and can lead to adaptations there. We analyze this inherent relationship because residual energy is one of the key factors for other layers' decisions regarding whether they can perform computationally intensive tasks (e.g., encryption at the application layer, eager rekeying at the middleware layer) to enhance a corresponding layer's own security level.

4 Cross-Layer Security

4.1 Understanding the Problem of Cross-Layer Security

To enhance security in the context of wireless mobile applications, researchers have proposed several techniques at various system layers, as described in Section 3. Note that one key performance metric for such techniques is how well they manage security under a multitude of constraints in a dynamic situation. Since security comes with cost in terms of performance, energy consumption, storage requirements, and bandwidth used, one needs to optimize security in the context of the operating conditions. However, most security techniques consider only a single system layer remaining unaware of the strategies employed in the other layers. A cross-layer approach that is cognizant of features, limitations, and dynamic changes at each layer enables better optimization than a straightforward composition of individual layers, because solutions for each individual layer can be globally suboptimal.

To coordinate the individual techniques in a cross-layer manner based on the operating condition, one needs to

- Quantify the effect of various security policies at each layer on system properties
- Explore methods of taking the impact of each policy into account and compensating for it at other layers

4.2 Supporting Security Composition within xTune

As explained in Section 1, we extend our earlier research on cross-layer system tuning. In particular, we extend the xTune framework [3] to accommodate

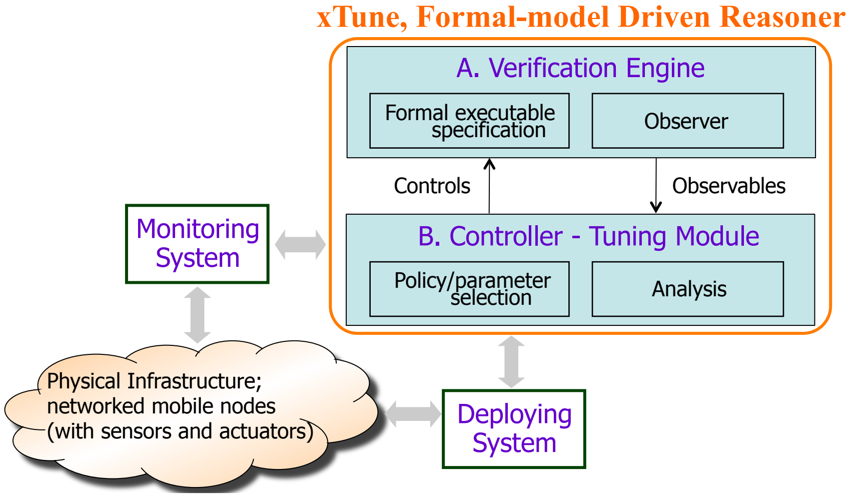

xTune, Formal-model Driven Reasoner
A. Verification Engine
Formal executable specification
Observer
Controls
Observables
B. Controller - Tuning Module
Policy/parameter selection
Analysis
Monitoring System
Physical Infrastructure; networked mobile nodes (with sensors and actuators)
Deploying System

Fig. 1. xTune Cross-layer System Tuning Framework

the cross-layer security concerns and evaluate various strategies for cross-layer adaptation.

Figure 1 illustrates how we envision the xTune framework. Here the monitoring system observes the current status of the nodes and environment that compose the physical infrastructure. The tuning module decides which strategy will be deployed for each node. The tuning module may consult the verification engine to ensure the quality of solution. The xTune framework supports a methodology for tuning that attempts cross-layer adaptation, and for verification that performs formal analysis to quantify utility and cost.

Initially, our framework performs property checking and quantitative analysis of candidate policy and parameter settings via formal executable specifications (i.e., formal models of the system that are executable) and statistical techniques. In particular, *Box A* in Figure 1 represents the formal modeling. The core of our formal modeling approach is to develop formal executable models of system components at each layer of interest. Our formal modeling is based on an executable specification language called Maude [10]. Its theoretical foundation is rewriting logic, a logic with operational as well as model-theoretic semantics. This formal prototyping enables us to experiment with an abstract mathematical but executable specification of the system.

Box B in Figure 1 shows the evaluation phase of given specifications that generate statistics for properties and values of interest, to come up with the cross-layer policies and parameters. The policy and parameter selection is achieved by the compositional method by constraining the behavior of the local optimizers working at each abstraction layer. As proposed in [11], we iteratively tune the system parameters by monitoring the current status of a system via the *observables* to generate the appropriate *control* of the corresponding subsystem.

Subsequently, each local optimizer uses the other optimizers' refinement results as its *constraints*.

Given an optimization problem with model $\mathcal{M}$ and parameter space $\mathbb{P}$ (e.g., $\mathbb{P} = \mathbb{P}_{App} \times \mathbb{P}_{HW}$ with $\mathbb{P}_{App} = \mathbb{R}$ and $\mathbb{P}_{HW} = \mathbb{N}$), the constraint refinement attempts to quickly find a region $P \in \mathcal{R}(\mathbb{P})$[4] containing a nearly optimal solution. For instance, a resulting region can be represented as $P = P_{App} \times P_{HW} = [Th_{min}, Th_{max}] \times [Tx_{min}, Tx_{max}] = [0.2, 0.3] \times [10, 25]$, where Th and Tx indicate paramaters of selective encryption and transmission range control policies, respectively. In particular, we obtain observables by Monte Carlo sampling over the current region $P_i \in \mathcal{R}(\mathbb{P})$ and subsequently refine P_i to P_{i+1}. The refinement such that the utility is maximized based on the samples available, and $size(P_{i+1}) = size(P_i) \cdot \tau_i$, where τ_i $(0.0 < \tau_i < 1.0)$ represents the i-th refinement ratio. The new region P_{i+1} is then used as the current region and the process is repeated.

The input P_i and output P_{i+1} of each refinement step are sets of feasible policies/parameters, and our approach treats P_i as *constraints* when we restrict the candidate policy/parameter space to find P_{i+1}. For example, if the application layer optimizer refines its parameters (e.g., threshold for intracoding; $P_{App} = [Th_{min}, Th_{max}] = [0.1, 0.8]$), then the hardware layer optimizer refines its parameters (e.g., transmission range; $P_{HW} = [Tx_{min}, Tx_{max}] = [1, 60]$), taking the application layer parameter ranges as constraints. The hardware layer results are transmitted to the application layer optimizer for further refinement. In [11], this process is referred to as *constraint refinement*. In this way, the constraint language can be used as the generic interface among different local optimizers, which enables cross-layer coordination of security policies by composition. We explore the above constraint refinement approach to determine the specifics of cross-layer security strategies.

5 Implementation

Figure 2 illustrates our system implementation. Our motivating scenario — distributed surveillance with mobile robots — is described in a declarative manner. The system goal is to collect information in areas where noise or motion is detected. The mobile robots have camera devices that can record short video footage of a target area. The raw video may be directly sent to other nodes if the network supports it, or it can be preprocessed, e.g., by encoding, and then communicated to other nodes. The encoded video can be further encrypted to enhance data confidentiality, which requires a key distribution among mobile robots. The application interacts with a logical engine to perform inferences based on forward and backward reasoning. A detailed explanation on the inference system and a declarative specification of a snapshot-based surveillance scenario can be found in [12].

[4] Region $P \in \mathcal{R}(\mathbb{P}) \iff P \subseteq \mathbb{P}$ is a closed convex set (i.e., if $(x, z \in P) \bigwedge (x < y < z)$, then $(y \in P)$) and P is finitely representable (e.g., interval based). For simplicity we use regions defined by the Cartesian product of intervals for each parameter.

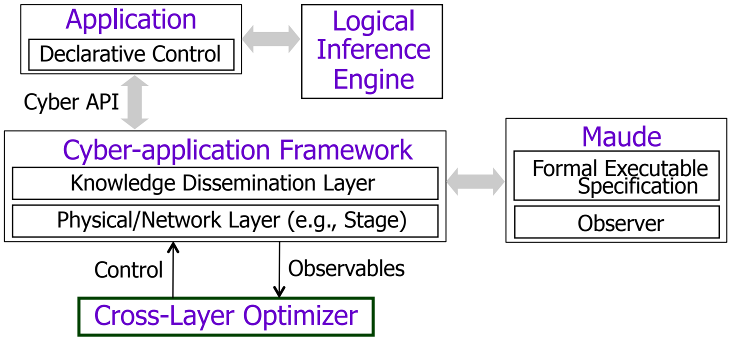

Application
Declarative Control
Logical Inference Engine
Cyber API
Cyber-application Framework
Knowledge Dissemination Layer
Physical/Network Layer (e.g., Stage)
Maude
Formal Executable Specification
Observer
Control
Observables
Cross-Layer Optimizer

Fig. 2. System Implementation

Our motivating scenario is programmed and executed on top of our cyber-application framework (cyber-framework, for short) via cyber APIs [4]. The cyber-framework implements the knowledge dissemination layer that propagates the current progress of the distributed surveillance, in the form of a knowledge unit, on top of an underlying physical/network layer. We use the Stage multirobot simulator [13] that provides both physical device models and wireless network models to simulate losses and delays associated with packet transmissions. The cyber-framework also interacts with Maude for formal prototyping of a group communication system among mobile robots. The cross-layer optimizer in Section 4.2 iteratively tunes the security policies/parameters (i.e., *control*) by monitoring the current status (i.e., *observables*)

At the application layer, a selective video encryption scheme based on secret key cryptography [5] is used to enhance multimedia data security. For this purpose, we build on our earlier work, PBPAIR (Probability Based Power Aware Intra Refresh) [14], which can effectively control partial intracoding, as a component of a selective encryption scheme. We manipulate the algorithmic parameters of PBPAIR to cope with security demands of the application and with the other layers' operating conditions. For example, PBPAIR can be controlled to insert more intramacro blocks (IMBs), leading to more encryption and less coding efficiency with the benefit of less coding energy, which in turn affects other layers' decisions. In this scheme, the number of IMBs can be an approximate measure of data security, bandwidth requirement, and encoding/encryption energy consumption.

As a middleware layer policy, we use the Maude formal specification presented in [6] for secure group communication with relaxed synchrony — virtual synchrony (VS) and extended virtual synchrony (EVS) — and various rekeying mechanisms explained in Section 3. In this work, we explore six different policies. In the case of VS, we use eager keying with i) blocking and ii) nonblocking data multicast while rekeying proceeds. In the case of EVS, iii) eager keying, iv) key caching, v) lazy keying, and vi) a combination of caching and lazy keying are explored with the nonblocking mode.

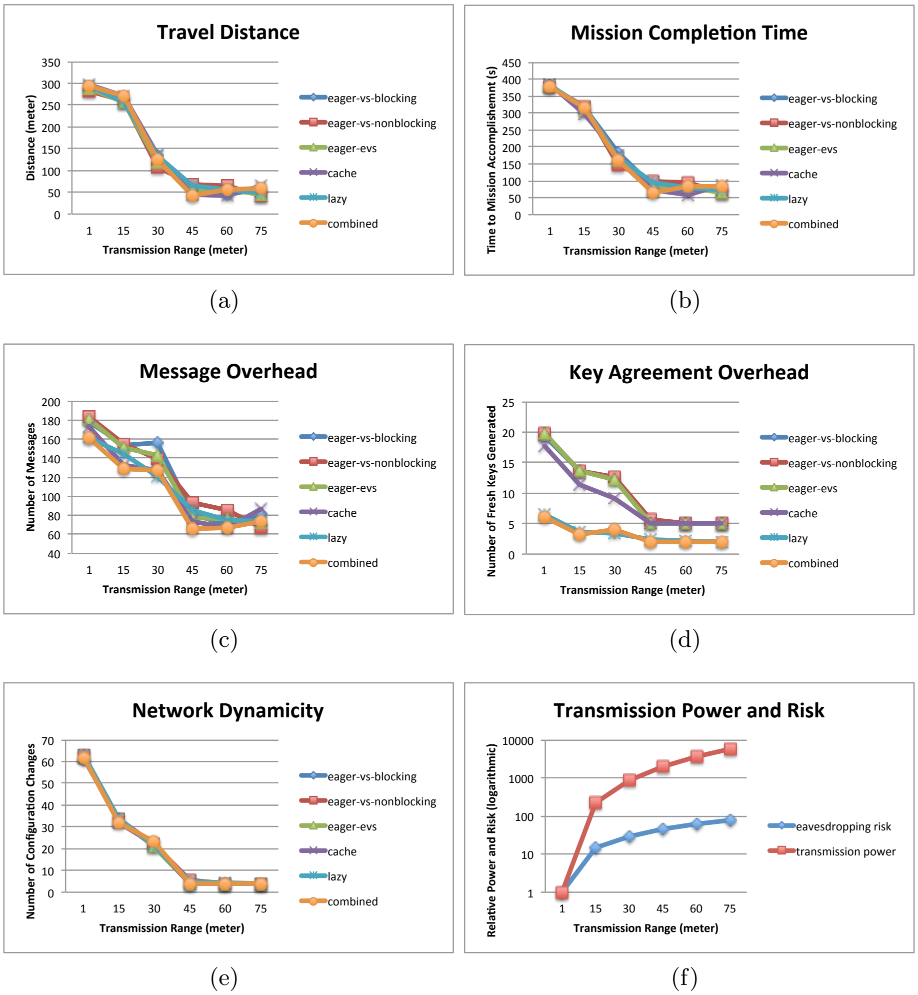
At the hardware layer, we consider transmission power control as an effective way to minimize the eavesdropping risk as proposed in [9]. The authors of [9] define the w-th eavesdropping risk as the maximum probability of packets being eavesdropped with w adversarial nodes present in an ad hoc wireless network. Subsequently, they prove that in an arbitrary random network consisting of n nodes, the 1st order eavesdropping risk is bounded below by $\frac{1}{3}r$, where r is the normalized transmission radius. The OS policy is ongoing research. We will discuss its implication in the concluding remarks (Section 8).

6 Experiments

We evaluated the effectiveness of our approach by carrying out a variety of experiments. Our first set of experiments is concerned with understanding the impact of various policies at different layers. In our second set of experiments, we focused on the effect of composition in the context of cross-layer optimization. Currently, cross-layer optimization is performed on the observables from all robots, and the same optimization results (i.e., parameter settings) are applied to all robots. In reality, each robot needs to autonomously tune its parameters at runtime, which is a topic of future research.

Given the inherent complexity due to dependencies among layers, the first goal is to perform quantitative analysis to determine the appropriate design tradeoff between security, QoS, and resource utilization. For this purpose, we perform an exhaustive exploration on two sublayer optimizers: the keying policy in group communication at the middleware layer and the policy for wireless transmission range control at the hardware layer. Figure 3 compares them in terms of security (eavesdropping risk), QoS (travel distance, mission completion time, communication overhead, network dynamicity), and resource (power consumption). Figure 3(a)-(e) show that the larger the communication range the better the QoS since the stable connectivity among mobile robots reduces the necessity of keying and also leads to immediate propagation of current progress toward mission completion. However, Figure 3(f) presents opposite results since the larger communication range requires more transmission power for wireless devices and higher chances of being eavesdropped, which indicates that single-layer policy often cannot accommodate the inherent complexity.

In the next set of experiments we study the effect of composition as a coordination mechanism for cross-layer security management. To capture the effectiveness of given parameter settings, we define a *utility* function based on the observables and user-defined *soft* and *hard* requirements. The situation that the system behavior resides below the soft requirement is most desirable. When the system is observed in between soft and hard requirement, however, the optimizer needs to tune the parameters. Hard requirement indicates an upper limit, above which a user cannot tolerate the quality degradation. We define the cost function



(a)

(b)

(c)

(d)

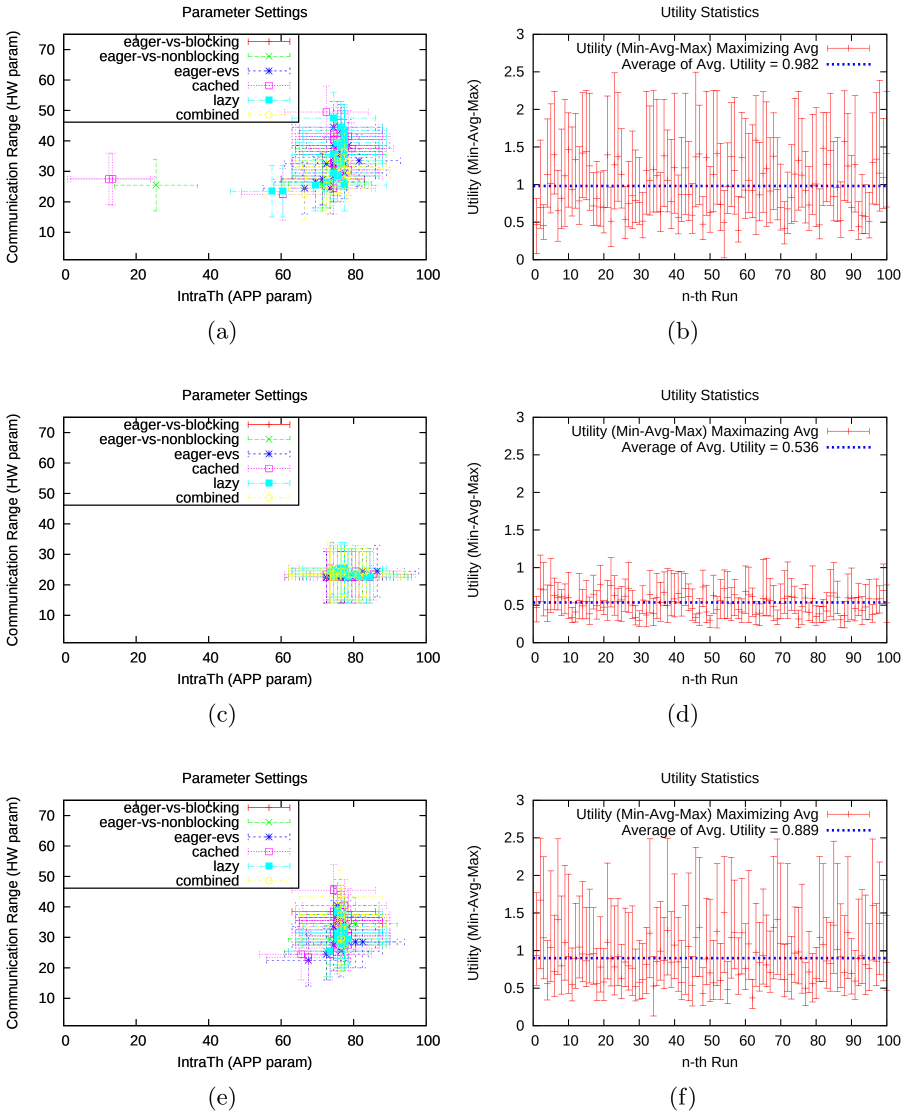
(e)

(f)

Fig. 3. Effect of Various Group Communication Schemes and Transmission Range Control (a) Total Travel Distance of All Robots, (b) Mission Completion Time, (c) Message Overhead for a Secured Dissemination, (d) Key Agreement Overhead for a Secure Group Communication, (e) Network Dynamicity in Terms of View Changes in a Group Communication, (f) Eavesdropping Risk [9] and Transmission Power Consumptions

of the observables $X = (x_e, x_t, x_m, x_k, x_b, x_r)$ as

$$cost(x) = \begin{cases} \infty & \text{if } x \geq h \\ \frac{1}{h-x} - \frac{1}{h-s} & \text{if } h > x \geq s \\ 0 & \text{otherwise} \end{cases}$$



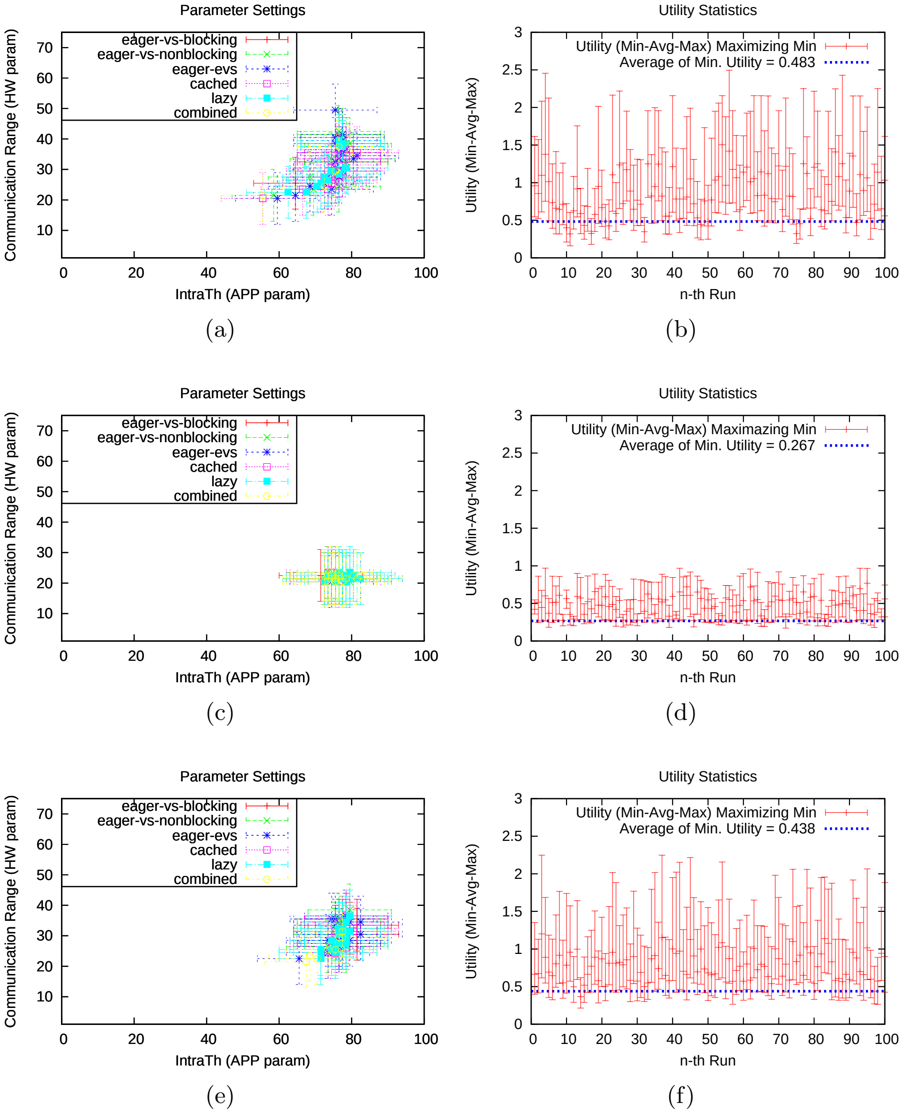
(a) (b)

(c) (d)

(e) (f)

Fig. 4. Effect of Cross-Layer Optimization (Constraint Refinement for Maximizing Average of Utilities) — Parameter Settings and Utility Statistics from 100 Runs (a),(b) Global Cross-Layer Optimization; (c),(d) Without Cross-Layer Optimization; (e),(f) Compositional Cross-Layer Optimization.

where h and s represent user-defined *hard* and *soft* requirement, respectively. The observables x_e, x_t, x_m, x_k, x_b concern energy consumption, mission completion time, messaging overhead, keying overhead, bandwidth, respectively. A risk



(a) (b)

(c) (d)

(e) (f)

Fig. 5. Effect of Cross-Layer Optimization (Constraint Refinement for Maximizing Minimum of Utilities) — Parameter Settings and Utility Statistics from 100 Runs (a),(b) Global Cross-Layer Optimization; (c),(d) Without Cross-Layer Optimization; (e),(f) Compositional Cross-Layer Optimization.

measure $x_r = \frac{r}{e}$ takes into account of eavesdropping risk r and the amount of encryption e. The utility is defined as follows:

$$utility(X) = \frac{1}{\sum_i cost(x_i)}$$

We use the constraint refinement in sublayer optimizers with two different objectives. First, the sublayer optimizer attempts to guarantee average-case performance. In this case, the average of utilities in a region is maximized. Second, to prevent worst-case performance, the sublayer optimizer maximizes the minimum of utilities in a region. For both cases, failures (i.e., zero utility) are avoided by the sublayer optimizers that refine towards a region with fewer numbers of parameter settings leading to a failure. We compare a compositional approach with two extreme techniques: *global* and *local* optimizations. Global optimization is fully aware of sublayers' decisions, which local optimization lacks. In this work, we use a uniform refinement ratio $\bar{\tau}$ for the recursion (i.e., $\tau_0 \cdots \tau_{t-1} = \tau$ and $\forall i \in [0, \cdots, t-1] : \tau_i = \bar{\tau}$) and a constant number of iterations t. For simplicity, we fix a sequential order for the sublayer optimizers. Arbitrary interleaving and distributed optimization are future research topics.

The results for improving average-case and worst-case performance are presented in Figure 4 and Figure 5, respectively. We evaluated the performance of compositional cross-layer optimization in terms of resulting utilities and parameter selections in solving the scenario in Section 5. We compare our compositional optimization (Figure 4(e)(f) and Figure 5(e)(f)) with the two extremes: without cross-layer optimization (i.e., *local* optimization) in Figure 4(c)(d) and Figure 5(c)(d) vs. *global* optimization in Figure 4(a)(b) and Figure 5(a)(b). In Figure 4(a)(c)(e) and Figure 5(a)(c)(e), the X-axis represents the application layer parameter while the Y-axis represents the hardware layer parameter. The various group communication schemes are pictured in different colors. The parameter settings are depicted as cross bars parallel with the x-y axes to represent a region. In Figure 4(b)(d)(f) and Figure 5(b)(d)(f), x-axis and y-axis represent n-th trial and utility distribution in terms of minimum-average-maximum utilities of a resulting region, respectively.

The *compositional* cross-layer optimization in Figure 4(e)(f) and Figure 5(e)(f) presents solutions reasonably close to the *global* approach since the values of resulting utilities reside between that of local and global optimization. By a blue dashed line in Figure 4(b)(d)(f), the average of the objective (i.e., maximizing the average utilities in a resulting region) from compositional optimization leads to 0.889, which resides between that of local (0.536) and global (0.982) optimization. Similarly in Figure 5(b)(d)(f), the compositional optimization leads to the average value of minimum utilities in a resulting region as 0.438, which resides much closer to that of global (0.483) than local (0.267) optimization. The relative execution time of our compositional approach is 8 times faster than the global approach. It should be noted that the speedup can be further improved because the compositional approach can be naturally parallelized. Finally, the refined solution region is very different from that of the global approach, while our compositional optimization gives similar results.

7 Related Work

The authors of [15] formulated the network resource allocation problem as a cross-layer decision of transmission strategies across the APP, MAC and PHY layers of a traditional network protocol stack to maximize multimedia quality with rate and delay constraints. In [16], the authors modeled the communication network as a generalized utility maximization problem to provide a systematic optimization method by analysis of layered decomposition, where each layer corresponds to a decomposed subproblem and the interfaces among layers are quantified as functions of the optimization variables coordinating the subproblems. Those efforts are, however, mainly focused on the architectural decisions in networking, not tuning the system parameters for energy-quality-security gain.

The author of [17] presented a quality-driven security design and resource allocation framework for wireless sensor networks with multimedia selective encryption and stream authentication schemes proposed at the application layer and network resource allocation schemes at low layers. In particular, an unequal (partial) error protection-based network resource allocation scheme is proposed by jointly designing selective multimedia encryption and multimedia stream authentication with communication resource allocation. Their cross layer resource management framework for secure multimedia streaming solves a global optimization problem requiring full awareness of the system dynamics while our compositional approach leads to acceptable solution quality at low complexity. Also, the composition can be fully distributed and capable of utilizing different even conflicting local objectives through the generic interface of constraint language.

8 Concluding Remarks

We have presented first steps toward situation- and resource-aware cross-layer security by investigating the security implications of existing policies and integrating them across all layers, which aims at automating verification and configuration of security policies to respond to cyber-attacks with minimum availability degradation. The existing xTune cross-layer optimization methodology, on which our approach is based, is general enough to handle simple versions of the cross-layer security problem. A compositional approach to enforcing security policies, while enabling desired activities, has the advantage of being agile and flexible. We build upon substantial preliminary efforts to facilitate the understanding of complex distributed multilayered systems. We believe that our work has broad implications for security analysis of application protocol optimization. In principle, it facilitates security-aware tradeoffs to improve system availability and the utilization of limited resources. We have attempted to illustrate the benefit through a sample scenario that involves cooperative operations of mobile nodes and physical infrastructure.

While this work focuses specifically on the surveillance of physical infrastructure, the approach directly applies to many distributed, dynamically reconfigurable architectures. Furthermore, the techniques described in this paper can apply to broader domains, such as vehicular networks (for both civilian and military usage), instrumented cyber-physical spaces [18], and search/rescue operations by first responders who carry mobile devices. More broadly, the proposed approach gives the basic methodology for future work in understanding runtime assurance techniques because model-based distributed control and optimization methods are especially useful for phenomena with significant uncertainty or failure in input. We are currently extending our models to include provenance at the OS layer, which will enable the detection of malicious nodes based on the examination of provenance data as described in our threat model (Section 2). We also plan to improve our composition methods to handle horizontal compositions and more complex constraint solving needed in many cyber-attack scenarios. Accommodating alternative ways of defining utility (e.g., by enforcing an ordering among the observables and using the induced lexicographic ordering or by adapting the concept of a Pareto front) with composite evaluation metrics for security is another interesting research avenue since our approach does not rely on a specific of the utility function.

Acknowledgments. This material is based in part upon work supported by the U.S. Department of Homeland Security under Grant Award Number 2006-CS-001-000001, under the auspices of the Institute for Information Infrastructure Protection (I3P) research program. The I3P is managed by Dartmouth College. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security, the I3P, or Dartmouth College.

Additional support from National Science Foundation Grant 0932397 (A Logical Framework for Self-Optimizing Networked Cyber-Physical Systems) is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF.

References

1. K. Pister, “From smart dust to smart plants the evolution of wireless sensor networking, available at http://www.dustnetworks.com,” in *ISA EXPO ’08: Keynote speech*.
2. S. Mohapatra, N. Dutt, A. Nicolau, and N. Venkatasubramanian, “DYNAMO: A cross-layer framework for end-to-end QoS and energy optimization in mobile handheld devices,” *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 4, pp. 722–737, 2007.
3. xTune Framework, http://xtune.ics.uci.edu.
4. M. Kim, M.-O. Stehr, J. Kim, and S. Ha, “An application framework for loosely coupled networked cyber-physical systems,” in *8th IEEE Intl. Conf. on Embedded and Ubiquitous Computing (EUC-10), Hong Kong, December, 2010*, available at http:/ncps.csl.sri.com/papers/cyber-framework.pdf.
5. B. Bhargava, C. Shi, and S.-Y. Wang, “MPEG video encryption algorithms,” *Multimedia Tools Appl.*, vol. 24, pp. 57–79, September 2004.

6. S. Gutierrez-Nolasco, N. Venkatasubramanian, M.-O. Stehr, and C. Talcott, “Exploring adaptability of secure group communication using formal prototyping techniques,” in *ARM ’04: Workshop on Adaptive and Reflective Middleware*, pp. 232–237.
7. SPADE Project, http://spade.csl.sri.com.
8. J. Bethencourt, A. Sahai, and B. Waters, “Ciphertext-policy attribute-based encryption,” in *SP ’07: Proceedings of the 2007 IEEE Symposium on Security and Privacy*, 2007, pp. 321–334.
9. J.-C. Kao and R. Marculescu, “Minimizing eavesdropping risk by transmission power control in multihop wireless networks,” *IEEE Trans. Comput.*, vol. 56, no. 8, pp. 1009–1023, 2007.
10. Maude System, http://maude.csl.sri.com.
11. M. Kim, M.-O. Stehr, C. Talcott, N. Dutt, and N. Venkatasubramanian, “Constraint refinement for online verifiable cross-layer system adaptation,” in *DATE ’08: Proceedings of the Design, Automation and Test in Europe Conference and Exposition*, 2008.
12. M.-O. Stehr, M. Kim, and C. L. Talcott, “Toward distributed declarative control of networked cyber-physical systems,” in *7th Intl. Conf. onUbiquitous Intelligence and Computing (UIC-10), Xi’an, China, October, 2010*, ser. LNCS, vol. 6406. Springer, 2010, pp. 397–413.
13. R. B. Rusu, A. Maldonado, M. Beetz, and B. Gerkey, “Extending Player/Stage/Gazebo towards cognitive robots acting in ubiquitous sensor-equipped environments,” in *IEEE Intl. Conf. on Robotics and Automation Workshop for Network Robot Systems*, 2007.
14. M. Kim, H. Oh, N. Dutt, A. Nicolau, and N. Venkatasubramanian, “PBPAIR: An energy-efficient error-resilient encoding using probability based power aware intra refresh,” *ACM SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 10, no. 3, pp. 58–69, 2006.
15. M. V. D. Schaar and S. Shankar, “Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms,” *IEEE Wireless Communications*, vol. 12, pp. 50–58, 2005.
16. M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, “Layering as optimization decomposition:a mathematical theory of network architectures,” in *Proceedings of the IEEE*, vol. 95, no. 1, Jan. 2007, pp. 255–312.
17. W. Wang, “Quality-driven cross layer design for multimedia security over resource constrained wireless sensor networks,” University of Nebraska, Lincoln, Dept. of Computer and Electronics Engineering, Ph.D. Dissertation, 2009.
18. M. Kim, D. Massaguer, N. Dutt, S. Mehrotra, S. Ren, M-O. Stehr, C. Talcott, N. Venkatasubramanian, “A semantic framework for reconfiguration of instrumented cyber physical spaces,” in *Workshop on Event-based Semantics, CPS Week*, 2008.