Fundamental Challenges Toward Making the IoT a Reachable Reality: A Model-Centric Investigation

YUANKUN XUE, JI LI, SHAHIN NAZARIAN, and PAUL BOGDAN,
University of Southern California

Constantly advancing integration capability is paving the way for the construction of the extremely large scale continuum of the Internet where entities or things from vastly varied domains are uniquely addressable and interacting seamlessly to form a giant networked system of systems known as the Internet-of-Things (IoT). In contrast to this visionary networked system paradigm, prior research efforts on the IoT are still very fragmented and confined to disjoint explorations of different applications, architecture, security, services, protocol, and economical domains, thus preventing design exploration and optimization from a unified and global perspective. In this context, this survey article first proposes a mathematical modeling framework that is rich in expressivity to capture IoT characteristics from a global perspective. It also sets forward a set of fundamental challenges in sensing, decentralized computation, robustness, energy efficiency, and hardware security based on the proposed modeling framework. Possible solutions are discussed to shed light on future development of the IoT system paradigm.

CCS Concepts: • **Mathematics of computing** → *Hypergraphs*; • **Computing methodologies** → *Modeling methodologies*; • **Hardware** → *Emerging architectures*;

Additional Key Words and Phrases: Internet of Things, mathematical modeling, optimization, challenges

ACM Reference Format:

Yuankun Xue, Ji Li, Shahin Nazarian, and Paul Bogdan. 2017. Fundamental challenges toward making IoT a reachable reality: A model-centric investigation. *ACM Trans. Des. Autom. Electron. Syst.* 22, 3, Article 53 (April 2017), 25 pages.

DOI: <http://dx.doi.org/10.1145/3001934>

1. INTRODUCTION

Recent astonishing advances on sensing modalities, energy-sensitive data processing, communication, and storage system design methodologies have led to capabilities for encoding the tremendously complicated interactions among (i) physical entities, (ii) humanity-related activities, and (iii) man-made devices of enormous heterogeneity by constructing a seamless continuum of networks. We are now embracing a physical world unprecedentedly tight-linked via an extremely extended web of smart devices that are rich in capabilities of sensing, computing, communicating, and control. As envisioned, objects ranging from tiny physical entities such as cells and microbial clusters to giant facilities such as power plants and nationwide power grids will be uniquely addressable and identifiable [Zanella et al. 2014]. Their interactions are captured (i.e., sensing), exchanged (i.e., communication), and analyzed (i.e., computation)

This work was supported by the National Science Foundation (NSF) under grants 1331610 and 1453860. Authors' addresses: Y. Xue, J. Li, S. Nazarian, and P. Bogdan, Ming Hsieh Department of Electrical Engineering, University of Southern California, CA, 90089; emails: {yuankunx, jli724, shahin.nazarian, pbogdan}@usc.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2017 ACM 1084-4309/2017/04-ART53 \$15.00

DOI: <http://dx.doi.org/10.1145/3001934>

by devices that are enormously heterogeneous yet synergistically collaborating through networked systems, hence forming an Internet of Things (IoT).

Apparently, such a vision has not yet become a close reality that is accessible on an everyday basis. However, the IoT is already unveiling its transformative role in rebooting the fundamental ways of our interaction with the physical world [Stankovic 2014; Pajic et al. 2015; Fawzi et al. 2014; Clark et al. 2016]. For instance, smart grids that consist of numerous networked power plants are already extensively deployed to dynamically capture the spatiotemporal variations in user demands and optimize the power plant management for maximization of the efficiency and robustness; wearable physiological sensors and actuators are actively adopted to monitor the crucial biomarkers and provide prompt medical response to alleviate or reverse potential health risks [Bogdan 2015; Bogdan and Xue 2015; Ghorbani and Bogdan 2013; Xue et al. 2016b; Kumar et al. 2013]; smart portable devices are able to run highly varied applications with full awareness of user contexts, which is enabled by vastly varied embedded ambience sensors (e.g., gyroscope, accelerometer, GPS, temperature, and lighting sensors) as well as advanced data processing technologies (e.g., data mining, deep learning).

The physical world is deeply rooted in seamless interactions among entities from vastly varied domains. In contrast to an intrinsically connected physical world, the prior research efforts on the IoT are still quite fragmented and very often confined to disjoint explorations in different application, architectural, security, service, protocol, and economical domains, which poses the following challenges in making IoT a reachable reality in near future:

- (1) How are the disjoint research efforts in relevant fields unified under the awareness of global constraints, objectives, and application scenarios?
- (2) How can we advocate for mathematical models rich in expressivity, thus enabling the exploration, identification, and formulation of IoT system optimization challenges from a global perspective that goes well beyond the reach of fragmented research efforts in specific fields?
- (3) How do such models lift us to reach the possible solutions to these challenges and make predictions of new IoT paradigms?

To discuss the preceding questions, we position this article not as a simply summarized list of well-known problems in the IoT community but as a brief investigation of unexplored challenges based on our proposed mathematical model of the IoT system, hence it is a model-centric investigation. It should be noted that even though these identified challenges exist objectively and do not change as a function of how we describe them, the challenges as well as their nature might not be well understood or even recognized without the help of a proper mathematical characterization (i.e., modeling and formulation based on it) that encodes a sufficient set of elements in related domains (i.e., rich expressivity). The formulation of these challenges is tightly connected to our understanding of the problem space, which further directly translates to how we could reason about the feasibility domain of the problem, the possible design of algorithms for solving it, and the quality of the solutions. All of these are added to the importance of the aforementioned challenges, thus motivating our discussion throughout this work.

The article is organized as follows. We first formulate the IoT system in Section 2 as a set of multiscale time-dependent random hypergraphs with multivalued attributes that is able to encode IoT entities and their interactions at different dimensionalities. It ranges from simple identifiers (e.g., RFID tags), self-powered sensors/actuators, and end-user devices (e.g., smart portable devices), all the way to the mega computing powers (e.g., clusters or data centers). In Section 3, we identify a list of fundamental challenges to be addressed and formulate them as optimization problems by examining

our proposed mathematical model. We also discuss the possible solutions by reviewing prior research milestones. Section 4 summarizes and concludes our work.

2. MATHEMATICAL FRAMEWORK FOR IOT DESIGN AND OPTIMIZATION

2.1. Vision of the Model

The IoT consists of huge collections of entities that possess widely ranged functionalities and unique identifiers for addressability and interoperability through an extensively networked infrastructure. In the IoT, such entities with disparate capabilities (e.g., sensing, actuation, computation, communication, storage) synergically sustain widely ranged applications, which leads to fiercely wide cross dependencies over different domains. Moreover, such cross dependencies evolve as a function of space and time, hence they are nonstationary. The understanding of such dynamic interactions requires an aggregation of information from all related domains. As a simple yet illustrative example, let us look at the metropolitan traffic sensors/controllers, air quality sensors, and biomarker sensors. Sensors and controllers are actively deployed for monitoring the traffic in real time to maximize road network throughput. Although the sensors deployed for air quality surveillance and biomarkers are quite different from those used for traffic monitoring, there are strong geometrical and temporal interdependencies between traffic data, air pollution index, and pathological pathways. The heavily congested points usually overlap with regions with worse air quality, and the most congested hours usually match with the time horizon with worse air quality in a city. Meanwhile, the air quality index and the duration of days with poor air quality usually have strong positive correlation with the number of patients suffering from a variety of cardiovascular and respiratory diseases [Lanki et al. 2015; Brunekreef et al. 2009]. The major pollutants might come from vehicular exhaust fumes, and one possible solution is to smartly control the traffic signals to reduce the congestion, thus minimizing the pollutant exposure to the environment and individuals. In addition, this correlation might not be stationary but stochastic in nature such that the congestion-pollution relationship does not hold over a certain period of time or some regions.

This simple case aptly demonstrates a cross coupling among a set of physical processes that require jointly organized data collection and analysis. With combined information obtained from an integrated IoT system joined by different sensing networks (i.e., traffic, air quality, and medical sensing), the IoT enables joint translation of the sensed data into knowledge, which could otherwise hardly be learned from a disjoint investigation of either system. This example emphasizes the need for unified models in the general context. Such unified models enable the capture of complex interdependencies and stochasticity in IoT systems for the betterment of design, deployment, and control of it, which goes beyond the capability of prior fragmented characterizing methodologies.

Added to the spatiotemporal dynamics of IoT systems is the exorbitantly intensive involvement of objects and their resulting extremely huge datasets. As envisioned, the IoT system will be the largest network ever built, closely interfacing with the physical world in very different ways. The monitored physical process digitalizes the information into a bitstream that flows across a widely spanned network of sensors, storage nodes, and processing centers for interpretation into actionable knowledge. Medical devices constantly capture the critical physiological, proteomic, metabolic, and genomic biosignatures for health care and prompt therapeutic advisement. Various applications running on smart devices leverage large sets of data for interactive gaming, creative activities, or multimedia entertainment. The ubiquitous presence of data generation associated with enormous amounts of entities necessitates the exascale capabilities for data exchange, storage, and process.

In response to all of these unresolved complexities toward the science of IoT system design and optimization, we advocate for a mathematical modeling framework as a sufficient representative of the IoT from a system-of-systems perspective. We resort not on the reductionist approach by fragmenting the complex dynamics of an integrated system into disjoint components. Rather, we seek a generalized mathematical model encoding the stochastic spatiotemporal interactions with no loss of richness in expressivity for characterizing the IoT of enormous heterogeneity.

2.2. Mathematical Modeling

As its stochasticity vastly spans both the spatial and temporal dimensions driven by the heterogeneous entities and exascale interaction involvement, an IoT system could hardly fit into a stationary characterization immune to the rich randomness in all possible ways that they operate, interact, and communicate. Both the structural composition of an IoT system (e.g., the involved IoT devices and their interconnections) and its dynamics (e.g., cross-entity data and control interactions) evolve over time. The failure of an IoT model to capture the time-dependent behaviors fundamentally limits its applicability.

Added to the time-dependent behaviors of an IoT system are the demands to analyze it in multiple resolutions or, alternatively stated, in a hierarchical way. Take the hydropower plant cyberphysical system as an example. Globally, we are interested in analyzing and operating a system that is able to rapidly gather all relevant information (e.g., the electricity demands, irrigation requirement, instant weather condition, and regional climatic patterns) and make optimal global control decisions on the operation strategies of each plant in the hydroplant network. Locally, the detailed design and control of each plant is of primary importance. The deployment and scheduling of sensors, actuators, and computing resources need to be carefully designed such that a good decision could be promptly made at a perfect timing and low cost based on informatively collected data. To enable such hierarchical design and control of an IoT system, a mathematical model should be well set up to allow multiscale analysis. Therefore, we propose to model an IoT system as an array of evolving random hypergraphs at multiscales where IoT entities and their interactions, individually or collectively, are encoded through multivalued time-dependent attributes for insights into both its deterministic and stochastic nature.

Definition 1. The IoT is an array of *time-dependent multivalued random hypergraphs*: $\Gamma(t) = \{G(t) | G(t) = (V, E, t), t \in T\}$. Each node $v_i \in V$ represents a node in the IoT as an active entity. Each hyperedge $e_i(t) = \{v_k | v_k \in V\}$ encodes (i) the physical connections and (ii) multientity interactions among a set of entities at time t . To differentiate between the two types of hyperedges, we define the mapping function $I(i, j)(t) = 1$ if a physical link exists between v_i and v_j . Otherwise, $I(i, j)(t) = 0$. Therefore, the physical link $e_{i,j}(t)$ could be further defined as

$$e_{i,j}(t) \subseteq e_k(t) \in E, |e_{i,j}(t)| = 2 \wedge I(i, j) = 1, \quad (1)$$

where $e_{i,j}(t)$ represents the physical link between v_i and v_j at time t . The nodes and edges of $G(t)$ change over time, forming an array of hypergraphs $\Gamma(t)$. The evolution of $\Gamma(t)$ over the temporal horizon constitutes the dynamics of an IoT system.

Moreover, to allow multiscale analysis, we introduce the scaling function $S : G(V, E, t) \rightarrow G(V', E', t)$. Each upscaling function S_k induces a partition $\pi_k = \{G_1, G_2, \dots, G_n\}$ of $G(V, E, t)$ such that

$$\cap G_i = G \wedge G_l \cap G_k = \emptyset, \forall l \neq k, \quad (2)$$

Fig. 1. Multiscale time-dependent graphical models encode the structural and dynamical features of the IoT system at different resolutions.

for $\forall v'_i \in V'$, $\exists G_l \in \pi_k$ such that

$$v'_i = \cup v_j, v_j \in G_l, \quad (3)$$

and $\forall e'_k \in E'$, $\exists G_l \in \pi_k$ such that

$$e'_k = \cup e_k, G_l \cap e_k \neq \emptyset. \quad (4)$$

We define the inverse function S_k^{-1} as a downscaling function. By applying the scaling functions to the hypergraph instance $G(t) \in \Gamma(t)$, at time t , one can merge nodes and edges or do the opposite to study the IoT system at interested scales.

In particular, we should note that design of an IoT system is usually an application-driven process where the objectives are highly directed by application requirements. For instance, given an IoT system characterized by $G(t)$ at fixed scale and time point, the optimization objectives, associated IoT entities, and their interplays could be vastly varied based on specific application context. To enrich the expressivity of the proposed model, we thus introduce the *editing function* $\mathcal{E} : G(V, E, t) \rightarrow G(V', E', t)$ such that

$$V \cap V' \neq \emptyset \wedge E \cap E' \neq \emptyset. \quad (5)$$

Applying the editing function to $G(t)$ at a given scale and application works like a “trimming process” where only structural details that matter in the context of the given application will be preserved. Take again the hydropower plant as an example. If we were to figure out minimal deployment of sensors in a single plant (i.e., the scale is fixed) to gain information at the required confidence level, it is unnecessary to keep the details of actuators and computing nodes in the $G(t)$. In such a case, we could apply $\mathcal{E}$ after $G(t)$ is correctly scaled by S . To give an intuition, we show a series of scaling and editing operations to the same graph as in Figure 1 at a given time instance, forming an array of hypergraphs of different scales under various interested applications.

By introducing the time-dependent hypergraph, scaling functions, and editing function, we construct the backbone of the proposed model for IoT systems. Given a fixed scale, the evolution of structural features (i.e., nodes and hyperedges) of $G(t)$ captures

the dynamics of IoT entities (e.g., sensors, controllers, smart devices, computers, data centers) and their interactions. Changing the scales of an IoT system at a given time point and trimming it given the application context enable performing analysis, development, and optimization of the IoT system of interest with a sufficiently right amount (neither too many nor too little) of modeling details. In what follows, we discuss in detail the proposed model taxonomically.

2.2.1. Entities in the IoT. It is noted that the functionalities of different nodes in the IoT are highly diversified and changing over time. Moreover, we might be interested in only a subset of them given the specific application context. To capture this variability, with this vision, we introduce the dynamical capabilities signature system (DCSS) for accurate modeling.

Definition 2. The dynamical capabilities signature system (DCSS) $\Lambda = (\Sigma, \mathcal{F}, \mathcal{C}, V, t)$ is a quintuple, where Σ is the finite functionality alphabet given a specific application context. It contains symbols that characterize the functionality of interest given an application. Functionality profiling function $\mathcal{F} : V \rightarrow 2^\Sigma$ relates each entity v_i in an IoT system at time t with a finite set of symbols defined in Σ as a capability signature set $\mathcal{C}$. To give an intuition, Σ can be as simple as an integer set $\{0, 1, 2, 3\}$ in the context of numeric computation, where 2 and 3 represents “addition” and “multiplication.” The 0 and 1 represent “integer” and “floating point,” respectively. $\mathcal{F}$ will induce a capability signature set $\mathcal{C}$ for each node. A node with $\mathcal{C} = \{1, 2, 3\}$ is capable of floating-point addition and multiplication, whereas one node with $\mathcal{C} = \{0, 2\}$ is only capable of integer addition. In particular, DCSS $\Lambda = (\Sigma, \mathcal{F}, \mathcal{C}, V, t)$ changes over time, thus capturing the functional variability of heterogeneous IoT devices.

The benefits of introducing DCSS are manifold. First, DCSS could provide a detailed and accurate characterization of available functionalities of nodes in $G(t)$ provided properly defining $(\Sigma, \mathcal{F}, \mathcal{C})$ given the application of interest. Second, DCSS is not only applicable to modeling of functionalities of IoT entities but also could be used to characterize the application tasks. As we will show later, rather than defining a functionality profiling function, we could define an application profiling function to characterize the application task based on Σ . Third, DCSS gives functional implications for how applications could be distributed for processing over time by comparison of DCSSs of both the applications and IoT entities. DCSS shows the capability of node v_i at time t such that we could capture the dynamics of heterogeneous nodes of different functionalities over time.

Combined with DCSS, we consider the following key attributes for accurately characterizing IoT entities in the proposed model:

- Current energy level* $E_0^{v_i}(t)$: $E_0^{v_i}(t)$ represents the current energy level of node v_i and could be used to categorize the nodes in an IoT system in terms of lifespan: (i) the energy-rich nodes (e.g., nodes with constant power supply or with sufficiently harvested energy) and (ii) energy-starved nodes (e.g., battery-powered or self-powered nodes) that might drop out due to the power failure:
  - v_i is an energy-rich node if $\frac{dE_0(t)}{dt} = 0, \forall t \in T$, and
  - v_i is an energy-starved node, otherwise.
- Critical energy bound* $E_{crt}^{v_i}(t)$: $E_{crt}^{v_i}(t)$ denotes the minimal energy of node v_i to keep it alive in the network. Node v_i will drop out if $E_0^{v_i}(t)$ goes below $E_{crt}^{v_i}(t)$. Thus, we define the $E_{bgt}(v_i, t) = E_0^{v_i}(t) - E_{crt}^{v_i}(t)$ as the *energy budget* at time t , which quantifies the upper bound of energy that v_i could utilize.
- Node failure distribution* $P_{lost}^{v_i}(E_0^{v_i}, E_{crt}^{v_i}, t, \tau)$: $P_{lost}^{v_i}(E_0^{v_i}, E_{crt}^{v_i}, t, \tau)$ quantifies the probability that v_i will fail within $[t, t + \tau]$. It negatively correlates to $E_{bgt}(v_i, t)$.

—*Processing/memory access energy cost* ($pJ/byte$)

$$E_{proc}^{v_i}(t)/E_{mem}^{v_i}(t).$$

—*Estimated time to live* $T_{ttl}(E_0^{v_i}, E_{crt}^{v_i}, \Lambda, P_{lost}^{v_i}, t)$: T_{ttl} represents the lower bound for the lifespan of node v_i before being reactivated (e.g., under recharge or repair).

—*Data capacity* $S_{v_i}(t)$: The data capacity S_{v_i} models how much data node v_i could accommodate at time t .

2.2.2. Interactions of the IoT. The spatiotemporal interdependency of IoT entities never follows a deterministic law and sometimes could be unexpectedly unpredictable in that its evolution usually encompasses a much broader, sometimes even unanticipated, range of participants at different timestamps. This time-varying multiagent involvement spurs us onto modeling the correlations beyond the pairwise fashion and employ the stochastic hyperedges for inclusion of entities contributing collectively to certain or a combination of (i) application task computation, (ii) sensing, and (iii) control of particular processes. Therefore, we propose the use of hyperedges for encoding such multiagent interdependency. More precisely, a hyperlink $e_i(t) = \{v_k | v_k \in V\}$ contains a subset of IoT nodes given $G(t)$, a specific application domain that could be a combination of, yet not limited to, (i) sensing of particular events (e.g., either macroscale events like pollution and floods or microscale events like biostatus), (ii) the control of particular processes, and (iii) processing of a relevant dataset. Given the definition of links in the IoT, we assign the multivalued dynamic attributes to quantify the time-dependent correlations:

—*Link bandwidth* $b(e_{i,j}, t)$: $b(e_{i,j}, t)$ provides the upper bound of the link bandwidth between v_i and v_j at time t .

—*Link failure distribution* $P_{lost}(e_{i,j}, t, \tau)$: $P_{lost}(e_{i,j}, t, \tau)$ quantifies the probability distribution of the physical path that $e_{i,j}$ fails within interval $[t, t + \tau]$. The lower bound of $P_{lost}(e_{i,j}, t, \tau)$ is decided by the sum of the node failure probability for v_i and v_j such that

$$P_{lost}(e_{i,j}, t, \tau) \geq 1 - (1 - P_{lost}^{v_i})(1 - P_{lost}^{v_j}). \quad (6)$$

The node failure distribution is related to the node current energy level and critical energy bound such that $P_{lost}(e_{i,j}, t, \tau)$ also encodes implicitly the influence from, but not limited to, the energy level.

—*Path set* $\Pi(v_i, v_j, t)$: Path set $\Pi(v_i, v_j, t)$ is a set of ordered physical link sequences where $\Pi(v_i, v_j) = \{< e_{i,k_1}, \dots, e_{k_{n-1}, k_n} > | k_n = j \wedge \forall e_{k_p, k_q}, I(k_p, k_q) = 1\}$. $\Pi(v_i, v_j, t)$ shows all possible paths between v_i and v_j at time t . Its subset $\Pi(v_i, v_j, t, \gamma) \subseteq \Pi(v_i, v_j, t)$ shows the candidate paths chosen under certain routing policy γ .

—*Data movement cost* $C(v_i, v_j, t)(pJ/byte)$: $C(v_i, v_j, t)$ models the average cost to move one byte from v_i to v_j at time t . $C(v_i, v_j, t) = \infty$ is said to be *strictly* boundless if $\Pi^0(v_i, v_j, t) = \emptyset$ and *conditionally* boundless given $\Pi(v_i, v_j, t, \gamma) = \emptyset$.

2.2.3. Tasks of the IoT. The unpredictable heterogeneity embedded in an IoT system constantly drives the emergence of applications in a vastly distributed way over geometrical and temporal domains: applications are provoked, processed, and paused or terminated with nondeterministic data structures and volumes in random region(s) of a network at any point in time. Added to the ubiquitous presence of application tasks are the enormously varied application signatures. First, many applications are mostly context sensitive such that applications in disparate domains have quite distinct task architectures, constraints, and requirements (e.g., the data structure and size, execution time, real-time constraints, memory footprint, and communication overhead might fiercely vary for applications in a wireless sensor network and in a distributed computing network). Second, these differences combined with the heterogeneous IoT entities

and interaction patterns have nondeterministic implications on the ways application tasks are distributed, executed, and interact with endpoint users. Third, the time-varying input datasets lead to the rich stochasticity of applications (e.g., user-defined application input on smart devices or the sensed data from an unpredictable physical process). All of these variables add up to the demand for a nonstationary application profiling formalism. Therefore, we extend the DCSS for modeling the applications. We define the general DCSS $\Lambda(t)$ as a time-dependent set of elementary functionalities.

Definition 3. A task $t_i(t)$ is an *ordered DCSS* sequence,

$$t_i(t) = \langle \Lambda_0^i(t), \Lambda_1^i(t), \dots, \Lambda_n^i(t) \rangle. \quad (7)$$

Thus, an *ordered task* sequence is called as a *routine* $R(t)$, where $R(t) = \langle t_0, t_1, t_2, \dots, t_n \rangle$. We have the following definition.

Definition 4. An application in the IoT is a time-dependent set of *ordered routine* sequences where each routine contains tasks of interested granularity,

$$\mathcal{A}_i(t) = \{R_k^i(t) | k \in N\}. \quad (8)$$

We model each application as a time-dependent set of routines. For instance, $\{R_0, R_1, R_2\} = \{\langle t_1, t_2 \rangle, \langle t_2, t_3, t_4 \rangle, \langle t_2, t_5 \rangle\}$ represents a simple application that has three routines composed of ordered tasks. The ordering implies the data/control dependencies and execution order of each routine. With this definition, we quantify the application signatures with the following attributes:

- Execution constraint* $T_e(\mathcal{A}_i, t)$: $T_e(\mathcal{A}_i, t)$ models the upper bound of execution time of application $\mathcal{A}_i$ at time t .
- Processing workload* $S(R_k^i, t)$: $S(R_k^i, t)$, $R_k^i \in \mathcal{A}_i(t)$ represents the size of the dataset to be processed associated with routine $R_k^i \in \mathcal{A}_i$. We define $S(\mathcal{A}_i, t) = \sum_k R_k^i$ as the application processing workload.
- Application holder set* $H(\mathcal{A}_i, t)$: A holder of an application is an entity in the IoT that initiates the entire or a subset of the application and cares about the execution as well as results of the application. Therefore, an application could have multiple holders (e.g., multiple sensors collectively sense data for processing). $H(\mathcal{A}_i, t)$ is thus a collection of such holders. Each application holder set induces a mapping function $\mathcal{M}_h(R_k(t)) = v_k$ for $\mathcal{A}_i$, where routine $R_k(t)$ is initiated by an application holder $v_k \in H(\mathcal{A}_i, t)$. Thus, $H(\mathcal{A}_i, t) = \cup \mathcal{M}_h(R_k(t), \forall R_k(t) \in \mathcal{A}_i(t)$.
- Application carrier set* $C(\mathcal{A}_i, t)$: An application carrier is a set of IoT entities that execute it. One application may have one or several carriers, and each operates on a routine or a task. An application carrier set induces a task-to-entity mapping function denoted as application carrier mapping function $\mathcal{M}_c$. For each routine $R_k(t)$, there exists $\mathcal{M}_c(R_k(t)) \in V$. Thus, $C(\mathcal{A}_i, t) = \cup \mathcal{M}_c(R_k(t), \forall R_k(t) \in \mathcal{A}_i(t)$.
- Cost budget* $B(\mathcal{A}_i, t)$: Cost budget $B(\mathcal{A}_i, t)$ quantifies how much cost an application could afford to be executed. $B(\mathcal{A}_i, t)$ is of particular importance in context of IoT since it largely affects how this application gets executed in the presence of limited computing power, faulty data links, and execution units.

3. CHALLENGES TO DESIGN AND OPTIMIZATION OF THE IOT

As envisioned in Section 1, we target to investigate the underexplored challenges from a direction in which an integrated perspective is well respected throughout our analysis. By virtue of the proposed formalism, the problems addressed are not comprehensive, and we do not seek to address all research topics of the IoT with a single model but to

present the underexplored representative research problems toward the science of IoT analysis, design, and optimization. More precisely, we highlight the research challenges in four topics: efficient sensing, decentralized computation for the IoT, energy-efficient IoT, and security.

3.1. Efficient Sensing

3.1.1. Problem Description. The unprecedented number of tightly coupled smart devices firmly implies that enforcement of the IoT will reshape the very way we interface the human activities with the physical world of vast stochasticity and randomness. Through interpretation of the collected information into actionable knowledge, we are experiencing the rocketing elevation of human capabilities in all possible ways brought by the information gain. In the context of the IoT, we are able to link our vision with not only the perceivable physical world but also the amorphous entities formed through human interactions such as social networks. Therefore, the concept of sensors are extended to any spots from which we have the access to the arbitrary data generation processes of our interest. These processes could be traditional physical processes like weather and air pollution, seismic processes, or social processes like rumor spreading or an evolving social network. Sensing becomes a general data acquisition process in which IoT probing entities of greatly varied forms have been actively deployed. Consequently, from a design standpoint, the following research questions could be posed. How can we obtain sufficient amount of information about the process of interest given a minimal number of probes (e.g., minimal deployment of sensors)? How can we maximize the information about the dynamics of a specific process given only a limited sensor budget? Formally, we could formulate these challenges using the proposed model.

Problem Description 1 (Minimal observability): Given an IoT system description $\Gamma = \{G_i(t) | G_i(t) = (V, E, t), t \in T\}$ and process of interest Y . Find hyperedge $e_k(t)$ such that

$$e_k(t) = A^*(t) = \operatorname{argmin}_{A(t) \in \Gamma} |A(t)| \quad (9)$$

Subject to

$$\mathbf{E}[I_G(t, A; Y)] \geq K, \quad (10)$$

where $I_G(t, A; Y)$ is a general information utility function defined under a specific metric, which quantifies the information of interest obtained by collecting data from sensor set $A \in \Gamma$ through observing a process Y . For instance, an information utility function in a temperature sensor deployment setting could be defined as the area coverage of the sensors. K is the lower bound on the amount of information required given a specific metric. $\mathbf{E}[I_G(t, A; Y)]$ is the statistical average of the information utility given chosen subset A with node failure distribution $P_{lost}^{v_i}(E_0^{v_i}, E_{crt}^{v_i}, t, \tau)$ and link failure distribution $P_{lost}(e_{i,j}, t, \tau)$, where v_i and $v_j \in A$ and $\Pi^0(v_i, v_j, t) \neq \emptyset$. This means that both the sensor failure to collect the information and the link failure or disconnectivity of the network (i.e., there exists at least one unreachable sensor either due to the network topology or communication protocols) will adversely affect the information gain.

Equation (9) gives a lower bound for the number of observed entities in A required to obtain statistically sufficient information about the interested process Y , which considers faulty sensors, lossy links, and communication protocol constraints. Following the similar line, when it comes to the situation where only a limited set of entities are accessible or the deployment of sensors is limited by the affordable resources, it is preferable to obtain as much information as we can. Thus, it leads to the following problem.

Problem Description 2 (Constrained Sensing). Given IoT system $\Gamma = \{G_i(t) | G_i(t) = (V, E, t), t \in T\}$ and a process of interest Y . Find hyperedge $e_k(t)$ such that

$$e_k(t) = A^*(t) = \operatorname{argmax}_{A(t) \in \Gamma} \mathbf{E}[I_G(t, A; Y)] \quad (11)$$

Subject to

$$|A(t)| \leq K, \quad (12)$$

where K is the predefined sensor budget or the maximum number of entities that are accessible during the observation. It should be noted that both problems are not conventional minimal observability or constrained observability problems in control or combinatorial optimization. The metric that defines the information is vastly varied in the context of the IoT, hence the problem space could be even more complicated. For instance, the information utility function can be defined as the prediction accuracy of a nonlinear model that describes the dynamic behaviors of a data generation process (e.g., evolution of temperatures, seismic activities, social interactions). In such cases, the optimization problem translates to how we can learn a data-driven model from the measurements that accurately captures the dynamics of the observed network. Both the choice of model and the sensor have impact on the information utility function, which is in contrast to a traditional sensor deployment problem in combinatorial optimization, where only the location of sensors is considered.

3.1.2. Problem Approaches. The problem described in Equations (9) and (11) are generally NP-hard, as the set cover problem is a subset of them. Prior efforts have been majorly focusing on solving them under different problem and information metric settings. Given the information utility function defined over metrics that lead to a sub-modular property, a greedy heuristic algorithm that maximizes the information gain at each step of choice could provide a solution with guaranteed optimality under both unconstrained cases [Krause and Guestrin 2012] and cases with connectivity constraints [Kuo et al. 2015].

From a control perspective, the observability problem is the dual problem of controllability problem. If the system dynamics could be described by a linear time-invariant (LTI) model (i.e., the dynamics of the system states evolve as a linear function of interactions among a set of contributing agents $\{v_i | v_i \in e_k(t)\}$), the information utility function can be understood as the prediction accuracy of system states based on the retrieved parameters of the linear model learned through a finite length of observations on a subset of nodes in $e_k(t)$. Observability aspects of linear and nonlinear systems have been amply studied since the introduction of the concept by Kalman [1963]. To identify the minimal subset of sensors required to reliably retrieve the system dynamics, or similarly to maximize the information gain given a fixed sensor budget or limited access to the process, the control-based analytical approaches usually maximize the rank of observability matrix and have been well applied to engineered biological circuits [Marucci et al. 2009], biochemical reaction networks [Liu et al. 2013], systems biology [Rajapakse et al. 2012], smart grid [Chakraborty and Ilic 2012; Zhu 2015] and traffic systems [Liu et al. 2013]. To investigate long-range dependent processes, the minimal observability in systems with fractional-order dynamics has been also explored in the context of physiological systems like task-induced brain dynamics [Pequito et al. 2015] and neuromuscular and vascular systems [Xue et al. 2016a]. The observability problem has also been studied in systems with structural features (e.g., the sparsity) [Pequito et al. 2014]. For disruptive scenarios that consider (i) the malfunction/loss of one arbitrary sensor and (ii) the failure of links between a pair of agents, a two-step heuristic is usually adopted [Liu et al. 2015]. The choice of sensors is first

made to ensure that the system is structurally observable. Then additional sensors need to be considered for robust observability.

Despite prior efforts and theoretical results in separate problem settings, it is still an open problem in a general IoT scenario where the aforementioned approaches might be invalidated considering that (i) the interactions of IoT entities is rich in randomness and usually nonlinear, (ii) the information utility function is not monotonic and nonsubmodular, and (iii) the minimal observability problem could be multiobjective and needs to consider a multitude of heterogeneous processes for which prior studies have very different mathematical tools and models, making the joint optimization very difficult.

3.2. Cost-Aware Decentralized IoT Computation

3.2.1. Problem Description. Given an application of interest and an IoT system consisting of an array of entities with different capabilities, a fundamental question would be: What is the best-fit architecture of IoT that supports the application given varying contexts? One possibility is a cloud-centric architecture where the data collected in distributed way is merged into energy-rich powerful data centers for processing and storage. At the heart of such an IoT paradigm lies the omnipotent hub nodes that are supposed to be responsive to the data processing and storage requests from all nodes in the network. But is cloud-centric paradigm the only solution? It is noted that not all IoT applications are closely related to big data analytics, which entails the exascale data computing capabilities usually possessed only by the mammoth data centers. Actually, the data movement to distantly located data centers for processing consumes an unnecessarily huge amount of energy in view of the abundant presence of locally available computing resources that could collectively contribute to a sufficient processing power with reduced response time and consumption of less power in shuttling the data among sensors, data centers, and end-point users. Smart phones, laptops, cars, and even refrigerators—all possible smart devices consisting of a web of smart things that could provide computing and data-storage capabilities shall be recruited in a distributed way for application task processing and storage. More importantly, a centralized computing paradigm introduces robustness issues into the IoT infrastructure as the failure of data centers and networking infrastructures that provide us access to them means the failure of the entire IoT system. Worse still, it is in such scenarios that we might desperately need the power of the IoT. For instance, it is of critical importance to gather information and process it after a major military attack or a devastating natural disaster where giant computing facilities and networking resources are easily targeted or damaged. If the IoT were built over such a centralized paradigm, there is great chance that the system might malfunction in such least expected situations. Consequently, a decentralized computing paradigm is not only preferred but also necessary for the IoT.

However, to distribute the storage and computing demands locally introduces new challenges, as these nodes cannot be considered constantly reliable due to (i) energy constraints $E_{bgt}(v_i, t)$ (e.g., self-powered devices might go offline to harvest the energy or smart devices need recharge); (ii) limited capabilities, such as limited DCSS Λ (e.g., only a subset of functionalities required by certain application could be accommodated and the distributed task cannot interrupt the applications running locally on smart devices); and (iii) lossy communication data path and faulty entities (i.e., $P_{lost}^{v_i}(E_0^{v_i}, E_{crt}^{v_i}, t, \tau)$, $P_{lost}(e_{i,j}, t, \tau)$, data links, and nodes might fail with much greater chance compared to that of data centers). All of these limitations serve as constraints for the minimization of execution costs given a specific application and an IoT system with a set of entities. Formally, we can define the optimization problem as follows.

Problem Description 3 (Cost-Aware Decentralized Computation). Given an IoT system description $\Gamma(t) = \{G_i(t) | G_i(t) = (V, E, t), t \in T\}$, an application $\mathcal{A}(t) = \{R_k(t) | k \in N\}$, and its induced application holder mapping function $\mathcal{M}_h$.

Find hyperedge $e_i(t)$ and application carrier mapping function $\mathcal{M}_c$ to

$$\text{Minimize} \quad \text{Cost}(\mathcal{A}(t)) = \left\{ \sum_k S(R_k, t) C(\mathcal{M}_h(R_k), \mathcal{M}_c(R_k), t) \right. \quad (13)$$

$$\left. + \sum_{k'} \left(E_{proc}^{\mathcal{M}_c(R_{k'})}(t) + E_{mem}^{\mathcal{M}_c(R_{k'})}(t) \right) S(R_{k'}, t) \right\} \quad (14)$$

Subject to

$$e_i(t) = H(\mathcal{A}, t) \cup C(\mathcal{A}, t) \quad (15)$$

$$R_k \subseteq \mathcal{F}(\mathcal{M}_c(R_k)), \forall R_k \in \mathcal{A}(t) \quad (16)$$

$$\text{Cost}(\mathcal{A}(t)) \leq B(\mathcal{A}, t) \quad (17)$$

$$\left(E_{proc}^{\mathcal{M}_c(R_k)}(t) + E_{mem}^{\mathcal{M}_c(R_k)}(t) \right) S(R_k, t) \leq E_{bgt}(\mathcal{M}_c(R_k), t), \forall R_k \in \mathcal{A}(t) \quad (18)$$

$$S(R_k, t) \leq S_{\mathcal{M}(R_k)}. \quad (19)$$

Equation (14) is the objective function that quantifies the overall energy consumed by moving, storing, and processing the datasets associated with application $\mathcal{A}$, where $S(R_k, t)$ denotes the size of the dataset associated with routine $R_k \in \mathcal{A}$ at time t . This optimization problem can be understood as finding a best way to distribute the tasks belonging to an application such that the overall energy consumed is minimized and constrained by the following: (i) the available functionalities should cover the requirements of the application (Equation (16)), (ii) the overall cost does not exceed the application budget (Equation (17)), (iii) the energy consumed by the task assigned to an entity does not exceed what it is allowed to offer (Equation (18)), and (iv) the data storage requirement of the task cannot exceed the data capacity of the entity to which it is mapped. In particular, the IoT task mapping optimization problem formulated as Equation (14) does not make the specific assumption that the application carrier can only be an omnipotent node like data centers (i.e., v_i is an omnipotent node if $\mathcal{F}(v_i) \cap \mathcal{A} = \mathcal{A}$), for any choice of $\mathcal{A}$), thus allowing a new IoT paradigm where tasks are distributed across available computing resources, including data centers. Intuitively, an application given a limited energy budget might choose to take advantage of locally available resources instead of moving its data to data centers as a result of an energy penalty imposed by the data movement cost.

3.2.2. Problem Approaches. The problem described by Equation (14) calls for a paradigm shift away from conventional centralized cloud computing to an architecture where computing resources of high heterogeneity can be organized and assigned in a distributed way. The conventional cloud computing paradigm frees the enterprise and the end user from the specification of both architectural and deployment details via a unified and centralized service model (e.g., PaaS), but at the expense of costly data movement. This becomes a problem for latency-sensitive applications requiring nodes in the vicinity to meet their delay requirements and constituting a major part of the IoT domain. In sharp contrast to a centralized cloud, Bonomi et al. [2012] propose fog computing, which provides mobility support and geodistribution in addition to location awareness and low latency. Fog is positioned as a virtualized platform between the traditional cloud and end devices and enables the ingestion and processing of the data close to the source. In this paradigm, the cost function as shown in Equation (14) is decreased by setting dedicated computing nodes close to data generation sources.

Nevertheless, this paradigm is valid as long as we have prior knowledge about these data sources and they are relatively stationary. In many scenarios, such as monitoring physical processes, pervasive health care, or mobile applications, either (i) it is difficult to predict the location of the source of data to be processed and set up computing nodes accordingly in advance or (ii) it is impractical to set up such dedicated nodes given a limited budget or unexpected emergencies. In response to this problem, mobile cloud computing (MCC) has been emerging as a new cloud computing paradigm that takes advantage of distributed mobile computing resources. It allows the offloading of computation not only to the powerful data centers but also to the locally available computing resources. Its deployment has proved useful in numerous application scenarios, such as multimedia [Kumar and Lu 2010; Zhuang et al. 2014; Canepa and Lee 2010], location-based services [Vallina-Rodriguez and Crowcroft 2011; Li et al. 2015a], collective sensing and monitoring [Cheng et al. 2010; Thiagarajan et al. 2009; Herring et al. 2010; Lane et al. 2010], and health care [Cimler et al. 2014; Wu et al. 2013; Hoang and Chen 2010]. In contrast to the actively advancing research on MCC, its connection to the IoT has not been richly observed in state-of-art academic literature. In the work of Hasan et al. [2015], Aura is proposed as an IoT-based cloud service for localized mobile computation outsourcing. It allows the clients to create ad hoc clouds using the IoT and other computing devices in the nearby physical environment while maintaining the flexibility of cloud computing. To improve the efficiency of offloading a computation to an IoT-based MCC environment, Kim [2015] proposes a nested game model that approaches an optimal solution for the offloading computation in the distributed IoT system.

It should be noted that the cost-aware decentralized computation problem in the IoT described by Equation (14) is still a challenge that is open to proposals addressing the following critical research problems. First, how can data be securely exchanged in a distributed way without compromising its privacy? Second, how can end users be motivated to share their resources for building up such an IoT cloud? Third, how can such an IoT cloud stay reliable and robust in the presence of time-varying user behaviors, unreliable computing nodes, and connectivity, as well as rich stochasticity, in system composition? Fourth, how can such an IoT cloud provide guaranteed quality of service (QoS), especially to applications with a high responsiveness requirement? All of these unresolved questions call for joint research efforts in a cross-disciplinary context based on expressive models that are able to provide a system-of-systems-level perspective.

3.3. Robustness

3.3.1. Problem Description. Robustness is of great importance, especially for mission-critical applications such as automobiles and health care IoT systems. The IoT robustness challenges can be subdivided into device level, circuit level, and system level. The main objective of device-level robustness is to devise robust devices that are immune to errors, whereas circuit-level robustness mainly focuses on minimizing the failure rate of the circuits. In this article, we are more interested in system-level robustness, as it collectively considers all low-level robustness aspects and encodes their contribution in the node failure distribution rate $P_{lost}^{v_i}(E_0^{v_i}, E_{crt}^{v_i}, t, \tau)$.

Following the same vision of an emerging decentralized IoT computing paradigm, let us assume that we are able to assign application tasks to a collection of IoT entities in a distributed way. A subsequent research challenge is how such assignment strategy is designed in a robust way such that the probability of application failure is minimized given faulty nodes and limited energy budget. We should note that such failure might originate from two sources: (i) the failure to execute the application tasks due to the node failure or (ii) the failure to deliver in a timely fashion (i.e., violating the timing

constraint). The complexity of this problem is rooted in the fact that it is not a pure mapping problem, because both the execution time and node failure distribution are influenced by how the mapping strategy is designed. Intuitively, the heavier a task is, the more energy it will consume. The energy level of a node is one of the deciding factors for failure distribution. To formally describe this problem, we formulate this challenge as follows.

Problem Description 4 (Robust Application Assignment). Given an IoT system description $\Gamma(t) = \{G_i(t) | G_i(t) = (V, E, t), t \in T\}$, node failure distribution $P_{lost}^{v_i}(E_0^{v_i}, E_{crt}^{v_i}, t, \tau)$ for a subset of nodes $K(v_i)$, and an application set $\mathcal{A}(t) = \{\mathcal{A}_k(t) | k \in N\}$.

Find hyperedge $e_i(t)$ and mapping function $\mathcal{M}_c : R \rightarrow V$ such that

$$\begin{aligned} & \text{Minimize} \\ & \text{Max} \quad P_{failure}^{\mathcal{A}_i(t)} = P(T_f(\mathcal{A}_i) > T_e(\mathcal{A}_i, t)) + P_{lost}^{\mathcal{A}_i(t)} \end{aligned} \quad (20)$$

Subject to

$$P_{lost}^{\mathcal{A}_i(t)} = 1 - \prod_{R_k^i} \left(1 - P_{lost}^{\mathcal{M}_c(R_k^i)}(E_0^{\mathcal{M}_c(R_k^i)}, E_{crt}^{\mathcal{M}_c(R_k^i)}, t, T_f(\mathcal{A}_i)) \right) \quad (21)$$

$$e_i(t) = (\cup H(\mathcal{A}_i, t)) \cup (\cup C(\mathcal{A}_i, t)) \quad (22)$$

$$R_k^i \subseteq \mathcal{F}(\mathcal{M}_c(R_k^i)), \forall R_k^i \in \mathcal{A}_i(t), \forall \mathcal{A}_i \in \mathcal{A}(t) \quad (23)$$

$$S(R_k^i, t) \leq S_{\mathcal{M}_c(R_k^i)}, \quad (24)$$

where the term $P_{failure}^{\mathcal{A}_i(t)}$ represents the probability of failure of an application $\mathcal{A}_i(t)$. An application fails due to either (i) the timing failure (i.e., the finish time of the application indicated by $T_f(\mathcal{A}_i)$ is larger than the execution constraint $T_e(\mathcal{A}_i, t)$) or (ii) the failure of any one of the nodes to which an application is mapped (i.e., $P_{lost}^{\mathcal{A}_i(t)}$). Thus, the optimization problem can be understood as a searching process for an optimal task assignment strategy as well as a collection of nodes such that (i) this collection of nodes is sufficiently capable of executing the application and (ii) the risk of the application failure is minimized in the worst case for a set of applications considered given faulty nodes, hence maximizing the robustness from a system level.

3.3.2. Problem Approaches. From a device-level perspective, great efforts have been spent dealing with variations and noises, which harm the robustness of the device. Devices featured with high robustness (e.g., TEFT [Morris et al. 2015]) have been developed, and robustness-enhancing techniques have been applied to new devices, such as carbon nanotube [Hills et al. 2014]. From a circuit-level perspective, several techniques have been developed to improve circuit robustness (or resilience) by dealing with specific robustness threats, such as crosstalk [Sengupta and Barman 2015], soft error [Wu and Marculescu 2014], and thermal noise [Liu et al. 2014]. From a system-level perspective, improved robustness can be achieved from leveraging the decentralized nature of the IoT [Petersen et al. 2014]. For instance, a low-complexity distributed data replication mechanism could be applied to the distributed storage in an IoT system to minimize the probability of node failure for the improvement of system robustness [Gonizzi et al. 2015].

3.4. Energy Efficiency

3.4.1. Problem Description. Although a single IoT entity has limited energy consumption, a one trillion node IoT system can consume a considerable amount of energy

[Klinefelter et al. 2015]. The energy efficiency problem tends to trail novelty in gauging the appeal of IoT devices [Poindexter 2014]. Therefore, it is imperative to improve energy efficiency from device entity design up to IoT system architecture development. Two important challenges arise. First, given the energy budget of nodes in the IoT network, how can we distribute the application such that the negative effect on their time-to-live (TTL) is minimized while the application execution constraints are met? Second, how can we improve the TTL for nodes or power efficiency of the IoT network? The first challenge requires an architecture-level improvement, whereas the second challenge is more related to circuit-, device-, and protocol-level optimization. The second challenge could be further divided into two subproblems. First, how can we control the power consumption using more efficient nano devices featuring low power consumption (FinFET, nanotube, etc.), more efficient circuit-level optimization (efficient storage like SRAM, efficient RF transceivers and baseband), and improved heterogeneous network protocols for different applications (smart grid, EVs, etc)? Second, how can we absorb power from the environment or how can we efficiently harvest power from IoT entities?

Apart from the aforementioned challenges, SRAM design also requires attention from the energy-efficiency point of view. SRAMs are required by many nodes in the IoT system to serve as cache memories, scratchpad buffers, register files, and tables. SRAM capacity affects a node's data capacity $S_{vi}(t)$ directly; the access speed of SRAM has a great influence on the node's computation speed; reliability performance of SRAM, as a part of the overall resilience of the IoT entity, plays an important role in the node failure distribution $P_{lost}^{vi}(E_0^{vi}, E_{crt}^{vi}, t, \tau)$; and most importantly, the power-performance portfolio of SRAM greatly affects energy-related properties of each node, such as $E_0^{vi}(t)$, $E_{crt}^{vi}(t)$ and $E_{proc}^{vi}(t)/E_{mem}^{vi}(t)$, affecting the TTL of each IoT node. In an IoT world driven by a new generation of medical devices, handheld devices, consumer electronics products, communication systems, and industrial controllers, there are growing demands for SRAMs that have high performance and small footprints while running with a minimized power supply. Such SRAMs should be able to support high-performance devices that have overall low power consumption and can perform complex operations within a limited area budget.

Problem Description 5 (Energy Efficiency). Given an IoT system description $\Gamma = \{G_i(t) | G_i(t) = (V, E, t), t \in T\}$, an application set $\mathcal{A}(t) = \{\mathcal{A}_k(t) | k \in N\}$ such that for $\forall v_i \in V$,

Maximize

$$\text{Min}\{T_{ttl}(E_0^{vi}, E_{crt}^{vi}, f_{vi}(t, e_k(t)), P_{lost}^{vi}, t)\}. \quad (25)$$

$$v_i \in e_k(t), \forall k$$

As indicated in Equation (25), the energy efficiency problem aims to maximize the minimum of the estimated time-to-live T_{ttl} among all nodes related to the application set $\mathcal{A}(t)$ in the IoT system (i.e., maximize the lower bound for the lifespans of the related nodes), considering the effects of energy-related properties as well as node failures.

3.4.2. Problem Approaches. The IoT energy efficiency solution is developed in a hierarchical manner. At the system level, energy-efficiency improvement techniques developed for the smart grid [Li et al. 2015b], electric vehicles [Chang et al. 2014], and cloud computing [Li et al. 2016] can be transformed to fit to the IoT energy efficiency problem, and protocols can be improved by reducing the number of transmissions and the amount of data in each transmission so as to cut down the power consumption for communication. At the circuit level, there are many promising low power techniques,

such as power gating [Arora et al. 2014], dynamic voltage and frequency scaling [Wu et al. 2014], and near- and subthreshold computing [Pinckney et al. 2016; Li et al. 2015]. At the device level, efforts have been spent on reducing leakage power consumption of transistors [Bardine et al. 2014], and a great amount of power-harvesting devices are being designed and implemented to avoid the cost and scalability challenge of battery replacement in such large numbers [Klinefelter et al. 2015; Battista et al. 2014; Fadhil et al. 2014]. Further efficiency can be achieved by using a storageless and convert-erless energy harvesting technique, where the harvested energy is directly supplied to the target IoT device [Lee and Chang 2015], and combining multiple harvesting sources to provide a more stable power supply [Heidari et al. 2015; Ding et al. 2016]. Improved battery charging and discharging control also helps overall energy efficiency [Babazadeh et al. 2014; Kawahara et al. 2014]. Enhanced SRAM designs like those of Chang et al. [2015] have been developed specifically for IoT applications, where the main goal is to reduce power, area, and voltage supply while maintaining a sufficient speed to support real-time computation. Advanced read and write assist techniques for low-voltage SRAMs are promising for improvement of SRAM energy efficiency [Yahya et al. 2015]. In addition, STT-RAM and RRAM, which enable an area reduction of four to five times compared to traditional 6T SRAM [Steegen 2015] could be used for future IoT applications. The TFET memory technique is also an option on the table [Morris et al. 2015]. Further energy efficiency can be achieved by using low-power asynchronous SRAMs [Abinaya and Sophia 2015]. IoT could indeed be memory centric in the sense that the memory transfer would possibly dominate the logic transfer in IoT links. Several content-centric architectures have been proposed [Song et al. 2013; Datta et al. 2014; Baccelli et al. 2014], which improve energy efficiency of information-centric networking in the IoT. Reducing data transmission costs in the IoT is another promising direction. A compressed sensing framework is proposed for the IoT, in which in-network compression is put into place to achieve accurate data reconstruction and lower energy efficiency [Li et al. 2013].

3.5. Security

3.5.1. Problem Description. The IoT is a heterogeneous complex distributed system with vastly ranged resources across tiers, description languages, OSes, and specialized hardware [Levis 2015], and IoT converges operational technology networks and information technology networks in addition to potentially billions of smart objects, such as sensors, actuators, and other networked devices [Cisco 2015; Hewlett Packard 2015]. The security requirements vary considerably from one network to another, and entities within the same network can have very distinct security needs [Cisco 2015; Hewlett Packard 2015; Infineon 2015]. The heterogeneity and complexity of the IoT, as well as the convergence of various networks that are built on different technologies and managed with different priorities [Cisco 2015], have significantly expanded existing security challenges and have led to the surface of enormous novel types of attacks. Along with the rapid growth of the IoT, the breadth and depth over existing network connectivity has also increased significantly, and this further exacerbates the IoT security problem by requiring more effort and resources in security development and enhancement to compensate for the increased problem scale. Instead of targeting traditional IoT security challenges such as confidentiality, identification, authentication, authorization, privacy, malware, and software vulnerability, which have been well summarized in previous works [Aggarwal et al. 2013; Zhang et al. 2014; Pohls et al. 2014], we focus on two novel categories of security challenges: (i) the category of attacks that are close to the physical layer of IoT entities and (ii) the category that targets vulnerability in the IoT distribution computation architecture.

It is expected that the majority of IoT attacks will occur at the network level, such as eavesdropping, data modification, and denial of service attacks [Xu et al. 2014]. However, network-level attacks are not the only IoT security concerns. The end entities in an IoT system are billions of hardware devices, which could be RFID tags, sensor devices, barcodes, or 2D codes residing in personal area networks, home area networks, or industrial automation networks using quite different network protocols [Levis 2015]. There are a variety of existing methods of hardware attacks targeting these IoT end entities, including glitching [Korczyk and Krasniewski 2012], power analysis [Lerman et al. 2014], UV attacks [Wang et al. 2014], microscopy [Stellari et al. 2014], fault injection [Delvaux and Verbauwhede 2014], voltage contrast [Kison et al. 2015], magnetic scan [Skorobogatov 2014], and reverse engineering [Wei et al. 2014]. Most of the aforementioned attacks can also be categorized into side-channel attacks, which break the system security based on an extra source of information from hardware, such as power consumption, electromagnetic leaks, or even sound [Rostami et al. 2014]. For instance, unloopers can be used to jeopardize smartcards of pay TV by causing the cards to skip one or more important instructions that are related to the security procedure [Noel and Miller 2012]. Another example is the hacking of the Content Scramble System (CSS), a digital rights management scheme for DVDs that aims to prevent the copying of material via encryption [Watson et al. 2013], leading to a large segment of the global public to make illegal copies of DVD movies [Rawlings 2008]. Traditional network-level security techniques cannot prevent all types of hardware attacks to these physical entities in the IoT network, which poses severe hardware security challenges.

Apart from hardware security challenges, there is an emerging IoT architecture that utilizes available IoT device computation resources to process computation tasks [Kleyman 2013]. More specifically, in such an architecture, application tasks are distributed among nearby IoT entities with processing capacity and collect the results back to the senders. The owners of those IoT devices involved can get paid based on the amount of tasks processed on their devices. Novel IoT computing architectures like this are becoming a reality; however, this trend poses new security challenges. First, data streams in the hyperlinks among involved computation IoT entities lack the same security protection compared to current Internet network links and may be jeopardized by side-channel attacks. Second, owners of IoT computing devices may seek valuable information based on the part of computation tasks sent to their devices. Third, owners of IoT computing devices may compromise the data integrity maliciously. Unlike a cloud-based computation architecture that has a relatively stable physical implementation in terms of links, processors, and storage, the IoT computing architecture is built on a dynamic IoT network where the hyperlinks among entities are unstable and changing, and both performance and storage capacity of IoT computing devices may alter. The preceding uncertainties can be found in timing-related parameters, such as $e_{i,j}(t)$, $b(e_{i,j}, t)$, $\Pi^0(v_i, v_j, t)$, $S_{vi}(t)$, and $f_{vi}(t)$. In addition, unlike cloud-based computation, the IoT computing architecture lacks control over all IoT entities to implement unified security protocols, as the devices are owned by individuals or companies and may not have the capacity to process the protocols. Therefore, the security challenge in the IoT computing architecture cannot be addressed properly with a traditional security scheme. The solution should consider the heterogeneity and dynamics of IoT networks as well as IoT entities. We formulate the preceding security challenges as follows.

Problem Description 6 (Hardware Security). Given an IoT system description $\Gamma(t) = \{G_i(t) | G_i(t) = (V, E, t), t \in T\}$, and an overall cost budget $B_{security}$.

Find the set of entities and links $M(t) = \{v_i, e_i | v_i, e_i \text{ face hardware security threats}\}$, and the corresponding security improvement strategy that contains W actions such

that

Maximize

$$\begin{aligned} & \text{Min}\{\zeta_{security}(t) | \zeta_{security}(t) = \zeta_{base}(t) \\ & + \sum_{w=1}^W \zeta^w(t)\}, \forall t \in T \end{aligned} \quad (26)$$

Subject to

$$\sum_{w=1}^W B(w) \leq B_{security} \quad (27)$$

The $\zeta_{security}(t)$, $\zeta_{base}(t)$, and $\zeta^w(t)$ terms in Equation (26) represent the system-level security, previous system level of security without additional actions, and the increased level of security after applying the w -th action at time t , whereas the term $B(w)$ in Equation (27) is the cost of investigating and implementing the w -th action for hardware security improvement. The security improvement strategy could be in the form of enhanced or new hardware for a subset of entities or links in $M(t)$, new or updated security protocols in one or several levels adopted by part of the IoT system $\Gamma(t)$, or even an improved architecture of the IoT system that has a higher degree of hardware security tolerance (i.e., valuable information cannot be easily gathered with a subset of hardware entities hacked).

3.5.2. Problem Approaches. There are many techniques for enhancing hardware security in the IoT world. Obfuscation and camouflaging are two effective techniques to thwart reverse-engineering security threats. More specifically, obfuscation inserts additional gates, states, or transitions into the hardware to hide the original functionality and implementation of the design [Wendt and Potkonjak 2014; Frey and Yu 2015; Zhang 2015], whereas camouflaging is a layout information hiding technique that prevents the extraction of netlists from the layout by inserting additional cells in empty spaces or making layouts of different gates look identical [Cocchi et al. 2014; El Massad et al. 2015; Karri et al. 2015]. The overhead associated with obfuscation and camouflaging needs to be carefully analyzed so that the most cost effective technique can be selected to increase the hardware security.

Side-channel attacks are powerful, as they gather information from the leakage in environmental channels. One category of countermeasures targeting side-channel attacks is the leakage reduction approach, which smooths the leaking information, such as power consumption trace and electromagnetic trace [Braun et al. 2015; Fang et al. 2015]. Adding noise into the traces is an alternative approach, as the injected artificial noise makes information extraction more difficult for attackers [Demme et al. 2012]. Another approach is to frequently derive and update key information such that the side-channel attacker cannot gather sufficient information for breaking the key value before the next key update [Oswald et al. 2014; Xiao et al. 2014]. Physical unclonable functions (PUFs) are an effective hardware security solution [Suh and Devadas 2007; Holcomb and Fu 2014]; however, side-channel attacks can still break PUFs [Rostami et al. 2014]. Therefore, improvements have been made in PUF design to reduce threats from side-channel attacks [Shiozaki et al. 2015; Rührmair et al. 2014]. Scan chains are accessible for testing purposes, yet the information from the scan chain combined with other side-channel leakage presents a great security threat. Several techniques have been developed to secure the scan chain by controlling the accessibility and content of the chain such that an attacker cannot easily gain access to the scan chain and the

information extracted would not be useful for breaking the hardware system [Fan et al. 2014; Oya et al. 2014].

In addition to the countermeasures for specific attack types, enhanced security protocols have been developed for IoT systems, such as the end-to-end security protocol in Sahraoui and Bilami [2014] and the heterogeneous non-IP security protocol in Giuliano et al. [2014]. As for homomorphic encryption, Gentry [2009] proposed a fully homomorphic encryption scheme that enables us to evaluate arbitrary functions over encrypted data without being able to decrypt. A double decryption method was proposed by Yang et al. [2014] to satisfy a fully or targeted fully homomorphic property. Several security frameworks were developed for sensor networks, such as SensorSafe [Choi et al. 2012] and SensorAct [Arjunan et al. 2012]. For resource-limited IoT devices, minimal trusted computing base and lightweight cryptography are desirable as security solutions. For instance, SANCUS provides a low hardware cost security architecture for networked embedded devices [Noorman et al. 2013]. Advanced algorithms have been proposed to accelerate partial homomorphic and order-preserving encryptions for IoT devices [Shafagh et al. 2015].

4. CONCLUSION

The IoT will consist of billions of heterogeneous sensing, processing, and actuation nodes. This high degree of heterogeneity and complexity leads to the urgent need for a model rich in expressivity and capable of unifying the previously fragmented research from a system-of-systems perspective. In response to such need, this work proposes a mathematical model characterizing the IoT system as multiscale time-dependent random hypergraphs. Based on the proposed model, we identified a list of fundamental research challenges in sensing, the computing paradigm, robustness, energy efficiency, and hardware security.

REFERENCES

- B. Abinaya and S. Sophia. 2015. An event based CMOS quad bilateral combination with asynchronous SRAM architecture based neural network using low power. In *Proceedings of the 2015 2nd International Conference on Electronics and Communication Systems (ICECS'15)*. IEEE, Los Alamitos, CA, 995–999.
- Charu C. Aggarwal, Naveen Ashish, and Amit P. Sheth. 2013. The Internet of Things: A survey from the data-centric perspective. In *Managing and Mining Sensor Data*. Springer, 383–428.
- Pandarasamy Arjunan, Nipun Batra, Haksoo Choi, Amarjeet Singh, Pushpendra Singh, and Mani B. Srivastava. 2012. SensorAct: A privacy and security aware federated middleware for building management. In *Proceedings of the 4th ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*. ACM, New York, NY, 80–87.
- Manish Arora, Srilatha Manne, Yasuko Eckert, Indrani Paul, Nuwan Jayasena, and Dean Tullsen. 2014. A comparison of core power gating strategies implemented in modern hardware. In *Proceedings of the 2014 ACM International Conference on Measurement and Modeling of Computer Systems*. ACM, New York, NY, 559–560.
- Hamed Babazadeh, Babak Asghari, and Ritu Sharma. 2014. A new control scheme in a multi-battery management system for expanding microgrids. In *Proceedings of the 2014 IEEE PES Innovative Smart Grid Technologies Conference (ISGT'14)*. IEEE, Los Alamitos, CA, 1–5.
- Emmanuel Baccelli, Christian Mehlis, Oliver Hahm, Thomas C. Schmidt, and Matthias Wählisch. 2014. Information centric networking in the IoT: Experiments with NDN in the wild. arXiv:1406.6608.
- Alessandro Bardine, Manuel Comparetti, Pierfrancesco Foglia, and Cosimo Antonio Prete. 2014. Evaluation of leakage reduction alternatives for deep submicron dynamic nonuniform cache architecture caches. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 22, 1, 185–190.
- Luigi Battista, Laura Mecozzi, Sara Coppola, Veronica Vespini, Simonetta Grilli, and Pietro Ferraro. 2014. Graphene and carbon black nano-composite polymer absorbers for a pyro-electric solar energy harvesting device based on LiNbO₃ crystals. *Applied Energy* 136, 357–362.
- Paul Bogdan. 2015. A cyber-physical systems approach to personalized medicine: Challenges and opportunities for NOC-based multicore platforms. In *Proceedings of the 2015 Design, Automation, and Test in Europe Conference and Exhibition*. 253–258.

- Paul Bogdan and Yuankun Xue. 2015. Cyber-physical systems for personalized and precise medicine. In *Proceedings of the 2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWS-CAS'15)*. IEEE, Los Alamitos, CA, 1–4.
- Flavio Bonomi, Rodolfo Milito, Jiang Zhu, and Sateesh Addepalli. 2012. Fog computing and its role in the Internet of Things. In *Proceedings of the 1st Edition of the MCC Workshop on Mobile Cloud Computing*. ACM, New York, NY, 13–16.
- Benjamin A. Braun, Suman Jana, and Dan Boneh. 2015. Robust and efficient elimination of cache and timing side channels. arXiv:1506.00189.
- B. Brunekreef, R. M. J. Beelen, G. Hoek, L. J. Schouten, S. Bausch-Goldbohm, P. Fischer, B. Armstrong, et al. 2009. Effects of long-term exposure to traffic-related air pollution on respiratory and cardiovascular mortality in the Netherlands: The NLCS-AIR study. *Research Report (Health Effects Institute)* 139139, 5–71.
- H. G. Canepa and D. Lee. 2010. A virtual cloud computing provider for mobile devices. In *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing and Services: Social Networks and Beyond (MSC'10)*. Article No. 6.
- Aranya Chakraborty and Marija D. Ilic. 2012. *Control and Optimization Methods for Electric Smart Grids*. Springer.
- Meng-Fan Chang, Chien-Fu Chen, Ting-Hao Chang, Chi-Chang Shuai, Yen-Yao Wang, and Hiroyuki Yamauchi. 2015. 17.3: A 28nm 256kb 6T-SRAM with 280mV improvement in VMIN using a dual-split-control assist scheme. In *Proceedings of the 2015 IEEE International Solid-State Circuits Conference (ISSCC'15)*. IEEE, Los Alamitos, CA, 1–3.
- Naehyuck Chang, Donkyu Baek, and Jeongmin Hong. 2014. Power consumption characterization, modeling and estimation of electric vehicles. In *Proceedings of the 2014 IEEE/ACM International Conference on Computer-Aided Design (ICCAD'14)*. IEEE, Los Alamitos, CA, 175–182.
- Heng-Tze Cheng, Feng-Tso Sun, Senaka Buthpitiya, and Martin Griss. 2010. SensOrchestra: Collaborative sensing for symbolic location recognition. In *Mobile Computing, Applications, and Services*. Springer, 195–210.
- Haksoo Choi, Supriyo Chakraborty, and Mani B. Srivastava. 2012. Design and evaluation of SensorSafe: A framework for achieving behavioral privacy in sharing personal sensory information. In *Proceedings of the 2012 IEEE 11th International Conference on Trust, Security, and Privacy in Computing and Communications (TrustCom'12)*. IEEE, Los Alamitos, CA, 1004–1011.
- Richard Cimler, Jan Matyska, and Vladimir Sobeslav. 2014. Cloud based solution for mobile healthcare application. In *Proceedings of the 18th International Database Engineering and Applications Symposium*. ACM, New York, NY, 298–301.
- Andrew Clark, Basel Alomair, Linda Bushnell, and Radha Poovendran. 2016. *Submodularity in Dynamics and Control of Networked Systems*. Springer.
- Ronald P. Cocchi, James P. Baukus, Lap Wai Chow, and B. Jiangyun Wang. 2014. Circuit camouflage integration for hardware IP protection. In *Proceedings of the 2014 51st ACM/EDAC/IEEE Design Automation Conference (DAC'14)*. IEEE, Los Alamitos, CA, 1–5.
- Cisco. 2015. Build a Secure, Manageable Network for Tomorrow, Today. Retrieved January 28, 2017, from <http://www.cisco.com/web/solutions/trends/iot/security.html>.
- Soumya Kanti Datta, Christian Bonnet, and Navid Nikaein. 2014. An IoT gateway centric architecture to provide novel M2M services. In *Proceedings of the 2014 IEEE World Forum on Internet of Things (WF-IoT'14)*. IEEE, Los Alamitos, CA, 514–519.
- Jeroen Delvaux and Ingrid Verbauwhede. 2014. Fault injection modeling attacks on 65 nm arbiter and RO sum PUFs via environmental changes. *IEEE Transactions on Circuits and Systems I: Regular Papers* 61, 6, 1701–1713.
- John Demme, Robert Martin, Adam Waksman, and Simha Sethumadhavan. 2012. Side-channel vulnerability factor: A metric for measuring information leakage. *ACM SIGARCH Computer Architecture News* 40, 106–117.
- Caiwen Ding, Soroush Heidari, Yanzhi Wang, Yongpan Liu, and Jingtong Hu. 2016. Multi-source in-door energy harvesting for non-volatile processors. In *Proceedings of the 2016 IEEE International Symposium on Circuits and Systems (ISCAS'16)*. IEEE, Los Alamitos, CA, 173–176.
- Mohamed El Massad, Siddharth Garg, and Mahesh V. Tripunitara. 2015. Integrated circuit (IC) decamouflaging: Reverse engineering camouflaged ICs within minutes. In *Proceedings of the 2015 Network and Distributed System Security Symposium (NDSS'15)*.
- Nazar Fadhil, Dleer Saber, Paris Cox, Kripa Vanashi, and Prabir Patra. 2014. Energy Harvesting Using Nano-Fibers PVDF/Graphene Composite for Medical Implanted Devices. Available at <https://scholarworks.bridgeport.edu/xmlui/handle/123456789/644>.

- Junfeng Fan, Hua Xie, and Yiwei Zhang. 2014. On the use of scan chain to improve physical attacks. In *Proceedings of the 2014 IEEE 23rd Asian Test Symposium (ATS'14)*. IEEE, Los Alamitos, CA, 354–357.
- Xin Fang, Pei Luo, Yunsu Fei, and Miriam Leeser. 2015. Balance power leakage to fight against side-channel analysis at gate level in FPGAs. In *Proceedings of the 2015 IEEE 26th International Conference on Application-Specific Systems, Architectures, and Processors (ASAP'15)*. IEEE, Los Alamitos, CA, 154–155.
- Hamza Fawzi, Paulo Tabuada, and Suhas Diggavi. 2014. Secure estimation and control for cyber-physical systems under adversarial attacks. *IEEE Transactions on Automatic Control* 59, 6, 1454–1467.
- Jonathan Frey and Qiaoyan Yu. 2015. Exploiting state obfuscation to detect hardware trojans in NoC network interfaces. In *Proceedings of the 2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS'15)*. IEEE, Los Alamitos, CA, 1–4.
- C. Gentry. 2009. *A Fully Homomorphic Encryption Scheme*. Ph.D. Dissertation. Stanford University, Stanford, CA.
- Mahboobeh Ghorbani and Paul Bogdan. 2013. A cyber-physical system approach to artificial pancreas design. In *Proceedings of the 9th IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*. IEEE, Los Alamitos, CA, 17.
- Romeo Giuliano, Franco Mazzenga, Alessandro Neri, and Anna Maria Vegni. 2014. Security access protocols in IoT networks with heterogenous non-IP terminals. In *Proceedings of the 2014 IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS'14)*. IEEE, Los Alamitos, CA, 257–262.
- Pietro Gonizzi, Gianluigi Ferrari, Vincent Gay, and Jérémie Leguay. 2015. Data dissemination scheme for distributed storage for IoT observation systems at large scale. *Information Fusion* 22, 16–25.
- Ragib Hasan, Md Mahmud Hossain, and Rasib Khan. 2015. Aura: An IoT based cloud infrastructure for localized mobile computation outsourcing. In *Proceedings of the 2015 3rd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud'15)*. IEEE, Los Alamitos, CA, 183–188.
- Soroush Heidari, Caiwen Ding, Yongpan Liu, Yanzhi Wang, and Jingtong Hu. 2015. Multi-source energy harvesting management and optimization for non-volatile processors. In *Proceedings of the 2015 6th International Green Computing Conference and Sustainable Computing Conference (IGSC'15)*. IEEE, Los Alamitos, CA, 1–2.
- R. Herring, A. Hofleitner, S. Amin, T. Nasr, A. Khalek, P. Abbeel, and A. Bayen. 2010. Using mobile phones to forecast arterial traffic through statistical learning. In *Proceedings of the 89th Transportation Research Board Annual Meeting*.
- G. Hills, M. Shulaker, H. Wei, H.-Y. Chen, H.-S. Philip Wong, and S. Mitra. 2014. Robust design and experimental demonstrations of carbon nanotube digital circuits. In *Proceedings of the 2014 IEEE Custom Integrated Circuits Conference (CICC'14)*. IEEE, Los Alamitos, CA, 1–4.
- Doan B. Hoang and Lingfeng Chen. 2010. Mobile cloud for assistive healthcare (MoCAsH). In *Proceedings of the 2010 IEEE Asia-Pacific Services Computing Conference (APSCC'10)*. IEEE, Los Alamitos, CA, 325–332.
- Daniel E. Holcomb and Kevin Fu. 2014. Bitline PUF: Building native challenge-response PUF capability into any SRAM. In *Cryptographic Hardware and Embedded Systems—CHES 2014*. Lecture Notes in Computer Science, Vol. 8731. Springer, 510–526.
- Infineon. 2015. The Right Security for the Internet of Things (IoT). Retrieved January 28, 2017, from <http://www.infineon.com/cms/en/applications/chip-card-security/internet-of-things-security>.
- Rudolf Emil Kalman. 1963. Mathematical description of linear dynamical systems. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control* 1, 2, 152–192.
- Ramesh Karri, Farinaz Koushanfar, Ozgur Sinanoglu, Yiorgos Makris, Ken Mai, Ahmad Reza Sadeghi, and Swarup Bhunia. 2015. Guest editorial special section on hardware security and trust. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 34, 6, 873–874.
- Youhei Kawahara, Ryouhei Nakao, Akihiko Kudo, Akihiko Emori, and Kei Sakabe. 2014. Battery control circuit for reducing power consumption. US Patent 8,736,229.
- Sungwook Kim. 2015. Nested game-based computation offloading scheme for mobile cloud IoT systems. *EURASIP Journal on Wireless Communications and Networking* 2015, 1, 1–11.
- Christian Kison, Jürgen Frinken, and Christof Paar. 2015. Finding the AES bits in the haystack: Reverse engineering and SCA using voltage contrast. In *Cryptographic Hardware and Embedded Systems—CHES 2015*. Lecture Notes in Computer Science, Vol. 9293. Springer, 641–660.
- Bill Kleymann. 2013. Welcome to Fog Computing: Extending the Cloud to the Edge. Retrieved January 28, 2017, from <http://www.datacenterknowledge.com/archives/2013/08/23/welcome-to-the-fog-a-new-type-of-distributed-computing>.

- Alicia Klinefelter, Nathan E. Roberts, Yousef Shakhsheer, Patricia Gonzalez, Aatmesh Shrivastava, Abhishek Roy, Kyle Craig, et al. 2015. 21.3: A $6.45\mu\text{W}$ self-powered IoT SoC with integrated energy-harvesting power management and ULP asymmetric radios. In *Proceedings of the 2015 IEEE International Solid-State Circuits Conference (ISSCC'15)*. IEEE, Los Alamitos, CA, 1–3.
- Jakub Korczyk and Andrzej Krasniewski. 2012. Evaluation of susceptibility of FPGA-based circuits to fault injection attacks based on clock glitching. In *Proceedings of the 2012 IEEE 15th International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS'15)*. IEEE, Los Alamitos, CA, 171–174.
- Andreas Krause and Carlos E. Guestrin. 2012. Near-optimal nonmyopic value of information in graphical models. arXiv:1207.1394.
- Karthik Kumar and Yung-Hsiang Lu. 2010. Cloud computing for mobile users: Can offloading computation save energy? *Computer* 44, 51–56.
- Santosh Kumar, Wendy Nilsen, Misha Pavel, and Mani Srivastava. 2013. Mobile health: Revolutionizing healthcare through transdisciplinary research. *Computer* 11, 28–35.
- Tung-Wei Kuo, Kate Ching-Ju Lin, and Ming-Jer Tsai. 2015. Maximizing submodular set function with connectivity constraint: Theory and application to networks. *IEEE/ACM Transactions on Networking* 23, 2, 533–546.
- Nicholas D. Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T. Campbell. 2010. A survey of mobile phone sensing. *IEEE Communications Magazine* 48, 9, 140–150.
- Timo Lanki, Regina Hampel, Pekka Tiittanen, Silke Andrich, Rob Beelen, Bert Brunekreef, Julia Dratva, et al. 2015. Air pollution from road traffic and systemic inflammation in adults: A cross-sectional analysis in the European ESCAPE project. *Environmental Health Perspectives* 123, 8, 785–791.
- Hyung Gyu Lee and Naehyuck Chang. 2015. Powering the IoT: Storage-less and converter-less energy harvesting. In *Proceedings of the 20th Asia and South Pacific Design Automation Conference*. IEEE, Los Alamitos, CA, 124–129.
- Liran Lerman, Gianluca Bontempi, and Olivier Markowitch. 2014. Power analysis attack: An approach based on machine learning. *International Journal of Applied Cryptography* 3, 2, 97–115.
- Philip Levis. 2015. Securing the Internet of Things. Retrieved January 28, 2017, from <http://iot.stanford.edu/seminar/sitp-w15-levis.pdf>.
- Ji Li, Yanzhi Wang, Xue Lin, Shahin Nazarian, and Massoud Pedram. 2015b. Negotiation-based task scheduling and storage control algorithm to minimize user's electric bills under dynamic prices. In *Proceedings of the 2015 20th Asia and South Pacific Design Automation Conference (ASP-DAC'15)*. IEEE, Los Alamitos, CA, 261–266.
- Ji Li, Yanzhi Wang, Xue Lin, Shahin Nazarian, and Massoud Pedram. 2016. Negotiation-based resource provisioning and task scheduling algorithm for cloud systems. In *Proceedings of the 2016 17th International Symposium on Quality Electronic Design (ISQED'16)*. IEEE, Los Alamitos, CA, 338–343.
- Ji Li, Qing Xie, Yanzhi Wang, Shahin Nazarian, and Massoud Pedram. 2015. Leakage power reduction for deeply-scaled FinFET circuits operating in multiple voltage regimes using fine-grained gate-length biasing technique. In *Proceedings of the 2015 Design, Automation, and Test in Europe Conference and Exhibition*. 1579–1582.
- Shancang Li, Li Da Xu, and Xinheng Wang. 2013. Compressed sensing signal and data acquisition in wireless sensor networks and Internet of Things. *IEEE Transactions on Industrial Informatics* 9, 4, 2177–2186.
- Wenzhong Li, Yuefei Hu, Xiaoming Fu, Sanglu Lu, and Daoxu Chen. 2015a. Cooperative positioning and tracking in disruption tolerant networks. *IEEE Transactions on Parallel and Distributed Systems* 26, 2, 382–391.
- Xiaofei Liu, Sergio Pequito, Soumya Kar, Bruno Sinopoli, and A. Pedro Aguiar. 2015. Minimum sensor placement for robust observability of structured complex networks. arXiv:1507.07205.
- Xiaowei Liu, Honglin Xu, Liang Yin, Zhiqiang Gao, and Mingyuan Ren. 2014. A low noise sigma-delta microaccelerometer interface circuit. *IEICE Electronics Express* 11, 12, 1–6.
- Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. 2013a. Observability of complex systems. *Proceedings of the National Academy of Sciences* 110, 7, 2460–2465.
- Lucia Marucci, David A. W. Barton, Irene Cantone, Maria Aurelia Ricci, Maria Pia Cosma, Stefania Santini, Diego di Bernardo, and Mario di Bernardo. 2009. How to turn a genetic circuit into a synthetic tunable oscillator, or a bistable switch. *PLoS One* 4, 12, e8083.
- Daniel H. Morris, Uygur E. Avci, and Ian A. Young. 2015. Variation-tolerant dense TFET memory with low V MIN matching low-voltage TFET logic. In *Proceedings of the 2015 Symposium on VLSI Technology (VLSI Technology'15)*. IEEE, Los Alamitos, CA, T24–T25.

- David E. Noel and M. Lisa Miller. 2012. A risk assessment technique: For predicting the magnitude of satellite programming theft. *Ethics and Critical Thinking Journal* 2012, 1, 28–37.
- Job Noorman, Pieter Agten, Wilfried Daniels, Raoul Strackx, Anthony Van Herreweghe, Christophe Huygens, Bart Preneel, Ingrid Verbauwhede, and Frank Piessens. 2013. Sancus: Low-cost trustworthy extensible networked devices with a zero-software trusted computing base. In *Proceedings of the 22nd USENIX Security Symposium (USENIX Security'13)*. 479–498.
- David Oswald, Daehyun Strobel, Falk Schellenberg, Timo Kasper, and Christof Paar. 2014. When reverse-engineering meets side-channel analysis—digital lockpicking in practice. In *Selected Areas in Cryptography—SAC 2013*. Lecture Notes in Computer Science, Vol. 8282. Springer, 571–588.
- Masaru Oya, Yuta Atobe, Youhua Shi, Masao Yanagisawa, and Nozomu Togawa. 2014. Secure scan design using improved random order and its evaluations. In *Proceedings of the 2014 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS'14)*. IEEE, Los Alamitos, CA, 555–558.
- Hewlett Packard. 2015. Internet of Things Research Study. Retrieved January 28, 2017, from <https://www.hpe.com/h20195/v2/GetPDF.aspx/4AA5-4759ENN.pdf>.
- Miroslav Pajic, Paulo Tabuada, Insup Lee, and George J. Pappas. 2015. Attack-resilient state estimation in the presence of noise. In *Proceedings of the 2015 54th IEEE Conference on Decision and Control (CDC'15)*. IEEE, Los Alamitos, CA, 5827–5832.
- Sergio Pequito, Paul Bogdan, and George J. Pappas. 2015. Minimum number of probes for brain dynamics observability. In *Proceedings of the 2015 54th IEEE Conference on Decision and Control (CDC'15)*. IEEE, Los Alamitos, CA, 306–311.
- Sergio Pequito, Soumya Kar, and A. Pedro Aguiar. 2014. Minimum number of information gatherers to ensure full observability of a dynamic social network: A structural systems approach. In *Proceedings of the 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP'14)*. IEEE, Los Alamitos, CA, 750–753.
- Hauke Petersen, Emmanuel Baccelli, Matthias Wählisch, Thomas C. Schmidt, and Jochen Schiller. 2014. The role of IoT in network resilience. arXiv:1406.6614.
- Nathaniel Pinckney, Lucian Shifren, Brian Cline, Saurabh Sinha, Supreet Jeloka, Ronald G. Dreslinski, Trevor Mudge, Dennis Sylvester, and David Blaauw. 2016. Near-threshold computing in FinFET technologies: Opportunities for improved voltage scalability. In *Proceedings of the 53rd Annual Design Automation Conference*. ACM, New York, NY, 76.
- Henrich C. Pohls, Vangelis Angelakis, Santiago Suppan, Kai Fischer, George Oikonomou, Elias Z. Tragos, Rodrigo Diaz Rodriguez, and Theodoros Mouroutis. 2014. RERUM: Building a reliable IoT upon privacy- and security-enabled smart objects. In *Proceedings of the 2014 IEEE Wireless Communications and Networking Conference Workshops (WCNCW'14)*. IEEE, Los Alamitos, CA, 122–127.
- Owen Poindexter. 2014. The Internet of Things Will Thrive on Energy Efficiency. Retrieved January 28, 2017, from <http://www.govtech.com/fs/news/The-Internet-of-Things-Will-Thrive-On-Energy-Efficiency-.html>.
- Indika Rajapakse, Mark Groudine, and Mehran Mesbahi. 2012. What can systems theory of networks offer to biology? *PLoS Computational Biology* 8, 6, e1002543.
- Craig Rawlings. 2008. Raising the bar for hardware security: Physical layer security in standard CMOS. In *Proceedings of the 2008 IEEE Conference on Technologies for Homeland Security*. IEEE, Los Alamitos, CA, 263–268.
- Mohamad Rostami, Farinaz Koushanfar, and Ramesh Karri. 2014. A primer on hardware security: Models, methods, and metrics. *Proceedings of the IEEE* 102, 8, 1283–1295.
- Ulrich Rüßmair, Ulf Schlichtmann, and Wayne Burleson. 2014. Special session: How secure are PUFs really? On the reach and limits of recent PUF attacks. In *Proceedings of the Conference on Design, Automation, and Test in Europe*. 346.
- Somia Sahraoui and Azeddine Bilami. 2014. Compressed and distributed host identity protocol for end-to-end security in the IoT. In *Proceedings of the 2014 5th International Conference on Next Generation Networks and Services (NGNS'14)*. IEEE, Los Alamitos, CA, 295–301.
- I. Sengupta and A. Das Barman. 2015. Investigation of inter-channel crosstalk mitigation by assist light in a TOAD switch by using electrical equivalent circuit model of SOA. *Optical and Quantum Electronics* 47, 2, 331–350.
- Hossein Shafagh, Anwar Hithnawi, Andreas Dröschner, Simon Duquennoy, and Wen Hu. 2015. Poster: Towards encrypted query processing for the Internet of Things. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, New York, NY, 251–253.
- Mitsuru Shiozaki, Takaya Kubota, Tsunato Nakai, Akihiro Takeuchi, Takashi Nishimura, and Takeshi Fujino. 2015. Tamper-resistant authentication system with side-channel attack resistant AES and PUF using MDR-ROM. In *Proceedings of the 2015 IEEE International Symposium on Circuits and Systems (ISCAS'15)*. IEEE, Los Alamitos, CA, 1462–1465.

- Sergei Skorobogatov. 2014. Tamper resistance and hardware security. Guest lecture in the Part II Security course.
- Yuning Song, Huadong Ma, and Liang Liu. 2013. Content-centric internetworking for resource-constrained devices in the Internet of Things. In *Proceedings of the 2013 IEEE International Conference on Communications (ICC'13)*. IEEE, Los Alamitos, CA, 1742–1747.
- John A. Stankovic. 2014. Research directions for the Internet of Things. *IEEE Internet of Things Journal* 1, 1, 3–9.
- An Steegen. 2015. Technology innovation in an IoT era. In *Proceedings of the 2015 Symposium on VLSI Technology (VLSI Technology'15)*. IEEE, Los Alamitos, CA, C170–C171.
- Franco Stellari, Peilin Song, and Herschel Ainspan. 2014. Functional block extraction for hardware security detection using time-integrated and time-resolved emission measurements. In *Proceedings of the 2014 IEEE 32nd VLSI Test Symposium (VTS'14)*. IEEE, Los Alamitos, CA, 1–6.
- G. Edward Suh and Srinivas Devadas. 2007. Physical unclonable functions for device authentication and secret key generation. In *Proceedings of the 44th Annual Design Automation Conference*. ACM, New York, NY, 9–14.
- Arvind Thiagarajan, Lenin Ravindranath, Katrina LaCurts, Samuel Madden, Hari Balakrishnan, Sivan Toledo, and Jakob Eriksson. 2009. VTrack: Accurate, energy-aware road traffic delay estimation using mobile phones. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*. ACM, New York, NY, 85–98.
- Narseo Vallina-Rodriguez and Jon Crowcroft. 2011. ErdOS: Achieving energy savings in mobile OS. In *Proceedings of the 6th International Workshop on MobiArch (MobiArch'11)*. ACM, New York, NY, 37–42.
- Hong Sheng Wang, Dao Gang Ji, Yang Zhang, Kai Yan Chen, and Kai Song. 2014. Optical fault injection attacks against cipher chip. In *Advanced Materials Research*, Vol. 1044. Trans Tech Publications, 1498–1502.
- David J. Watson, James Bielman, Phillip L. Barrett, and Nicole A. Hamilton. 2013. System and method for automatically creating a media archive from content on a recording medium. US Patent 8,582,954.
- Sheng Wei, James B. Wendt, Ani Nahapetian, and Miodrag Potkonjak. 2014. Reverse engineering and prevention techniques for physical unclonable functions using side channels. In *Proceedings of the 51st Annual Design Automation Conference*. ACM, New York, NY, 1–6.
- James B. Wendt and Miodrag Potkonjak. 2014. Hardware obfuscation using PUF-based logic. In *Proceedings of the 2014 IEEE/ACM International Conference on Computer-Aided Design*. IEEE, Los Alamitos, CA, 270–277.
- Chia-Ming Wu, Ruay-Shiung Chang, and Hsin-Yu Chan. 2014. A green energy-efficient scheduling algorithm using the DVFS technique for cloud datacenters. *Future Generation Computer Systems* 37, 141–147.
- Huaming Wu, Qiushi Wang, and Katinka Wolter. 2013. Mobile healthcare systems with multi-cloud off-loading. In *Proceedings of the 2013 IEEE 14th International Conference on Mobile Data Management (MDM'13)*, Vol. 2. IEEE, Los Alamitos, CA, 188–193.
- Kai-Chiang Wu and Diana Marculescu. 2014. Power-planning-aware soft error hardening via selective voltage assignment. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 22, 1, 136–145.
- Sheng Xiao, Weibo Gong, and Don Towsley. 2014. Dynamic key management in a smart grid. In *Dynamic Secrets in Communication Security*. Springer, 55–68.
- Teng Xu, James B. Wendt, and Miodrag Potkonjak. 2014. Security of IoT systems: Design challenges and opportunities. In *Proceedings of the 2014 IEEE/ACM International Conference on Computer-Aided Design*. IEEE, Los Alamitos, CA, 417–423.
- Yuankun Xue, Sergio Pequito, Joana R. Coelho, Paul Bogdan, and George J. Pappas. 2016a. Minimum number of sensors to ensure observability of physiological systems: A case study. In *Proceedings of the 2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, Los Alamitos, CA.
- Yuankun Xue, Saul Rodriguez, and Paul Bogdan. 2016b. A spatio-temporal fractal model for a CPS approach to brain-machine-body interfaces. In *Proceedings of the 2016 Design, Automation, and Test in Europe Conference and Exhibition (DATE'16)*. IEEE, Los Alamitos, CA, 642–647.
- Farah B. Yahya, Harsh N. Patel, Vikas Chandra, and Benton H. Calhoun. 2015. Combined SRAM read/write assist techniques for near/sub-threshold voltage operation. In *Proceedings of the 2015 6th Asia Symposium on Quality Electronic Design (ASQED'15)*. IEEE, Los Alamitos, CA, 1–6.
- Yatao Yang, Shuang Zhang, Junming Yang, Jia Li, and Zichen Li. 2014. Targeted fully homomorphic encryption based on a double decryption algorithm for polynomials. *Tsinghua Science and Technology* 19, 5, 478–485.

- Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, and Michele Zorzi. 2014. Internet of Things for smart cities. *IEEE Internet of Things Journal* 1, 1, 22–32.
- Jiliang Zhang. 2015. A practical logic obfuscation technique for hardware security. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 24, 3, 1193–1197.
- Zhi-Kai Zhang, Michael Cheng Yi Cho, Chia-Wei Wang, Chia-Wei Hsu, Chong-Kuan Chen, and Shiuhyng Shieh. 2014. IoT security: Ongoing challenges and research opportunities. In *Proceedings of the 2014 IEEE 7th International Conference on Service-Oriented Computing and Applications (SOCA'14)*. IEEE, Los Alamitos, CA, 230–234.
- Jizhong Zhu. 2015. *Optimization of Power System Operation*. Vol. 47. John Wiley & Sons.
- Yi Zhuang, Nan Jiang, Zhiang Wu, Qing Li, Dickson K. W. Chiu, and Hua Hu. 2014. Efficient and robust large medical image retrieval in mobile cloud computing environment. *Information Sciences* 263, 60–86.

Received June 2016; revised September 2016; accepted September 2016