1

Internet of Things-Enabled Systems and Infrastructure

1.1 Cyber–Physical Realm of IoT

Network-connected electronic devices are becoming an essential part of modern infrastructure systems to automate manual processes resulting in improved efficiency and productivity. The Internet of Things (IoT) is an interconnection of different types of devices (classified as sensors and actuators) using communication networks and computing systems to achieve such automated operation. The difference in IoT from traditional computing systems is their interaction with the physical world as opposed to just the cyber world. For instance, we have electronic devices controlling the temperature in smart buildings by sensing the environment and operating the heating, ventilation, and air conditioning (HVAC) systems. The IoT is in fact a massive network of cyber–physical systems (CPSs). Therefore, the cyber and physical components are an integral part of the emerging IoT ecosystem. The cyber and physical systems are coupled together in an intricate fashion where the cyber world influences decisions in the physical world and vice versa.

Figure 1.1 shows the structure of a typical IoT system. In essence, there are several actors involved in setting up the IoT ecosystem that includes sensing/actuating devices, firmware, radio access network (RAN), cloud server, mobile apps, and end user devices. The endpoint devices are made of embedded hardware that interact with the physical environment and are driven by software processes referred to as firmware or operating system. They make use of communication infrastructure, which is composed of access points, gateways, and core IP networks to connect to cloud servers, that in turn host applications and services, which are operated by users via computing devices, such as smart phones, smart watches, and voice assistants, etc.

Resource Management for On-Demand Mission-Critical Internet of Things Applications, First Edition.
Junaid Farooq and Quanyan Zhu.

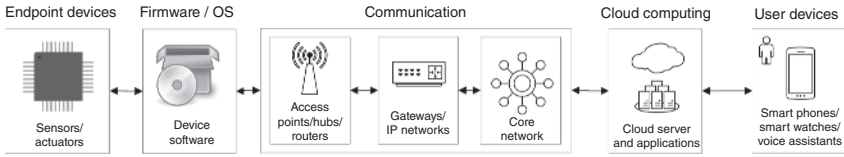

Figure 1.1 IoT technology stack describing the different actors involved in setting up the IoT ecosystem.

1.2 IoT in Mission-Critical Applications

IoT systems have a wide variety of application areas. Some of the IoT applications are highly delay-sensitive, e.g. real-time systems such as those involving artificial intelligence (AI), virtual reality (VR) and augmented reality (AR), real-time control loops, streaming analytics, etc. [121]. Such applications are referred to as *mission-critical* [158] not only due to conventional "life risk" interpretation but also pertaining to the risks of interruption of public services interruption, perturbing public order, jeopardizing enterprise operation and causing losses to businesses, etc. In mission-critical IoT (MC-IoT) applications, often a delay in communication may in fact fail the initial objective of the application. For instance, in a surveillance system where an unusual activity needs to be reported promptly to avoid any potential damage or loss of property and a report beyond a certain delay may be futile. Nevertheless, the traditional MC definition still holds and more so since IoT is also being rapidly integrated into these systems such as in public safety systems or other emergency networks [38] requiring dedicated resources at all times due to the unpredictability of unforeseen events.

1.3 Overview of the Book

The book is organized into six main parts. Part I provides a high level description of the IoT ecosystem and its main features; Par II provides an overview of the main design challenges facing the IoT systems and networks across different layers; Part III investigates and addresses the wireless connectivity challenge faced by the endpoint devices; Part IV tackles the networking layer challenges including information dissemination and message propagation over networks. Part V deals with the service provisioning and resource allocation problems in IoT for mission-critical service delivery. Finally, Part VI provides a broader view of the impact of this work along with a vision governing future research in this domain. Each part is further organized into multiple chapters. The main contributions of the book along with references to the relevant chapters are summarized in the following subsection.

1.3.1 Main Topics

This book takes a clean slate approach toward the design of IoT-enabled systems and networks. It uses a cross-layer perspective in decision-making across various avenues in the IoT ecosystem. Some of the main topics are presented in the following subsections.

1.3.1.1 Dynamic Reservation of Wireless Spectrum Resources

In Chapter 5, a dynamic mechanism for spectrum reservation considering the uncertainty in available spectrum at each time and the uncertainty in the requirement for spectrum access is developed. We make use of tools from sequential screening [26] and mechanism design literature to establish a dynamic menu of contracts which comprise of an advanced payment for spectrum reservation in the future along with a rebate policy if the spectrum is released before the time of spectrum access. This allows the network operator to discriminate the unknown application types and generate higher profits than the traditional auction mechanisms where every application is completely aware of its true utility. A two-type categorization of IoT applications is considered where they are classified as either MC or non-MC and consequently an optimal binary contract is designed by the service provider. Based on assumptions on the distribution of utility of the MC and non-MC applications, closed form results for the optimal contracts are derived and the effect of system parameters is analyzed to gain insights.

1.3.1.2 Dynamic Cross-Layer Connectivity Using Aerial Networks

In Chapter 6, a dynamic approach is used to configure robotic network nodes such as unmanned aerial vehicles (UAVs) to provide connectivity to IoT devices. Although the existing methodologies provide optimization based approaches to the UAV placement problem, this problem is dynamic in nature and hence a more holistic approach is required to obtain an efficient placement of the UAVs in real-time. In addition to effective initial deployment of UAVs, there is a need for an autonomic, self-organizing, and self-healing overlay network that can continuously adapt and reconfigure according to the constantly changing network conditions [23]. Therefore, a distributed and dynamic approach to providing resilient connectivity is essential to cope with the growing scale of the networks toward a massive IoT [44]. To this end, this book develops a feedback based distributed cognitive framework that maintains connectivity of the network and is resilient to the mobility of ground users and/or failures of the UAVs. The continuous feedback enables the framework to actively react to network changes and appropriately reconfigure the network in response to a failure event that has resulted in loss of connectivity. Simulation results demonstrate that if sufficient UAVs are available, they can be arranged into a desired configuration from arbitrary initial positions

and the configuration continuously adapts according to the movement of the ground users as well as recovers connectivity under varying levels of a random UAV failure event.

1.3.1.3 Dynamic Processes Over Multiplex Spatial Networks and Reconfigurable Design

In Chapter 7, a stochastic geometry (SG) based model is used to characterize the connectivity of wireless networks in adversarial environments such as battlefields. We then use an epidemic spreading model to capture the dynamic diffusion of multiple messages within the network of devices at the equilibrium state. A novel multiplex network model for Internet of battlefield things (IoBT) networks is proposed that helps in characterizing the intra-layer and network-wide connectivity of heterogeneous battlefield devices by considering the spatial randomness in their locations. A tractable framework is developed for quantification of simultaneous information dissemination in the multiplex IoBT network based on mathematical epidemiology that considers the perceived level of threat to the network from cyber–physical attacks. Approximate closed form results relating the proportion of informed devices at equilibrium and the network parameters are provided. The resulting integrated open-loop system model is used as a basis for reconfiguring the network parameters to ensure a mission-driven information spreading profile in the network. An optimization problem is formulated that can assist military commanders in identifying the physical network parameters that are required in order to sufficiently secure the network from the perceived attacks. It can also help in reconfiguring existing networks to achieve a desired level of communication reliability. A detailed investigation of the developed integrated framework is provided for particular battlefield missions, and the effect of threat level and performance thresholds is studied. This book bridges the gap between the spatial stochastic models for wireless networks and the dynamic diffusion models in contact-based biological networks to derive new insights that aid in the planning and design of secure and reliable IoBT networks for mission critical information dissemination. The developed framework, with some modifications, is also applicable to the more general class of heterogeneous ad-hoc networks.

In Chapter 8, novel methodologies are proposed to overcome the unique challenges of modeling and analyzing the crucial interplay between malware infection, control commands propagation, and device patching in wireless IoT networks. We leverage ideas from the theories of dynamic population processes [70] and point processes to setup a mean field dynamical system that captures the evolution of malware infected devices and control command aware devices over time. In general, obtaining tractable characterizations of the equilibrium state in such population processes is theoretically involved due to the self-consistent nature of the equations involved and the complex connectivity

profile of the network. However, we propose a variation of the mean field population process model based on a customized state space that allows us to analyze the formation of botnets in wireless IoT networks and helps in making decisions to control its impact.

1.3.1.4 Sequential Resource Allocation Under Spatio-Temporal Uncertainties

In Chapter 9, an adaptive and resilient dynamic resource allocation and pricing framework is developed for the context of cloud-enabled IoT systems. We present an optimal dynamic policy to filter incoming service requests by IoT applications based on the complexity of the tasks. The qualification threshold for tasks is adaptive to the number of available virtual machines (VMs), the arrival rate of requests, and their average complexity. The optimal policy can be dynamically updated in order to maintain high expected revenues of the cloud service provider (CSP). Furthermore, the proposed framework is also able to adapt according to the changing availability of the VMs due to reprovisioning of resources for other applications or due to the effect of malicious attacks.

In Chapter 10, a revenue maximizing perspective toward allocation and pricing in fog based systems designed for mission critical IoT applications is proposed. The quality-of-experience (QoE) resulting from the pairing of fog resources with computation requests is used as a basis for pricing. We develop a dynamic policy framework leveraging the literature in economics, mechanism design [51], and dynamic revenue maximization [52] to provide an implementable mechanism for dynamic allocation and pricing of sequentially arriving IoT requests that maximizes the expected revenue of the CSP. The developed optimal policy framework assists in both determining which fog node to allocate an incoming task to and the price that should be charged for it for revenue maximization. The proposed policy is statistically optimal, dynamic, i.e. adapts with time, and is implementable in real-time as opposed to other static matching schemes in the literature. The dynamically optimal solution can be computed offline and implemented in real-time for sequentially arriving computation requests.

In Chapter 11, the spatio-temporal aspect is combined with incomplete information about resource requests to devise an integrated resource provisioning framework. Ideas from the stochastic assignment of sequentially arriving tasks to workers [32] are enriched to encompass a more generic utility function that also incorporates the spatial dimension of the sequentially arriving requests. Statistical properties of utility maximizing spatial service requests are characterized using spatio-temporal Poisson processes and extreme value analysis. Analysis for a generalized utility function is done based on the distance from the source as well as the magnitude of the request. Special cases of the utility function are considered for numerical evaluations. An integrated and holistic policy framework is developed that is dynamically optimal and can act as the

foundation for allocation and pricing in a wide variety of applications in the context of smart city applications. Finally, a comparison of the performance of the proposed resource provisioning framework is provided with benchmark allocation strategies.

1.3.2 Notations

The book has used a breadth of different notations for studying various theoretical models. In general, there has been an attempt to keep the notations consistent unless otherwise highlighted individually on a case-by-case basis. Some common notations used are presented as follows:

Double-struck symbols, e.g. $\mathbb{R}, \mathbb{Z}, \mathbb{N}$, etc. generally indicate sets or spaces, except $\mathbb{P}(\cdot)$, which represents the probability measure, and $\mathbb{E}[\cdot]$, which represents the expectation operator. The associated density and distribution functions are denoted by $f_X(\cdot)$ and $F_X(\cdot)$, respectively. Upper case alphabets, e.g. K, M, etc. are generally used to represent random variables, while lower case alphabets indicate the value assumed by the random variables. The operator $|\cdot|$ represents the cardinality of the set, while $\|\cdot\|$ denotes the Euclidean norm. First derivative and second derivatives are denoted by $(\cdot)'$ and $(\cdot)''$, respectively. Other notations apply to the chapters within which they are defined.