Technical Report
1239

Survey of Data Fusion in IoT

K.J. Kratkiewicz
J. White Bear
B.A. Miller

23 January 2019

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

This material is based upon work supported under Air Force Contract
No. FA8702-15-D-0001.

This report is the result of studies performed at Lincoln Laboratory, a federally funded research and development center operated by Massachusetts Institute of Technology. This material is based upon work supported under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Air Force.

Massachusetts Institute of Technology
Lincoln Laboratory

Survey of Data Fusion in IoT

K.J. Kratkiewicz
B.A. Miller

Group 58

J. White Bear

Group 45

Technical Report 1239

23 January 2019

Case 19-545

Lexington Massachusetts

This page intentionally left blank.

TABLE OF CONTENTS

This page intentionally left blank.

LIST OF FIGURES

This page intentionally left blank.

LIST OF TABLES

This page intentionally left blank.

1. INTRODUCTION

With the advent of the Internet—and, in particular, the World Wide Web—came a massive increase in the amount of data we as a society create on a daily basis. The more recent proliferation of mobile devices and online social networking has increased the amount of user-generated content to an enormous level. The current spread of Internet of Things (IoT) devices is causing an even greater acceleration, with the number of IoT devices online already outnumbering the number of people in the world. Within all of this data is information of potential value to many missions of the U.S. government. While IoT devices are plentiful, they tend to be inexpensive and are sometimes unreliable in their measurements. A prudent consumer of IoT data would be skeptical of any information gleaned from a single IoT sensor. This, however, is where the sheer volume of IoT devices in the world can provide a substantial benefit. The low cost of IoT sensors may yield relatively low performance, but also makes it feasible to collect large amounts of data over a sea of devices. While, for example, a single traffic sensor may frequently malfunction, fusing data across a large number of sensors improves robustness to unreliable measurements from any given device.

A diverse array of sensor types, application areas, and data fusion methodologies form a large body of work in the academic literature. This report provides a survey of the recent literature on fusion techniques for IoT data, with an eye toward methods that may be interesting for U.S. government analysts, enabling them to augment their data most effectively and provide the highest possible force multiplier for their analysis products.

The remainder of this report is organized as follows. Section 2 provides a brief background on the Internet of Things. Section 3 formally states the objective of the current study. The findings from several literature surveys are summarized in Section 4. Section 5 highlights recent results in relevant application areas. Finally, in order to get a set of findings that can be compared to each other, Section 6 outlines results in a particular application: indoor positioning. Section 7 provides a discussion of the implications of the current literature on future research and development. Section 8 summarizes and concludes the report.

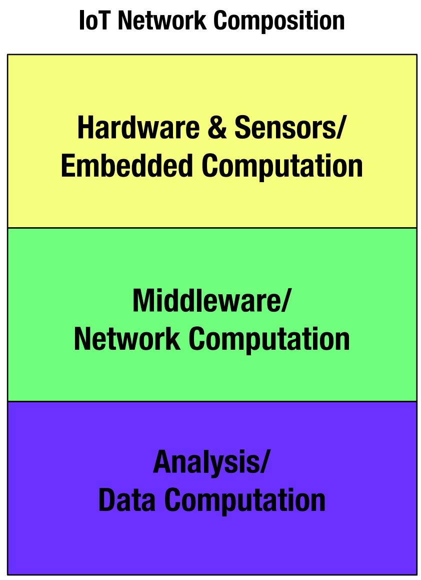
This page intentionally left blank.

2. BACKGROUND

The Internet of Things (IoT) was first notably mentioned in 1999 in private industry and subsequently incorporated into research at the Massachusetts Institute of Technology's (MIT) Auto ID Lab for use with radio frequency identification (RFID). IoT, emerging from the integration of both embedded devices and pervasive computing, aims to create a heterogeneous network architecture that interacts with the physical world. Physical objects are attached to this network as "things," which are often equipped with intelligence that can allow them to operate as autonomous agents [1]. Before the term was even coined, IoT technology was already in use in both industrial and government applications, such as Carnegie Mellon's Coke machine, which in 1982 became one of the first internet-connected appliances [2]. The overall goals of IoT have been summarized as providing global, real-time solutions and comprehensive data about the environments of attached objects [3]. However, the scale and diversity of sensors introduced since the original architecture was planned have increased a great deal. Market projections indicate about 75 billion devices will be connected by 2025 [4, 5].

Accordingly, the contemporary vision of IoT, along with the Internet and the variety of connected devices, has expanded significantly beyond its original vision of RFID and machine-to-machine (M2M) communication to address the scale and diversity of IoT networks [1]. This is an ongoing research challenge that combines problems from multiple disciplines. Starting from a hardware oriented perspective, this includes embedded devices, signal processing, network scaling and infrastructure, and energy constraints. Extracting meaning from these various sources introduces further challenges in data acquisition and downstream analysis involving integration of sensor data and analytic methods, including artificial intelligence (AI) and machine learning (ML). A firm grasp of data aggregation over the IoT space involves three key domains, as depicted in Figure 1, and their associated challenges: 1) hardware, 2) middleware, and 3) analysis.

Figure 1. IoT Network Composition



2.1 HARDWARE AND SENSORS

At the hardware level, IoT devices or "things" have the ability to access or communicate over a network. The types of devices or sensors included in the definition of "things" are not limited by any standard definition. One of the first "things," as discussed above, was a Coke machine classified as an appliance [2]. Modern technologies have vastly expanded the numbers and types of things with network access to include smart homes [6], mobile phones [7], smart offices [8], and, more recently, smart cities and smart cars. These "things" often incorporate multimodal sensors and have embedded processing; for instance, today's smart phones include gyroscopes, cameras, cellular signals, WiFi, Bluetooth, and Global Positioning System (GPS) as standard sensors. The data from these sensors is often aggregated for purposes such as distance or step tracking, navigation, and localization [9]. Data fusion with IoT devices often incorporates data from Internet devices including routers, mobile devices such as smart phones, and more recently body sensors and implantable devices that monitor vital signs such as heart rate [10]. "Things" at the hardware level have several key environmental, physical, and energy-related challenges. For instance, IoT devices used in industrial agriculture would be subject to rain and other weather phenomena. These same devices might suffer physical interference from birds and other wildlife. Finally, and probably one of the most significant challenges today, is the question of keeping remote devices powered on and operable for an undefined duration. This challenge has inspired its own areas of research on embedded computation, network communication protocols, and algorithms for analysis that optimize power usage for low power situations [11].

2.2 MIDDLEWARE: NETWORK TOPOLOGY & DISTRIBUTED COMPUTATION

Once data is generated at the sensor, it is important to understand how to communicate it across multiple protocols and standards that include different latencies, bandwidths, and frequencies, as well as how to maximize data extraction from these sensor networks [12]. This involves the creation of middleware that extracts meaning from sensor signals and communicates it to the network. For example, indoor positioning using a single wireless router would require some protocol that extracts the signal and processes it into an acceptable format for network communication to a node that has analysis capability. Usually the node is a server on the network, as most wireless routers are not equipped with embedded processing for target detection [13]. IoT devices require the implementation of more complex network topologies, such as mesh networks, to accomplish these tasks. These network topologies, which are much more dynamic than the traditional distributed computing model of a static centralized server processing sensor data, must effectively route communications across heterogeneous data and communication protocols while accounting for challenges such as limited node availability and delays in data transmission, particularly on low powered devices [14]. Table 1 gives an overview of several network topologies widely used for IoT networks. Among these, Zigbee is one of the more popular specifications for IoT devices addressing power management. Zigbee is a low-power, low data rate, and close-proximity wireless ad hoc network (WANET) standard [15, 16].

TABLE 1

Network Topologies for IoT Networks

Network Topology	Description	References
WANET/ MANET/ VANET/ WMN	A wireless/mobile/vehicular ad hoc network (WANET), inclusive of wireless mesh networks (WMNs), is a decentralized type of wireless network. An ad hoc network typically refers to any set of networks where all devices have equal status.	[7, 14, 17]
Wireless mesh network (WMN)	A wireless mesh network (WMN) is a form of wireless ad hoc network, which is a communications network made up of radio nodes organized in a mesh topology where all nodes cooperate in the distribution of data in the network (i.e., each node relays data for the network)	[18–20]
Wireless sensor network (WSN)	A wireless sensor network (WSN) of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure, etc. and to cooperatively pass their data through the network to a main location.	[21–25]
Smart phone ad hoc networks (SPAN)	Smart phone ad hoc networks leverage the existing hardware (primarily Bluetooth and WiFi) in commercially available smart phones to create peer-to-peer networks without relying on cellular carrier networks, wireless access points, or traditional network infrastructure.	[26–30]
Delay-tolerant networks (DTN)	DTN are such networks that may lack continuous connectivity. A lot of terminologies are used for DTNs in the literature: opportunistic networks, disconnected mobile ad hoc networks, time-varying network, intermittently connected network (ICNs), extreme networks, etc.	[26, 31–35]
Low-power wide-area network (LPWA)/ low-power network (LPN)	Low-power wide-area (LPWA) network or low-power network (LPN) are a type of wireless telecommunication wide area network designed to allow long range communications at a low bit rate among things (connected objects), such as sensors operated on a battery.	[11, 20, 36, 37]

Disparate network topologies, like WANETs, inherently increase the difficulty of responding in real-time outside of a localized range, and increase the complexity around decisions such as the size of the transmitted data and its persistence across servers. Constrained embedded processing on low-power devices and limited data transmission rates can make long distance data transmission infeasible. Furthermore, network topologies like wireless mesh networks (WMNs) or mesh networks, a form of WANET, require unique routing protocols to optimize across these spaces [38]. The introduction of fog, cloud, and edge computing techniques as middleware protocols to replace traditional communication protocols was especially important for IoT networks, as they make possible local and distributed computing that prioritizes data, communication, and power usage for diverse settings across heterogeneous networks [39]. Table 2 gives an overview of the types of middleware topologies in use for computation and communication across IoT networks.

TABLE 2

Compute Topologies for IoT Networks

Technology	Description	References
Edge	Edge computing is a method of optimizing cloud computing systems "by taking the control of computing applications, data, and services away from some central nodes (the "core") to the other logical extreme (the "edge") of the Internet" which makes contact with the physical world.[1] In this architecture, data comes in from the physical world via various sensors, and actions are taken to change physical state via various forms of output and actuators; by performing analytics and knowledge generation at the edge, communications bandwidth between systems under control and the central data center is reduced.	[40–43]
Fog	Fog computing or fog networking, also known as fogging, is an architecture that uses edge devices to carry out a substantial amount of computation, storage, and communication both locally and routed over the Internet backbone, and most definitively has input and output from the physical world known as transduction. Fog computing consists of edge nodes directly performing physical input and output of-ten to achieve sensor input, display output, or full, closed-loop process control, and may also use smaller edge clouds (often called cloudlets) at the edge or nearer to the edge than centralized clouds residing in very large data centers.	[44–48]
Cloud	Cloud computing is an information technology (IT) paradigm that enables ubiquitous access to shared pools of configurable system resources and higher-level services that can be rapidly provisioned with minimal management effort, often over the Internet. Cloud computing relies on sharing of resources to achieve coherence and economies of scale, similar to a public utility.	[48–52]

2.3 ANALYSIS: ALGORITHMS AND ANALYTICS

After the hardware and middleware challenges are met, the final domain of data fusion for multimodal sensors is analysis. The advent of autonomous or "smart" agents compounds existing challenges in the IoT space. Autonomous vehicles (AVs), smart cars, smart cities, and mission-focused data fusion prompt assurances around timely delivery of communications across dispersed, heavy-traffic networks, as well as processing the data once it has been received [7,53]. This problem is particularly challenging in the case of mobile "things" such as AVs, which often require near real-time responses. For example, if IoT data collected from AVs is analyzed and correlated with cell phone and weather data available as part of the public infrastructure, we can preemptively issue warnings in real-time about slippery intersections, accidents blocking traffic, or detect and prevent collisions with pedestrians or bicyclists based on their cell phone location data. As networks continue to evolve to include a wide variety of sensors, the rate at which data is produced, processed, and analyzed must scale accordingly. AI and ML algorithms are strong contenders for analysis of large data sets; however, traditional, supervised training methods for algorithms, like neural networks, are sensitive to conditions that are likely to exist in dispersed networks like IoT: the curse of dimensionality (high dimensional data); concept shifts, such as severe weather changes, that render existing models invalid; and missing and null values in feature sets [54–56]. However, many of these algorithms are relatively easy to implement and deploy, and provide clearly communicated results. Traditional signal processing, using methods like received signal strength indicator (RSSI) or trilateration, can be sensitive to outliers, increased noise, may require extensive investigation for the creation of each processing algorithm, can be more difficult to deploy, and the results may not be as easy to communicate [57]. On the other hand, signal processing algorithms are often designed to be robust to missing values and latency. Recent research combines these efforts to create algorithmic analysis pipelines with built-in redundancy that provides robustness and validation for comprehensive data fusion that optimizes across multiple parameters and expands into multi-objective output algorithms that account for underlying statistical correlations across diverse data sets. This layer of analysis is the culmination of data fusion for IoT networks. We explore these further in examining the applications and results from existing experiments in Applications (Section 5).

Data aggregation, data fusion, and data analysis may vary substantially across networks to account for scaling, infrastructure, and availability. Hardware, middleware, and analysis approaches for large scale data are inherent problems in meeting the goals of IoT data fusion. In this work, we survey a variety of sensors, algorithms, and applications for data fusion methods with a particular focus on indoor positioning.

3. STUDY OBJECTIVE

The academic literature contains a large body of work on Internet of Things technology and analysis of IoT data. As mentioned in Section 1, the focus of this study is on techniques for fusing data from IoT. The following questions are of particular interest:

- What methods comprise the state of the art (and practice) in IoT data fusion?
- How does performance of these algorithms improve with additional data?
- Are there particular applications in which there has been substantial progress on IoT data fusion?
- Are there applications in which high-fidelity information is available based on IoT data alone?
- What is the state of commercial IoT fusion services?

The objective of the present study is to make progress toward answering these questions, with a focus on those techniques that have been demonstrated in applications of potential relevance to various branches of government. Note that we refer to "fusion" rather than "aggregation" for consistency with [58], from which the following definitions were summarized:

- *Aggregation*: summarization techniques often used to reduce redundancy and conserve resources, such as suppression (e.g., discarding duplicates) and packaging (i.e., grouping multiple observations together)
- *Fusion*: the combination of multiple data sources to obtain improved information (cheaper, higher quality, or more relevant/useful)

As the focus of this study is providing enhanced information content, the methodologies of interest fall in the "fusion" category.

This page intentionally left blank.

4. FINDINGS FROM SURVEYS

TABLE 3

Surveys

Category	Description	References
Architectures	*Key ingredients in an IoT recipe: Fog Computing, Cloud computing, and more Fog Computing*: Good general reference for implementing IoT fog computing at scale. Discusses IoT challenges that make fog computing a natural solution. Outlines approaches, including mobile cloud computing (MCC) and HetNets (heterogenous networks). Emphasizes computational offloading in the fog and decentralization for disbursed networks.	2014 [46]
	A Survey of Fog Computing: Concepts, Applications and Issues: Good overview of how fog computing integrates IoT and the challenges in meeting the QoS critieria for fog computation.	2015 [39]
	Survey of real-time processing technologies of IoT data streams: Focused on distributed architectures and paradigms for real-time utilization of massive IoT data streams, and proposes Information Flow of Things (IFoT) framework based on distributed processing among IoT devices.	2016 [12]
Data Aggregation	*A Survey of Distributed Data Aggregation Algorithms*: Defines and examines standard data aggregation functions that IoT protocols would have to perform in a distributed environment. Describes main aggregation techniques, and provides guidelines for selection and use.	2015 [59]
	A review of aggregation algorithms for the Internet of Things: Provides an overview of IoT data aggregation approaches, and proposes a new class of consensus-based aggregation with fault tolerance to address node reliability issues.	2017 [60]
	Data aggregation mechanisms in the Internet of things: A systematic review of the literature and recommendations for future research: Reviews and categorizes IoT data aggregation mechanisms, comparing techniques in each class.	2017 [61]
Data Fusion	**Information fusion for wireless sensor networks: Methods, models, and classifications*: Surveys methods, algorithms, architectures, and models of information fusion, with applications, limitations, and trade-offs.	2007 [58]
	Data fusion and IoT for smart ubiquitous environments: A survey: Reviews IoT data fusion literature with focus on mathematical methods and specific IoT environments (e.g., object tracking). Discusses opportunities, challenges, and emerging areas. Includes 214 references, including 8 other data fusion surveys (in their Table 1).	2017 [62]

Continued on next page

* Discussed in text

TABLE 3

Surveys (continued)

Category	Description	References
Hardware	*The Internet of Things on Its Edge*: Review of commercial IoT edge device hardware characteristics with histograms, such as # of IoT devices containing each sensor type (e.g., temp, accelerometer, gyroscope, altimeter, microphone) or communication standard (e.g. USB, Ethernet, Modbus); resolution and accuracy for temp sensing; RF carrier frequencies used; RF receiver sensitivity, etc.	2018 [63]
Implementations	*Internet of Things (IoT) Operating Systems Support, Networking Technologies, Applications, and Challenges: A Comparative Review*: Detailed comparison of features across nine IoT operation systems (architecture, schedulers, network stack, memory management, energy efficiency, etc.). Extensive bibliography (over 300 entries).	2018 [64]
	Internet of Things: A survey on the security of IoT frameworks: Compares features of eight IoT frameworks, including architecture, app development, hardware, and security architecture/features (including comms, authentication, and authorization/access).	2018 [65]

* Discussed in text

We began our research by reviewing a variety of existing, broad, IoT surveys, such as the 2017 technology assessment published by the United States Government Accountability Office's (GAO) Center for Science, Technology, and Engineering, entitled "Internet of Things: Status and implications of an increasingly connected world." [66] While these provided some good background material, they did not necessarily lend much insight into the questions of interest surrounding data fusion techniques in IoT. We then broadened our review of surveys to include several more specific areas that brought us closer to our goal. Table 3 provides examples of these surveys, organized in categories with a summary description of each example survey.

Surveys on IoT architectures, included under the Architectures category in Table 3, are closely related to how data is processed and combined in IoT networks under different paradigms such as centralized cloud computing versus decentralized edge or fog computing [46] [39]. Real-time applications require consideration of additional constraints and requirements [12].

These architectural paradigms employ a variety of data aggregation techniques, for which several surveys are included in the Data Aggregation category of the surveys table [59] [60] [61]. However, as outlined in the Objective section, data aggregation techniques are largely aimed at data reduction for processing efficiency, which is distinct from the goals of data fusion.

Surveys such as those listed in the Implementations section of the surveys table review and compare realized IoT frameworks [65] and operating systems [64]. Other surveys of tangential interest include hardware-oriented surveys such as the 2018 Alioto and Shahghasemi review of the features, capabilities, and trends of IoT devices [63].

The 2007 survey by Nakamura et al. [58] is an excellent and widely-cited survey on information fusion for wireless sensor networks that provides comprehensive background information, such as terminology, definitions, and various ways of classifying approaches, an extensive set of references, and detailed descriptions of methods and techniques along with their applications and limitations. Here data fusion is distinguished from data aggregation, in that the goal is to combine multiple data sources to obtain improved information. One of the interesting ways of classifying fusion methods, presented in this survey, is based on the relationship among the data sources. Complementary fusion involves fusing different portions of a picture into a more complete picture (for instance, temperature readings from different locations within a greenhouse in order to better understand the temperature status of the greenhouse as a whole). In contrast, redundant fusion combines multiple instances of the same information to increase reliability, accuracy, and/or confidence (e.g., multiple temperature readings from the same location in a greenhouse to get a more reliable or accurate understanding of the temperature at that location). Finally, cooperative fusion combines multiple independent sources of information into new, more complex information (e.g., temperature combined with light yields a better understanding of overall growing conditions within the greenhouse).

Another useful contribution of the Nakamura et al. survey is the categorization and enumeration of information fusion methods, along with examples of each. For instance, higher level method categories include inference, estimation, feature maps, reliable abstract sensors, aggregation, compression, and information theory approaches. Within the inference category, enumerated methods and examples include Bayesian inference (e.g., fusing laser/radar/video to obtain improved information for driving assistance), Dempster-Shafer inference (e.g., UAV sensor nodes to obtain battlefield situational awareness), fuzzy logic (e.g., node positioning and autonomous navigation), neural networks (e.g., automatic target recognition), abductive reasoning (e.g., diagnosing problems, or event detection/explanation), and semantic information fusion (e.g., recognizing robot behavior based on trajectory). Another large category is estimation, which includes maximum likelihood (ML) (e.g., location discovery), maximum a posteriori (MAP) (e.g., track position of autonomously moving objects), least squares (e.g., reduce communications), moving average filter (e.g., reduce errors in target tracking), Kalman filter (e.g., refining location and distance estimates), and particle filter (e.g., node location discovery and target tracking).

The 2017 Alam et al. data fusion survey [62] cites eight other data fusion surveys as well (in their Table 1), and provides extensive references, especially in the area of object tracking, many of which are more recent than examples covered in the 2007 Nakamura et al. survey.

As we sought more specific and recent examples of academic research and experiments with these data fusion methods, we turned to the application areas in which the work is generally published, such as smart cities and transportation, industrial manufacturing and agriculture, public health, human activity detection and classification, and mapping and localization. Additional related surveys, as well as individual examples, are covered in Section 5, Findings from Application-Focused Literature, as well as in Section 6, Deep Dive: Indoor Positioning with IoT.

This page intentionally left blank.

5. FINDINGS FROM APPLICATION-FOCUSED LITERATURE

IoT has diverse applications across many fields, and in this section we survey some of the established and emerging applications where examples of data fusion abound.

5.1 SMART CITIES & TRANSPORTATION

Smart cities are incorporated into the landscape across a wide spectrum of technologies and are designed to effectively integrate urban areas for communication, power management, resource management, transportation, emergency services, law enforcement, and many more applications. IoT networks in smart cities further extend this to citizens through smart offices and buildings, and even mobile devices for pedestrian management. The term smart city encompasses a wide variety of technologies and applications with many significantly funded initiatives underway both internationally and in the United States [9, 67].

The ongoing focus in smart cities today includes both real-time responses and big data analytics. Notably, Barcelona's "CityOS" strategy, a smart city initiative, has implemented a number of sensors to meet their goals [68]. The irrigation system in a local park uses sensors transmitting real-time data to gardening crews about the level of water required for the plants. A new bus network based on data analysis of the most common traffic flows in Barcelona, integrates smart traffic lights that optimize the number of green lights. In an emergency, emergency vehicles are prioritized by the traffic light system, setting green lights as the vehicle approaches. This method integrates local sensors, GPS, and traffic management software. Much of this data is managed by the Sentilo Platform, an open source software that integrates sensor data, and allows for real time responses [69].

In further application review, the area of intelligent transportation systems showcased two related methods of integration in Madrid. Both use forms of Bayesian networks, where a set of states have their conditional probabilities altered based on observed evidence (e.g., observing that traffic jams are more likely to occur when it rains, and updating the predictive model accordingly). A 2018 paper trained a Bayesian network on complex events derived from Twitter, weather data, and traffic sensors around Madrid. These sources, along with temporal data, create a set of features that are used to derive a probability of traffic congestion. In simulation, the system achieves a 75% precision rate with 95% recall when classifying congestion vs. non-congestion, though there is no direct comparison to a baseline approach [70]. A similar approach is taken in a 2017 paper that uses a dynamic Bayesian network, in which state transitions are tracked over time. The method is meant to be general and take arbitrary evidence into account to determine the state of traffic (e.g., is what the sensors see consistent with construction or a vehicle breakdown). The empirical results in this paper focus on using frames from traffic cameras and demonstrate over 80% inferred probability of the correct condition. Again, no direct comparison to any existing approach is provided [71].

Table 4 provides a list of further applications and methodologies for smart cities.

TABLE 4

Smart Cities & Transportation

Title	Description	Reference
Internet of Things for Smart Cities	Survey of technologies, protocols, and architecture for urban IoT, along with discussion of Padova Smart City project.	[67]
CityPulse: Large Scale Data Analytics Framework for Smart Cities	CityPulse framework integrates multimodal, mixed quality, uncertain and incomplete data to create reliable, dependable information and continuously adapts data processing techniques.	[72]
An Information Framework for Creating a Smart City Through Internet of Things	Illustrates a new method for existing operations that can be adapted for the enhancement and delivery of important city services.	[73]
Urban Planning and Building Smart Cities Based on the Internet of Things Using Big Data Analytics	Open-source real time processing of IoT data for urban planning or city future development. The offline historical data generated by smart homes, smart parking weather, pollution, and vehicle data sets are used for analysis and evaluation.	[74]
Smart Cities and the Future Internet: Towards Cooperation Frameworks for Open Innovation	This paper explores "smart cities" as environments of open and user-driven innovation for experimenting and validating Future Internet-enabled services. Based on an analysis of the current landscape of smart city pilot programs.	[9]
City Saves Money, Attracts Businesses with Smart City Strategy	Barcelona's investment in data-driven city management results in improved city services and reduced expenditures.	[68]

5.2 INDUSTRIAL MANUFACTURING & AGRICULTURE

The IoT connectivity and "smart" paradigm has seen significant expansion and use in both industrial manufacturing and agriculture. Industry leaders like John Deere [75] and Deloitte [76] have already implemented the smart factory or Industry 4.0. Smart factories in Korea have already been shown to improve productivity [77]. Smart factory sensors allow for a myriad of improvements in production. They can control energy and water usage to reduce waste and optimize for environmentally sustainable operations. Quality control in the smart factory can be enhanced with real time analytics on the supply chain. The exchange of information from other systems and devices directly back into the production line, enables predictive maintenance and forecasts needed improvement or down times. Supply chain connectivity improves asset tracking and triggers orders when stock runs low, synchronizing stock delivery and the production line [78]. There is even some evidence of improved security from cyber attacks in smart factories [79].

Specifically in Korea, a forerunner of smart factory production for small and medium enterprises (SMEs), The Ministry of Trade, Industry, and Energy (MOTIE) has implemented smart factory production lines in 1,240 small and medium enterprises. The smart factories have improved the SMEs' productivity by approximately 25% including a 27.6% decrease in fraction defective, a cost reduction of 29.2%, and a 7.1% reduction in the length of time taken for prototype production. An additional 137 SMEs developed in 2015 resulted in quality improvement of 77%, and 139% productivity improvement [77]. Smart factory innovation in the USA, including Boeing and Ford are starting to gain footing as manufacturers realize the value of incorporating IoT into production [80, 81].

Similarly, large scale agricultural has many parallels with industrial manufacturing operations and has also started to adopt the smart paradigm and incorporate IoT into their business practices. Smart farming and precision agriculture utilize a wide array of sensor networks to incorporate soil, temperature, atmospheric and weather data, GPS, cumulative statistics for crop rotation, and other information. Agricultural sensors utilize both predictive approaches and control approaches that include information from static crop indicators. As with industrial manufacturing, they have incorporated robotics, such as John Deere's self-driving tractors [75]. Data is usually shared over WSNs and connected to supply chains for information like seed dispersal. Smart farming has been expanded into cattle care and beekeeping connected to IoT networks [78, 82].

John Deere, considered a pioneer in precision agriculture, utilizes a positioning system that combines GPS signals with ground-based base stations and IoT sensor networks to achieve position accuracy of about 2 cm for their self-driving tractors and agricultural systems. Described as similar to NASA's rover positioning systems, metrics have been withheld for proprietary reasons [83, 84]. However, similar technology is in use by other industry leaders including Monsanto and DuPont [85, 86].

Table 5 covers further examples of smart factories and smart farming/precision agriculture.

TABLE 5

Smart Factories & Farming

Title	Category	Description	Reference
Smart Factories in Industry 4.0: A review of the concept and of energy management approaches in production based on the Internet of Things paradigm	Smart Factory	IoT has stimulated the next phase of factories launching an evolution called Industry 4.0. Industrial production of the new era will be sustainable; highly flexible in production volume and customization; and will have extensive integration between customers, companies, and suppliers. This study reviews and analyzes the current initiatives and related studies of smart factories./Industry 4.0,	[87]
IoT-Based Intelligent Perception and Access of Manufacturing Resource Toward Cloud Manufacturing	Smart Factory	This study looks at IoT technologies in cloud manufacturing (CMfg), classifying manufacturing resources and services in a five-layered structure (i.e., resource layer, perception layer, network layer, service layer, and application layer) The key technologies for intelligent perception and access of various resources are explored.	[88]
Internet of Things for Enterprise Systems of Modern Manufacturing	Smart Factory	Investigates the impact of emerging Internet of Things (IoT) on enterprise systems in modern manufacturing by examining the evolution of manufacturing system paradigms to identify the requirements of decision support systems in dynamic and distributed environments.	[89]
IoT-Based Real-Time Production Logistics Synchronization System Under Smart Cloud Manufacturing	Smart Factory	A study examining the interlinkage of cloud manufacturing (CM) and Internet of things (IoT) as an inadequate method for a highly service-driven manufacturing execution system which entails systematic CM support to respond to the real-time dynamics captured from the IoT-enabled execution hierarchy.	[90]
Smart Agriculture Based on Cloud Computing and IOT	Smart Farming	A study examining smart farming development in China, and the introduction of cloud computing and IoT solutions. Using cloud computing and key IoT techniques, visualization and SOA technologies can take advantage of massive data involved in agricultural production.	[91]
The Study and Application of the IOT Technology in Agriculture	Smart Farming	Study examines the rapid development and wide application of IoT technology in agricultural automation, including the integration of control networks with IoT and actual greenhouse agricultural production.	[92]
Design of Intelligent Agriculture Management Information System Based on IoT	Smart Farming	Introduces the concept of agricultural information management and analyzes the features of agricultural data. Discusses Intelligent Agriculture MIS architecture, and presents an example related to crop production.	[93]
*John Deere and the Birth of Precision Agriculture	Smart Farming	Article looks at John Deere's pioneering efforts into smart farming and precision agriculture.	[75]

* Discussed in text

5.3 PUBLIC HEALTH

The introduction of IoT for public health has seen three major focus areas: wearable or bio-sensors for health monitoring and disease onset or progress [94, 95], smart homes and ambient assisted living for high risk populations [96], and more general climate monitoring for both disease and natural disaster [97–99]. Many of these applications involve extensive data fusion and aggregation across time, and often require real-time responses for medical emergencies, like heart attacks, or relief from natural disasters [99, 100].

Wearable or bio-sensors have become mainstream with technology like FitBit [101] and the Apple Watch [102]. Recreational exercisers have used these tools to monitor and improve performance and technique [103]. These devices and more specialized versions have recently entered the IoT domain, including applications like sharing jogging routes or weight-loss goals with friends and social media [104]. The medical community has further advanced the use of bio-sensors for disease progression and monitoring. Bachlin et. al.'s study looked at a wearable assistant for Parkinson's disease (PD) patients with the freezing of gait (FOG) symptom and used body acceleration sensors to measure the patients' movements. Time series and fractal analysis algorithms automatically detected FOG events by analyzing frequency components inherent in these movements. FOG events triggered real-time auditory responses that stimulated the patient to resume walking. Their methods detected FOG events with a sensitivity of 73.1% and a specificity of 81.6% and many patients expressed resumed mobility concurrent with the auditory stimulation [105].

The use of bio-sensors may be further extended by integration with smart homes and smart cities to provide necessary medical interventions. The area of ambient assisted living (AAL) has seen an increased usage of this IoT application for monitoring of high risk populations such as the elderly. An example application uses a wireless sensor network with visual and wearable sensors for smart home monitoring and event detection. This study highlighted fall detection of a person under care using a wearable wireless badge to connect the user and the network. The badge provided event detection and a voice communication channel for communication with emergency services without the use of a phone. The smart home implementation utilized received signal strength indication (RSSI) to track the approximate location of the user in the monitoring environment and provide additional information about the user's situation [106]. AAL applications have been extended in further studies to include medication management and other helpful services to users of assisted living services [107].

When we combine wearable and smart infrastructures with climate monitoring there are even greater implications for IoT in the public health sphere. A surprising 2014 study combined aggregated climate data gathered from weather sensors collecting temperature, humidity, and rainfall data with health data collected by the Ministry of Health in Chile to accurately predict outbreaks of Hantavirus in the region. Auto Regressive Integrated Moving Average (ARIMA) time-series models and regression models with ARIMA errors were used to predict outbreaks with greater than 95% confidence in many cases over a 12 year period between 2001-2012 [108]. Further studies integrating social media data with climate and health data have predicted disease outbreaks and have been used to locate power outages in disaster stricken areas [109].

Table 6 covers public health related examples.

TABLE 6

Public Health

Title	Category	Description	Reference
* Wearable Assistant for Parkinson's Disease Patients With the Freezing of Gait Symptoms	Disease Onset/ Progression	A wearable assistant for Parkinson's disease (PD) patients with the freezing of gait (FOG) symptom using acceleration sensors to measure the patients' movements, detect FOG events and provide an auditory response to enable patients to resume motion.	[105]
ECG signal analysis and arrhythmia detection on IoT wearable medical devices	Disease Onset/ Progression	Study using wearable IoT to monitor and control cardiovascular diseases (CVD). Uses Support Vector Machine (SVM) classifier for ECG analysis and heartbeat diagnosis. The best classification accuracy achieved is 98.9%, for a feature vector of size 18 and 2493 support vectors. The ECG analysis and classification can be performed in real time.	[100]
Technologies for an aging society: a systematic review of "smart home" applications	AAL	Comprehensive review of health related smart home projects and their challenges.	[94] Addl Applications [96, 110]
Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing: Opportunities and challenges	AAL	Discusses low power, wireless body area networks (WBAN) for localized IoT sensors in health monitoring. Introduces a hybrid cloud/cloudlet framework with context awareness, and outlines challenges of dealing with high dimensional, time series data.	[111] Addl. Applications [107]
* Smart home care network using sensor fusion and distributed vision-based reasoning	AAL	An example of sensor fusion for improved fall detection. Combines image-based sensing and reasoning (from multiple camera views) to corroborate accelerometer events and reduce false alarms.	[106] Addl Application [112]
*Modeling to Predict Cases of Hantavirus Pulmonary Syndrome in Chile	Disease/ Disaster	Hantavirus outbreaks are typically small and geographically confined. Few have considered climatological modeling of HPS incidence for monitoring and forecasting purposes. Methodology Monthly counts of confirmed HPS cases were obtained from the Chilean Ministry of Health for 2001–2012. The data suggested a seasonal trend, which appeared to correlate with changes in climatological variables such as temperature, precipitation, and humidity.	[108]
Wearable IoT sensor based healthcare system for identifying and controlling chikungunya virus	Disease/ Disaster	Wearable IoT sensors, fog computing, mobile technology, cloud computing, and better internet coverage have enhanced the quality of remote healthcare services. IoT-assisted fog health monitoring systems can be used to identify possibly infected individuals in an early phase of their illness so that the outbreak of CHV can be controlled.	[109]

Continued on next page

* Discussed in text

TABLE 6

Public Health (continued)

Title	Category	Description	Reference
A Semantic IoT Early Warning System for Natural Environment Crisis Management	Disease/ Disaster	An early warning system (EWS) is a core type of data-driven IoT system used for environment disaster risk and effect management. This study uses lightweight semantics for metadata to enhance rich sensor data acquisition. This approach is validated through determining both system-related metrics and a case study integrated with a deployed EWS infrastructure.	[98] Addl. Application [99]

* Discussed in text

5.4 HUMAN ACTIVITY DETECTION AND CLASSIFICATION

An area of growing interest, and specifically related to indoor positioning, is human activity detection and classification. This area has numerous applications, from smart home or building applications, like occupancy monitoring, to more involved human pose detection. As described in Akkaya et al. [113], occupancy monitoring is a multi-layered problem, with occupancy detection (the presence or absence of people in a space) being the simplest. Beyond that, occupancy counting seeks to determine how many people are in the space, while occupancy tracking seeks to discover where they are and where they move within the space. The latter overlaps heavily with work in indoor localization and tracking. At an even higher level is occupancy behavior recognition, with the goal of recognizing and understanding individual or collective activities within the space. Challenges mentioned include the desire to be minimally intrusive (i.e., use existing infrastructure), and to develop data fusion techniques to improve accuracy. To that end, occupancy detection often involves multi-modal data fusion including network nodes, cameras, environmental sensors, and more fine-grained wearable sensors to determine occupancy and location of individuals.

Ghai et al.'s work uses only "opportunistic context sources" or "soft sensors," such as access badges, WiFi access points, and applications such as calendaring and instant messaging to per-form occupancy detection. They deployed low-cost machine learning algorithms, including linear regression and C4.5 classification, to locate and detect occupants. Volunteers in the office were manually tagged using a desktop application installed on their laptops that updated location and status whenever they moved with annotations such as 'cubicle' or 'break'. This information was used to denote ground truth for the machine learning algorithm. The pilot study involved a single floor with 5 volunteers for 6 weeks, and demonstrated precision and recall between 80 and 89% for C4.5 classification for 'cubicle' and 'break' classes. Linear regression performed less reliably on this task with precision between 58% and 92% and recall at approximately 59%. However, there were no baseline methods for comparison [114]. In contrast, Nesa and Banerjee used Dempster-Shafer evidence theory (DSET) to fuse information from multiple heterogeneous data sources such as temperature, humidity, light, and CO2. They compared their results to other methods, and also showed how performance improved as the number of fusion parameters increased [115]. Occupancy behavior recognition techniques, such as those investigated by Dong and Andrews, are often used for more sophisticated automatic environment management to improve energy efficiency and comfort of building occupants. Events such as lights turning on and motion being detected are derived from raw data. Significant patterns in event sequences are then recognized based on length and periodicity, which then feed into a model of occupancy patterns and duration. This model informs environmental adjustments, such as heat and cooling. Simulations demonstrated that their dynamically-derived schedule would achieve comparable energy savings to a motion-only based strategy, but with less discomfort time [116].

Advanced human activity classification often involves multipath or through wall human activity detection including activities like sitting, standing, walking, or tracking. Kim et al.'s 2009 study used micro-doppler to classify human activity for running, walking, walking while holding a stick, crawling, boxing while moving forward, boxing while standing in place, and sitting. This study measured radar signals during each of these activities and then applied a short-time Fourier Transform (STFT) over an averaged window of time and then applied machine learning (support vector machines and decision trees) to the resulting data set. This method achieved 100% accuracy on classifying sitting activity and a more modest result of 78% for boxing while moving forward [117]. These results, while impressive, have seen improvement with more recent algorithms that isolate fine-grained human pose detection.

A recent MIT CSAIL study utilized wireless signals for two dimensional ($2D$) pose estimation. This study analyzed radio signals reflected off a person's body using frequencies in the WiFi spectrum. The signal analysis was used to build a confidence map and heat map of joints on a 2D figure to obtain the object keypoint similarity (OKS). Average precision for each OKS is roughly the distance between ground truth, or the actual activity pose, and detected keypoint from the sensor, or WiFi signal. Deep neural networks were trained using visual camera data and then used on pose detection from wireless data to implement a cross-modal learning and detection algorithm. This method achieved similar performance through walls to the existing OpenPose framework for human pose detection that relied primarily on visual data [118].

Table 7 provides a list of further applications and methodologies for human activity detection.

TABLE 7

Human Activity Detection & Classification

Title	Category	Description	Reference
* Occupancy detection in commercial buildings using opportunistic context sources	Occupancy Detection	Uses "opportunistic context sources" or "soft sensors" such as access badges, WiFi access points, and applications such as calendaring and instant messaging with ML techniques.	[114] Addl application [115]
Occupancy detection through an extensive environmental sensor network in an open-plan office building	Occupancy Counting	Discusses feature selection - largest info gain sources: multiple CO2-based values, acoustics, and PIR (passive infrared). Compares performance of SVM, NN, and HMM. HMM fares well - "successfully describes the major changes in occupancy while ignoring abrupt fluctuations of short duration"	[119], Addl application [120]
*IoT-based occupancy monitoring techniques for energy-efficient smart buildings	Occupancy Monitoring	Survey of multi-modal data fusion techniques for smart buildings using existing infrastructure; develop data fusion techniques to improve accuracy.	[113]
*Sensor-based occupancy behavioral pattern recognition for energy and comfort management in intelligent buildings	Occupancy Behavior Detection	Occupancy behavioral patterns used to develop dynamic schedule for building energy/comfort management. Detect events from raw data to discover significant patterns in event sequences and develop model of occupancy patterns/duration via HSMM (hidden semi-Markov model).	[116]
* Through-Wall Human Pose Estimation Using Radio Signals	Activity Classification	Wireless signals in the WiFi frequencies traverse walls and reflect off the human body. We introduce a deep neural network approach that parses such radio signals to estimate 2D poses.	[118]
* Human Activity Classification Based on Micro-Doppler Signatures Using a Support Vector Machine	Activity Classification	Measured 12 human subjects performing seven different activities running, walking, walking while holding a stick, crawling, boxing while moving forward, boxing while standing in place, and sitting using a micro-Doppler radar	[117]

* Discussed in text

5.5 MAPPING & LOCALIZATION

Simultaneous localization and mapping (SLAM) is a longstanding and key challenge in robotics that involves using combined robotic sensors to map the space that a robot occupies as well as understand the robot's positioning in that space. The SLAM problem spans multiple fields including physics, sensor fusion, and real-time computation. It has key applications in envi-ronments where global positioning systems (GPS) may be inaccurate, previously unmapped areas, or areas of high uncertainty such as poorly marked roads, adverse weather events, or unexpected visual interference. Ideally, the SLAM algorithm takes place in near real-time and allows for in-teraction with the environment, and encompasses a larger variation in sensor times and expected responses. SLAM is used in practice for things like autonomous vehicles and robot vacuums which build a map of the user's home and navigates collision avoidance [121–123].

The SLAM problem incorporates multiple modalities including visual, odometry, gyroscopic and many other variations. The problem is diverse enough to give key insight into the challenges we will face with autonomous vehicles (AV) and integration with smart factories, smart cities, smart homes, and the wider future of IoT. Integrating the AV and other robotics requiring indoor positioning into the IoT network will allow deeper integration into emerging infrastructures. Traditionally, SLAM has been performed used relatively isolated embedded computation. The introduction of fog, cloud, and edge computing and the expansion of IoT has been rapidly changing the way we approach the SLAM problem [44].

The canonical example of SLAM is for the use of smart or autonomous vehicles. Several commercial examples have already been released and are in limited use including Tesla [124], Uber [125], and Google [126]. Much of this development was realized from the DARPA Grand Challenge, which over the years has pioneered development in both robotics and autonomous vehicles. The 2007 DARPA Urban Grand Challenge was won by a vehicle named Boss, created by the Carnegie Mellon University (CMU) Team [127]. Boss was a massively complex undertaking fusing global positioning system, lasers, radars, and cameras to track other vehicles, detect static obstacles, and localize itself relative to a road model. It also utilized extensive AI/ML algorithms for mission, behavioral, and motion planning. The mission planning layer optimizes over a value function to determine which street to take to achieve a mission goal. The behavioral layer determines when to change lanes and precedence at intersections in addition to performing error recovery maneuvers. The motion planning layer selects actions to avoid obstacles while making progress toward local goals [128].

More recent SLAM variants include applications, like SmartSLAM, that focus on localization of pedestrians and mapping of spaces using devices like smart phones. Shin et. al. evaluated mapping variants with SmartSlam using a smart phone. Their experiment automatically constructed an indoor floor plan and radio fingerprint map for anonymous buildings using odometry tracing with inertial sensors, an observation model using WiFi signals, and a Bayesian estimation for floor-plan construction. Their method does not require additional devices and achieved an average error between 2.98 and 3.51 meters for localizing and mapping using a particle filter algorithm [129].

A relatively recent innovation in mapping and localization applications is radio tomography (RT), which refers to the application of tomographic techniques to radio waves. Tomography is a long-standing practice across many fields including medical magnetic resonance imaging (MRI), ultrasounds, and even archaeology. It refers to imaging by sections or sectioning, through the use of any kind of penetrating wave. Specifically, radio tomography has evolved in the past fifteen years from theory to practical application. Initially it was applied to geophysical investigations of solar-terrestrial processes and has recently expanded to the mapping and modelling of the ionized atmosphere with practical radio systems. Modern applications of interest in the IoT domain are the use of RT for mapping and localization across many radio or radar nodes. The method applies tomographic inversion using RSS and other wave frequency data to give a two-dimensional image of the space [130].

Wilson and Patwari demonstrated effective use of radio tomographic imaging (RTI). Their RTI system used 28 wireless nodes at 2.4 GHz. Signal processing assumed a Gaussian noise model, and image reconstruction used a maximum a posteriori (MAP) estimator. Results showed that the system was effective at creating attenuation images of humans standing in areas on the order of hundreds of square feet and was able to track targets that moved within a wireless network [131]. Through-wall tracking was further evaluated in a follow up paper that used variance based RTI (VRTI). Experimental results for a 34-node through-wall imaging and tracking system over a 780 square foot area saw respectable performance. For example, over a ten second average of human movement through 20 known positions using VRTI and Kalman filtering, the average error was approximately 1.46 feet. Overall, their experiments demonstrated accuracy between approximately 2 and 6 feet for localization of a target [132]. The further works by these authors are listed in Table 8.

In the commercial sector, Bastille Networks (whose business motto is "security for the internet of radios") markets Bastille Enterprise, a product aimed at enterprise security, which scans the radio spectrum to identify and locate radio emitters operating on frequencies between 60 MHz and 6 GHz. Among other things, this may be used to discover rogue WiFi access points or unauthorized Bluetooth tethering. The company employs several patented technologies, such as Bayesian device fingerprinting to detect and identify devices, and distributed tomographic localization to determine a device's position within one meter. The latter technology allows them to estimate the position of walls and other objects which are then factored into their device localization model [133]. Dr. Robert Baxley, Chief Engineer at Bastille and Adjunct Faculty member at Georgia Tech, has co-authored a number of academic research papers related to RTI [134–139].

TABLE 8

Mapping & Localization

Title	Category	Description	Reference
Toward a rapidly deployable radio tomographic imaging system for tactical operations	Localization: RT	A study of a rapidly deployed radio frequency (RF)-based tomographic imaging (RTI) system for use in tactical operations by Special Weapons and Tactics (SWAT) and other SOF, using low-power radio devices around a building to image and track the motion of humans inside the building.	[140]
*Radio tomographic imaging with wireless networks	Localization: RT	A linear model for using received signal strength (RSS) measurements to obtain images of moving objects. Noise models are investigated based on real measurements of a deployed RTI system. Mean-squared error (MSE) bounds on image accuracy are derived, which are used to calculate the accuracy of an RTI system for a given node geometry.	[131] Addl Applications [134, 137, 141–144]

Continued on next page

* Discussed in text

TABLE 8

Mapping & Localization (continued)

Title	Category	Description	Reference
Radio frequency tomography in mobile networks	Localization: Mobile RT	A study using received signal strength (RSS) measurements as used in ad-hoc networking to describe how information about the shadowing environment is encoded into RSS measurements. Utilizes a novel RF Exploitation for Tomographic Imaging and Non-cooperative Analysis (RETINA) algorithm to detect stationary obstacles and track moving objects.	[135] Addl Applications [139]
*See-through walls: Motion tracking using variance-based radio tomography networks	Localization: Multipath RT	A new method for imaging, localizing, and tracking motion behind walls in real time. The method takes advantage of the motion-induced variance of RSS measurements made in a wireless peer-to-peer network. Using a multipath channel model, it shows the signal strength on a wireless link is largely dependent on the power contained in multipath components that travel through space containing moving objects.	[132] Addl. Applications [145]
Improving radio tomographic images using multipath signals	Localization: Multipath RT	This paper presents a method by which multipath signals may be treated as useful measurements in an RTI system to informs the existing spatial loss field (SLF) image, improve image quality and reducing root-mean-squared-error (RMSE).	[138] Addl Applications [139]
Regularization techniques for floor plan estimation in radio tomographic imaging	Mapping RT	A potential application of RTI for estimation of building floor plans and interior features in a wireless network. Proposes a technique for the regularization of solutions to RTI problems with the goal of enhancing building floor plan images by exploiting a-priori information about the structure, and therefore spatial covariance, of typical buildings.	[136]
*Bastille Networks Internet Security, 'Product introduction.'	Localization & Mapping: RT	Bastille's technology scans the entire radio spectrum, identifying devices on frequencies from 60MHz to 6GHz. This data is then gathered and stored, and mapped so that you can understand what devices are transmitting data from your corporate airspace. This provides improved situational awareness of potential cyber threats and post-event forensic analysis.	[133] Addl Applications [146]
Simultaneous Localization and Mapping	Localization & Mapping: SLAM	A comprehensive introduction to the SLAM problem, addressing robot navigation in an unknown environment, map acquisition, and localization using the map. SLAM aims for detailed environment models, or an accurate sense of a mobile robot's location.	[123] Addl. Application [122, 147–149]
*Autonomous driving in urban environments: Boss and the Urban Challenge	Localization & Mapping: SLAM	Boss is an autonomous vehicle that uses on-board sensors (global positioning system, lasers, radars, and cameras) to track other vehicles, detect static obstacles, and localize itself relative to a road model.	[128] Addl. Application [149–152]

Continued on next page

* Discussed in text

TABLE 8

Mapping & Localization (continued)

Title	Category	Description	Reference
Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds	Localization & Mapping: SLAM	The car is now a formidable sensor platform, absorbing information from the environment (and from other cars) and feeding it to drivers and infrastructure to assist in safe navigation, pollution control, and traffic management. The Internet of Vehicles will be a distributed transport fabric like other important instantiations of the Internet of Things (e.g., the smart building), with storage, intelligence, and learning capabilities.	[53] Addl. Application [149–152]
Human SLAM, Indoor Localisation of Devices and Users	Localization & Mapping: Smart-SLAM	The indoor localisation problem is more complex than just finding whereabouts of users. Finding positions of users relative to the devices of a smart space is even more important. This study proposes a new system to address the problem of locating devices and users relative to those devices, and combine this problem into a single estimation problem using (RSSI) of devices and motion data from users.	[153] Addl. Application [122, 147–149]
Google Cartographer	Localization & Mapping: Smart-SLAM	Cartographer is a system that provides real-time simultaneous localization and mapping (SLAM) in 2D and 3D across multiple platforms and sensor configurations.	[154] Addl. Application [122, 147–149]
*Unsupervised construction of an indoor floor plan using a smartphone	Localization & Mapping: Smart-SLAM	Indoor pedestrian tracking extends location-based services to indoor environments. Typical indoor positioning systems employ a training/positioning model using WiFi fingerprints that require an in-door map, which is typically not available to the av-erage user and involves significant training costs.This paper presents an indoor pedestrian tracking system, called SmartSLAM, which automatically constructs an indoor floor plan and radio fingerprint map for anonymous buildings using a smartphone employing odometry tracing using inertial sensors, an observation model using WiFi signals, and a Bayesian estimation for floor-plan construction.	[129] Addl. Application [122, 147–149]

* Discussed in text

The various application areas covered in this section offer a broad array of data fusion methodologies, including techniques for fusing both homogeneous and heterogeneous data. Very few, however, make comparisons against a baseline or compare results across several methodologies. In addition, one of our objectives is to understand the value of data fusion, and very few papers consider performance as the amount and type of data increases. The application area that best approaches our goals has had considerable attention, even prior to the dawn of IoT: indoor positioning. Section 6 provides a more in-depth look at this particular application, in which methods can be compared to some degree with respect to common metrics.

This page intentionally left blank.

6. DEEP DIVE: INDOOR POSITIONING WITH IOT

As described by Sakpere et al. [155] and others, while GPS has become the standard for outdoor navigation, there is no one, widely-adopted equivalent for indoor navigation, a relatively newer and evolving field which has grown along with the proliferation of personal mobile devices such as smartphones. In many indoor positioning or navigation scenarios, a moving or "target" device is being localized with respect to fixed infrastructure (e.g., wireless access points at known locations). However, in an indoor environment, it is rare to have unobstructed line of sight (LOS) between transmitters and receivers; this is one reason, along with signal attenuation through structures, that indoor localization using GPS is unreliable. Indoor localization difficulties cited by Liu et al. [156] include multipath propagation, wherein the same signals may reach a receiver via two or more paths, and the lack of good models for such indoor multipath propagation. Both static and dynamic aspects of the indoor environment contribute to the difficulty of this problem, including varying floor plans and fixed reflective surfaces, as well as moving objects, such as people and doors that may be open or closed.

As a result, a large body of research exists in attempting to solve this problem in a number of ways. Researchers have considered using and combining a wide variety wireless technologies and indoor sensors, in complementary, redundant, or cooperative schemes (as defined in the previously described information fusion survey by Nakamura et al. [58]). Indeed, such an explosion of research has occurred that a plethora of surveys have been written attempting to categorize, summarize, and compare the works according to various criteria.

In Section 6.1, we describe a number of such surveys and their contributions, while in Section 6.2 we provide a discussion and comparison of numerous works from the past ten years wherein the authors measured and provided some performance information.

6.1 INDOOR POSITIONING SURVEYS

Table 9 provides examples of surveys and comparisons of indoor positioning techniques and systems. Each one has a slightly different focus and provides some insights not found in the others. Each is described in more detail below.

The 2007 survey [156], cited by most of the others, provides an overview of various positioning algorithms within three main categories: triangulation (including five lateration as well as angulation techniques), scene analysis (including multiple probabilistic and machine learning methods), and proximity. The authors further evaluate 20 different systems or solutions based on a va-riety of wireless technologies, such as Bluetooth, WiFi, and RFID. These systems are compared across six performance metrics: accuracy, precision, complexity, robustness, scalability, and cost. This comparison is summarized in Table 1 of their paper. Note that we use a similar format for Table 10 in this report in order to extend their summary with additional work that was published after their 2007 survey.

As indicated by the title, the 2011 survey [110] looks at indoor positioning technologies in the context of ambient assisted living applications. After an overview of the application area and

TABLE 9

Indoor Positioning Surveys and Comparisons

Title	Year	References
Survey of Wireless Indoor Positioning Techniques and Systems	2007	[156]
Overview of Indoor Positioning Technologies for Context Aware AAL Applications	2011	[110]
Overview of RFID-Based Indoor Positioning Technology	2012	[157]
Performance Evaluation of Localization Algorithms for WSNs	2015	[158]
A state-of-the-art survey of indoor positioning and navigation systems and technologies	2017	[155]
Robustness, security and privacy in location-based services for future IoT: A survey	2017	[159]
Evolution of indoor positioning technologies: A survey	2017	[160]
Location of Things (LoT): A review and taxonomy of sensors localization in IoT infrastructure	2018	[161]

technique basics, the authors discuss metrics; in addition to accuracy, precision, and cost, they also consider privacy and "possible context information" that could be gained with sensor fusion. They go on to examine six major single-medium systems (e.g., WLAN, Bluetooth, and RFID), but also emphasize seven examples of newer, multi-medium approaches (e.g., WLAN + RFID, and WLAN+ Altimeter + Image), focusing on works published after the 2007 survey.

The 2012 survey [157] focuses specifically on RFID-based technology, providing an overview of principles, characteristics (e.g., frequency ranges), and application areas before delving further into RFID-based indoor positioning techniques. Also discussed are hybrid techniques, which in-tegrate RFID with other sensors and techniques such that the integrated capabilities compensate for the weaknesses of the component techniques, resulting in more reliable and accurate position-ing determinations. Table 4 in their survey summarizes the advantages and disadvantages of four different measurement models (all of which fall under the triangulation methods described in the 2007 survey [156]). Table 5 in their survey summarizes the advantages and disadvantages of three different RFID-based algorithms (Cell of Origin, lateration, and fingerprinting) and also assesses their accuracy, with fingerprinting judged as the most accurate. The survey concludes with remarks on future trends.

The 2015 performance evaluation [158] focuses on RSS-based localization algorithms. Received signal strength (RSS), or more generally signal attenuation-based methods fall under the lateration techniques described in the 2007 survey [156]. Here the authors classify them into two main types: range-free and range-based (which are generally more accurate). When trying to localize a node of interest (NOI), range-based techniques estimate the distance between the NOI (which may be mobile) and other fixed reference nodes in the network. Range-based techniques include weighted least squares (WLS) multilateration, as well as circular and hyperbolic position-ing algorithms. In contrast, range-free techniques do not base the NOI position on such a distance estimation. Examples include several variations of weighted centroid and K-nearest neighbor localization algorithms, some combined with fuzzy logic. The authors propose yet another variation on WLS, and then report extensively on the results of simulation-based evaluations of the various techniques. Performance metrics include accuracy, precision, and computational complexity.

The 2017 state-of-the-art survey [155] provides a comprehensive overview of indoor positioning techniques, suggesting two different ways of classifying them: by signal property (angle of arrival, time of arrival, time difference of arrival, and received signal strength), and by positioning algorithm (triangulation, trilateration, proximity, and scene analysis/fingerprinting). Their Table 2 summarizes the pros and cons of these approaches. The authors also discuss nine different technology categories (infrared, ultrasound, audible sound, magnetic, optical and vision, radio frequency, visible light, dead reckoning, and hybrid). For each category, several papers or systems are cited and discussed, along with an assessment of the challenges and drawbacks of the technology category. Table 3 in their survey presents a summary and comparison of the technologies across several aspects: technique, algorithm, accuracy, cost, complexity, scalability, privacy/security, and real-time suitability. This survey includes many works that had not been published at the time of the 2007 survey.

The 2017 robustness survey [159] is unique in that it evaluates threats (and mitigations) to the accuracy, security, and privacy of both global navigation satellite system (GNSS) and non-GNSS-based localization solutions including, for example, radio frequency interference (RFI) and database-related vulnerabilities.

The 2017 evolution survey [160] purports to differentiate itself from other surveys by the breadth of its coverage, and by the structure in which it separates and presents the various techniques (such as multilateration and fingerprinting) versus the technologies (such as optical, sound, RF, passive, and hybrid). Their Table 1 provides a comparison to five prior surveys, including the 2007 Liu survey [156], detailing which technologies and sensors are considered in each. Their Table 2 presents a comparison of indoor positioning technologies across multiple facets (accuracy, coverage, cost, strengths, and weaknesses), with accuracies ranging from 1 centimeter (cm) for both ultrasonic and computer vision technologies, to 25 cm for Zigbee, to up to several meters for the majority of others. The authors further discuss evolving trends, such as taking full advantage of existing infrastructure, hybrid technology approaches, the use of mobile devices, and crowdsourcing.

The 2018 survey [161] provides an extensive analysis of localization techniques and compares itself to 28 other localization surveys (in their Table II). They classify techniques accord-ing to the presence or absence of offline training, including a variety self-determining (geometric, mobility model, path planning, and statistical approximation) and training-dependent (fingerprint-ing, stochastic modes/Markov process, and machine learning) approaches, which are diagrammed in their Figure 3. Each of these seven approaches is examined in detail, including a separate ta-ble summarizing and comparing numerous published works for each approach. Also included are several figures plotting accuracy versus power consumption for the various approaches.

6.2 INDOOR POSITIONING METRICS ACROSS APPROACHES

TABLE 10

Indoor Positioning Metrics

Sensors/ Protocols	Data	Algorithms	Accuracy/ Precision	†Complexity	Scalability/ Space dimension	†Robustness	Notes
*Real & simulated WINS, Medusa, RF, Ultrasound, RSSI [162]	50-300 simulation, 5-9 real	Atomic, iterative, and collaborative multilateration	WINS 2-4m, Ultrasound 2-20cm	§ moderate / moderate	30m distance, height of 1.5m; 2D,3D	moderate	Foundational literature on fine grained indoor localization
TelosB 2.4GHz CC2420 radio using received signal strength (RSS) [163]	5 nodes	Specialized localization using proximity vectors and sequence building	$\leq$ 50cm, 90% with 16 channels, 0% with 1 channel	low/medium	§8 x 8m; 24m, nodes 3m apart; 1D	high for linear sequences (noise and multipath)	Linear node discovery using minimal communication, non linear patterns not evaluated
ZigBee wireless sensor networks, TI 2.4 GHz CC2530 nodes, using RSSI [164]	8 nodes	Kalman filter method, Triangle Centroid Location Algorithm (TCLA) based on weighted feature points	Kalman 9.5%;TCLA 13%; Gaussian filter 29%; Mean filter 29.5%	§ low/ moderate	24m; 1m and 2m spacing	moderate (noise)	Focuses on accuracy of RSSI value and optimization of localization algorithm
WiFi RSS, inertial navigation sensor, gyroscope, a biaxial accelerometer, atmospheric pressure sensor [165]	Pedestrian at 1m/s, WiFi at Ts = 300ms, and a new INSmeasurement at δ T = 40 ms.	KNN, Kalman filter, particle filter, statistical algorithm/ML	Mean error lowest being 1.53m Table 1 and Table II full results	§ high/high	moderate, 40m x 40m; 2D, 3D potential	moderate (INS sensor errors)	Dead-reckoning navigation self localization of an autonomous mobile device by fusing pedestrian WiFi signal strength measurements.
ZigBee RSSI [166]	4 nodes emitting @0.67 Hz	Weighted centroid	RMSE 10m, std. dev. 2m	§ Linear in # of beacons	2D	moderate	Tour guide robot application
GPS,ZigBee RSSI [167]	5 sensors	Distributed barycentric coord. est.	1 m	no analysis	10m×10m	moderate (robust to white noise)	Accuracy based on reported initial error and “error ratio”

Continued on next page

* Denotes results obtained from simulation.

† Complexity and robustness apply to the algorithm. When hardware complexity is also known, it is reported as hardware/algorithm complexity.

§ Indicates solution is designed for or feasible in real time

TABLE 10

Indoor Positioning Metrics (continued)

Sensors/ Protocols	Data	Algorithms	Accuracy/ Precision	†Complexity	Scalability/ Space dimension	†Robustness	Notes
TelosB RSSI [168]	20 nodes, 16 channels	Adaptive Kalman filter, weighted RSSI fingerprinting, optimized by memetic alg.	0.6m	§ $<$ 50 ms per generation (Matlab implementation)	10m×10m	high	Requires calibration
ZigBee RSSI [169]	≤ 8 beacons, 60 points	probability map from RSSI	1.5m err., std. dev. 0.8m	§ Linear in # of beacons	30–600 m^2	moderate	Requires training of prob. map
RSS [170]	17 nodes, 100 packets each	Kalman/ moving avg. filt.	1.5m	§ linear in # nodes	$7 \times 3m^2$	high	Kalman filter performed better than moving average
WLAN,RSS and Bluetooth RSSI [171]	4 WiFi and 3 Bluetooth access points	fingerprinting (Bluetooth zone followed by WiFi)	2.3m 1/6: 1.75-2m; 2/3: 2-2.5m; 1/6: 2.5-2.75m	no analysis	2D	moderate	Requires Bluetooth RSSI and WiFi RSS maps
Zigbee RSSI [172]	8 reference and 1 blind nodes	Environment adaptive distance matrix	2.75m	no analysis	16m x 27m,2D	moderate	Requires calibration
*Zigbee RSSI and UWB TOA [173]	number of simulations, nodes not specified	RINPS, RCAPS with 1-hop ranging/ Least Squares	RMSE: Zigbee RSSI <0.4m; UWB TOA <0.2m	no analysis	100m x 100m, 2D, 3D	moderate	Focuses on improving accuracy in cluster-tree (vs mesh) topologies
WiFi RSS and opportunistic GPS [174]	small: 50 locations, 48 visible APs; large: 100 locations, 156 visible APs	hybrid genetic algorithm and gradient descent to compute RSS model; trilateration/Least Squares for mobile devices	mobile device in small: 1.8m/50th percentile, 2.8m/80th; AP small: 2m/50th percentile, 3.3m/80th; AP large: 4m/50th percentile, 7m/80th	§ mobile node; mins to hours for model estimation	small: 27m x 18m; large: 90m x 140m; 2D	moderate; RF model more robust than RF map; robust to gain differences across devices	Does not require collaboration or prior environment knowledge

Continued on next page

* Denotes results obtained from simulation.

† Complexity and robustness apply to the algorithm. When hardware complexity is also known, it is reported as hardware/algorithm complexity.

§ Indicates solution is designed for or feasible in real time

TABLE 10

Indoor Positioning Metrics (continued)

Sensors/ Protocols	Data	Algorithms	Accuracy/ Precision	†Complexity	Scalability/ Space dimension	†Robustness	Notes
WiFi RSS and acoustic ranging (AR) [175]	8 laptops	Bayesian inference	0.7m - 3.5m	secs to mins for 2-5 nodes depending on number of edges	27m x 8m; 2D	makes AR more robust in non-LOS settings; large error for one node can increase error for other nodes	EchoBeep, DeafBeep, and Centaur framework
2.4GHz + 5GHz WiFi; WiFi + accelerometer + gyroscope [176]	7 APs, RSSI measurements every 0.5m	particle filter	1.7m mean, 1.6m median, 4.5m max	no analysis	100m path; 2D	use of dual band increases robustness	2.4 and 5GHz signal strength model errors not well correlated so using both provides more info
	20 APs; hand-carried smartphone; RSSI every 4 secs; 20Hz accelerometer sampling freq	particle filter	1.3m mean, 5.8m max	no analysis	273m path; 2D	combo increases robustness since PDR performed between WiFi measurements	combines WiFi and PDR
Zigbee RSSI [177]	3-16 reference nodes, with coordinates in interval (0m,63.75m)	modified trilateration with proximity learning	0.2-0.6m	§ 50μs-13ms to calculate node position	64m x 64m positioning range; 2D	moderate	focuses on faster computation
Zigbee RSS and hop count [178]	9 reference nodes and 1 target; 2 hops avg; 5sec sampling rate; 7m transmission range	hop count used to determine area, then RSS-based trilateration to refine	1.42m	§ ~35% fewer packets exchanged than for single methods	12m x 8m; 2D	no analysis	fuses hop count and RSS for improved accuracy and lower power consumption than either method alone

Continued on next page

* Denotes results obtained from simulation.

† Complexity and robustness apply to the algorithm. When hardware complexity is also known, it is reported as hardware/algorithm complexity.

§ Indicates solution is designed for or feasible in real time

TABLE 10

Indoor Positioning Metrics (continued)

Sensors/ Protocols	Data	Algorithms	Accuracy/ Precision	†Complexity	Scalability/ Space dimension	†Robustness	Notes
Zigbee fine-grained RSS and time [179]	4 anchor nodes (USRP N210), 1 reference node (TelosB); TelosB target nodes broadcasting 5 packets/sec, 20-30 test positions	learning-based fingerprinting, fusing RSS and DT-DOA; KNN-RF pattern matching	1.61m mean (range 1.09m - 1.86m); 36% improvement over traditional RSS-based fingerprinting	high	16m x 18m; 2D	moderate	requires site survey for RSS fingerprint collection, and synchronization of anchor nodes (with GPS receivers via extended antennas)

* Denotes results obtained from simulation.

† Complexity and robustness apply to the algorithm. When hardware complexity is also known, it is reported as hardware/algorithm complexity.

§ Indicates solution is designed for or feasible in real time

As mentioned in Section 6.1, the widely cited Liu 2007 survey [156] compared (in their Table 1) approximately 20 different systems or solutions across a variety of performance metrics, including accuracy, precision, complexity, robustness, and scalability. In our Table 10, we use a similar format to provide a representative sample of fusion-oriented work published after their 2007 survey which also included some experimental results. Our columns are explained below, along with some observations and discussion about the information provided in each.

Sensors/Protocols: indicates which types of technologies, sensors, and/or protocols are used in the cited work's experiments, while an asterisk denotes whether the experiments were simulated or real-world. Due largely to attempts to take advantage of ubiquitous, existing wireless infrastructure, the preponderance of solutions use received signal strength (RSS) or received signal strength indicator (RSSI) as a primary feature, regardless of the specific RF technology used. Note that while sometimes used interchangeably, there is a technical difference between RSS, which is a raw, power value measured in decibel-milliwatts (dBm), and RSSI, which is a unitless, relative indicator derived from raw RSS values for which there is no standard (and therefore may vary across hardware vendors/implementations). Regardless, both RSS and RSSI vary with distance between the transmitter and receiver. In an indoor environment with limited line of sight (LOS), multipath propagation limits the accuracy of RSS-based methods. Therefore, in the quest for greater accuracy, we see the realization of prior survey predictions that research would expand into hybrid methods and fusing the RSS feature with information from other sources (such as navigational and acoustic sensors) that may help pinpoint a device's location. Notable since the Liu 2007 survey [156] is the increase in Zigbee-related research, as this relatively newer technology becomes more popular. Given that Zigbee is intended for lower-power, mesh network applications in which individual fixed nodes may be in closer proximity to each other than, for instance, WiFi access points, it presents an opportunity for finer-resolution localization.

Data: includes details authors provided about experimental data, including number of nodes deployed, channels and frequencies used, sampling rates, number of simulations performed, etc. Reported details vary widely across publications; many authors provide no indication of actual hardware/channels/frequencies used, the number of experiments run, quantity of samples obtained, or attempts to vary the quantity or quality of the data. While most report how many fixed and target nodes were deployed (often under 10 combined), most do not investigate the effects of more or fewer fixed nodes on the results.

Algorithms: describes the algorithms/methods used in the work. RSS-based trilateration is popular since it requires neither angle measurements nor time synchronization of devices as other methods may require. Variations on weighted centroid algorithms try to improve accuracy by giving higher weight to closer or more reliable reference nodes as determined by higher RSS values or higher link quality, for instance. Due to the previously mentioned multipath propagation issue, and the lack of good models for this phenomenon, fingerprinting-based methods are often used with RSSI data. With this technique, RSSI values from nodes within range are measured and recorded (such as in a database) at numerous locations within the indoor space to form an RSSI fingerprint map. The RSSI readings of a target device then may be compared (e.g., via a pattern matching algorithm) to this map in order to determine its most likely location. Kalman filters are often used to reduce the relatively high noise in RSSI values, which are heavily influenced by the indoor environment. Particle filters and dead reckoning techniques are often used in dynamic location tracking (rather than static location determination) applications.

Accuracy/Precision: summarizes the quantitative results reported by the authors which, again, is highly variable. Some results are reported as straight Euclidean distance error in meters or as a percentage, while others report root-mean-square (RMSE) values. Some include standard deviations, percentiles, and/or percent improvement over comparison methods, while other do not. In addition, note that in most cases the solution is trying to determine the location of a static, target node, but in a few cases they are tracking the position of a moving target along a path. In the latter case, accuracy may be determined by examining cumulative error along a navigation path and final destination. Another point to consider when comparing accuracy across solutions is the scope of the problem being solved. For instance, the 50 cm error reported by [163] seems out-of-the-ordinary compared to others until noting that it is applicable to linear node discovery (1D) only. Also of import to accuracy expectations is the up-front calibration or mapping required (versus systems designed to operate with no prior knowledge or site survey). Most localization errors reported here are in the range of one to three meters, with only a small number reporting sub-meter error results.

Complexity: where provided by authors, includes reported details about algorithmic complexity or runtimes. Qualitative assessments such as low, moderate, and high may be provided when published information adequately supported such an evaluation. In the case of two such assessments provided and separated by a forward slash, the first applies to algorithmic and the second to hardware complexity. The special symbol § denotes solutions that are designed for or feasible in real time. Most authors provide no analysis of the algorithmic complexity of their solution, nor do they necessarily publish enough detail for others to perform this analysis. Those solutions intended for real-time applications are more likely to come with runtime measurements or estimates.

Scalability/Space Dimension: includes details authors provided about the size of the space in which experiments were conducted. Where possible to assess, we also include an indication of the dimensions (1D, 2D, and/or 3D) in which the approach was tested or to which it may extend. The vast majority of approaches tackle the 2D problem; those attempting 3D localization generally require additional sensors (such as the atmospheric pressure sensor considered in [165], which is a good example of multi-sensor data fusion for both real-time and static indoor localization). Some 3D localization solutions attempt only to discern the floor of a building on which a device is located rather than an exact height for the third dimension.

Robustness: summarizes authors' comments on robustness, and/or includes a qualitative assessment (such as low, moderate, or high), if possible, of the solution's robustness to interference, noise, incomplete information, etc. The correlation between RSS/RSSI values and distance, upon which the majority of approaches are based, has been experimentally evaluated by several researchers, such as Dong and Dargie in 2012 [180] who concluded that this signal alone is unacceptably error-prone and unreliable for indoor localization. An interesting 2017 study by Konings et al. [181] investigated how microwave ovens and WiFi interference may affect Zigbee RSSI values, and determined that neither affect the RSSI values of packets that are correctly received, but WiFi interference may cause increased packet loss or corruption as data rates increase. RSSI-based fingerprinting methods have their own robustness issues as well; for instance, the number of people present and moving in an indoor environment can affect location fingerprints. Some authors attempt to address this issue by creating separate daytime and nighttime fingerprint maps, while others don't address it at all.

Notes: includes other thoughts relevant to the work, such as take-aways, caveats, the goal/focus of the work, or notes on calibration/preparation required. Since most methods rely on cooperation and/or calibration, it is interesting to note Chintalapudi et al. 2010 [174] as an example that requires neither pre-deployment calibration nor cooperation/collaboration. It does, however, rely on an occasional "opportunistic GPS" location of an edge node in order to then determine the absolute locations of other nodes for which relative positions have been previously calculated.

As we've described, comparing the performance of differing approaches is particularly challenging, since no official benchmark or standard exists against which authors evaluate their solutions and report their results. To address this issue, beginning in 2014 Microsoft has held the yearly Microsoft Indoor Localization Competition, in which teams from both academia and industry deploy their solutions in a somewhat realistic environment which allows a direct comparison of their performance. Lymberopoulos and Liu provide an excellent and thorough accounting of the 2014-2017 competitions and results in [182], while the 2018 results are available online [183]. The competition includes both infrastructure-free and infrastructure-based categories. Infrastructure-free solutions don't require the deployment of additional hardware, relying solely on pre-existing WiFi and inertial sensors available on most mobile phones; the best results for these systems have been in the one to two meter range. In contrast, infrastructure-based solutions rely on additional technologies and custom-deployed hardware associated with, for example, sound/ultrasound, UWB, and visible light; many of these systems consistently achieve sub-meter accuracy, as low as 0.17 m for one 3D UWB approach. During the 2014 competition, the competition-achieved accuracies were compared to the self-tested/self-reported accuracies in an attempt to quantify the bias of the highly controlled test environments used in most self-evaluations. Indeed, the accuracy of most entries declined by 1.5 to 4 meters. Additionally, the 2014 competition investigated the impact of furniture setup on the competing solutions by changing the layout after calibration in one room (which did decrease accuracy) and not in the other. Interestingly, the fact that the venue varies year-to-year means that the indoor evaluation space varies as well, and this has proven to affect the overall results such that they cannot be compared from year to year.

Despite another decade of research into improving the accuracy of largely RSS-based indoor localization techniques, the WiFi-based results reported here don't differ tremendously from those reported in the Liu 2007 survey [156]. For the most part, accuracy remains in the one-to-three meter range for WiFi RSSI approaches, while some Zigbee-based solutions claim slightly better accuracy of under one meter. As reported previously by other surveys, and borne out by this sampling, achieving substantially smaller errors (e.g., under 200 centimeters) requires incorporating relatively more expensive and less ubiquitous technology such as ultrasound or ultra-wideband (UWB).

7. DISCUSSION

This survey considers many application areas in which the recent expansion of IoT provides information about the environment in which it operates that would not be available without these new, prolific, inexpensive sensors. The applications are as diverse as the devices themselves, spanning several questions of relevance to government analysts, including:

- How is traffic in a given area different than what is typical?
- What is the occupancy of this building (or room)?
- Where is this device located with respect to known locations?
- Are conditions in this area ideal for the spread of a particular illness?

Throughout the broad spread of work considered over the course of this study, much of it relies on methodologies that have been proven in numerous contexts over the years. The indoor positioning material in particular makes heavy use of tried and true methodologies from signal and information processing (e.g., ARIMA modeling and Kalman filtering) that have been studied over several decades. Indeed, many of these methodologies are applicable for arbitrary RF emitters and are not IoT specific. The same can be said for technologies such as simultaneous localization and mapping, which draw heavily on important research problems in the robotics area. Robotics is an especially relevant area to consider for multimodal data fusion, as there are frequently multiple sensor modalities used on a platform, such as RF, imagery, and lidar.

Ambient assisted living and occupancy detection are two other areas in which there is frequently multimodal data fusion. In addition to using RF methods like RSSI, positioning methods in these application areas often incorporate other sources of information, such as accelerometer readings, to provide a more reliable model of the scene. Indeed, the winners of the 2D (infrastructure-free) positioning challenge in the Microsoft Indoor Localization Challenge have been based on pedestrian dead reckoning [183], which often uses accelerometer measurements. These multimodal methodologies typically use off-the-shelf machine learning methods like decision trees, support vector machines, or deep neural networks, and the novelty is primarily in the application and data fusion.

The very ubiquity and diversity of IoT, however, presents research challenges when considering broad questions such as, "How much information can be obtained from IoT sources?" As is clear when viewing the literature at a high level, it is often difficult to obtain a direct comparison between candidate fusion methods without explicitly setting out to make such a comparison in a research program. To reliably gauge how these solutions would perform in practice, a rigorous suite of experiments must be run in which a controlled context is defined and each method is evaluated in the same context. For military and intelligence applications, testing in stressed conditions—e.g., lack of calibration, node dropouts, interference—would be essential to fully evaluate applicability to high-consequence, uncooperative scenarios. The Microsoft challenge could be a useful model, and experimentation could be run for a set of representative applications in a few well-defined scenarios under conditions of varying difficulty.

Recent projections suggest the number of IoT devices in operation will soon exceed the number of people on Earth by an order of magnitude [4]. As this trend continues, greater processing capability at the edge will likely increase as well [63], presenting new challenges and opportunities. Where possible, experimentation such as that recommended here could make use of new edge processing technology, such as existing system-on-chip solutions used as a baseline in [169]. This would provide an additional layer of information about the capabilities of IoT: What information (and how much) can be inferred in the field based exclusively on local measurements, and what requires fusion with broader data? This is a truly rich area for research and development with many candidate solutions ready for testing, as outlined in this report.

8. SUMMARY

This report documents the findings of a six-month literature review of the state of the art in data fusion methods for data available from Internet of Things sources. Such methods have significant potential for use by government analysts to gather information supplemental to that obtained via specialized sensors and tradecraft. IoT can provide information on conditions on the ground at a specific time and place, such as traffic, weather, and human activity, all of which can be essential to describe an area of interest. While there is tremendous potential with this data, these prolific devices can be unreliable and information from any one device may not be useful. The fusion methods described in this report aim to combine measurements to provide greater robustness, leveraging the volume of data to overcome noise and errors arising from using a vast sea of inexpensive, uncontrolled sensors.

The literature survey turned up a substantial body of research covering numerous application domains, all of which benefit from fusion of multimodal data made available through the proliferation of IoT. To focus on methods that are potentially relevant to government analysts, the focus of the study was directed at a representative sample of applications, including smart cities ("What are traffic conditions in this area?"), public health ("Is a disease outbreak likely in this area?"), occupancy detection ("How many people are in this building?"), and mapping and localization ("What are typical paths taken through this area?"). As there is a particularly large body of work on indoor positioning, this area was researched in greater depth to determine what features provide greater performance.

While a plethora of methods have been applied to the various applications, a consistent challenge in conducting the survey was providing a baseline point of comparison and fixed sets of metrics on which the solutions can be evaluated. In the indoor positioning literature, where the literature is focused enough that this comparison is somewhat possible, it is clear that using additional modalities (such as ultrawideband signals and pedestrian dead reckoning) beyond the traditional received signal strength indicator can substantially improve performance (to errors as small as 0.5 m in laboratory settings).

An ideal next step in this research area would be to choose a broad set of analytics where IoT data are useful and run experiments in a setting where the environment (including obstructions and interference) can be controlled and the effects of these variables can be rigorously quantified. IoT will only become more prevalent in the years to come, and such a suite of experiments would put government professionals in the best possible position to take advantage of the vast sea of data it provides.

9. GLOSSARY

AAL Ambient assisted living, a sub-area of ambient intelligence which seeks to enable elderly people or those with disabilities to live independently

actuator A device that provides the mechanism for a control system to affect the physical environment (e.g. open or close a valve)

AI Artificial intelligence, an area of computer science that seeks to build intelligent machines that can "think" or work like humans

anchor node In an indoor positioning context, a node (such as a *WAP*) in a fixed, known location that serves as a reference when trying to determine the position of a *target node*)

ANN Artificial *neural network* (see *neural network*)

AOA Angle of arrival, a positioning method that requires knowing or calculating the angle and distance to multiple reference points

AP Access point, a networking hardware device, typically connected to a wired network, that allows *WiFi* devices to connect to that wired network

ARIMA Autoregressive integrated moving average, a class of model fitted to time series data either to better understand the data or to predict future points in the series (forecasting)

AV Autonomous vehicle

BAN Body area network, a wireless network of wearable or body-embedded devices, which may consist of multiple, miniaturized body sensor units (BSUs) together with a single body central unit (BCU)

Bayes' rule An equation describing the probability of an event based on prior knowledge of conditions that might be related to the event: (P(A|B) = P(B|A)*P(A)/P(B)

Bayesian inference A statistical method of combining evidence according to probability theory, based on *Bayes' rule*

Bayesian networks A type of statistical model that represents variables and their dependencies with a directed acyclic graph (DAG) that can be used to compute the probabilities of various outcomes

Bluetooth A wireless technology standard (IEEE 802.15.1), managed by the Bluetooth Special Interest Group, for exchanging data over short distances using short-wavelength *UHF* radio waves

BSN Body sensor network (see *BAN*)

C4.5 An *ML* algorithm used to build decision trees from training data

calibration In an indoor positioning context, refers to preparatory work necessary, such as surveying a site to collect and store signal measurements, to serve as reference point for accurate future operation

cloud In a cloud computing paradigm, devices forward data to shared, centralized storage and compute services

complementary fusion Fusing data from different portions of a picture into a more complete picture (for instance, temperature readings from different locations within a greenhouse in order to better understand the temperature status of the greenhouse as a whole) [58]

cooperative localization A node localization paradigm in which nodes share information to help each other determine their locations

cooperative fusion Fusing multiple independent sources of information into new, more complex information (e.g., temperature combined with light yields a better understanding of overall growing conditions within the greenhouse) [58]

DARPA Defense Advanced Research Projects Agency

data fusion Combining multiple data sources to obtain improved information (cheaper, higher quality, or more relevant/useful) [58]

data aggregation Combining data via summarization techniques often used to reduce redundancy and conserve resources, such as in suppression (e.g., discarding duplicates) and packaging (i.e., grouping multiple observations together) [58]

dead reckoning A method of current position estimation and future position prediction based on a previously known position and direction, the time at which it was obtained, and travel speed over the intervening period

Dempster-Shafer theory A framework for reasoning with uncertainty involving belief and plausibility

Dempster-Shafer inference A math-based belief accumulation and evidential reasoning approach based on *Dempster-Shafer theory*, used to fuse data from different types of sensors without requiring a-priori probabilities

DFL Device-free localization, a method for detecting and tracking entities that are not carrying any radio device or participating actively in the localization process, generally by using signal perturbations from existing infrastructure, such as *WiFi* networks

DNN Deep *neural network*, a *neural network* with many layers between input and output (see *neural network*)

DSET Dempster-Shafer evidence theory, (see *Dempster-Shafer theory*)

DSN Distributed sensor network, a distributed, non-centralized version of a *WSN* (see *WSN*)

DST See *Dempster-Shafer theory*

DTDOA Differential time difference of arrival, a variation on *TDOA* (see *TDOA*)

DTN Delay-tolerant networks lack continuous connectivity and employ store-and-forward strategies to compensate; variations in literature include opportunistic networks, disconnected mobile ad hoc networks, time-varying networks, intermittently connected networks (ICNs), and extreme networks

edge In an IoT context, edge devices are the "things" containing *sensors*, *actuators*, and communications capabilities, while the edge computing paradigm emphasizes processing data directly on the edge devices or the devices to which they are directly attached in order to reduce communications bandwidth to any centralized storage or compute services

fingerprinting A localization method that matches against a previously collected database of location-dependent characteristics (e.g., *RSSI* values)

fixed node See *anchor node*

fog Fog computing, fog networking, and fogging all refer to an IoT architecture that is essentially similar to *edge* computing, but in which processing may also occur slightly further away from the *edge* devices, such as in an IoT gateway on the *LAN*

fuzzy logic An inference method of approximate reasoning, drawing imprecise conclusions from imprecise evidence; useful in real-world situations that are difficult to model precisely such as node positioning and navigation

Gaussian noise Statistical noise with a probability density function (PDF) equal to that of the normal distribution (i.e., bell curve)

GPS Global positioning system, a satellite-based system that provides geolocation and time information to a GPS receiver within unobstructed line of sight of a GPS satellite

HetNet Heterogeneous network, connecting devices with different operating systems, access technologies, and/or protocols

HMM Hidden Markov model, a tool for representing probability distributions over sequences of observations

ICN Information-centric networking, an information-centric rather than host-centric networking paradigm that addresses intermittent connectivity (such as in *DTNs*) by relying on in-network caching and replication

INS Inertial navigation system, a navigation system using a variety of self-contained sensors (such as accelerometers, gyroscopes, and magnetometers) to continuously calculate position by *dead reckoning*

IoT Internet of Things, a network of physical devices equipped with *sensors*, *actuators*, embedded software, and connectivity which enables them to connect and exchange data

IR Infrared radiation, a type of electromagnetic radiation with wavelengths longer than those of visible light making it generally invisible to the human eye

Kalman filter A type of estimation algorithm used to fuse low-level redundant data and remove outliers, useful when a linear model can describe observations over time with statistical (Gaussian) noise; often used in guidance and navigation applications

KNN K-nearest neighbor, a non-parametric pattern recognition algorithm used for classification and regression, where the input consists of the k closest training examples in the feature space, and the output is a class determination (in the case of classification) or the property value for the object (in the case of regression)

LAN Local area network, connects computers within a limited area (such as a building), typically via Ethernet or *WiFi*

least squares A method of estimation that searches for a function that best fits a set of input samples, most suitable when the estimated parameter does not have substantial uncertainty, as the method is strongly affected by noise

LIDAR Light detection and ranging, a detection system that works on the principle of radar, but uses light from a laser to measure the distance to a target by illuminating it with pulsed laser light and measuring the reflected pulses with a sensor

linear regression A statistical method that models the relationship between two continuous quantitative variables

LPWA(N) Low-power, wide-area network, designed to allow long-range communications at a low bit rate among low-power connected objects (such as battery-operated *sensors*)

M2M Machine-to-machine, as in communications

MANET Mobile ad hoc network (see *WANET*)

MAP Maximum a posteriori, a method for estimating an unknown quantity with a known prior distribution

MBAN Medical body area network (see *BAN*)

MCC Mobile *cloud* computing (see *cloud*)

mesh network A network topology in which there is a high degree of connection between all nodes and all nodes can route data, reducing dependency on any one node and increasing robustness

middleware In an IoT context, software that serves as an interface or bridge between the sensor/device layer and the application layer, providing basic messaging, routing, and data transformation services

ML Machine learning, a sub-field of *AI* that uses statistical techniques to automatically build and improve models from data

MLE Maximum likelihood estimation, a method of estimating the parameters of a statistical model, given observations

modality In the *data fusion* context, refers to data obtained from a *sensor* of a particular type (e.g., temperature data obtained from a temperature sensor)

multilateration The process of determining the location of an object using the difference between distances to or arrival times of signals from known reference points (such as in *TDOA*), based on the intersection of hyperbolic curves and generally requiring four reference points

multimodal In the *data fusion* context, refers to data obtained from multiple *modalities*, or types of *sensors*, (e.g., temperature, light, and sound), often used in *cooperative fusion* in order to gain a more complete understanding than could be obtained from one *modality* alone

multipath propagation A phenomenon wherein the same signals may reach a receiver via two or more paths due to bouncing off reflective surfaces, posing significant challenges to indoor positioning methods

neural network An *ML* paradigm or set of algorithms inspired by the human brain, involving multiple layers of densely connected nodes with weighted edges through which data propagates to eventually yield a single result

NN *Neural network* (see *neural network*)

NOI Node of interest (see *target*)

PAN Personal area network, connects an individual person's electronic devices (such as laptops, smart phones, and tablets) to each other and/or the Internet

particle filter A type of recursive, statistical signal processing (aka Sequential Monte Carlo) which uses a genetic mutation-selection sampling approach with a set of particles (or samples), useful in noisy systems and/or with partial observations; often used in target tracking and node location discovery applications

PDR Pedestrian dead reckoning, *dead reckoning* for pedestrians (see *dead reckoning*)

QoS Quality of service, an indicator of network performance taking into account aspects such as packet loss, throughput, latency, and availability

range-free localization A localization method that uses non-distance based information, such as hop count, to infer the proximity of a node to some reference points and typically uses centroid calculations to estimate a node's position, generally producing less accurate results than *range-based localization* methods

range-based localization A localization method that uses timing, direction, or signal strength-based techniques to estimate distance and calculate a node's position using trilateration or multilateration, generally producing more accurate results than *range-free localization* methods

redundant fusion Fusing multiple instances of the same type of data to increase reliability, accuracy, and/or confidence (e.g., multiple temperature readings from the same location in a greenhouse to get a more reliable or accurate understanding of the temperature at that location) [58]

RF Radio frequency

RFID Radio frequency identification, a technology that uses electromagnetic fields to identify and track passive or active electronic tags (containing information) attached to or embedded within objects

RMSE Root-mean-square error, the standard deviation of prediction errors, a scale-dependent measure of accuracy used to compare prediction errors within a dataset

RSS Received signal strength, a raw, power value measured in decibel-milliwatts (dBm), which varies with distance between a transmitter and receiver

RSSI Received signal strength indicator, a unitless, relative indicator derived from raw *RSS* values for which there is no standard (and therefore may vary across hardware vendors/implementations), which, like *RSS*, varies with distance between a transmitter and receiver

RT Radio tomography, a recent application of tomographic techniques (imaging by sections or sectioning through the use of any kind of penetrating wave) to radio waves, used for mapping and localization across many radio or radar nodes

RTI Radio tomographic imaging (see *RT*)

sensor A device that detects and/or measures events or changes in its environment

SLAM Simultaneous localization and mapping, the problem of building a map of an unknown environment while navigating through it and keeping track of one's location within it

SOA Service-oriented architecture, a type of distributed software architecture wherein independent components (services) communicate over a network to provide data or functionality to each other

SPAN Smart phone ad hoc network, a type of *MANET* in which smart phones form peer-to-peer networks leveraging their own, non-cellular communications capabilities (typically *Bluetooth* or *WiFi*)

SVM Support vector machine, a supervised *ML* method used for classification, regression and outlier detection, representing the training samples as points in space, mapped so that separate categories are divided by as wide a gap as possible, and predicting the category of a new sample based upon which side of the gap it falls

target or target node In an indoor positioning context, the node (typically mobile) for which we seek to determine a location

TDOA Time difference of arrival, a positioning method based on the difference in propagation time of a signal from a transmitter to multiple reference points, which requires knowing only arrival times and not transmission time

TelosB An IEEE 802.15.4 compliant wireless sensor node based on the original open-source TelosB platform design developed and published by the University of California, Berkeley, including sensors that measure temperature, relative humidity, and light

TOA Time of arrival, a positioning method that uses the time taken for a signal to arrive at receiver to calculate its distance from the transmitter, which requires knowing both the time of transmission and receipt

triangulation The process of determining the location of an object by forming triangles to it from at least two known references points with measured angle information

trilateration The process of determining the location of an object by measuring distances (not angles) to at least three known reference points and using the geometry of intersecting circles, spheres or triangles

UHF Ultra high frequency

ultrasound Sound waves with frequencies higher than those detectable by human hearing

UWB Ultra-wideband, a radio technology operational over a large portion of the radio spectrum, using very low energy for high-bandwidth communications over short distances; traditionally used in non-cooperative radar imaging, and recently used in tracking and localization applications

VANET Vehicular ad hoc network, a *MANET* in the context of a vehicle (see *MANET*)

WANET Wireless ad hoc network, a decentralized, dynamic, wireless network in which nodes of equal status may move around and self-organize instead of relying on fixed infrastructure

WAP Wireless access point (see *AP*)

WBAN Wireless body area network (see *BAN*)

weighted centroid A variation on the centroid localization algorithm (based on *trilateration*) in which weighting factors are added in order to increase or decrease the influence of the distance information obtained from one of the reference points

WiFi Wireless *RF* technology based on the IEEE 802.11 standards for local area networking of devices

WLAN Wireless local area network, a *LAN* that uses wireless communications technology

WLS Weighted least squares, a generalization of ordinary *least squares* (see *least squares*)

WMN Wireless *mesh network* (see *mesh network*)

WPAN Wireless personal area network, low-powered, short-distance (up to a few meters) *PAN* using wireless network technology such as *Bluetooth* or *Zigbee*

WSN Wireless sensor network, a network of spatially dispersed *sensor* nodes, not limited to a particular topology

Zigbee A specification (based on IEEE 802.15.4) for a low-power, low data rate, and close proximity wireless ad hoc network (see *WANET*)

REFERENCES

[1] K. Ashton, "That 'Internet of Things' thing," *RFID Journal*, 2009.

[2] Carnegie Mellon Computer Science Department, "The "only" Coke machine on the internet." https://www.cs.cmu.edu/~coke/history_long.txt, (accessed September 25, 2018).

[3] Z. Yang, Y. Yue, Y. Yang, Y. Peng, X. Wang, and W. Liu, "Study and application on the architecture and key technologies for IoT," in *2011 International Conference on Multimedia Technology*, pp. 747–751, July 2011.

[4] Louis Columbus, "2017 Roundup of Internet of Things Forecasts." https://www.forbes.com/sites/louiscolumbus/2017/12/10/2017-roundup-of-internet-of-things-forecasts, December 2017 (accessed October 13, 2018).

[5] A. C. Sarma and J. Girão, "Identities in the future Internet of Things," *Wireless Personal Communications*, vol. 49, pp. 353–363, May 2009.

[6] B. L. Risteska Stojkoska and K. V. Trivodaliev, "A review of Internet of Things for smart home: Challenges and solutions," *Journal of Cleaner Production*, vol. 140, pp. 1454–1464, 2017.

[7] P. Bellavista, G. Cardone, A. Corradi, and L. Foschini, "Convergence of MANET and WSN in IoT urban scenarios," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3558–3567, 2013.

[8] M. Ryu, J. Kim, J. Yun, M. Ryu, J. Kim, and J. Yun, "Integrated semantics service platform for the Internet of Things: A case study of a smart office," *Sensors*, vol. 15, no. 1, pp. 2137–2160, 2015.

[9] H. Schaffers, N. Komninos, M. Pallot, B. Trousse, M. Nilsson, and A. Oliveira, "Smart Cities and the Future Internet: Towards Cooperation Frameworks for Open Innovation," in *The Future Internet* (J. Domingue, A. Galis, A. Gavras, T. Zahariadis, D. Lambert, F. Cleary, P. Daras, S. Krco, H. Müller, M.-S. Li, H. Schaffers, V. Lotz, F. Alvarez, B. Stiller, S. Karnouskos, S. Avessta, and M. Nilsson, eds.), Lecture Notes in Computer Science, pp. 431–446, Springer Berlin Heidelberg, 2011.

[10] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," *Information Fusion*, vol. 35, pp. 68–80, 2017.

[11] A. Augustin, J. Yi, T. Clausen, W. Townsley, A. Augustin, J. Yi, T. Clausen, and W. M. Townsley, "A study of LoRa: Long range & low power networks for the Internet of Things," *Sensors*, vol. 16, no. 9, p. 1466, 2016.

[12] K. Yasumoto, H. Yamaguchi, and H. Shigeno, "Survey of real-time processing technologies of IoT data streams," *Journal of Information Processing*, vol. 24, pp. 195–202, 3 2016.

[13] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Commun. ACM*, vol. 43, no. 5, pp. 51–58, 2000.

[14] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[15] J. Lee, Y. Su, and C. Shen, "A comparative study of wireless protocols: Bluetooth, UWB, Zigbee, and Wi-Fi," in *IECON 2007 - 33rd Annual Conference of the IEEE Industrial Electronics Society*, pp. 46–51, 2007.

[16] Z. J. Haas, J. Deng, B. Liang, P. Papadimitratos, and S. Sajama, "Wireless ad hoc networks," *Wiley Encyclopedia of Telecommunications*, 2003.

[17] R. Bruzgiene and L. N. a. T. Adomkus, "MANET network in Internet of Things system," *Ad Hoc Networks*, 2017.

[18] J. Kim and J. Lee, "Cluster-based mobility supporting WMN for IoT networks," in *2012 IEEE International Conference on Green Computing and Communications*, pp. 700–703, 2012.

[19] Y. Liu, K.-F. Tong, X. Qiu, Y. Liu, and X. Ding, "Wireless mesh networks in IoT networks," in *2017 International Workshop on Electromagnetics: Applications and Student Innovation Competition*, pp. 183–185, 2017.

[20] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, "Long-range communications in unlicensed bands: the rising stars in the IoT and smart city scenarios," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 60–67, 2016.

[21] L. Li, H. Xiaoguang, C. Ke, and H. Ketai, "The applications of WiFi-based wireless sensor network in Internet of Things and smart grid," in *2011 6th IEEE Conference on Industrial Electronics and Applications*, pp. 789–793, 2011.

[22] S. Li, L. D. Xu, and X. Wang, "Compressed sensing signal and data acquisition in wireless sensor networks and Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2177–2186, 2013.

[23] Q. Zhu, R. Wang, Q. Chen, Y. Liu, and W. Qin, "IoT gateway: Bridging wireless sensor networks into Internet of Things," in *2010 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing*, pp. 347–352, 2010.

[24] L. Mainetti, L. Patrono, and A. Vilei, "Evolution of wireless sensor networks towards the Internet of Things: A survey," in *SoftCOM 2011, 19th International Conference on Software, Telecommunications and Computer Networks*, pp. 1–6, 2011.

[25] F. L. Lewis, "Wireless sensor networks," in *Smart Environments*, pp. 11–46, Wiley-Blackwell, 2005.

[26] P. Schulz, M. Matthe, H. Klessig, M. Simsek, G. Fettweis, J. Ansari, S. A. Ashraf, B. Almeroth, J. Voigt, I. Riedel, A. Puschmann, A. Mitschele-Thiel, M. Muller, T. Elste,

and M. Windisch, "Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 70–78, 2017.

[27] A. Faggiani, E. Gregori, L. Lenzini, V. Luconi, and A. Vecchio, "Smartphone-based crowdsourcing for network monitoring: Opportunities, challenges, and a case study," *IEEE Communications Magazine*, vol. 52, no. 1, pp. 106–113, 2014.

[28] R. Golchay, F. L. Mouël, S. Frénot, and J. Ponge, "Towards bridging IoT and cloud services: Proposing smartphones as mobile and autonomic service gateways," *arXiv:1107.4786 [cs] http://arxiv.org/abs/1107.4786*, 2011.

[29] L. Zhang, J. Liu, and H. Jiang, "Energy-efficient location tracking with smartphones for IoT," in *2012 IEEE SENSORS*, pp. 1–4, 2012.

[30] C. Pereira, J. Rodrigues, A. Pinto, P. Rocha, F. Santiago, J. Sousa, and A. Aguiar, "Smartphones as M2M gateways in smart cities IoT applications," in *2016 23rd International Conference on Telecommunications (ICT)*, pp. 1–7, 2016.

[31] Z. Zhang, "Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges," *IEEE Communications Surveys Tutorials*, vol. 8, no. 1, pp. 24–37, 2006.

[32] K. Fall, "A delay-tolerant network architecture for challenged internets," in *Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, SIGCOMM '03, pp. 27–34, ACM, 2003.

[33] E. P. C. Jones, L. Li, J. K. Schmidtke, and P. A. S. Ward, "Practical routing in delay-tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 8, pp. 943–959, 2007.

[34] F. M. Al-Turjman, A. E. Al-Fagih, W. M. Alsalih, and H. S. Hassanein, "A delay-tolerant framework for integrated RSNs in IoT," *Computer Communications*, vol. 36, no. 9, pp. 998–1010, 2013.

[35] S. Cc, V. Raychoudhury, G. Marfia, and A. Singla, "A survey of routing and data dissemination in delay tolerant networks," *Journal of Network and Computer Applications*, vol. 67, pp. 128–146, 2016.

[36] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 855–873, 2017.

[37] K. E. Nolan, W. Guibene, and M. Y. Kelly, "An evaluation of low power wide area network technologies for the Internet of Things," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 439–444, 2016.

[38] Jae-Hwan Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, vol. 1, pp. 22–31, IEEE, 2000.

[39] S. Yi, C. Li, and Q. Li, “A survey of fog computing: Concepts, applications and issues,” in *Proceedings of the 2015 Workshop on Mobile Big Data - Mobidata ’15*, pp. 37–42, ACM Press, 2015.

[40] M. Samaniego and R. Deters, “Blockchain as a service for IoT,” in *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pp. 433–436, 2016.

[41] F. Samie, V. Tsoutsouras, L. Bauer, S. Xydis, D. Soudris, and J. Henkel, “Computation offloading and resource allocation for low-power IoT edge devices,” in *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, pp. 7–12, 2016.

[42] M. Satyanarayanan, “The emergence of edge computing,” *Computer*, vol. 50, no. 1, pp. 30–39, 2017.

[43] W. Shi and S. Dustdar, “The promise of edge computing,” *Computer*, vol. 49, no. 5, pp. 78–81, 2016.

[44] M. Chiang and T. Zhang, “Fog and IoT: An overview of research opportunities,” *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 854–864, 2016.

[45] P. Mishra, S. Kumar, and S. Mishra, “Significance of fog computing in IoT,” SSRN Scholarly Paper ID 3170299, Social Science Research Network, 2018.

[46] M. Yannuzzi, R. Milito, R. Serral-Gracià, D. Montero, and M. Nemirovsky, “Key ingredients in an IoT recipe: Fog computing, cloud computing, and more fog computing,” in *2014 IEEE 19th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 325–329, 2014.

[47] A. V. Dastjerdi and R. Buyya, “Fog computing: Helping the Internet of Things realize its potential,” *Computer*, vol. 49, no. 8, pp. 112–116, 2016.

[48] M. Aazam and E. Huh, “Fog computing and smart gateway based communication for cloud of things,” in *2014 International Conference on Future Internet of Things and Cloud*, pp. 464–470, 2014.

[49] C. Doukas and I. Maglogiannis, “Bringing IoT and cloud computing towards pervasive healthcare,” in *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pp. 922–926, 2012.

[50] S. Nastic, S. Sehic, D. Le, H. Truong, and S. Dustdar, “Provisioning software-defined IoT cloud systems,” in *2014 2nd International Conference on Future Internet of Things and Cloud (FiCloud)(FICLOUD)*, pp. 288–295, 2014.

[51] A. Sajid, H. Abbas, and K. Saleem, “Cloud-assisted IoT-based SCADA systems security: A review of the state of the art and future challenges,” *IEEE Access*, vol. 4, pp. 1375–1384, 2016.

[52] H. Truong and S. Dustdar, "Principles for engineering IoT cloud systems," *IEEE Cloud Computing*, vol. 2, no. 2, pp. 68–76, 2015.

[53] M. Gerla, E. Lee, G. Pau, and U. Lee, "Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds," in *2014 IEEE World Forum on Internet of Things (WF-IoT)*, pp. 241–246, 2014.

[54] J. Zhang, "Multi-source remote sensing data fusion: status and trends," *International Journal of Image and Data Fusion*, vol. 1, no. 1, pp. 5–24, 2010.

[55] G. Ditzler, M. Roveri, C. Alippi, and R. Polikar, "Learning in nonstationary environments: A survey," *IEEE Computational Intelligence Magazine*, vol. 10, no. 4, pp. 12–25, 2015.

[56] Y. Wu, E. Y. Chang, K. C.-C. Chang, and J. R. Smith, "Optimal multimodal fusion for multimedia data analysis," in *Proceedings of the 12th Annual ACM International Conference on Multimedia*, MULTIMEDIA '04, pp. 572–579, ACM, 2004.

[57] G. Deak, K. Curran, and J. Condell, "A survey of active and passive indoor localisation systems," *Computer Communications*, vol. 35, no. 16, pp. 1939–1954, 2012.

[58] E. F. Nakamura, A. A. F. Loureiro, and A. C. Frery, "Information fusion for wireless sensor networks: Methods, models, and classifications," *ACM Comput. Surv.*, vol. 39, Sept. 2007.

[59] P. Jesus, C. Baquero, and P. S. Almeida, "A survey of distributed data aggregation algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 381–404, 2015.

[60] F. Al-Doghman, Z. Chaczko, and J. Jiang, "A review of aggregation algorithms for the Internet of Things," in *Systems Engineering (ICSEng), 2017 25th International Conference on*, pp. 480–487, IEEE, 2017.

[61] B. Pourghebleh and N. J. Navimipour, "Data aggregation mechanisms in the Internet of Things: A systematic review of the literature and recommendations for future research," *Journal of Network and Computer Applications*, vol. 97, pp. 23–34, 2017.

[62] F. Alam, R. Mehmood, I. Katib, N. N. Albogami, and A. Albeshri, "Data fusion and IoT for smart ubiquitous environments: A survey," *IEEE Access*, vol. 5, pp. 9533–9554, 2017.

[63] M. Alioto and M. Shahghasemi, "The Internet of Things on its edge: Trends toward its tipping point," *IEEE Consumer Electronics Magazine*, vol. 7, no. 1, pp. 77–87, 2018.

[64] F. Javed, M. K. Afzal, M. Sharif, and B.-S. Kim, "Internet of Things (IoTs) operating systems support, networking technologies, applications, and challenges: A comparative review," *IEEE Communications Surveys & Tutorials*, 2018.

[65] M. Ammar, G. Russello, and B. Crispo, "Internet of Things: A survey on the security of IoT frameworks," *Journal of Information Security and Applications*, vol. 38, pp. 8–27, 2018.

[66] United States Government Accountability Office, "Internet of Things: Status and implications of an increasingly connected world." `https://www.gao.gov/assets/690/684590.pdf`, 2017 (accessed October 9, 2018).

[67] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22–32, 2014.

[68] L. Laursen, "City Saves Money, Attracts Businesses with Smart City Strategy." https://www.technologyreview.com/s/532511/barcelonas-smart-city-ecosystem/, November 2014 (accessed October 1, 2018).

[69] "The Urban Internet of Things." https://datasmart.ash.harvard.edu/news/article/the-urban-internet-of-things-727, August 2015 (accessed October 11, 2018).

[70] A. Akbar, G. Kousiouris, H. Pervaiz, J. Sancho, P. Ta-Shma, F. Carrez, and K. Moessner, "Real-time probabilistic data fusion for large-scale IoT applications," *IEEE Access*, vol. 6, pp. 10015–10027, 2018.

[71] D. Goel, S. Chaudhury, and H. Ghosh, "An IoT approach for context-aware smart traffic management using ontology," in *Proceedings of the International Conference on Web Intelligence*, WI '17, (New York, NY, USA), pp. 42–49, ACM, 2017.

[72] D. Puiu, P. Barnaghi, R. Tönjes, D. Kümper, M. I. Ali, A. Mileo, J. X. Parreira, M. Fischer, S. Kolozali, N. Farajidavar, F. Gao, T. Iggena, T. L. Pham, C. S. Nechifor, D. Puschmann, and J. Fernandes, "CityPulse: Large Scale Data Analytics Framework for Smart Cities," *IEEE Access*, vol. 4, pp. 1086–1108, 2016.

[73] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, "An Information Framework for Creating a Smart City Through Internet of Things," *IEEE Internet of Things Journal*, vol. 1, pp. 112–121, Apr. 2014.

[74] M. M. Rathore, A. Ahmad, A. Paul, and S. Rho, "Urban planning and building smart cities based on the Internet of Things using Big Data analytics," *Computer Networks*, vol. 101, pp. 63–80, June 2016.

[75] A. Marsh, "John Deere and the Birth of Precision Agriculture." https://spectrum.ieee.org/tech-history/silicon-revolution/john-deere-and-the-birth-of-precision-agriculture, February 2018 (accessed October 8, 2018).

[76] R. Burke, A. Mussomeli, S. Laaper, M. Hartigan, and B. Sniderman, "The smart factory." https://www2.deloitte.com/insights/us/en/focus/industry-4-0/smart-factory-connected-manufacturing.html, August 2017 (accessed October 8, 2018).

[77] J. Min-hee, "Smart factories improving productivity of SMEs." http://www.businesskorea.co.kr/news/articleView.html?idxno=14073, March 2016 (accessed October 8, 2018).

[78] "Digital disruption on the farm." https://www.economist.com/business/2014/05/24/digital-disruption-on-the-farm, May 2014 (accessed October 8, 2018).

[79] A. Sadeghi, C. Wachsmann, and M. Waidner, "Security and privacy challenges in industrial Internet of Things," in *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp. 1–6, June 2015.

[80] J. Blackman, "BAE Systems automates manufacturing with Boeing IoT platform." https://enterpriseiotinsights.com/20180411/channels/news/bae-systems-automates-manufacturing-with-boeing-iot-platform-tag40, April 2018 (accessed October 8, 2018).

[81] M. Martinez, "Ford's smart factory approach yields savings." http://www.autonews.com/article/20180730/OEM01/180739933/ford-smart-factory-tim-geiger, July 2018 (ac-cessed October 8, 2018).

[82] "Precision beekeeping with wireless temperature monitoring." https://www.iotone.com/casestudy/precision-beekeeping-with-wireless-temperature-monitoring/c918, 2018 (accessed October 8, 2018).

[83] A. Scroxton, "Interview: How John Deere uses connectivity to make farms more efficient." https://www.computerweekly.com/feature/Interview-How-John-Deere-uses-connectivity-to-make-farms-more-efficient, (accessed October 8, 2018).

[84] NASA, "GPS Correction Technology Lets Tractors Drive Themselves." https://spinoff.nasa.gov/Spinoff2017/ee_1.html, (accessed October 8, 2018).

[85] Monsanto, "Modern Agriculture Practices." https://monsanto.com/company/sustainability/articles/innovation-agriculture-practices/, May 2017 (accessed October 8, 2018).

[86] DuPont, "The Science Behind Precision Agriculture | DuPont USA." http: //www.dupont.com/corporate-functions/our-approach/global-challenges/food/articles/the-science-behindprecisionagriculture.html, 2018 (accessed October 8, 2018).

[87] F. Shrouf, J. Ordieres, and G. Miragliotta, "Smart factories in industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm," in *2014 IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 697–701, 2014.

[88] F. Tao, Y. Zuo, L. D. Xu, and L. Zhang, "IoT-based intelligent perception and access of manufacturing resource toward cloud manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1547–1557, 2014.

[89] Z. Bi, L. D. Xu, and C. Wang, "Internet of Things for enterprise systems of modern manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1537–1546, 2014.

[90] T. Qu, S. P. Lei, Z. Z. Wang, D. X. Nie, X. Chen, and G. Q. Huang, "IoT-based real-time production logistics synchronization system under smart cloud manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 84, no. 1, pp. 147–164, 2016.

[91] F. TongKe, "Smart agriculture based on cloud computing and IoT," *Journal of Convergence Information Technology*, vol. 8, no. 2, pp. 210–216, 2013.

[92] J. Zhao, J. Zhang, Y. Feng, and J. Guo, "The study and application of the IoT technology in agriculture," in *2010 3rd International Conference on Computer Science and Information Technology*, vol. 2, pp. 462–465, 2010.

[93] D. Yan-e, "Design of intelligent agriculture management information system based on IoT," in *2011 Fourth International Conference on Intelligent Computation Technology and Automation*, vol. 1, pp. 1045–1049, 2011.

[94] G. Demiris and B. K. Hensel, "Technologies for an aging society: a systematic review of "smart home" applications," *Yearbook of medical informatics*, vol. 17, no. 01, pp. 33–40, 2008.

[95] N. Bui and M. Zorzi, "Health care applications: A solution based on the Internet of Things," in *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*, ISABEL '11, pp. 131:1–131:5, ACM, 2011.

[96] J. Anuradha and S. Vandhana, "Aspects of ambient assisted living and its applications," in *Internet of Things (IoT): Technologies, Applications, Challenges and Solutions* (B. Tripathy and J. Anuradha, eds.), p. 139, CRC Press, 2017.

[97] H. Petersen, E. Baccelli, M. Wählisch, T. C. Schmidt, and J. Schiller, "The role of the Internet of Things in network resilience," in *Internet of Things. IoT Infrastructures* (R. Giaffreda, D. Cagáňová, Y. Li, R. Riggio, and A. Voisard, eds.), Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, pp. 283–296, Springer International Publishing, 2015.

[98] S. Poslad, S. E. Middleton, F. Chaves, R. Tao, O. Necmioglu, and U. Bügel, "A semantic IoT early warning system for natural environment crisis management," *IEEE Transactions on Emerging Topics in Computing*, vol. 3, no. 2, pp. 246–257, 2015.

[99] T. Gomes, J. Tapolcai, C. Esposito, D. Hutchison, F. Kuipers, J. Rak, A. d. Sousa, A. Iossifides, R. Travanca, J. André, L. Jorge, L. Martins, P. O. Ugalde, A. Pašić, D. Pezaros, S. Jouet, S. Secci, and M. Tornatore, "A survey of strategies for communication networks to protect against large-scale natural disasters," in *2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM)*, pp. 11–22, 2016.

[100] D. Azariadi, V. Tsoutsouras, S. Xydis, and D. Soudris, "ECG signal analysis and arrhythmia detection on IoT wearable medical devices," in *2016 5th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, pp. 1–4, May 2016.

[101] G. Aloi, G. Caliciuri, G. Fortino, R. Gravina, P. Pace, W. Russo, and C. Savaglio, "Enabling IoT interoperability through opportunistic smartphone-based mobile gateways," *Journal of Network and Computer Applications*, vol. 81, pp. 74–84, 2017.

[102] S. Cirani and M. Picone, "Wearable computing for the Internet of Things," *IT Professional*, vol. 17, no. 5, pp. 35–41, 2015.

[103] A. Kos and A. Umek, "Wearable sensor devices for prevention and rehabilitation in healthcare: Swimming exercise with real-time therapist feedback," *IEEE Internet of Things Journal*, pp. 1–1, 2018.

[104] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32–39, 2011.

[105] M. Bachlin, M. Plotnik, D. Roggen, I. Maidan, J. M. Hausdorff, N. Giladi, and G. Troster, "Wearable assistant for Parkinson's disease patients with the freezing of gait symptom," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 436–446, 2010.

[106] A. M. Tabar, A. Keshavarz, and H. Aghajan, "Smart home care network using sensor fusion and distributed vision-based reasoning," in *Proceedings of the 4th ACM International Workshop on Video Surveillance and Sensor Networks*, VSSN '06, (New York, NY, USA), pp. 145–154, ACM, 2006.

[107] G. Yang, L. Xie, M. Mäntysalo, X. Zhou, Z. Pang, L. D. Xu, S. Kao-Walter, Q. Chen, and L. Zheng, "A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2180–2191, 2014.

[108] E. O. Nsoesie, S. R. Mekaru, N. Ramakrishnan, M. V. Marathe, and J. S. Brownstein, "Modeling to predict cases of hantavirus pulmonary syndrome in Chile," *PLoS Neglected Tropical Diseases*, vol. 8, no. 4, p. e2779, 2014.

[109] S. K. Sood and I. Mahajan, "Wearable IoT sensor based healthcare system for identifying and controlling chikungunya virus," *Computers in Industry*, vol. 91, pp. 33–44, 2017.

[110] C. K. Schindhelm and A. MacWilliams, "Overview of indoor positioning technologies for context aware AAL applications," in *Ambient Assisted Living* (R. Wichert and B. Eberhardt, eds.), pp. 273–291, Springer, 2011.

[111] M. Hassanalieragh, A. Page, T. Soyata, G. Sharma, M. Aktas, G. Mateos, B. Kantarci, and S. Andreescu, "Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing: Opportunities and challenges," in *2015 IEEE international conference on services computing (SCC)*, pp. 285–292, IEEE, 2015.

[112] G. Diraco, A. Leone, P. Siciliano, M. Grassi, and P. Malcovati, "A multi-sensor system for fall detection in ambient assisted living contexts," in *SENSORNETS*, pp. 213–219, 2012.

[113] K. Akkaya, I. Guvenc, R. Aygun, N. Pala, and A. Kadri, "IoT-based occupancy monitoring techniques for energy-efficient smart buildings," in *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 58–63, March 2015.

[114] S. K. Ghai, L. V. Thanayankizil, D. P. Seetharam, and D. Chakraborty, "Occupancy detection in commercial buildings using opportunistic context sources," in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pp. 463–466, IEEE, 2012.

[115] N. Nesa and I. Banerjee, "IoT-based sensor data fusion for occupancy sensing using Dempster–Shafer evidence theory for smart buildings," *IEEE Internet of Things Journal*, vol. 4, pp. 1563–1570, Oct 2017.

[116] B. Dong and B. Andrews, "Sensor-based occupancy behavioral pattern recognition for energy and comfort management in intelligent buildings," in *Proceedings of Building Simulation*, pp. 1444–1451, 2009.

[117] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1328–1337, May 2009.

[118] M. Zhao, T. Li, M. A. Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-wall human pose estimation using radio signals," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7356–7365, 2018.

[119] K. P. Lam, M. Höynck, B. Dong, B. Andrews, Y.-S. Chiou, R. Zhang, D. Benitez, J. Choi, *et al.*, "Occupancy detection through an extensive environmental sensor network in an open-plan office building," *IBPSA Building Simulation*, vol. 145, pp. 1452–1459, 2009.

[120] S. Meyn, A. Surana, Y. Lin, S. M. Oggianu, S. Narayanan, and T. A. Frewen, "A sensor-utility-network method for estimation of occupancy in buildings," in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceed-ings of the 48th IEEE Conference on*, pp. 1494–1500, IEEE, 2009.

[121] E. Ackerman and E. Guizzo, "iRobot brings visual mapping and navigation to the roomba 980." `https://spectrum.ieee.org/automaton/robotics/home-robots/irobot-brings-visual-mapping-and-navigation-to-the-roomba-980`, September 2015 (accessed October 6, 2018).

[122] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I," *IEEE Robotics & Automation Magazine*, vol. 13, no. 2, pp. 99–110, 2006.

[123] S. Thrun, "Simultaneous localization and mapping," in *Robotics and Cognitive Approaches to Spatial Mapping* (M. E. Jefferies and W.-K. Yeap, eds.), Springer Tracts in Advanced Robotics, pp. 13–41, Springer Berlin Heidelberg, 2008.

[124] M. R. Endsley, "Autonomous driving systems: A preliminary naturalistic study of the Tesla Model S," *Journal of Cognitive Engineering and Decision Making*, vol. 11, no. 3, pp. 225–238, 2017.

[125] L. Johnson and M. Fitzsimmons, "Uber self-driving cars: everything you need to know." `https://www.techradar.com/news/uber-self-driving-cars`, May 2015 (accessed October 6, 2018).

[126] Waymo, "A new way forward for mobility." `https://waymo.com/`, (accessed October 6, 2018).

[127] DARPA, "DARPA Urban Challenge." `http://archive.darpa.mil/grandchallenge/`, (accessed October 6, 2018).

[128] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. N. Clark, J. Dolan, D. Duggins, T. Galatali, C. Geyer, M. Gittleman, S. Harbaugh, M. Hebert, T. M. Howard, S. Kolski, A. Kelly, M. Likhachev, M. McNaughton, N. Miller, K. Peterson, B. Pilnick, R. Rajkumar, P. Rybski, B. Salesky, Y.-W. Seo, S. Singh, J. Snider, A. Stentz, W. Whittaker, Z. Wolkowicki, J. Ziglar, H. Bae, T. Brown, D. Demitrish, B. Litkouhi, J. Nickolaou, V. Sadekar, W. Zhang, J. Struble, M. Taylor, M. Darms, and D. Ferguson, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics*, vol. 25, no. 8, pp. 425–466, 2008.

[129] H. Shin, Y. Chon, and H. Cha, "Unsupervised construction of an indoor floor plan using a smartphone," *IEEE Trans. Systems, Man, and Cybernetics, Part C*, vol. 42, no. 6, pp. 889–898, 2012.

[130] S. E. Pryse, "Radio tomography: A new experimental technique," *Surveys in Geophysics*, vol. 24, no. 1, pp. 1–38, 2003.

[131] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 5, pp. 621–632, 2010.

[132] J. Wilson and N. Patwari, "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 5, pp. 612–621, 2011.

[133] Bastille Networks Internet Security, "Product introduction." `https://www.bastille.net/product/introduction/`, 2018 (accessed September 28, 2018).

[134] B. R. Hamilton, X. Ma, R. J. Baxley, and S. M. Matechik, "Propagation modeling for radio frequency tomography in wireless networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 1, pp. 55–65, 2014.

[135] B. R. Hamilton, X. Ma, R. J. Baxley, and S. M. Matechik, "Radio frequency tomography in mobile networks," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, pp. 508–511, IEEE, 2012.

[136] B. Beck, R. Baxley, and X. Ma, "Regularization techniques for floor plan estimation in radio tomographic imaging," in *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE*, pp. 177–180, IEEE, 2013.

[137] B. R. Hamilton, X. Ma, and R. J. Baxley, "SAL: Shadowing assisted localization," in *Statistical Signal Processing Workshop (SSP), 2011 IEEE*, pp. 229–232, IEEE, 2011.

[138] B. Beck, R. Baxley, and X. Ma, "Improving radio tomographic images using multipath signals," in *Wireless Information Technology and Systems (ICWITS), 2012 IEEE International Conference on*, pp. 1–4, IEEE, 2012.

[139] B. Beck, X. Ma, and R. Baxley, "Ultrawideband tomographic imaging in uncalibrated networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6474–6486, 2016.

[140] D. Maas, J. Wilson, and N. Patwari, "Toward a rapidly deployable radio tomographic imaging system for tactical operations," in *Local Computer Networks Workshops (LCN Workshops), 2013 IEEE 38th Conference on*, pp. 203–210, IEEE, 2013.

[141] N. Patwari and J. Wilson, "RF sensor networks for device-free localization: Measurements, models, and algorithms," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1961–1973, 2010.

[142] J. Wilson and N. Patwari, "A fade-level skew-Laplace signal strength model for device-free localization with wireless networks.," *IEEE Transactions on Mobile Computing*, vol. 11, no. 6, pp. 947–958, 2012.

[143] J. Li, R. L. Ewing, C. A. Berdanier, D. Wetzel, and X. Shen, "Sparse reconstruction of RF tomography with dynamic dictionary," in *Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS), 2016 IEEE National*, pp. 391–395, IEEE, 2016.

[144] D. Lee, D. Berberidis, and G. B. Giannakis, "Adaptive Bayesian radio tomography," *arXiv preprint arXiv:1804.02084 (https://arxiv.org/abs/1804.02084)*, 2018.

[145] H. Liu, S. Zhan, R. Zhou, W. Sun, J. Cui, Z. Xie, and X. Su, "An FRFT-based approach to eliminate multipath interference in radio tomographic imaging networks," in *Communications and Networking in China (CHINACOM), 2013 8th International ICST Conference on*, pp. 302–307, IEEE, 2013.

[146] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54–69, 2005.

[147] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): part II," *IEEE Robotics Automation Magazine*, vol. 13, no. 3, pp. 108–117, 2014.

[148] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 2, pp. 125–137, 2014.

[149] J. Guo, H. Zhang, Y. Sun, and R. Bie, "Square-root unscented Kalman filtering-based localization and tracking in the Internet of Things," *Personal and Ubiquitous Computing*, vol. 18, no. 4, pp. 987–996, 2014.

[150] L. A. Grieco, A. Rizzo, S. Colucci, S. Sicari, G. Piro, D. Di Paola, and G. Boggia, "IoT-aided robotics applications: Technological implications, target domains and open issues," *Computer Communications*, vol. 54, pp. 32–47, 2014.

[151] N. F. Thusabantu and G. Vadivu, "Adoption of big data streaming techniques for simultaneous localization and mapping (SLAM) in IoT-aided robotics devices," in *Cognitive Informatics and Soft Computing* (P. K. Mallick, V. E. Balas, A. K. Bhoi, and A. F. Zobaa, eds.), Advances in Intelligent Systems and Computing, pp. 315–320, Springer Singapore, 2019.

[152] H. A. Ahmed and J.-W. Jang, "Design of cloud based indoor autonomous navigation with TurtleBot3," *International Conference on Future Information & Communication Engineering*, vol. 10, no. 1, pp. 118–122, 2018.

[153] W. Bulten, A. C. V. Rossum, and W. F. G. Haselager, "Human SLAM, indoor localisation of devices and users," in *2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI)*, pp. 211–222, 2016.

[154] D. Kohler, W. Hess, and H. Rapp, "Introducing cartographer." `https://opensource.googleblog.com/2016/10/introducing-cartographer.html`, October 2016 (accessed October 6, 2018).

[155] W. Sakpere, M. Adeyeye-Oshin, and N. B. Mlitwa, "A state-of-the-art survey of indoor positioning and navigation systems and technologies," *South African Computer Journal*, vol. 29, pp. 145–197, December 2017.

[156] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, vol. 37, pp. 1067–1080, November 2007.

[157] Y. B. Bai, S. Wu, H. R. Wu, and K. Zhang, "Overview of RFID-based indoor positioning technology.," in *GSR*, Citeseer, 2012.

[158] C. Vargas-Rosales, J. Mass-Sanchez, E. Ruiz-Ibarra, D. Torres-Roman, and A. Espinoza-Ruiz, "Performance evaluation of localization algorithms for WSNs," *International Journal of Distributed Sensor Networks*, vol. 11, no. 3, p. 493930, 2015.

[159] L. Chen, S. Thombre, K. Jarvinen, E. S. Lohan, A. Alen-Savikko, H. Leppakoski, M. Z. H. Bhuiyan, S. Bu-Pasha, G. N. Ferrara, S. Honkala, *et al.*, "Robustness, security and privacy in location-based services for future IoT: A survey," *IEEE Access*, vol. 5, pp. 8956–8977, 2017.

[160] R. F. Brena, J. P. García-Vázquez, C. E. Galván-Tejada, D. Muñoz-Rodriguez, C. Vargas-Rosales, and J. Fangmeyer, "Evolution of indoor positioning technologies: A survey," *Journal of Sensors*, vol. 2017, 2017.

[161] R. C. Shit, S. Sharma, D. Puthal, and A. Y. Zomaya, "Location of things (LoT): A review and taxonomy of sensors localization in IoT infrastructure," *IEEE Communications Surveys & Tutorials*, 2018.

[162] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, MobiCom '01, (New York, NY, USA), pp. 166–179, ACM, 2001.

[163] M. O. Ergin, V. Handziski, and A. Wolisz, "Node sequence discovery in wireless sensor networks," in *Distributed Computing in Sensor Systems (DCOSS), 2013 IEEE International Conference on*, pp. 394–401, IEEE, 2013.

[164] Z. Yu and G. Guo, "Improvement of positioning technology based on RSSI in ZigBee networks," *Wireless Personal Communications*, vol. 95, no. 3, pp. 1943–1962, 2017.

[165] F. Evennou and F. Marx, "Advanced integration of WiFi and inertial navigation systems for indoor mobile positioning," *Eurasip journal on applied signal processing*, vol. 2006, pp. 164–164, 2006.

[166] J. MacDougall and G. S. Tewolde, "Tour guide robot using wireless based localization," in *IEEE International Conference on Electro-Information Technology , EIT 2013*, pp. 1–6, May 2013.

[167] P. Cheng, T. Han, X. Zhang, R. Zheng, and Z. Lin, "A single-mobile-anchor based distributed localization scheme for sensor networks," in *2016 35th Chinese Control Conference (CCC)*, pp. 8026–8031, July 2016.

[168] X. Fang, L. Nan, Z. Jiang, and L. Chen, "Multi-channel fingerprint localisation algorithm for wireless sensor network in multipath environment," *IET Communications*, vol. 11, no. 15, pp. 2253–2260, 2017.

[169] C. Lin and K. Song, "Location estimation using probability map of a ZigBee sensor network," in *2013 International Conference on Advanced Robotics and Intelligent Systems*, pp. 11–16, May 2013.

[170] A. Koubâa, M. B. Jamâa, and A. AlHaqbani, "An empirical analysis of the impact of RSS to distance mapping on localization in WSNs," in *Third International Conference on Communications and Networking*, pp. 1–7, March 2012.

[171] S. Aparicio, J. Pérez, A. M. Bernardos, and J. R. Casar, "A fusion method based on Bluetooth and WLAN technologies for indoor location," in *Multisensor Fusion and Integration for Intelligent Systems, 2008. MFI 2008. IEEE International Conference on*, pp. 487–491, IEEE, 2008.

[172] J. Larranaga, L. Muguira, J.-M. Lopez-Garde, and J.-I. Vazquez, "An environment adaptive ZigBee-based indoor positioning algorithm," in *Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on*, pp. 1–8, IEEE, 2010.

[173] A. Moragrega and C. Ibars, "Performance analysis of cooperative and range based localization algorithms for Zigbee and 802.15. 4a wireless sensor networks," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, pp. 1996–2001, IEEE, 2010.

[174] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, pp. 173–184, ACM, 2010.

[175] R. Nandakumar, K. K. Chintalapudi, and V. N. Padmanabhan, "Centaur: locating devices in an office environment," in *Proceedings of the 18th annual international conference on Mobile computing and networking*, pp. 281–292, ACM, 2012.

[176] F. Karlsson, M. Karlsson, B. Bernhardsson, F. Tufvesson, and M. Persson, "Sensor fused indoor positioning using dual band WiFi signal measurements," in *Control Conference (ECC), 2015 European*, pp. 1669–1672, IEEE, 2015.

[177] C. Ou, C. Chao, F. Chang, S. Wang, G. Liu, M. Wu, K. Cho, L. Hwang, and Y. Huan, "A ZigBee position technique for indoor localization based on proximity learning," in *Mechatronics and Automation (ICMA), 2017 IEEE International Conference on*, pp. 875–880, IEEE, 2017.

[178] T. Alhmiedat and A. A. Salem, "A hybrid range-free localization algorithm for ZigBee wireless sensor networks.," *International Arab Journal of Information Technology (IAJIT)*, vol. 14, 2017.

[179] Z. Li, T. Braun, X. Zhao, Z. Zhao, F. Hu, and H. Liang, "A narrow-band indoor positioning system by fusing time and received signal strength via ensemble learning," *IEEE Access*, vol. 6, pp. 9936–9950, 2018.

[180] Q. Dong and W. Dargie, "Evaluation of the reliability of RSSI for indoor localization," in *Wireless communications in unusual and confined areas (ICWCUCA), 2012 international conference on*, pp. 1–6, IEEE, 2012.

[181] D. Konings, N. Faulkner, F. Alam, F. Noble, and E. M. Lai, "The effects of interference on the RSSI values of a ZigBee based indoor localization system," in *24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, pp. 1–5, Nov 2017.

[182] D. Lymberopoulos and J. Liu, "The Microsoft Indoor Localization Competition: Experiences and lessons learned," *IEEE Signal Processing Magazine*, vol. 34, no. 5, pp. 125–140, 2017.

[183] Microsoft, "Microsoft indoor localization competition – ipsn 2018." `https://www.microsoft.com/en-us/research/event/microsoft-indoor-localization-competition-ipsn-2018/`, 2018 (accessed October 2, 2018).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
23-01-2019

2. REPORT TYPE
Technical Report

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Survey of Data Fusion in IoT

5a. CONTRACT NUMBER
FA8721-05-C-0002 & FA8702-15-D-0001

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
K.J. Kratkiewicz, J. White Bear, and B.A. Miller

5d. PROJECT NUMBER
3118

5e. TASK NUMBER
271

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02421-6426

8. PERFORMING ORGANIZATION REPORT NUMBER
TR-1239

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Geospatial-Intelligence Agency
7500 GEOINT Drive
Springfield, Virginia 22150
571-557-5400

10. SPONSOR/MONITOR'S ACRONYM(S)
NGA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
With the advent of the Internet—and, in particular, the World Wide Web—came a massive increase in the amount of data we as a society create on a daily basis. The more recent proliferation of mobile devices and online social networking has increased the amount of user-generated content to an enormous level. The current spread of Internet of Things (IoT) devices is causing an even greater acceleration, with the number of IoT devices online already outnumbering the number of people in the world. Within all of this data is information of potential value to many missions of the U.S. government. While IoT devices are plentiful, they tend to be inexpensive and are sometimes unreliable in their measurements. A prudent consumer of IoT data would be skeptical of any information gleaned from a single IoT sensor. This, however, is where the sheer volume of IoT devices in the world can provide a substantial benet. The low cost of IoT sensors may yield relatively low performance, but also makes it feasible to collect large amounts of data over a sea of devices. While, for example, a single traffic sensor may frequently malfunction, fusing data across a large number of sensors improves robustness to unreliable measurements from any given device.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified	Same as report	75	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18