

A Survey of Data Fusion in Smart City Applications

Billy Pik Lik Lau, Sumudu Hasala Marakkalage, Yuren Zhou, Naveed Ul Hassan, Chau Yuen, Meng Zhang,
U-Xuan Tan

Abstract—The advancement of various research sectors such as Internet of Things (IoT), Machine Learning, Data Mining, Big Data, and Communication Technology has shed some light in transforming an urban city integrating the aforementioned techniques to a commonly known term - Smart City. With the emergence of smart city, plethora of data sources have been made available for wide variety of applications. The common technique for handling multiple data sources is data fusion, where it improves data output quality or extracts knowledge from the raw data. In order to cater evergrowing highly complicated applications, studies in smart city have to utilize data from various sources and evaluate their performance based on multiple aspects. To this end, we introduce a multi-perspectives classification of the data fusion to evaluate the smart city applications. Moreover, we applied the proposed multi-perspectives classification to evaluate selected applications in each domain of the smart city. We conclude the paper by discussing potential future direction and challenges of data fusion integration.

Index Terms—Data Fusion; Sensor Fusion; Urban Computing; Smart City; Big Data; Internet of Things; Multi-Perspectives Classification

I. INTRODUCTION

According to UN estimates [1], 68% of the world population would be living in cities by 2050. Hence, managing the existing resources and infrastructure to cater sustainable urban living conditions for the growing needs of the urban population has become ever more challenging. Fortunately, the advancement in Information and Communication Technologies (ICT), Internet of Things (IoT), Big Data, Data Mining, and Data Fusion is gradually paving path for the emergence of smart cities [2]–[4]. In this paper, we adopt the following definition of smart city [5]:

“A city combining ICT and Web 2.0 technology with other organizational, design and planning efforts to de-materialize and speed up bureaucratic processes and help to identify new, innovative solutions to city management complexity, in order to improve sustainability and livability”

The integration of aforementioned technologies into various urban domains enables city managers to equip with the necessary information for better planning and resource management. Several cities around the world have already been leveraging these technologies to improve the comfort, security, mobility, health, and well-being of their citizens. To better evaluate rapid progress and to recognize the efforts of urban planners, smart

city ranking systems have been established. For instance, IESE cities in motion index [6] has suggested 83 indicators to rank 165 cities over 80 countries. New York, London, and Paris are the top three smart cities in 2018. Smart city projects in New York [7] aim to consistently improve the quality of residents’ life, reduce the environmental impacts, increase the street light efficiency, and enhance the water quality. Meanwhile, the focus of Smart London Projects [8] is to collect city wide data to provide world class connectivity, security, and smarter streets to its residents. Digital transformation, sustainability, and urbanization for improving citizen services are at the cores of Paris Smart City Projects [9]. The following up of the top smart city list includes Singapore and Tokyo, which are some other notable smart cities in the world. In Singapore, Smart Nation Project [10] has been proposed, which includes e-payment systems, smart nation sensor platform, smart urban mobility, and smart community initiatives, with the aim to enhance the national digital identity of its citizens. On the other hand, Tokyo [11] aims to become the greenest city in Asia Pacific by improving the transportation and other sectors of their economy. Local governments in several Chinese cities [12], such as, Shenzhen, Shanghai, Hangzhou, and Beijing are also shaping up their cities to facilitate economic and social development to build high income smart cities. In addition, there are several research institutes and laboratories focusing on developing smart city applications, which are currently leading the worldwide effort in smart domains. These include MIT Senseable Lab [13], Future Cities Laboratory [14], SINTEF Smart Cities [15], SMART [16], etc.

Nowadays, communication technology is the backbone for the smart city applications as it provides a channel for applications to transfer data effortlessly. The ongoing quest for novel, more efficient, low-latency, and cost-effective communication technologies and networks, such as, 5G [27]–[29], wireless sensor networks (WSN) [30]–[32], Low Power Wide Area Network (LPWAN) [33], [34], and Narrow Band IoT (NB-IoT) [35], [36] and their integration in smart city projects is also relentless.

These advancement has made many data sources available due to the potential of sensors collecting data with better coverage and power efficiency of the communication platform. With the large amounts of data becoming readily available in a smart city, data mining techniques [37], [38] are commonly used in the collected data. It helps in identifying the essential and important data sources in the smart city applications such as monitoring, control, resource management, anomaly detection, etc. With the availability of parallel data sources in various smart city domains, data fusion techniques that combine multiple data sources, lie at the heart of smart city platform integration. The major objectives of data fusion are to

B. P. L. Lau, S. H. Marakkalage, Y. Zhou, N. Ul Hassan, C. Yuen, U-X. Tan are with the Engineering Product Development, Singapore University of Technology and Design, 8 Somapah Rd, Singapore 487372.

N. Ul Hassan is also with Lahore University of Management Sciences (LUMS), Lahore, Pakistan 54792

Meng Zhang is with Southeast University, Nanjing, China 210096

TABLE I
LITERATURE REVIEW FOR DATA FUSION ON SMART CITY

Surveys	Objectives and Topics Covered
Khaledgi et al. [17]	Provides insights on the different types of data fusion techniques by exploring their concept, benefits, and challenges.
Castanedo [18]	Provides an overall view on the different data fusion techniques and methods. The author also reviewed common algorithms such as data association, state estimation, and decision fusion.
Alam. et al. [19]	Provides a comprehensive survey on the mathematical model used in data fusion for specific IoT environments.
Wang et al. [20]	Proposes an IoT architecture concept to survey on the different sensor data fusion techniques and also provides an overall view on their evaluation framework.
Zheng [21]	Discusses about differences on fusing sources and varying techniques for cross domain data fusion.
El et al. [22]	Provides a survey on the intelligent transportation systems, which use data fusion techniques.
Esmailian et al. [23]	Provides a throughout study on waste management for smart city aspects with three categories: (1) infrastructure for the collection of product lifecycle data, (2) new adapting business model, and (3) waste upstream separation techniques.
Da Xu et al. [24]	Provides an overall view on the current state of the industries for IoT and discusses key enabling technologies such as communication platforms, sensing technologies, and services.
Chen et al. [25]	Reviews the building occupancy estimation and detection techniques while providing a comparison between different sensor types for cost, detection and estimation accuracy, and privacy issues.
Qin and Gu [26]	Introduces the data fusion algorithms in IoT domains and data acquisition characteristics.

address problematic data while enhancing the data reliability and extracting knowledge from multiple data sources. The existing survey papers related to smart city applications or data fusion classification are summarized in Table I. Majority of these review papers [18], [39]–[41] strictly focus on one particular smart city domain or one genre of classification perspective. In [19], Alam et al. have conducted a review on data fusion technique based on mathematical model in IoT environment. Alternately in [20], Wang et al. have described the frameworks of data fusion within the smart city application. Interested readers can follow these references for additional technical details. However, there is only a handful of limited work to provide a multi-perspectives approach for data fusion problems in smart cities and this literature gap further motivates our study.

Therefore, a different perspective to look at data fusion in smart city domains is necessitated by the expanding scale and scope of data sources, data collection techniques, and data processing system architectures. In order to cater evergrowing highly complicated applications, studies in smart city have to utilize data from various sources and evaluate their performance based on multiple aspects. To this end, we propose multiple generic perspectives with the ability to cover the

entire depth and breadth of data fusion problems in smart city. These perspectives include data fusion objectives, data fusion techniques, data input and data output types, data sources, data fusion scales, and platform architectures for data processing. Utilizing proposed perspectives, we provide an overall view of classification techniques found in the seven domains of smart city applications such as: Smart Living, Smart Urban Area Management, Smart Environment, Smart Industry, Smart Economics, Smart Human Mobility, and Smart Infrastructure. A simple illustration of seven application domains discussed in this paper can be found in the Figure 1. In each domain, we only select notable papers to demonstrate the universality and effectiveness of our multi-perspective approach on evaluating the data fusion techniques. Please note that we do not provide a comprehensive review of all the smart city applications. Afterwards, we talk about emerging data fusion trends in smart cities, while outlining the best practices for deploying a smart city application. In addition, data fusion challenges in different smart city applications are also identified and discussed.

To summarize, our novel contributions in this paper are three-fold as shown below:

- We propose a multi-perspectives classification to evaluate common data fusion techniques in smart city applications.
- We provide an overview of smart city application domains and discuss the common trend of data fusion techniques in each domain utilizing proposed multi-perspectives classification.
- We list down the future challenges and the ideal scenario for deploying data fusion techniques in a smart city application.

Overall, we believe that with these contributions, the readers would have a quick grasp on the current data fusion trends in smart city research without extensively going through all the details.

The rest of the paper is organized as follows: in Section II, we define the data fusion classification using multi-perspectives to evaluate a smart city application. This lays a foundation for evaluating the smart city applications leveraging data fusion techniques. In Section III, different application domains of smart city based on data fusion techniques are evaluated using the proposed multi-perspectives classification of data fusion. In addition, a brief overall view of the current research trend of respective domain is presented. Subsequently in Section IV, we discuss the ideal data fusion scenario along with potential research directions/opportunities based on speculations of smart city applications from previous section. Lastly, we conclude our works in Section V.

II. DATA FUSION CLASSIFICATION USING MULTI-PERSPECTIVES

In this section, we identify multiple generic perspectives with the ability to cover the entire depth and breadth of data fusion literature in smart city applications. We use smart city single perspective data fusion review papers [19], [26] and non-smart city data fusion classification papers [18], [39]–[41] as references. In non-smart city literature, there are four well-known data fusion classification techniques, which are

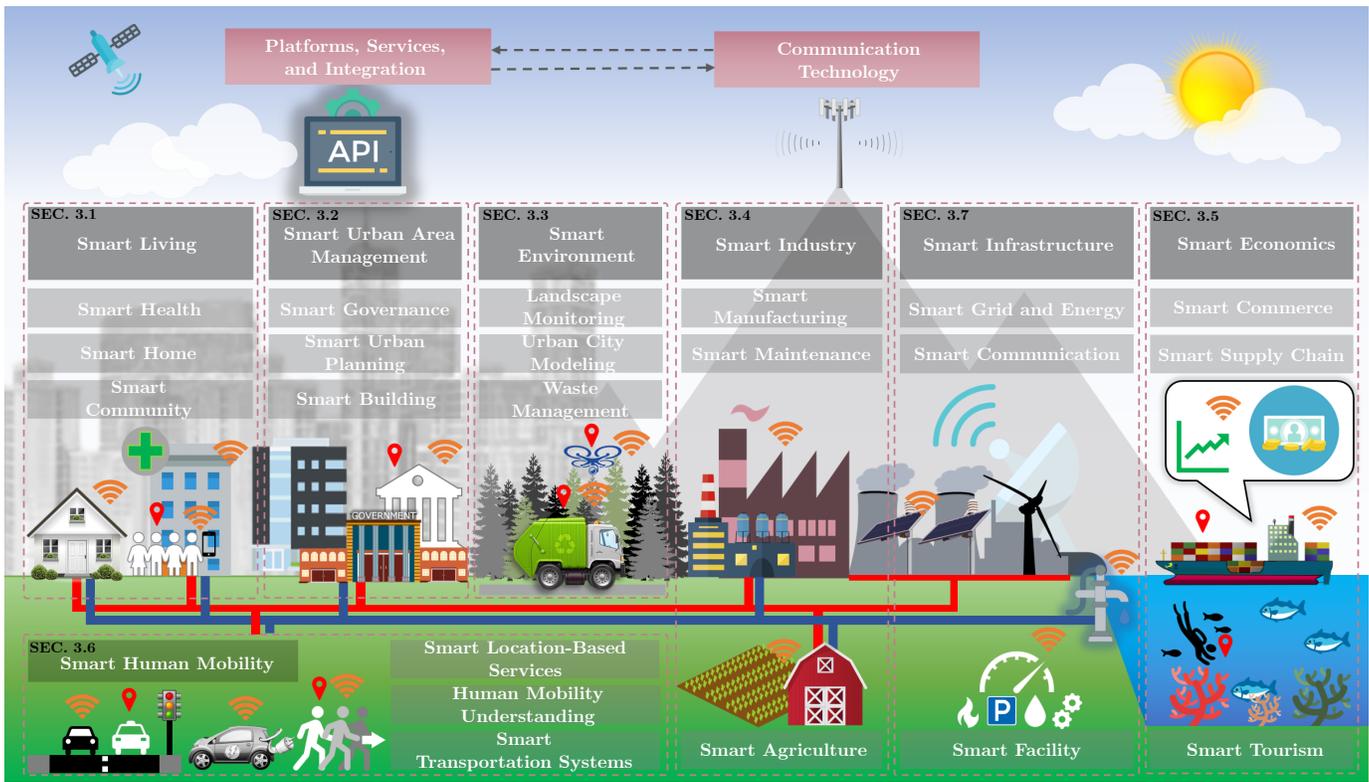


Fig. 1. List of smart city applications domain, where data fusion is commonly applied (each domain is enclosed in the dotted pink box).

Dasarathy’s Classification [39], Whyte’s Classification [40], Fusion Architecture’s Classification [18], and US Joint Directories of Laboratories (JDL) data fusion classification [41]. Dasarathy’s Classification is based on the data input and output types between data, where Whyte’s Classification focuses on the relationship between the data. JDL focuses on classifying the fusion process according to five processing levels. Meanwhile, the architecture-based classification only captures the system design level and does not consider data relationships and types. Most of the aforementioned classification of the data fusion techniques are not suitable for evaluating the applications of a smart city.

Our proposed data fusion classification approach for smart city comprises of six different perspectives (also called categories): i) data fusion objectives (*O*), ii) data fusion techniques (*T*), iii) data input and output types (*D*), iv) data source types (*S*), v) system scales (*L*), and vi) platform architectures (*P*). Within each category, we further identify various sub-categories (also called classes). Overall, there are 30 different classes. The complete list of the adopted classification indicating all the categories and their classes is shown in Table II. Short reference codes (*O, T, D, S, L, P*) for each class are also included in the table for further use in the paper. For example, *O1* refers to the data fusion objective category and problematic data fusion class. Similarly, *S3* refers to data source types category and participatory class.

Note that, there could be potentially more than one perspectives (other than data sources, fusion scales, and platform architecture) for smart city application depending on the complexity and fusion objective itself. Below, we provide

TABLE II
DATA FUSION CLASSIFICATIONS FOR SMART CITY APPLICATIONS USING MULTI-PERSPECTIVES

Perspective/Category	Code	Classes
Data Fusion Objectives	O1	Fixing Problematic Data
	O2	Improving Data Reliability
	O3	Extracting Higher Level Information
	O4	Increasing Data Completeness
Data Fusion Techniques	T1	Data Association
	T2	State Estimation
	T3	Decision Fusion
	T4	Classification
	T5	Prediction / Regression
	T6	Unsupervised Machine Learning
	T7	Dimension Reduction
	T8	Statistical Inference and Analytics
	T9	Visualization
Data Input and Output Types	D1	Data in Data Out (DAI-DAO)
	D2	Data In Feature Out (DAI-FEO)
	D3	Feature in Feature Out (FEI-FEO)
	D4	Feature in Decision Out (FEI-DEO)
	D5	Decision in Decision Out (DEI-DEO)
Data Source Types	S1	Physical Data Sources
	S2	Cyber Data Sources
	S3	Participatory Data Sources
	S4	Hybrid Data Sources
Data Fusion Scales	L1	Sensor Level Fusion
	L2	Building Wide Fusion
	L3	Inter-Building Fusion
	L4	City Wide Fusion
	L5	Inter-City Fusion (or Larger)
Platform Architectures	P1	Edge Computation
	P2	Fog / Mist Computation
	P3	Cloud Computation
	P4	Hybrid Computation

further details of all the perspectives and classes adopted in this paper.

A. Data Fusion Objectives (O)

The data fusion techniques deployed in a smart city project is influenced by the objective of applications. In this paper, we have summarized the four objectives as follows:

- **O1: Fixing Problematic Data**
‘Problematic Data’ class refers to the case when the data source is having quality issues such as, inconsistency, imperfection, disparateness, etc. Data fusion could be used as an easy approach to overcome such problems. Examples of O1 can be found in [27], [42]–[44].
- **O2: Improving Data Reliability**
Data may suffer from reliability issues when it is collected in a less ideal (less controlled) environment with high presence of noise. In such situation, additional data sources are required to add redundancy for increasing data quality to enhance data reliability. Such situations are identified as ‘Data Reliability’ class and [45]–[48] exhibits such pattern. In addition, security enhancement through the data fusion also belongs to this category and examples of such objectives can be found in [49]–[51].
- **O3: Extracting Higher Level Information**
Data mining advancement has contributed to many different architectures of data fusion in order to obtain knowledge from multiple data sources. For instance, the occupancy of a building can be detected using a combination of few ambient sensors with data fusion, where occupancy information cannot be directly inferred from the raw data sources. We classify these approaches as ‘Higher Level Information Extraction’ class and examples can be found in [52]–[54].
- **O4: Increasing Data Completeness**
In a situation of coverage limitations, an individual data source is insufficient to provide complete details of the desired output. Therefore, in ‘Data Completeness’ class, data fusion is performed across multiple data sources to obtain a complete picture of the overall system such as [55]–[57].

B. Data Fusion Techniques (T)

In this category, we present the data fusion techniques in two different information enrichment obtained after data fusion. The $T1$ until $T3$ are the common data fusion techniques and the further details can be found in [19], [39], where it describes the lower level information being fused to generate identical level of information. The techniques $T4$ – $T8$ are associated with data mining [38], [58], where simple input data from multiple sources is fused to generate higher level information enrichment. Brief description of these classes is given below:

- **T1: Data Association**
Data association refers to data fusion technique that fuse data based on similarity between at least two or more data sources. Common techniques for data association include Nearest Neighbors [59], Probabilistic Data Association [60], and Multiple Hypothesis Test [61].

- **T2: State Estimation**

State estimation indicates the usage of multiple data sources to achieve higher state estimation accuracy. Common techniques under this category are Maximum Likelihood [62], Kalman Filter [63], Particle Filter [64], and Covariance Consistency Model [65].

- **T3: Decision Fusion**

Decision fusion is a technique that is used to fuse the decisions made by various sub-components of a system to achieve a certain overall objective. For instance, a robot can fuse different decisions from the modules to perform an actuation (direction, events, or actions). General techniques include Bayesian inference [66], Dempster-Shafer Inference [67], and semantic approaches [68].

- **T4: Classification**

Classification technique denotes methodology of grouping objects into different classes based on their unique characteristics. In-depth details of generic classification techniques can be found in [38], [58].

- **T5: Prediction**

Prediction techniques are used to forecast output based on single or multiple different data sources. Note that, this covers simple methods such as regression and as well as complicated methods such as forecast modeling. Examples of such can be found in [69]–[71].

- **T6: Unsupervised Machine Learning**

Unsupervised machine learning tries to automate the knowledge discovery without relying on the data labels. Examples of such methods involves clustering [72], anomaly detection [73] and others [38]. Note that, semi-supervised machine learning approach [74] is also categorized under this class.

- **T7: Dimension Reduction**

Dimension reduction refers to the method of reducing data sources’ dimensions for features extraction or visualization purposes. Examples of dimension reduction techniques are Principal Component Analysis (PCA) [75], and others [38]. The aim is to preserve the characteristic of the data sources while reducing the complexity of processing high dimensional data.

- **T8: Statistical Inference and Analysis**

Statistical inference and analysis is used for outlining certain information along with some common knowledge / hypothesis from the input data sources. Examples of papers using such approaches can be found in [76], [77].

- **T9: Visualization**

Visualization is a technique used for the presentation of output to the end users via some platform. The end result often requires human intervention. Examples of such techniques can be referred to the following papers [78]–[80].

C. Data Input and Output Types (D)

Dasarathy’s classification [39] is based on input and output of fusion technique to determine the relation between input and output data. There are five classes in data input and output perspective. Brief details are given below:

- **D1: Data In Data Out (DAI-DAO)**

Data In Data Out (DAI-DAO) refers to the situation when multiple raw data sources are fused to increase data reliability and the output after fusion is still a raw data.

- **D2: Data In Feature Out (DAI-FEO)**

Data In Feature Out (DAI-FEO) refers to the situation when multiple raw data sources are fused to extract some unique feature of the observed system. The output feature describes certain aspect of the system and it could be further used for more feature extraction or to make certain decisions.

- **D3: Feature In Feature Out (FEI-FEO)**

Feature In Feature Out (FEI-FEO) refers to the situation when multiple unique features from different sensors are combined to generate new features. This class is commonly known as feature fusion.

- **D4: Feature In Decision Out (FEI-DEO)**

Feature In Decision Out (FEI-DEO) refers to the situation when certain features of the system are fused to make certain decisions, e.g. actuation of various system components.

- **D5: Decision In Decision Out (DEI-DEO)**

Decision In Decision Out (DEI-DEO) refers to the situation when different decision sources (maintenance status, events, etc.) are combined to obtain a final output decision.

D. Data Source Types (S)

There are four types of generic data sources in smart city applications and we categorize them based on the data sources regardless of the communication medium. Details of each category can be found as follows:

- **S1: Physical Data Sources**

The physical data sources are collected from sensors that are being deployed to capture information of a particular space, area, or even city wide. Examples of the physical sensors include temperature [81], air quality [82], camera [83], ultrasonic [84], LiDAR [85], and etc. Note that, we categorize smart city application based on the data sources rather than the method they are acquired. For instance, a temperature probe in a sensor nodes of a wireless sensor network (WSN) transmits data through gateway to cloud database is considered as physical data source, S1.

- **S2: Cyber Data Sources**

Cyber data sources denote datasets which are commonly obtained from the Internet domain such as social media information [76], [86], web access data [87], [88], and opinion based datasets [89]. Social media information involves major social media platforms such as Twitter, Facebook, LinkedIn, Weibo, and others. Note that, usually the data is acquired through data mining techniques. Meanwhile, the web access data can be obtained from web applications programming interface (API), such as transportation tickets information and online customer records. Apart from that, open datasets refer to data from third party vendors such as telecom operator or a company with readily available data.

- **S3: Participatory Data Sources**

Participatory data sources include crowdsensing [90], [91] and crowdsourcing [92], [93] data contributed by the personal devices, e.g. mobile phones, wearable devices, tablets, etc. of the users in smart city. Users provide the data voluntarily or through some incentive mechanisms.

- **S4: Hybrid Data Sources**

The hybrid data sources include data obtained from mixed data sources [94], [95], e.g. by combining the participatory and physical sensor data. As pointed in [21], hybrid data sources can achieve more insights as compared to single data sources.

E. Data Fusion Scales (L)

The scale of data fusion is also an important classification perspective. Please note that data fusion scale is based on sensor coverage rather than sensor deployment. There are four different classes, which are described below:

- **L1: Sensor Level Fusion**

At the sensor scale, data from various physical sensors is fused to form an output such as [53], [96]. For instance, fusion of data collected by various smartphone sensors is an example of data fusion at sensor level.

- **L2: Building Wide Fusion**

At the building wide scale, data sources collected within a premise or building is fused to form an output. For instance, fusion of building energy and building security data to develop a building management system [97]–[99] is an example of data fusion at building level.

- **L3: Inter-Building Fusion**

In the inter-building scale, the data sources collected over several buildings are fused to form an output, where the scale of deployment normally includes small area. For example, data sources of several buildings within a university are used to generate a particular output is considered as inter-building scale. Other examples of this data fusion scale also can be found in [100], [101].

- **L4: City Wide Fusion**

In the case of city wide fusion, data sources that involve whole city's area as input for the data fusion architecture fall under this class such as [102]–[104]. For instance, the study of citizen behavior involves fusion of data gathered in different areas of the city is considered city wide data.

- **L5: Inter-City Fusion (or larger)**

At the inter-city fusion (or larger) scale, data from large areas involving one or more cities or terrains (mountains, sea, forests, etc.) is fused to form an output. Examples of this scale involve comparing one smart city to another city or data of a city outskirts and its surrounding areas. More examples of inter-city fusion (or larger) can be referred to [43], [105], [106].

F. Platform Architectures (P)

The architecture of computational platform involved in data fusion is another important classification perspective. In this category, we identify four generic classes:

- **P1: Edge Computation Platform**

In edge computation platform, data sources are processed and fused at the edge (i.e. very close to the physical location, where data is actually collected). Edge computation devices include micro-controller, computing devices (Raspberry pi), computers, etc. Such architecture can be found in works such as [96], [99], [101]. With this architecture, communication overheads and latency can be significantly reduced.

- **P2: Fog Computation Platform**

In fog computation platform, data sources are processed and fused at the middle layer, i.e. between the edge and the cloud. In this architecture, data is periodically or continuously sampled at the edge (without processing) and is then forwarded to a gateway (that acts as a fog device). At the gateway, computing resources are provided for data processing. Both fog computing and edge computing platforms provide similar benefits of offloading computation as shown in [102], [107], [108]. However, fog computing architecture should be preferred when it is difficult to find stable power sources at the edge.

- **P3: Cloud Computation Platform**

In cloud computation platform, data sources are processed and fused in the cloud. This is the most common technique practiced by industry and research institutes for processing big data. Examples of this architecture being used are [56], [87], [109]. The advantages of cloud computing architecture includes ready access to the data and both online and offline for further processing or fusing. The disadvantages include increased communication overheads and costs.

- **P4: Hybrid Computation Platform**

In hybrid computation platform, processing is distributed among two or more layers (edge, fog and cloud) as shown in [105], [110], [111]. In this architecture, depending on the available resources or application objectives, some low level data fusion and processing is done at the edge or fog, while high level information is extracted in the cloud.

III. SMART CITY APPLICATIONS OVERVIEW

Smart city applications tend to have extremely diverse requirements, which contribute to a large variety of different techniques and requirements as stated previously in Section 2 for different domains. Thus, it is necessary to evaluate the smart city applications from a more generic perspectives rather than one specific perspective. In this section, we select smart city applications with data fusion techniques from different domains listed in Figure 1, and evaluate them based on multi-perspectives from the Section II. Note that, there exist some literatures that are cross-disciplinary, which may involve more than one domain. In order to address the cross-disciplinary smart city applications, we have grouped them into their closest relevant domain. In each application domain, we outline sub-domains and present works related to data fusion techniques. Using the proposed data fusion classification based

on multi-perspectives, we discuss the common data sources and fusion techniques, along with the current research trends in each domain.

A. Smart Living

Smart living concerns with the life of the urban citizens and revolves around the concept of improving live-ability in urban area. In the literature, the general objectives of utilizing the smart living domain involve data being used to extract higher level information or increasing the data completeness. In addition, smart city applications in this domain often leverage the cloud or hybrid platform architecture. In this domains, we have studied three different aspects of smart living, namely, (1) Smart Health, (2) Smart Home, and (3) Smart Community.

1) *Smart Health*: Healthcare is a crucial component in everyday life concerning medical and public practices using devices as defined by Lee and Co-authors [144], [145]. The rapid development of technology (e.g. smartphones and their in-built sensing devices such as heart rate sensors) provides more opportunities to adopt technology in healthcare applications pervasively. For telehealth application in smart city, Hossain et al. [112] have used electroencephalographic (EGG) signals and voice to monitor a specific user's health with the support of cloud technology and doctor's advices. In [113], work has shown to monitor elderly at home based on fuzzy fusion model using behavioral and acoustical environment data. Similarly, Noury [146] also monitors the activities and fall detection of elderly through fuzzy logic by fusing accelerometer, vibration, and orientation sensor. In [91], Marakkalage et al have used crowd-sensing data from a smartphone application (location, noise, light, etc.) and introduced sensor fusion based environment classification (SFEC) to profile elderly people for understanding their daily lifestyle. In addition, Dawar and Kehtarnavaz in [52] have implemented a Convolution Neural Network (CNN) to combine both depth camera and wearable devices to detect the transition of movements to fall. Apart from that, Hondori et al. [111] have proposed using sensor fusion between depth images and inertia to perform tele-rehab in the home. The main challenge occurs in pervasive smart healthcare data fusion is discussed in [147] as the need of a higher accuracy to improve sensing robustness against uncertainty and unreliable integration.

2) *Smart Home*: The concept of Smart Homes plays an important role nowadays in contemporary urban areas. According to Jiang et al. [148], the definition of a smart home provides the capability of controlling, monitoring, and accessed appliances & services through implementation of ICT. There are currently many big players in developing the smart home appliances such as Amazon, Google, Apple, IBM, Intel, Microsoft, Xiaomi, and others. The challenge faced by manufacturers are related with service integration and formulating software ontology platform. These are necessary for implementing the services through different vendors and allow for a better integration. Meanwhile in [114], physical sensors (soil moisture) and cyber (weather, traffic) have been fused to control home appliances such as alarm clock and water sprinkle. The study of user daily activity is yet

TABLE III
LIST OF SMART CITY APPLICATIONS USING DATA FUSION TECHNIQUE(S)

Domain	Sources	O	S	D	T	L	P	Remarks
Smart Living	[52]	3	1	2	4	1	4	Smart Healthcare
	[112]	3	1	1	4	4	4	Voice Pathology Detection
	[113]	3	3	4	3	2	4	Smart Home Healthcare Monitoring
	[110]	3	1	2	4	1	4	Daily Activity Classification
	[45]	2	1	2	4	2	4	Smart Home Activity Recognition
	[111]	3	1	3	4	2	4	Tele-Rehabilitation
	[114]	4	4	4	3	2	4	Smart Home Control System
	[79]	3	4	3	9	4	4	Intelligent Video Surveillance
	[115]	3	3,4	4,5	4,5	5	3	Distance Learning
[116]	3	2	3	1	4	3	Smart Community	
Smart Urban Area Management	[94]	4	4	4	5	2	3	Building Management
	[117]	4	1	2	2	1	1	Fire Detection System
	[118]	3	3	4	3	5	3	Lean Government
	[43]	1	1	1	1	5	1	Urban Planning with Satellite Images
	[95], [119]	3	4	1,2	8,9	4	3	Urban Space Utilization Detection
	[56]	4	3	1	9	4	3	Fault Reporting Platform
	[76]	3	4	3	8,9	5	3	Landscape Rating Systems
Smart Environment	[106]	3	1	1	9	5	3	City Environment Monitoring
	[78]	3	1	2	2,9	1	1	City Building Map Modeling
	[120]	3	4	4	4	5	1	Forest Types Classification
	[46]	2	1	1	1	5	1	Long Term Landscape Monitoring
	[121]	4	1	4	4	5	1	Forest Species Classification
	[122]	3	4	2	4	1	1	Waste Water Treatment
	[102]	4	1	2,3	2,9	4	2	Urban Solid Waste Management
Smart Industry	[96], [123]	4	1	2,4	2,4	1	1	Fault Detection
	[124], [125]	3,4	1	3,4	5	1	1	Tools Life Prediction
	[98]	2	4	4,5	3	2	1	Decision Support in Manufacturing
	[51]	2	1	2,4	2,3	1	1	Autonomous Robots and Security
	[126]	3	1	2,4	4,7	1	1	Seafood Freshness Classification
	[127], [128]	3,4	1	2,4	2,4,5	1	1	Agriculture Plant Disease Classification
Smart Economics	[87]	4	2,3	1,3	1,8	5	3	Customer Profiling
	[129]	4	4	1,4	5	5	3	Consumer Awareness
	[107]	4	4	1	9	5	2	Blockchain and Supply Chain
	[130]	3,4	4	2,4	1,5,8	5	3	Supply Chain Management
	[77]	3	2,3	2,3	5,8	4	3	Tourist Behavior Analysis
	[57]	4	4	2,3	6	4	3	Travel Recommendation System
[131]	3	1,2,3	2	1,4	4	1,3	Tourist Tracking Application	
Smart Human Mobility	[132], [133]	2,3	1	1	5,1	4	3	Outdoor Positioning
	[134], [135]	2,4	1	1	1,2	2	3	Indoor Positioning
	[136], [137]	4	1	4	5,1	4,2	3	Location-based Services
	[103]	3	3	2	1	4	3	Obtaining Origin-Destination Matrices
	[54]	3	3	2	4	4	3	Identifying Transportation Modes
	[138]	3	3	2	1	2	3	Monitoring Visitors Inside a Building
	[109]	4	1	4	3	4	3	Traffic Signal Controlling
	[139]	3	3	2	1	4	3	Analyzing Public Transport Services
	[140]	4	1	4	4	1	3	Autonomous Vehicle Controlling
Smart Infrastructure	[55], [88]	3,4	1	2,4	5	4,1	1	Smart Grid and Power Utilities
	[101], [141]	3	1,4	1	4,5	3	1	Solar Farm
	[105]	3	3	2	2	5	4	Smart Metering
	[27], [42]	1,2	1	1	1,2	1	1,3	Communication (5G)
	[47], [48]	2	1	1	1,5	1	1	Communication (WSN)
	[142]	4	1	2,3	4	1	1	Drone Detection
	[143]	3	4	1,2	2	4	3	Smart Parking System
	[99]	4	1	1	2	2	1	Bridge Monitoring Platform
	[104]	3	1	2,3	4,5	4	3	Water Distribution System

another important aspect to understand urban citizen well-being. In [45], Hong et al. have combined series of life activities to understand the lifestyle pattern depends on the equally weighted sum operation and Dempster-Shafer theory. Also, similar study on the user daily activity patterns can be found in [110]. Combination of house environmental sensor (infrared, door contact, temperature, hygrometry sensor, microphone) and wearable devices (kinematic sensors) using support vector machine (SVM) can be used to identify the user activity patterns. In addition, the modeling of human behavior in a smart home [149] in order to generate learning situation models have proven the efficiency of context-aware services. In addition, smart home security is yet another study field for many researchers [150]–[152] due to increased usage of IoT devices in normal household. The research challenges is to develop the applications for the smart houses while retaining the privacy and security of the end user.

3) *Smart Community*: According to Smart Communities Guidebook [153], a smart community is described as “a geographical area ranging in size from neighborhood to a multi-county region whose residents, organizations, and governing institutions are using information technology to transform their

region in significant ways”. There is only a handful of cities focus on this aspect as majority are still in the stage of transforming from facility to community welfare. First world countries such as USA, Canada, Australia, European Union, and Singapore shown in [154] have started up initiatives to create smart communities. Information fusion for smart community video surveillance system is performed in [79] to aid neighborhood in terms of security. The combination of the different modal surveillance camera provides a vast amount of visual information extraction such as video summarization for highlighting certain events. A distance learning framework is proposed in [115], which enables personalized learning to cater what is best for each individual user. It uses data fusion to understand user environment and their activities by means of hybrid data sources. Real-time community monitoring also helps to prevent emergency situations and it ensures the safety of community citizens. A good example for a smart community application in large-scale is the Social Credit System in China [155]. It is a state-owned system to collect data from both public (traffic cameras, transit data etc.) and private (online shopping, fitness trackers etc.) data sources to monitor and analyze user behaviour to generate a single “credit

score” for each person, which helps in community well-being. The techniques fuse these data sources and remains a back box to the general public. However, the effect on user privacy with the rise of “data state” remains a debate for some [156]. A mature citizen should be on alert and always responds to any potential threat, while spreading the awareness to build a safer community in the urban city.

B. Smart Urban Area Management

Smart urban area management denotes the managing of urban area using ICT. Sub-domains in this regime composed of urban planning, governance, and smart buildings. For an application to fit into this definition, the minimum scale would be at the building level (e.g. a building management system). The main trend of data fusion techniques being applied in this domain mostly consists of objectives of extracting higher level information or increasing the data completeness. The end product of data fusion include visualization of information for respective authorities.

1) *Smart Governance*: In smart governance, managing a city is considered as a complex task as the integration of different domains and services is proven to be challenging. Transparent services integration is an example of why many governance authorities are having difficulties to sort it out. It is hard to strike a balance in developing a transparent governance policy with consideration of sensitive information. Therefore, there is only limited study materials available to the best of our knowledge. Janssen and Estevez [118] have proposed a centralized platform for cutting down government staff by shifting existing organization to rely on integration of platforms. The disaster response management is also considered as another vital element for a smart city to carry out any potential counter measurements towards disaster as shown in [157]. Apart from that, urban reporting system [56] has collected report from the city wide region on the faulty infrastructure so that immediate actions can be taken to remedy the situation. It uses cloud technology and focuses on the display of fused data report, which it also describes the location and types of infrastructure. Example of research challenges is to remove any potential fake report to prevent misuse of the reporting platform. Another example of smart governance that involves city safety can be found in [158], where it can act as an emergency aid application (light pulse on emergency through mesh network) while providing energy efficient lighting to urban area. Moreover, there are cities also working on governance platform such as New York [7], Singapore [10], Tokyo [11], Oslo [159], and others. The potential research opportunity is to propose consensus protocols within the city for better integration of services.

2) *Smart Urban Planning*: Urban planning plays an important role in developing the city economy by taking account of well-being of the urban residents. Traditionally in urban planning, aerial photography and statistical data sources (building size, population number, public amenities, etc.) are combined to understand the current development state of the city. The downside of such method is data sources frequently lacks of fine details, which resulting the output result is not

representative. To address such issue, Cheng and Toutin [43] have combined various satellite and aerial images to generate details for the exiting urban structures. Alternately, low power sensors are capable to provide a larger coverage with lower deployment cost, which give researchers the opportunity to study different points of interest in the urban area. In [81], [95], [119], a bottom up urban planning method is implemented, where sensors are installed in a designated region to capture space utilization. From the collected data, urban planners can study public space utilization pattern using an integrated portal. Here, a hybrid processing method is proposed, where the data processing and fusion occur in different stages of data pipeline. In addition, a large variety of data sources can be used for urban planning such as physical sensors [160], photography [76], [161], or hybrid data sources [85]. Despite wide variety of data sources, human interpretation is required when it comes to make decision on a proposed urban design. The need of full automated planning system would further benefit the urban planners to combine different data sources in order to achieve a more ideal city planning.

3) *Smart Building*: Urban building management provides building owner a platform to understand building’s energy consumption rate while automating building resources management. It has been extensively studied in [25], [162]–[164] and the current trend is to optimize the building resources such as hot water systems, electrical consumption, and heating ventilation & air conditioning (HVAC). In [94], Aftab et al. have combined four different parameters to predict building occupancy to control HVAC using low-cost embedded systems. Some other works such as [97], [165], [166] also have the same objectives but using different types of data sources. The potential solution for better building management system is to rely on fusing weather, human feedback, and electricity price to fine tune the building resources in order to maximize human comfort, while minimizing the energy consumption. Apart from that, fire alarm system is considered another important features of the smart building management system. Luo and Su [117] have fused three different data sources (flame, smoke, and temperature sensor) to detect any potential fire outbreak and reduce false alarms. In addition, a notification-based system is implemented to notify the property owner and manager in case of emergency. In future, potential building safety features may include a group of robots to deal with fire hazards and double duty as building security patrols.

C. Smart Environment

Smart environment studies the surrounding of a given area of interest, which covers the internal and external surrounding of a city. From the literature, we observed that majority of the data sources consist of physical and hybrid data sources, while the data scale often represent a large spatial coverage. Nowadays, the most common surrounding effects studied in the smart city include urban heat island (UHI), green house effect, and global warming. In addition, we have grouped urban waste management under this domain because it also has an environmental impact.

1) *Landscape Monitoring*: The main challenge of landscape monitoring in smart city is the sensing coverage of the data sources. To address such issue, two different sensing approaches have been used such as relying on mobile sensing or satellite-based data. Mobile sensing [106], [167] offers greater sensing capability by leveraging the mobility of moving objects (vehicles or humans). The mobile sensing technique provides a large spatial coverage, but it is not suitable for real-time applications unless there are multiple data sources to compensate the lack of spatial resolution concurrently. The output type of this mobile sensing includes combination of different spatial data in order to complete the data sources before proceed to data processing stage. Mobile sensing works such as [82], [168] utilized different data sources to complete spatial resolution and visualized the ambient changes across the city. The common characteristic of aforementioned works is feature extraction, which they visualize the processed features from the raw data sources. Majority of data input and output types in this domain are DAI-DAO and DAI-FEO since physical sensors are the common data sources. Using the satellite-based data sources, Shen et al. [46] have studied the UHI effect in a city using data sources collected over 26 years. The UHI index changes are measured through the combination of Landsat and MODIS images data. Mobile sensing offers a lower deployment cost, where it sacrifice the spatial resolution given there is limited number of sensors. Also, it has a lower coverage compared to satellite data sources. In contrast, satellite data has a wider coverage of spatial resolution but it frequently needs data enhancement and lacks of finer details.

2) *Urban City Modeling*: The surrounding natural resources of an urban city such as mountains and forests are considered as important assets of a city. The most common data sources in modeling the city area are satellite images, which as stated before, it requires data enhancement such as [169], [170] before using it. Therefore, prior work of data fusion [171] was focused on improving the satellite images quality. Only until recently, the emergence of machine learning algorithms and faster computers have created new ways to extract large variety of satellite image features. For instance in [120] and [121], forest types classification have been conducted in order to understand the variety of tree species in a specific region of interest. Both methods involve region-wide data sources and classification techniques, which are used to identify the tree species based on the forest types. With a lower deployment cost, small satellite (smallsat) and nano satellite (nanosat) could improve spatial coverage to generate a better data sources. Smart city applications leveraging satellite data will also beneficial from these deployment.

3) *Waste Management*: With astonishing rate of garbage being generated daily, waste management for an urban city can be rather challenging. Thus, it is essential to handle the waste efficiently to improve on sustainability of a city. An example of such effort could be found in [23], where they have proposed three new aspects of a smart waste management system such as: (1) infrastructure to overlook the overall life cycle of the product, (2) new business models revolving the product life cycle for preventing any waste generation, and (3)

intelligent sensor networks for waste management facilities. In [102], Catania and Ventura have combined the proximity reading and weight sensor from garbage bin to estimate the garbage capacity of a typical household. Afterwards, rubbish categories collected from user mobile devices and garbage trucks are combined to keep track of residential participation in recycling scheme. On the other hand, waste water treatment helps to manage liquid waste of urban city before discharging to river or reuse. Chang et al. [122] have combined landsat and MODIS dataset in order to trace the water pollution level of a lake. On top of that, a web portal has been deployed to visualize and monitor the water pollution region over the time. Currently, many researchers are working together to develop an efficient waste management system since there is only limited resources available on earth. The goal is to adopt the 3R (Reduce, Reuse, and Recycle) concept with the help of ICT to improve city resource sustainability.

D. Smart Industry

With the upcoming Industry 4.0 standards [172] touted as the gold standard of the future, various industries have been experiencing transformation with automation and data driven approaches. The majority of smart industry applications often leverage data collected from physical sensors while data fusion techniques are often performed at sensor or building level. Here, smart industry can be divided into three sub-domains, which are Smart Manufacturing, Smart Maintenance, and Smart Agriculture.

1) *Smart Manufacturing*: Smart manufacturing denotes the factory that depends on ICT to optimize the manufacturing process by increasing the production throughput. In [98], De Vin et al. have proposed a simulation tool to test out the management decision support by fusing undisclosed data entries and manufacturing process events. Similar to the aforementioned approach, decision based fusion can also be seen in [173], [174], which combines different machinery sensors data and data warehouse entries. The data fusion integration also considers supply chain demand in order to further optimize the manufacturing process. The challenge in this domain is to develop a self-optimizing manufacturing process while delivering the products to meet the demand of supply chain. Therefore, smart manufacturing frequently has a high correlation with the supply chain and attempts to deliver the market needs. In addition, the robotics usage in the smart manufacturing domain is nothing new. Guo et al. [51] have proposed an anomaly detection to combat potential security aspects in the robots using sensor fusion technique such as state estimation.

2) *Smart Maintenance*: The reliability and stability of the equipment and machinery is vital to all the industries to ensure smooth operation in production. Without the guarantee of smooth operation, any downtime can cost damages to reputation and also loses profit. Thus, preventive maintenance has been studied in [124], [125], [175], [176] and attempts to predict the remaining useful life (RUL) of a machine accurately. By accurately predicting the RUL, maintenance can be carried out on time to save cost only when needed.

The common data fusion techniques for predicting RUL are neural network (NN) based model such as CNN and Deep NN (DNN). Please note that, common data source in this sub-domain is physical data source such as machine states, sensors readings, and related parameters. Nonetheless on the fault detection domain, machine fault detection can be found in [96], [123], where they describe the problem of fault diagnosis and apply data fusion techniques to overcome. State estimation and classification have been used to detect the current state of the machinery. The data sources share some similarity with the preventive maintenance, where lower level of data information is preferred. This yields a faster fault detection when compared to a complex data pipeline. The research challenge here is to develop a generic and a flexible maintenance system for different scale of applications adhering to the goal of accurate fault detection.

3) *Smart Agriculture*: In order to produce sustainable food resources in smart city, smart farming [177], [178] has become a trend to meet the food supply demand in a smart city. There are two different sub-domains in smart farming such as land and sea agriculture. In the land agriculture aspect, planting crops using controlled environment has shed some light in fulfilling the city needs of fresh supplies. However, plant disease remains a potential threat to a highly-dense plantation crop framing. In [127], Moshou et al. have classified the plant disease infection through Self Organizing Map (SOM) by fusing the spectral reflection and fluorescence imaging data. This helps to isolate infected crops while it focuses on the production of healthy plants. Apart from that, electromagnetic induction sensors, vegetation index, water stress level, and radiance data are combined in [179] to better determine the partition of the crop field. Similar work also can be found in [128], where Khanum et al. propose an ontology-based fuzzy logic to classify plant disease. The research gaps in this domain involve improving live stock management as well as optimizing smart farm. On the other hand, sea agriculture is responsible for supplying the seafood supplies in a city. Obtaining fresh seafood supplies in an urban city sometimes can be rather difficult due to various factors such as delivery, city location, weather, seasonal pricing, etc. Therefore, a fresh seafood supply in a city is often not guaranteed. In order to address such issue, Huang et al. [126] have provided a solution by integrating two types of cameras for seafood freshness inspection. Camera and near infrared spectroscopy are fused through PCA and use NN to classify the freshness index. The research gaps in this domain involve developing large scale fish breeding and also wide varieties of seafood product such as calm, mussels, abalone, etc. A potential solution such as smart fish breeding with IoT has been proposed in [180], where it suggests using a moving pod to breed fishes while transporting them to destination in a particular destination simultaneously.

E. Smart Economics

Smart economics can be defined as the generic commercial activities in an urban city ranging from supply chain, logistic, finance center, to tourism. All these activities yield potential

commercial value to a city, which it depends on the unilateral or bilateral trading relationship. In this subsection, we discuss smart economics in three major sub-domains, namely, (1) Smart Commerce, (2) Smart Supply Chain, and (3) Smart Tourism.

1) *Smart Commerce*: Today, modern e-commerce platforms use multi modal data sources to reach and better understand their customers. This helps e-commerce vendors to give better product recommendations for their customers and it helps customers to make their decisions easily. Fusing customer data such as mobility, credit card purchases, and social media interactions is commonly used in modern recommender systems. In [87], Breur introduced the fusion of customer behavior data and market research data to obtain a holistic picture of the customer. Investors can leverage financial data to make investment decisions, as Hassan et al. [181] have introduced a fusion model of Hidden Markov Model (HMM), NN, and Genetic Algorithm (GA) for stock market prediction. Improving the consumer awareness is conducted in [129], by fusing real world (weather, geographical) and cyber world (Twitter, Facebook) data. The proposed system has two levels of fusion, which relies on hierarchical-based processing architecture. The data combined bottom level input and fed it into upper level for further processing to achieve its objectives.

2) *Smart Supply Chain*: In a smart supply chain, it often involves sources and destination tracking in order to understand the flow / processing of the objects. As discussed in [182], supply chain management and logistic are the fundamental of modern supplies on fulfilling the needs of an urban city. For instance in food supply chain, three tiers information fusion framework is proposed in [130] such as: (1) to accelerate data processing, (2) shelf life prediction, and (3) real-time supply chain planning. The proposed hierarchical information fusion architecture (HIFA) includes a process that is intelligently transforming the sensor's data sources into usable decision-making information. Recently, combination of blockchain technology has paved a new way for revolutionizing the existing supply chain. In [107], Tian has shown the integration of blockchain and supply chain in the agri-food supply application. It aids consumers to trace the origin of food using Radio Frequency Identification (RFID) along with database or WSN. The information also includes food origin to help consumers to identify the brand authenticity and avoids consuming counterfeit products. The research gap in this sub domain concerns with the implementation of smart supply and it needs the involvement from various commercial organizations. The consensus and national regulations are also parts of the critical factors of smart supply implementation.

3) *Smart Tourism*: The advancement of transportation technology has granted accessibility for the humans to move around the globe with ease. This phenomenon has caused rapid expansion of the tourism commercial values contributed to a city side income. Since then, Internet resources such as travel blogs and recommendation systems have influenced public to venture different locations. For instance, recommendation system [57] has been developed to recommend the place to travel based on user's information such as socioeconomic (e.g. age, education, and income) and psychological and cognitive

(experience, personality, involvement, and so forth) groups. User choices are used as feedback to further fine-tune the recommendation system using Rocchio's method. Apart from that, Miah et al. [77] have combined social media-generated big data (geo-tagged photos of tourist attraction places) to predict tourist behavioral patterns. Alternately, Viswanath et al. [131] used a smartphone based mobile application to passively track tourist location data and obtain user ratings for tourist attraction places to better understand the preferences of tourists when they visit tourist attractions. The potential research development for smart travel is to focus on using a smartphone application for improving travel experience by relying on real-time translation and augmented reality (AR) navigation.

F. Smart Human Mobility

Human mobility has been an important research area as commuting and traveling play big roles in modern life. With the help of advanced ICT, plentiful data sources related to human mobility have been collected and accessible to researchers, which yields deeper insights into the nature of human mobility as well as better improvement strategies for transportation systems. Smart human mobility, therefore, means collecting, managing, and analyzing (fusing) various data sources related to different aspects of residents' movement in order to better understand and improve the way people move. Depending on the purpose of different applications, smart human mobility domain can be further divided into three sub-domains: (1) Smart Location-Based Services, (2) Human Mobility Understanding, and (3) Smart Transportation Systems.

1) *Smart Location-Based Services*: This sub-domain aims to get the accurate position of individuals and further to provide services, such as route planning and navigation, to help them travel efficiently and comfortably, in both outdoor and indoor environments. For outdoor positioning, Global Positioning System (GPS) has been the most accurate, reliable and dominant technology since it was allowed for civilian use in 1980s [183], [184]. Less-accurate non-GPS positioning approaches, such as wifi-based localization and cell-tower triangulation, are sometimes used instead of (or together with) GPS, because they consume less energy [132], [133]. For indoor positioning, since GPS does not work well indoors, other positioning approaches have been proposed. The data collection technologies used for these approaches mainly include Wi-Fi (WLAN), inertial measurement unit (IMU), RFID tags, Bluetooth, global system for mobile communications (GSM), frequency modulation (FM), and ultra-wide band (UWB) [185], [186]. Meanwhile, multiple data sources are often fused to achieve more accurate localization results [134], [135]. Once accurate locations are obtained, either indoors or outdoors, location-based services (e.g. route planning and navigation) can be provided to end users by fusing the location sequences with other information sources such as geographic information system (GIS) data, real-time traffic data, and user preference data [136], [137], [187]–[190]. Since the outdoor positioning and location-based services have been well

developed and commercialized, the current research trend in this field is mainly focused on improving the performance (accuracy, deployment cost, and energy cost) of indoor systems and services.

2) *Human Mobility Understanding*: Positioning systems not only enable the location-based services for individuals but also provide data sources for further monitoring and understanding human mobility in a larger and more comprehensive scale. By aggregating and analyzing (fusing) the location data of residents along with GIS data of the environment, various aspects of human mobility can be monitored and the hidden patterns can be obtained. As summarized in [100], the most common subjects of monitoring and understanding human mobility include distance and duration distributions [191], origin-destination matrices [103], individual activity-based mobility patterns [192], transportation mode identification [54], and densities and flows within a building (or a cluster of buildings) [138], [193]. Results obtained from these subjects provide clues for improving transportation system [194], urban planning [195], and communication network [196]. Typical studies in this sub-domain usually fuse one data source of people's movement trajectories with the environment information, such as GIS data of the city or floor plan of a building. Although this type of approach has produced much deeper insights compared with traditional approach relied on survey data, there is a trend to fuse multiple data sources related to people's movement and obtain a more comprehensive picture of human mobility [197], [198]. Moreover, social media data sources, such as Tweets, also bring in more information regarding the mobility status in cities due to the combination of spatio-temporal data and descriptive text [86], [199].

3) *Smart Transportation Systems*: Another large part of smart mobility is the improvement of transportation systems, which mainly comes from three aspects: relieving traffic congestion, improving public transportation, and introducing new transport systems. To relieve traffic congestion, effective light control plays an important role. While existing light control systems are usually based on hand-crafted rules and do not adjust to the rapid dynamics of traffic flows, intelligent light control approaches have been proposed using different data sources, data fusion techniques, and decision making (optimization and control) algorithms [109], [200].

Challenges in this aspect mainly come from the implementation of such intelligent light control approaches. Improvement of the public transportation system is mainly conducted through the network and schedule optimization [201]. Although these two topics have been thoroughly discussed in the literature, new insights related to the public transit system (e.g. origin-destination matrices and service level obtained from big data) [139], [202] and more advanced transport modeling tools enabled by big data [203] have brought new opportunities.

Even if the existing transportation manner has been optimized, there are still problems that cannot be solved, such as last mile issue and driving accidents. Therefore, new transport systems, such as bike sharing systems and autonomous vehicle systems, are introduced. Advanced ICT and data fusion techniques are the core of the realization of these systems. For a bike sharing system, data fusion and analysis helps

to understand how the system works and evaluate different operational strategies [204], [205]. As for the autonomous vehicle system, the control of an autonomous vehicle itself is a complex data fusion process, fusing various data sources about the vehicle and the road by advanced machine learning and control algorithms [140], [206]. Security plays an important role in the autonomous vehicles deployment to ensure reliability of the autonomous driving. Examples of such techniques can be found in [50], [207].

G. Smart Infrastructure

In a smart city, infrastructure aims to provide convenience for the public by supplying resources (electricity, gas, and water) or providing services (public facility or communication systems). Here, we outline four different sub-domains for discussion, which are (1) Smart Grid, (2) Smart Energy, (3) Smart Facility, and (4) Smart Communication.

1) *Smart Grid*: The electrical grid provides an intermediate platform for relaying the electricity from the power plant to residential and industrial area. The common goal in this sub-domain is to provide reliable and stable electricity supply with the integration of ICT, which is commonly known as smart grid. Smart grid has been extensively studied in [55], [208]–[210] and the goal is to address on load and demand balancing of electricity in a particular area, building, or even household. Common technique applied in this sub-domain is forecasting, and example of such application can be found in [55], which it combines the information received from residential meters and predicts the electricity consumption load. Wang et al. [88] have proposed a different approach, where the concept of multi agent systems (MAS) is used to predict building energy consumption by denoting each meter as an agent. The common goal is to use a higher information extraction technique such as prediction, where it allows grid operators to forecast the grid demand to ensure sufficient electricity load. Test bed currently is the common method for testing out the smart grid use case and has been studied in [211]. Another common research topic is security and reliability of the smart grid system. Li et al. [49] have proposed a secure state estimation, which it can be used to address single sensor or multi-sensor scenarios. Similar works addressing smart grid security also can be found in [212], [213]. On the other hand, advanced metering infrastructure (AMI) has been studied along with the smart grid to ensure the electrical metering is tamper-proof while able to accurately measure energy consumption. For instance, work in [214] uses the clustering algorithm to identify energy theft accurately while reducing potential false positives. Meanwhile, work in [105] has presented a real-time price estimation by fusing local power and global power consumption to understand real-time electric load of the grid. In future, prosumers (producer and consumer) will emerge in the smart grid market and sole distributor paradigm will be no longer valid. This scenario greatly increase the difficulty of the energy demand and load when accounting the energy as a live market

2) *Smart Energy*: The search for clean energy resources has been an ongoing effort for many researchers in order to

cut down the dependency on the fossil fuels. Therefore, the clean energy research direction mostly focuses on renewable energy, which propose to go for a green and less carbon footprint energy producing approach. Nowadays, the most common renewable energy sources emerged in the market are solar farm [101], [141] and wind power [215], [216]. Solar energy is generated based on the conversion of the sunlight into electricity, but the energy harvesting technique suffers from limited energy harvesting time. Thus, solar irradiance prediction is crucial to ensure maximum energy throughput in the solar farm within the limited time. Huang et al. [141] have proposed to use data driven algorithms such as ABB, SVM, BRT, and Lasso, in which the information from neighboring solar plants are combined to accurately predict the solar irradiance. Meanwhile in [217], Jung and Broadwater have implemented a statistical model to fuse wind speed, direction, temperature from forecast station and online measurement to determine the total power output of the wind farm. Most of the aforementioned methods focus on improving efficiency of the existing energy harvesting methodology. Future research on the clean energy relies on various data and energy sources in order to construct a high efficiency energy harvesting model.

3) *Smart Facility*: Smart facility denotes access of physical facility that provides services to the public such as parking facility, water supply, etc. The most vital facility in a smart city would be water treatment center as clean water is an important necessity for the urban citizens. Any potential leakage or downtime of water supply in a city would be proven troublesome. Mounce et al. [104] propose a water leakage detection using classification technique, which combines all the district water meter data. Similar concept can be applied on other resources such as gas pipe leakage detection or electricity theft in smart grid. In the public facility, the wear and tear of structures can be a major issue due to the frequent rate of public usage. Hence in [99], Park et al. have combined multi-metric sensors to estimate the bridge displacement. Through this, a rough estimation of the structural health can be determined. Alternately, Khoa et al. [218] have proposed a tensor decomposition approach using the facility data sources in order to understand the facility usage details. In addition, the emergence of data centers providing various functionalities to the smart city applications such as [219]–[221] also one of the focuses for the ongoing efforts of smart city. There is also a few domains that is highly correlated with smart facility such as Smart Maintenance and Governance [56], where integration of a web portal is used to report potential damages.

4) *Smart Communication*: Communication in an urban city remains an essential infrastructure for various application platforms to communicate with each other. Not all communication platforms and standards are designed equally as each of them serve different purposes. Therefore, different standards and protocols to meet varying requirements have been established. Currently, the upcoming 5G technology [28], [222] has promised to bring integration of 5G interface with support for older generation spectrum such as LTE and Wi-Fi in order to provide seamless user experience. The common data source in 5G standards is raw signal, and that is the reason why data fusion only happens at the edge level. For

example, Huang et al. [27] and Rappaport et al. [222] have fused raw signals that are divided through multiple antenna during transmission. The receivers will receive multiple signal sources and reconstruct the original information being transferred. Further discussion of the energy efficient trade-off in wireless communication technology can be found in [223], [224]. In the IoT domain, wireless sensor network (WSN) is considered a common communication platform because of its wide coverage and low power consumption. WSN is built on top of nodes' network, which is smaller than a wireless ad hoc network. Hence, multiple nodes can be combined for encoding and decoding the packets received. Kreibich et al. [47] and Luo et al. [48] have proposed approaches to improve communication between WSN focusing on the communication mechanism between nodes. The main objective is to focus on the reliability of communication channel while maintaining the coverage (from relay to sink nodes) and also low power consumption. The research significance of communication is undoubtedly a necessity in smart city as it benefits all domains leveraging communication technology. The main goal is to design efficient and reliable communication protocols to meet different requirements of applications. Alternately, low power communication is yet another goal for IoT in order to achieve long sensing operation.

IV. CHALLENGES AND OPEN RESEARCH DIRECTIONS

After outlining the applications of the smart city that use data fusion, we discuss the potential aspects to improve the data fusion in the smart city applications observed from previous section. These aspects include potential categories or perspectives that are not discussed in Section II and III. As shown in Figure 2, we identify four major research directions, which are (1) data quality, (2) data representation, (3) data privacy and security, and (4) data fusion technique.

A. Data Quality

Quality of the data sources directly determine the quality of output results since processing module follows the "garbage in and garbage out" theorem in fusing data sources. Thus, we discuss two aspects to improve the data sources in the smart city applications, which are sensing coverage and sensing longevity.

1) *Sensing Coverage*: Sensing coverage is one of the important factors to determine the quality of data sources. Insufficient data coverage will generate a result that is not representative, and often it implies more sensors need to be installed to increase the sensing coverage. This indirectly affects the deployment cost since more physical hardware is required to compensate the sensing coverage. Apart from that, it also affects the design of communication architecture because more physical sensors are required to transmit data, and thus potentially congests the communication platform. These factors are common obstacles for a large-scale deployment in smart city applications and getting worse when increasing the deployment scale. There are two commonly used approaches to address the aforementioned issues, which are crowdsensing and mobile sensing platform.

As shown in [225], crowdsensing is one of the most cost-efficient method as personal mobile devices such as smartphones. Smartphones offer wide variety of sensors such as vibration, magnetic field, IMU, GPS, and others. The problems with crowdsensing are related to user privacy intrusion and high battery consumption when actively collecting data. User privacy is a challenge in collecting data as regulations in many countries have been facilitated to prevent applications to collect any sensitive information. This issue will be further discussed in user privacy and security sub-section. Another problems with crowdsensing are the unavailability of geolocations information or random distribution of geographical located data. These scenarios lead to inconsistent data quality. Potential way to resolve this limitation is to collect data at a fixed time and location only when needed, where incentive is provided for valid participants. Also, the trade-off problem of the mobile sensing can be further found in [226]. Through this method, only qualified data will be included as data sources, while invalid information will be automatically filtered.

Using similar concept as crowdsensing, mobile sensing has offered the same data sensing approach but only follows designated route to collect data. The idea is to leverage the mobility of the transportation (normally public transports, cabs, and garbage trucks) to conduct data collection, where the vehicles are traveling across the city. Example of mobile sensing platform can be found in [106], where garbage trucks on duty will collect the ambient data across different parts of the city weekly. An identical concept can also be implemented with the public transport systems, since majority of them follow fixed schedules. The challenge with the mobile sensing is that spatial resolution of the data may not have a finer detail when compared to crowdsensing due to fixed data collection schedule. The main cause is due to the limited accessibility of the vehicles in certain areas (pedestrian path and residential area). Potential workaround of this limitation would be combining the mobile sensing and crowdsensing data sources to generate data that covers large area within the urban city. Services integration also plays an important role in supplying platforms alternate data sources to perform data enrichment. By simulating the different IoT services in smart city as shown in [227], potential limitation or bottlenecks of smart services can be avoided in order to design a better smart city application.

2) *Data Sensing Longevity*: Long term data collection offers different aspects of knowledge discovery as data is able to cover more detail in a larger temporal resolution. The advancement of miniaturization has greatly reduced the power consumption of the sensors and IoT devices while maintaining the same sensing performance. As a result, combining both energy harvesting techniques and low energy devices are able to create a long self-sustaining sensing approach. This breakthrough allows physical sensors to run independently without the need of external power sources.

In order to preserve the longevity of physical sensors' sensing capability, energy harvesting is one of the common approach in large area networks. It allows sensors to draw energy from solar energy, vibration, or temperature difference. The most widely available energy harvesting technique is solar

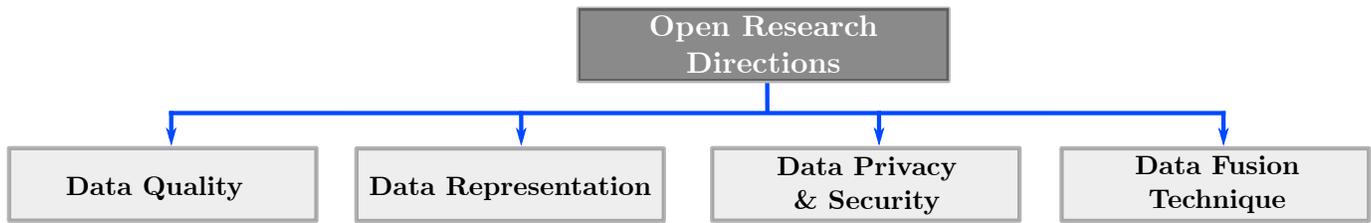


Fig. 2. Open Research Directions for Data Fusion in Smart City Applications.

panels and it can be easily obtained. Solar panel is affected by the presence of solar irradiance, where the energy harvested varies throughout the different time of the day. Contrast to solar farm, the goal here is to conserve as much energy, while maintaining the sensing capability of the physical sensors. The most notable influence would be the energy management architecture as well as the battery capacity and the solar panel efficiency. Apart from that, although temperature difference and sensor vibration are capable of harvesting energy but it is limited to certain use case and not suitable for general usage.

Alternately, potential replacement of the traditional energy harvesting technique is wireless power transfer. As shown in [228], [229], this method offers power to be transferred wirelessly without battery and energy harvesting module. Currently, there are different types of wireless power transfer technologies such as inductive coupling, capacitive coupling, magnetodynamics coupling, microwaves, and light-waves. Each of them has their limitation such as inductive coupling only has limited range of transferring energy. That being said, this technology is still relatively new, and it requires further investigation in order to guarantee its minimum working efficiency for smart city applications.

Other than using external power sources, low power sensing for carrying out the sensing tasks. In order to drive different smart city applications, various standards have been proposed for LPWAN, such as LoRaWAN by LoRa Alliance and NB-IoT Release 13 by 3GPP. LoRaWAN focuses on the long range IoT connectivity for industrial applications while the NB-IoT focuses on the indoor coverage, low cost, long battery life, and stable communication in high density communication channel. The main reason to use low power sensing approach is due to the high compatibility with large scale deployment relying on the low bit rate communication channel usage. However, standardization of these protocols remains a challenge in LPWAN due to the possibly of using unlicensed spectrum, where organizations may choose not to follow the agreed spectrum. In future, low power communication will ensure the long term sensing capability of physical sensors in the smart city applications and therefore will improve data sources quality.

B. Data Representation

A high speed Internet connection provides easy access to many genres of data sources and creates opportunity to study wide variety of different data sources. However, large variety of data sources frequently indicate the incompatibility of data formats. The problem becomes more obvious when

there is no standardization on the data format. To tackle such problem, data ontology is the building block to represent the data sources to connect different sources of data for seamless services integration. If the format of the data source cannot be interpreted, it will be marked as useless for the platform integrator. Therefore, semantic web has been proposed as an extension to WWW web services utilizing Resource Description Framework (RDF) to provide standard data exchange formats. It opens the path to create different solutions for the IoT applications and it supports the Open Government Data (OGD) principles [230]. To date, there are few common ontology languages have been developed such as Delivery Context (DCN) [231], Web Ontology Language (OWL) [232], Resource Description Framework Schema (RDFS) [233], Semantic Sensor Network (SSN) [234], and others. Majority of the ontology languages only focus on one application domain because they are not suitable for representing the metadata from other domains. This causes data segmentation in the smart city applications, where further increases the gap between different smart city domains. Thus, DBpedia [235] is designed to address the aforementioned issue using public and private stocks of semantic web. DBpedia has provided solutions for the ontology software as it offers different classes and types that are available on the Wikipedia. That being said, not all applications adopt the idea of DBpedia and there is a fraction of applications remain conservative using proprietary data representation. Apart from that, Message Queuing Telemetry Transport (MQTT) [236] v3.1 protocol has been introduced as one of the protocols to address ontology problems between brokers. It offers machine to machine (M2M) communication by providing lightweight publish and subscribe messaging services, where network bandwidth limitation is one of the main constraint. It is possible to combine the aforementioned technologies in order to generate a better data integration for data fusion purposes across different domains. Therefore, the future agenda for the ontology language is to encourage integration of different levels of data sources using different system architecture such as edge, fog, and cloud computing.

C. Privacy and Security

1) *Privacy*: Collecting urban residents' data in a smart city application can be challenging due the nature of sensitive data that can be misused if poorly managed. As privacy issue has been discussed extensively by the authors in [116], [237], [238], misuse of private information may lead to catastrophic events such as information theft, or identity fraud. Currently in Europe, General Data Protection Regulation (GDPR) as dis-

cussed in [239], [240] has been proposed to better address the data privacy concern of the Internet. In other countries, there are also similar efforts to enforce data privacy protection such as Canada's Personal Information Protection and Electronic Documents Act (PIPEDA), China's China Data Protection Regulations (CDPR), Singapore's Personal Data Protection Act (PDPA), Japan's Personal Information Protection Commission (PIPC), etc. Meanwhile in USA, Health Insurance Portability and Accountability Act of 1996 (HIPAA), the Children's Online Privacy Protection Act of 1998 (COPPA), and the Fair and Accurate Credit Transactions Act of 2003 (FACTA) have been introduced to improve with the information flow efficiency across agencies. This is also a part of the efforts to prevent sensitive information being available for unauthorized parties.

Majority of the policies and regulations emphasize on the users' consent for collecting personal data and this can be problematic as not all platforms provide ample security for data storage. With insufficient security measurements, the data collected may be compromised, which may lead to tainted reputation and loss of public faith. For instance, Facebook and Cambridge Analytica scandal [241] has shown potential misuse of user data collected. With that in mind, potential right of accessing data sources could be revoked if the data source is not handled properly by the right person. Hence, privacy and security should be the responsibility for both platforms and users. A thorough review has been conducted in [242], which works on the IoT requirements to address privacy issues. Potential solution for the aforementioned problem is to use hybrid data fusion technique in a smart city application. The idea here is to locally fuse the sensitive information (user identity, phone number, bank account number) into generic information, before uploading to the cloud for further processing. The benefits of such approach are two-folds, which are the ability to offload computational cost and to preserve sensitive information at the physical sensor only. In addition, we can leverage machine learning approaches such as [243], [244] to generate synthetic datasets with identical data characteristic for study purpose. This eliminates the chances of private data been leaked out and encourage the openness of datasets to be studied by different researchers and data scientists. To draw a clear line between generic and sensitive data remains a debate among researchers. In future, the data fusion can be applied at the lower level to remove any potential sensitive data.

2) *Security*: According to Kitchin [245], there are two general security concerns in the smart city applications, which are security of technology/infrastructure (data center, services, and system architecture) and data security (data generation, storage, and communication).

The security of the technology and infrastructure highly relies on the design architecture of the system being deployed. Depending on the application requirements, it varies from traditional client server architecture to decentralized architecture. The main objective is to deploy a hack-proof/exploit-less system architecture. Alternately, there are also ways of improving security of system architecture such as incentive/bounty for reporting flaws, simulating injection attacks, security assessment from third party, etc. Nowadays, the security enhancement

focuses towards continuous effort as the technology has been changing rapidly. For instance, security works in [246]–[248] have proposed different strategies to enhance the security of the smart city application's architecture by focusing on the common security standards/practices/protocols. This shows that as the number of smart city applications increase rapidly, system architectures implemented with the security design in mind become apparent with good practices and standard architecture design. Subsequently, regular security assessment and auditing also pave way for a safer smart city applications deployment.

Meanwhile, data security also contributes to the significant part of smart city applications ecosystem from generation, storage, and communication. The common method to combat such issue is leveraging encryption techniques, where it encodes the data so that only the authorized parties have access to it. For instance in [249], Wang et al. have introduced an attribute based encryption scheme, which it allows fine-grained access control, scalable key management, and flexible data distribution. In addition, encryption also can be used in the communication platform between IoT devices in smart city application as shown in [250], [251] to prevent information hijacking.

Despite constant effort of cyber security researchers developing new security schemes, the numbers of data breaches and cyber threats increase every year according to David et al. [252]. The main culprit of such occurrence is due to negligence of data security practices/implementation. Security often appears to be an afterthought in deployment of a smart city application. Thus, in order to combat such threat, the smart city application should comply with security standards as shown in [253] to mitigate the chances of becoming a victim.

D. Data Fusion Techniques

Extracting knowledge from a smart city application frequently involves data mining techniques in order to fuse different data sources. Lower tier data fusion techniques have been well explored in [39] and the current research trend focuses more on the machine learning approach. The main reason why machine learning approach has gained so much attention is due to its capability of handling high dimensional data. The problem of high dimensional data is also known as curse of dimensionality as described by Bellman [254]. In this context, we discuss two research trends on applying machine learning techniques in data fusion as follows:

1) *Explainable Deep Neural Network*: Lately, supervised machine learning techniques focus on the DNN, where the in-depth reviews of the recent development can be found in [255]–[257]. Major research efforts aim to increase the explainability of the model such as NN, CNN, and DNN rather than using them as black box models. To this end, explainable AI (XAI) [258] is the new motivation for data scientists to explore the interpretable learning paradigm of the modeling in order to provide a semantic meaning behind modeling logic. This new learning process has driven three big fields in the deep learning domains, which are (1) Deep Explanation, (2)

Interpretable Model, and (3) Model Induction. To develop a deep explanation on the model interpretation, the cognitive layers will act as an intermediate layer between learning and explanation layer in order to cast the learned abstractions, policies, and clusters information into an explainable format. Subsequently, the interpretable model such as Bayesian learning [259] can be built to explain the uncertainties required when developing the deep learning models to learn the choices of a learning process. Alternate approach has proposed to use subspace approximation with an adjusted bias technique [260] to build interpretable CNN, which uses feed forward design to better explain the model's choice in allocating certain hyper-parameters. Meanwhile, model induction refers to the technique used for inferring the model's decision and learning progress. Through a thorough understanding of the model, parameters can be fine-tuned to increase the learning optimization rate in a long-term application deployment. Hence, the search of XAI is an important milestone for the data scientists, which can be used to explain the learning process and the decision machine learning made. An example of potential use case would be trying to understand the reason behind (also known as reasoning in some literatures) the predictive maintenance decision machine learning rather than performing maintenance due to the result of predictive algorithm.

2) *Unsupervised Data Fusion*: In the smart city applications, collecting the ground truth could be proven challenging due to the uncertainties and errors in the collected data sources. Hence, obtaining labels or data annotation are another problems with certain data sources. Despite the rapid development of advanced modeling tools like DNN, it still requires labels and data annotation in order to achieve objectives of extracting higher information. There are a few approaches that address the lack of labels such as manual annotation, crowd labeling, software annotation, and pattern labeling. However, manual annotation only works well with a small dataset while other approaches do not guarantee the correctness of end result. This shows a big research gap to seek a better way to label data sources accurately.

Research works such as Zhou et al. [261], [262] have attempted to fix unlabeled data by transforming them into useful features to achieve certain objectives. Traditionally, raw data is required to be preprocessed into something meaningful, but it still suffers from the need of data cleansing and amputation. The simplest method would be to solely depend on the filtering technique. However, aggressive filtering may remove large amount of raw data resulting potential loss of knowledge. Another simple solution is to increase the number of reliable data sources to be fused to create potential annotation. Increasing data sources often indicates an increment of the overall deployment cost. Alternative solution to the increased deployment cost is to use transfer learning [263], where the knowledge from existing domain can be transferred to other domain to learn from it.

3) *Emergence of Hybrid Model*: The emergence of the hybrid models has become common due to wide variety of data sources available. It allows different levels of data sources (high, low, or both) to combine in order to create potential insights in a particular domain. It also helps to solve the

data privacy problem along with machine learning technique, which has opened up many opportunities for researchers and data scientist to study on these big data collected. One example of the hybrid model is shown as follows: an urban planning system has different data sources as input such as human comfort factor index (environmental ambient sensors), positive urban city factor (feedback data on urban area such as greenery, surrounding amenities, recreational parks, and others), and cyber data (social media input) to design a fully automated urban planning system by fulfilling predefined criteria. The result from the data fusion needs to be explainable as discussed in the previous XAI for understanding choices made by the automation software. In this example, different tiers of data sources are fused using data sources types ($D1, D2$) and the result is some features. Eventually, these features will be combined to generate a potential plan for city through computation modeling ($D3, D4$). By joining different data sources, simulation can be used concurrently to verify the performance of urban planning system before deploying to the city. In future, implementation of the hybrid model will become a general trend due to wide availability of the data sources and processing platforms. As mentioned in the discussion, data ontology is another key factor to allow data sources to be connected from different platforms to provide knowledge for the smart city applications.

V. CONCLUSION

This paper presents an overall view of the data fusion techniques found in the smart city applications. Easy accessibility of the data sources has paved way for data fusion in different smart city applications in various forms. The increasing trends of data fusion in the smart city applications create the need for a new evaluation method. Therefore, we propose a multi-perspectives classification for the smart city applications that involve data fusion techniques. The data fusion classification based on multi-perspectives introduced in this paper are: (1) Fusion Objectives, (2) Fusion Techniques, (3) Data Input and Output Types, (4) Data Source Types, (5) Data Fusion Scales, and (6) System Architecture. Using the proposed multi-perspectives, we evaluated some selected works in the smart city applications and we also discussed the research trend for each domain respectively. Next, we also discuss four open research directions of data fusion in a smart city application such as data quality, data representation, data privacy & security, and data fusion technique. Overall, we are certain that generic nature of the multi-perspectives classification is able to perform well with various smart city applications for different domains that leverage the data fusion techniques. In addition, an in-depth analysis can be further extended onto individual domain to study the common requirements and techniques applied, which we do not include in this paper due to limited paper length. A successful smart city application is built on top of the data (also known as data-driven architecture) and data fusion has provided a wide variety of techniques to improve the input data for an application. Therefore, data fusion has opened the path for various applications to gain insights about the city. This also holds the key for a smart

city to further understand and improve the domains that it is lacking.

ACKNOWLEDGMENT

The research work was supported in part by the National Research Foundation (NRF) of Singapore via the Green Buildings Innovation Cluster (GBIC) administered by the Building and Construction Authority (BCA) Green Building Innovation Cluster (GBIC) Program Office; in part, by the SUTD-MIT International Design Center (IDC; idc@sutd.edu.sg); in part by Natural Science Foundation of China (NSFC) through Project No. 61750110529, 61850410535 and Higher Education Commission (HEC) Pakistan through grant number NRPUP#5913. We thank our colleagues and reviewers, who have provided insight and expertise that greatly assisted with improving the context of this survey paper.

REFERENCES

- [1] UN, *68% of the world population projected to live in urban areas by 2050, says UN*. UN, 2018. [Online]. Available: <https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html>
- [2] A. Boulton, S. D. Brunn, and L. Devriendt, "18 cyber-infrastructures and smartworld cities: physical, human and soft infrastructures," *International Handbook of Globalization and World Cities*, p. 198, 2011.
- [3] R. G. Hollands, "Will the real smart city please stand up? intelligent, progressive or entrepreneurial?" *City*, vol. 12, no. 3, pp. 303–320, 2008.
- [4] T. Nam and T. A. Pardo, "Smart city as urban innovation: Focusing on management, policy, and context," in *2011 Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance*. ACM, 2011, pp. 185–194.
- [5] D. Toppeta, "The smart city vision: how innovation and ict can build smart, livable, sustainable cities," *The Innovation Knowledge Foundation*, vol. 5, pp. 1–9, 2010.
- [6] P. Berrone and J. Enric, "Iese cities in motion index 2018," *IESE Business School, University of Navarra, España*, 2018.
- [7] C. of New York, "Nyc sustainability," 2018. [Online]. Available: <https://www1.nyc.gov/site/sustainability/index.page>
- [8] G. L. Authority, "Smart london," 2018. [Online]. Available: <https://www.london.gov.uk/what-we-do>
- [9] M. D. Paris, "Paris smart and sustainable," 2018. [Online]. Available: <https://api-site-cdn.paris.fr/images/99354>
- [10] S. Nation and D. G. Office, "Smart nation singapore," 2018. [Online]. Available: <https://www.smartnation.sg/>
- [11] T. M. Government, "New tokyo. new tomorrow. the action plan for 2020," 2016. [Online]. Available: http://www.metro.tokyo.jp/english/about/plan/documents/pocket_english.pdf
- [12] D. T. T. Limited, "Super smart city - happier society with higher quality," *Deloitte CN Public Sector*, 2018. [Online]. Available: <https://www2.deloitte.com/cn/en/pages/public-sector/articles/super-smart-city.html>
- [13] MIT, "Mit senseable city laboratory," 2018. [Online]. Available: <http://senseable.mit.edu/>
- [14] E. Zurich, "Future cities laboratory," 2018. [Online]. Available: <http://www.fcl.ethz.ch/>
- [15] SINTEF, "Sintef smart cities," 2018. [Online]. Available: <https://www.sintef.no/en/smartcities/#/>
- [16] SMART, "Future urban mobility," 2018. [Online]. Available: <https://fm.smart.mit.edu/>
- [17] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion*, vol. 14, no. 1, pp. 28–44, 2013.
- [18] F. Castanedo, "A review of data fusion techniques," *The Scientific World Journal*, vol. 2013, 2013.
- [19] 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.
- [20] M. Wang, C. Perera, P. P. Jayaraman, M. Zhang, P. Strazdins, R. Shyamsundar, and R. Ranjan, "City data fusion: Sensor data fusion in the internet of things," *International Journal of Distributed Systems and Technologies (IJ DST)*, vol. 7, no. 1, pp. 15–36, 2016.
- [21] Y. Zheng *et al.*, "Methodologies for cross-domain data fusion: An overview," *IEEE Transactions on Big Data*, vol. 1, no. 1, pp. 16–34, 2015.
- [22] N.-E. El Faouzi, H. Leung, and A. Kurian, "Data fusion in intelligent transportation systems: Progress and challenges—a survey," *Information Fusion*, vol. 12, no. 1, pp. 4–10, 2011.
- [23] B. Esmailian, B. Wang, K. Lewis, F. Duarte, C. Ratti, and S. Behdad, "The future of waste management in smart and sustainable cities: A review and concept paper," *Waste Management*, vol. 81, pp. 177–195, 2018.
- [24] L. Da Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Transactions on industrial informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.
- [25] Z. Chen, C. Jiang, and L. Xie, "Building occupancy estimation and detection: A review," *Energy and Buildings*, 2018.
- [26] X. Qin and Y. Gu, "Data fusion in the internet of things," *Procedia Engineering*, vol. 15, pp. 3023–3026, 2011.
- [27] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using lse and sparse message passing for mmwave mimo systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 1, pp. 245–259, 2019.
- [28] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [29] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5g," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.
- [30] M. Erol-Kantarci and H. T. Mouftah, "Wireless sensor networks for cost-efficient residential energy management in the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 314–325, 2011.
- [31] A. A. Sreesh, S. Somal, and I.-T. Lu, "Cognitive radio based wireless sensor network architecture for smart grid utility," in *2011 IEEE Long Island Systems, Applications and Technology Conference*. IEEE, 2011, pp. 1–7.
- [32] Y.-G. Yue and P. He, "A comprehensive survey on the reliability of mobile wireless sensor networks: Taxonomy, challenges, and future directions," *Information Fusion*, vol. 44, pp. 188–204, 2018.
- [33] O. Georgiou and U. Raza, "Low power wide area network analysis: Can lora scale?" *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 162–165, 2017.
- [34] 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.
- [35] M. Chen, Y. Miao, Y. Hao, and K. Hwang, "Narrow band internet of things," *IEEE Access*, vol. 5, pp. 20557–20577, 2017.
- [36] Y.-P. E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, and H. S. Razaghi, "A primer on 3gpp narrowband internet of things," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 117–123, 2017.
- [37] I. A. T. Hashem, V. Chang, N. B. Anuar, K. Adewole, I. Yaqoob, A. Gani, E. Ahmed, and H. Chiroma, "The role of big data in smart city," pp. 748–758, 2016.
- [38] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.
- [39] B. V. Dasarathy, "Sensor fusion potential exploitation-innovative architectures and illustrative applications," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 24–38, 1997.
- [40] H. F. Durrant Whyte, "Sensor models and multisensor integration," *The International Journal of Robotics Research*, vol. 7, no. 6, pp. 97–113, 1988.
- [41] A. N. Steinberg and C. L. Bowman, "Revisions to the jdl data fusion model," in *Handbook of Multisensor Data Fusion*. CRC Press, 2008, pp. 65–88.
- [42] S. Grime and H. F. Durrant-Whyte, "Data fusion in decentralized sensor networks," *Control Engineering Practice*, vol. 2, no. 5, pp. 849–863, 1994.
- [43] P. Cheng and T. Toutin, "Urban planning using data fusion of satellite

- and aerial photo images," in *1997 IEEE International Geoscience and Remote Sensing, 1997. IGARSS'97. Remote Sensing-A Scientific Vision for Sustainable Development.*, vol. 2. IEEE, 1997, pp. 839–841.
- [44] C. Huang, G. C. Alexandropoulos, C. Yuen, and M. Debbah, "Deep learning for ul/dl channel calibration in generic massive mimo systems large intelligent surfaces," 2019, pp. 1–6. [Online]. Available: [Online]https://arxiv.org/abs/1903.02875, toappearICC2019
- [45] X. Hong, C. Nugent, M. Mulvenna, S. McClean, B. Scotney, and S. Devlin, "Evidential fusion of sensor data for activity recognition in smart homes," *Pervasive and Mobile Computing*, vol. 5, no. 3, pp. 236–252, 2009.
- [46] H. Shen, L. Huang, L. Zhang, P. Wu, and C. Zeng, "Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: A 26-year case study of the city of wuhan in china," *Remote Sensing of Environment*, vol. 172, pp. 109–125, 2016.
- [47] O. Kreibich, J. Neuzil, and R. Smid, "Quality-based multiple-sensor fusion in an industrial wireless sensor network for mcm," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 9, pp. 4903–4911, 2014.
- [48] X. Luo, D. Zhang, L. T. Yang, J. Liu, X. Chang, and H. Ning, "A kernel machine-based secure data sensing and fusion scheme in wireless sensor networks for the cyber-physical systems," *Future Generation Computer Systems*, vol. 61, pp. 85–96, 2016.
- [49] H. Li, L. Lai, and W. Zhang, "Communication requirement for reliable and secure state estimation and control in smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 476–486, 2011.
- [50] J. Petit, B. Stottelaar, M. Feiri, and F. Kargl, "Remote attacks on automated vehicles sensors: Experiments on camera and lidar," *Black Hat Europe*, vol. 11, p. 2015, 2015.
- [51] P. Guo, H. Kim, N. Virani, J. Xu, M. Zhu, and P. Liu, "Roboads: Anomaly detection against sensor and actuator misbehaviors in mobile robots," in *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. IEEE, 2018, pp. 574–585.
- [52] N. Dawar and N. Kehtarnavaz, "A convolutional neural network-based sensor fusion system for monitoring transition movements in healthcare applications," in *2018 IEEE 14th International Conference on Control and Automation (ICCA)*. IEEE, 2018, pp. 482–485.
- [53] L. Jayasinghe, N. Wijerathne, C. Yuen, and M. Zhang, "Feature learning and analysis for cleanliness classification in restrooms," *IEEE Access*, vol. 7, pp. 14 871–14 882, 2019.
- [54] A. Ghorpade, F. C. Pereira, F. Zhao, C. Zegras, and M. Ben-Akiva, "An integrated stop-mode detection algorithm for real world smartphone-based travel survey," in *Transportation Research Board 94th Annual Meeting*, vol. 15-6021, 2015, pp. 1–16.
- [55] W. Luan, D. Sharp, and S. Lancashire, "Smart grid communication network capacity planning for power utilities," in *2010 IEEE PES Transmission and Distribution Conference and Exposition*. IEEE, 2010, pp. 1–4.
- [56] S. Consoli, D. Reforgiato Recupero, M. Mongiovi, V. Presutti, G. Cataldi, and W. Patatu, "An urban fault reporting and management platform for smart cities," in *2015 Proceedings of the 24th International Conference on World Wide Web*. ACM, 2015, pp. 535–540.
- [57] F. Ricci, "Travel recommender systems," *IEEE Intelligent Systems*, vol. 17, no. 6, pp. 55–57, 2002.
- [58] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerging Artificial Intelligence Applications in Computer Engineering*, vol. 160, pp. 3–24, 2007.
- [59] T. M. Cover, P. E. Hart *et al.*, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [60] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Systems Magazine*, vol. 29, no. 6, pp. 82–100, 2009.
- [61] J. P. Shaffer, "Multiple hypothesis testing," *Annual review of psychology*, vol. 46, no. 1, pp. 561–584, 1995.
- [62] I. J. Myung, "Tutorial on maximum likelihood estimation," *Journal of mathematical Psychology*, vol. 47, no. 1, pp. 90–100, 2003.
- [63] G. Welch, G. Bishop *et al.*, "An introduction to the Kalman filter," 1995.
- [64] B. Ristic, S. Arulampalam, and N. Gordon, "Beyond the kalman filter," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 7, pp. 37–38, 2004.
- [65] J. K. Uhlmann, "Covariance consistency methods for fault-tolerant distributed data fusion," *Information Fusion*, vol. 4, no. 3, pp. 201–215, 2003.
- [66] G. E. Box and G. C. Tiao, *Bayesian inference in statistical analysis*. John Wiley & Sons, 2011, vol. 40.
- [67] H. Wu, M. Siegel, R. Stiefelhagen, and J. Yang, "Sensor fusion using Dempster-Shafer theory [for context-aware hci]," in *IMTC/2002. Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference (IEEE Cat. No. 00CH37276)*, vol. 1. IEEE, 2002, pp. 7–12.
- [68] F. Herrera, E. Herrera-Viedma, and L. Martinez, "A fusion approach for managing multi-granularity linguistic term sets in decision making," *Fuzzy sets and systems*, vol. 114, no. 1, pp. 43–58, 2000.
- [69] D. A. Pacyga, *Applied linear regression models*. Chicago: University of Chicago Press, 1996.
- [70] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 561–580, 1975.
- [71] C. Lork, B. Rajasekhar, C. Yuen, and N. M. Pindoriya, "How many watts: A data driven approach to aggregated residential air-conditioning load forecasting," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2017, pp. 285–290.
- [72] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.
- [73] H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 16–24, 2013.
- [74] X. J. Zhu, "Semi-supervised learning literature survey," University of Wisconsin-Madison Department of Computer Sciences, Tech. Rep., 2005.
- [75] I. Jolliffe, *Principal component analysis*. Springer, 2011.
- [76] F. Zhang, B. Zhou, L. Liu, Y. Liu, H. H. Fung, H. Lin, and C. Ratti, "Measuring human perceptions of a large-scale urban region using machine learning," *Landscape and Urban Planning*, vol. 180, pp. 148–160, 2018.
- [77] S. J. Miah, H. Q. Vu, J. Gammack, and M. McGrath, "A big data analytics method for tourist behaviour analysis," *Information & Management*, vol. 54, no. 6, pp. 771–785, 2017.
- [78] J. Nichol and M. S. Wong, "Modeling urban environmental quality in a tropical city," *Landscape and Urban Planning*, vol. 73, no. 1, pp. 49–58, 2005.
- [79] C.-T. Fan, Y.-K. Wang, and C.-R. Huang, "Heterogeneous information fusion and visualization for a large-scale intelligent video surveillance system," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 4, pp. 593–604, 2017.
- [80] C. Ware, *Information visualization: perception for design*. Elsevier, 2012.
- [81] B. P. L. Lau, T. Chaturvedi, B. K. K. Ng, K. Li, M. S. Hasala, and C. Yuen, "Spatial and temporal analysis of urban space utilization with renewable wireless sensor network," in *2016 IEEE/ACM 3rd International Conference on Big Data Computing, Applications and Technologies*. ACM, 2016, pp. 133–142.
- [82] Y. Zheng, F. Liu, and H.-P. Hsieh, "U-air: When urban air quality inference meets big data," in *2013 ACM SIGKDD Proceedings of the 19th International Conference on Knowledge Discovery and Data Mining*. ACM, 2013, pp. 1436–1444.
- [83] L. Spinello and K. O. Arras, "People detection in rgb-d data," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2011, pp. 3838–3843.
- [84] S. Lee, D. Yoon, and A. Ghosh, "Intelligent parking lot application using wireless sensor networks," in *2008 International Symposium on Collaborative Technologies and Systems*. IEEE, 2008, pp. 48–57.
- [85] L.-C. Chen, T.-A. Teo, Y.-C. Shao, Y.-C. Lai, and J.-Y. Rau, "Fusion of lidar data and optical imagery for building modeling," *International Archives of Photogrammetry and Remote Sensing*, vol. 35, no. B4, pp. 732–737, 2004.
- [86] S. Suma, R. Mehmood, and A. Albeshri, "Automatic event detection in smart cities using big data analytics," in *International Conference on Smart Cities, Infrastructure, Technologies and Applications*. Springer, 2017, pp. 111–122.

- [87] T. Breur, "Data analysis across various media: Data fusion, direct marketing, clickstream data and social media," *Journal of Direct, Data and Digital Marketing Practice*, vol. 13, no. 2, pp. 95–105, 2011.
- [88] Z. Wang, L. Wang, A. I. Dounis, and R. Yang, "Multi-agent control system with information fusion based comfort model for smart buildings," *Applied Energy*, vol. 99, pp. 247–254, 2012.
- [89] J. A. Balazs and J. D. Velásquez, "Opinion mining and information fusion: a survey," *Information Fusion*, vol. 27, pp. 95–110, 2016.
- [90] B. Guo, Z. Wang, Z. Yu, Y. Wang, N. Y. Yen, R. Huang, and X. Zhou, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Computing Surveys (CSUR)*, vol. 48, no. 1, pp. 1–31, 2015.
- [91] S. H. Marakkalage, S. Sarica, B. P. L. Lau, S. K. Viswanath, T. Balasubramaniam, C. Yuen, B. Yuen, J. Luo, and R. Nayak, "Understanding the lifestyle of older population: Mobile crowdsensing approach," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 1, pp. 82–95, 2019.
- [92] E. Estellés-Arolas and F. González-Ladrón-De-Guevara, "Towards an integrated crowdsourcing definition," *Journal of Information science*, vol. 38, no. 2, pp. 189–200, 2012.
- [93] J. Howe, "The rise of crowdsourcing," *Wired magazine*, vol. 14, no. 6, pp. 1–4, 2006.
- [94] M. Aftab, C. Chen, C.-K. Chau, and T. Rahwan, "Automatic hvac control with real-time occupancy recognition and simulation-guided model predictive control in low-cost embedded system," *Energy and Buildings*, vol. 154, pp. 141–156, 2017.
- [95] L. You, B. Tunçer, and H. Xing, "Harnessing multi-source data about public sentiments and activities for informed design," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 2, pp. 343–356, 2018.
- [96] F. Serdio, E. Lughofer, K. Pichler, T. Buchegger, M. Pichler, and H. Efcendic, "Fault detection in multi-sensor networks based on multivariate time-series models and orthogonal transformations," *Information Fusion*, vol. 20, pp. 272–291, 2014.
- [97] W. Tushar, N. Wijerathne, W.-T. Li, C. Yuen, H. V. Poor, T. K. Saha, and K. L. Wood, "Internet of things for green building management: Disruptive innovations through low-cost sensor technology and artificial intelligence," *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 100–110, 2018.
- [98] L. J. De Vin, A. H. Ng, J. Oscarsson, and S. F. Andler, "Information fusion for simulation based decision support in manufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 22, no. 5-6, pp. 429–436, 2006.
- [99] J.-W. Park, S.-H. Sim, and H.-J. Jung, "Wireless displacement sensing system for bridges using multi-sensor fusion," *Smart Materials and Structures*, vol. 23, no. 4, p. 045022, 2014.
- [100] Y. Zhou, B. P. L. Lau, C. Yuen, B. Tuncer, and E. Wilhelm, "Understanding urban human mobility through crowdsensed data," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 52–59, 2018.
- [101] S. Katoch, G. Muniraju, S. Rao, A. Spanias, P. Turaga, C. Tepedelenlioglu, M. Banavar, and D. Srinivasan, "Shading prediction, fault detection, and consensus estimation for solar array control," in *2018 IEEE Industrial Cyber-Physical Systems (ICPS)*. IEEE, 2018, pp. 217–222.
- [102] V. Catania and D. Ventura, "An approach for monitoring and smart planning of urban solid waste management using smart-m3 platform," in *2014 Proceedings of 15th Conference of Open Innovations Association FRUCT*. IEEE, 2014, pp. 24–31.
- [103] J. L. Toole, S. Colak, B. Sturt, L. P. Alexander, A. Evsukoff, and M. C. González, "The path most traveled: Travel demand estimation using big data resources," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 162–177, 2015.
- [104] S. R. Mounce, A. Khan, A. S. Wood, A. J. Day, P. D. Widdop, and J. Machell, "Sensor-fusion of hydraulic data for burst detection and location in a treated water distribution system," *Information Fusion*, vol. 4, no. 3, pp. 217–229, 2003.
- [105] S. Izumi and S.-i. Azuma, "Real-time pricing by data fusion on networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 1175–1185, 2018.
- [106] A. Anjomshoaa, F. Duarte, D. Rennings, T. Matarazzo, P. de Souza, and C. Ratti, "City scanner: Building and scheduling a mobile sensing platform for smart city services," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4567–4579, 2018.
- [107] F. Tian, "An agri-food supply chain traceability system for china based on rfid & blockchain technology," in *2016 13th International Conference on Service Systems and Service Management*. IEEE, 2016, pp. 1–6.
- [108] R. Mehmood, M. A. Faisal, and S. Altowaijri, "Future networked healthcare systems: A review and case study," in *Handbook of Research on Redesigning the Future of Internet Architectures*. IGI Global, 2015, pp. 531–558.
- [109] F. Ahmed and Y. Hawas, "An integrated real-time traffic signal system for transit signal priority, incident detection and congestion management," *Transportation Research Part C: Emerging Technologies*, vol. 60, pp. 52–76, 2015.
- [110] A. Fleury, M. Vacher, and N. Noury, "Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 274–283, 2010.
- [111] H. M. Hondori, M. Khademi, and C. V. Lopes, "Monitoring intake gestures using sensor fusion (microsoft kinect and inertial sensors) for smart home tele-rehab setting," in *2012 IEEE 1st Annual Healthcare Innovation Conference*, 2012, pp. 36–39.
- [112] M. S. Hossain, G. Muhammad, and A. Alamri, "Smart healthcare monitoring: A voice pathology detection paradigm for smart cities," *Multimedia Systems*, pp. 1–11, 2017.
- [113] H. Medjahed, D. Istrate, J. Boudy, J.-L. Baldinger, and B. Dorizzi, "A pervasive multi-sensor data fusion for smart home healthcare monitoring," in *2011 IEEE International Conference on Fuzzy Systems*. IEEE, 2011, pp. 1466–1473.
- [114] L. Zhang, H. Leung, and K. C. C. Chan, "Information fusion based smart home control system and its application," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 3, pp. 1157–1165, 2008.
- [115] R. Mehmood, F. Alam, N. N. Albogami, I. Katib, A. Albeshri, and S. M. Altowaijri, "Utilearn: a personalised ubiquitous teaching and learning system for smart societies," *IEEE Access*, vol. 5, pp. 2615–2635, 2017.
- [116] A. B. Chan, Z.-S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–7.
- [117] R. C. Luo and K. L. Su, "Autonomous fire-detection system using adaptive sensory fusion for intelligent security robot," *IEEE/ASME Transactions on Mechatronics*, vol. 12, no. 3, pp. 274–281, 2007.
- [118] M. Janssen and E. Estevez, "Lean government and platform-based governance doing more with less," *Government Information Quarterly*, vol. 30, pp. S1–S8, 2013.
- [119] B. P. L. Lau, N. Wijerathne, B. K. K. Ng, and C. Yuen, "Sensor fusion for public space utilization monitoring in a smart city," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 473–481, 2018.
- [120] M. Lu, B. Chen, X. Liao, T. Yue, H. Yue, S. Ren, X. Li, Z. Nie, and B. Xu, "Forest types classification based on multi-source data fusion," *Remote Sensing*, vol. 9, no. 11, p. 1153, 2017.
- [121] P. T. Wolter and P. A. Townsend, "Multi-sensor data fusion for estimating forest species composition and abundance in northern minnesota," *Remote Sensing of Environment*, vol. 115, no. 2, pp. 671–691, 2011.
- [122] N.-B. Chang, C. Mostafiz, Z. Sun, W. Gao, and C.-F. Chen, "Developing a prototype satellite-based cyber-physical system for smart wastewater treatment," in *2017 IEEE 14th International Conference on Networking, Sensing and Control*. IEEE, 2017, pp. 339–344.
- [123] H. Zhang and Y. Deng, "Engine fault diagnosis based on sensor data fusion considering information quality and evidence theory," *Advances in Mechanical Engineering*, vol. 10, no. 11, p. 1687814018809184, 2018.
- [124] L. Jayasinghe, T. Samarasinghe, C. Yuen, and S. S. Ge, "Temporal convolutional memory networks for remaining useful life estimation of industrial machinery," in *2019 IEEE International Conference on Industrial Technology*, 2018, pp. 1–6.
- [125] N. Ghosh, Y. Ravi, A. Patra, S. Mukhopadhyay, S. Paul, A. Mohanty, and A. Chattopadhyay, "Estimation of tool wear during cnc milling using neural network-based sensor fusion," *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 466–479, 2007.
- [126] X. Huang, H. Xu, L. Wu, H. Dai, L. Yao, and F. Han, "A data fusion

- detection method for fish freshness based on computer vision and near-infrared spectroscopy,” *Analytical Methods*, vol. 8, no. 14, pp. 2929–2935, 2016.
- [127] D. Moshou, C. Bravo, R. Oberti, J. West, L. Bodria, A. McCartney, and H. Ramon, “Plant disease detection based on data fusion of hyperspectral and multi-spectral fluorescence imaging using kohonen maps,” *Real-Time Imaging*, vol. 11, no. 2, pp. 75–83, 2005.
- [128] A. Khanum, A. Alvi, and R. Mehmood, “Towards a semantically enriched computational intelligence (seci) framework for smart farming,” in *International Conference on Smart Cities, Infrastructure, Technologies and Applications*. Springer, 2017, pp. 247–257.
- [129] A. Sato, R. Huang, and N. Y. Yen, “Design of fusion technique-based mining engine for smart business,” *Human-centric Computing and Information Sciences*, vol. 5, no. 1, p. 23, 2015.
- [130] Z. Pang, Q. Chen, W. Han, and L. Zheng, “Value-centric design of the internet-of-things solution for food supply chain: Value creation, sensor portfolio and information fusion,” *Information Systems Frontiers*, vol. 17, no. 2, pp. 289–319, Apr. 2015.
- [131] S. K. Viswanath, C. Yuen, X. Ku, and X. Liu, “Smart tourist-passive mobility tracking through mobile application,” in *International Internet of Things Summit*. Springer, 2014, pp. 183–191.
- [132] K. Lin, A. Kansal, D. Lymberopoulos, and F. Zhao, “Energy-accuracy trade-off for continuous mobile device location,” in *2010 Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*. ACM, 2010, pp. 285–298.
- [133] E. Wilhelm, S. Siby, Y. Zhou, X. J. S. Ashok, M. Jayasuriya, S. Foong, J. Kee, K. L. Wood, and N. O. Tippenhauer, “Wearable environmental sensors and infrastructure for mobile large-scale urban deployment,” *IEEE Sensors Journal*, vol. 16, no. 22, pp. 8111–8123, 2016.
- [134] R. Liu, C. Yuen, T.-N. Do, and U.-X. Tan, “Fusing similarity-based sequence and dead reckoning for indoor positioning without training,” *IEEE Sensors Journal*, vol. 17, no. 13, pp. 4197–4207, 2017.
- [135] R. Liu, C. Yuen, T.-N. Do, D. Jiao, X. Liu, and U.-X. Tan, “Cooperative relative positioning of mobile users by fusing imu inertial and ubw ranging information,” in *2017 IEEE International Conference on Robotics and Automation*. IEEE, 2017, pp. 5623–5629.
- [136] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik, “Dynamic route planning with real-time traffic predictions,” *Information Systems*, vol. 64, pp. 258–265, 2017.
- [137] T.-A. Teo and K.-H. Cho, “Bim-oriented indoor network model for indoor and outdoor combined route planning,” *Advanced Engineering Informatics*, vol. 30, no. 3, pp. 268–282, 2016.
- [138] Y. Yoshimura, S. Sobolevsky, C. Ratti, F. Girardin, J. P. Carrascal, J. Blat, and R. Sinatra, “An analysis of visitors’ behavior in the louvre museum: A study using bluetooth data,” *Environment and Planning B: Planning and Design*, vol. 41, no. 6, pp. 1113–1131, 2014.
- [139] H. Poonawala, V. Kolar, S. Blandin, L. Wynter, and S. Sahu, “Singapore in motion: Insights on public transport service level through farecard and mobile data analytics,” in *2016 ACM SIGKDD Proceedings of the 22nd International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 589–598.
- [140] Q. Li, L. Chen, M. Li, S.-L. Shaw, and A. Nüchter, “A sensor-fusion drivable-region and lane-detection system for autonomous vehicle navigation in challenging road scenarios,” *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 540–555, 2014.
- [141] C. Huang, L. Wang, and L. Lai, “Data-driven short-term solar irradiance forecasting based on information of neighboring sites,” *IEEE Transactions on Industrial Electronics*, 2018.
- [142] S. Abeywickrama, L. Jayasinghe, H. Fu, S. Nissanka, and C. Yuen, “RF-Based Direction Finding of UAVs Using DNN,” in *2018 IEEE International Conference on Communication Systems*, 2018, pp. 157–161.
- [143] R. Salpietro, L. Bedogni, M. Di Felice, and L. Bononi, “Park here! a smart parking system based on smartphones’ embedded sensors and short range communication technologies,” in *2015 IEEE 2nd World Forum on Internet of Things*. IEEE, 2015, pp. 18–23.
- [144] J. H. Lee, “Smart health: Concepts and status of ubiquitous health with smartphone,” *2011 International Conference on ICT Convergence*, pp. 388–389, 2011.
- [145] T. Muhammed, R. Mehmood, A. Albeshri, and I. Katib, “Ubehealth: A personalized ubiquitous cloud and edge-enabled networked healthcare system for smart cities,” *IEEE Access*, vol. 6, pp. 32258–32285, 2018.
- [146] N. Noury, “A smart sensor for the remote follow up of activity and fall detection of the elderly,” in *IEEE-EMBS Special Topic 2nd Annual International Conference on Microtechnologies in Medicine & Biology*. IEEE, 2002, pp. 314–317.
- [147] H. Lee, K. Park, B. Lee, J. Choi, and R. Elmasri, “Issues in data fusion for healthcare monitoring,” in *2008 Proceedings of the 1st International Conference on Pervasive Technologies Related to Assistive Environments*. ACM, 2008, p. 3.
- [148] L. Jiang, D.-Y. Liu, and B. Yang, “Smart home research,” in *2004 International Conference on Machine Learning and Cybernetics*, vol. 2. IEEE, 2004, pp. 659–663.
- [149] O. Brdiczka, J. L. Crowley, and P. Reignier, “Learning situation models in a smart home,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 1, pp. 56–63, 2009.
- [150] E. Fernandes, J. Jung, and A. Prakash, “Security analysis of emerging smart home applications,” in *2016 IEEE symposium on security and privacy (SP)*. IEEE, 2016, pp. 636–654.
- [151] N. Komninos, E. Philippou, and A. Pitsillides, “Survey in smart grid and smart home security: Issues, challenges and countermeasures,” *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1933–1954, 2014.
- [152] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, “Blockchain for iot security and privacy: The case study of a smart home,” in *2017 IEEE international conference on pervasive computing and communications workshops (PerCom workshops)*. IEEE, 2017, pp. 618–623.
- [153] S. D. S. U. I. C. for Communications and C. D. of Transportation, *Smart Communities Guidebook: Building Smart Communities, how California’s Communities Can Thrive in the Digital Age*. International Center for Communications, College of Professional Studies and Fine Arts, San Diego State University, 1997.
- [154] H. Lindskog, “Smart communities initiatives,” in *Proceedings of the 3rd ISONeWorld Conference*, vol. 16, 2004, pp. 14–16.
- [155] F. Liang, V. Das, N. Kostyuk, and M. M. Hussain, “Constructing a data-driven society: China’s social credit system as a state surveillance infrastructure,” *Policy & Internet*, 2018.
- [156] A. Cheung and Y. Chen, “The rise of the data state: Chinas social credit system,” in *Emerging Technologies and the Future of Citizenship Workshop*. Berlin Social Science Centre., 2018.
- [157] Z. Alazawi, O. Alani, M. B. Abdjljbar, S. Altowajjri, and R. Mehmood, “A smart disaster management system for future cities,” in *Proceedings of the 2014 ACM international workshop on wireless and mobile technologies for smart cities*. ACM, 2014, pp. 1–10.
- [158] D. Jin, C. Hannon, Z. Li, P. Cortes, S. Ramaraju, P. Burgess, N. Buch, and M. Shahidehpour, “Smart street lighting system: A platform for innovative smart city applications and a new frontier for cyber-security,” *The Electricity Journal*, vol. 29, no. 10, pp. 28–35, 2016.
- [159] “The City of Oslo: Oslo Smart City Strategy,” 2018. [Online]. Available: <https://www.oslo.kommune.no/\protect\unhbox\voidb@x\hbox{english/politics-and-administration/smart-oslo/smart-oslo-strategy/}>
- [160] G. Sohn and I. Dowman, “Data fusion of high-resolution satellite imagery and lidar data for automatic building extraction,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 1, pp. 43–63, 2007.
- [161] P. Xu, F. Davoine, J.-B. Bordes, H. Zhao, and T. Dencœur, “Multimodal information fusion for urban scene understanding,” *Machine Vision and Applications*, vol. 27, no. 3, pp. 331–349, 2016.
- [162] M. Q. Raza and A. Khosravi, “A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings,” *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352–1372, 2015.
- [163] R. Baetens, B. P. Jelle, and A. Gustavsen, “Properties, requirements and possibilities of smart windows for dynamic daylight and solar energy control in buildings: A state-of-the-art review,” *Solar Energy Materials and Solar Cells*, vol. 94, no. 2, pp. 87–105, 2010.
- [164] W.-T. Li, K. Thirugnanam, W. Tushar, C. Yuen, and K. L. Wood, “Optimizing energy consumption of hot water system in buildings with solar thermal systems,” in *SMARTGREENS*, 2017, pp. 266–273.
- [165] E. McKenna, M. Krawczynski, and M. Thomson, “Four-state domestic building occupancy model for energy demand simulations,” *Energy and Buildings*, vol. 96, pp. 30–39, 2015.
- [166] H. Chen, P. Chou, S. Duri, H. Lei, and J. Reason, “The design and

- implementation of a smart building control system,” in *2009 IEEE International Conference on E-Business Engineering*. IEEE, 2009, pp. 255–262.
- [167] G. Cardone, A. Cirri, A. Corradi, and L. Foschini, “The participact mobile crowd sensing living lab: The testbed for smart cities,” *IEEE Communications Magazine*, vol. 52, no. 10, pp. 78–85, 2014.
- [168] A. Antonić, V. Bilas, M. Marjanović, M. Matijašević, D. Oletić, M. Pavelić, I. P. Žarko, K. Pripužić, and L. Skorin-Kapov, “Urban crowd sensing demonstrator: Sense the zagreb air,” in *International Conference on Software, Telecommunications and Computer Networks*. IEEE, 2014, pp. 423–424.
- [169] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, “A new look at ihs-like image fusion methods,” *Information fusion*, vol. 2, no. 3, pp. 177–186, 2001.
- [170] M. Wu, C. Wu, W. Huang, Z. Niu, C. Wang, W. Li, and P. Hao, “An improved high spatial and temporal data fusion approach for combining landsat and modis data to generate daily synthetic landsat imagery,” *Information Fusion*, vol. 31, pp. 14–25, 2016.
- [171] Y. Zeng, W. Huang, M. Liu, H. Zhang, and B. Zou, “Fusion of satellite images in urban area: Assessing the quality of resulting images,” in *2010 18th International Conference on Geoinformatics*. IEEE, 2010, pp. 1–4.
- [172] S. Wang, J. Wan, D. Zhang, D. Li, and C. Zhang, “Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination,” *Computer Networks*, vol. 101, pp. 158–168, 2016.
- [173] C. Gröger, F. Niedermann, and B. Mitschang, “Data mining-driven manufacturing process optimization,” in *Proceedings of the World Congress on Engineering*, vol. 3, 2012, pp. 4–6.
- [174] J. Lee, “E-manufacturing – fundamental, tools, and transformation,” *Robotics and Computer-Integrated Manufacturing*, vol. 19, no. 6, pp. 501–507, 2003.
- [175] G. Niu, B.-S. Yang, and M. Pecht, “Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance,” *Reliability Engineering & System Safety*, vol. 95, no. 7, pp. 786–796, 2010.
- [176] B. Schmidt and L. Wang, “Cloud-enhanced predictive maintenance,” *The International Journal of Advanced Manufacturing Technology*, vol. 99, no. 1-4, pp. 5–13, 2018.
- [177] S. Wolfert, L. Ge, C. Verdouw, and M.-J. Bogaardt, “Big data in smart farming—a review,” *Agricultural Systems*, vol. 153, pp. 69–80, 2017.
- [178] A. Walter, R. Finger, R. Huber, and N. Buchmann, “Opinion: Smart farming is key to developing sustainable agriculture,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 24, pp. 6148–6150, 2017.
- [179] D. De Benedetto, A. Castrignano, M. Diacono, M. Rinaldi, S. Ruggieri, and R. Tamborrino, “Field partition by proximal and remote sensing data fusion,” *Biosystems engineering*, vol. 114, no. 4, pp. 372–383, 2013.
- [180] Atlas, “Free-range fish farming - aquapod,” 2018. [Online]. Available: <https://atlasofthefuture.org/project/aquapod-fish-farm/>
- [181] M. R. Hassan, B. Nath, and M. Kirley, “A fusion model of hmm, ann and ga for stock market forecasting,” *Expert Systems with Applications*, vol. 33, no. 1, pp. 171–180, 2007.
- [182] M. Christopher, *Logistics & supply chain management*. Pearson UK, 2016.
- [183] B. W. Parkinson, P. Enge, P. Axelrad, and J. J. Spilker Jr, *Global positioning system: Theory and applications, Volume II*. American Institute of Aeronautics and Astronautics, 1996.
- [184] P. Misra and P. Enge, *Global Positioning System: signals, measurements and performance second edition*. Massachusetts: Ganga-Jamuna Press, 2006.
- [185] A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, R. Raulefs, and E. Aboutanios, “Recent advances in indoor localization: A survey on theoretical approaches and applications,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2016.
- [186] Z. B. Tariq, D. M. Cheema, M. Z. Kamran, and I. H. Naqvi, “Non-gps positioning systems: a survey,” *ACM Computing Surveys (CSUR)*, vol. 50, no. 4, pp. 57:1–57:34, 2017.
- [187] J. Schlingensiepen, F. Nemtanu, R. Mehmood, and L. McCluskey, “Autonomic transport management systems enabler for smart cities, personalized medicine, participation and industry grid/industry 4.0,” in *Intelligent transportation systems—problems and perspectives*. Springer, 2016, pp. 3–35.
- [188] D. Dellinger, M. Goldszmidt, A. V. Goldberg, J. Krumm, and R. F. F. Werneck, “Controlling travel route planning module based upon user travel preference,” Apr. 4 2017, uS Patent 9,612,128.
- [189] D. Han, S. Jung, M. Lee, and G. Yoon, “Building a practical wi-fi-based indoor navigation system,” *IEEE Pervasive Computing*, vol. 13, no. 2, pp. 72–79, 2014.
- [190] Y. Arfat, R. Mehmood, and A. Albeshri, “Parallel shortest path graph computations of united states road network data on apache spark,” in *International Conference on Smart Cities, Infrastructure, Technologies and Applications*. Springer, 2017, pp. 323–336.
- [191] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, “Understanding individual human mobility patterns,” *Nature*, vol. 453, no. 7196, p. 779, 2008.
- [192] S. Jiang, J. Ferreira, and M. C. González, “Activity-based human mobility patterns inferred from mobile phone data: A case study of singapore,” *IEEE Transactions on Big Data*, vol. 3, no. 2, pp. 208–219, 2017.
- [193] T. S. Prentow, A. J. Ruiz-Ruiz, H. Blunck, A. Stisen, and M. B. Kjær-gaard, “Spatio-temporal facility utilization analysis from exhaustive wifi monitoring,” *Pervasive and Mobile Computing*, vol. 16, pp. 305–316, 2015.
- [194] M. G. Demissie, S. Phithakkitnukoon, T. Sukhivibul, F. Antunes, R. Gomes, and C. Bento, “Inferring passenger travel demand to improve urban mobility in developing countries using cell phone data: a case study of senegal,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2466–2478, 2016.
- [195] M. W. Horner and M. E. O’Kelly, “Embedding economies of scale concepts for hub network design,” *Journal of Transport Geography*, vol. 9, no. 4, pp. 255–265, 2001.
- [196] D. Karamshuk, C. Boldrini, M. Conti, and A. Passarella, “Human mobility models for opportunistic networks,” *IEEE Communications Magazine*, vol. 49, no. 12, pp. 157–165, 2011.
- [197] D. Zhang, J. Huang, Y. Li, F. Zhang, C. Xu, and T. He, “Exploring human mobility with multi-source data at extremely large metropolitan scales,” in *2014 Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*. ACM, 2014, pp. 201–212.
- [198] D. Zhang, J. Zhao, F. Zhang, and T. He, “coMobile: Real-time human mobility modeling at urban scale using multi-view learning,” in *2015 SIGSPATIAL Proceedings of the 23rd International Conference on Advances in Geographic Information Systems*, ACM, 2015, pp. 40:1–40:10.
- [199] E. Alomari and R. Mehmood, “Analysis of tweets in arabic language for detection of road traffic conditions,” in *International Conference on Smart Cities, Infrastructure, Technologies and Applications*. Springer, 2017, pp. 98–110.
- [200] H. Wei, G. Zheng, H. Yao, and Z. Li, “Intellilight: A reinforcement learning approach for intelligent traffic light control,” in *2018 ACM SIGKDD Proceedings of the 24th International Conference on Knowledge Discovery & Data Mining*. ACM, 2018, pp. 2496–2505.
- [201] B. Yao, P. Hu, X. Lu, J. Gao, and M. Zhang, “Transit network design based on travel time reliability,” *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 233–248, 2014.
- [202] M. A. Munizaga and C. Palma, “Estimation of a disaggregate multi-modal public transport origin–destination matrix from passive smart-card data from santiago, chile,” *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 9–18, 2012.
- [203] R. Mehmood, R. Meriton, G. Graham, P. Hennelly, and M. Kumar, “Exploring the influence of big data on city transport operations: a markovian approach,” *International Journal of Operations & Production Management*, vol. 37, no. 1, pp. 75–104, 2017.
- [204] S. A. Shaheen, H. Zhang, E. Martin, and S. Guzman, “China’s hangzhou public bicycle: understanding early adoption and behavioral response to bikesharing,” *Transportation Research Record*, vol. 2247, no. 1, pp. 33–41, 2011.
- [205] J. Schuijbroek, R. C. Hampshire, and W.-J. Van Hove, “Inventory rebalancing and vehicle routing in bike sharing systems,” *European Journal of Operational Research*, vol. 257, no. 3, pp. 992–1004, 2017.
- [206] P. Falcone, F. Borrelli, J. Asgari, H. E. Tseng, and D. Hrovat, “Predictive active steering control for autonomous vehicle systems,” *IEEE Transactions on Control Systems Technology*, vol. 15, no. 3, pp.

- 566–580, 2007.
- [207] A. Ferdowsi, U. Challita, W. Saad, and N. B. Mandayam, “Robust deep reinforcement learning for security and safety in autonomous vehicle systems,” in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 307–312.
- [208] J. Gao, Y. Xiao, J. Liu, W. Liang, and C. P. Chen, “A survey of communication/networking in smart grids,” *Future Generation Computer Systems*, vol. 28, no. 2, pp. 391–404, 2012.
- [209] M. Kordestani and M. Saif, “Data fusion for fault diagnosis in smart grid power systems,” in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 2017, pp. 1–6.
- [210] K. Thirugnanam, S. K. Kerk, C. Yuen, N. Liu, and M. Zhang, “Energy management for renewable microgrid in reducing diesel generators usage with multiple types of battery,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 8, pp. 6772–6786, 2018.
- [211] W. Tushar, C. Yuen, B. Chai, S. Huang, K. L. Wood, S. G. Kerk, and Z. Yang, “Smart grid testbed for demand focused energy management in end user environments,” *IEEE Wireless Communications*, vol. 23, no. 6, pp. 70–80, 2016.
- [212] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, “State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid,” *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 33–43, 2012.
- [213] T. Liu, Y. Sun, Y. Liu, Y. Gui, Y. Zhao, D. Wang, and C. Shen, “Abnormal traffic-indexed state estimation: A cyber–physical fusion approach for smart grid attack detection,” *Future Generation Computer Systems*, vol. 49, pp. 94–103, 2015.
- [214] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, “A multi-sensor energy theft detection framework for advanced metering infrastructures,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1319–1330, 2013.
- [215] G. Sideratos and N. D. Hatziaargyriou, “An advanced statistical method for wind power forecasting,” *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 258–265, 2007.
- [216] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, “Current methods and advances in forecasting of wind power generation,” *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012.
- [217] J. Jung and R. P. Broadwater, “Current status and future advances for wind speed and power forecasting,” *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 762–777, 2014.
- [218] N. L. D. Khoa, A. Anaissi, and Y. Wang, “Smart infrastructure maintenance using incremental tensor analysis,” in *2017 ACM Proceedings of the Conference on Information and Knowledge Management*. ACM, 2017, pp. 959–967.
- [219] T. Cioara, I. Anghel, I. Salomie, M. Antal, C. Pop, M. Bertoncini, D. Arnone, and F. Pop, “Exploiting data centres energy flexibility in smart cities: Business scenarios,” *Information Sciences*, vol. 476, pp. 392–412, 2019.
- [220] Y. Li, Y. Zhang, K. Luo, T. Jiang, Z. Li, and W. Peng, “Ultra-dense hetnets meet big data: Green frameworks, techniques, and approaches,” *IEEE Communications Magazine*, vol. 56, no. 6, pp. 56–63, 2018.
- [221] F. Kong and X. Liu, “A survey on green-energy-aware power management for datacenters,” *ACM Computing Surveys (CSUR)*, vol. 47, no. 2, pp. 30:1–30:38, 2015.
- [222] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez Jr, “Millimeter wave mobile communications for 5g cellular: It will work!” *IEEE access*, vol. 1, no. 1, pp. 335–349, 2013.
- [223] R. Mahapatra, Y. Nijssure, G. Kaddoum, N. U. Hassan, and C. Yuen, “Energy efficiency tradeoff mechanism towards wireless green communication: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 686–705, 2016.
- [224] S. Wang, X. Zhang, Y. Zhang, L. Wang, J. Yang, and W. Wang, “A survey on mobile edge networks: Convergence of computing, caching and communications,” *IEEE Access*, vol. 5, pp. 6757–6779, 2017.
- [225] H. Ma, D. Zhao, and P. Yuan, “Opportunities in mobile crowd sensing,” *IEEE Communications Magazine*, vol. 52, no. 8, pp. 29–35, 2014.
- [226] X. Wang, Y. Sui, J. Wang, C. Yuen, and W. Wu, “A distributed truthful auction mechanism for task allocation in mobile cloud computing,” *IEEE Transactions on Services Computing*, pp. 1–1, 2018.
- [227] D. N. Jha, S. Garg, P. P. Jayaraman, R. Buyya, Z. Li, and R. Ranjan, “A holistic evaluation of docker containers for interfering microservices,” in *2018 IEEE International Conference on Services Computing (SCC)*. IEEE, 2018, pp. 33–40.
- [228] S. Bi, C. K. Ho, and R. Zhang, “Wireless powered communication: Opportunities and challenges,” *IEEE Communications Magazine*, vol. 53, no. 4, pp. 117–125, 2015.
- [229] T. Sekitani, M. Takamiya, Y. Noguchi, S. Nakano, Y. Kato, T. Sakurai, and T. Someya, “A large-area wireless power-transmission sheet using printed organic transistors and plastic MEMS switches,” *Nature Materials*, vol. 6, no. 6, p. 413, 2007.
- [230] B. Ubaldi, “Open government data,” *Open government data 2013*, 2013.
- [231] J. M. Cantera and R. Lewis, “Delivery context ontology,” *W3C Working Group Note*, 2010.
- [232] G. Antoniou and F. Van Harmelen, “Web ontology language: Owl,” in *Handbook on Ontologies*. Springer, 2004, pp. 67–92.
- [233] D. Brickley, “Resource description framework (rdf) schema specification 1.0,” <http://www.w3.org/TR/rdf-schema>, 2000.
- [234] H. Neuhaus and M. Compton, “The semantic sensor network ontology,” in *AGILE Workshop on Challenges in Geospatial Data Harmonisation, Hannover, Germany*, 2009, pp. 1–33.
- [235] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, “Dbpedia: A nucleus for a web of open data,” in *The Semantic Web*. Springer, 2007, pp. 722–735.
- [236] M. Organization, “Message Queuing Telemetry Transport,” 2018.
- [237] A. Martínez-Ballesté, P. A. Pérez-Martínez, and A. Solanas, “The pursuit of citizens’ privacy: a privacy-aware smart city is possible,” *IEEE Communications Magazine*, vol. 51, no. 6, pp. 136–141, 2013.
- [238] Y. Li, W. Dai, Z. Ming, and M. Qiu, “Privacy protection for preventing data over-collection in smart city,” *IEEE Transactions on Computers*, vol. 65, no. 5, pp. 1339–1350, 2016.
- [239] C. Tankard, “What the gdpr means for businesses,” *Network Security*, vol. 2016, no. 6, pp. 5–8, 2016.
- [240] J. P. Albrecht, “How the gdpr will change the world,” *Eur. Data Prot. L. Rev.*, vol. 2, p. 287, 2016.
- [241] C. Cadwalladr and E. Graham-Harrison, “Revealed: 50 million facebook profiles harvested for cambridge analytica in major data breach,” *The Guardian*, vol. 17, 2018.
- [242] W. Ding, X. Jing, Z. Yan, and L. T. Yang, “A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion,” *Information Fusion*, 2018.
- [243] B. K. Beaulieu-Jones, Z. S. Wu, C. Williams, and C. S. Greene, “Privacy-preserving generative deep neural networks support clinical data sharing,” *BioRxiv*, p. 159756, 2017.
- [244] C. Esteban, S. L. Hyland, and G. Rätsch, “Real-valued (medical) time series generation with recurrent conditional GANs,” *arXiv preprint arXiv:1706.02633*, 2017.
- [245] R. Kitchin, “Getting smarter about smart cities: Improving data privacy and data security,” *Data Protection Unit, Department of the Taoiseach, Dublin, Ireland.*, 2016.
- [246] S. Chakrabarty and D. W. Engels, “A secure iot architecture for smart cities,” in *2016 13th IEEE annual consumer communications & networking conference (CCNC)*. IEEE, 2016, pp. 812–813.
- [247] B. Mocanu, F. Pop, A. Mihaita, C. Dobre, and A. Castiglione, “Data fusion technique in spider peer-to-peer networks in smart cities for security enhancements,” *Information Sciences*, vol. 479, pp. 607–621, 2019.
- [248] A. Talaş, F. Pop, and G. Neagu, “Elastic stack in action for smart cities: Making sense of big data,” in *2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*. IEEE, 2017, pp. 469–476.
- [249] X. Wang, J. Zhang, E. M. Schooler, and M. Ion, “Performance evaluation of attribute-based encryption: Toward data privacy in the iot,” in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 725–730.
- [250] M. Singh, M. Rajan, V. Shivraj, and P. Balamuralidhar, “Secure mqtt for internet of things (iot),” in *2015 Fifth International Conference on Communication Systems and Network Technologies*. IEEE, 2015, pp. 746–751.
- [251] M. Elhoseny, G. Ramírez-González, O. M. Abu-Elnasr, S. A. Shawkat, N. Arunkumar, and A. Farouk, “Secure medical data transmission

- model for iot-based healthcare systems,” *IEEE Access*, vol. 6, pp. 20 596–20 608, 2018.
- [252] M. David, E. Tom, B. Paul, and T. Stephanie, “World’s biggest data breaches & hacks,” 2019. [Online]. Available: <https://informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/>
- [253] A. Bartoli, J. Hernández-Serrano, M. Soriano, M. Dohler, A. Kountouris, and D. Barthel, “Security and privacy in your smart city,” in *Proceedings of the Barcelona smart cities congress*, vol. 292, 2011, pp. 1–3.
- [254] R. Bellman, *Dynamic programming*. Courier Corporation, 2013.
- [255] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, “A survey on deep learning for big data,” *Information Fusion*, vol. 42, pp. 146–157, 2018.
- [256] R. Miikkulainen, J. Liang, E. Meyerson, A. Rawal, D. Fink, O. Francon, B. Raju, H. Shahrzad, A. Navruzyan, N. Duffy *et al.*, “Evolving deep neural networks,” in *Artificial Intelligence in the Age of Neural Networks and Brain Computing*. Elsevier, 2019, pp. 293–312.
- [257] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, “A survey of deep neural network architectures and their applications,” *Neurocomputing*, vol. 234, pp. 11–26, 2017.
- [258] D. Gunning, “Explainable artificial intelligence (xai),” *Defense Advanced Research Projects Agency (DARPA)*, 2017.
- [259] A. Kendall and Y. Gal, “What uncertainties do we need in bayesian deep learning for computer vision?” in *Advances in Neural Information Processing Systems*, 2017, pp. 5574–5584.
- [260] C.-C. J. Kuo, M. Zhang, S. Li, J. Duan, and Y. Chen, “Interpretable convolutional neural networks via feedforward design,” *Journal of Visual Communication and Image Representation*, vol. 60, pp. 346 – 359, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S104732031930104X>
- [261] Q. Da, Y. Yu, and Z.-H. Zhou, “Learning with augmented class by exploiting unlabeled data,” in *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014, pp. 1760–1766.
- [262] Y.-F. Li and Z.-H. Zhou, “Towards making unlabeled data never hurt,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 1, pp. 175–188, 2015.
- [263] S. Hoo-Chang, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Noguez, J. Yao, D. Mollura, and R. M. Summers, “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, p. 1285, 2016.