

Received January 13, 2020, accepted February 3, 2020, date of publication February 12, 2020, date of current version February 26, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2973461

Cyber Physical and Social Networks in IoV (CPSN-IoV): A Multimodal Architecture in Edge-Based Networks for Optimal Route Selection Using 5G Technologies

ANSIF AROOJ¹, MUHAMMAD SHOAIB FAROOQ², TARIQ UMER³, (Senior Member, IEEE), GHULAM RASOOL³, AND BO WANG⁴

¹University of Education, Division of Science and Technology, Lahore 54000, Pakistan

²Department of Computer Science, University of Management and Technology, Lahore 54000, Pakistan

³Department of Computer Science, COMSATS University Islamabad, Lahore Campus, Lahore 54000, Pakistan

⁴School of Electrical Engineering, Wuhan University, Wuhan 430070, China

Corresponding author: Bo Wang (whwdwb@whu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 51207113.

ABSTRACT Humans are blessed with the intelligence to create links, develop semantic metaphors and models for reasoning; construct rules for decision making; and to form bounded loops for interaction, socialization and knowledge sharing. But machines are inadequate with these extraordinary abilities rather, numerous algorithms and mathematical models can be used to connect physical resources with cyberspaces to control objects and, develop cognitive learning for optimal decision making. Connected users and devices in closed virtual and physical proximity give direction towards the plethora of real-world applications for physical, social and, cyber computing. Because of the increase in social media networking and 5G communication links offer real-time crowdsourcing and sensing as a complementary base for information. Proceeding this idea, in this study we have proposed Cyber-Physical and Social Networks (CPSN) for two fundamental operations in IoV (Internet of Vehicles) as CPSN-IoV; (1) to define conceptual architecture of CPSN-IoV for data-oriented network for smart infrastructure and, (2) to create the significant virtual space where the instances of smart vehicles, devices, and things will have meaningful links with the real world objects where, CPSN-IoV will evolve, emerge, compete, and collaborate with all connected objects to strengthen the decision making process. To investigate the potential impact of our proposed study, we have simulated the taxicab trajectory data of the urban city of Portugal in OMNeT++ for the in-depth understanding of road topology, connected vehicles and things, and their traffic trends; and users' social media streams in respective edge for efficient route planning. The results of simulation demonstrate that our proposed framework has the ability to achieve human-machine intellectual association for managing the smart environment.

INDEX TERMS CPSN, data fusion, IoV, IoT, knowledge discovery.

I. INTRODUCTION

A. BACKGROUND AND MOTIVATION

Online socialization is a dominant human activity over the internet for social networking, collaborating purchasing, discussion groups, recommendations and social bookmarking, etc [1] However, a major concern of these platforms are the mean of initial contact for users with each other. Because in

The associate editor coordinating the review of this manuscript and approving it for publication was Junhui Zhao¹.

the real world people's interactions are heavily influenced by the physical context in closed proximity. Moreover, the emergence of technology has encouraged ubiquitous networks that enabled computing capabilities and control for every human 24/7, remotely [2]. This is also evident that the interaction between peers can often happen due to shared geolocation. For example, people visiting the same university, museum, shopping mall, sports event, or people travelling by the same bus, train and, flight. These series of geo and social interactions and the sharing of decisions and recommendations,

lead towards the emergence of two computational technologies, Cyber-Physical Networks (CPN) with Cyber Social Networks (CSN) [3]–[5] as the next generation of conventional internet of things and, termed as Cyber-Physical–Social Networks (CPSN).

The term CPS was initially introduced by [6], and since the time of initialization, it has become a famous area of research and development in many real-world applications. The virtual world is linked with different types of physical devices and, has the capability to build intelligent systems with cognitive abilities. And the variety of sensory devices monitor real-time activity and the behaviour of physical objects and their observations are continuously stimulated and moved towards the cyber world. The virtual space where these resultant streams are analyzed and interpreted to deduce the state of the physical object and to identify their digital representations in order to derive knowledge for optimal control [7], [8]. Similarly, CSS emphasizes human behaviour and interaction analysis by using data analytics and mining techniques on social media posts [9], [10]. Historically, CPSN evolution is based on mechatronics systems, which is a combination of electrical, mechanical and control engineering [11]–[13]. Well-defined processes are designed to support procedures into the real-world physical system [11], [12]. The next generation of CPSN was embedded systems, where applications are installed in a physical system for managing processes. The successor of embedded solutions is CPSN frameworks with extraordinary computing and control capabilities. These networks are concerned with sensing ability from the physical world by using extraordinary communication abilities for control and management [14].

The concept of CPSN is related to IoT; some researchers have views about the distinctions between two of the famous terms and only assert as the bridge between physical and virtual/cyber world and declared IoT as the communication identifier of all smart objects in the network [15]. On the other hand, some authors believe that both terms are similar in nature, i.e., IoT provides the horizontal prospect consisting of hardware components with interconnection capabilities whereas, CPSN has a vertical view containing computational and control processes [8]. In recent studies, the term human-in-the-loop (HiTL) is associated with CPSN which recognizes the importance of human physiological and psychological studies with objectives to control objects in the real world by using brain-computer solutions with cognitive abilities [8], [16]. Finally, CPSN is a set of several physical, cyber and, social integration by using data fusion techniques for creating intelligent systems to facilitate humans. Moreover, smart and connected objects with self-configuration capabilities have enabled the convergence of physical, cyber and, social domains for secure information exchange, data processing, cognitive learning in many real-world applications. The confluence of CPSN have a powerful impact by extensive cognitive learning algorithms and computational processes, in order to provide effective process control, management and decision support [17], [6], [5], [3].

The idea of CPSN comprises of both the physical and virtual context of users, users at the same geolocation have some virtual context and existence at respective edge and cloud. Furthermore, IEEE 802.11p has transfigured the IoV connections as well and this advancement in information and communication technologies have encouraged us to develop a CPSN based architecture specifically in the domain of IoV as CPSN-IoV. Thus, our proposed term CPSN-IoV is the deep interplay of physical and smart objects, their association and, rule-based reasoning for collective and cognitive intelligence. It also has the ability to fuse diverse information that is initializing and observing from the physical world. Cognitive intelligence and CPSN-IoV technologies can provide human-centric services to cumbersome chores and tasks in physical worlds as well as in social worlds. But in smart cities, open access city-data by distributed sensors and the social media data of citizens can offer real-time, heterogeneous and large scale observation and sense for effective and intelligent services by intelligence mining [18]. These potential capabilities have encouraged different research dimensions in IoV, for example, collision detection, avoidance, precision, operation and resource management as some of the major research attractions. Furthermore, Big data analytics for knowledge extraction in CPSN-IoV networks can facilitate urban dynamics for decision making [19]. For instance, traffic signal adjustments can be managed by vehicle arrival data with intelligent cyber analytics [20], [21]. Event detection and management can easily be dealt with user preferences and road ambient environmental constraints. [22]–[24].

Application Scenario: But before getting into details let's consider a real-time application scenario of CPSN-IoV. Suppose we are travelling in a car and are trying to reach as early as possible. Finding the appropriate driving direction is a key concern for every driver and his route selection consists of a number of parameters. For instance, traffic and road condition, no of road signals, the capacity of road, rush hour, and the number of turns. Google, Open-Street and Bing maps, etc. provide navigation and thus not only save time but reduce energy consumption as well. Higher gas consumption and huge time wastage can cause fatigue during traffic jams. Therefore, there is a need for a system where users and government can get ease from traffic problems, in addition to planning better traffic management, as just navigation alone is not enough to achieve these extraordinary capabilities, thus, the term CPSN-IoV can be defined as:

Definition 1: Cyber-physical and social network in IoV (CPSN-IoV) is the next generation of vehicular networks where every connected vehicle, user and device has virtual existence in a respective edge. This virtual location is like a web, where the drivers and cars cannot only navigate but can also communicate and learn from other nodes and achieve potential benefits in closed proximity. (figure no 1)

In this study, we have proposed edge and cloud-based multimodal architecture as CPSN-IoV, where all aspects of real-time navigation can strengthen by identifying related patterns and associations using mining and aggregation of

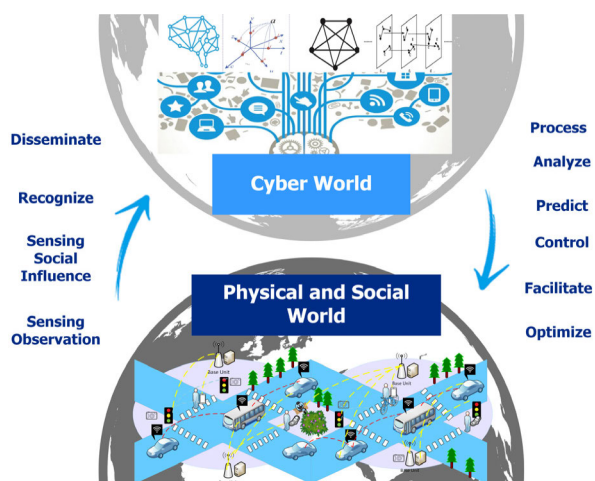


FIGURE 1. Cyber Physical Social Network for Internet of Vehicles (CPSN-IoV).

data with respect to human physical and social behaviour. Moreover, a case study and simulation in OMNET++ have been developed and tested in order to test the strength of the proposed architecture. Specifically, for the test case of this study, we have selected optimal route selection during navigation. We have examined how sensory observation from vehicular sensory observation, driver's social media stream, logs of traffic directions, and smartphone data etc. can collectively work together to get optimal route during travel.

B. HISTORICAL NOTES AND RELATED WORK

As described earlier, the development of technical, physical and social networks, the idea of remote observation and controlling objects came into existence. IoT, cloud computing, social networks, big data and smart devices work together in real-world activities. From the inception of the CPS in the year 2000 [6], it has become a hot research and development (R&D) area of interest in several disciplines and application. Moreover, many reports by the President's Council of Advisors on Science and Technology (PCAST) in the US, have described CPSN as a national interest [25] and a report by McKinsey Global Institute has mentioned CPSN as a global revolution paradigm by 2025 in life, industry and, economy [26]. Real-world objects will directly pertinent with cyber objects using robotics, cloud/edge computing, and ubiquity. IMC-AESOP and partner's project as 'Industry 4.0' is an effort to inspire the industry to strengthen conventional systems by using CPSN capabilities. SAP, Honeywell, Microsoft, Schneider Electric, etc. are major contributors to this project. National Science Foundation (NSF) has also identified CPSN as a significant research area since 2006 and has sponsored more than 300 projects for research and development in CPSN for uniform and standardized solutions [27].

So, the framework of CPSN is tightly integrated with the physical objects and this integration intended to facilitate a wide range of state-of-the-art applications and services. CPSN are increasingly embedding in several kinds

of physical systems for intelligence, comfortability, energy efficiency and, resource management etc. Similarly, CPSN has potential to influence IoV applications as well, with many benefits as intervention, precision, operations, collaboration and coordination etc. Analogously in IoV, smart vehicles not only share navigational streams to broadcasting satellites and cyber units rather, people on vehicles also share real-time information about events on online social media platforms. These events are occurring in the surroundings of their environment, such as traffic accidents [19] and earthquakes [28], etc. This information can act as a collaborative information repository and source for smart city dynamics and social behaviours [6], [29], [30]. Collision detection and avoidance, nano-level resource management, rescue services, security, traffic control, service management etc. [11], [31]–[33] are some of the extensive applications in IoV.

This is also evident that due to tremendous growth of urbanization, the roads in big cities are getting congested and if any application can predict and suggest shortest and most appropriate path can help to manage many resources from energy management to reducing fatigue causes. Because, if bottlenecks will be early detected and information regarding the traffic issues will early disseminate to the concerns will ensure resource optimization. TOMTOM [34] and GARMIN [35] is the famous solutions in automobile navigation market, these solutions provide GPS localization and valuable feedback system for route guidance. Despite the fact that these solutions provide maximum guidance but still the rapid change in traffic patterns are needed to address. Because, routing requires complex integration of several aspects and only applying group theory on navigational patterns [36]–[38] will not be enough. Route planning and selection must update as soon as a new vehicle will enter into road segment or any other change of aspect [31].

Dynamic routing control algorithm has also been proposed by [39]–[41] to address this problem. During the literature review, we have explored different dynamic routing algorithms as optimal, heuristics and, hybrid algorithms. Optimal algorithms include Dijkstra [42] and incremental graphs [43] and, A* [44], Genetic Algorithms (GA) [45], Ant colony optimization [46], and Tabu search [47] are some of the heuristic approaches for route selection and planning. Lastly, using Dijkstra and Genetic Algorithms (GA) together form as hybrid algorithms. Moreover, Bellman Ford's Algorithm [48] and Floyd–Warshall's Algorithm [49] are also important solutions for finding the best path. Each approach has its own advantages and disadvantages as described in table no 1. Furthermore, research by [31] has proposed that there is a need for centralized and integrated architecture for load balancing and resource management in a dynamic environment. But efficiency is still the challenging issue and using sensing information will not be enough in future. Because drivers' preference and road infrastructure are also important in that [50], [51]. Recent researches by [41], [52]–[56] have also addressed the route planning and selection.

TABLE 1. Comparison of algorithms for shortest path method.

Algorithm	Advantages and features	Disadvantages
Dijkstra's Algorithm	<ul style="list-style-type: none"> • Exhaustive Search • Always retrieve optimal path • Performance Efficient • Provide Global Knowledge of network • Don't need to calculate repeatedly because the shortest distance is already calculated from source to destination. • Use the Greedy approach for searching • Simple to implement 	<ul style="list-style-type: none"> • Time-consuming for large networks • Doesn't Work with Non-negative graphs
Incremental Graphs	<ul style="list-style-type: none"> • Uses archival information for speed up path selection • Faster processing to research and finding a better path • Complexity $O(a(n))$ • Develop a priority queue • Work with static nodes 	<ul style="list-style-type: none"> • Doesn't search from heuristic data • Update priority queues periodically thus computationally expensive.
A*	<ul style="list-style-type: none"> • Uses the lower bound on the total path length • Assured to expand fewer nodes 	<ul style="list-style-type: none"> • Efficiency depends on heuristic algorithms' accuracy. • Only work with the finite nodes and fixed cost for every operation. • Complexity issues
Genetic Algorithms	<ul style="list-style-type: none"> • Faster Dynamic Convergence • Strong Versatility 	<ul style="list-style-type: none"> • Complexity • Lack of pre-mature convergence • Highly depended on the initial population
Ant Colony Optimization	<ul style="list-style-type: none"> • Inherent parallelism processing • Rapid discovery of changing nodes • Ensure convergence 	<ul style="list-style-type: none"> • Sequencing issues for random decisions • Analysis for processing is complex • Require iteration for decision making • Convergence from one to another node is uncertain
Tabu	<ul style="list-style-type: none"> • Have short, intermediate and, long term memory structure. • Provide better results with local and neighbour nodes. • Prefer to store the solution attributes 	<ul style="list-style-type: none"> • Re-initialization of the path is a time-consuming effort • Storing complete problem is difficult and insufficient
Bellman Ford's Algorithm	<ul style="list-style-type: none"> • Work with Negative Graphs • Information exchange to direct neighbours • Update information based on neighbour information • Time Complexity $O(VE)$ 	<ul style="list-style-type: none"> • Poor change management • Complex computation • Slower
Floyd-Warshall's Algorithm	<ul style="list-style-type: none"> • Better for multi-stop trips • Fast but require complex computation • Can reconstruct path • Time Complexity $O(V^3)$ 	<ul style="list-style-type: none"> • Static weight for each node • Doesn't update information during flow

The above studies have revealed that many algorithms and architectures have been applied to get the optimal route selection and planning individually, yet an appropriate way to achieve the best route by using cyber, physical and social aspects does not exist. This gap has given the rise to the demand for an inclusive solution for best route planning and selection by using CPSN-IoV.

In this work, we have focused on defining a comprehensive architecture for CPSN-IoV, with all in-depth aspects and relevant details. The novelty of our proposed solution is that apart from CPSN-IoV architecture, we have also tested a case study by simulation in OMNET++ for examining and evaluating our proposed architecture. Moreover, a detailed description

of each component of the proposed study has been proposed. Lastly, the performance evaluation of the proposed study has also been added at the end of this study.

C. PURPOSE OF THE STUDY

The purpose of this paper has three major aspects, *First*, to construct a context-aware and pervasive CPSN-IoV architecture for smart vehicles, users (drivers, passengers and, traffic authorities) and infrastructure (RSU and traffic signals, etc.) with edge and cloud capability for network management and driver's optimized services and solutions. *Second*, we analyze three remarkable aspects, cyber, physical and,

social data streams for cognition and intelligence and *third*, we exemplified an application scenario by dynamic trajectories of taxis and social media streams of drivers for optimal route selections by using CPSN-IoV.

D. ORGANIZATION OF THE PAPER

The format of this study is presented as follows. In Section 1, we explore and discuss the introduction, background, history, motivation, and intent of the study. In Section 2, we introduce the main architecture of CPSN-IoV and describe the key enabling technologies in domain. In this section, we present different layers of CPSN-IoV architecture including; data collection, communication, knowledge discovery, data fusion and, application layer for optimized solutions in real-time scenarios. In Section 3, we simulate and examine one real-time example in OMNeT++ for evaluation of our proposed architecture. In Section 4, we discuss the overall findings of the study and, in Section 5 conclude our study with proposed future research directions and, Section 6 describes challenges and criticalities of CPSN-IoV domain

II. ARCHITECTURE

A. LAYER 1: DATA COLLECTION

The major sources of CPSN-IoV data collection are the set of heterogeneous devices with inherent sensing and observing capabilities for fixed, mobile and user-contributed data. People (P), vehicle (V), and sensors (S) are the main data contributors those cover the broad field to capture passive and explicit data and can be defined as Equation (1):

$$D(\text{CPSN-IoV}) = \sum_{i=1}^n \{P_i, V_i, S_i\} \quad (1)$$

where P_i represents the set of people enclosed in the physiological spaces and can have wearable sensory devices and smartphones and have the ability to share sensory data (S_n), geospatial data (GS) and social media feed (SM). Thus can be described as Equation (2):

$$S(P_i) = \{S_n, GS, SM\} \quad (2)$$

Furthermore, the Vehicle (V_i) dataset consist of the set of all smart vehicles having inherent sensing capability by accelerometer, gyroscope, pressure, Magnetic field sensory data etc. ($I-Sen$), poisoning and navigational data (GPS) and in build camera images and videos (IMG), and can represent as Equation (3):

$$S(V_i) = \{i - Sen, GPS, IMG\} \quad (3)$$

On the one hand, both people and vehicles cover a wide range of mobile sensing and observation data in sensory and social media streams. On the other hand, sensors (S_i) are fixed and mounted devices by a regulatory authority and contribute to the fixed sensory observation and imagery data. The example of these sensors are smart meters, GPS sensors at public places, smart bins, sensors in tarmac to identify the available parking spaces, air pollution, quality and irrigation sensors to

check the water levels at specific location such as public fountains, ultrasonic and infrared detectors, night vision camera and roadside units (RSU) etc. Physical space of CPSN-IoV sense, observe, retrieve and disseminate data with-in respective edges. For understanding the domain of physiological space, let's consider a relevant example, the driver with the smartphone and wearable sensory devices collectively can sense onboard activities, for instance, the fuel consumption prediction or an air pollution alert. Thus, by mean of these readings, the driver can change the operations of the vehicle by using a smartphone or can allow applications to handle autonomously. Moreover, the driver may observe the surroundings whilst moving as many are posting about events happening nearby. Similarly, in the respective edge, people on-board may be sharing and reporting the same event and can interact with other physical objects such as vehicles, people and RSUs as described in figure no 2. This complete communication stream is continuously sharing and communicating with the respective servers in edges as well. Hence, this observed sensory data is an asset for our proposed study and it can be derived that the corresponding dataset about the same physical object, location or phenomenon may have different aspects. Consequently, it is necessary to combine all these resources for collaborative sensing and observation. This observation can ultimately be providing the food for data harvesting and knowledge discovery. This multidimensional space, with the features of human interaction and participation, lead towards social relations and characteristics. In particular, it is evident that human social relations exert influence on the cyber world.

Social interaction between different nodes in CPSN-IoV can be modelled as the undirected graph without weights can represent as $G(V_i, E_i)$ where V_i is the set of smart vehicles and E_i represents the set of corresponding edges that presents the social bound for interaction between all objects in closed edge-based networks. In equation 4, $P(N_i)$ is denoted as the probability for every V_i in the edge-based network and N represents the set of friend nodes with a degree i . Assume the probability as a power-law distribution [57], then the probable curve can be derived as Equation (4):

$$P(N_i) = \begin{cases} 0 & N < S \\ T(S, \gamma)N^{-\gamma} & S \leq N \leq i \end{cases} \quad (4)$$

where S is denoted as the minimal degree node in the edge-based network and γ is the skewness of distribution and can have positive and negative skewness in a hyperplane. $T(S, \gamma)$ represents as normalized constant and, S and γ are inversely proportional to each other. For instance, the social graph implies stronger ties between nodes of the network with larger S and minimal γ .

Furthermore, the dimensions of this hyperspace consist of cyber, physiological, physical and social space and create the complex linkage between various different aspects related to human psychology and mental space [58]. Because decisions in CPSN will ultimately be dependent on human thinking and

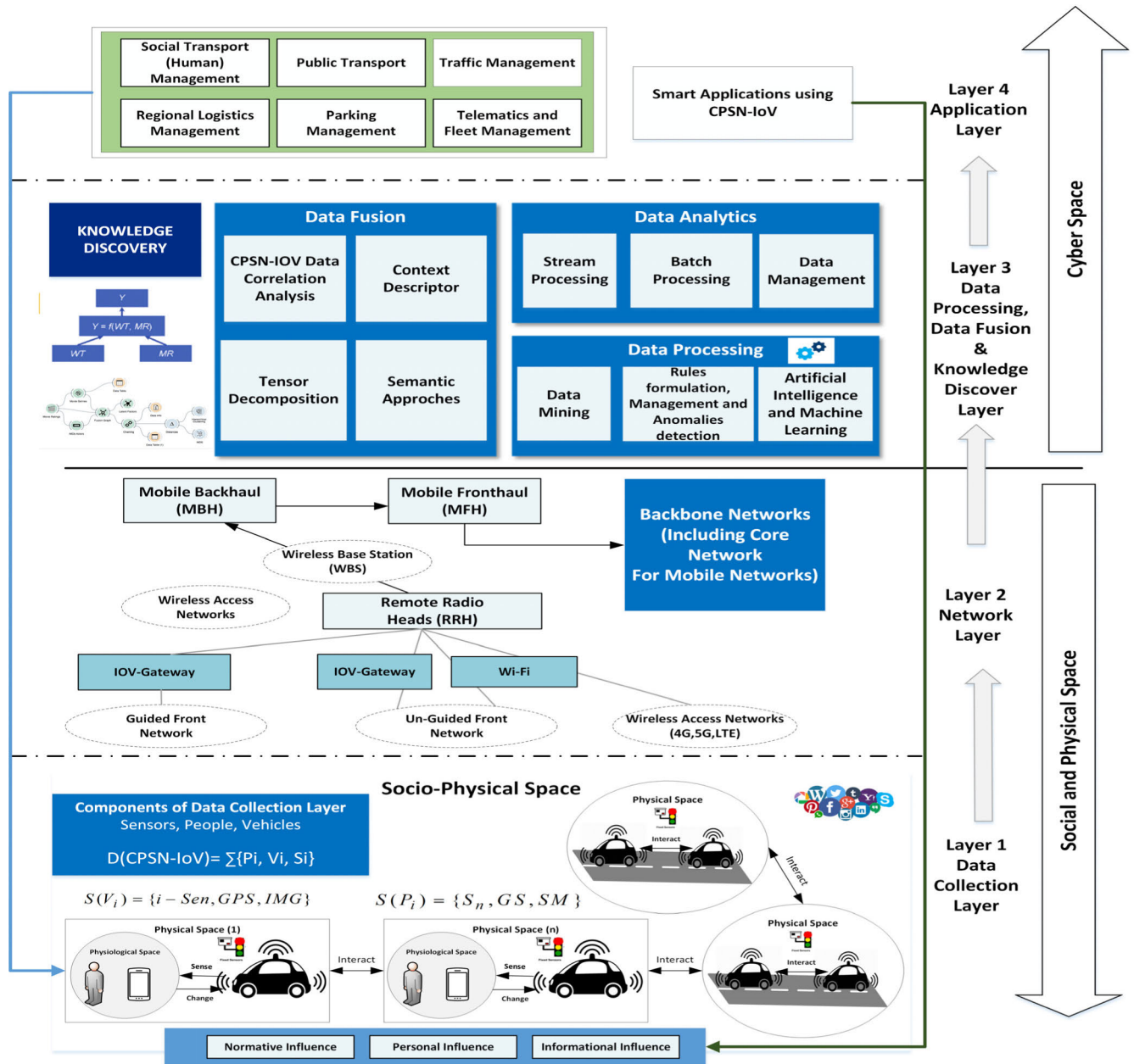


FIGURE 2. CPSN-IoV Architecture (Layer 1) Data Collection Layer: addresses the data acquisition methods from heterogeneous devices with inherent sensing and observation capabilities from fixed, mobile and user-contributed data sources. (Layer 2) Network Layer: addresses the communication technologies specifically 5G networks to disseminate heterogeneous data. (Layer 3) Data Processing, Data Fusion and Knowledge Discovery Layer: discusses the data processing, analytics, and data fusion techniques for optimized knowledge discovery and resource management. (Layer 4) Application Layer: focuses the fundamental enabling technologies in CPSN-IoV for number of real time applications.

influence their physical actions. (i) Informational, (ii) personal and (iii) normative influences define different cognitive and semantic channels during communication. (i) *informational influence*: analyzes real-world information and choice as the best option. A major concern of this influence is “what is best and right?” but this decision-making process is amazingly fragile [59]. For instance, we are travelling by a smart car and the engine fault light is switched on, this information will lead us to stop the car immediately. Thus the

decision was reliant on the information we received. (ii) *personal Influence*: the interpersonal influence that involves forming reliable relationships and relating their decisions by mutual-approach to cope with emotional arousal. The key questions of this facet are: who should I trust and who should I like? [60]. In CPSN-IoV, during travel people would like to listen to the same songs as their friends listen and for navigation people like to use the same application as their friend or relative may have in their smart devices.

(iii) *normative influence*: is based on the group or community and being part of the community one can feel more protected and that is easy to decide what others are doing? And such kind of influences builds the means of crowdsensing and very beneficial in resource and services management in the respective edge [61]. Traffic congestion monitoring is the best example of normative influence. Where onboard sensors, mobile sensors, and fixed sensors are reporting the same thing as the people may be sharing by social media. But the conflicts may arise during the communication due to liking and disliking of individuals such as the bad reviews about any application at social media form the conflict of interest. Thus there is a need for human and cognitive intelligence in order to manage these conflicts.

B. LAYER 2: NETWORK LAYER

In today's world, the use of the guided medium is really minimal but selective connections with fixed devices such as smart street lights, speed sensors, cameras, radar sensors and laser beam sensors may require a wired medium for connections. To accommodate connections with the fixed devices in smart cities, IOV-Gateways (IOV-GW) are integrated with the guided front networks. Further, the short-range wireless networks, for instance, Ultra-Wide-Band (UWB), NFC, Bluetooth, Zigbee and, RPL, etc. are connected with the IOV-GW and WiFi through the un-guided front network.

Figure no 2 (Layer 2) represents the network layer model of CPSN-IoV for every object in IoV that have the capability to connect as V2X, where X consists of infrastructure, vehicle, mobile, and devices, etc. and short-range communication channels are helpful for managing onboard services and connection of devices nearby such as smart street light to smart vehicle. Wireless networks in CPSN-IoV can be categorized into three different groups. (i) To communicate and coordinate with the centralized nodes in the respective edge. [17], (ii) Ad-hoc and non-structured communication channels to assign roles to all nodes within the network [62], [63], (iii) hybrid networks, where all nodes can communicate with different devices with the core networks over the cloud. These three networking channels enable Road Side Unit (RSU) and Onboard Unit (OBU) to communicate and share information for diverse applications. For instance, traffic monitoring and control, mitigation and prevention of road accidents, crowd management, infotainment and enhancing the driving experiences.

Moreover, technology enhancement in mobile networks such as 5G networks with ultra-reliable and low latency services enhance the abilities of 4G or 3G wireless networks and can support highly diversified wireless channels, more sophisticatedly. 5G can use different spectrum bands such as millimetre wave (mmWave) and have the ability to carry a massive amount of data from multiple inputs sources and produces resultant multiple outputs (MIMO) in full-duplex mode. The area of transmission describes as small cells and also enable software-defined networks (SDN). Furthermore, beamforming is another feature of 5G technology that ensures

to identify the most efficient route for data transmission and with the properties of narrowing beams, vertical and horizontal transmission and, alteration of beam direction many times in milliseconds have strengthened the conventional means of mobile communications. And this has encouraged the biggest companies such as Apple, Google, and Microsoft to create an operating system and digital content manager for vehicular networks. CarPlay by Apple, Android Auto by Google and Windows Mobile by Microsoft are examples of onboard infotainment systems [64].

As illustrated in figure no 2 (Layer 2) all networks are associated with the backbone and core networks are linked with the Gateways or Remote Radio Heads (RRH), those are responsible for providing a connection between Wireless Base Station (WBS) and cloud networks by the guided and unguided medium. This is a matter of fact, that massive data traffic will require to increase the capacity and speed of Mobile Backhaul (MBH), MBH is the process of creating intermediate links between core networks and cells within respective edges of the network. Whereas mobile Front-haul (MFH) is the part of centralized radio access networks that provide connections between centralized radio controllers and their corresponding Remote RRH. Considering the 5G properties, it is also evident that in Massive Machine Type Communication (M-MTC) a large number of devices and their controlling connections (for instance sensors and actuators, etc.) can accommodate by using efficient call processing up to 100 folds. 5G networks offer the following features: (i) Heterogeneous network connections (ii) Dynamic network configuration management for all users and devices (iii) Easy to manage the commoditized devices and Open Source Software (OSS) (iv) Highly scalable networks with the ability to program the diverse applications. With these extensive communication capabilities, the dream of an autonomous vehicle with sensing, observing, navigation, context-aware and being sociable can be possible. But this is also a concern of implementation to modify the performance and quality parameters of preceding physical infrastructure too.

C. LAYER 3: DATA PROCESSING AND KNOWLEDGE DISCOVERY LAYER

The third layer of the proposed architecture presents different knowledge discovery aspects applied to the collected data from physical and cyberspaces. But before applying different data fusion and analytical techniques it is partially expected to preprocess the CPSN-IoV dataset to remove noise and redundancy to eliminate and estimate the missing values [24]. Data processing require cross-space collaborative series of execution for the understanding of context. For instance, on-board GPS sensory information and GPS data obtained from driver mobile phones both are numerical and related as well. But to construct the multimodal rational cross-space data structure, the physical and social world data is required to extract the related patterns, identify association and abstractions for knowledge discovery. As illustrated in figure no 3, edge-based processing and computation engines are

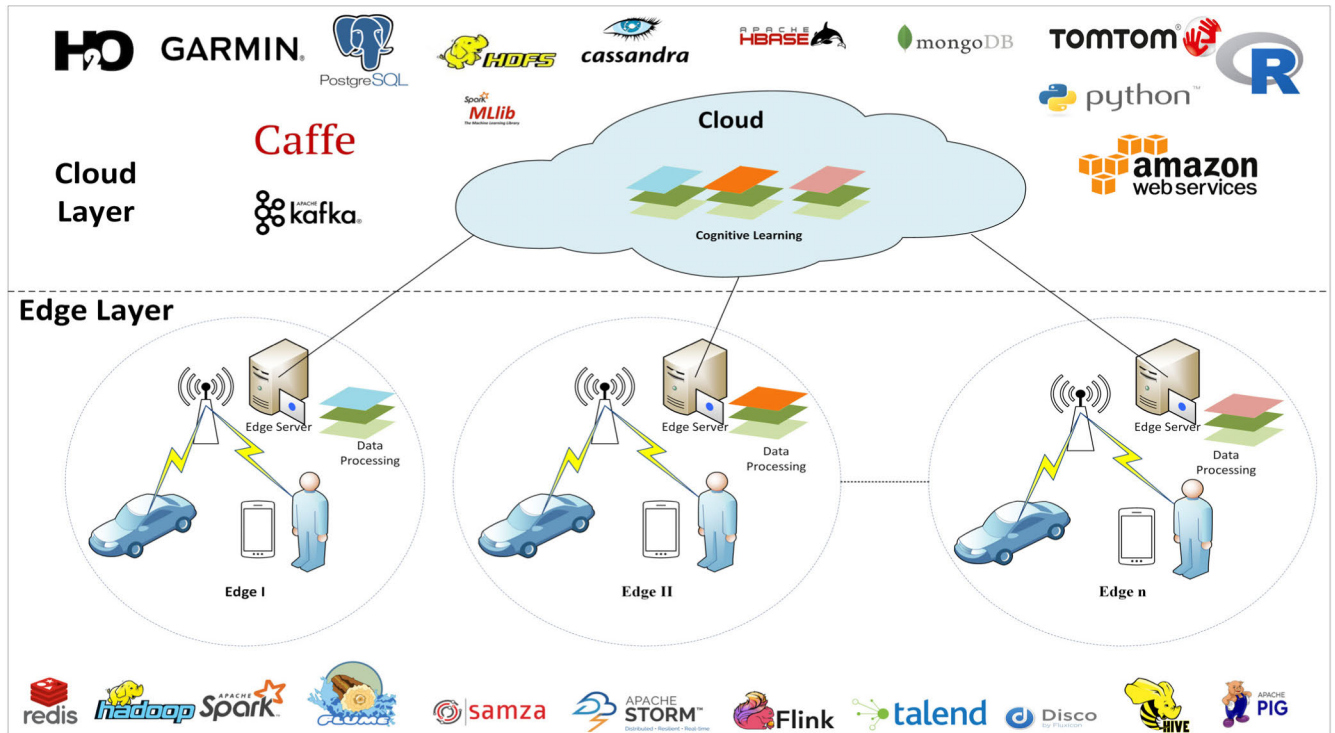


FIGURE 3. Data Processing services and solutions at edge/cloud Server- An illustration of edge and cloud based processing and computation engines for real time physiological and physical observation for instant learning and knowledge discovery.

distributed in closed proximities for real-time physiological and physical observation is being gathered and processed for instant learning and knowledge discovery. Further, this live observation and processing to the centralized cloud-based systems for ontology and rule-based engines for process learning and automation. Thus, the taxonomy of this layer can be divided into different aspects as described in this section.

1) DATA PROCESSING

The ecosystem of CPSN-IoV is highly heterogeneous, distributed, and complex. Sensory observation is like a nonentity without data processing thus, it requires different prospects to analyze the same data for discovering optimum knowledge. The computational platform of CPSN-IoV requires to address different layers of data processing and management that can be divided as (i) Stream processing and (ii) Batch processing and, (iii) Data Management.

a: STREAM PROCESSING

comprises real-time data processing for reducing the processing load of the database. A major feature of this processing is to provide preprocessing to the live data streams as, without such monitoring, these applications bring huge query load to the database due to raw and noisy data. Since the uploading and writing operation is already continuous load to the database, it will cause the overhead for accessing the same targeted object. In order to reduce this conflict, the subset of raw data can be extracted and maintained separately and

further it can be incrementally updated in the data store by data-driven methodologies. The extrication of read and write operations in the database will reduce the overhead of query processing and add scalability, and efficiency in applications. The real-time streaming events in CPSN-IoV are collection, extraction, filtering, monitoring, regulation, and automation of alert messages. Therefore, real-time efficient responsiveness is required for better integration and execution in existing infrastructure. Strom [65], JStrom, Spark [66], Flume [67], Apache S4 [68], Samza [69], and flink [70] are several technological platforms available for real-time stream processing with scalability, fault tolerance, high performance, high throughput, and low latency.

b: BATCH PROCESSING

is employed for complex data processing requirements and require archive data that has been loaded into databases iteratively by orchestration workflow. Batch processing leads for further data exploration by different data mining, machine learning, and knowledge engineering techniques to optimize analysis and visualization. Hadoop [71], Talend [72], Disco [73], Hive [74] and, pig [75] are different commercial tools available for batch processing.

c: DATA MANAGEMENT

is the third consideration of CPSN-IoV networks and addresses the storage and management of databases, caches, file systems, data warehouses, and lakes, etc. Data storage

for CPSN is heterogonous in nature, including SQL and NoSQL structures. The data store designs are dependent on comprehensive processing and analysis for efficient access requirements of the domain. Cassandra, HBase, MongoDB, PostgreSQL and HDFS are the most recurrently used solutions for data management [76]–[80].

2) DATA ANALYTICS

The major methodologies for data processing in CPSN-IoV can be categorized as follows:

a: DATA MINING

In CPSN-IoV, the real-time and archive data is generating continuously with several data irregularities and issues thus intelligent algorithms are required to process data. It is a well-known quote that, it is much easier to generate data than to process data. This explosion of data will certainly cause a serious problem in real-world data processing and analysis. Historically, it has tried to solve many big data processing problems [81], [82] but the need for effective and efficient analysis tools will definitely be required to stop the submergence of this huge amount of data. Real stream data mining and batch mining, problems should be addressed separately due to the nature of process handling. Data mining algorithms not only are the part of the hardware domain for cloud computing and distributed data services, but it is also required for software integration and process. However, in the current situations, most data mining technologies are developed to process on a single machine.

In the continuously changing circumstances, it is highly likely that the conventional mining algorithm will not be applied to execute the CPSN-IoV data. Generally, the procedures for preprocessing of data will need to be redefined otherwise these technologies cannot handle such unprecedented big data. For the understanding of the KDD process, three technological considerations are required for solving any problem in CPSN-IoV using data mining; goal (g), features (f), and mining algorithm (m). Data Processing for any problem domain can be described as Equation (5):

$$(X_i) = \{g_i, f_i, m_i\} \quad (5)$$

Goal (g_i): The aim of data processing should clearly be defined with all assumptions and limitations. Every problem should address the explanatory, confirmatory and exploratory parameters for better domain understanding.

Features (f_i): The features of data consist of data size, distribution nature, and visual representation. Different agents usually need different interpretations of the same data. Although data is coming from different sources as described in figure 2 (layer 1), it may be possible that one data source is sharing the same information as the other source. For instance, the geo-coordinates of the mobile phone of a driver are similar to the coordinates of the smart vehicle that he is driving. Thus, for both agents, different data processing will be required on similar data.

Mining Algorithm (m_i): Goal (g_i) and features (f_i) of the problem specify which algorithm will be required for the solution for any problem. There are several algorithms with the categorization of classification, clustering, frequent pattern and association analysis to satisfy knowledge discovery needs. The DIKW (Data, Information, Knowledge, and Wisdom) represents the data mining core principles and the end goal of every process in data mining is to facilitate the agents in closed proximity.

Most significantly in Equation (5), the selection of mining algorithm (m_i) depends on the nature of the problem domain and objective of the research. Several mining techniques have been used in CPSN for instance, the Density-based Spatial Clustering Algorithm (DBSCAN) was used for feature extraction of different caller records to assign the most appropriate activity label. Cellular towers were also attributed to geographically cluster the network by using the mining algorithm [83], [84]. In another research [85], KNN was used to form clusters and groups by modelling user behaviour and their preferences. Furthermore, a parking space management model was introduced by using a density-based clustering algorithm with GPS and POI data vectors. This recommendation system was used to generate candidates, to identify the available parking slots for public transportations [86]. The identified vector location was grouped with passenger's geolocation for smart pedestrian activities such as bus arrival, departure, drop-off, and pick-up location. Another model was mapped using publication/subscription architecture with the clustering algorithm to create an early warning system [10]. Association rules, data aggregation, tensor modelling, Local Indicators of Spatial Association (LISA) and, density-based clustering, are key algorithms used for data processing in CPSN [22], [87]–[91].

However, it is a matter of fact that data mining in CPSN is the ongoing process and more computationally efficient algorithms are still required for futuristic infrastructures for optimum processing. Commercially available systems expose details for a single process, depending on the devices used to gather and sense information in smart cities.

b: RULES FORMULATION, MANAGEMENT, AND ANOMALIES DETECTION

In CPSN-IoV, cross-space raw data are generated by heterogeneous and distributed agents. This multiscale data requires efficient data processing and rule mapping for state and domain understanding. For instance, consider the CPSN-IoV application, where physical objects such as smart vehicles, fixed sensory devices, and humans, etc. exist as agents. These have a structural and parametric understanding of every concept in smart environments.

Structural knowledge refers to the tasks that have to be done by agents for situation awareness and this learning can be achieved by different properties described as the parametric understanding of the context. Importantly, to satisfy the need for relevant data, parametric understanding is subjected to characterize by the operational framework in

smart environments. For instance, in smart environments the parametric set of smart vehicle (agent) can be described as geolocation, contacts, sensory and visual observation, communication channels, etc. and possible behaviour can be stopped, start, speed, increase speed, decrease speed, brake, short brake, show alert message and, send message, etc. Furthermore, onboard and cloud-based storage and the set of predefined rules lead to the understanding of a state and can help to identify a solution in any problem domain.

Moreover, an ontological model requires two aspects of understanding a) user model and, b) a smart environment that is based on user behaviour and preferences. Thus, rules can generate by the semantic understanding of any context. For instance, Wi-City plus has introduced a smart city observation system where the rules for customer satisfaction and quality of services are generated by using SPARQL and RDF queries [22]. Furthermore, iConAwa is another example of a context-aware model, where the semantic web is used to represent ontology for context agents and formalize the probabilistic approach, to handle the uncertainty about different events in pervasive environments [92]. It is evident by these historical approaches that data analysis and processing have to serve the need of all agents working in smart environments. Open Connectivity Foundation (OCF) has introduced a standard model for the representation of smart devices having the capabilities as IoT [93]. Additionally, Web of Things Model also has been proposed to develop a standard for construction of semantic models to define the physical objects into the virtual environments [94].

After context awareness and understanding of state, the forth next task is to detect anomalous events happening in close proximity that have a spatiotemporal state. This processing can be done by using social and traffic data analysis that have been continuously shared over the web by all agents. Computationally, it is required to fasten the data processing for irregular event detection because this may cause damages in real life. Thus, it is necessary to not only process the live data stream but also to disseminate the corresponding information for appropriate actions. Machine learning algorithms have been used in many case studies for event detection. For instance, route management and prediction of trajectory and road segment data are processed and mapped for efficient route planning. And the traffic route span is matched with the predefined scale and the traffic route having longer span than the predefined scale is detected as an anomaly [95].

Another method is proposed by the use of multiple sources of data such as road network, mobile calls, taxi trajectory, and car/bike renting, etc. This data set is modelled into LDA for distribution analysis and the rational test model is used for identification of irregular events [96]. Moreover, data processing and analysis is not only limited to anomaly detection but also refers to the mapping of situational reasoning. Spatiotemporal conditions with social network data can collectively be used for reducing the upcoming events such as traffic jams, accidents, and any catastrophic situation [97], [98].

c: ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Machine learning is a subset of Artificial Intelligence. Where, intelligence becomes feasible by mathematical, statistical models and algorithms to develop cognitive learning in processes. The process of choosing the best algorithm for any problem domain is actually the intensity and speed of required computation. Novel multimodal algorithms are vital for cognitive pattern recognition and to develop predictive data models for efficient decision support systems.

Genetic algorithms, artificial neural networks, deep learning, regression, evolutionary algorithms probabilistic and non-probabilistic models, etc. are some well-known algorithms for cognitive learning. For instance, drivers and smart cars form a crowdsensing and crowdsourcing network with a large amount of real-time data and observations. Machine learning algorithms provide multilayer contextual understanding and precise results for different problems in CPSN-IoV including; optimal route selection, smart parking, hindrance and accident detection, optimal path navigation, and smart infrastructure [99], [30], [100]–[102].

3) DATA FUSION

This layer addresses the statistical and logical methods for the integration of outputs and data of real-world objects. The data that has been collected from different sources as described in figure 2 (layer 1). The intent of this integration is to attain a cohesive and comprehend data understanding of the same object in heterogeneous, high-dimensional cross space. Situation awareness is the main goal of data fusion in CPSN-IoV. It allows disseminating all possible events affecting infrastructures and human. Additionally, it clarifies the actual synthesis of concepts in great depth and detail level of understanding for rigorous results. An extensive concern has been dedicated to the selection of algorithms and unified techniques to designate the principal elements of situation awareness in CPSN-IoV. However, integration of data is a prerequisite of high dimensional, multimodal and noisy collection of hybrid data sets [103]. Data fusion is far more than multi-sensor data fusion, as it is not only responsible for combining multisensory observation data but also to provide context description to the agents with intelligent data processing features. Following are some data fusion techniques from informational deluge for the construction of data-centric algorithms.

a: CORRELATION ANALYSIS

The architecture of CPSN-IoV is highly heterogeneous and distributed thus can be categorized into two approaches for processing a) *Centralized data fusion*, where the unprocessed data from multiple sensory nodes have been collected and then fused into centralized data processing nodes for event estimations and, b) *Distributed data fusion*, for independent data processing at different edges of network for getting awareness about the closed and local proximities. It is a matter of fact that processing all data sources at the

centralized node is neither feasible and nor efficient. This is due to the fact that the more number of nodes will add overhead to communication gateways and nodes. On the other hand, distribution will reduce the communication channels and the cost of deployment and integration.

The cross-space fusion of data uses statistical approaches for identifying the correlation between different data streams in CPS space. These include the identification of semantic cross space correlation in conjunction with sensory data of agents as described in figure 2 (layer 1). Pearson algorithm, Spearman, Kendall, Goodman & Kruskal, Kalman Filter (KF) for minimal mean squared error, Covariance and, Bar-Shalom Campo (BC) are different correlation types for cross-correlation identification and analysis [104].

Several techniques have been used for correlation analysis in CPSN-IoV. For instance, Patras, Greece's foursquare data was correlated with traffic volume for identifying the causes of pollutant objects. This correlation analysis was processed by diurnal pollutant averages for oxides of nitrogen (NO_x) and carbon monoxide (CO) with check-ins data [105]. Additionally, in CleanSpace [106] project different aggregation techniques were used to correlate pollution and temperature for smart cities. In research by Nest [107], different correlation and machine learning techniques were used to learn the habits of the inhabitant by using different sensory observation data. In research [20], V2I and parking space sensor data were used to predict the indent derivation for traffic influence and congestion control. Moreover, traffic patterns were identified and correlated with different temperature sensors in the smart city[87]. The suitability of correlation techniques in data fusion depends on the problem domain and appropriate assumptions for every case study.

b: CONTEXT DESCRIPTOR

Context is the hidden ingredient of data and it can turn into meaningful behaviour by the exploratory clarification of data from different channels to improve user experience. The context model is a series of different components; sensing, recognition, information distribution, cognitive reasoning, and context information model development[108]. In CPSN-IoV, situation awareness and detection methodology for action invocation is the ultimate goal. Where the aim is to accelerate the easier understanding and development of model while processing real-time data. This understanding leads towards social intelligence by using pattern recognition algorithms for social media data as explanatory and descriptive evidence to support the processing of physical sensory observation. Context descriptor not only enables the validation of the context model but also provides a mechanism to process the semantics as well. A number of researchers have used different context awareness and descriptor model techniques for dynamic rule analysis and management. For instance, in research [109], the physical sensory observation was used to monitor the environmental parameters for air quality by four different distinguished groups. The analysis was based on correlation analysis between the physical sensory data and

social information of these groups based on different characteristics. The resultant indicators have pledged the public grouping and involvement in air quality aspects. In another research [96], traffic anomalies were detected by using different data mining algorithms on social media data. These different terms are then interlinked for the identification of similar contextual constraints from different datasets. In addition, a traffic prediction model was developed by data analysis on weather, social media news feed, scheduled events and, static city geographical map [110].

c: TENSOR DECOMPOSITION

In CPSN-IoV, tensor decomposition is used for processing the high-dimensional data, for dynamic and multi-modal data process flows. Multivariate characteristics are analyzed to process the dynamic latent for the association, separation of concern and, anomaly pattern recognition. Tensors deals with the representation of different data sources, those are indexed and accessed more than one mode for the same instance. Typically, this decomposition reduces and retains the complexity of dimensionality conversions and captures, the high-level intelligence among system components. Tensor decomposition is helpful in the integration of future data processing by probabilistic models to create a unified framework based on group invariance for global convergence. Today, a smart vehicle can capture and disseminate approximately 250 GB data/hour and, it is predicted that there will be 50 billion smart objects integrated into CPSN by 2020 [111].

Thus, there is a need for multidimensional and complex data representation in terms of process, network, and memory-based hierarchy management in a distributed environment. Tensors create linear and multilinear associations among algebraic notions and often presents complex data in a more compact and array format. Tensor decomposition can fundamentally improve efficiency by applying numerical and multilinear libraries at lower level computation [112].

CANDECOMP/PARAFAC tensor decomposition is a well-known tool for high order and multidimensional data. Decomposition uses the factorization of tensors as a summation of the dot product of vectors and can be optimized by the least square method iteratively. Tensor decomposition with machine learning is a famous application in CPSN for object recognition, pattern discovery, clique detection, and prediction [113]–[115]. HOSVD used a distributed tensor decomposition method for object recognition and social connections with smart objects, such as smartphones and vehicles, etc. [116], but this understanding interprets the hidden community patters with IHOVD for better learning. Another instance of mapping the high dimensional data by denser approximate tensors' method is to probe the user behaviour similarity and prevalence for each instance of the group identified, by assigning a weight to each class of user [10].

d: SEMANTIC APPROACHES

Sensory observation and live stream data from all components of CPSN-IoV is heterogeneous, distributed, multimodal and,

multidimensional in nature. Knowledge extraction, integration and, dissemination lead towards common representation for complex associations between objects. Thus, there is a need for context annotation, feature extraction and, ontology-based methodologies for capturing the actual understanding of the situation. Automation in processes can be pledged to handle different contexts and states of objects as, the mapping of datasets into interconnected objects enables inference thus, provides the reasoning for decisions and actions. RDF based formats with SPARQL are well-known formats for representing the semantics and ontologies for decision support. For instance, in [22], different data sets comprising of; weather, live traffic, environment monitoring, sensory observation; user social media data are collectively integrated by semantic-based data fusion for path and event recommendation, generation of alert messages, elderly health assistance and DSS. In another instance [117], vehicular observation, user's social media, and fixed sensory data were processed by semantic modelling integrated with Dempster-Shafer theory to customize the user services.

Google cloud M3 frame is developed by Apache to facilitate the semantic reasoning and data fusion for vehicular sensory observation and smart environment data for edge-based computing platforms. Jena and AndroJena are lightweight libraries to empowering the process of reasoning engines [118].

D. LAYER 4: APPLICATION LAYER

The emergence of technology has been prompted by advancement in communication technologies, network diversity and vehicular onboard facilities. These enhancements have encouraged the industries for an extensive range of applications to facilitate humans. Billions of smart devices, vehicles, and on-road sensors have the ability to connect and, communicate for real-time information sharing. It is also a matter of fact that, during travel every user wants updated information and government authorities want to manage the traffic congestion, pollution, and event management, etc. remotely. Thus, every component of CPSN needs to collaborate, in order to create intelligent applications for the better quality of services. Mobile users on road with smartphones and, two or more smart vehicles in closed proximity may have a common concern thus, building complex topologies and social metrics. More specifically, the CPSN-IoV application metric can be generalized to the user on the basis of behaviour, association, routine events and activities and this metric can collect useful crowdsourcing data for social intelligence.

Parking management, public transport, traffic management, social transportation, regional logistics management, telematics, and fleet management are some of the required applications for CPSN-IoV. Researchers, industry, and authorities are collaborating in working for creating efficient and reliable solutions. In research [22], vehicular and infrastructure sensors collectively integrated for traffic awareness, management, reporting and have instant results using edge

computing. Moreover, a smart vehicular solution has been developed by using the camera, sensors, infrared detectors and, user data for parking management, traffic signals automation, trip prediction, and controlling system [20]. Wi-City project has also used traffic real-time data and weather update for route planning and recommendations [22].

III. SIMULATION & RESULTS

For validation and evaluation of our proposed model, we integrated several simulation tools over OMNet++ [119] as SUMO [120] and VSimRTI [121] for real-time traffic simulation and urban mobility analysis including crowd behaviour analysis. For CPSN-IoV real-time traffic scenarios have been studied that contain vehicle mobility in an urban city of Portugal, where the original data set contained instances of 800 vehicles including 15,434 observations. After closed review and preprocessing, we have only included 120 vehicles' data with 412 observational instances in 2 hours, respectively. This dataset also integrated with the city map for visualization and spatiotemporal analysis of events. Furthermore, the human mobility data including the GPS and social media feeds are integrated to relate in a given scenario.

We have categorized CPSN-IoV objects into three categories:

- (i) Road Side Units (RSU), those responsible for capturing the vehicular and GPS data of the mobile users.
- (ii) Vehicles (V) are the agents sharing the concurrent observation and GPS location.
- (iii) Walkers (W) are human, those are moving in respective edge and sharing GPS data and social media feeds.

A. DATA OBSERVATION AND DISSEMINATION

For our study, we have exemplified that RSU in edges is responsible for all types of information dissemination so the components share real-time data to the RSU and data centre. But before simulation, we have developed algorithms used for domain understanding. Algorithm 1 represents the process of RSU network distribution, working, and control in an edge. Where length, width, and radius are categorized to represent the edge boundary and the algorithm ensures all respective RSUs to cover geographical dimensions on the map. n_{rsu} and n_{GPS} are declared to store the actual location and coordinates of each RSU (line 7, line 8). This relevant information will store (line 14) and disseminate (line 16) by the set of all RSUs for edge-based computing.

Furthermore, vehicle observation and information sharing are described in Algorithm 2. Where real-time data radar reading, event data recorders, pictures, videos, route and GPS location (line 7-8) is captured by each vehicle (line 8- 11) and disseminate to respective RSU and Edge Data Centre (line 14) for further processing. Algorithm 3 describes the data observation of mobile user; this user may be on-board or wandering around. Thus, mobile data including the GPS and social media feed is periodically disseminating to the RSU

Algorithm 1 RSU Observation and Information Sharing

```

1 Input: Width  $\leftarrow w$  // Width of edge map
2 Length  $\leftarrow l$  // Length of edge map
3 Radius  $\leftarrow r$  // Radius of operating range
4 SetofVehicles  $\leftarrow X$  // Set of Vehicles in edge
5 CellularGPS  $\leftarrow g$  // Set of active Cellular devices
in edge
6 Output: Coordinates of objects for each RSU in edge
7 nsu  $\leftarrow$  GetRSU (w, l, r)
//returns the total no of RSUs in the edge
8 nGPS  $\leftarrow$  GetGPS (latidute,longitude)
//returns the total no of active cellular devices GPS
data in edge
9 for each (n  $\in$  nsu)  $\wedge$  (s  $\in$  nGPS) do
10 nsu_Coordinates.add(getCoordinates(n))
//returns the two dimension coordinates (x, y) of RSUs
11 nGPS_Coordinates.add(getLocation(s))
//returns the latitude and longitude of cellular phones
12 for each (V  $\in$  X)  $\wedge$  (G  $\in$  g) do
13 VData  $\leftarrow$  (V.getData ())  $\wedge$  (GPS_Data  $\leftarrow$  (G.getData (
)))
// returns observed data and GPS data of objects in
edge.
14 Y  $\leftarrow$  nsu_Coordinates.getRSU(V.getPosition())  $\wedge$ 
nGPS_Coordinates.getGPSData(G.getPosition())
//returns the coordinates (x, y) of RSUs and coordi-
nates (latitude and longitude) of cellular devices.
15 Message  $\leftarrow$  CreateMessage (VData  $\wedge$  GPS_Data)
16 Message.SendTo(Y)
//creates an information message contains vehicular
data and sends it to the corresponding RSU in the
edge.
```

in range as well as to the edge servers for interpretation and further analysis.

B. DATA CONTROL AND MANAGEMENT

For understanding the proposed study, we have selected one instance, to select and manage the shortest and efficient path by using CPSN-IoV characteristics and features. From algorithm 1, it has been finalized that each RSU has identified communication range within the bounded area. This property was ensured by the information that has been shared by a set of vehicles and mobile users in range. By concluding the understanding, each RSU is able to create a graph as $G = (R, C)$ where R is the set of all roads and C represents the crossroads within the boundary of each RSU that indicates the graph edges. We assume that each road (R) has the maximum and average speed to travel, so we have assigned the weights for each road reciprocally, depending on the prescribed and maximum allowed speed.

Thus, if the vehicle has a closer speed to the allowed road speed, it will lead to a lower weight, and if the speed of

Algorithm 2 Vehicle Observation and Information Sharing

```

1 Input: Radar-Readings  $\leftarrow RR$ 
//Readings of the vehicle
forward and
read Radar by Continuous
Wave
Frequency Modulation
(CWFM)
EDR-Readings
2  $\leftarrow$  EDRR
// Readings of vehicle speed,
braking, and throttle, etc.
3 Images  $\leftarrow$  IMG
// Capturing the real-time
environment images in the
surroundings of the vehicle.
4 GPS  $\leftarrow$  g
//Set of vehicle location
5 SetofVehicles  $\leftarrow$  X
//Set of Vehicles in the edge
6 Output: Real-time vehicle observation
7 for each (Vi  $\in$  X) do
//returns the observation of
vehicle
8 RR  $\leftarrow$  Vi.getRadarReading()
9 EDR  $\leftarrow$  Vi.getEDRReading()
10 IMG  $\leftarrow$  Vi.getIMG()
11 g  $\leftarrow$  Vi.getLocation()
12 R  $\leftarrow$  Vi.getRoute();
13 Y  $\leftarrow$  VD_ Data.add (getVehicleObservationData
(Vi))
//returns the Real time observation
to instance Y
14 Message  $\leftarrow$  CreateMessage (Y)
//Create a message containing the
observation
15 Message.SendTo (RSU,EDC)
//Send VD to respective RSU and
Edge Data Centre.
```

the vehicle is lower than the maximum allowed limit, it will have a higher weight. Selection of the best path and rerouting accordingly require RSU and edge servers to analyze the graph, that is periodically being shared by the set of all vehicles in the boundary. During analysis, the graph (G) indicates all roads with congestion, because of the weight assigned to the roads as described previously. Furthermore, social media feeds also are being analyzed for identifying the cause of congestion.

The causes of road congestion are due to many reasons for instance; obstacles as double parking, roadside or utility work, accidents and, may have mass transit due to some protest or road blockage, rush hours, and huge pedestrians crossing at the road, etc. Mobile users are sharing feeds at

Algorithm 3 Walkers (W) Data Observation and Sharing

```

1 Input:      SetofWalkers ← W
                // Set of users
2 CellularGPS ← g
                // returns the user mobile GPS data
3 SocialFeeds ← SF
                // returns the user social media
                Feeds
4 Output:    Real-Time mobile GPS coordinates
                and Social Media Feeds
5 nGPS ← GetPhoneGPS (latidute,longitude)
                // returns the cellular phone GPS
                data in edge
6 nSMF ← GetSocialMediaFeeds(UserID,
                timestamp,messageType,message)
                // returns the Social Media data in
                edge
7 for each (gp ∈ nGPS) ^ (sm ∈ nSMF) do
8 nGPS_Coordinates.add(getCoordinates(gp))
                // returns the latitude and
                longitude of cellular phones.
9 nSMF_Feeds.add(getSocialMediaFeed(sm))
                // returns the Social Media Feeds.
10 for each (Wi ∈ W) ^ (G ∈ g) do
11 UD ← nSMF_Feeds.getSocialMediaFeed
                (Wi.getSMF()) ^
                nGPS_Coordinates.getGPSData(G.getPosition())
                // returns the social media feeds
                and GPS coordinates of user mobile
                phone. Where mobile user can be
                on-board or walking around.
12 Message ← CreateMessage (UD)
13 Message.SendTo(RSU,EDC)
                // creates an information message
                contains vehicular data and sends
                it to the corresponding RSU in the
                edge.

```

social media regarding the events happening at the road and this will help to strengthen the graph for rerouting the option to the management of route.

Figure 4 illustrates the proposed simulation map of two RSUs and a set of vehicles, with all possible routes in a bounded two-dimensional plane. where X_i is the vehicle entering into the range of the edges; the entry point is donated as source node (S_n) in map. Vehicle X_i has three main routes as route 1 to 3 and one sub-route as 3B to reach the point of destination (D_n). Selection and re-routing occur at the current geo-location of the vehicle to the last road in the range of RSUs and edge servers. In this study, we have used the shortest path metric for rerouting and selection of the set of K paths to identify the optimal route. It is possible that same route may select as an optimal route for the set of the vehicle

in the edge. Hence, it is required to provide more than one alternative route, if possible to evade this issue.

Therefore, the Dijkstra shortest path method is used to calculate more than one route for each vehicle [122]. The reason to select Dijkstra other than available shortest path algorithms is the nature of selecting the shortest route between two nodes by calculating all nodes until reaching the destination node. Hence, the vehicle has global knowledge of network, and this feature will enable to identify the corresponding shortest route until the final destination. Moreover, Map Suite and distance matrix are used to store distances for each vertex from one node to another with less storage requirement. Dijkstra graph edges are non-negative and can save a lot of processing time and re-traversing nodes more than one time. Thus, Dijkstra provides support for time constraint and critical applications with less delay by minimizing the number of nodes. Comparison matrix for shortest path selection methods is illustrated as the table no 1.

Algorithm 4 describes the selection and re-routing process for vehicles, where X is the set of vehicles in the boundary of each RSU and G represents the graph created by each RSU for mapping the available routes for each vehicle. There may be more than one RSU in edge boundary but for better understanding, we have selected only one RSU for our study. For each vehicle on the map, there may be multiple alternative routes with better speed and minimal coverage time. K denotes the set of alternative routes for each vehicle (line 3). Moreover, two variables as α and ω are representing threshold and weight respectively. The value of α is dependable on the nature of the route and described as how many numbers of vehicles are presented in the boundary of RSU at a specific time and if the number of vehicles is more than α this will add weight to reroute the vehicles at source (S_n).

Furthermore, ω is initialized for adding the weight of social media feeds, it is assigned as; if social media feed is relevant to the road situation, route or any other relevant hashtag, it will assign 1, otherwise 0. Thus, RSU is responsible for calculating the path of each vehicle from the current location along with the last node of the graph as described D_n in figure 4. After perceiving the D_n , the K will be calculated as the shortest path by using the Dijkstra algorithm. Here the shortest route will be selected for balancing the route load. Importantly, each route has assigned weight as the sum of all sub-paths of the route as described in the previous section. Routes with lower resultant weight will be mostly requested thus, there is a need for such a constant measure that will reduce the likelihood of being chosen for every vehicle in edge boundary. The decision will be chosen on the set of different parameters as:

$$\begin{aligned}
 X &= \text{set of vehicles in the edge} \\
 \Delta X &= \text{set of routes for vehicle } V_i \text{ such that } (V_i \in X) \\
 \sigma_x^t &= \text{route } t \text{ of } V_i \text{ such that } (V_i \in X) \\
 \omega_x^t &= \text{weight of } \sigma_x^t
 \end{aligned}$$

Algorithm 4 Route Selection and Re-Routing

```

1 Input: X
  // Set of vehicles in the boundary
  of each RSU
2 Graph→G
  // graph created by each RSU for
  mapping the available routes in the
  boundary
Alternative-Routes
3 → K
  // Set of alternative routes for
  each vehicle
4 Threshold→ α
  //threshold for identifying
  traffic congestion using user
  mobile GPS.
5 Weight→ ω
  //weight for social media
  feeds after analysis store as
  1=relevance 2=not relevance for
  identifying the cause of traffic
  congestion.
6 for each(Vi ∈ X) do
  //set of vehicles representing
  each edge to create a route of the
  vehicle
7   route← Vi.getRoute( );
  //returns the present location
  and direction of vehicle Vi in
  graph.
8   Sn← Vi.getlocation( );
  //returns the current position
  of vehicle.
9   Dn← G.getLastNode(route);
  //returns the shortest path
  of route from source (Sn) to
  destination (Dn) in graph.
10  Ar←G.getKShortestRoute(Sn, Dn, K);
  //calculate and select the path
  of the vehicle from a set of
  alternative routes available
  in the graph (G), also assign
  and concatenate routes with
  vehicle (Vi)
11  Nr←Dijkstra(Ar, G);
  //instance NR will store a new
  route for the vehicle by
  Dijkstra shortest path
  distribution in the graph
  of edge.
12 If (Dn!=route.getEndNode( ) then
  //calculates the remaining routes
  of vehicle Vi from the destination
  node (Dn)

```

Algorithm 4 (Continued.) Route Selection and Re-Routing

```

13   If (nGPS ≥ α) then
  //check the total number
  of mobile GPS are more than
  the threshold
14   If (ω.nSMF=1) then
  // check the relevance of
  social media feeds to the
  road condition
15   Rn← route.split(Dn);
  // Split the remaining nodes
  before destination node
  (Dn)
16   Nr.add(Rn);
  // concatenate the new route
  (NR) with the remaining node
  (Rn) for vehicle (Vi)
17   Vi.setRoute(Nr);
18   end if
19 end if
20 end if
21 end

```

For simplicity and execution weights should be normalized, defined as Equation (6):

$$\mathbf{Z}(\omega_x^t) = \frac{(\sigma_x^t)}{\max\{(\sigma_x^t) | \forall (\sigma_x^t) \in \Delta X\}} \quad (6)$$

there is a number of possible routes with assigned weights thus the proposed equation of chosen route can be denoted by Equation (7):

$$R(\Delta X) = \max\{a^* P((\sigma_x^t) | \forall (\sigma_x^t) \in \Delta X), a \sim U([0, 1])\} \quad (7)$$

where $P(\sigma_x^t)$ is the probability of chosen path for vehicle Vi and variable a is the normalized weight assigned as 0 or 1. (Line 14) After route selection, it is checked whether the observed route is the end node of the route for vehicle Vi (Line 12). If this being correct, then the route will be recalculated for the updated road data in the respective edge. Moreover, mobile GPS and social media feed relevance will be checked for the relevance of the route situation and will split the remaining nodes before the end node. (Line 15) and concatenate the new route NR with the remaining nodes (Rn) for vehicle Vi. (Line 16,17).

C. SIMULATION AND EVALUATION

For CPSN-IoV real-time traffic scenario has been analyzed that contains vehicle mobility in an urban city of Portugal, where the original data set was containing instances of 800 vehicles including 15, 434 observations. After closed review and preprocessing, we have only included 120 vehicles' data with 412 observational instances in 2 hours, respectively. The reason for selecting the 2-hours simulation is the simplicity of network length. As defined earlier we have

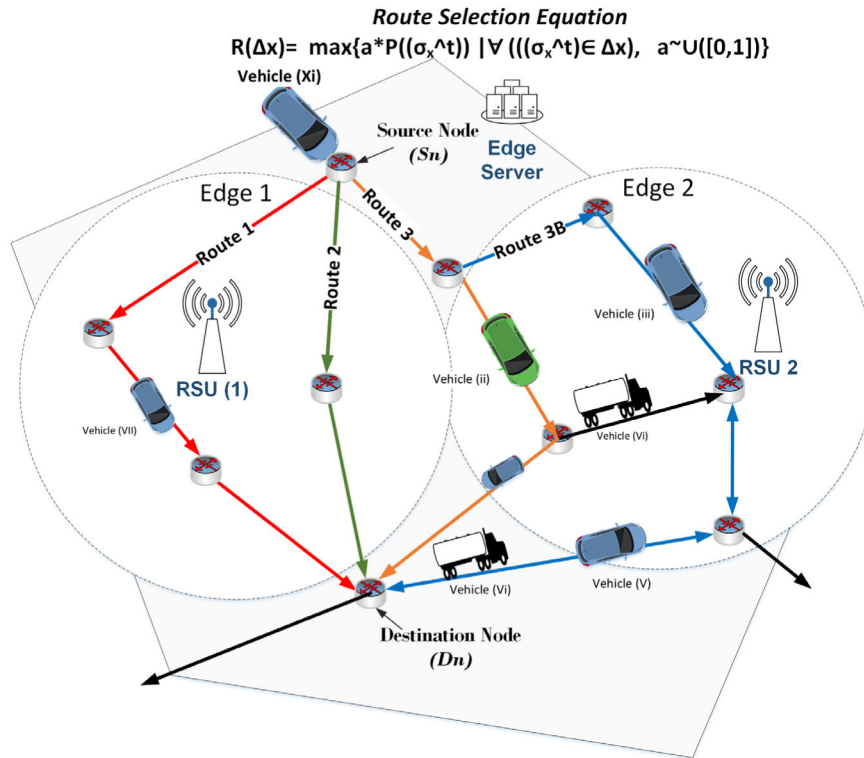


FIGURE 4. An illustration of proposed simulation-The simulation setting the proposed setup, where the S_n and D_n are source and destination respectively and this plane of two RSUs described as two information dissemination units in one edge.

generated nodes by using SUMO [120] with different parameters as, the maximum and average speed limit for each route. By running the simulation, we have executed the following tasks: (i) Introduce edge network to the new vehicle entering in range. (ii) assign weight to each route depending upon the availability of optimized path and, (iii) assign route to the vehicle. Network parameters for the simulation are enlisted in Table 2.

In this section, we represent simulation results of the study, where figure 5 (a) is showing the normal behaviour of traffic for the same source (S_n) and destination (D_n) using the routes as λ_1 to λ_3 represents route 1 to route 3 respectively. The normal behaviour of vehicles under these routes are showing fluctuation depending upon different parameters as rush hours, vehicle speed, driving behaviour, etc. Figure 5 (b) shows the results after running the simulation algorithms and street view of simulated results are illustrated by google maps in figure no 6. Figure 5(c) represents the routing overhead of simulation with respect to realistic and CPSN-IoV minimum, maximum and mean value. The routing overhead is gradually decreasing for specific simulation setup. Figure 5(d) illustrates the end to end delay of the number of packets during simulation routing. Moreover, figure 5(e) and figure 5(f) represent the packet delay ratio and throughput rate of transmission. Importantly, we have tested different parameters for variation as described in the previous section. The evaluation has been analyzed for the resultant specifics such as, it has

TABLE 2. Network parameters setting for simulation.

Parameters	Values
MAC Protocol	IEEE 802.11p
Transmission Range	500m
Bit rate	18Mbps
Transmission Rate	6Mbps
Vehicle Speed	15-45 m/s
Map	Grid 2 x5 (2km x 2.5 km)
Number of streets	21
Number of intersection	15
Number of vehicles	120
Observation instances	412
Packet size	512 bytes
Simulation time	21600s for each test
Departure/Arrival Position	Randomly Chosen
Number of Traffic lights	20
Social Media Feeds	Twitter
Network Interface	OMNET++
Traffic Simulator	SUMO

been noted that after rerouting, the time for route coverage has been decreased by 20 percent approximately. Moreover, better communication range can have better route predictions and does not require changing paths unnecessarily. Thus, efficient traffic will result in lesser fuel consumption, fatigue, and proper management.

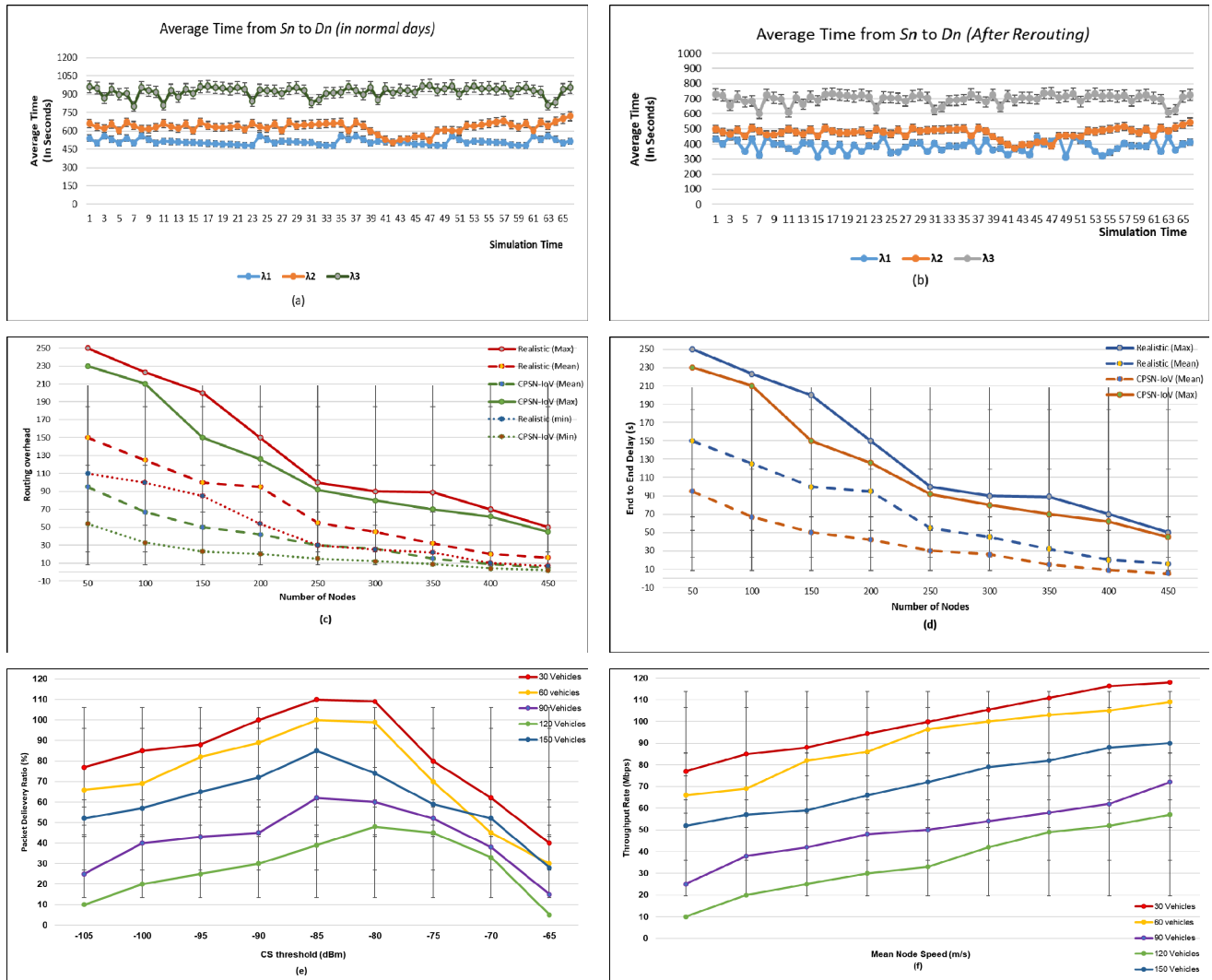


FIGURE 5. Simulation Results: (a) The graph shows the normal behavior of traffic for same source (S_n) and destination (D_n) using the routes as λ_1 to λ_3 (b) After implication of proposed rerouting algorithm the simulation results for the same routes (c) Illustration of routing overhead with respect to realistic min, realistic mean and CPSN-IoV min and mean (after applying algorithms) (d) End to End delay of realistic min, mean and CPSN-IoV min mean (e) The representation of packet delay ratio from 30 vehicles to 150 vehicles (f) Throughput rate of end to end transmission from 30 vehicles to 120 vehicle.

IV. DISCUSSION

Internet of Vehicles (IoV) especially the emerging technologies for autonomous vehicles are key technologies for improving the conventional traffic behaviour including security, congestion and, efficiency. IoV can use cloud services for accessing the data from V2I and V2V network components, thus the vehicle capabilities can expand by CPSN extensions. In this study, we have proposed the CPSN-IoV, where intelligent vehicles, traffic control and management system, data processing solutions and, cloud computation services work together to manage the complete transportation hierarchy. Cloud services offer assistance to vehicle and people on board as cyberspace parallelism by using a wide range of applications. Moreover, CPSN is not only limited to intelligent services for transportation rather people’s transportation is also part of the concern. Because people’s data bring other social bonding parameters and can improve the vehicle’s

autonomy and lead to developing a solid base for the next generation of IoV as CPSN-IoV. The human behaviour factors, experiences and learning have introduced new modelling and analysis methods as personal assistance style, software tools and application choices, navigational schemes, and top searched locations etc. In future, the collective information of cyber, social and, physical observation system can help to reduce traffic by intelligent navigation, emergency control and management, Intersection and traffic signals control and many more. Optimal route selection for vehicles already exists in literature for instance [91], [123]–[127] but these proposed methods only have included the navigational patterns from archival data of vehicles but social behaviour of people on board or walkers didn’t contribute to identifying the causes of congestion. In this research, we have introduced how all components of CPSN can work together in order to facilitate the navigation of vehicles. We have developed

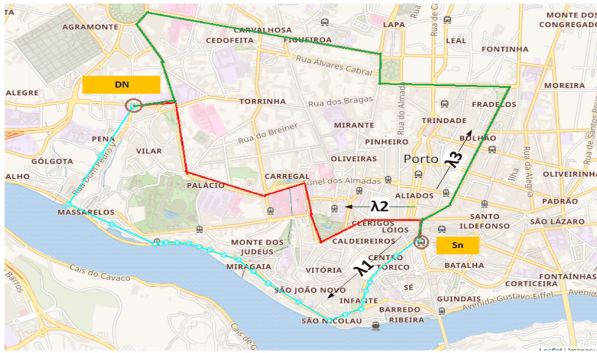


FIGURE 6. Street view of the simulated results in google map.

algorithms for getting the GPS positions, vehicle and RSU observations and, social media feeds of people on board and walkers as well, in closed proximity and all parameters assisted in finding the optimal route. Drivers, passengers and people on roads exhibit the agents in the controlled and smart environment and have capabilities to block the roads for specific time-space, the transportation control and management based CPSN will manage the services and hence provide guidance for traffic avoidance. By simulation we have tested for specific case study as described in section III of this study. However, this is just the start of next generation of IoV where social factor will also be included in IoV applications and thus have opened the continuum of possibilities for instance assistance, control and autonomous driving. Thus, there is need for academic and market researchers to work in this domain for more outcomes.

V. CONCLUSION AND, FUTURE WORK

In this study, we have proposed the cross-space, integrated and multimodal architecture of CPSN-IoV and discussed the key features to address different related aspects in order to reduce the complexity between heterogeneous data sources. Data processing for smart cities requires smart infrastructure to handle massive amounts of data in real-time. Therefore, this domain requires innovative ideas, techniques, and algorithms for efficient data execution in real-time and this study is an effort to initiate the implication of CPSN-IoV for pattern engineering, crowdsources and, social intelligence. Combining the sensory observation with social media feeds will not only be helpful in operational management and communication rather, but this will also enable a whole new paradigm that can implicate citywide services and solutions. The rising prevalence of numerical and multimodal data fusion has also strengthened the semantic reasoning and knowledge discovery to encourage social intelligence for efficient and optimized decision making and resource management.

Moreover, we also have discussed some of the available commercial solutions to integrate sensory observation with social media feeds and have also simulated and analyzed a trajectory dataset to strengthen our proposed study. The proposed architecture is simulated in OMNeT++ for vehicle mobility and dynamic traceability for routes management with the help of sensory observation and social media feeds.

Based on the significant findings of the study, future work is an open domain for all collaborative research areas in order to explore data fusion, correlation and statistical relationships between sensory data streams, crowdsources, social media feeds, efficient communication protocols.

VI. CHALLENGES

Considering the advantages of CPSN-IoV, it provides new opportunities for academia, research and, industry. However, it may include many challenges and research problems. This section follows some of the related issues in CPSN-IoV as hereunder:

(i) *Privacy and Security*: One of the most vigorous challenges is the privacy and security of the architecture because the risk of being hacked at different levels of communication links is a real concern of every stakeholder as the geographically distributed and unattended devices are the common vulnerable objects for cyber-attacks [128]–[130]. Moreover, high management cost restricts the upgrading ratio of devices and applications using in CPSN-IoV and hacking methodologies are getting more advanced day by day due to better computation devices and faster communication channels. Therefore, there is a need for IT companies to understand the importance of risk-free applications in order to build confidence among all users and reduce any privacy violation impact.

(ii) *Legislation and Regularization*: In CPSN-IoV, communication technologies have the required number of dynamic ranges and characteristics of channels. It may be a complicated task for the government to create regulations and law as well, with other serious problems in the real world. However, we cannot deny the importance of being able to regularize all available protocols and standards for managing V2V, V2I, and V2X connections with security, efficiency, and reliability [118], [131], [132].

(iii) *Integration and Interoperability*: System integration is also a problem due to data sharing among heterogeneous devices using multiscale platforms as a result, the performance of applications in CPSN-IoV will be reduced. Data interchange, portability and, interoperability are serious issues thus require to scale according to the system design and usage. Optimization of operational costs, increase in bandwidth and range, power consumption management for all applications are some of the open research dimensions in CPSN-IoV [118], [24].

(iv) *Data Analytics*: with the gradual change in technology, the huge amount of observation and data gathered from billions of objects e.g., sensors, actuators, mobile phones, smart cars, etc. can paralyze the big data processing and analytical companies. Currently, many companies are facing difficulties in data management, storage, and analytics for existing data. Consequently, conventional data analytics are not enough to address the big data processing and visualization issues [133]. Thus, data reduction methodologies may require us to rebuild statistical and, sampling methodologies and models for better prediction and interpretation [134].

REFERENCES

- [1] J. Poushter, J. Carle, and B. James. (Mar. 19, 2015). *Internet Seen as Positive Influence on Education But Negative on Morality in Emerging and Developing Nations* | Pew Research Center. Accessed: Apr. 2, 2019. [Online]. Available: <https://www.pewglobal.org/2015/03/19/internet-seen-as-positive-influence-on-education-but-negative-influence-on-morality-in-emerging-and-developing-nations/>.
- [2] K. Lyytinen and Y. Yoo, "Ubiquitous computing," *Commun. ACM*, vol. 45, no. 12, pp. 63–96, 2002.
- [3] J. Zeng, L. T. Yang, M. Lin, H. Ning, and J. Ma, "A survey: Cyber-physical-social systems and their system-level design methodology," *Future Gener. Comput. Syst.*, vol. 105, pp. 1028–1042, Apr. 2020, doi: [10.1016/j.future.2016.06.034](https://doi.org/10.1016/j.future.2016.06.034).
- [4] H. Ning, H. Liu, J. Ma, L. T. Yang, and R. Huang, "Cybermatics: Cyber-physical-social-thinking hyperspace based science and technology," *Future Gener. Comput. Syst.*, vol. 56, pp. 504–522, Mar. 2016, doi: [10.1016/j.future.2015.07.012](https://doi.org/10.1016/j.future.2015.07.012).
- [5] H. Zhuge, "Cyber-physical society—The science and engineering for future society," *Future Gener. Comput. Syst.*, vol. 32, pp. 180–186, Mar. 2014, doi: [10.1016/j.future.2013.10.008](https://doi.org/10.1016/j.future.2013.10.008).
- [6] Z. Liu, D.-S. Yang, D. Wen, W.-M. Zhang, and W. Mao, "Cyber-physical-social systems for command and control," *IEEE Intell. Syst.*, vol. 26, no. 4, pp. 92–96, Jul. 2011.
- [7] M. Conti, A. Passarella, and S. K. Das, "The Internet of people (IoP): A new wave in pervasive mobile computing," *Pervasive Mobile Comput.*, vol. 41, pp. 1–27, Oct. 2017, doi: [10.1016/j.pmcj.2017.07.009](https://doi.org/10.1016/j.pmcj.2017.07.009).
- [8] B. Bordel, R. Alcarria, T. Robles, and D. Martín, "Cyber-physical systems: Extending pervasive sensing from control theory to the Internet of Things," *Pervasive Mobile Comput.*, vol. 40, pp. 156–184, 2017, doi: [10.1016/j.pmcj.2017.06.011](https://doi.org/10.1016/j.pmcj.2017.06.011).
- [9] L. Atzori, A. Iera, G. Morabito, and M. Nitti, "The social Internet of Things (SIoT)—When social networks meet the Internet of Things: Concept, architecture and network characterization," *Comput. Netw.*, vol. 56, no. 16, pp. 3594–3608, Nov. 2012, doi: [10.1016/j.comnet.2012.07.010](https://doi.org/10.1016/j.comnet.2012.07.010).
- [10] Y. Zhang, "GroRec: A group-centric intelligent recommender system integrating social, mobile and big data technologies," *IEEE Trans. Serv. Comput.*, vol. 9, no. 5, pp. 786–795, Sep. 2016, doi: [10.1109/TSC.2016.2592520](https://doi.org/10.1109/TSC.2016.2592520).
- [11] P. Hehenberger, B. Vogel-Heuser, D. Bradley, B. Eynard, T. Tomiyama, and S. Achiche, "Design, modelling, simulation and integration of cyber physical systems: Methods and applications," *Comput. Ind.*, vol. 82, pp. 273–289, Oct. 2016, doi: [10.1016/j.compind.2016.05.006](https://doi.org/10.1016/j.compind.2016.05.006).
- [12] D. Moldovan, G. Copil, and S. Dustdar, "Elastic systems: Towards cyber-physical ecosystems of people, processes, and things," *Comput. Standards Interfaces*, vol. 57, pp. 76–82, Mar. 2018, doi: [10.1016/j.csi.2017.04.002](https://doi.org/10.1016/j.csi.2017.04.002).
- [13] F. Harashima, M. Tomizuka, and T. Fukuda, "Mechatronics- 'what is it, why, and how?' An editorial," *IEEE/ASME Trans. Mechatronics*, vol. 1, no. 1, pp. 1–4, Mar. 1996, doi: [10.1109/TMECH.1996.7827930](https://doi.org/10.1109/TMECH.1996.7827930).
- [14] E. M. Frazzon, J. Hartmann, T. Makuschewitz, and B. Scholz-Reiter, "Towards socio-cyber-physical systems in production networks," *Procedia CIRP*, vol. 7, pp. 49–54, May 2013, doi: [10.1016/j.procir.2013.05.009](https://doi.org/10.1016/j.procir.2013.05.009).
- [15] F. Salim and U. Haque, "Urban computing in the wild: A survey on large scale participation and citizen engagement with ubiquitous computing, cyber physical systems, and Internet of Things," *Int. J. Hum.-Comput. Stud.*, vol. 81, pp. 31–48, Sep. 2015, doi: [10.1016/j.ijhcs.2015.03.003](https://doi.org/10.1016/j.ijhcs.2015.03.003).
- [16] I. Lee, "Challenges and research directions in medical cyber-physical systems," *Proc. IEEE*, vol. 100, no. 1, pp. 75–90, Oct. 2011.
- [17] A. El-Hoiydi and J.-D. Decotignie, "WiseMAC: An ultra low power MAC protocol for the downlink of infrastructure wireless sensor networks," in *Proc. 9th Int. Symp. Comput. Commun. (ISCC)*, vol. 1, Dec. 2004, pp. 244–251, doi: [10.1109/iscc.2004.1358412](https://doi.org/10.1109/iscc.2004.1358412).
- [18] B. Guo, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Comput. Surv.*, vol. 48, no. 1, p. 7, Sep. 2015, doi: [10.1145/2794400](https://doi.org/10.1145/2794400).
- [19] Y. Zhou, S. De, and K. Moessner, "Real world city event extraction from Twitter data streams," *Procedia Comput. Sci.*, vol. 98, pp. 443–448, Jan. 2016, doi: [10.1016/j.procs.2016.09.069](https://doi.org/10.1016/j.procs.2016.09.069).
- [20] W. Guo, Y. Zhang, and L. Li, "The integration of CPS, CPSS, and ITS: A focus on data," *Tinshua Sci. Technol.*, vol. 20, no. 4, pp. 327–335, Aug. 2015, doi: [10.1109/tst.2015.7173449](https://doi.org/10.1109/tst.2015.7173449).
- [21] T. Anagnostopoulos, D. Ferreira, A. Samodelkin, M. Ahmed, and V. Kostakos, "Cyclist-aware traffic lights through distributed smartphone sensing," *Pervasive Mobile Comput.*, vol. 31, pp. 22–36, Sep. 2016, doi: [10.1016/j.pmcj.2016.01.012](https://doi.org/10.1016/j.pmcj.2016.01.012).
- [22] A. Costanzo, A. Faro, D. Giordano, and C. Spampinato, "An ontological ubiquitous city information platform provided with cyber-physical-social-systems," in *Proc. 13th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2016, pp. 137–144, doi: [10.1109/ccnc.2016.7444746](https://doi.org/10.1109/ccnc.2016.7444746).
- [23] A. Costanzo, A. Faro, D. Giordano, and C. Spampinato, "Implementing cyber physical social systems for smart cities: A semantic Web perspective," in *Proc. 13th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2016, pp. 274–275.
- [24] B. Guo, Z. Yu, and X. Zhou, "A data-centric framework for cyber-physical-social systems," *IT Prof.*, vol. 17, no. 6, pp. 4–7, Nov. 2015, doi: [10.1109/mitp.2015.116](https://doi.org/10.1109/mitp.2015.116).
- [25] R. Baheti and H. Gill, "Cyber-physical systems," *Impact Control Technol.*, vol. 12, pp. 161–166, Mar. 2011.
- [26] J. Manyika, *Disruptive Technologies: Advances that will Transform Life, Business, and the Global Economy* | McKinsey. Accessed: Sep. 2, 2019. [Online]. Available: <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/disruptive-technologies>
- [27] H. Lasi, P. Fettke, H. G. Kemper, T. Feld, and M. Hoffmann, "Industry 4.0," *Bus. Inf. Syst. Eng.*, vol. 6, no. 4, pp. 239–242, 2014.
- [28] *Workload Model Based Dynamic Adaptation of Social Internet of Vehicles—PubMed—NCBI*. Accessed: Apr. 4, 2019. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/26389905>
- [29] P. Anantharam, P. Barnaghi, K. Thirunarayan, and A. Sheth, "Extracting city traffic events from social streams," *ACM Trans. Intell. Syst. Technol.*, vol. 6, no. 4, pp. 43:1–43:27, Jul. 2015, doi: [10.1145/2717317](https://doi.org/10.1145/2717317).
- [30] G. Fusco, C. Colombaroni, L. Comelli, and N. Isaenko, "Short-term traffic predictions on large urban traffic networks: Applications of network-based machine learning models and dynamic traffic assignment models," in *Proc. Int. Conf. Models Technol. Intell. Transp. Syst. (MT-ITS)*, Jun. 2015, pp. 93–101, doi: [10.1109/mtits.2015.7223242](https://doi.org/10.1109/mtits.2015.7223242).
- [31] S. Djahel, R. Doolan, G.-M. Muntean, and J. Murphy, "A communications-oriented perspective on traffic management systems for smart cities: Challenges and innovative approaches," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 125–151, Mar. 2015.
- [32] C. G. Cassandras, "Smart cities as cyber-physical social systems," *Engineering*, vol. 2, no. 2, pp. 156–158, Jun. 2016.
- [33] G. Xiong, F. Zhu, X. Liu, X. Dong, W. Huang, S. Chen, and K. Zhao, "Cyber-physical-social system in intelligent transportation," *IEEE/CAA J. Autom. Sinica*, vol. 2, no. 3, pp. 320–333, Jul. 2015.
- [34] TomTom. *TomTom | Home*. Accessed: Dec. 20, 2019. [Online]. Available: <https://www.tomtom.com/>
- [35] Garmin International | Home. Accessed: Dec. 20, 2019. [Online]. Available: <https://www.garmin.com/en-US/>
- [36] V. T. N. Nha, S. Djahel, and J. Murphy, "A comparative study of vehicles' routing algorithms for route planning in smart cities," in *Proc. 1st Int. Workshop Veh. Traffic Manage. Smart Cities (VTM)*, Nov. 2012, pp. 1–6.
- [37] S. Wang, S. Djahel, J. Mcmanis, C. McKenna, and L. Murphy, "Comprehensive performance analysis and comparison of vehicles routing algorithms in smart cities," in *Proc. Global Inf. Infrastruct. Symp. (GIIS)*, Oct. 2013, pp. 1–8, doi: [10.1109/giis.2013.6684365](https://doi.org/10.1109/giis.2013.6684365).
- [38] D. Smith, S. Djahel, and J. Murphy, "A SUMO based evaluation of road incidents' impact on traffic congestion level in smart cities," in *Proc. 39th Annu. IEEE Conf. Local Comput. Netw. Workshops*, Sep. 2014, pp. 702–710.
- [39] I. Chabini and S. Lan, "Adaptations of the A* algorithm for the computation of fastest paths in deterministic discrete-time dynamic networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 1, pp. 60–74, Mar. 2002.
- [40] B. Huang, Q. Wu, and F. B. Zhan, "A shortest path algorithm with novel heuristics for dynamic transportation networks," *Int. J. Geograph. Inf. Sci.*, vol. 21, no. 6, pp. 625–644, Jul. 2007.
- [41] M. G. Bell, V. Trozzi, S. H. Hosseinloo, G. Gentile, and A. Fonzone, "Time-dependent hyperstar algorithm for robust vehicle navigation," *Transp. Res. A. Policy Pract.*, vol. 46, no. 5, pp. 790–800, 2012.
- [42] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.*, vol. 1, no. 1, pp. 269–271, Dec. 1959.
- [43] J. C. Pemberton and R. E. Korf, "Incremental path planning on graphs with cycles," in *Artificial Intelligence Planning Systems*. Amsterdam, The Netherlands: Elsevier, 1992, pp. 179–188.
- [44] D. Delling, P. Sanders, D. Schultes, and D. Wagner, "Engineering route planning algorithms," in *Algorithmics of Large and Complex Networks*. Berlin, Germany: Springer, 2009, pp. 117–139.

- [45] Z. Michalewicz and S. J. Hartley, *Genetic Algorithms + Data Structures = Evolution Programs* (Mathematical Intelligence), New York, NY, USA: Springer, vol. 18, no. 3, 1996, p. 71.
- [46] J.-B. Waldner, *Nanocomputers and Swarm Intelligence*. Hoboken, NJ, USA: Wiley, 2013.
- [47] F. Glover, "Future paths for integer programming and links to artificial intelligence," *Comput. Oper. Res.*, vol. 13, no. 5, pp. 533–549, 1986.
- [48] J. Bang-Jensen and G. Gutin, "Section 2.3.4: The Bellman-Ford-Moore algorithm," *Digraphs: Theory, Algorithms and Applications*, 1st ed. 2000.
- [49] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. Cambridge, MA, USA: MIT Press, 2009.
- [50] H. Li and D. K. Irick, "Collaborative spectrum sensing in cognitive radio vehicular ad hoc networks: Belief propagation on highway," in *Proc. IEEE 71st Veh. Technol. Conf.*, May 2010, pp. 1–5.
- [51] J. Gong, C.-Z. Xu, and J. Holle, "Predictive directional greedy routing in vehicular ad hoc networks," in *Proc. 27th Int. Conf. Distrib. Comput. Syst. Workshops (ICDCSW)*, 2007, p. 2.
- [52] C. L. P. Chen, J. Zhou, and W. Zhao, "A real-time vehicle navigation algorithm in sensor network environments," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1657–1666, Dec. 2012.
- [53] T.-Y. Liao and T.-Y. Hu, "An object-oriented evaluation framework for dynamic vehicle routing problems under real-time information," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 12548–12558, Sep. 2011.
- [54] M. Gendreau, J.-Y. Potvin, O. Bråumlays, G. Hasle, and A. Løkketangen, "Metaheuristics for the vehicle routing problem and its extensions: A categorized bibliography," in *The Vehicle Routing Problem: Latest Advances and New Challenges*. Boston, MA, USA: Springer, 2008, pp. 143–169.
- [55] H. Kanoh and K. Hara, "Hybrid genetic algorithm for dynamic multi-objective route planning with predicted traffic in a real-world road network," in *Proc. 10th Annu. Conf. Genetic Evol. Comput.*, 2008, pp. 657–664.
- [56] B. Chakraborty, T. Maeda, and G. Chakraborty, "Multiobjective route selection for car navigation system using genetic algorithm," in *Proc. IEEE Midnight-Summer Workshop Soft Comput. Ind. Appl. (SMCia)*, 2005, pp. 190–195.
- [57] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *SIAM Rev.*, vol. 51, no. 4, pp. 661–703, Nov. 2009, doi: 10.1137/070710111.
- [58] H. Zhuge, "Semantic linking through spaces for cyber-physical-socio intelligence: A methodology," *Artif. Intell.*, vol. 175, nos. 5–6, pp. 988–1019, Apr. 2011, doi: 10.1016/j.artint.2010.09.009.
- [59] A. Whitworth, J. Webster, and D. Howard, *A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia: A Clinician's Guide*. Abingdon, U.K.: Taylor & Francis, 2005, pp. 1–292, doi: 10.4324/9780203493502.
- [60] G. R. H. Gray, "The social psychology of telecommunications," *Telecommun. Policy*, vol. 1, no. 2, pp. 175–176, Mar. 1977.
- [61] R. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: Current state and future challenges," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 32–39, Nov. 2011, doi: 10.1109/mcom.2011.6069707.
- [62] P. Caballero-Gil, C. Caballero-Gil, and J. Molina-Gil, "How to build vehicular ad-hoc networks on smartphones," *J. Syst. Archit.*, vol. 59, no. 10, pp. 996–1004, Nov. 2013.
- [63] A. Boukerche, H. A. B. F. Oliveira, E. F. Nakamura, and A. A. F. Loureiro, "Vehicular ad hoc networks: A new challenge for localization-based systems," *Comput. Commun.*, vol. 31, no. 12, pp. 2838–2849, Jul. 2008.
- [64] D. Gruyer, R. Belaroussi, and M. Revilloud, "Accurate lateral positioning from map data and road marking detection," *Expert Syst. Appl.*, vol. 43, pp. 1–8, Jan. 2016, doi: 10.1016/j.eswa.2015.08.015.
- [65] *Apache Storm*. Accessed: Jul. 11, 2019. [Online]. Available: <https://storm.apache.org/>
- [66] *Apache Spark™—Unified Analytics Engine for Big Data*. Accessed: Jul. 11, 2019. [Online]. Available: <https://spark.apache.org/>
- [67] *Welcome to Apache Flume—Apache Flume*. Accessed: Jul. 11, 2019. [Online]. Available: <https://flume.apache.org/>
- [68] *S4 0.5.0 Overview—Apache S4—Apache Software Foundation*. Accessed: Jul. 11, 2019. [Online]. Available: <https://cwiki.apache.org/confluence/display/S4/S4+0.5.0+Overview>
- [69] *Samza*. Accessed: Jul. 11, 2019. [Online]. Available: <http://samza.apache.org/>
- [70] *Apache Flink: Stateful Computations Over Data Streams*. Accessed: Jul. 11, 2019. [Online]. Available: <https://flink.apache.org/>
- [71] *Apache Hadoop*. Accessed: Jul. 11, 2019. [Online]. Available: <https://hadoop.apache.org/>
- [72] *Industry-Leading Data Management Platform—Talend, Talend Real-Time Open Source Data Integration Software*. Accessed: Jul. 11, 2019. [Online]. Available: <https://www.talend.com/products/data-integration/data-management-platform/>
- [73] *DISCO*. Accessed: Jul. 11, 2019. [Online]. Available: <https://support.csdisco.com/hc/en-us>
- [74] *Apache Hive™*. Accessed: Jul. 11, 2019. [Online]. Available: <https://hive.apache.org/>
- [75] *Welcome to Apache Pig!* Accessed: Jul. 11, 2019. [Online]. Available: <https://pig.apache.org/>
- [76] *Apache Cassandra*. Accessed: Jul. 10, 2019. [Online]. Available: <http://cassandra.apache.org/>
- [77] *Apache HBase—Apache HBase™ Home*. Accessed: Jul. 10, 2019. [Online]. Available: <https://hbase.apache.org/>
- [78] MongoDB. *The Most Popular Database for Modern Apps*. Accessed: Jul. 10, 2019. [Online]. Available: <https://www.mongodb.com>
- [79] PostgreSQL. *The World's Most Advanced Open Source Database*. Accessed: Jul. 10, 2019. [Online]. Available: <https://www.postgresql.org/>
- [80] *HDFS Architecture Guide*. Accessed: Jul. 10, 2019. [Online]. Available: https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html
- [81] V. Cantoni, L. Lombardi, and P. Lombardi, "Challenges for data mining in distributed sensor networks," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, vol. 1, 2006, pp. 1000–1007, doi: 10.1109/icpr.2006.359.
- [82] E. Masciari, "A framework for outlier mining in RFID data," in *Proc. 11th Int. Database Eng. Appl. Symp.*, Washington, DC, USA, Jun. 2007, pp. 263–267, doi: 10.1109/IDEAS.2007.5.
- [83] A. Noulas, C. Mascolo, and E. Frias-Martinez, "Exploiting foursquare and cellular data to infer user activity in urban environments," in *Proc. IEEE 14th Int. Conf. Mobile Data Manage.*, Washington, DC, USA, vol. 1, 2013, pp. 167–176, doi: 10.1109/MDM.2013.27.
- [84] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with Noise," in *Proc. 2nd Int. Conf. Knowl. Discovery Data Mining*, Portland, OR, USA, 1996, pp. 226–231.
- [85] S. Yang, M. A. Cheema, X. Lin, and W. Wang, "Reverse k nearest neighbors query processing: Experiments and analysis," *Proc. VLDB Endowment*, vol. 8, no. 5, pp. 605–616, Jan. 2015, doi: 10.14778/2735479.2735492.
- [86] C. Zhou, P. Dai, F. Wang, and Z. Zhang, "Predicting the passenger demand on bus services for mobile users," *Pervasive Mobile Comput.*, vol. 25, pp. 48–66, Jan. 2016, doi: 10.1016/j.pmcj.2015.10.003.
- [87] A. J. Jara, D. Genoud, and Y. Bocchi, "Big data for smart cities with KNIME a real experience in the SmartSantander testbed," *Softw. Pract. Exper.*, vol. 45, no. 8, pp. 1145–1160, Aug. 2015, doi: 10.1002/spe.2274.
- [88] X. Du, O. Emebo, A. Varde, N. Tandon, S. N. Chowdhury, and G. Weikum, "Air quality assessment from social media and structured data: Pollutants and health impacts in urban planning," in *Proc. IEEE 32nd Int. Conf. Data Eng. Workshops (ICDEW)*, May 2016, pp. 54–59, doi: 10.1109/icdew.2016.7495616.
- [89] Z. C. M. Candra, H.-L. Truong, and S. Dustdar, "On monitoring cyber-physical-social systems," in *Proc. IEEE World Congr. Services (SERVICES)*, Jun. 2016, pp. 56–63.
- [90] D. E. Difallah, P. Cudre-Mauroux, and S. A. Mckenna, "Scalable anomaly detection for smart city infrastructure networks," *IEEE Internet Comput.*, vol. 17, no. 6, pp. 39–47, Nov. 2013, doi: 10.1109/mic.2013.84.
- [91] D. Puiu, P. Barnaghi, R. Tonjes, D. Kumper, M. I. Ali, A. Mileo, J. X. Parreira, M. Fischer, S. Kolozali, N. Farajidavar, F. Gao, T. Ilgen, T.-L. Pham, C.-S. Nechifor, D. Puschmann, and J. Fernandes, "CityPulse: Large scale data analytics framework for smart cities," *IEEE Access*, vol. 4, pp. 1086–1108, 2016, doi: 10.1109/access.2016.2541999.
- [92] Ö. Yılmaz and R. C. Erdur, "iConAwa—An intelligent context-aware system," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 2907–2918, Feb. 2012, doi: 10.1016/j.eswa.2011.08.152.
- [93] Open Connectivity Foundation. *OCF—Specifications*. Accessed: Jun. 24, 2019. [Online]. Available: <https://openconnectivity.org/developer/specifications>
- [94] *Web Thing Model*. Accessed: Jun. 24, 2019. [Online]. Available: <https://www.w3.org/Submission/wot-model/>
- [95] R. Szabo, "Framework for smart city applications based on participatory sensing," in *Proc. 4th IEEE Int. Conf. Cognit. Infocommun. (CogInfoCom)*, Dec. 2013, pp. 295–300, doi: 10.1109/CogInfoCom.2013.6719260.

- [96] B. Pan, Y. Zheng, D. Wilkie, and C. Shahabi, "Crowd sensing of traffic anomalies based on human mobility and social media," in *Proc. 21st ACM SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, New York, NY, USA, 2013, pp. 344–353, doi: [10.1145/2525314.2525343](https://doi.org/10.1145/2525314.2525343).
- [97] S. De, Y. Zhou, I. Larizgoitia Abad, and K. Moessner, "Cyber–physical–social frameworks for urban big data systems: A survey," *Appl. Sci.*, vol. 7, no. 10, p. 1017, Oct. 2017, doi: [10.3390/app7101017](https://doi.org/10.3390/app7101017).
- [98] D. Hussein, S. Park, S. N. Han, and N. Crespi, "Dynamic social structure of things: A contextual approach in CPSS," *IEEE Internet Comput.*, vol. 19, no. 3, pp. 12–20, May 2015, doi: [10.1109/mic.2015.27](https://doi.org/10.1109/mic.2015.27).
- [99] J. Yang, Y. Han, Y. Wang, B. Jiang, Z. Lv, and H. Song, "Optimization of real-time traffic network assignment based on IoT data using DBN and clustering model in smart city," *Future Gener. Comput. Syst.*, to be published, doi: [10.1016/j.future.2017.12.012](https://doi.org/10.1016/j.future.2017.12.012).
- [100] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Apr. 2015, doi: [10.1109/TITS.2014.2345663](https://doi.org/10.1109/TITS.2014.2345663).
- [101] A. Ghosh, T. Chatterjee, S. Samanta, J. Aich, and S. Roy, "Distracted driving: A novel approach towards accident prevention," *Adv. Comput. Sci. Technol.*, vol. 10, no. 8, pp. 2693–2705, 2017.
- [102] G. Amato, F. Carrara, F. Falchi, C. Gennaro, C. Meghini, and C. Vairo, "Deep learning for decentralized parking lot occupancy detection," *Expert Syst. Appl.*, vol. 72, pp. 327–334, Apr. 2017, doi: [10.1016/j.eswa.2016.10.055](https://doi.org/10.1016/j.eswa.2016.10.055).
- [103] L. Kuang, L. Yang, and Y. Liao, "An integration framework on cloud for cyber physical social systems big data," *IEEE Trans. Cloud Comput.*, to be published, doi: [10.1109/tcc.2015.2511766](https://doi.org/10.1109/tcc.2015.2511766).
- [104] A. Lehman, N. O'Rourke, L. Hatcher, and E. Stephens, *JMP for Basic Univariate and Multivariate Statistics: A Step-by Step Guide*. Cary, NC, USA: SAS Institute, 2005.
- [105] A. Komninos, V. Stefanis, A. Plessas, and J. Besharat, "Capturing urban dynamics with scarce check-in data," *IEEE Pervas. Comput.*, vol. 12, no. 4, pp. 20–28, Oct. 2013, doi: [10.1109/mprv.2013.42](https://doi.org/10.1109/mprv.2013.42).
- [106] *Home—CleanSpace*. Accessed: Jul. 17, 2019. [Online]. Available: <https://our.cleanspace/>
- [107] *Nest | Create a Connected Home*. Accessed: Jul. 17, 2019. [Online]. Available: <https://nest.com/>
- [108] A. Otebolaku and G. M. Lee, "A framework for exploiting Internet of Things for context-aware trust-based personalized services," *Mobile Inf. Syst.*, vol. 2018, Apr. 2018, Art. no. 6138418. Accessed: Jul. 19, 2019. [Online]. Available: <https://www.hindawi.com/journals/misy/2018/6138418/>
- [109] F. K. Ozenc and S. D. Farnham, "Life 'modes' in social media," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2011, pp. 561–570.
- [110] F. Lécué, S. Tallevi-Diotallevi, J. Hayes, R. Tucker, V. Bicer, M. Sbodio, and P. Tommasi, "Smart traffic analytics in the semantic Web with STAR-CITY: Scenarios, system and lessons learned in Dublin City," *J. Web Semantics*, vols. 27–28, pp. 26–33, Aug. 2014, doi: [10.1016/j.websem.2014.07.002](https://doi.org/10.1016/j.websem.2014.07.002).
- [111] (Jan. 17, 2018). *Internet of Things: The Top 5 IoT Trends for 2018—Hub:raum IoT Academy*. Accessed: Jul. 23, 2019. [Online]. Available: <https://medium.com/iot-creators/internet-of-things-the-top-5-iot-trends-for-2018-a2e244513a6e>
- [112] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009, doi: [10.1137/07070111x](https://doi.org/10.1137/07070111x).
- [113] L. Xiong, X. Chen, T.-K. Huang, J. Schneider, and J. G. Carbonell, "Temporal collaborative filtering with Bayesian probabilistic tensor factorization," in *Proc. SIAM Int. Conf. Data Mining*, Apr. 2010, pp. 211–222.
- [114] E. Papalexakis, K. Pelechrinis, and C. Faloutsos, "Spotting misbehaviors in location-based social networks using tensors," in *Proc. 23rd Int. Conf. World Wide Web (WWW)*, 2014, pp. 551–552.
- [115] D. M. Dunlavy, T. G. Kolda, and E. Acar, "Temporal link prediction using matrix and tensor factorizations," *ACM Trans. Knowl. Discov. Data*, vol. 5, no. 2, pp. 1–27, Feb. 2011.
- [116] X. Wang, L. T. Yang, J. Feng, X. Chen, and M. J. Deen, "A tensor-based big service framework for enhanced living environments," *IEEE Cloud Comput.*, vol. 3, no. 6, pp. 36–43, Nov. 2016, doi: [10.1109/mcc.2016.130](https://doi.org/10.1109/mcc.2016.130).
- [117] A. Gaur, B. Scotney, G. Parr, and S. McClean, "Smart city architecture and its applications based on IoT," *Procedia Comput. Sci.*, vol. 52, pp. 1089–1094, Jan. 2015.
- [118] S. K. Datta, R. P. F. Da Costa, J. Harri, and C. Bonnet, "Integrating connected vehicles in Internet of Things ecosystems: Challenges and solutions," in *Proc. IEEE 17th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2016, pp. 1–6, doi: [10.1109/wowmom.2016.7523574](https://doi.org/10.1109/wowmom.2016.7523574).
- [119] *OMNeT++*. Accessed: Aug. 6, 2019. [Online]. Available: <https://omnetpp.org/>
- [120] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO—Simulation of urban mobility," *Int. J. Adv. Syst. Meas.*, vol. 5, nos. 3–4, p. 89, 2012.
- [121] D.-T. Berlin. *Simulation of Vehicle-2-X Communication*. Accessed: Aug. 6, 2019. [Online]. Available: <https://www.dcaiti.tu-berlin.de/research/simulation/>
- [122] U. Brandes, "A faster algorithm for betweenness centrality," *J. Math. Sociol.*, vol. 25, no. 2, pp. 163–177, 2001.
- [123] A. Golubev, I. Chechetkin, K. S. Solnushkin, N. Sadovnikova, D. Parygin, and M. Shcherbakov, "Strategway: Web solutions for building public transportation routes using big geodata analysis," in *Proc. 17th Int. Conf. Inf. Integr. Web-Based Appl. Services*, 2015, p. 91.
- [124] A. Attanasio, E. Silvestri, P. Meschini, and G. Gentile, "Real world applications using parallel computing techniques in dynamic traffic assignment and shortest path search," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 316–321.
- [125] K. Lin, C. Li, G. Fortino, and J. J. P. C. Rodrigues, "Vehicle route selection based on game evolution in social Internet of vehicles," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2423–2430, Aug. 2018.
- [126] M. T. Abbas, A. Muhammad, and W.-C. Song, "Road-aware estimation model for path duration in Internet of vehicles (IoV)," *Wireless Pers. Commun.*, vol. 109, no. 2, pp. 715–738, Nov. 2019.
- [127] M. Banikhalaf, A. M. Manasrah, A. F. AlEroud, N. Hamadneh, A. Qawasmeh, and A. Y. Al-Dubai, "A reliable route repairing scheme for Internet of vehicles," *Int. J. Comput. Appl. Technol.*, vol. 61, no. 3, pp. 229–238, 2019.
- [128] Y. Sun, L. Wu, S. Wu, S. Li, T. Zhang, L. Zhang, J. Xu, and Y. Xiong, "Security and privacy in the Internet of vehicles," in *Proc. Int. Conf. Identificat., Inf., Knowl. Internet Things (IKI)*, Oct. 2015, pp. 116–121, doi: [10.1109/iki.2015.33](https://doi.org/10.1109/iki.2015.33).
- [129] Y. Lin, P. Wang, and M. Ma, "Intelligent transportation system (ITS): Concept, challenge and opportunity," in *Proc. IEEE 3rd Int. Conf. Big Data Secur. Cloud (Bigdatasecurity)*, *IEEE Int. Conf. High Perform. Smart Comput. (HPSC)*, *IEEE Int. Conf. Intell. Data Secur. (IDS)*, May 2017, pp. 167–172.
- [130] S. Tbatou, A. Ramrami, and Y. Tabii, "Security of communications in connected cars modeling and safety assessment," in *Proc. 2nd Int. Conf. Big Data, Cloud Appl.*, Mar. 2017, p. 56.
- [131] N. Lu, N. Cheng, N. Zhang, X. Shen, and J. W. Mark, "Connected vehicles: Solutions and challenges," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 289–299, Aug. 2014.
- [132] S. Fischer-Hübner, *IT-Security and Privacy: Design and Use of Privacy-enhancing Security Mechanisms*. Berlin, Germany: Springer-Verlag, 2001.
- [133] B. McQueen, *Big Data Analytics for Connected Vehicles and Smart Cities*. Norwood, MA, USA: Artech House, 2017.
- [134] D. Kwon, S. Park, and J.-T. Ryu, "A study on big data thinking of the Internet of Things-based smart-connected car in conjunction with controller area network bus and 4G-long term evolution," *Symmetry*, vol. 9, no. 8, p. 152, Aug. 2017.



ANSIF AROOJ received the bachelor's degree in information technology from the Punjab University College of Information Technology (PUCIT), in 2006, and the M.S. degree in information technology from the University of Management and Technology (UMT), Lahore, Pakistan, in 2014. She has joined the University of Education, Division of Science and Technology, Lahore, Pakistan as a Lecturer in computer science. Since 2009, she has been continuing teaching and involved in research activities. She has received the IEEE Travel Grant Award from Gwangju Institute of Science and Technology (GIST), South Korea, and the Young Researcher Award from Heidelberg Laureate Forum, Heidelberg, Germany. She has attended several conferences and workshops in national and international institutes. Her research interests are data mining, the Internet of Things, data science, cyber-physical and social networks, and the Internet of Vehicles.



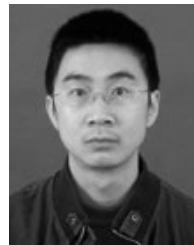
MUHAMMAD SHOAIB FAROOQ was born in Lahore, Pakistan. He received the M.Sc. degree from Quaid-e-Azam University, Pakistan, in 1995, the M.Phil. degree in computer science from Government College University, in 2007, and the Ph.D. degree from Abdul Wali Khan University, Pakistan, in 2015. He possesses more than 24 years of teaching experience in the field of computer science. He is currently an Associate Professor with the Department of Computer Science, University of Management and Technology, Pakistan. His research interests include the theory of programming languages, big data, the IoT, and computer science education. He is a member of the IEEE Systems, Man, and Cybernetics Society.



TARIQ UMER (Senior Member, IEEE) received the Ph.D. degree in communication systems from the School of Computing & Communication, Lancaster University, U.K, in 2012, and the master's degree in computer science from Bahauudin Zakariya University, Multan, Pakistan, in 1997. He served for IT Education Sector in Pakistan for more than 13 years. Since January 2007, he has been working as an Assistant Professor with the CS Department, COMSATS University Islamabad, Lahore Campus, Lahore. His research interests include vehicular ad-hoc networks, the Internet of Vehicles (IoV), the Internet of Things (IoT), wireless sensor networks, and telecommunication network design. He served in the TPC for the different International and national level conferences. He is the Active Member of the Pakistan Computer Society and the Internet Society Pakistan. He is currently serving as an Editor for *Future Generation Computer Systems* (Elsevier) and *IEEE ACCESS*. He is serving as a Reviewer for the *IEEE COMMUNICATIONS LETTERS*, *IEEE ACCESS*, *Computers and Electrical Engineering* (Elsevier), the *Journal of Network and Computer Applications* (Elsevier), *Ad Hoc Sensor Wireless Networks*, *Wireless Networks* (Springer) journal, and the *Journal of Communications and Networks*.



GHULAM RASOOL received the M.Sc. degree in computer science from BZU, Multan, Pakistan, in 1998, the M.S. degree in CS from the University of Lahore, in 2008, and the Ph.D. degree in reverse engineering from the Technical University of Ilmenau, Germany, in 2011. He is currently an Associate Professor and the Head of the Computer Science Department, COMSATS University Islamabad, Lahore Campus. He has more than 20 years of teaching and research experience at national and international levels. His research interests include reverse engineering, design patterns and antipatterns recovery, program comprehension, and source code analysis.



BO WANG received the Ph.D. degree in computer science from Wuhan University, Wuhan, China, in 2006. From 2007 to 2009, he was a Postdoctoral Researcher with the School of Electrical Engineering, Wuhan University, where he is currently an Associate Professor. His research interests include power system online assessment, big data, integrated energy systems, and smart city.

...