Computer Networks 101 (2016) 63–80

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Urban planning and building smart cities based on the Internet of Things using Big Data analytics

M. Mazhar Rathore[a], Awais Ahmad[a], Anand Paul[a,*], Seungmin Rho[b]

[a] The School of Computer Science and Engineering, Kyungpook National University, Daegu 702-701, Republic of Korea
[b] Department of Media Software, Sungkyul University, Anyang, Republic of Korea

ARTICLE INFO

Article history:
Received 31 July 2015
Revised 22 December 2015
Accepted 23 December 2015
Available online 11 January 2016

Keywords:
Big Data
Smart City
Urban Planning
IoT

ABSTRACT

The rapid growth in the population density in urban cities demands that services and an infrastructure be provided to meet the needs of city inhabitants. Thus, there has been an increase in the request for embedded devices, such as sensors, actuators, and smartphones, leading to considerable business potential for the new era of the Internet of Things (IoT), in which all devices are capable of interconnecting and communicating with each other over the Internet. Thus, Internet technologies provide a way of integrating and sharing a common communication medium. With this knowledge, in this paper, we propose a combined IoT-based system for smart city development and urban planning using Big Data analytics. We propose a complete system consisting of various types of sensor deployment, including smart home sensors, vehicular networking, weather and water sensors, smart parking sensors, and surveillance objects. A four-tier architecture is proposed that includes 1) Bottom tier-1, which is responsible for IoT sources and data generation and collection, 2) Intermediate tier-1, which is responsible for all types of communication between, for instance, sensors, relays, base stations, and the Internet, 3) Intermediate tier 2, which is responsible for data management and processing using a Hadoop framework, and 4) Top tier, which is responsible for application and usage of the data analysis and the results generated. The system implementation consists of various steps that begin with data generation and move to collection, aggregation, filtration, classification, preprocessing, computing and decision making. The proposed system is implemented using Hadoop with Spark, voltDB, Storm or S4 for real time processing of the IoT data to generate results to establish the smart city. For urban planning or city future development, the offline historical data are analyzed with Hadoop using MapReduce programming. IoT datasets generated by smart homes, smart parking weather, pollution, and vehicle data sets are used for analysis and evaluation. This type of system with full functionality does not currently exist. Similarly, the results demonstrate that the proposed system is more scalable and efficient than existing systems. Moreover, system efficiency is measured in terms of throughput and processing time.

1. Introduction

An emergent number of objects is being connected to the Internet at an extraordinary rate, comprising the knowledge of the Internet of Things (IoT). In 2008, CISCO reported that the number of things connected to the Internet surpassed the number of people living on earth,

* Corresponding author. Tel.: +82539507547.
E-mail addresses: rathoremazhar@gmail.com (M.M. Rathore), awais.ahmad@live.com (A. Ahmad), paul.editor@gmail.com (A. Paul), smrho@sungkyul.edu (S. Rho).

http://dx.doi.org/10.1016/j.comnet.2015.12.023
1389-1286/

whereas in 2020, it will reach the limit of 50 billion, resulting in the enrichment of the digital world [1]. There is a different domain in which IoT plays a vital role and improves the quality of human life. People are also now using capillary devices in IoT for health applications [2]. Similarly, there are many other domains in which IoT facilitates human life in a noteworthy way, including healthcare, automation, transportation, and emergency responses to manmade and natural disasters, under which circumstances it is difficult to make decisions.

IoT empowers an object to hear, see, listen and communicate at the same time. Thus, IoT transforms those objects from being traditionally smart by incorporating its ubiquitous and pervasive computing, embedded devices (e.g., actuators, smartphones, tablets, and other networked-enabled devices), communication technologies, sensor networks, Internet protocols and applications to revolutionize human life. The Internet will be no longer considered a network of computers. However, it will involve billions of smart devices along with embedded systems. As a result, the Internet of Things (IoT) will significantly increase in size and scope, providing new opportunities as well as challenges [3]. The majority of countries have formed longstanding national strategies for the implementation of IoT after completing the intangible stage of service level. For instance, Japan's broadband access facilitates communication between people, people and things, and between things and things [4]. Similarly, South Korea's smart home enables their citizens to access things remotely [5]. Singapore's next generation I-Hub [6] intends to comprehend the next generation "U" type network through a secure and ubiquitous network [7]. The stated initiatives laid the foundation of IoT [8]. Moreover, the efforts in Tag free activity sensing using RFID [9], evidence theory [10], and mobile ad-hoc social networking [11] leads us towards advancement in the IoT.

The IoT is considered the next large prospect for the world of the Internet. Thus, this leads us to the concept of smart homes where different electronic appliances are interconnected with each other and achieve high-quality two-way interactive multimedia services. In such a system where a large number of devices are communicating with each other, a massive volume of data (called Big Data) is generated. To enrich smart home technology, the better analytics of Big Data could play a vital role in the advancement of Information and Communications Technologies (ICTs). This type of Big Data analysis provides a better understanding and useful information about the future as well as about planning and development, thus providing us insight into Big Data.

Otherwise, to make the IoT more appealing, traditional applications can be considered, i.e., a smart home where embedded devices, such as sensors and actuators, are self-configurable and can be controlled remotely with the help of Internet technology. This type of technology is used to enable a large variety of security as well as monitoring applications. A large number of the devices involved sense the surrounding activities and transmit a massive amount of data to the remote station where it can be processed, analyzed, and predict or give a response to the user for his/her convenience based on the received data. In the literature, extensive research has been performed on smart home technology [12]. This research focused on individual homes. Similarly, the idea of the smart home is extends to the Smart Community where the Home Domain, Community Domain, and Service Domain are integrated to benefit people. However, this technology is lacking in various factors, such as how to connect vehicles, roadside units, GPS, and others to the same infrastructure, i.e., the web.

Seventy percent of the world's population (more than six billion) will live in cities and neighboring regions by 2050 [13]. With this massive volume of the population, billions of devices will also communicate with each other, thus producing overwhelming Big Data. Hence, in analyzing the data based on user needs and choices, cities would become even smarter. Consequently, powered by the variation of enabling technologies and their data analytics, the IoT has come out of its early stages and is entering into the era of revolutionizing the traditional network infrastructure into a fully integrated future with the Internet. The Wireless Sensor Network (WSN) and its related technologies are flawlessly unified into an urban infrastructure, establishing a digital skin [13]. The massive amount of information generated by the embedded and pervasive devices will be shared across assorted platforms and applications to enrich cities and predict planning and development.

Traditionally, for urbanization, it is of utmost importance to comprehend the demand for service profiling to enhance efficiency and may advance city management. Presently, few organizations are on their way with their platforms to live monitoring, planning and gathering urban process parameters. These activities are followed by collecting data offline and real-time, Big Data processing and analysis, and decision making. Usually, data collection techniques are costly and difficult. Therefore, there is a need to incorporate smart technology that can efficiently and quickly collect a vast amount of data, perform analyses on Big Data, and predict the future to facilitate better planning and development [14, 15].

In understanding the feasibility and potential of the IoT and the smart home, in this paper, we propel the concept of the smart home to the smart city with the idea of urban planning and development based on Big Data analytics. In the paper, we propose a complete architecture to develop the smart city and conduct urban planning using IoT-based Big Data analytics. The 4-tier architecture is proposed, which has the capability to analyze the large amount of IoT datasets generated from various sources of the smart system in the city, such as smart homes, smart car parking, vehicular traffic, and others. In addition, the complete system implementation model guides various municipalities to implement the system. Moreover, the analysis is performed on the IoT datasets to make smart city decisions using the proposed system. Finally, the system is tested and evaluated with respect to efficiency measures in terms of throughput and processing time.

2. Motivation

As mentioned earlier, smart cities become smarter due to the enriched nature of digital technology, in which the smart city is equipped with different electronic equipment utilized by the various applications, such as street cameras for the surveillance system, sensors for the transportation system, and so on. However, there are also initiatives that use objects to provide different value-added services, such as Google street view, the global positioning system (GPS), and others. Furthermore, this extends to the usage of individual mobile devices, contributing to the abovementioned scenario. Consequently, in this heterogeneous environment in term of objects' features, contributors, motivations, security rules, and so on, different queries arise in a city environment that must be addressed [16]. These are as follows:

- How to tackle uncertainty induced due to real-time and offline dynamics and ensure the quality of information.
- How to make existing objects smarter. Alternatively, how to design new objects to be smarter based on user choice.
- How to enable objects to react accordingly with respect to context.
- How to minimize the cost of data collection that is being generated by some devices.
- How to obtain insight into the data if the data are collected and going to the processing stage in real time.

Based on the above questions, the smart city concept utilizes ICT in a way that could help citizens in daily life using limited resources. Moreover, various organizations aim to develop a system that uses advanced technology by providing efficient services to their citizens. The majority of these recent technologies consist of advanced sensing capabilities, storage capability for an unprecedented volume of data, and finally, an insight into the voluminous data.

The rationale behind our intentions is to enrich the vast deployment of ICT resources in developing the entire system. For this reason, we know that the advancement of recent technology in the embedded system depicts the trends of ICT. Therefore, a system is required that could inhale all of the recent developments in the field of ICT, due to which remarkable growth can be observed in the near future. The design of this system requires all of the capabilities of sensing the environment and analyzing the sensing information. Therefore, various real-time actions could occur due to these technological resources. Moreover, it can be observed that integrating a large amount of data to perform an efficient analysis is already being performed as best as possible. However, with a large-scale environment, it is unavoidable that a large portion of the data is left disjointed. As a result, such data cannot provide us a better understanding of the situation so that we may plan for the future. For this reason, urban planning and development provides a new way to the field of the IoT, in which devices are integrated by means of their geographic location and are then analyzed by means of a newly designed system for various services in a city.

Because urban planning and development applications can benefit from a smart city, IoT capabilities can be grouped into impact areas [13]. This includes the effect on the citizen in terms of health and safety, the transportation system in terms of mobility and pollution, and so on. Various projects related to the monitoring of cyclists, cars, public car parking, and so on are occurring that utilize sensor services for the collection of specific data. Apparently, other service domain applications have been identified that utilize a smart city IoT infrastructure to provide operations in air and noise pollution, vehicle mobility, and surveillance systems in cities. The recent research consists of very few research findings in the field of the smart city as well as in urban areas. Similarly, a compact system has not yet been built that is scalable and efficient. Big Data are used to analyze different aspects of the smart city and then uses the knowledge obtained from past generated data for the betterment of cities. A similar concept follows using the IoT paradigm and the Big Data concept for urban planning. Thus, we attempted a solution that can be used in the smart city as well as in urban areas. The proposed system is implemented and tested on the Hadoop framework with Spark to obtain real-time effects in the case of real-time smart city decisions. Moreover, Hadoop and MapReduce is used for large historical data for urban planning and future enhancements.

3. Urban planning and development smart cities based on IoT

The key concept of the smart city is to obtain the right information at the right place and on the right device to make a city-related decision with ease and to aid citizens more quickly. To develop the IoT-based smart city concept and urban planning system, we deployed several wireless and wired sensors, surveillance cameras, emergency buttons in streets, and other fixed devices. The main challenge in this regard is to achieve a smart city system and link IoT information together. We do this by providing relay nodes, aggregation classifiers, and so on. Moreover, all sensors generate abundant data with high speed, which is called Big Data. To process that data in an efficient way, the Hadoop system is employed. In this section, we provide a complete architecture of how the sensors are deployed and generate data. Similarly, we proposed an IV-tier architecture and system implementation to clearly show the workings of the proposed system.

3.1. IoT-based smart city

One of the core challenges is how IoT can be used to established and build a smart city. IoT is the interlinking of heterogeneous devices with each other together over the Internet. We are moving towards the digital era in our homes, cities and so on; therefore, the devices that are available in our homes and surroundings should be linked to the Internet for fast accessibility. In order to achieve our target, we deployed many sensors at different places to collect and analyze data for better usage. The ultimate goal is to achieve smart homes, parking, weather and water systems, vehicular traffic, environmental pollution and surveillance systems.

In a smart home, the home is continuously monitored by sending data generated from the sensors measuring the smoke and temperature. Similar to detecting a fire in real time, electricity and gas companies effectively manage the power, gas, and water consumption to houses and different areas of the city. Monitoring pollution helps in the health care of the citizens and alerts them when the pollution increases over a particular threshold.

Smart parking helps by tracking the vehicles coming and going out of different car parking zones. Thus, smart car parking can be designed considering the number of vehicles in a region, or new car parking can be developed where there are more cars overall. Similarly, smart car parking data facilitates citizens and merchants daily lives in a smart city. Citizens can easily obtain information on the nearest free slot of parking or more suitable places to park their vehicles. This system reduces the fuel consumption of vehicles. Moreover, other applications may include saving time, as a person would be able to spend more time in a marketplace or doing other activities than parking.

Weather and water information also increases the efficiency of the smart city by providing weather-related data such as temperature, rain, humidity, pressure, wind speed and water levels at rivers, lakes, dams, and other reservoirs. All of this information is collected by placing the sensors in water reservoirs and other open places. In the world, most floods occur due to rain and a few from snow melting and dam breakage. Therefore, we can use rain measuring sensors and snow melting parameters in order to predict floods earlier. We can also predict information pertaining to water reservoirs in advance to meet the citizens' water needs.

Vehicular traffic information is the most significant source of data in a smart city. Through this type of data source and with useful real-time analysis, citizens and the government will benefit greatly. Citizens can determine how long it will take to reach a destination based on the current intensity of traffic and the average speed of the vehicles. The traffic can be diverse in all cities, and knowing traffic patterns will reduce fuel consumption as well as decrease pollution occurring from heavy traffic. Government authorities can also obtain real-time information about road blockages due to accidents or other issues and can thus take necessary action to manage traffic. In our smart city system, we obtain traffic information by GPRS and vehicular sensors, as well as sensors placed on the windshield of the car. We obtain the location of each vehicle and the number of vehicles between two pairs of sensors placed at various locations in the city. Moreover, if any accident has happened, the windshield will be damaged and the sensor will send an alert to the police, traffic authorities, and the hospital. Additionally, other real-time actions will make this process more efficient in the future.

Moreover, for people with health care conditions, monitoring environmental pollution and delivering this information to those people is also vital. A city can never be smart with unhealthy citizens. Therefore, while designing the smart city, we designed a separate module to obtain environmental data that includes information on gases, such as particular metals, carbon monoxide sulfur di-oxide, and ozone, as well as noise. These gases are very dangerous to human health and can cause liver disorders, coughing, and heart disease. People should not go outside when these gases are at high levels in the environment, especially children, the elderly, those who are engaging in physical exercise, or the sick. This will only be possible when there is real-time access to all of this information and when alerts are generated when any of the gases exceeds a particular threshold. Moreover, in more populated places, the government should attempt to reduce the causes of pollution, such as by moving industries to other areas, diverting traffic to other routes, and so on.

Last but not least, the most important thing for the citizens of a smart city is security. Security is achieved by the proposed system through continuous video monitoring of the whole city. However, it is very challenging to analyze the video and detect crimes through the system. To overcome this limitation, we propose new scenarios that increase the entire city's security system. We placed various emergency buttons, including microphones, in various places with surveillance cameras. When any crime occurs, such as a robbery, a car or purse being stolen, a fight, or illegal activity, a witness can push the emergency button, and it will send a message to the nearest police station. Thus, police or security agencies can start monitoring nearby locations through surveillance cameras and can easily locate the perpetrator. Moreover, the information collected from different sensors can be used to avoid future security issues, providing a more secure environment to the citizens of the proposed smart city.

The complete IoT objects deployment is shown in Fig. 1. There is one aggregation server that collects and aggregates the data from all smart systems. The data are received with high speed. Therefore, the aggregation process is powerful enough to aggregate the data and send it for analysis through IoT systems.

3.2. IoT-based urban planning

For urban planning, the same IoT scenario is considered with the same devices and sensors as shown in Fig. 1. The only difference in the urban planning system is the use of sensor-generated data and the purpose of analysis. In a smart city, we perform real-time decision making on real-time data. In urban planning, we use previous historical data generated from the same smart city's IoT devices and plan for the future. For example, by analyzing electricity consumption from previous years, we predict the demand for next year and take necessary action to fulfill the demand.

Using smart home-generated data, government authorities can analyze previous energy consumption data and growing needs and build new dams to produce more energy. Moreover, they can also analyze the pattern of energy usage at different periods and manage electricity and gas bills to help citizens. Additionally, they can make energy plans for various periods of the year. In terms of smart parking and vehicular traffic–generated data, a need for new parking lots and new buildings or places to build new roads or extend roads can be predicted for the future. Data on increases or decreases in pollution due to traffic changes is analyzed for the causes of the pollution and used to plan accordingly. Similarly, analyzing weather and water consumption data sets, we can plan for agriculture, safety from floods, safe drinking water, and so on. Based on temperature data and electricity consumption, we can better plan for high-temperature seasons to reduce

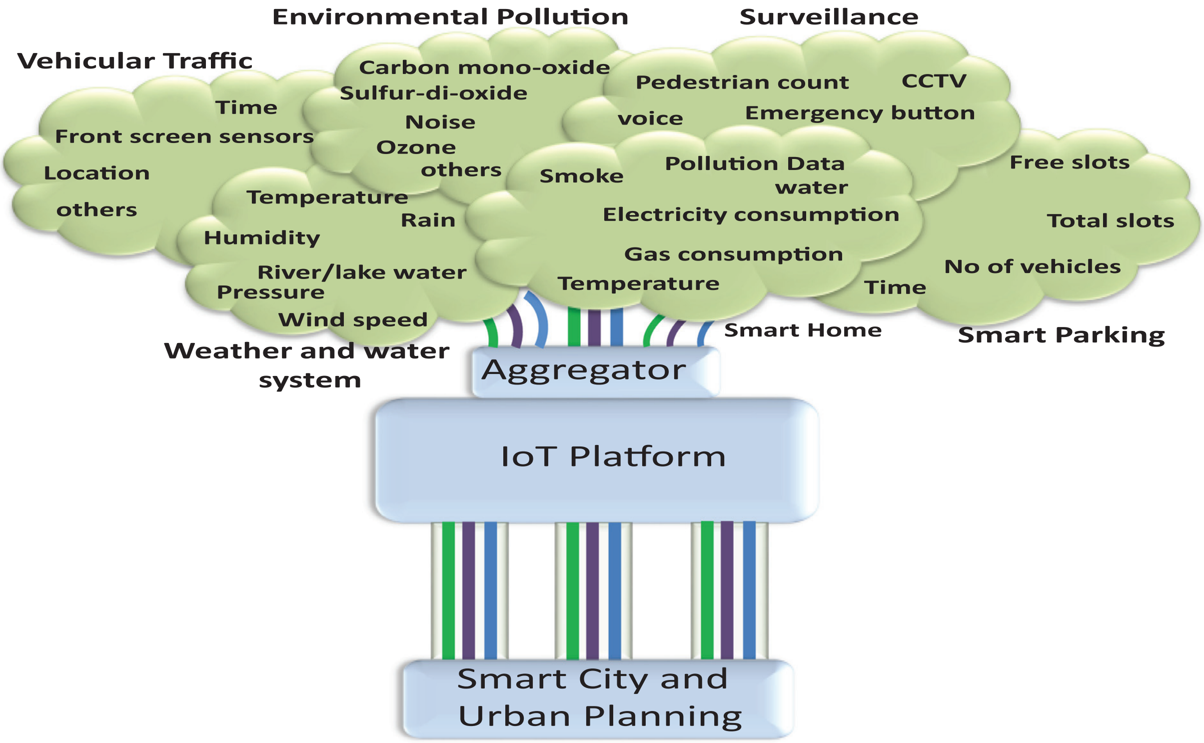

Fig. 1. Sensors deployment.

the consumption of electricity. Similarly, from surveillance data sets, we can analyze the number of criminal events, as well as determine which places are most dangerous, in which places the most people are affected, and where crime is spreading; based on these data, we can prepare locations with higher security for the next year or even for next month.

3.3. The big data analytical architecture and implementation model

Based on the needs of the smart city and urban planning, we propose a 4-tier architecture to analyze IoT Big Data to establish smart cities. The complete architecture is shown in Fig. 2; the 1st tier is the bottom tier, then the two intermediate tiers, and finally the top tier. The functionality of each tier is described below:

Tier 1. Bottom tier: this layer handles data generation through various IoT sources and then collects and aggregates that data. Because many IoT sensors participate in the generation of data, there is a significant amount of heterogeneous data produced with varying formats, a different point of origin and periodicity. Moreover, various data have security, privacy, and quality requirements. Additionally, with regard to sensor data, the metadata are always greater than the actual measure. Therefore, early registration and filtration techniques are applied at this layer, which filters the unnecessary metadata, and repeated data are also discarded.

Tier-II. Intermediate tier-I: this tier is responsible for the communication between sensors and from sensors to relay nodes through ZigBee technology, and relies on GW or base station and then on the Internet using various communication technologies, such as Wi-Fi, WiMAX, LTE, 3G, etc. Ethernet is used between various analysis servers.

Tier-III. Intermediate tier-II: this layer is the main layer of the entire analytical system and is responsible for data processing. We need real-time analysis for the smart system; therefore, we need a third party real-time tool to combine with Hadoop to provide real-time implementation. To provide real-time implementation, Strom, Spark, or VoltDb could also be used. However, for system evaluation, we implemented the system with Spark. At the lower layer of Hadoop, the same structure of MapReduce and HDFS is used. With this system, we can also use HIVE, HBASE, and SQL for managing Database (in-memory or Offline) to store historical information. For urban planning because real-time results were unimportant, we used Hadoop with MapReduce programming.

All data will be stored in Hadoop using HDFS, and analyses are performed at intermediate tier-II. The last tier is the interpretation tier, which is the usage of the results of analyzed data and the generation of reports. Here, the generator results are announced and used for many applications, such as flood detection, security, and city planning.

We also designed an implementation model of the system, shown in Fig. 3, that outlines the complete details of all the steps performed while implementing the system. Initially, every system generates their data, such as smart home-generated data, vehicular data, smart parking data, and so on. In every system, there is a relay node that is responsible for collection of data from all the sensors

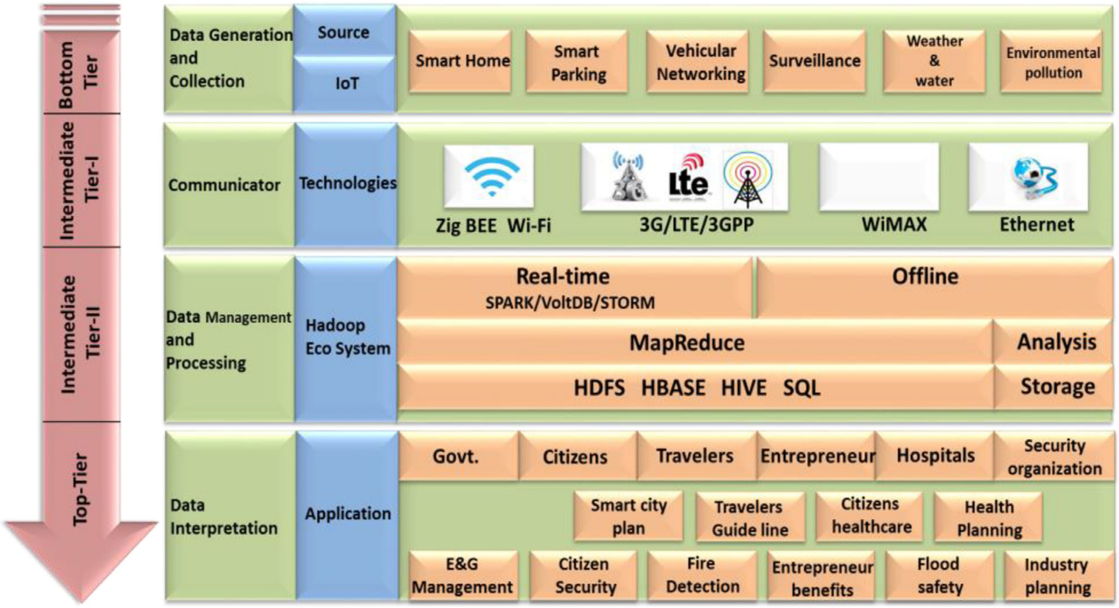

Fig. 2. IV-tier architecture for IoT Big Data analytics for remote smart city and urban planning.

in the system. It uses ZigBee technology to communicate with the sensors. The relay handles collecting data from all sensors and then sends the data to the analytical system through GW and the Internet. The sensors contain a significant amount of metadata; therefore, all of the unnecessary metadata and redundant data are discarded. Moreover, the data are classified by message type and identifier. After classification, the classified data are converted to another format, i.e., it is understandable to the Hadoop ecosystem, such as a sequence file.

Because we are dealing with a large amount of data (Big Data), we need a system that can efficiently process a large set of huge datasets. To meet these requirements, we used the Hadoop ecosystem, which contains master nodes and various data nodes under the master node. The Hadoop ecosystem has HDFS file storage, which divides the data into an equal amount of chunks and stores them on various data nodes. Later, the parallel processing is performed on these chunks using the MapReduce system. All processing calculations and results generation are performed in the Hadoop ecosystem. Finally, decision making is performed based on the results generated by the Hadoop ecosystem. The decision-making approach uses machine learning, pattern recognition, soft computing and decision models.

4. Urban data analysis and discussion

To perform the feasibility study and understand the importance of the system, a detailed analysis is performed on various IoT datasets. The analysis is performed to show that how a smart city can be built by using the proposed system, how the deployment of sensors matters for building a smart city, and how we can use historic sensor data to perform Big Data analytics for urban planning. This section also illustrates how we can use the same IoT generated data for real-time decision making to make a city smarter as well as performing offline analysis on historical data to conduct urban planning. In this section, we describe the details of the datasets used for analysis as well as for evaluation and discussion on analysis to establish the smart city and perform urban planning for the future.

4.1. Dataset description

We take large IoT-generated datasets from various reliable resources. The datasets include 1) the data of floods occurring all over the world, 2) smart home temperatures, including the water usage of each house and so on, 3) vehicular datasets, including all details of the vehicles traveling between many source and destination points at various places in the city, 4) parking place datasets, including the current status of number of vehicles in the parking area, 5) pollution datasets, including various gases and noise pollution, 6) social media datasets, such as Twitter, including daily tweet records, 7) weather datasets, including continuous measurements of temperature, humidity, rain, and so on, outside and inside the home, 8) other data common city datasets, such as cultural and library events. Complete dataset details, including dataset size, the number of parameters, and the source, are shown in Table 1.

Brakenridge [17] generated the flood dataset by collecting news from official and TV news channels of the flooded country. The data contains the date of flood, area of flood, damage, intensity, death, and so on.

Water usage data for each household of Surrey, Canada, was taken for household analysis. A total of 61,263 houses' water meter readings were measured. The data contain the complete address and water usage of the house. The 3rd dataset that we analyzed for smart city and urban planning is the vehicular traffic on the Madrid Highway. This dataset

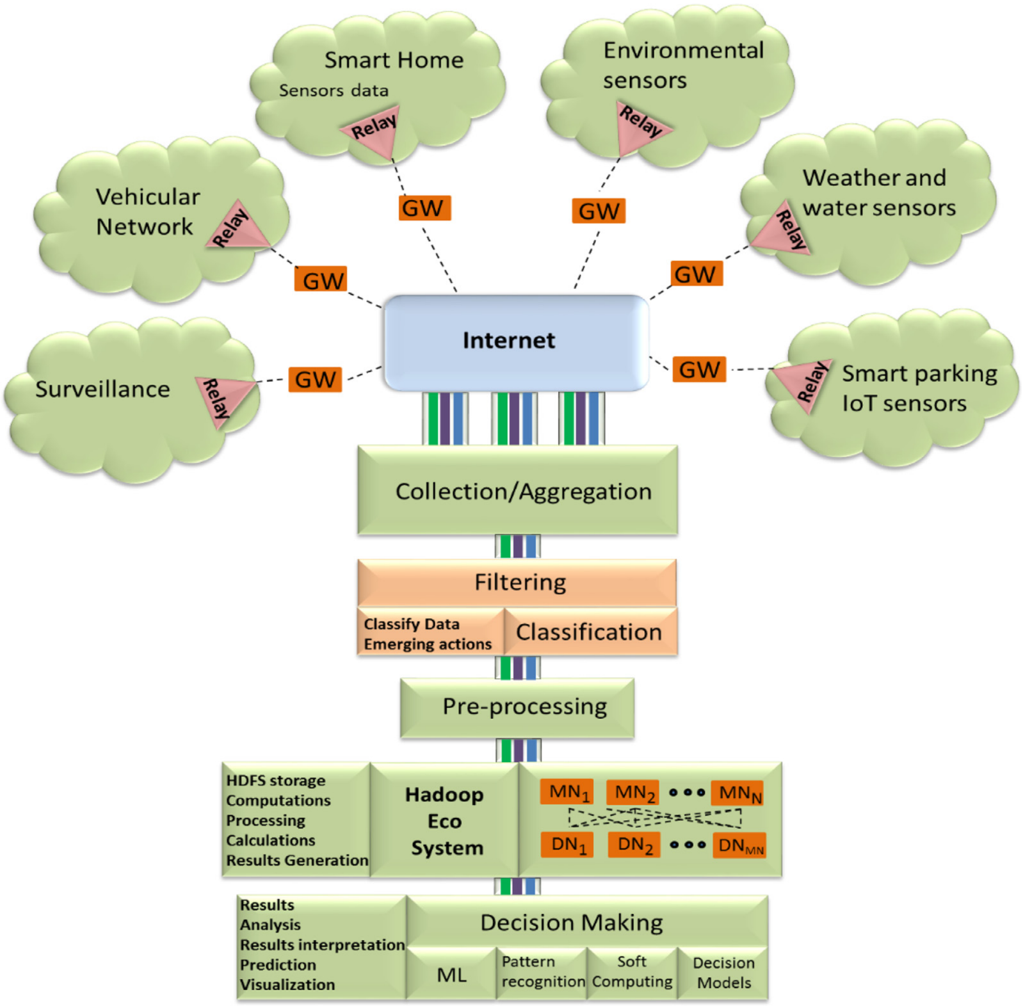

Fig. 3. Implementation model.

Table 1
Dataset details.

S#	Datasets	Size	No. of parameters	Source
1	Floods	16MB	30	[17]
2	Water usage	5MB	11	[18]
3	Madrid Highway vehicular traffic	450MB	5	[19]
4	Vehicular mobility traces	4.03GB	5	[20–22]
5	Parking lots	294KB	7	[23–25]
6	Pollution	32GB + 570MB	8	[23–25]
7	Social network (twitter)	8 + 8MB	7	[23–25]
8	Aarhus city traffic	33GB	9	[23–25]
9	Weather	3MB	7	[23–25]

is more important for smart cities to facilitate daily life as well as for urban planning in constructing new roads and buildings. It contains the location of each vehicle between the two ends of the Madrid Highway as well as the speed of the vehicle. We also tested the vehicular mobility dataset generated by the Institute of Transportation Systems, German Aerospace Center (ITS-DLR) as the TAPAS-Cologne project containing the mobility of all the cars in Cologne, Germany, and covering 400 km^2 in 24 h with 700 cars. Next, all other datasets covered Aarhus, Denmark. The parking lot dataset covers the continuous monitoring of eight parking lots of the city with respect to usage. It contains data from May 22, 2014, to Nov 4, 2014, by capturing data from 55 points. The pollution datasets and Aarhus city vehicular datasets were generated by placing sensors at the same location at the same times to find the effects of traffic on the environment. Both datasets contain various periods of data from 2014 and generate data by placing 449

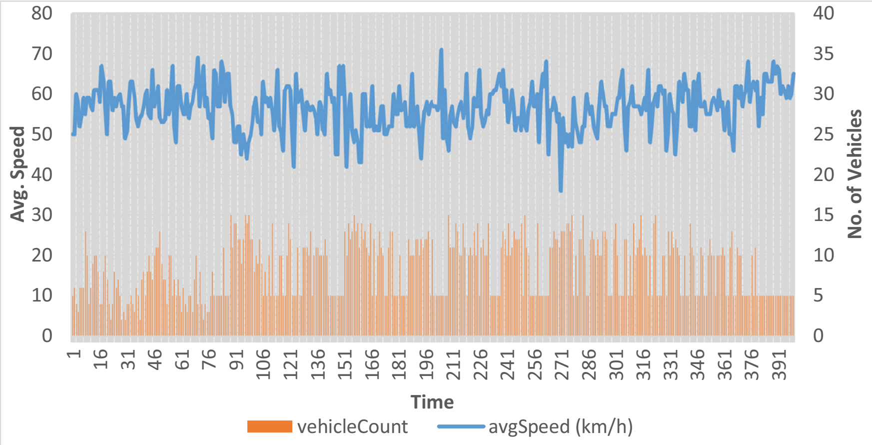

Fig. 4. The speed of vehicles at low intensity of traffic between two points.

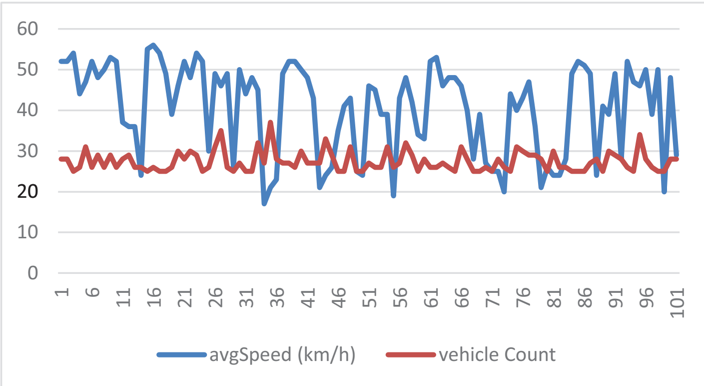

Fig. 5. Speed of vehicles at high traffic intensity between two points.

sensors at different locations in the city. For vehicular data generation, they placed source and destination pair sensors in different locations to estimate the traffic between two points that contained various information about the average speed of vehicles between two points, average speed, and time to reach the 2nd point. The pollution data had various measures, including ozone, nitrogen dioxide, nitrogen oxide, particle matters, carbon dioxide, and so on. Moreover, social network data are also important for smart city real-time decisions and urban planning. Thus, we took into account Twitter data that included tweets including city, location, and time. Twitter data contains information from 13,674 tweets from September 23, 2013, to December 17, 2013. Finally, weather data consisting of temperature, humidity, rain, pressure, and wind were also considered for analysis and evaluation, covering the period from February to June 2014 and August to September 2014.

4.2. Analysis and discussion

The main challenge in smart city development is the analysis of real-time data to enable urgent actions. Enabling smart cities not only benefits the government but also the citizens, as it helps citizens to save fuel by efficiently managing routes to reach destinations as well as to protect themselves from higher levels of environmental pollution. Here, we analyze various types of data and direct authority as to how they can use IoT technologies and Big Data generated from IoT for smart cities and urban planning. We present an analysis of vehicular traffic, parking lots, smart home of water usage by each house, flood, and pollution in Aarhus.

4.2.1. Vehicular traffic analysis

As a use case scenario, we use publically available traffic data from Aarhus, Denmark, which contains information on geographical location, timestamp, and traffic intensity, such as average speed and vehicle count. Moreover, we also used the vehicular datasets of Madrid City as we mentioned earlier. The analysis of Aarhus city traffic is presented only by the data taken from the two sensors placed at 1-km distance in the “A rhusvej” street of Hinnerup.

The number of vehicles in a particular area plays a vital role in society. For instance, during on hours, the traffic intensity on particular roads is higher than during off times. Similarly, the road management system can be affected by the number of vehicles at a particular time and on a particular road. In Figs. 4 and 5, we carefully analyzed

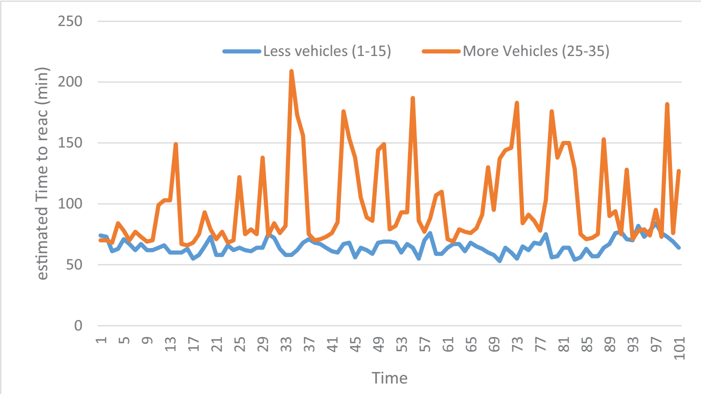

Fig. 6. Estimated time to reach the destination depending on traffic intensity.

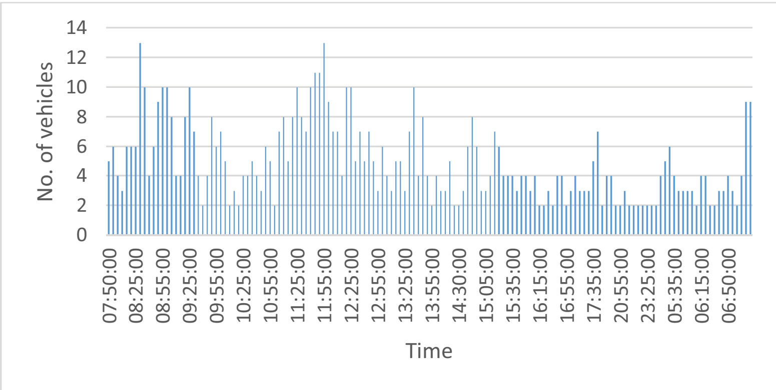

Fig. 7. No. of vehicles between sources and destination pairs at various times of day.

the traffic intensity on different roads in a society. For instance, if vehicle speed is low on some roads, it means that the intensity of the cars is high on those roads. Moreover, in Fig. 4, when the number of vehicles is higher, for example, at 106 and 121, the vehicle speed is less, for instance, 45 and 42. Therefore, maintaining this relationship between vehicle and vehicle speed, we can design roads for better vehicular management. Similarly, in Fig. 5, the number of vehicles is between 25 and 35, which is considered high-intensity traffic. We can see that when the number of vehicles is high, for example, at 37, the vehicle speed decreases to 18. Thus, the statistics in Figs. 4 and 5 can be used to design wide roads where the intensity of vehicles is high and vice versa when planning for the future.

In Fig. 6, two types of traffic classes are used, i.e., 1–15 and 25–35 cars. We performed an experiment testing the movement between two points. We begin by assuming a car is moving from point A to point B on the road with the number of cars being between 1 and 15. The figure shows that the time required for the car to reach its destination is less compared to the same road with the number of cars being between 25 and 35. This estimation is taken in real time with the average speeds of the cars on the roads. Thus, we can design wider roads in those areas where the intensity of cars is high. For example, if on a road the number of schools, colleges, universities, and so on is high, then using statistics, we can consider designing a wider road. Similarly, in areas where the number of buildings is lower, the roads can be designed with fewer lanes. However, we are avoiding scalability for now and will consider it in our future work.

In Fig. 7, we test the intensity of vehicles along a road at different durations. For example, we can see from the graph at 08:25 and 11:55 that the number of vehicles is very high, i.e., more than 12. Thus, an efficient road system can be designed that can dynamically change routes during the rush hour time. Similarly, sensors can be installed at different locations that can communicate with vehicles in the case of accidents and congestion. Thus, various conclusions can be drawn from the statistics of Fig. 7. For example, the engineer can be provided with better information about road design and construction.

From the above IoT-based network traffic analysis, we can predict the estimated time to reach one point from the other point. The smart city analyzes vehicular traffic data in real time and facilitates finding how much time it will take to reach a destination by following alternative routes depending on the current intensity of the traffic. It provides updated information regarding travel so that people can plan to reach the destination by following the most convenient route. Moreover, it helps governmental traffic authorities to control traffic and create an optimized plan

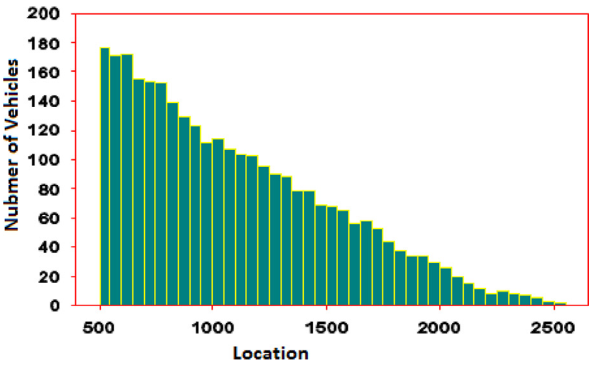

Fig. 8. Intensity of traffic at various locations on the Madrid Highway.

when the intensity of traffic becomes higher or the road is blocked due to accidents, strikes, and so on. This type of traffic management not only helps citizens and the government in saving fuel but also provides safety from pollution generated by an abundance of traffic at a single point. Thus, the smart city aids in the diversion of traffic from busy roads to free roads to obtain equal usage of all alternative roads.

In the next phase of vehicular traffic analysis, a slightly different dataset covering Madrid is addressed. The traffic intensity of the first 2500 locations for a particular time is shown in Fig. 8. The figure shows the congested locations where the intensity of the traffic is higher. We can easily observe that at the starting position, there are more vehicles, and when we move forward, the number of vehicles begins to reduce. The figure shows that location 500 is the central location, where most of the vehicles pass through. On the other hand, location 2500 is very far from the city, where fewer cars are. Therefore, on the basis of this analysis, we can plan for the road by building more lanes where the traffic is more intense. We can also assume that at location 2500, the number of people, houses, shops, and buildings is lower. Therefore, we can plan to build more houses and buildings there to reduce the burden of traffic, pollution, and crown.

For the data regarding traffic in Madrid, we have also analyzed the speed of the vehicles on the highway. The average speed is 90 km/h. When measuring speed, we can estimate the conditions of the road by identifying regions where the speed is lower due to poor road structure or road damage. Similarly, in a smart city, we can identify speed violations and charge for them accordingly. We identified regions where most of the vehicles cross the maximum speed limit, shown in Fig. 9. On the Madrid Highway, most of the vehicles crossed the maximum limit from location 5000 to 1100. Most of the violations occurred in lane three (the fastest lane) of the highway. These violations may be due to a lower number of vehicles on the road. This can be addressed by signage or by placing speed barkers in that location, whichever is more suitable. This may be a better option for the smart city and urban planning as well. Moreover, in the smart city, the accident ratio is also monitored with respect to the speed and violation data for the area.

4.2.2. Use of parking lot data analysis

By analyzing the current usage of parking lots, citizens can select the closest parking lot to their location. Fig. 10 shows the number of free spaces in various parking garages in Aarhus, and Fig. 11 shows the current use of parking garages. According to this study, users are updated about free parking in real time and can thus save fuel without manually searching for free parking. Moreover, this also creates a profit equilibrium for vendors in the city by benefitting shop owners who not making much profit. Generally, citizens prefer to go shopping where it is less congested and where parking is readily available, resulting in greater profit for vendors in such areas. The parking analysis also provides direction to governmental authorities in urban planning to build more parking areas near places with a higher volume of consumers. Fig. 11 shows that the Bruuns has ample parking with a capacity of 931 cars, but parking is still generally unavailable there. This conveys the need for more parking lots at that location. Similarly, we obtain the same results through an analysis of vendors operating near the garage.

4.2.3. Smart home data analysis

While analyzing the smart home data, one use case was taken into consideration to analyze the current usage of water consumption in each house. This helps smart cities to manage their water resources with respect to the current data usage, and next year's water needs can also be

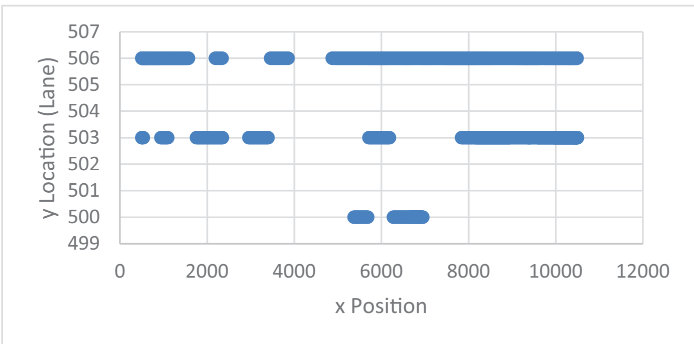

Fig. 9. Location of speed violations on the Madrid Highway.

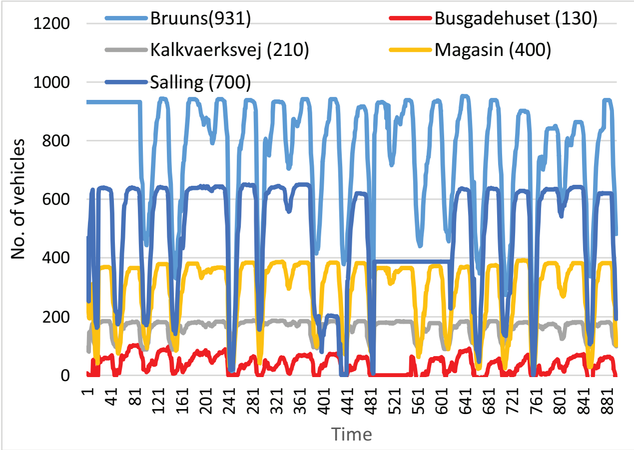

Fig. 10. Free spaces at various parking lots at different times.

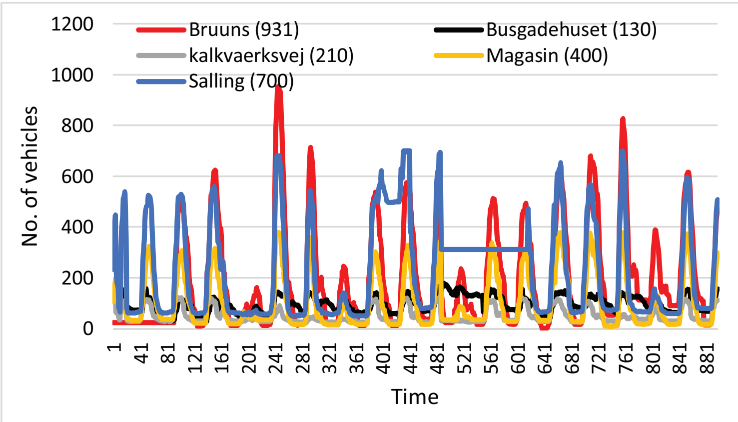

Fig. 11. Usage of various parking lots at different times.

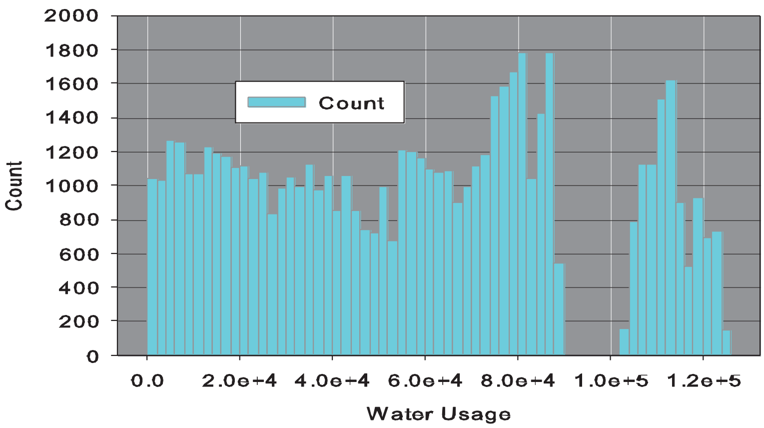

Fig. 12. Total water usage for Surrey.

predicted. Moreover, the flow of water to various areas depending on the needs of the area can also be controlled. The water consumption of each house in Surrey, Canada, was analyzed for that purpose. Fig. 12 shows a histogram of water usage by cubic meter for all houses in the city. It shows that more than 6000 houses consume more than 8000–9000 cubic meters of water, which is the normal water usage maximum houses. This study can help authorities to determine water billing rates based on normal water usage.

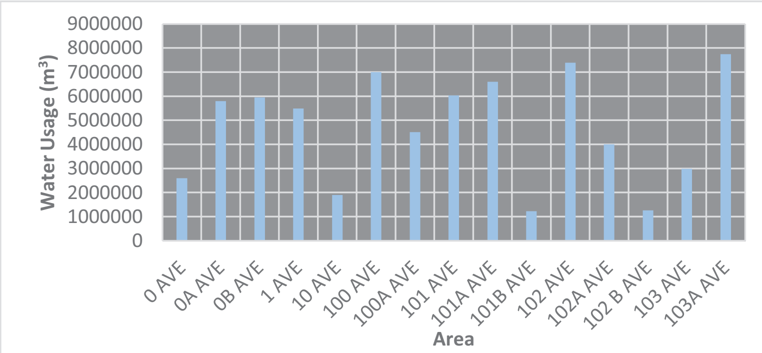

Fig. 13. Water usage of various areas of Surrey City.

Table 2
World Flood Report from 1985–2014.

Flood type	Total floods	Duration	Total deaths	Total ($M > 4$)	Total ($M > 6$)	% of total floods
Avalanche	3	11	33	14.02157794	0	0.005970149
Rain	3657	41,637	190,426	17830.89731	6539.589962	35.48437276
Snow	134	2404	851	776.500426	416.4602809	1.54527448
Storm	83	981	6320	473.2605046	229.0867418	0.941811949
Dam breaking	54	568	3600	163.5712257	44.54054417	0.325514877
Typhoon	5	38	1486	28.63278646	12.34100746	0.05698067

In general, every city and each street or home in a city uses a different amount of water. Water consumption directly depends on the number of people in a city. Similarly, some cities provide fewer services, such as industries, hospitals, universities, and schools; therefore, the population of these cities is fewer compared to other cities. The statistics in Fig. 12 for Surrey help us to design a water usage system for the houses in a city. Similarly, fresh water consumption can be maintained; for example, if a house needs more fresh water and another needs less, a balance can be created amongst them. Moreover, the authorities can better control water resources depending on the reservoirs in a city. For instance, if there are more reservoirs, the required amount of water can be stored by finding the overall water consumption parameters in a smart city. Likewise, if there is a scarcity of reservoirs, water needs can be predicted beforehand, and consumption can be planned accordingly.

We also noticed that the water usage in some areas, such as cities and industrial zones, is higher than the water usage in the residential areas. In Fig. 13, we show the average water consumption in different areas. For instance, in areas 101B and 102B, the average water consumption is very low. In areas 102 and 103A, the average water consumption is very high. This helps us to design a system by increasing or decreasing the flow and level of water in different areas. Similarly, an efficient drainage system can be designed while keeping the above statistics in mind. Thus, we can draw a conclusion regarding water consumption in a particular city by planning a billing system of water usage. To test the authenticity of the statistics regarding water usage and make predictions for future needs, we use the skewness measure. The univariate usage of water consumption is W1, W2, W3,… WN using the following skewness formula.

$$\mathrm{Sk} = \sum_{i=1}^{N} \frac{1(Wi - W)3/N}{s^3} \tag{1}$$

where W^{-} is the mean, s is the standard deviation, and N is the number of data points. While computing the skewness, s is computed with N, rather than $N-1$.

We observed that in total 61,263 houses, the average consumption of the house is 57877.937. However, 50% of the citizens consume less than 58,186 cubic meters of water (determined using a median analysis), and 25% of citizens use less than 26,893 cubic meters of water and 75% use less than 81,983 cubic meters of water. The data are positively skewed, which means that more than 50% use more than the average consumption of water. Using this analysis of water consumption, authorities can manage the billing system by choosing a limit for fewer fixed bill payments and by charging extra to those who consume more water than others.

Through this type of water management of water, we can similarly manage energy, such as electricity and gas.

4.2.4. Flood data analysis

The flood due to rainwater normally happens more often and more intensively compared to other types of floods, such as floods due to snow melting, storms, etc. In Table 2, we examine different types of floods, with the results that rainwater produces higher chances of floods, followed by snow. M represents the magnitude of the flood, which is calculated as the log (duration × severity × area affected). For example, if M is greater than 4, it means the flood is of a higher intensity. Approximately 50,250 floods at a higher intensity have occurred at various locations in the world. Similarly, if the value of M is greater than 6, the

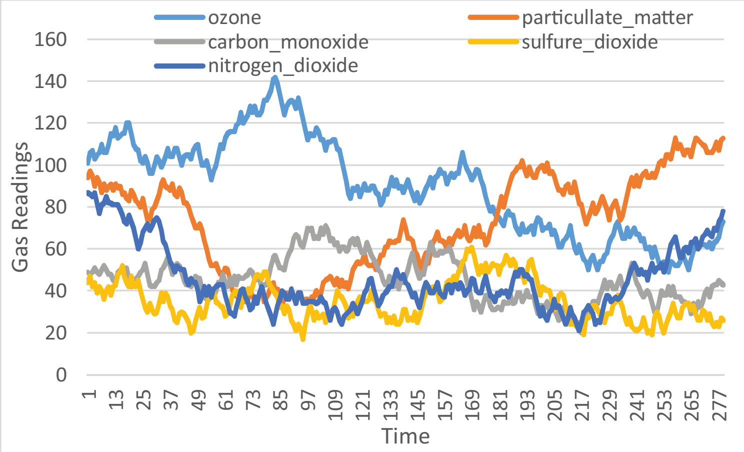

Fig. 14. Pollution levels at different times of day

intensity of the flood is dangerous. A total of 13,751 floods have been recorded at this intensity. The flood ratio in the case of both of these magnitudes is greater in the event of rain. We can see that 35% of the floods have occurred due to rain, followed by snow at 1.5%. Thus, we can predict predefined thresholds of rain for the smart city. For instance, if rain in an area crosses a predefined threshold, then a warning signal or alert can be broadcasted to the public. Society can be made safer by installing high diameter drainage pipes in an area where rain levels are high. Moreover, rain measurements also used to manage the water reservoirs in a smart city. Similarly, melting snow is also a cause of floods but not as often. Sensors can be placed at stations on hills to predict floods due to melting snow.

4.2.5. Environmental data analysis for pollution

Transportation is the main daily activity for Europeans. Each citizen travels at least one hour per day [26]. Therefore, many transportation means, such as buses, trains, and cars, exist in cities. These forms of transportation produce emissions of 12% Co_2 [27]. Moreover, the road population is more than twice as deadly as traffic accidents [28]; additionally, pollution from cars also damages the health of youths and increases the risk of earlier deaths [29], showing how much awareness and safety regarding pollution are important. The most important gases that affect human health are ozone (O_3), carbon monoxide, sulfur dioxide (SO_2), nitrogen oxide, and particulate matter. The existence of these gases in the environment can be analyzed to deliver their current intensity so that more people protect themselves.

Ozone (O_3) is formed with three oxygen molecules. It is too dangerous to contact the living tissues of humans, as it can have the effect of a sunburn on the lungs; produce a cough, irritated throat or uncomfortable feeling in the chest; worsen asthma, emphysema and bronchitis; and may reduce the body's ability to fight infections in the respiratory system. It is formed by the reaction of volatile organic compounds (VOC), nitrogen oxide (NO), and nitrogen dioxide (NO_2). Therefore, nitrogen dioxide is also dangerous. A higher number of VOCs and NO_2 create more ozone; thus, sunny weather, less wind, and high levels of traffic can cause an increase in ozone. The adverse respiratory effects of sulfur dioxide (SO_2) include bronchoconstriction and increased asthma symptoms. Particulate matter is a complex fusion of extremely small particles and liquid droplets. These particles can be formed by acids (such as nitrates and sulfates), organic chemicals, metals, and soil or dust particles. They are so small that they can be absorbed deep into the lungs and cause serious health problems.

For the purposes of analysis and to keep the gas values within a limit, the calculations of gas values are slightly modified [23–25]. However, this does not affect the analysis or the reality and effect of the gases. The values of carbon monoxide, nitrogen dioxide, sulfur dioxide, particulate matter, and ozone index level gas values are calculated as follows:

- Initially assigned a value between 25 and 100. Every 5 min, the values is updated as follows:
- If the value was below 20 before, it would now be the last value + random integer between 1 and 10.
- If the value was higher than 210, it would now be the last value − random integer between 1 and 10.
- Otherwise, the value is the last value + a random integer between −5 and 5.

These gases are dangerous when their values are greater, such as shown in Fig. 14, where the pollution data of Aarhus is depicted. The maximum values for all gases are shown as ozone value at times 70–90, particulate matter values at time 185–215 and more than 245, nitrogen dioxide at the start and end of the time interval, and carbon monoxide at 90–115, are all dangerous for health. Therefore, children should not be allowed to spend time outdoors during these times. Moreover, adults should not exercise outdoors at these times, as people engaged in physical activity breathe faster and more deeply, causing a flow of ozone into the lungs. People with the respiratory diseases should also be careful when ozone values are

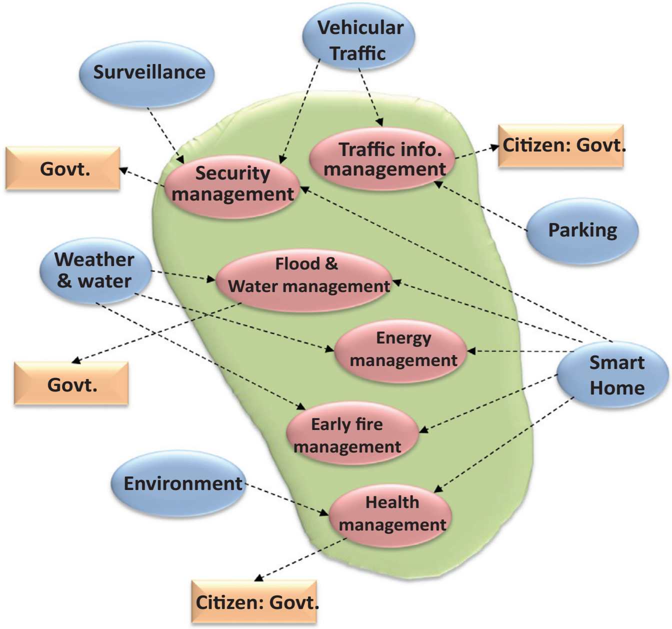

Fig. 15. Smart city system implementation scenario.

higher, as ozone can further damage the lungs of people with respiratory diseases.

For a daily pollution analysis, we advised people about the intensity of the pollution and suggested they not go outside and did not allow children, the elderly, or people with respiratory diseases to go outside when the intensity of any of the gases was higher. Authorities can also take action and alert the public when the air pollution exceeds a certain limit. The government can also use these data for urban planning by analyzing the history and changes in pollution during different seasons and months. A yearly analysis can be used to plan for traffic, cities and industrial buildings and can be used to shift industries to places outside the city or to build new industries further from the cities when gas levels begin to increase.

5. System implementation

Based on the datasets collected, the analyses made, and the proposed system architecture, the system was developed using a Hadoop single node at Ubuntu 14.04 LTS with 3.2 GHz × 4 processors and 4GB memory. The PCaP format traffic was processed by Hadoop-pcap-lib and Hadoop-pcap-scr-de libraries. These traffic data were then converted into sequence files to be able to process them using Hadoop. The system was implemented by two major modules, i.e., the smart city and urban planning. These two modules have other sub-modules for various functionalities.

5.1. Smart city implementation

The input source remains the same as described previously, as shown in Fig. 15, with circles outside the boundary of the system, i.e., the smart home, parking, etc. Each facility of the smart city is implemented as a separated class or sub-module that takes data from various sources. Traffic information measurements take data from vehicular traffic and parking. The security management module takes data from surveillance, smart homes, and vehicular traffic in the case that the government needs to monitor stolen vehicles. The flood and water management modules take water usage data from smart home data for rain and ice storms also predict floods in real time. Similarly, energy consumption management takes electricity and gas data from smart homes as well as dam and water-related data. This module manages and saves extra energy that is not being used by homes. It also distributes the energy to various areas according to the need. Similarly, the early fire management program performs fire detection. Finally, the health management module makes decision regarding pollution data. Citizens have limited access to the results of these modules, but the government has full access to them. The complete flow of data, modules, and actors is shown in Fig. 15.

5.2. Urban planning system implementation

Urban planning system implementation is conducted at three levels, i.e., the physical, intermediate and upper level, as shown in Fig. 16. The physical level is called the storage level and is based on the Hadoop HTFS system. All historical data are stored at the physical level. Each dataset is given a number in the figure, such as vehicular data at number 1, energy data at number 2, and so on. The intermediate level is the second level, which is also called the processing level. All processing is performed at this level using the data stored on the physical level. At this level, statistical calculation, computation, graph analysis, and other computations are performed. The third level is

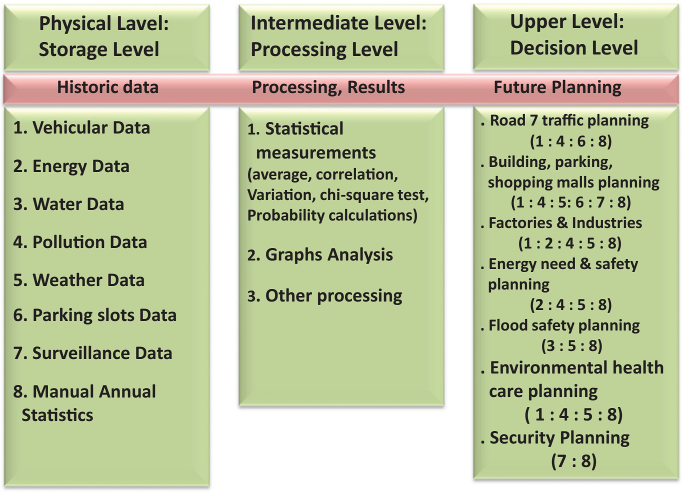

Fig. 16. Urban planning system implementation scenario.

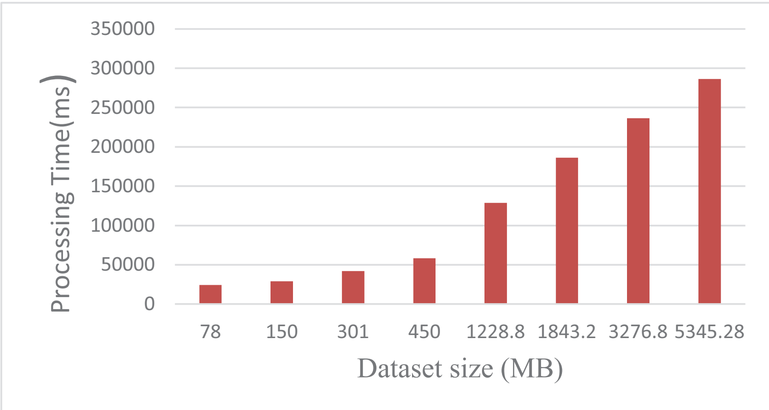

Fig. 17. Processing time of various size vehicular datasets.

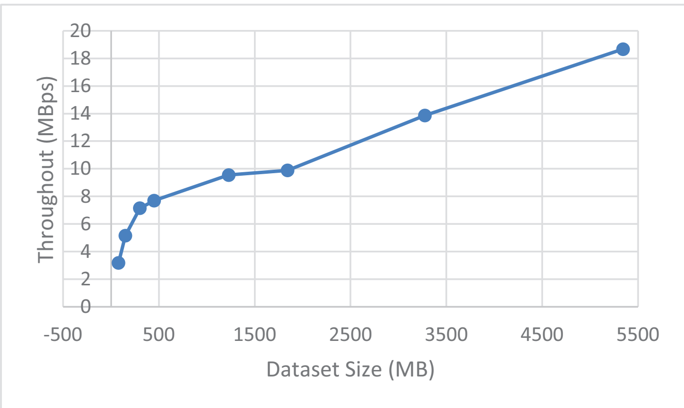

Fig. 18. Throughput of datasets depending on the size of data.

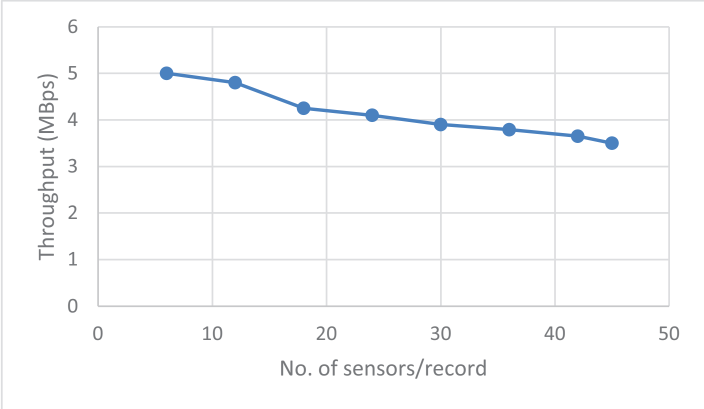

Fig. 19. Throughput of the system by increasing the number of sensors per record for 1GB of data.

the upper level, which is also called the decision level. Decisions regarding urban planning are made at this level. The decision level has various modules for each type of planning, for example, road planning and building planning. The number written under the planning module is the number of datasets from which the module takes data for input.

6. System evaluation

The proposed algorithm is implemented using the Hadoop single node setup on UBUNTU 14.04 LTS coreTMi5 machine with a 3.2 GHz processor and 4GB memory. For real-time traffic, we generated Pcap packets using Wireshark libraries and retransmitted them using other systems to develop our system. Hadoop-pcap-lib, Hadoop-pcap-serde, and Hadoop Pcap Input libraries were used for network packet processing and generating Hadoop Readable (sequence files) at collection and aggregation units such that Spark can process the data. MapReduce programming was used to perform offline analysis for urban planning. The datasets mentioned in section IV were used to perform an efficiency evaluation of the system.

Because the system is based on Big Data analytics, it was evaluated with respect to efficiency and response time. System performance was measured for various sized datasets considering the processing time (in milliseconds) and throughput (Mbps). The processing time results are shown in Fig. 17, and the throughput analysis results are shown in Fig. 18. It is obvious from the graph that when data size is increased, the processing time also proportionally increased, as both data size and processing time are directly proportional to each other. However, we can examine the processing at higher (larger) dataset, i.e., 5345MB; the processing time for this dataset is 300,000, which is far better than other systems. Moreover, when we analyzed the throughput corresponding to the data size, we identified that the throughput was also directly proportional to data size because of the parallel processing nature of the Hadoop system. This is the major achievement of the system, as when there is an increase in data size, the throughput is also increased.

We also tested the performance of the system by increasing the number of sensors for a single record. We kept the data size constant, i.e., 2GB, and raised the number of sensors per record; we thus learned that with an increase in the number of sensors, the throughput was decreased because when we increased the sensors, it took significant time for classification filtrations and processing, as there were many comparisons due to the large number of sensors in a single record. The throughput of the system with respect to the number of sensors is shown in Fig. 19.

7. Conclusions

Smart cities and urban planning can have a major impact on national development. These efforts can increase the decision-making power of society by allowing them to make intelligent and effective decisions at appropriate times. In this paper, we propose a system for smart cities and urban planning by using an IoT-generated Big Data analysis. The proposed architecture consists of four tiers that have functionalities including collection, aggregation, communication, processing, and interpretation. The complete system is developed using Hadoop technologies with Spark to achieve real-time processing. The simple IoT-based smart city datasets, such as vehicular networks, smart parking, smart home, weather, pollution, surveillance, and so on are analyzed for developing a smart city as well as for urban planning decisions. The proposed system not only benefits citizens but also authorities while providing them with the facilities to make intelligent and quick decisions. The system is finally tested based on efficiency performance considering processing time and throughput. The system gives efficient results on even larger data sets. The system throughput is also increased with an increase in data size.

Acknowledgment

This study was supported by the Brain Korea 21 Plus project (SW Human Resource Development Program for Supporting Smart Life) funded by Ministry of Education, The School of Computer Science and Engineering, Kyungpook National University, Korea (21A20131600005).

This work was supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIP). [No. 10041145, Self-Organized Software platform (SoSp) for Welfare Devices]. This research was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2013R1A1A2061978).

References

[1] CISCO, The Internet of Things, Infographic, available online at http://blogs.cisco.com/news/the-internet-of-things-infographic, 2015.
[2] A. Ahmad, A. Paul, M. Mazhar Rathore, H. Chang, Smart cyber society: integration of capillary devices with high usability based on cyber–physical system, Futur. Gener. Comput. Syst. 56 (2016) 493–503.
[3] D. Zeng, S. Guo, Z. Cheng, The web of things: a survey, J. Commun. 6 (6) (2011) 424–438.
[4] L. Srivastava, Japan's ubiquitous mobile information society, Info 6 (4) (2004) 234–251.
[5] S. Giroux, H. Pigot, From smart homes to smart care: ICOST 2005, in: Proceedings of the 3rd International Conference on Smart Homes and Health Telematics, vol. 15, IOS Press, 2005.
[6] S.S. Han, Global city making in Singapore: a real estate perspective, Prog. Plan. 64 (2) (2005) 69–175.
[7] M. O'droma, I. Ganchev, The creation of a ubiquitous consumer wireless world through strategic ITU-T standardization, IEEE Commun. Mag. 48 (10) (2010) 158–165.
[8] F. Xia, L.T. Yang, L. Wang, A. Vinel, Internet of things, Int. J. Commun. Syst. 25 (9) (2012) 1101.
[9] D. Zhang, J. Zhou, M. Guo, J. Cao, T. Li, TASA: tag-free activity sensing using RFID tag arrays, IEEE Trans. Parallel Distrib. Syst. 22 (4) (2011) 558–570.
[10] D. Zhang, M. Guo, J. Zhou, D. Kang, J. Cao, Context reasoning using extended evidence theory in pervasive computing environments, Futur. Gener. Comp. Syst. 26 (2) (2010) 207–216.
[11] D. Zhang, D. Zhang, H. Xiong, C.-H. Hsu, A.V. Vasilakos, BASA: building mobile Ad-Hoc social networks on top of android, IEEE Netw. 28 (1) (2014) 4–9.
[12] Sudhir Dixit, Ramjee Prasad (Eds.), Technologies for Home Networking, John Wiley & Sons, 2007.
[13] J. Jin, J. Gubbi, S. Marusic, M. Palaniswami, An information framework for creating a smart city through Internet of things, IEEE Internet Things J. 1 (2) (2014) 112–121.
[14] A. Ahmad, A. Paul, M. Mazhar Rathore, An efficient divide-and-conquer approach for big data analytics in machine-to-machine communication, Neurocomputing 174 (2016) 439–453.
[15] M.M.U. Rathore, A. Paul, A. Ahmad, B. Chen, B. Huang, W. Ji, Real-time big data analytical architecture for remote sensing application, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (10) (2015) 4610–4621.
[16] D. Kyriazis, T. Varvarigou, A. Rossi, D. White, J. Cooper, Sustainable smart city IoT applications: heat and electricity management & Eco-conscious cruise control for public transportation, in: Proceedings of the 2013 IEEE 14th International Symposium and Workshops on World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2013, pp. 1–5.
[17] G.R. Brakenridge, Global Active Archive of Large Flood Events, Dartmouth Flood Observatory, University of Colorado, 2015 http://floodobservatory.colorado.edu/Archives/index.html (accessed 15.06.30).
[18] http://data.surrey.ca/dataset/water-meters, (accessed 15.06.30).
[19] H.M. Gramaglia, O. Trullols-Cruces, D. Naboulsi, M. Fiore, M. Calderon, Vehicular networks on two Madrid, in: Proceedings of IEEE Conference on Sensor and Ad Hoc Communications and Networks (SECON), Singapore, 2014.
[20] S. Uppoor, M. Fiore, Large-scale urban vehicular mobility for networking research, in: Proceedings of Vehicular Networking Conference (VNC), IEEE, 2011, pp. 62–69.
[21] D. Naboulsi, M. Fiore, On the instantaneous topology of a large-scale urban vehicular network: the cologne case. in: Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 2013), Bangalore, India, 2013.
[22] S. Uppoor, O. Trullols-Cruces, M. Fiore, J.M. Barcelo-Ordinas, Generation and analysis of a large-scale urban vehicular mobility dataset, IEEE Trans. Mob. Comput. 13 (5) (2014) 1061–1075.
[23] T. Bischof, A. Karapantelakis, C.-S. Nechifor, A. Sheth, A. Mileo, P. Barnaghi, Semantic modeling of smart city data, in: Proceedings of Position Paper in W3C Workshop on the Web of Things: Enablers and Services for an Open Web of Devices, Berlin, Germany, 2014.
[24] R. Tönjes, P. Barnaghi, M. Ali, A. Mileo, M. Hauswirth, F. Ganz, S. Ganea, B. Kjærgaard, D. Kuemper, S. Nechifor, D. Puiu, A. Sheth, V. Tsiatsis, L. Vestergaard, Real time iot stream processing and large-scale data analytics for smart city applications, in: Proceedings of European Conference on Networks and Communications, Poster Session, 2014.
[25] S. Kolozali, M. Bermudez-Edo, D. Puschmann, F. Ganz, P. Barnaghi, A knowledge-based approach for real-time IoT data stream annotation and processing, in: Proceedings of the 2014 IEEE International Conference on Internet of Things (iThings 2014), Taipei, Taiwan, 2014.
[26] Eurostat, "Passenger mobility in Europe." European Commission, 2007
[27] Eurostat, "Energy, transport and environment indicators", European Commission, 2011
[28] S. Yim, S S. Barrett, Public Health Impacts of Combustion Emissions in the United Kingdom, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, United States, 2012.
[29] US Environmental Protection Agency (EPA), Car Pollution Effects, 2012.

Muhammad Mazhar Ullah Rathore received the Master's degree in computer and communication security from the National University of Sciences and Technology, Islamabad, Pakistan, in 2012, and is currently pursuing the Ph.D. degree atKyungpook National University, Daegu, Korea. His research interests include IoT, Big Data analytics, network traffic analysis and monitoring, intrusion detection, and computer and network security.

Awais Ahmad received the B.S. degree (CS) from the University of Peshawar, Peshawar, Pakistan, and the M.S. degree (telecommunication and networking) from Bahria University, Islamabad, Pakistan, in 2008 and 2010, respectively. Currently, he is pursuing the Ph.D. degree at Kyungpook National University, Daegu, Korea. During his research work, he worked on energy efficient congestion control schemes in Mobile Wireless Sensor Networks (WSN). There he got research experience on Big Data analytics, machine-to- machine communication, and wireless sensor network. Mr. Ahmad was the recipient of three prestigious awards: (1) Research Award from President of Bahria University Islamabad, Pakistan in 2011, (2) best Paper Nomination Award in WCECS 2011 at UCLA, USA, and (3) best Paper Award in 1st Symposium on CS&E, Moju Resort, Korea, in 2013.

Anand Paul received the Ph.D. degree in electrical engineering from the National Cheng Kung University, Tainan, Taiwan, in 2010. He is currently working as an Associate Professor with the School of Computer Science and Engineering, Kyungpook National University, Daegu, Korea. He is a delegate representing Korea for M2M focus group and for MPEG. His research interests include algorithm and architecture reconfigurable embedded computing. Prof. Paul has Guest Edited various international journals and he is also part of Editorial Team for Journal of Platform Technology and Cyber Physical Systems. He serves as a Reviewer for various IEEE/IET journals. He is the track Chair for smart human computer interaction in ACMSAC 2015, 2014. He was the recipient of the Outstanding International Student Scholarship Award in 20042010, the Best Paper Award in National Computer Symposium, Taipei, Taiwan, in 2009, and UWSS 2015, in Beijing , China. He is also IEEE Senior Member

Dr. Seungmin Rho, Ph.D. is a faculty of Department of Media Software at Sungkyul University in Korea. In 2012, he was an assistant professor at Division of Information and Communication in Baekseok University. In 2009-2011, he had been working as a Research Professor at School of Electrical Engineering in Korea University. In 2008-2009, he was a Postdoctoral Research Fellow at the Computer Music Lab of the School of Computer Science in Carnegie Mellon University. He gained his B.Science. (2001) in Computer Science from Ajou University, Korea (South), M.Science. (2003) and Ph.D. (2008) in Information and Communication Technology from the Graduate School of Information and Communication at Ajou University. He visited Multimedia Systems and Networking Lab. in Univ. of Texas at Dallas from Dec. 2003 to March 2004. Before he joined the Computer Sciences Department of Ajou University, he spent two years in industry. His current research interests include database, big data analysis, music retrieval, multimedia systems, machine learning, knowledge management as well as computational intelligence. He has published more than 180 papers in refereed journals and conference proceedings in these areas. He has been involved in more than 20 conferences and workshops as various chairs and more than 30 conferences/workshops as a program committee member. He has been appointed as an Editor-in-Chief in Journal of Platform Technology (JPT) since 2013. He has edited a number of international journal special issues as a guest editor, such as Enterprise Information Systems, Multimedia Systems, Information Fusion, ACM Transactions on Embedded Computing, Journal of Real-Time Image Processing, Future Generation Computer Systems, Engineering Applications of Artificial Intelligence, New Review of Hypermedia and Multimedia, Multimedia Tools and Applications, Personal and Ubiquitous Computing, Telecommunication Systems, Ad Hoc & Sensor Wireless Networks and etc. He has received a few awards including Who's Who in America, Who's Who in Science and Engineering, and Who's Who in the World in 2007 and 2008, respectively.