

Designing Smart Control Systems Based on Internet of Things and Big Data Analytics

Murad Khan¹ · Kijun Han² · S. Karthik³

Published online: 22 January 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract The Internet of Things (IoT) lay down a platform for global communication among millions of electronic devices connected to the internet. These devices and electronic appliances included both wireless and wired sensors, home appliances such as Television, refrigerator, etc., radio frequency identifications (RFID), and so. Similarly, heterogeneous networks provide a platform for media independent communications. However, there exist several issues in using heterogeneous technologies for IoT communications. These challenges include the co-existence of wireless technologies such as ZigBee, Bluetooth and WIFI, cross-layer communications, high packet loss due to interferences with electronic devices, etc. Similarly, IoT is a new paradigm for interconnecting electronic devices, thus, it needs major refinement for standardising it for various services. In order to address the aforementioned challenges, we proposed a scheme for enabling a generic IoT framework and platform for various IoT embedded devices. The working of the proposed scheme is twofold, (1) discovering and collecting information from IoT enabled devices using sensors and (2) scheduling the working of these appliances based on the data collected using sensors attached to these devices. Further, the data is transferred to Hadoop ecosystem for processing and analyzing to disseminate the relevant information to the citizen. Moreover, the proposed system is tested in a smart home scenario by installing sensors attached to various home appliances. The energy consumption and packet loss

✉ Kijun Han
kjhan@knu.ac.kr

Murad Khan
murad.csit@suit.edu.pk

S. Karthik
profskarthik@gmail.com

¹ Department of Computer Science and Information Technology, Sarhad University of Science and Information Technology, Peshawar, KPK, Pakistan

² School of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea

³ SNS College of Technology, Coimbatore, India

occur due to available electronic appliances, and heterogeneous devices are computed and analyzed for planning an optimal scheduling scheme and load balancing. Similarly, data from various authentic sources is analyzed using Hadoop ecosystem and disseminate it to the citizen in real-time.

Keywords Wireless sensor networks · Hadoop ecosystem · Internet of Things · Scheduling · Load balancing

1 Introduction

The Internet of Things (IoT) has progressively developed with the passage of connecting heterogeneous technologies. The word IoT was first announced by two students working at MIT in late 2000 while working on Radio Frequency Identifiers (RFID). However, the IoT has been used in various fields of technology and communication for connecting embedded internet devices. Both short and long-range communication technologies is used for data transfer in an IoT environment with sophisticated protocols and algorithms. The data transfer from a sensor attached to an IoT device to an end user is made possible with processing data in real-time. Various flavours of IoT have been developed and designed with the passage of time such as the Internet of Every Things (IoET), Internet of Social Things (IoST), Internet of Vehicle Things (IoVT), fog for everything etc. [1, 2]. These flavours help in identifying the communication protocols, underline technologies, and privacy of data to be used for any specific purpose in an IoT environment. Similarly, the classifications of such huge amount of data over the internet also need sophisticated technologies such as cognitive radio networks, software-defined networking, etc. The most important goal of IoT is to enable human beings to do their work in the shorter amount of time with less possible congestion and overhead in communication networks. The betterment of human beings can be made possible with implementing IoT in a way that can be simple and user-friendly to use. For example, if a user wants to wash his clothes back in a home with the help of IoT which is further connected both the user and washing machine at home. The user can send commands over the internet to control the machine at home. Similarly, other examples of betterment can be seen by controlling the IoT devices such as lights, television, refrigerator, etc. at home. Such controlling of home appliances come under the category of smart homes which is an application of IoT. Other example consists of smart parking, smart health management, smart traffic and vehicular communications, etc. In addition, other new technologies and concepts such as Information-Centric Networking (ICN), Name Data Networks (NDN), and Wireless Sensor Network (WSN) based data dissemination, etc. is progressively developing due to the IoT [3–7]. Further, the literature consists of various examples and innovative technologies which are using IoT in some way.

The progressive research in IoT identifies many challenges and issues in IoT. These challenges and issues include the huge amount of data processing generated through various connected devices in real-time [8, 9]. Similarly, other challenges include designing sophisticated and efficient protocols for data collection and acquisitions, data dissemination and transfer over existing networks, controlling IoT devices with the help of Power Line Communications (PLC), and so on [10–12]. The literature still lacks a complete standard fulfilling and addressing the demands and challenges present in an IoT based infrastructure such smart homes, cities, etc. These challenges can be addressed with designing a layering architecture for IoT communications. Similarly, the data is transferred

with the help of an intermediate layer consisting of various technologies to the data processing layer which is further connected with an application layer. However, it is still a challenging job how these layers can communicate with each other for efficient data transfer. The existing layering architecture have several problems, and issues such as poorly addressing the cross-layer communications, processing the data in offline mode, communication protocols for seamless transfer of connections among mobile IoT devices, and so on [13, 14]. In addition, the short-range communication technologies need much assistance as it is mainly operated on batteries. Thus, a communication overhead or other problem can consume much battery during peak hour time which can make it imperative to improve energy efficiency and increase the lifetime of an IoT device with limited battery capacity. Further, the IoT environment needs separate addressing of IoT devices which can help in efficient data transfer in a local IoT environment. For example, a sensor device can be named such as an IoT-enabled device to make it separate from existing WSN based sensors. One of the reasons for separate naming schemes can help in designing standard communication protocols for IoT [15]. However, it is impossible until a standard body such as IEEE, ACM, Microsoft, etc. define such standard names [16].

The aforementioned challenges are addressed by various researchers from various fields of information technology and engineering. One of the steps was taken by the European Communication Council (ECC) in 2007 to present a standard for IoT technology. However, it is still going through refinement for final use. Similarly, other market firms such as Qualcomm, ACM, etc. come up with several ways and research ideas to use IoT in various fields such as smart homes, cities, and so on. According to the US National Intelligence Council (NIC) by 2025, various things such as food items, furniture, electronic machines and so on will be connected with IoT. The literature consists of various schemes for efficient data gathering and collection from various IoT devices. These schemes include the collecting data using clustering approach, service platform approach, etc. [17–19]. The said protocols are able to place sensors at various locations such as different roadsides to count the number of cars on the roads, smart homes, smart health clinics, etc. and transfer it to citizens for a better selection of less congested roads in real-time. The real-time processing of such huge amount of data is carried out using Hadoop ecosystem alongside SPARK and GRAPHX. However, it is still a challenging job to disseminate the correct and accurate information among the citizens in real-time. Therefore, researchers develop new technologies such as ICN and NDN for the said purpose. However, it is not yet fully operated for disseminating data in heterogeneous networks [20]. A home automation system has been developed for efficient collection of data from various home devices and sent it to a management station to process it for optimising energy consumption of home appliances in real-time [21]. The proposed scheme presented a case of study of a single user home system followed by a multi-user smart home system. A scheduling scheme is designed based on game theory and dynamic programming to address the consumption of electricity and calculate the electricity pricing with various home appliances turned-on at different interval of time. A similar scheme is presented to the optimised electricity consumption of light sources available in a smart home by integrating the sunlight with the light sources [22]. The proposed scheme optimises the energy consumption by dividing the entire room into different zones and then tune the light sources according to the sunlight. Similarly, various schemes are presented for specific operations in a smart home, smart city, smart community, etc. However, they do not address the challenges present in an IoT environment in a generic way.

In order to address the problems and challenges present in IoT environment in a generic way, we come up with scheme presented a model of installing sensors in an IoT environment to disseminating data with citizens in real-time. The proposed scheme broadly

targets the urban planning with the help of sensors networks, ICN, NDN, and big data analytics. The proposed scheme is further supported by a case study scenario of a smart home and a smart city data analytics using Hadoop alongside SPARK and GRAPHX. The data of various authentic sources such as parking, environmental, etc. are tested in real-time, and the results are disseminated to citizens for better use and selection of services. Further, the system is consisted of a sensor deployment phase to efficiently and accurately deploy sensors at various locations in a smart city to collect data from various IoT-enabled devices. The data is then passed to Hadoop ecosystem with the help of underlying technologies. The processed data is transferred to an application layer to distribute it among required and concerned departments. The results of various smart homes and other services and departments such as smart parking management, environmental department, etc. of a smart city are tested using proposed scheme. The results show that the proposed scheme is able to process the data in real-time and the citizens can select the required service with less amount of time.

The rest of the paper is further divided into following sections. Section 2 explains the related work and back ground studies of the recent work published in the said domain. The proposed scheme is explained in Sect. 3 following by results and discussion in Sect. 4. Finally, the paper is concluded in Sect. 5.

2 Related Work

IoT (Internet of Thing) system has many vital factors; low power layer communication is one of them. Hence, IEEE802.15.4 display a complete detail of physical layer for IoT system, which facilitates short-term communication services. Thus, IoT prime function is to direct communication in low coverage area between two objects. Network interface card (NIC) accomplish multiple tasks. Conversion of digital into an analog signal or electromagnetic is one of the key task assigned to NIC. Modulation techniques are used for conversion digital bytes into analog streams. Analog signals, being very low power and cannot be sent directly to the antenna by NIC. Firstly, these signals are converted to high power in the air for efficient transformation. The power amplifier receives signals from NIC and sends it to the antenna. While converting and sending from the system, the signal agonised to noise and other challenges like riding and volatizing. Now, the low noise amplifier is used to make robust signals which can be managed by the demodulator. Further, weak signal range from-80- to 90-dBm is received by radio interface. Highest energy consumption elements are noise amplifier, demodulator and modulator. Therefore, IoT infrastructure should design in such a smart way that it will perform radio operation in lowest possible energy. To overcome power consumption challenges, IoT Enabled Device IoTED needs a mature scheduling mechanism which auto turn on/off the radio.

To keep the above mechanism of energy consumption in mind, the power if IoTED is mostly affected by the scheduling module. In the last couple of year, different scheduling mechanisms have been proposed and target the round trip of the IoTED. Lower than 1% round trip is claimed in these schemes as shown in [23], totally based on priority queue model. To handle IoT scheduling among the IoT clients and IoT devices, the concept of agent is made known to IoT setup. As it's hard to figure out the cyclic duration of IoT client and concern device within conventional network. By using M/M/1 queue method, traffic among participant devices will flow without any noise. Traffic flow categorisation into a diverse module (k) is carried out on behalf of messages. The author focusses on

request time and gives a detail description of how the request time can be processed and manage using the traffic flow methodologies. As random device request may lead to queue congestion in IoT structure and down grade performance. In [24], the author gives the same idea built on round-trip radio interface of a device. An eminent traffic responsive algorithm is used in the proposed scheme to minimize the response time up to certain limit. Moreover, time slots allocation and channel offsets are done through a network topology. Finally, the node was made responsible for scheduling with the help of a data flow paradigm. Furthermore, the projected scheme achieved the milestone, minimise delay and maximise throughput using a smaller number of channels.

Power saving is one of the key parameters which can be achieved by scheduling IoT devices. There should be a global standard for IoT hardware vendors to manufacture lightweight, energy aware and small devices in terms of dimensions comparatively other conventional devices. By doing this, a variety of devices are designed using constrain energy and dimensions of IoT devices which deliver refined IoT devices for future generation aspects. For IoT devices, Different statements are documented in IEEE802.15.4, in which the most common state using by IoTED are Idle, receiving, transmitting, sleep, awake, etc. Many of the research studies focus on these states of IEEE802.15.4 for better management of energy consumption and saving. In the case of listening state, the node starts transmission if and only if the channel is free and the availability of channel is carried out by the transceiver, as transceiver continuously listens to the available channels. In the case of the idle state, even though the transceiver always in the awake state but the interface is not able to process incoming data and request from the channel. Finally, in the active state, both transmission and receiving take place. It is concluded that round-trip time can be minimized by keeping the radio interface in an idle state. Nowadays, researchers are working on duty cycle to overcome energy and enhance the operational activities of an IoT device. In order to explain the working of a duty cycle, let take an example of reducing the duty cycle of a node contain battery power of 3000 mAh and let assume that an LTC5800 radio consumes 10 mA electric current to receive the data. If we keep the radio interface on permanently, then its duty cycle will be 100% and the consumed battery will be $3000/10 = 300$ h which means 12.5 days. Similarly, by keeping the round-trip 1% the battery will work up to 1217 days. The IEEE802.15.4 power consumption constraints have been intimidated in the upgraded model, i.e. IEEE802.15.4e where it defines the MAC layer and introduces a direct protocol communication with the interface. In addition, it relays on MAC header and defines well the communication methodology among different nodes. The communicating strategy of IEEE802.15.4e can be simulated by Bluetooth v3.0 where nodes are obliging to communicate with the master one and Bluetooth v4.0 where nodes are communicating with each other. Moreover, considering wireless networking methodology, the dedicated medium is never guaranteed to be stable and free of any interferences because the signals are mostly exposed to any external noise, jitter, and acknowledgement delays. Hence, single-channel operations will most likely results in performance decreasing over networks. The term synchronisation is one of the necessary key perimeters to keep establishing the communication between devices. The synchronisation process is performed in two different methods called Frame and Acknowledgment. The acknowledgement based synchronisation is referring to the process where receiver calculates time difference of sending and receiving frames. While in the case of frame synchronisation process, the receiver calculates the time difference between arrival and actual arrival time. As the Model IEEE802.15.4e is unlike of Model IEEE802.15.4 in MAC protocol, hence, the hardware change is not needed at all. Model IEEE802.15.4e attain a tiny duty cycle which results in a significant decrease in node energy consumption because this model uses the time based synchronised channel hopping

technique named as TSCH. Furthermore, the prior MAC protocol version was adopted as time synchronised mesh protocol (TSMP) since 2006. TSMP is adopted for Wireless HART [25] because of its substantial success in various products. The utmost aim of IEEE802.15.4e model is to effectively well-organise the nodes' energy by synchronising the nodes along with TSCH for channel hopping to make it reliable. Figure 1 shows TSCH graph where two nodes are perfectly synchronized with each other by listening to the advertisement packets [26]. As the synchronisation process is ended, each node is labeled with time slots and channel offsets. The same nodes information (slots and channel slots) are used to communicate with neighbours.

A significant feature of MAC layer in IEEE802.15.4e is to intensify the scheduling methods. The scheduling needs a great care of attention to be shaped because if a node is intended to communicate with another node, the receiving one should be on the listening node. There are two different types of scheduling under IEEE802.15.4e, i.e. Centralized and distributed. The centralised scheduling approach is specifying a manager node, which is responsible for maintaining and controlling the distribution of schedules. In addition, the available nodes in the network assist the Manager Node with the newbie, and the one cannot listen to the manager. The distributed approach refers to link establishment among nodes to freely schedule with other nodes. Similarly, interested readers are referred to [27] for scheduling based on the link capacity. The authors also referred to the Aloha-based scheduling model that devote the nodes with the channel for broadcasting advertisement. Moreover, the reservation-based scheduling is dedicatedly used to hold timeslots for target advertisement. Considering IoT area, a number of scheduling algorithms have appointed based on QoS-aware scheduling [28]. While sadly saying, these algorithms have focused on generic network environment and elsewhere. For instance, in [28], the author alters the existing layering structure with QoS-aware layering structure. Likewise, the authors introduced a scheduling algorithm in [23] while the mentioned proposed scheme is notably upgraded. For a generic scheduling mechanism, one should stay with above-mentioned constraints in mind. Otherwise, this algorithm can be limited to some scenarios. Following are the other factors that concern the power consumption of IEEE802.15.4e.

3 Proposed IoT Architecture

The proposed scheme is in two folds. In the initial phase identifies the appliances that are always or frequently in the state "on". Upon identifying those appliances, the second phase proposes a switching mechanism to minimise the energy consumption.

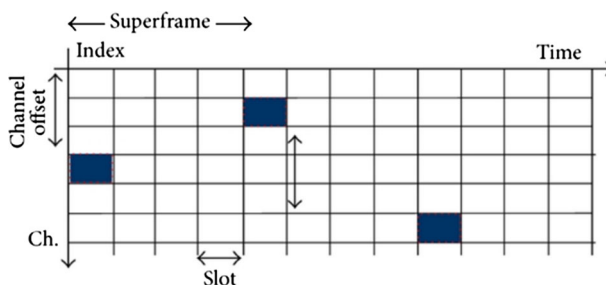


Fig. 1 TSCH: time synchronisation channel hopping

3.1 Overview of the Proposed System

The IoT notion enables communication among multi-vendor, multi-purpose devices that are deployed in IoT environment via heterogeneous communication technologies. Optimal communication capabilities are achievable by considering various aspects and arranging them accordingly. Sensor technology plays a key role in IoT. Hence, extra cautious mechanisms should be followed to facilitate energy-aware communication. Sensors consume considerable energy to be in on state. Thus, we propose a scheduling mechanism to manage the energy consumption of sensors. The sensors are turned to sleep mode by the scheduling mechanism, upon noticing no activity from a sensor or a group of sensors. In the deployment of sensors, a set of sensors are attached to a coordinator (management station). The coordinator processes data sent from the sensors attached to it. In addition to data processing, coordinators communicate with control server and send commands to sensors. The control server hosts a web service that performs automatic actions and records all the actions performed on each sensor.

3.2 Appliances Discovery

In this phase, the scheme identifies all appliances that are switched on during user absence. For example, assume a scenario where a resident is at home and watching television, but the computer is in on-state at the same time. Another common example could be a house with children where they turn on various electronic devices, i.e. television, light bulbs, etc. without actually using them. Henceforth, we categorised the appliances in an IoT environment on the basis of user dependability. Table 1 describes the appliances categorisation.

Afore stated appliances are commonly found in most of the IoT environments. Nevertheless, devices are not limited to aforementioned ones. Therefore, we built a generic architecture that caters heterogeneous devices.

3.3 Sensor Configuration and Deployment

The sensor deployment mainly depends on the IoT environment. However, the distribution in an IoT environment is carried out using the Poisson distribution. The proposed architecture occupied two types of sensors i.e. Farthest sensor (F_R) and Relay sensor (R_S). Each sensor manages communication between user and communication between a Management Station (MS), the MS has more battery power than a typical sensor. It is further connected with the web service via cellular network or Wi-Fi. Since F_R are not directly communicating with MS, R_S nodes act as mediators between F_R and MS. Contrasting with F_R , R_S

Table 1 Appliances categorization

User availability	Usage time	Appliances
High	User dependent	Television, light bulb, etc.
Medium	Partially user dependent	Toilet, scanner, etc.
Low	Less user dependent	Washing machine, rice cooker, etc.
Needless	User depends less	Street lights, microwave Owen, etc.

have both sensing and processing capability. The key responsibility of these sensor nodes is to determine whether the devices are in on-state when the users are not available. Information collected at F_S travels through $k - 1$ hops. Since we need to allow multiple hops, we employed multi-hop technology between F_S and corresponding MS. R_S act as the intermediate node between MS and F_R , therefore, R_S nodes consume more energy compared to F_S nodes. Figure 2 clearly illustrates the communication pathway of MS, F_R , and R_S .

3.4 Event Management and Scheduling

As previously stated, sensors consume considerable energy to operate in on state. For example, 37 mA energy is consumed by CC2530 sensor for both data receiving and data transfer operations. Hence, we introduced an Appliance Sleep-Scheduling (ASS) mechanism that can significantly improve the battery life of sensors. We categorized the sensors into groups and attached every group with an MS, in order to manage ASS mechanism efficiently. The MS offers two types of services i.e. send information collected from sensors to the web service and also sends information to the user. A form of message event M_{EU} is occupied to transfer information to the user. The M_{EU} contains the identification number of the device that is in on-state (S_{ID}) and turned-on time for that particular device (T_{ON}). Owing to essential information provided by M_{EU} , user can decide on the succeeding action that can be either turning-off the device or let the device to be in on-state ($R_{ON/OFF}$).

On the other hand, if the user intends to change sensor state into sleep mode, R_{SLEEP} response message is forwarded to the MS. Whenever the user is unable to receive the M_{EU} or the user is unable to respond, we introduced an automatic response A_R generator with the web service. Once the MS receives response message from the user, MS sends U_{STOP} message to the web service, which restricts A_R response to the sensor. However, if the web service does not receive U_{STOP} , it initiates A_R upon the expiration of elapsed time E_T . Subsequent to the expiration of E_T , A_R changes the corresponding sensor into sleep-state. A confirmation message (Ack) is sent to the MS, Web and user for at the end of performing

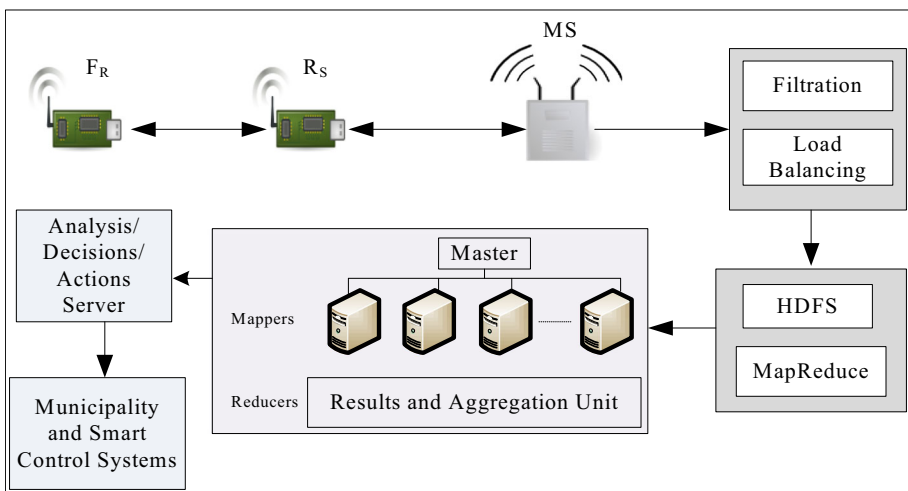


Fig. 2 Communication of F_R and R_S with MS

the necessary action. Figure 3 illustrates the entire sleep-scheduling mechanism using a flow diagram between sensors and a user.

3.5 Information Gathering and Processing

The required data is collected by the sensors and forwarded towards the MS. Each sensor collects two types of data related to a device i.e. (1) S_{ID} of the device and (2) turned-on time of the device. A sensor stores these information for T_S time. If the device is turned-off during T_S time, stored information is deleted from the sensor to release memory for the next event. If the device remains in on-state after T_S time, sensor generates M_{EU} message and transfers it to the corresponding MS. The sensor waits for W_T time to get a response from user or MS. Upon receiving a response, the sensor deletes stored information and performs necessary action accordingly i.e. switch to sleep state or remain in on- state. A sensor consumes its energy on three main tasks namely; (1) Information sensing, (2) Transferring information to user and MS, and (3) receive commands from user and MS.

The web service collects data stored in sensors on an hourly basis. Data aggregation from sensors assist in managing energy consumption of an IoT environment. The Hadoop system tests and evaluates collected data considering the sensor ID. Subsequently, results are presented to the users via the web service. Incorporating this information, users can manage the appliances to operate in off-peak hours. Figure 2 illustrates an example scenario of the Hadoop system. All data are collected in packet form, which contains sensor ID and the operation time during an hour. According to the ID, the collection unit categorizes data. Management of packet loading to the Hadoop system is handled by the load-balancing module. The processing level transforms sensor data into meaningful information. Noteworthy that large amount data consumes more time and power for processing. Hence, processing enormous data in real-time has become a major challenge. The proposed scheme occupies Hadoop distributed file system (HDFS) for data storing and MapReduce paradigm for data analysis. Hadoop ecosystem incorporates multiple clusters to process data bulks. Existing literature have identified a variety of mechanisms that allocate jobs among multiple clusters in a Hadoop ecosystem. In fact, dividing jobs into sub-jobs is essential for scheduling of MapReduce tasks. However, after loading to MapReduce system, the maximum number of jobs remains unchanged. Considering the fact, we employed a dynamic scheduling mechanism to load jobs into MapReduce adaptively. When transferring a part of job from one node to another, the job tracker evaluates CPU utilization and memory requirements. The job-shifting task is solely depended on the load of the cluster and manages in real-time basis. Nevertheless, job transferring during

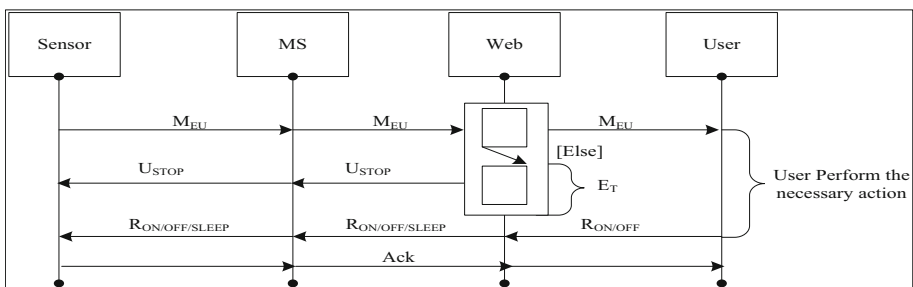


Fig. 3 Working of the ASS mechanism

processing time is not allowed in typical Hadoop ecosystem. Even though in-progress shifting is not allowed, that compromises optimal performance as high performing nodes remain idle and low performing nodes remain in busy state. Hadoop ecosystem become unstable as a result of this disagreement between high performing and low performing nodes. The proposed scheme monitors each node during runtime and allows any node that utilize below 75% capacity to request for new jobs. Owing to this behavior, the job scheduler allocates jobs according to the workload of each node. Consequently, every node utilizes the maximum capacity, while improving the system performance. Moreover, MapReduce output is transferred to the HDFS to perform storing and manipulation tasks.

The analysis and decision server show the Hadoop output via the web service. The threshold levels are separately defined for energy consumption of the building, home, and office. Once the web service notices an energy consumption record that surpasses the threshold, it immediately notifies the user.

4 Results and Data Analysis

The proposed system is tested using various data sets from authentic sources for various services such as parking and roads congestion data, environmental data, energy consumption data of a smart home. Various authentic sources obtained These data sets by installing sensors at various places and location around the world. For example, the data collected for parking lots is obtained from Aarhus city, Denmark at various daytime. This dataset is collected from 125 measuring points using a Bluetooth sensor attached to the mobile car from one particular point to another point. Similarly, the environmental data obtained from different sources consisted of various information of different gasses such as Ozone, Sulphur dioxide, etc. Similarly, another dataset contains buried container data at various locations in Aarhus city, Denmark. The road congestion data can help the citizens of selection the best possible route to a particular destination in real-time. However, it is a difficult job to process such huge amount of data in real-time i.e. more than 100 GBs in real-time. The data sets of road congestion are obtained from various roads of Madrid city, Spain. The data is collected by installing various sensors at different locations on roads. Disseminating the information of intensity of traffic with citizen can help the citizens of planning which road can be used to any particular destination. Similarly, the smart home energy consumption data is obtained by placing and installing sensors with various appliances in different rooms and kitchen. The data obtained through a single smart home is tested for designing sophisticated algorithms for optimal energy consumption. The researchers can process the data in real-time to test various existing algorithms for better appliances usage and load balancing techniques. However, it is only possible with the help of analyzing data from a number of smart homes in a city. Therefore, in this research work, we used the proposed scheduling algorithm on the data obtained from various houses using programmable sensors. The dataset is investigated for energy consumption of various appliances such as television, refrigerator, microwave oven, etc. as shown in Fig. 4. Figure 4 shows the energy consumption of heavy duty appliances such as refrigerator and television are high during specific daytimes. Similarly, the energy consumption of the appliances such as washing machine and microwave oven is high at some specific time but zero at other time of the day. The results of Fig. 4 can be used for designing sophisticated and efficient algorithms for load balancing among smart home appliances in real-time.

The packet loss rate is calculated using different sensors in a smart home scenario with different rooms and a kitchen. The results are calculated in two different scenarios, i.e. pure WSN and WSN with relay nodes support. Both pure and relay based WSN follows the IEEE802.15.4e protocol for data exchange. Figure 5 shows that the pure WSN requires more energy as the sensors are assumed in on-state most of the time. However, the relay support decreases the energy consumption because of transferring most of the data to the MS with the help of relay nodes. Similarly, we performed an experiment of grouping the sensors based on the distance from the MS. The sensors near MS is less affected by the congestion and, therefore, the packet failure rate of such sensors is less compared to the farthest sensors.

In order to test the Hadoop ecosystem functionality, a parking dataset is investigated for two different purposes, i.e. vehicles on different roads in Aarhus city, Denmark and empty parking lots in Fig. 6a, b as an example to check the performance of the proposed architecture of disseminating the data with the citizens in real-time. The data collected from various authentic sources of smart cities [29–31]. The number of vehicles can be used for disseminating the information of congestion on roads. For example, if there are four different lanes between two cities. Then a citizen can choose the road with less congestion based on the processing of data obtained from different roads in real-time. Similarly, a citizen can check free parking lots in various parking lots based on the data obtained from parking lots located in a city.

Further, the information of vehicles congestion on different roads can be used for planning roads and streets. Moreover, in future, the municipality and highway construction authority can design efficient vehicular transportation system based on analysing data from various data authentic sources.

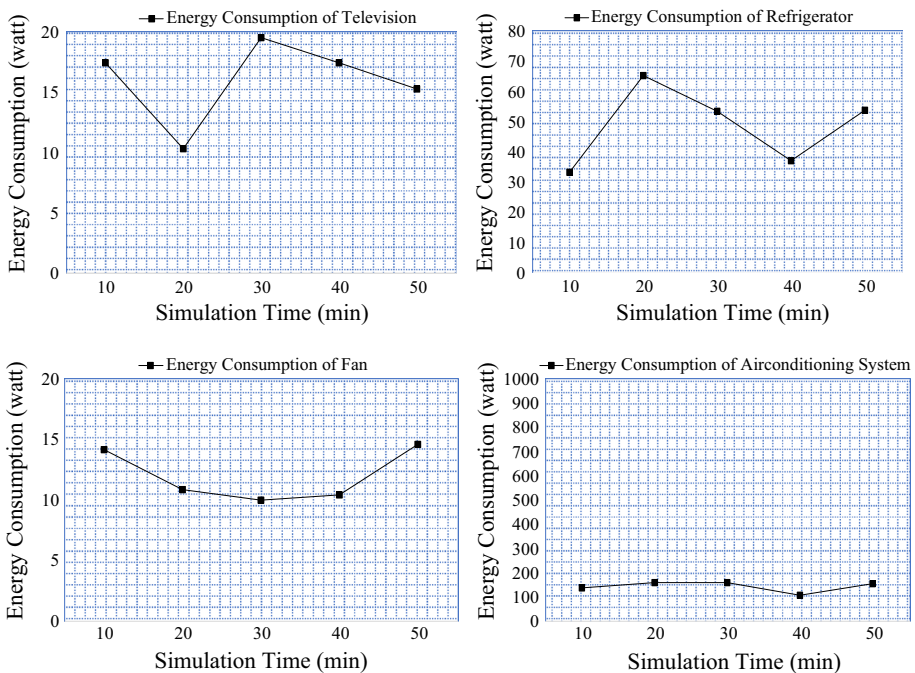


Fig. 4 Appliances energy consumption after testing and applying load balancing module

Fig. 5 Packet failure rate analysis of relay based WSN and pure WSN

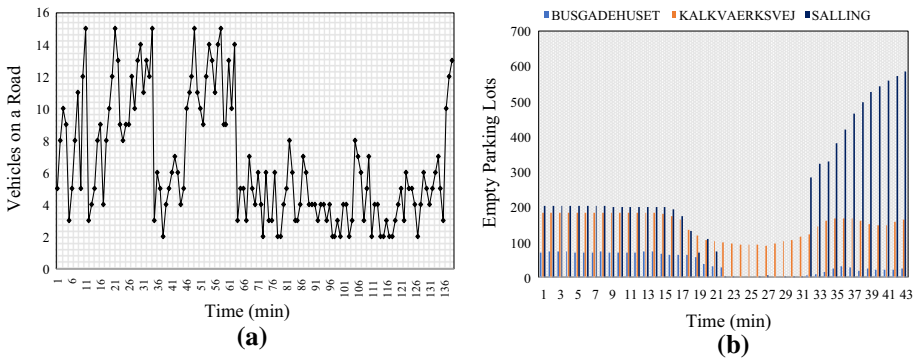
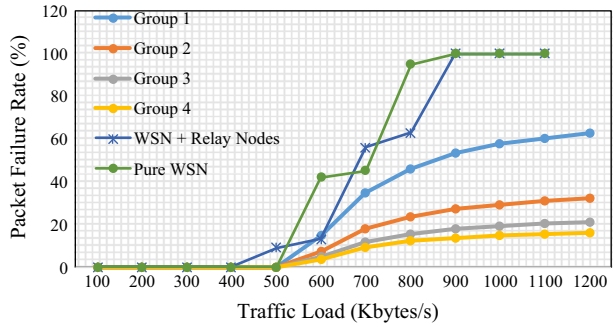


Fig. 6 **a** Number of vehicles on different roads, **b** empty parking lots in various locations in Aarhus city, Denmark

5 Conclusion and Future Directions

In this article an analysis has been carried out for checking the performance of IEEE802.15.4e protocol in a smart home scenario. Similarly, various datasets are investigated from authentic sources of smart cities for vehicular data. These datasets are thoroughly investigated using Hadoop ecosystem for disseminating the congestion of roads and empty parking lots with the citizen in real-time. Similarly, a detail description of the existing MAC protocols based on IEEE802.15.4 is investigated for choosing the right protocol for IoT scenarios. An architecture for IoT environment is designed to handle the energy consumption and processing the big data in real-time. The proposed architecture consists of four main steps i.e. discovering IoT appliances in a smart home, deployment and configuration of sensors as a management station and normal sensors, load balancing and scheduling, and information gathering and processing using Hadoop ecosystem. In future, we will implement the system on real sensors and hardware to get accurate results both in term of efficiency and performance.

Acknowledgements This work was supported by Institute for Information and communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No. 2017-0-00770). This study was supported by the BK21 Plus project (SW Human Resource Development Program for Supporting Smart Life) funded by the Ministry of Education, School of Computer Science and Engineering, Kyungpook National University, Korea (21A20131600005).

References

1. Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., & Zhao, W. (2017). A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications. *IEEE Internet of Things Journal*, 4(5), 1125–1142.
2. Baccarelli, E., Naranjo, P. G. V., Scarpiniti, M., Shojafar, M., & Abawajy, J. H. (2017). Fog of everything: Energy-efficient networked computing architectures, research challenges, and a case study. *IEEE Access*, 5, 9882–9910.
3. Bari, M. F., Chowdhury, S. R., Ahmed, R., Boutaba, R., & Mathieu, B. (2012). A survey of naming and routing in information-centric networks. *IEEE Communications Magazine*, 50(12), 44–53.
4. Bouk, S. H., Ahmed, S. H., Kim, D., & Song, H. (2017). Named-data-networking-based ITS for smart cities. *IEEE Communications Magazine*, 55(1), 105–111.
5. Gou, L., Zhang, G., Bian, D., Zhang, W., & Xie, Z. (2016). Data dissemination in wireless sensor networks with instantly decodable network coding. *Journal of Communications and Networks*, 18(5), 846–856.
6. Silva, B. N., Khan, M., & Han, K. (2017). Internet of Things: A comprehensive review of enabling technologies, architecture, and challenges. *IETE Technical Review*, 1–16.
7. Shojafar, M., Cordeschi, N., & Baccarelli, E. (2016). Energy-efficient adaptive resource management for real-time vehicular cloud services. *IEEE Transactions on Cloud Computing*, 99, 1.
8. Xia, F., Wang, W., Bekele, T. M., & Liu, H. (2017). Big scholarly data: A survey. *IEEE Transactions on Big Data*, 3(1), 18–35.
9. Jabbar, S., Khan, M., Silva, B. N., & Han, K. (2016). A REST-based industrial web of things' framework for smart warehousing. *The Journal of Supercomputing*, 2016, 1–15.
10. Barker, S., Irwin, D., & Shenoy, P. (2017). Pervasive energy monitoring and control through low-bandwidth power line communication. *IEEE Internet of Things Journal*, 4(5), 1349–1359.
11. Daneels, G., Municio, E., Spaey, K., Vandewiele, G., Dejonghe, A., Ongenaes, F., Latré, S., & Famaey, J. (2017). Real-time data dissemination and analytics platform for challenging IoT environments. In *Global information infrastructure and networking symposium (GIIS)*.
12. Khan, M., Silva, B. N., & Han, K. (2017). A web of things-based emerging sensor network architecture for smart control systems. *Sensors*, 17(2), 332.
13. Amadeo, M., Campolo, C., Quevedo, J., Corujo, D., Molinaro, A., Iera, A., et al. (2016). Information-centric networking for the internet of things: Challenges and opportunities. *IEEE Network*, 30(2), 92–100.
14. Zhang, Y., Duan, L., Sun, C.-A., Cheng, B., & Chen, J. (2017). A cross-layer security solution for publish/subscribe-based IoT services communication infrastructure. In *IEEE international conference on web services (ICWS)*.
15. Kim, S., Lee, K., & Jeong, J. P. (2017). DNS naming services for service discovery and remote control for Internet-of-Things devices. In *International conference on information and communication technology convergence (ICTC)*.
16. Gazis, V. (2017). A survey of standards for machine-to-machine and the internet of things. *IEEE Communications Surveys & Tutorials*, 19(1), 482–511.
17. Ju, Q., & Zhang, Y. (2017). Clustered data collection for internet of batteryless things. *IEEE Internet of Things Journal*, 4(6), 2275–2285.
18. Maiti, P., Sahoo, B., Turuk, A. K., & Satpathy, S. (2017). Sensors data collection architecture in the Internet of Mobile Things as a service (IoMTaaS) platform. In *International conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC)*.
19. Luong, N. C., Hoang, D. T., Wang, P., Niyato, D., Kim, D. I., & Han, Z. (2016). Data collection and wireless communication in Internet of Things (IoT) using economic analysis and pricing models: A survey. *IEEE Communications Surveys & Tutorials*, 18(4), 2546–2590.
20. Trihinas, D., Pallis, G., & Dikaiakos, M. D. (2017). ADMIN: Adaptive monitoring dissemination for the Internet of Things. In *IEEE INFOCOM 2017—IEEE conference on computer communications*.
21. Liu, L., Liu, Y., & Wang, L. (2014). Economical and balanced energy usage in the smart home infrastructure: A tutorial and new results. *IEEE Transactions on Emerging Topics in Computing*, 3(4), 556–570.
22. Li, M., & Lin, H.-J. (2015). Design and implementation of smart home control systems based on wireless sensor networks and power line communications. *IEEE Transactions on Industrial Electronics*, 62(7), 4430–4442.
23. Leu, J.-S., Chen, C.-F., & Hsu, K.-C. (2013). Improving heterogeneous SOA-based IoT message stability by shortest processing time scheduling. *IEEE Transactions on Services Computing*, 7(4), 575–585.

24. Palattella, M., Accettura, N., Grieco, L., Boggia, G., Dohler, M., & Engel, T. (2013). On optimal scheduling in duty-cycled industrial IoT applications using IEEE802.15.4e TSCH. *IEEE Sensors Journal*, 13(10), 3655–3666.
25. Zand, P., & Shiva, M. (2008). Defining a new frame based on IEEE 802.15.4 for having the synchronized mesh networks with channel hopping capability. In *11th IEEE international conference on communication technology*, Hangzhou.
26. Mehta, A., Kerkez, B., Glaser, S., & Pister, K. (2012). TDMA-based dual-mode communication for mobile wireless sensor networks. *Sensors*, 12(12), 16194–16210.
27. Tinka, A., Watey, T., & Pister, K. S. J. (2010). A decentralized scheduling algorithm for time synchronized channel hopping. *Ad Hoc Networks*, 4(49), 201–216.
28. Li, L., Li, S., & Zhao, S. (2014). QoS-aware scheduling of services-oriented internet of things. *IEEE Transactions on Industrial Informatics*, 10(2), 1497–1505.
29. Uppoor, S., & Fiore, M. (2011). Large-scale urban vehicular mobility for networking research. In *IEEE vehicular networking conference (VNC)*.
30. Naboulsi, D., & Fiore, M. (2013). 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*.
31. Uppoor, S., Trullols-Cruces, O., Fiore, M., & Barcelo-Ordinas, J. M. (2014). Generation and analysis of a large-scale urban vehicular mobility dataset. *IEEE Transactions on Mobile Computing*, 13(5), 1061–1075.



Murad Khan received the B.S. degree in computer science from university of Peshawar Pakistan in 2008. He completed his Ph.D. degree from School of Computer Science and Engineering, Kyungpook National University, Daegu, Korea. Dr. Khan is currently working as an assistant professor in the Department of Computer Science and Information Technology, Sarhad University of Science and Information Technology, Peshawar, Pakistan. Dr. Khan published over 40 International conference and Journal papers along with two books chapters in Springer and CRC press. He also served as a TPC member in world reputed conferences and as a reviewer in numerous journals such as Future Generation Systems (Elsevier), IEEE Access, etc. In 2016, he was awarded with Qualcomm innovation award at Kyungpook National University for designing a Smart Home Control System. He was also awarded with Bronze Medal in ACM SAC 2015, Salamanca, Spain, on his distinguished work in Multi-criteria based Handover Techniques. He is a member of various communities such as

ACM and IEEE, CRC press, etc. His area of expertise includes ad-hoc and wireless networks, architecture designing for Internet of Things, and Communication Protocols designing for smart cities and homes, Big Data Analytics, etc.



Kijun Han received the B.S. degree in electrical engineering from Seoul National University, Korea, in 1979 and the M.S. degree in electrical engineering from the KAIST, Korea, in 1981 and the M.S. and Ph.D. degrees in computer engineering from the University of Arizona, in 1985 and 1987, respectively. He has been a professor of department of electrical engineering and computer science at the Kyungpook National University, Korea since 1988.



Dr. S. Karthik is presently Professor and Dean in the Department of Computer Science and Engineering, SNS College of Technology, affiliated to Anna University- Chennai, Tamilnadu, India. He received the M.E. and Ph.D. degree from Anna University, Chennai. His research interests include network security, bigdata, cloud computing, web services and wireless systems. In particular, he is currently working in a research group developing new Internet security architectures and active defense systems against DDoS attacks. Dr. S. Karthik published more than 96 papers in refereed international journals and 125 papers in conferences and has been involved many international conferences as Technical Chair and tutorial presenter. He is an active member of IEEE, ISTE, IAENG, IACSIT and Indian Computer Society.