Peer-reviewed version available at *Sensors* **2018**, *18*, 3980; doi:10.3390/s18113980

*Article*

# A Novel Framework and Enhanced QoS Big Data Protocol for Smart City Applications

**Shalli Rani [1,*] and Sajjad Hussain Chauhdary [2]**

[1] Chitkara University Institute of Engineering and Technology, Chitkara University,Rajpura(140401), Punjab,India

[2] Department of Information and Computing Technology, University of Jeddah, Kingdom of Saudi Arabia; chauhdary.sajjad@gmail.com

* Author to whom correspondence should be addressed; shalli.rani@chitkara.edu.in, 9914080394.

**Abstract:** Various heterogeneous devices or objects shall be integrated for transplant and seamless communication under the umbrella of internet of things (IoT). It would facilitate the open accession of data for the growth of a glut of digital services. To build a general framework of IoT is very complex task because of heterogeneity in devices, technologies, platforms and services, operating in the same system. In this paper, we mainly focus on the framework for big data analytics in smart city applications , which being a broad category specifies the different domains for each application. IoT is intended to support the vision of Smart City, where advance technologies will be used for communication for the quality life of citizens. A novel approach used in this paper, is for enhancing the energy conservation and to reduce the delay in big data gathering at tiny sensor nodes used in IoT framework. To implement the smart city scenario in terms of big data in IoT, an efficient (optimized in quality of service) WSN is required where communication of nodes is energy effcient. That is why, a new protocol QoS-IoT is proposed on the top layer of the architecture which is validated over the traditional protocols.

**Keywords:** Energy Efficiency;Big Data Analytics; QoS-IoT; Internet of Things; Smart City ; WSN ; Green Computing

**1. Introduction** Since the last decade, tremendous efforts have been made to explore the Internet of Things (IoT) applications, where communication between heterogeneous physical objects, and sensors could inter-operate in integrated environment or platforms. The contemporary advancements have given a boost to the idea of smart city. It has been perceived as the intellectual and open environment where citizens can enjoy the quality life, further giving an enhanced research in IoT technology and standards providing a building block to these application scenarios. To facilitate an intelligent communication among the different objects of the smart city, there is growing need of gathering, analyzing, processing and presenting data continuously from all the devices with partaking sensing mechanisms.

This bustle requires the substantial efforts to tackle the heterogeneity in perspectives, to deal with open, dynamic and hazardous deployment conditions. A pertinent objective is to develop a novel, autonomous and compliant model for smart cities, which may interact with diverse domains of the applications like security surveillance, environmental monitoring, structural monitoring etc. which could facilitate citizenship and urban living. Powered by the adaptation of various technologies such as sensor nodes (SNs), the IoT has stridden out of its beginning phase and is revolutionizing the current frameworks into integrated structures of IoT and wireless sensor networks (WSNs, strong and base layer of IoT framework). The huge amount of data (Big Data) gathered by the sensing nodes will be distributed across the varied applications and platforms for the public picture of smart city. There is no formal definition of smart city, however we can presume its aim is to develop a city for the efficient use of resources and to enhance the quality of service for the people. This objective can be gained with the help of IoT where WSN plays a crucial role.

A metropolitan IoT will bring a lot of benefits in the administration of public and traditional services like, transportation, parking, lightning, agriculture, safety etc. The availability of diverse types of data gathered by metropolitan IoT will be exploited to enhance the transparency and local government would be able to take the actions accordingly. Participation of the citizens will be enhanced and will stimulate the invention of novel services. Moreover, as in the industrial survey [1], authors have reviewed the literature from the industrial perspectives and have mentioned the need of metropolitan IoT for the small to large business organizations which range beginners to the well settled corporations. Need of the IoT in industrial marketplace, trends and openings are elaborated in it.

IoT is transforming heterogeneous objects from being traditionally intelligent by integrating its abundant and prevalent computing, embedded devices (e.g., actuators, smart phones, networked enabled smart devices etc.), sensor networks, Internet/WSN protocols and applications to transform human life. In the upcoming times, it will integrate billions of devices along with objects and embedded systems. As a consequence, the Internet of Things (IoT) will considerably increase in scope and size, presenting new challenges as well as opportunities [2] .

The IoT is reflected as the future prospect for the realm of the Internet. Consequently, it leads to the rise of the concepts like smart homes, smart transportation, smart agriculture, smart health, smart buildings etc. converting the traditional cities into smart cities where diverse electronic appliances and objects are interconnected with each other to gain high-quality 2-way interactive multimedia services. In these cities, where huge number of devices will be connected, a massive quantity of data will be generated known as Big Data. To enrich the smart cities applications the Big Data analytics can play a crucial role in the advancements of information and communication technologies (ICT). The Big Data gathered from various places gives a better understanding of present as well as future and leads to accurate planning and development. Analysis of Big Data provides an insight into the major

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 3980; doi:10.3390/s18113980



perspectives of the smart applications. Smart city applications can consider the integration of sensors and actuators which are auto configurable which can be controlled remotely with the assistance of internet. A large number of sensors sense the huge amount of data and transmit it to the remote center where it is processed, analyzed and predicted or used to respond the userâĂŹs query. A lot of studies [3-8] have been conducted in recent past on Big Data. The notion of the Smart City is extendable to the Smart World where the Countries Domain, Cities Domain, and Service Domain are integrated to benefit people all around the world. However, for the Smart City concept, various other concepts like inter-connected vehicles, GPS, roadside units etc. are still under development. In future, around 70

Therefore, data analysis will be based on the userâĂŹs requirements and priorities which will make cities smarter. Hence, diverse technologies used in applications and their data analytics are driving the IoT out of its initial stages and revolutionizing it, therefore IoT is migrating from traditional framework to fully integrated novel future network.

The WSN and its various technologies are soundly integrated into a metropolitan infrastructure, forming a digital skin [9]. The huge amount of information generated by the prevalent and embedded devices will be shared across the different platforms and applications to convert the cities into smart cities and forecast the development. Typically to achieve the urbanization, it is important to realize the requirement of service outlining to improve the efficiency and city management schemes. Recently many organizations are in the queue for setting up platforms for live monitoring, scheduling and collecting urban process attributes. All these activities are accomplished by gathering data both in both ways; offline and real-time (for Big Data processing, analysis and decision making). Data gathering techniques are costly and complex. Consequently, there is a requirement to include intelligent technology which can proficiently and speedily gather a huge amount of data, accomplish analyses on Big Data, and envisage the future to enable improved planning and development [10-11].In the consideration of the above facts about IoT and the role of Big Data in smart city applications, in this paper, we propose a smart city big data protocol for M-IoT where WSN plays a fundamental role in gathering the huge amount of data. A complete architecture to develop the smart city applications is presented which is based on IoT based Big Data Analytics.

This protocol will work on the basis of data which will be managed and gathered by the sensor nodes in WSN. However, WSN is constrained by many issues, which should be resolved for the real time communication in M-IoT. Most of the applications (environment monitoring, border security surveillance, transportation etc.) in M-IoT, require the energy efficient and real time communication. To overcome the various challenges faced by the Big Data, a lot of research has already been done but not from the perspective of scalability (IoT applications will cover a huge area of city). Traditional protocols developed for WSN will not provide the strong platform to M-IoT; therefore a novel approach is required to cater present need. In the previous work of WSN [12], areas are monitored on the basis of queries. Queries based applications are the most prominent part of the M-IoT. However, queries (users queries from the data server i.e. Big Data) are not handled in the accurate way and it is not suitable for real world scenarios for the two reasons. 1) They have not considered the priority of the query messages. As a consequence query responses will be dropped off along with query requests. Due to this whole system become malfunctioned. Most of the work of smart cities is based on the Big Data only. Smart city citizens require the accession of real time Big Data, hence loss of messages and delay in response will

Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



not be tolerable caused due to WSN constraints. 2) The header nodes are dynamic which are responsible for the forwarding of data. Furthermore, all the nodes have different levels of energy. However, static and homogeneous nodes can simplify the processes in the real world applications.

In this paper we have proposed a framework for Smart City Big Data analytics and an enhanced quality of service based protocol (QoS-IoT) for IoT applications. 5 layer architecture is proposed, which is capable of analyzing the huge amount of dataset of IoT. Moreover, the timely delivery of data makes the decision making process efficient. The proposed system is evaluated in terms of throughput, energy and transmission time.

Rest of the paper is organized as follows: in the 2nd section background and motivation is discussed. In section 3, we have presented the technologies used in Big Data. Framework for of IoT based Smart City with Big Data management is provided in section 4 followed by the proposed model in section 5. In, Section 6 concluding observation is presented.

## 2. BACKGROUND AND MOTIVATION

Due to the digitization, cities are converting into smart cities, where cities are equipped with various electronic equipment used by many applications like street surveillance cameras, transportation monitoring, intelligent health system etc. There are some application which use already established systems to handle the smart concepts like, Global positioning System (GPS) and others. It adds on the power to the handheld devices to generate a scenario where data is exchanged on the basis of locations. Therefore in this type of environment, the queries based on features, objectives, security etc. required to be resolved [16]. These can be outlined as below:

1. Reliability: How to handle the uncertainty due to real time constraints and offline dynamics. How the quality of the data can be ensured.

2. Intelligent Transmission: How objects can become intelligent to transmit the data and designing of the new objects.

3. Less Delay: How the response time can be reduced.

4. Data Gathering: Data should be gathered in the lowest possible time and at low cost (in terms of energy, fault tolerance, throughput, standard deviation etc.). How this objective can be gained when there is large number of sensors is deployed around the environment.

5. Processing of Big Data: Which schemes can be implied for the big data processing and analytics (online data analytics (OLAP), online data transaction processing (OLTP) for real time communication.

Based on these questions, the Smart City applications exploit the capabilities of ICT which help the citizens in the efficient utilization of limited resources. Various organizations are using the advance systems and majority of these technologies are comprised of efficient storage skills and advance sensing schemes for exceptional quantity of data. The fundamental objective behind the present approach is the vast deployment of sensor nodes to gather the data in an efficient way. The design of this system needs all the features of WSN to analyze, gather and transmit the data of all the objects. Integration of large amount of data poses many constraints and efficient techniques of Big Data analysis are already developed. However, for large scale applications, in some circumstances (posed by environment) a major portion of the data can be disjointed. It can lead to the unreliability and it requires better urban planning and a novel scheme to avoid it. Smart M-IoT framework can provide a new way to integrate the resources on the basis of geographical locations and can then be analyzed by novel system to provide diverse services to the citizens. Citizens can be benefited in terms of health, security, mobility, pollution etc. Many projects are available in literature like smart car parking, dynamic lightning system, car locking system etc. where sensors are playing the major role (specific to the applications). Other service oriented applications for smart city applications in IoT are vehicle to vehicle communication, air/noise pollution, health care systems, real time driver assistance, potholes monitoring etc. Recent research is very constrained due to simple design issues of IoT. No single system is fully developed and authorized to support the scalability and efficient communication (Big Data transmission and gathering). To derive the knowledge about the various aspects of Smart City Big Data analytics of old as well as new information is required. Therefore the role of WSN in Big Data analytics cannot be ignored.

WSN provides all the necessary data to the infrastructure of IoT which is used in all the smart city applications. This integration is particularized in [13]; however, focus of the paper was on network only. Data, which has the most crucial role and around which whole system of IoT revolves, is not considered. The context and instructions of communication for data are emphasized by the authors [14]. The necessity of semantic gloss of IoT is outlined in it, which motivated us to propose a model approach where sensor nodes can slog with IoT framework and attention is given to the data. Overlays are formed over WSNs in [15], for the upliftment of urban data in IoT and proposed a solution for the delivery of urgent data. But they have not considered the auto response from the nodes, which can reduce the further delay in the process. A framework for urban system is proposed in [16] based on the support of sensors and network, goes through the data and cloud integration of heterogeneous devices. It focused on the noise based smart city application and can be applied to other services. A survey on the protocols, standards, and architecture for urban IoT has been made in [17]. It presented the technical solutions and guidelines for the Padova smart city project. Looking into the problem of manhole cover, authors have presented the solution with an automated monitoring system which is a fragment of smart city and IoT [18]. Underground security issues for manhole cover are conversed with automated and non-automated system. A distributed sensing algorithm which deploys the sensor network, on the top of IoT devices is proposed in [19] which are based on parameters of sensors. Event driven architecture for IoT domain is proposed by authors in [20]. Different approaches of event triggered methodologies are also conferred to address the efficient phases of event driven data in IoT. In this paper, we are presenting the framework for the continuous monitoring of smart city with reliable and fast communication system, which will function over the top of IoT in energy efficient way (essential requirement of Big Data). The comparison with energy efficient algorithm for IoT (ME-CBCCP [21]) and other traditional algorithms as presented in [22] provides the validated insight into the proposed scheme.

## 3. CHALLENGES AND MANAGEMENT OF BIG DATA USING EMERGING TECHNOLOGIES

Peer-reviewed version available at *Sensors* **2018**, *18*, 3980; doi:10.3390/s18113980



In the processing, analyzing, filtering etc. of gathered data management is vital process. The problem which was raised a long ago for the first time during the efforts of UK e-science where data was distributed over geographical regions and was owned by various entities [4]. To handle this problem, scientific data life cycle management (SDLM) model was proposed. This model analyzed the already existing approaches from different perspectives. The traditional model follows the same steps, to handle the data, planning, data gathering, filtering, processing, feedback and documentation [58-60]. Following is the discussion of the generic stages used in the Big Data handling:

*3.1. Raw Data* Agencies, data server centers, researchers and organizations incorporate the gathered raw

data and enhance the value of data by using inputs of individual developed programs and novel reach projects. The data is converted from the early stage and stored in the form of value added services. No standard has been globally accepted to store and administer the data. The data is generated with specific attributes which depends upon the programming done to handle the data.

*3.2. Data Processing* Sensor nodes deployed in WSN, gather the data which is the very first stage of the

Big Data life cycle. In IoT data is also gathered from the mobile phones, satellites, laboratories, blog messages etc. Specific techniques are implied to gather the raw data from the integrated environment of IoT. Data can be processed before transmission (sensor nodes) or at data center (data server or base station (BS)). Transmission of processed data consumes less energy and time in data transmission which are foremost constraints of WSN. Data generation in IoT depends upon the requirements and life of the citizens. Data in IoT is also based on expressions, habits, emotions which are accumulated with the help of internet and sensing technology again plays an important role at this place. This type of data is required in health departments. Scientific Data Infrastructure (SDI) organization must [58, 59] must consider the problem of heterogeneous data. Generally following methods are used to gather the data [23]:

(i) Sensing: Sensors are deployed randomly or manually at fixed locations to gather the data to measure the physical attributes, and converted into the digital signals for transmission and storage. This data may be of any type temperature, humidity, pollution, chemical, medical, security etc. Sensed data is transmitted to the BS with the help of wired or wireless networks. For smart city projects, both networks are required. The wired network is the network of base stations or the data centers where the gathered data (Video surveillance, patient health monitoring etc.) is processed and data is gathered with the help of wireless networks from remote areas. That is why WSN in known as the backbone of IoT. WSN has gained a lot of attention and has been used in many fields like underwater monitoring, civil engineering, wildlife, health care, surveillance etc. Various techniques are developed to utilize the capabilities of WSN in efficient ways.

(ii) Data Capturing Techniques: Data in IoT is captured by integration of task, word segmentation, and index and web crawler. In search engines, web crawler stores/download the pages and other linked information is accessed through the uniform resource locater (URL). Data from various applications is accessed by caching and search engine optimization techniques. Many techniques have been proposed

Peer-reviewed version available at *Sensors* **2018**, *18*, 3980; doi:10.3390/s18113980


in recent past for efficient extraction strategies (for high quality data) to address the numerous internet applications.

(iii) Log Files: Data is collected automatically by recording the manipulation steps and processing via data source management system. Log files are maintained by all the companies (IBM, Infosys, and TCS etc.) where the work is carried out on digital data. The advantage of log files is in back up of the system, in case of any failure. Timestamps are used to roll back the crucial transactions. Data is also recorded on the web by userâĂŹs visits on the web pages, clicks on links etc. Data is in all the formats of ASCII to enhance the query efficiency in huge data warehouses.

(iv) Zero-Copy (ZC) Packets: Data copies exchanged between the nodes are not copied due to this technology. Data is generated from the usersâĂŹ nodes and routed through the network interfaces and sent to the BS. Data is transmitted to the usersâĂŹ nodes which are accessed from the BS. Data is not copied in between the system calls. Direct memory access (DMA) is used to avoid the data copies and reduces the number of system calls, hence time.

(v) Smart Phones: Mobile phones are converted into smart phones and gradually they are becoming more powerful. Data acquisition techniques are enhanced and various different parameters are produced. These devices are capable of gathering information about weather, health conditions, location, and can capture multimedia data. Mobile internet technology is becoming popular among the people due to its capability of gathering and transmitting Big Data in smarter way.

Along with various methods mentioned above, other technologies such as cloud computing is also assisting in handling the data processing.

*3.3. Analysis of Data* This process empowers any organization to gather the plentiful information, which

can affect the various processes of business. This process is very complex due to heterogeneity of data and scalability of algorithms. It helps the in understanding the relationships among the data and their features. It enables to develop the new methods of data mining to predict the future interpretations. Over the time, new techniques have facilitated the speedy accession and mining of both types of data; structured and unstructured. Analytical techniques can be categorized into data mining, statistical analysis, visualization and machine learning. Data mining is useful in many application of medical and engineering. Gathering of Big Data involves different formats of data and data should be accessed in minimum possible time. An efficient architecture should support all the techniques of data analysis and formats of data. High performance of algorithms and protocols which are developed for WSN or for IoT can help in achieving these objectives. Many traditional data analysis techniques can be applied on Big Data Analytics such as: Cluster analysis, data mining techniques, correlation analysis, statistical analysis and regression analysis.

With the increase in the size of data, new methods should be developed to gather, store, analyze and process the Big data in efficient ways. Various challenges in the Big Data analytics can be described as below:

1. Heterogeneity of Data Formats: Algorithms of data mining search the unknown patterns and various homogeneous formats for the analysis in structured way. Still, analysis of semi/unstructured formats is complicated. Data should be structured before the analysis process. Variety of data is the big

Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



issue in case of Big Data. Information of data may not be structured and may not be well organized (relational database) as data is collected from the various sources. 2. Correctness: Data gathered from the small sets of applications can be seen or perceived as the accurate data. However, Big Data collected from the various resources cannot be considered as the accurate data as volume of data adversely affects it. 3. Scalability: A huge amount of data, itself is the issue of Big Data analysis. This issue can be mitigated with help of processing speed. Still, amount of data increases than the CPU and resources speed. Many computing resources are shared by the nodes along with memory and processor. Scalable systems are required to deal with these kinds of issues. 4. Data Complexity: Different types of data, such as structured, unstructured and semi structured pose another challenge in front of Big Data Analysis. As opposite to unstructured, structured data have similar format, predefined range/domain and processes. Data is generated with the help of sensors or computers without human intervention. Data is processed with database query languages such as SQL. However, Big Data gathered from the multimedia resources is unstructured and softwareâĂŹs like Hadoop can handle this type of data. Hadoop analyzes and clusters the unstructured and semi structured data with the help of MapReduce. Extraction of important data is serious challenge. It is difficult to validate all the data items in Big data because data sources are varied temporally as well as spatially in gathering and format. The metadata description of data cannot be controlled and may or may not be accurate. It requires inspection and critical analysis of data.

*3.4. Publishing and Sharing of Data* To benefit the public, governments, researchers, agencies etc. data

and its resources are gathered, analyzed and published. Large and wide datasets of Big data must be stored with easy accessibility, reliability and availability. Storage space should be managed to apply DBMS techniques efficiently. Stored data should be consistent (with the help of data center), and easily available. Data should not be lost due to network obsoleting (main constraint of WSN).

*3.5. Privacy and Security* Citizens of smart city, require that their data on cloud should be safe and

private. Intellectual property rights should be developed to maintain security, privacy and confidentiality of data. E.g. data of patients in the hospitals is confidential and if the Big data techniques and IoT protocols are involved in the patient data monitoring, then some security techniques are required to manage their digital data. Data integrity is also an important factor for joint analysis where data analysts and decision makers share information. Data mining techniques are required to improve decision making processes and cooperative tasks on Big Data.

*3.6. Discovery and Re-usability of Data* Data accession schemes are required to ensure the quality,

validation, value addition, and data conservation by reuse of existing data and discovery of new data. It involves many fields within it, like archiving, representation and data management. Schema and relational model are used for the structured data management and re-usability of data.

As aforementioned section discusses about the various techniques and issues involved in Big Data, some of the issues can be solved by the use of novel methods and techniques. In the next section we

Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



are presenting the proposed framework to handle the data in energy efficient (in energy constrained environment of WSN) ways for Smart City applications.

## 4. A FRAMEWORK FOR BIG DATA ANALYTICS IN SMART CITIES APPLICATIONS OF IoT

IoT based Smart City should facilitate the fast accession of the data along with accurate information. For the Smart City development, several sensors are required to gather the information along with other IoT devices. To connect the Smart city devices and IoT system, other devices are required like Edge nodes, aggregators, data server etc. Sensors generate the information at high speed and Hadoop system is required for the same. As per the requirement and constraints of the WSN and IoT, we have deployed the QoS novel protocol to gather the data efficiently (Big Data). Architecture of the proposed system is discussed below to show the efficiency of the system.

### *4.1. IoT Based Smart City*

The main challenge in front of Smart City projects is how to use the IoT system in these projects. IoT is connecting various heterogeneous objects and data is coming from the many heterogeneous resources. In the digital era, all the devices are required to connect with high speed internet and data from sensors should come with in shortest possible time. The main aim of deployment of sensor nodes is to gather the information even from remote areas. Smart parking, pollution measurements, detection of unsafe events, smart driving, smart homes, smart streets, smart health monitoring etc. is some of the applications which are useful for citizens of smart cities. Sensors play a major role in implementation of these applications. Sensors are constrained by many features like energy, scalability, processing and low memory. Some techniques are required to overcome these constraints, as in Iota system these constraints pose major challenges. Data for all the applications is required for real time analysis. However, for real time data and processing, applications should gather the data in minimum possible time and it should be reliable. These applications not only require the location information (GPS), but also require other information as temperature, humidity, patient symptoms, security intruders etc. Proposed approach gathers the information in energy efficient way, without network obsoleting. The architecture of the smart city can be perceived in the figure 1, where utility programs are installed on the topmost layer of framework. These protocols exploit the benefits of IoT by optimizing the existing techniques of data gathering, transmission, processing and analysis. These protocols are required for the accurate working of the applications by reducing the standard deviation and enhancing the reliability. Hadoop and Mapreduce are required for Big Data analytics along with supporting technologies. Data sources are the base of this framework where sensors are deployed to generate the data.

### *4.2. IoT-Based Smart City Planning and Implementation*

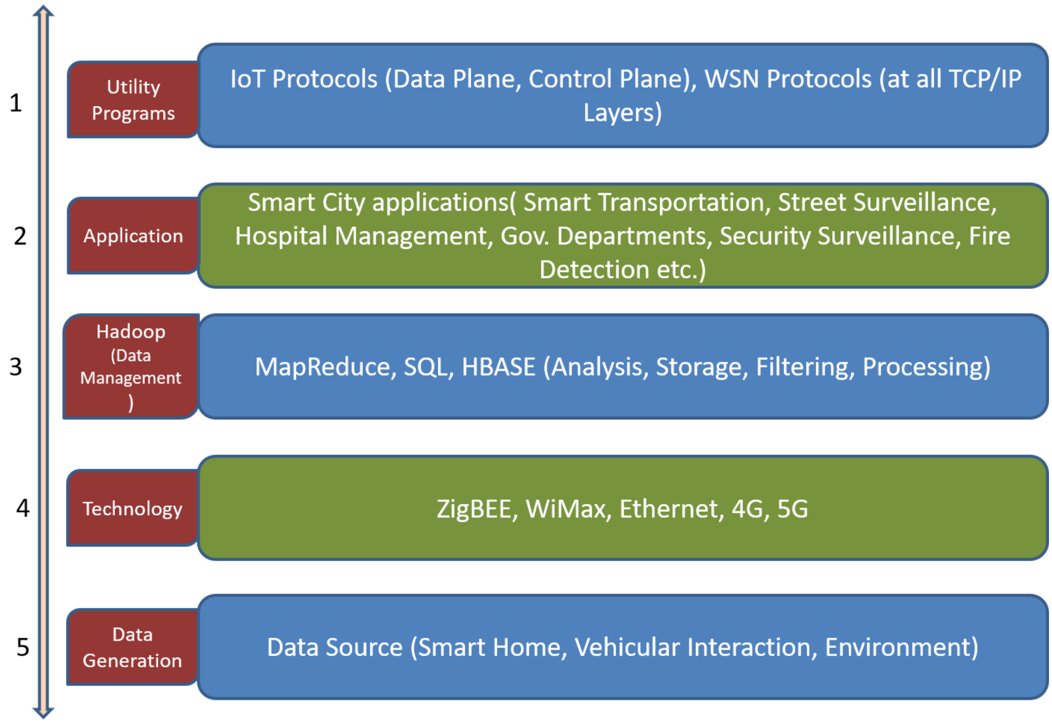
For planning and implementation of the novel framework, we have considered layers proposed in the figure 1, with one exception that data is generated by the sensor nodes (not from log files, auto systems etc. as mentioned in section III). To propose the optimized protocol as a utility for IoT smart city projects,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 3980; doi:10.3390/s18113980



**Figure 1.** Structure of Smart City Big Data Analytics




a novel QoS based protocol is developed for the sensor nodes for the cooperative data transmission, for energy efficient communication (green computing in IoT). This approach is necessary to fulfill the demands of the citizens. E.g. to build the smart streets and smart homes scenarios, government analyze the needs of people by analyzing the energy consumption from historical data and can plan for better services in future. In another application, for the healthy environment, information about pollutants is required which should be available with the organizations in continuous time which is possible only when sensors are deployed randomly, and data is gathered and transmitted without human intervention in optimal way. In dynamic traffic light system, traffic is controlled by the real time data, where reliability of the information is crucial. All these constraints are related to the sensor nodes and in another way to the IoT as WSN is the fundamental stone of IoT system. In IoT abundant number of sensors is required to control the information and this information acts as the part and parcel for Smart City scenarios. Therefore in this paper, we have proposed a novel approach to transmit the data in efficient way to optimize the energy parameter as shown in figure 2.

### *4.3. Big Data Analytical Architecture*

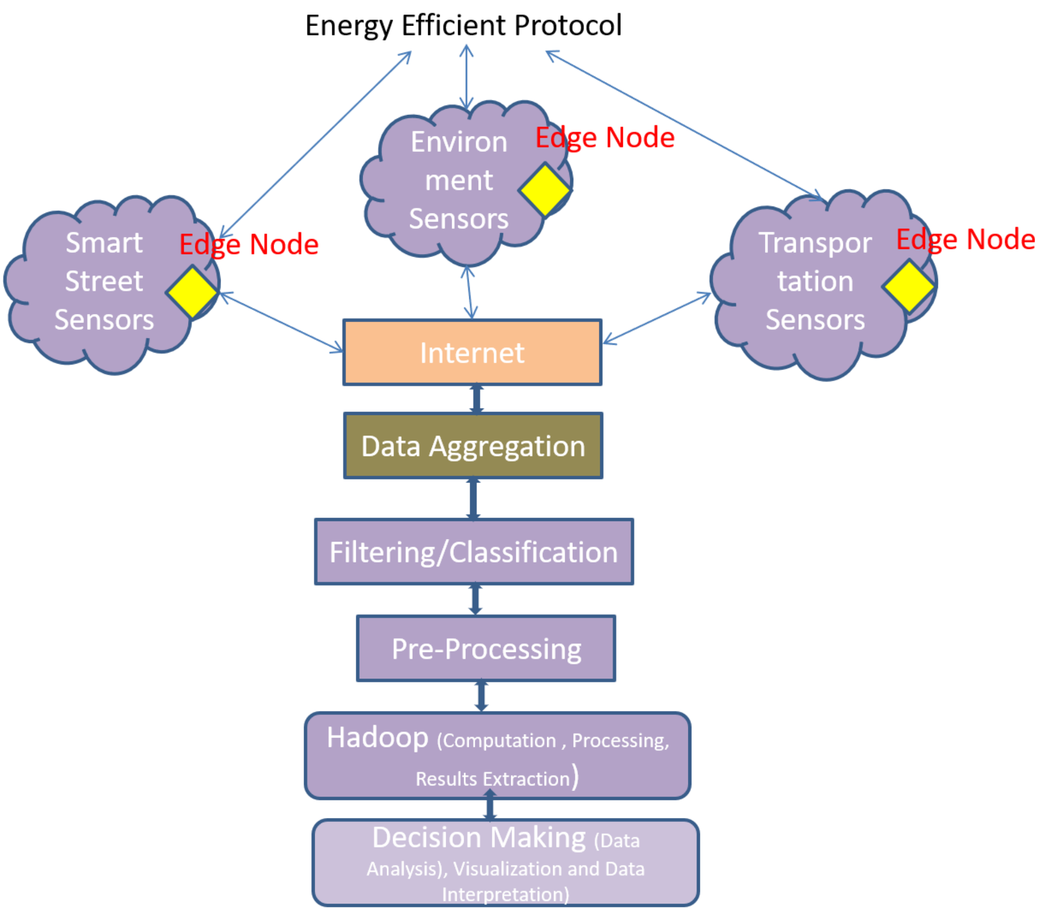
The data is gathered by the sensor nodes from data sources where they are deployed. This module gathers and transmits the data to aggregator system (BS) via internet. Citizens do not have or limited access to the results of data. A five layer system is involved in the Big Data processing as shown in figure 1. The bottom layer is the data source layer, which gathers the data from the smart city resources. This layer is based on the Hadoop system. The data storage is done on the Hadoop system by using the MapReduce . The data is transmitted to the third layer by using the one of technology from ZigBee, WiMax, 4G or 5G. The data is managed by the third layer which is used by the various smart applications at the top. However, data gathering and transmission requires the inclusion of WSN/IoT protocols which

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



**Figure 2.** Implementation Framework



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



have been shown in the first layer. These protocols act as the utility softwareâĂŹs because they reduce the standard deviation, enhance the energy management, network lifetime, scalability and reliability. Afterwards, decision making process is executed on the data which is collected at the data center /aggregator (BS) after the process of data filtering as shown in figure 2. Five tier architecture can be explained in detail as below:

1. Data Source: This layer handles the data, generated by various IoT sources like sensors and objects. Data is gathered at the sensor nodes and this data is processed before data transmission to remove the errors (if any) before data transmission. Data compression techniques can also be applied on the sensor nodes to reduce the required bandwidth for the data and to enhance the data transmission from other nodes (reduce network congestion). The data produces at this layer is heterogeneous data. Along with all other techniques, security can also be maintained at this layer by encryption/ decryption techniques. Data filtration at this layer removes the unnecessary and redundant data. 2. Technology: This layer is responsible for communication among the sensor nodes, the edge nodes and BS, and it relies on the ZigBee, WiMax etc. In IoT there are several BSsâĂŹ connected with each other through internet. 3. Data Management: This layer is required for the data analytics and is responsible for data management. Third party tool is required to integrate Hadoop within the system for implementing the model. All gathered data must be stored at Hadoop by using programming at MapReduce. Analysis of data is performed at this layer. 4. Application layer: The analytical data and reports generated at the third layer are used by the application layer by end users. 5. Utility Programs: WSN and IoT protocols are used at the topmost layer for the efficient working of other layers (data gathering, transmission, encryption, processing etc.) Traditional protocols do not suit the scalable network of IoT. New techniques are required for the efficient working of the architectures. As it is shown in the figure 2, that protocol is implemented for the efficient working of the sensors. Data is collected by the sensor nodes and transmitted to the edge nodes which further transmit the data to the BS. Data from all the network areas and subareas is crucial for excellent and accurate decision making. It requires fault tolerance and energy efficient communication (to conserve energy for long network lifetime). The proposed protocol is implemented in MATLAB and it is validated over BDEG [22] protocol along with other traditional protocols. As IoT deals with a large amount of Big Data, we require a system that will proficiently process and execute large collections of huge datasets. To fulfill these necessities the Hadoop system can be used, which contains various master nodes along with other nodes under it. In Hadoop data is divided into equal amount of portions and stored on data nodes. Parallel processing is performed on these nodes by MapReduce. All the reports are generated at the Hadoop and decision is taken on these results.

## 5. NETWORK MODEL and RESULTS DISCUSSIONS Of PROPOSED PROTOCOL

In this section, the proposed protocol model for smart applications is being discussed. The network model and the radio energy model are two standard models that are used in the data processing in any Smart City application.

*5.1. Network Model* It is assumed that sensor node deployment is randomly uniform in a square shaped area (table 1). Following assumptions are made for all nodes in the network:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



i. All nodes are considered to be static, which means there is no movement of nodes once they are deployed. The main objective of sensor network is that nodes collect data from the environment periodically and send to the base station. ii. All the nodes links are symmetric and have the same initial energy at the beginning. iii. Each node is having the ability to merge the redundant data. All sensor nodes are assumed to have limited batteries and recharging them is really infeasible. iv. Nodes do not possess any GPS equipment and their relative distances are calculated on the basis of received signal strength.

*5.2. Energy Consumption Model* The radio model is used for reception and transmission of an l-bit

message. Communication energy is consumed at much lower level as compare to the energy consumed on computing and storage process. Energy consumption on communication is considered for simplicity by using equations 1 and 2 [21].

$$Etx(l,d) = lEelec + lEefsd^2 Ford < d_o(1)$$

$$Etx(l,d) = lEelec + lEefsd^4 Ford > d_o(2)$$

For reception of messages the radio expands (equation 3):

$$Erx(l) = lEelec(3)$$

To merge the number m number of message, the energy consumption is computed as (equation 4):

$$Edx(l) = mlEda(4)$$

In Equations (1), (2) and (3) , Eelec represents the energy consumption of transmit or receive 1 bit message. In equation (4) , Eda represents the energy consumption of merge 1 bit message.

Here do shows the threshold value, when the distance is less than do , the free space channel model is used (d2 power loss) ; When the distance is more than do , the multi-path fading channel model (d4 power loss) is used.

*5.3. Results Discussion* We have created the network scenario as shown in figure 3 to gather the data

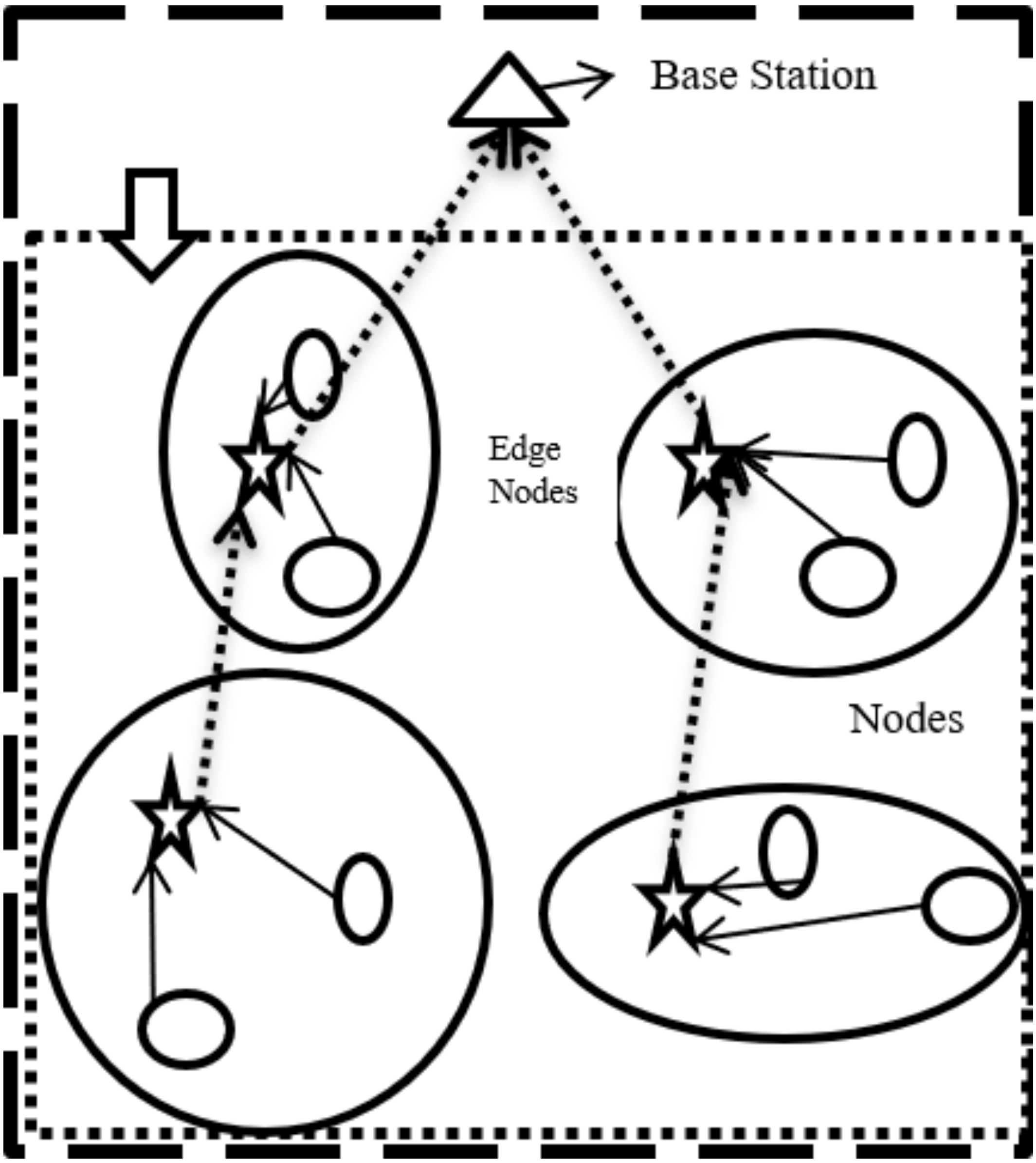
from the sensor nodes at the BS. The data is transmitted through predefined algorithm which has divided the area into four subareas and all of same size. Data is transmitted to the BS with the help of edge node which is placed in each subarea. Edge nodes help in maintenance of security while transmitting data to the BS/aggregator through internet. Figure 4 shows the network lifetime of the sensors with remaining

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 3980; doi:10.3390/s18113980



**Figure 3.** Network Scenario

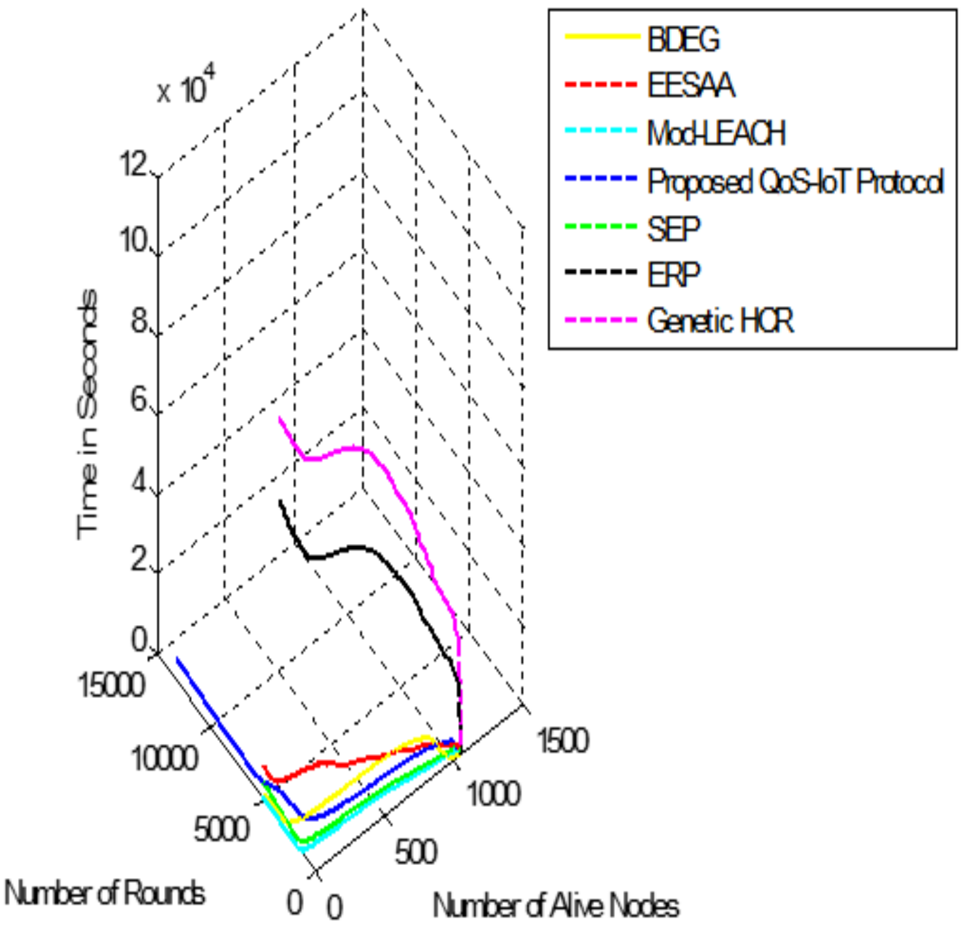


Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 3980; doi:10.3390/s18113980



**Figure 4.** Network Lifetime in reference to the Remaining Energy of Nodes



Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



**Table 1.** Parameters used in QoS-IoT

| Parameter | Value |
|---|---|
| **Network coverage** | (200,200)m |
| **BS location** | (100,200)m |
| **Node Number** | 10055 |
| **Initial energy (Quantity) of Normal nodes** | 0.5 joules |
| **Eelec** | 50nJ/bit |
| **Efs** | 10pJ/bit/m$^2$ |
| **Emp** | 0.0013pJ/bit/m$^4$ |
| **$d_0$** | 87m |
| **Eda** | 5nJ/bit/signal |
| **Data packet size** | 4000bits |

energy after transmission of data. The proposed protocol (QoS-IoT) is optimized over the BDEG (Big Data Protocol) and other traditional protocols. Novel protocol has still 7 nodes alive as comparative to the BDEG (where only 3 nodes are alive)

To make this clearer another comparison of the network lifetime has been shown in figure 5, with dead nodes. Figure 6 and Figure 7, are depicting the standard deviation and variance in network throughput (number of transmitted packets in one time).

Variance and standard deviations are used to show the distribution of data over population. The variance gives the results in square while standard deviation simple squares the variance. Variance of the network throughput of the protocols can be observed in figure 6; the higher variance shows the throughput away from the mean. But still, as it gives the results in squares, so it is not the very useful measure of the transmitted packets. E.g. in figure 6, the variance of QoS-IoT protocols is 2

High standard deviation means values are far away from the mean. In network throughput, we require low value of standard deviation because we want that packets should be transmitted in uniform manner to utilize the bandwidth in efficient manner and for long term working on network. If its value is high, it shows some of the nodes in the network will deplete their energy a long before than other nodes and it can make network obsolete (network partitioning).

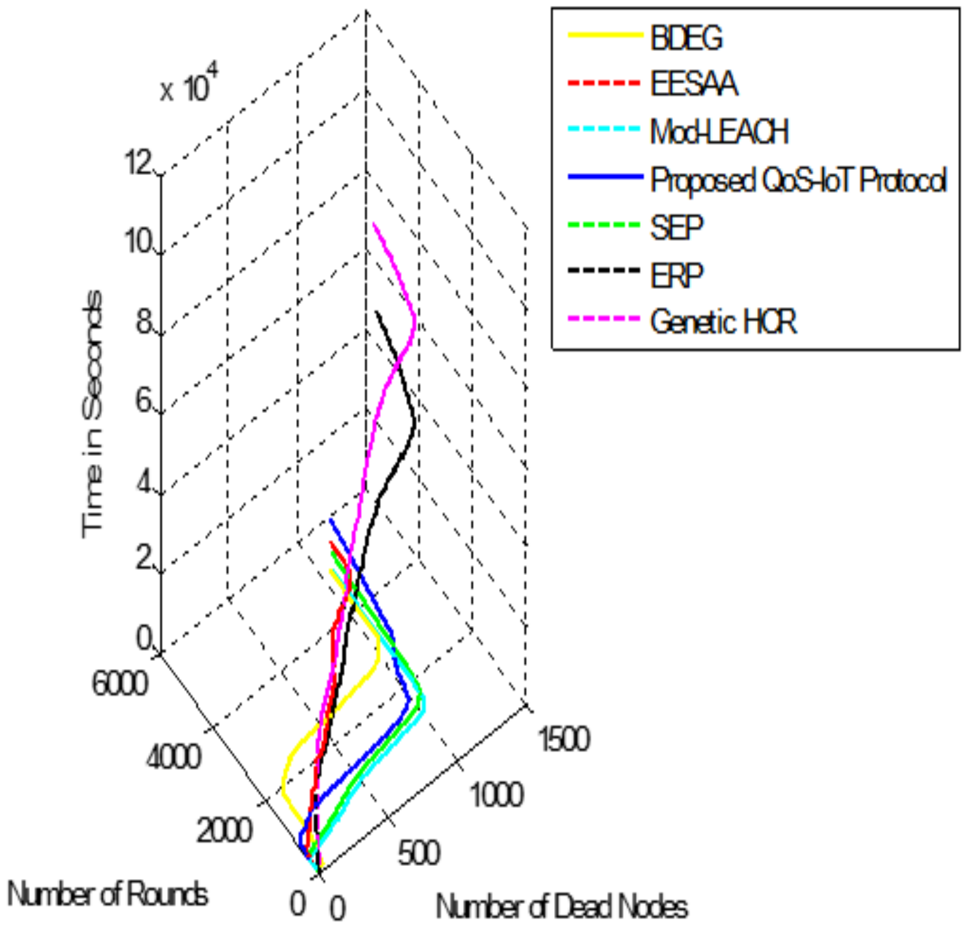
Figure 7 shows the standard deviation of the comparative protocols. According to the network lifetime, proposed protocol is validated over all the protocols but in network throughput, BDEG and EESAA are working better than it. It is due to the fact that, sensors of two subareas of the network are located far away from the BS and if the BS will be in center then it will show the improvement over

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at *Sensors* **2018**, *18*, 3980; doi:10.3390/s18113980



**Figure 5.** Network Lifetime in reference to the Energy Depletion of Nodes

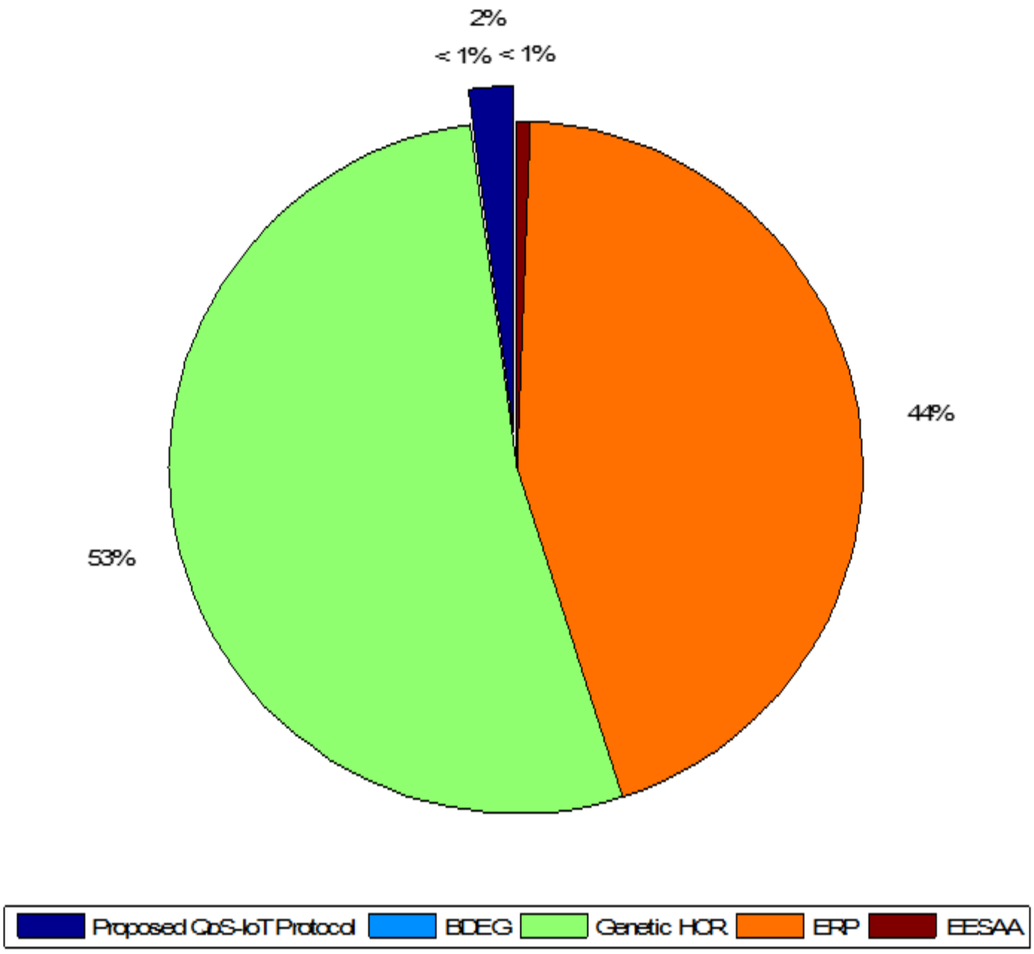


Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 October 2018 doi:10.20944/preprints201810.0469.v1

Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



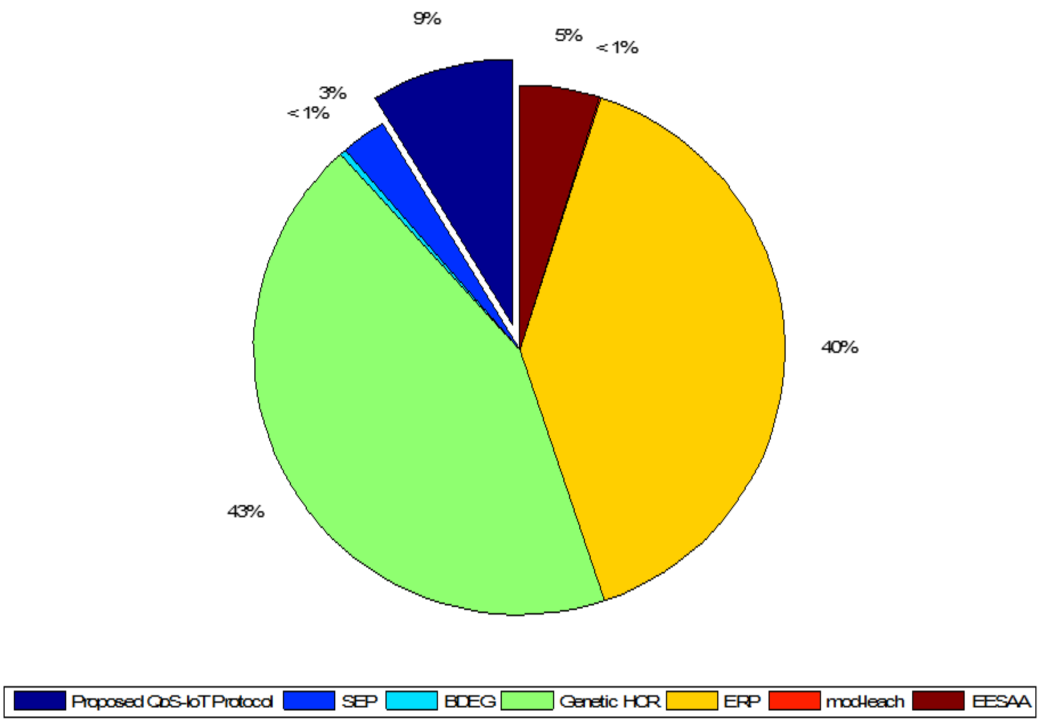
**Figure 6.** Comparison of Variance of the Protocols in Network Throughput



Peer-reviewed version available at Sensors 2018, 18, 3980; doi:10.3390/s18113980



**Figure 7.** Comparison of Standard Deviation of the Network Throughput




both protocols. As compare to the figure 6, figure 7 depicting the true variation of the protocols where QoS-IoT protocolâĂŹs value is 9

## 6. Concluding Observation

Smart city applications have major impact on development of nation. These applications will enhance the decision making power of the citizens. Big Data analytics will facilitate the intelligent and effective decision making policies. In this paper, we proposed a framework for Smart City applications by use of Big Data generated from the IoT system. The proposed system composed of five layer model, where data is aggregated, transmitted, analyzed, filtered and processed. The major role in this framework is played by the WSN. Sensors nodes generate and transmit the data to the BS. However, they are constrained by many features. To achieve the green computing in IoT, it is necessary to conserve energy by using new methods or protocols. That is why along with the proposed framework, a new protocol, QoS-IoT is developed to save maximum energy of the nodes for the long network lifetime of WSN. QoS-IoT is validated in terms of time, throughput and energy. Throughput of the network is enhanced but with greater variation over time. To reduce the variation of the throughput, a cross layer model will be considered in future work.

## References

1. erera, C., Liu, C. H., & Jayawardena, S. (2015). The emerging internet of things marketplace from an industrial perspective: A survey. IEEE Transactions on Emerging Topics in Computing, 3(4), 585-598.
2. . Zeng , S. Guo , Z. Cheng , The web of things: a survey, J. Commun. 6 (6), (2011), pp. 424âĂŞ438 .
3. grawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U., Franklin,M., & Widom, J. (2012). Challenges and Opportunities with Big Data: A white paper prepared for the Computing Community Consortium committee of the Computing Research Association. Available: http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf
4. BM, Armonk, NY, USA. Four Vendor Views on Big Data and Big Data Analytics: IBM [Online]. 2012 Available: http://www-01.ibm.com/software/in/data/bigdata/
5. ing, X., Tian, Y., & Yu, Y. A real-time big data gathering algorithm based on indoor wireless sensor networks for risk analysis of industrial operations. IEEE Transactions on Industrial Informatics, 12(3), 1232-1242, 2016.
6. itchin, R. (2014). The real-time city? Big data and smart urbanism. GeoJournal, 79(1), 1-14.
7. han, Z., Anjum, A., Soomro, K., & Tahir, M. A. (2015). Towards cloud based big data analytics for smart future cities. Journal of Cloud Computing, 4(1), 2.
8. athore, M. M., Ahmad, A., Paul, A., & Rho, S. (2016). Urban planning and building smart cities based on the internet of things using big data analytics. Computer Networks, 101, 63-80.
9. . Jin , J. Gubbi , S. Marusic , M. Palaniswami , An information frame- work for creating a smart city through Internet of things, IEEE Inter- net Things J. 1 (2) (2014) 112âĂŞ121 .
10. . Ahmad , A . Paul , M. Mazhar Rathore , An efficient divide-and- conquer approach for big data analytics in machine-to-machine communication, Neurocomputing 174 (2016) 439âĂŞ453 .
11. .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) 4610âĂŞ4621 .
12. ashaf, A., Javaid, N., Khan, Z. A., & Khan, I. A. (2012, December). TSEP: Threshold-sensitive stable election protocol for WSNs. In Frontiers of Information Technology (FIT), 2012 10th International Conference on (pp. 164-168). IEEE.
13. . Alcaraz, P. Najera, J. Lopez, and R. Roman, âĂIJWireless sensor networks and the internet of things: Do we need a complete integration?âĂİ in 1st International Workshop on the Security of the Internet of Things(SecIoTâĂŹ10), IEEE. Tokyo (Japan): IEEE, December 2010, p. xxxx.
14. . Barnaghi, W. Wang, C. Henson, and K. Taylor, âĂIJSemantics for the Internet of Things,âĂİ International Journal on Semantic Web and Information Systems, vol. 8, no. 1, pp. 1âĂŞ21, 2012.
15. ellavista, P., Cardone, G., Corradi, A., & Foschini, L. (2013). Convergence of MANET and WSN in IoT urban scenarios. IEEE Sensors Journal, 13(10), 3558-3567.
16. in, J., Gubbi, J., Marusic, S., & Palaniswami, M. (2014). An information framework for creating a smart city through internet of things. IEEE Internet of Things Journal, 1(2), 112-121.
17. anella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of things for smart cities. IEEE Internet of Things journal, 1(1), 22-32.

18. ly, H. H., Soliman, A. H., & Mouniri, M. (2015, October). Towards a fully automated monitoring system for Manhole Cover: Smart cities and IOT applications. In Smart Cities Conference (ISC2), 2015 IEEE First International (pp. 1-7). IEEE.
19. bdelwahab, S., Hamdaoui, B., Guizani, M., & Znati, T. (2016). Cloud of things for sensing-as-a-service: Architecture, algorithms, and use case. IEEE Internet of Things Journal, 3(6), 1099-1112.
20. olios, P., Panayiotou, C., Ellinas, G., & Polycarpou, M. (2016). Data-driven event triggering for IoT applications. IEEE Internet of Things Journal, 3(6), 1146-1158.
21. ani, S., Talwar, R., Malhotra, J., Ahmed, S. H., Sarkar, M., & Song, H. (2015). A novel scheme for an energy efficient Internet of Things based on wireless sensor networks. Sensors, 15(11), 28603-28626.
22. ani, S., Ahmed, S. H., Talwar, R., & Malhotra, J. (2017). Can Sensors Collect Big Data? An Energy Efficient Big Data Gathering Algorithm for WSN. IEEE Transactions on Industrial Informatics.
23. han, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M & Gani, A. (2014). Big data: survey, technologies, opportunities, and challenges. The Scientific World Journal, 2014.
24. Lastname, F.; Author, T. The title of the cited article. *Journal Abbreviation* **2008**, *10*, 142-149.
25. Lastname, F.F.; Author, T. The title of the cited contribution. In *The Book Title*; Editor, F., Meditor, A., Eds.; Publishing House: City, Country, 2007; pp. 32-58.