Journal of Advances in Science and Technology
Vol. 21, Issue No. 1, March-2024, ISSN 2230-9659

An Investigation on Handling IoT Big Data with RBSEE Architecture

Rajendra Mahto[1]*, Dr. Nidhi Mishra[2]

[1] Research Scholar, Kalinga University, Raipur, Chhattisgarh, India

Email: rmahto2250@gmail.com

[2] Assistant Professor, Dpt. of Computer Science, Kalinga University, Raipur, Chhattisgarh, India

Abstract- Internet of Things (IoT) devices have resulted in an unprecedented surge in data generation, presenting both opportunities and challenges in data management and analytics. This paper investigates the feasibility and effectiveness of utilizing the RBSEE (Resource-Based Storage, Edge Computing, and Elasticity) architecture for handling IoT Big Data. Firstly, the paper elucidates the fundamental concepts of IoT and Big Data, highlighting the unique characteristics and challenges associated with managing large volumes of heterogeneous data streams generated by IoT devices. Subsequently, it introduces the RBSEE architecture, which integrates resource-based storage, edge computing, and elasticity to address the scalability, latency, and resource constraints inherent in IoT environments. The architecture's ability to distribute data storage and processing tasks across edge devices and cloud infrastructure while dynamically adapting to fluctuating workloads makes it particularly suitable for handling IoT Big Data. In any typical IoT ecosystem, real-time and instantaneous handling of request and dataset is pivotal and a fundamental property. This aims to propose energy efficient distributed system architecture for persuasive, privacy-preserved and real time handling of data generated from IoT devices.

Keywords- IoT, Big data, RBSEE, Handling, Security

----------X----------

INTRODUCTION

This study deals with those big data which are generated from IoT devices and systems. Since the inception of the IoT technology, the range and domain of datasets in many cases has expanded in multiple dimensions (Buyya 2009). Along with such expansions, IoT system involves synchronization, collaboration and extension of thousands of sensors, actuators and transponders. To adapt the data produced from such varied devices into a standard format possesses baffling challenges. Furthermore, with the amount of information generated and transmitted from these devices (which is largely unstructured) it becomes really hard for the system to be instantaneously responsive with minimum possible latencies. Today, IoT technology is slowly but surely penetrating in every aspect of daily life. No doubt, IoT has great potential to provide modern consumer with convenience but at the same time due to its restrictive and constrained nature its potential vulnerabilities increase. Therefore, there is an urgent need to address these vulnerabilities in order to fully exploit the potentials of IoT technology in a secured and privacy-preserved manner (Al-Fuqaha 2013; Ashton 2009; Atzori 2010, Gubbi 2013, Jara 2013, Kopetz 2011, Lin, J 2017, Miorandi 2012, Ray 2016, Sethi 2017, Weber 2010, Xia, F., Yang 2012., Xuan 2017, Zanella 2014). In recent times we have seen an increasing number of a variety of tools and systems for the purpose of effectively handling big data. Among them, Hadoop has emerged as one of the most prominent and favoured frameworks. Although Hadoop outshines all other frameworks available for handling large volumes of data, however, it has severe limitations while handling real-time requests. In any typical IoT ecosystem, real-time and instantaneous handling of request and dataset is pivotal and a fundamental property. This study aims to propose energy efficient distributed system architecture for persuasive, privacy-preserved and real time handling of data generated from IoT devices. This architecture is called by the name “Request-Based, Secured and Energy Efficient (RBSEE)” architecture (Ahad 2018). The current study is divided into five sections. The second section talks about the preliminaries about the concept along with the need of securing IoT systems and devices. The third section gives an introduction about the wireless sensor networks (WSN). The fourth section provides a detailed overview of the proposed architecture.

SECURING IOT DEVICES

The constrained and limited memory, low power consumption and limited processing capabilities are some of the restrictions that are attached to any typical sensor network. If the IoT devices are left unsecured there are chances that the hackers can cause adverse effects which may directly or

indirectly affect the user or consumers. Let us take few examples to understand the need of securing IoT devices and system.

Example 1:

Suppose that we have a person having an IoT enabled pace maker embedded in the body. If a hacker hacks the devices or the system and stops the pace maker then in this situation what will happen to the person? This can cause severe effects and even cause the death of the person!

Example 2:

A Person is having an IoT enabled Smart home automation system. He left for work leaving the kids alone at home. Suppose that some hacker hacks into the system and opens the gate and get hold of the kids. In this situation what can happen? The kids can be kidnapped by the intruders apart from burglary.

Example 3:

Suppose the home is equipped with cameras. You are spending some leisure time with the family members, and suppose that someone hacks into the system and uploads the video on the internet. This can cause adverse after-effects and can compromise the privacy and security of the individuals.

Thus, securing the IoT device is highly crucial in order to protect the privacy of the users and their data and information.

SECURITY ISSUES IN IOT DEVICES

Although the security of IoT devices is of utmost importance, yet there are several hindrances in the path of securing these devices. Some of these are briefly explained below.

Inability to understand the potential risks involved in adopting IoT technology

Most of the manufacturers are not well aware of the potential threats and vulnerabilities associated with adopting IoT technology for their business. Most of the time they rely on third-party solutions to overcome these issues. Primarily there are two kinds of risks involved. These are mentioned below.

(i) Risks from the manufacturer's point of view

For the sake of business and increasing the profits and market shares, most of the manufacturers are blindly developing IoT devices without implementing proper security mechanisms. Their primary aim is to capture the market for selling the product as fast as possible. The security issues generally neglected or take a back seat. These IoT devices sometimes collect highly sensitive and private data without the knowledge of the consumers which must not be exposed to the outside world. The manufacturer of an IoT device which measures the temperature, humidity and pressure of the environment often argue "what is the need for implementing security features in such IoT devices" without understanding that these data in combination with other data from other sensors can provide several insights into the user's routine and tastes which can be used by the promotional companies for user profiling.

(ii) Risks from the consumer's point of view

The consumers don't want to waste their time to configure the security modules of the IoT devices and generally opt for plug and play devices with minimal or no external interventions. It is often observed that even the passwords of these devices are left to default values and never changed by majority of the users. This gives the hackers a cakewalk in hacking the devices as most of the default password are in public domain which can be easily cracked. Most of the times the consumers don't care about what information are being captured by the sensors of the IoT devices, rather they are interested only in the service they have asked for. We have often observed that whatever we talk about on social media, the advertisements of same things start pouring in our inboxes. All this happens because of the information captured by these social networking websites are shared with the promotional companies which perform the profiling of the consumers and on that basis, post the relevant advertisements to them.

WIRELESS SENSOR NETWORKS

Big data constitutes huge datasets which cannot be effectively handled by classical data management systems and tools. The term "big data" has a lot of room for interpretation since different companies use it in different ways. One company might consider data larger than 100 GB to be big data, while another might consider data larger than 500 TB to be big data. The emergence of IoT and its adoption in almost all fields of computing have exponentially increased the generation of data. Although these data are individually very small in size but collectively becomes enormous. The sensors, actuators and transponders constitute the core elements of the IoT ecosystem (Akkaya 2005, Al-Karaki 2004, Broch 1998, Lewis 2004, Matin 2012, Raghavendra 2006, Rani 2015, Tang 2006). The interconnections of these devices are governed by wireless sensor networks (WSN). Figure 1 presents the architecture of wireless sensor network.

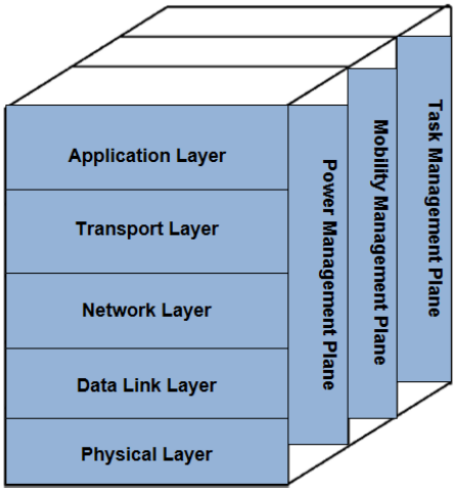

Figure 1: Layered Architecture of WSN

A WSN may be defined as a collection of huge number of miniature sensors. These sensors are self-directed and consume very low power. These sensors can be embedded into network devices or any other equipment or object (Ehyaie 2009, Huang 2014, Pelegri-Sebastia 2017, Swan 2012). They are responsible for sensing, collecting, processing and transferring the data about its surroundings (like temperature, pressure, humidity, blood pressure, heart beat, pulse rate etc) to the base station or operators. A WSN is highly useful in situation where we cannot create a typical wired network like terrains, sea, underground etc. A WSN can be considered as the backbone of the IoT ecosystem. Figure 2 shows an example of a typical WSN. Some of the advantages of WSN are given below:

- Ability to create and manage moveable network.
- Low-cost setup.
- No need for complex wiring.
- Easily extensible and scalable.
- Support variety of network topologies.

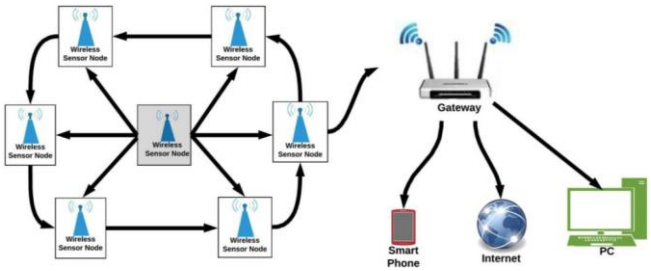

Figure 2: Wireless Sensor Network

RBSEE ARCHITECTURE

Our proposed RBSEE architecture adopts Atrain Distributed System (ADS) as the base of the architecture (Ahad 2018). It also uses "Software Defined Networking (SDN)" and "Two fish cryptographic technique". SDN is incorporated here to provide dynamic network management capabilities while Two fish is used to protect the data stored in the system as well data in transit. For providing energy efficiency, RBSEE introduces the concept of Request-Type Identifying (RTI) unit. Using the RTI unit, the requests are characterized on the source of certain criteria 7 parameters, which are explained here in the later section. These components are incorporated in ADS to construct a unified framework for efficiently handling big data generated from IoT devices. The primary aim of RBSEE is to provide a responsive system with an ability to handle requests and queries in real time. The Pilot Computer (PC) of ADS acts as the control unit while the DCs act as IoT devices in the proposed RBSEE architecture.

COMPONENTS OF RBSEE ARCHITECTURE

RBSEE architecture is aimed at dynamically servicing the requests of the user on the basis of its type rather than adopting a generalized strategy for servicing the requests. Security, reliability, consistency and energy efficient routing of the data in transit constitute the core characteristics of RBSEE. It provides a generalized framework which can be adopted in any domain of IoT like healthcare, transport, home automations, education etc. Security remains the primary concern when we talk about adopting IoT technology in any confidentiality specific domain like healthcare or defence. With RBSEE we aim to address this concern along with faster and energy efficient servicing of requests. Figure 3 shows the architectural framework of the RBSEE system. It can be observed from figure 3, that there are eight components in RBSEE architecture viz User, Request-type Identifying Unit (RTIU), Data Encryption and Standardization unit, System controller, pilot computer, IoT devices, relay nodes and software defined networking unit. The functional description of each of these components is given below

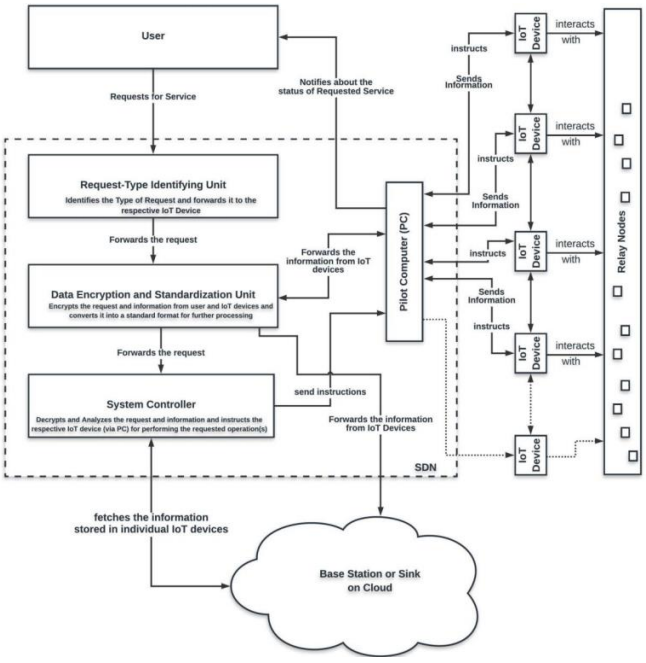

Figure 3: RBSEE Architecture

User

They are the individuals or systems that interact with the RBSEE. In order to get any work done, the users request the RBSEE system thorough the user interface to perform the desired operation.

Request-Type Identifying Unit (RTIU)

This unit is responsible for differentiating among the different types of requests. The various types of request categories are given below:

(i) Requests demanding Instant Response

(ii) Requests demanding Scheduled Response

(iii) Requests demanding Delayed Response

(iv) Requests demanding Continuous response

(v) Event Driven Response (vi) Requests where No Response Required

The details about these requests' types are already presented in above section.

Data Encryption and Standardization (DES) Unit

The DES unit is responsible for encrypting the requests from the users as well as the data captured by the sensors of the IoT devices. This is done to ensure that the data and requests always remain secure whether in transit or at rest in storage devices on the cloud. The data standardization component is responsible for converting the data collected from the IoT devices in a common standard format so that we can perform appropriate and consistent analysis of the data. The complex & diverse nature of the interconnected IoT devices makes data standardization an absolute necessity. IoT gadgets that make up a big ecosystem are likely to be made by numerous companies and have varied models, storage formats, and makes. Consequently, a uniform format must be established to transform the data collected by these different devices.

System Controller

It is the brain of the RBSEE system. It controls every other component by providing instructions to perform as per the need and requirements of the requests from the users.

Pilot Computer (PC)

It is the central node which consists of the metadata of all the connected IoT devices of the system. More specifically it stores information like "device ids" of the IoT individual devices, individual data formats, size of data formats, distance of each device from the PC and other important information like respective functionalities of individual IoT devices.

IoT Devices

They are miniature devices which consist of the sensors embedded in them. These devices are installed at the desired locations by the users, sometimes they are also used as wearable devices which the user can wear on their body like (smart band, smart watch etc). The sensors embedded in these devices captures the data of the surroundings and forward the data to the base station either continuously or periodically depending upon the type of sensor and requests.

Relay Nodes

In order to perform energy efficient routing of data and information from one place to another in RBSEE system, the concept of relay nodes has been incorporated (Ehyaie 2009, Lloyd 2007, Tang 2006). These are tiny sensor nodes which act as a forwarding node and are responsible for forwarding the data captured by the neighbouring sensor nodes to the base station in case the sensor node are at a distant place from the base station.

Software Defined Networking (SDN) Unit

RBSEE system incorporates SDN in order to dynamically control the network with respect to the traffic and congestion in an IoT ecosystem.

WORKING OF RBSEE APPROACH

Every IoT device in RBSEE system needs to register itself with the system. This registration is done by sending the device related information like (device_id, processing power, make, model, data format) to the PC. The PC stores this information as a metadata for each device. The registration is essential in order to identify, connect and communicate with the individual IoT device in future references. The RBSEE system begins by sending user service requests to the RTIU, which then

analyses the data to determine the request's categorization based on its type. Once this analysis is complete, the current request is placed into one of the categories that were previously explained. In addition, before sending the request to the DES unit, the RTIU appends the request id (req_id) & category identification (cat_id). After converting the request to a predetermined format, the DES unit encrypts it. The request is thereafter forwarded to the controller of the IoT system. In order to process the request, the IoT system controller decrypts it. The request is analyzed and its category is determined after decryption. After that, the controller of the system tells the computer to find which IoT devices are suitable to handle the request. The PC then tells the relevant Internet of Things devices how to process the request. In response to commands sent from a personal computer, the Internet of Things device or devices carries out the specified action, either alone or in conjunction with data from other connected devices. While processing a request that calls for data from numerous IoT devices, the RBSEE makes sure that not a single one of those devices is aware of any such integration. In order to facilitate an anonymous data connection amongst numerous participating IoT devices in order to fulfill the user's request, the data integration unit within the pilot computer employs the pseudonymization approach to substitute the identifying fields of specific IoT devices with fictitious identifiers. Distributed network functions (SDN) control the flow of requests, data, & information throughout the RBSEE system. It is the responsibility of SDN to critically monitor and analyse the traffic and congestion in the network and choose the optimal route for transferring these data and requests from one unit to another. The IoT devices that constitute the RBSEE system are wirelessly connected together with each other and with the PC. Every IoT device has a direct connection with the PC. The different IoT devices and the PC are connected together using multi-horse cart topology (Biswas 2015). When the requested operation is successfully completed, the sensor embedded in the IoT device(s) notifies the PC. The PC on the other hand acknowledges the IoT device(s) and prompts the user about the successful completion of the requested operation (this is done only in case it is an explicit service requested by the user, otherwise, no notification is sent to the user, only the requested service gets completed). Figure 4 depicts the flowchart presenting the overview of RBSEE approach (Ahad 2018). The following algorithm shows the working of RBSEE Architecture:

Algorithm : Working of RBSEE Architecture

1. **FOR** every explicit request from the user, **DO**
2. Identify the request-type and categorise it.
3. Forward it to the DES unit.
4. Convert the request into a predefined standard format.
5. Forward the converted request to IoT System controller.
6. Decrypt the request and analyse it.
7. Instruct the PC to identify the IoT device(s) for performing the requested service.
8. Instruct the identified IoT device(s), to perform the requested Service.
9. **END FOR**

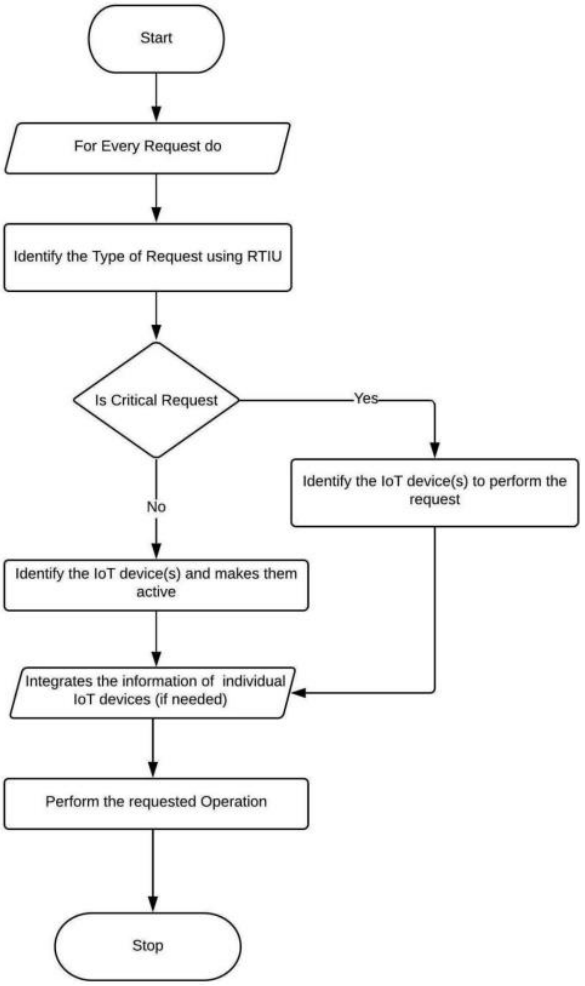

Figure 4: Illustration of the RBSEE Approach

For the sake of understanding let us take one simple example. Suppose that Mr. X has several IoT devices installed at his home automation system. Now Mr. X (user) who was out for work is coming back to home. The proximity sensor and the route map attached to the IoT device in the car senses that the user will reach home in next 15 minutes depending upon the current speed and traffic on the road. It (sensor) immediately communicates with the PC and forwards this information to it. The PC analyses the information and forwards it to the DES for encrypting it and then forwards it to the cloud

storage. After this, it informs the system controller about the information received from the sensors. The system controller fetches the information stored on the cloud storage and after analysing it, instructs the sensor embedded in the Air-Conditioner to switch it on. Furthermore, since the user has a habit of drinking coffee after reaching home. The PC instructs the coffee maker to get the coffee ready. Similarly, the PC instructs the geyser to set the optimal temperature for the user to take a bath. As soon as the user reaches home, the sensors embedded in the Gate, open it and the user gets in. Once the user is in front of the main door, the camera embedded on the door performs the facial recognition of the users and opens the gate.

EFFICIENT USE OF ENERGY IN RBSEE ARCHITECTURES

This study focuses on the energy efficiency of the RBSEE system. As we are aware that every wireless sensor network (WSN) is constituted of several miniature nodes (sensors) which are capable of detecting any event of interest or any change in its surrounding environment and send this data to the base station (or controller node) (Ehyaie 2009, Huang 2017, Pelegri-Sebastia 2017, Swan 2012). These sensors are generally self-sustained and work without any external interventions. In order to perform their respective functionalities, these sensors consist of miniature battery or alternate energy (e.g. solar energy) source to provide appropriate energy to keep them alive and functioning. The kinds of processing & operations carried out, along with the battery's capacity, directly affect how long the sensors last (energy source). Thus, it is essential to develop new methods, technologies, & apparatus that use the least amount of energy possible without sacrificing processing power, extending the life of the sensors in the process. The three major factors that affect the amount of energy consumed by the sensors are mentioned below (Ahad 2018, Akkaya 2005, Al-Karaki 2004).

- The distance between the source and destination: As we are already aware, that a sensor senses the surroundings and forwards this information to the processing node (base station) in the network. The longer the distance between the sensor and the base station, more power is consumed for transferring the data from source to destination.

- Type of activity performed by the sensor: The dissipation of energy (power) is directly proportion to the nature and amount of activity performed by the sensor.

- Traffic and congestion in the network: If the network is densely populated, there are chances of increased network traffic and congestion in routes. This makes the sensor wait until a congestion free route is discovered for the data transfer. This unnecessary waiting of the sensors to get the free route also consumes significant amount of energy. This is because of the reason that in most of the proprietary network devices, the data transfer routes are hardcoded within the device firmware and thus are prefixed. This means that the similar types of data travel through the same path irrespective of the network traffic or congestion.

In order to overwhelmed these limiting factors, the RBSEE approach adopts the concept of relay nodes & multi-hop (Akkaya 2005, Al-Karaki 2004, Broch 1998, Huang 2017, Jiang 2017, Johnson 2001, Lewis 2004, Matin 2012, Raghavendra 2006, Rani 2015, Tang 2006) communication strategy using SDN to transfer the sensor data from one node to another and to dynamically manage the network configuration in real-time. Figure 5 given below shows a direct connection between source and destination.

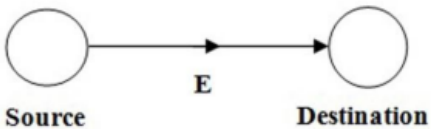

Figure 5: Direct Communication between source & destination

Relay nodes are a unique kind of wireless nodes that are used in wireless sensor networks (WSNs) to transmit data produced by other sensor nodes to the subsequent node in the network, without the need for the relay nodes to sense their own surroundings (Ehyaie 2009, Lloyd 2007, Tang 2006). The concept is analogous to the relay race in which we have multiple racers. Each racer covers some distance and passes the baton to the next racer and so on till the last racer reaches the finish point. The advantage of deploying relay nodes in the network is that it can act as a hop between two or more sensor nodes and can shorten the transit distance between two or more sensor nodes. With the help of relay nodes, the energy (power) required by the sensors to transfer the data can be reduced to a large extent. Thus, the relay nodes help in relaxing the over-burdened sensor nodes within the network and consequently play a vital role in enhancing the lifetime of the sensor nodes. Figure 6 shows the deployment of relay nodes in a typical WSN.

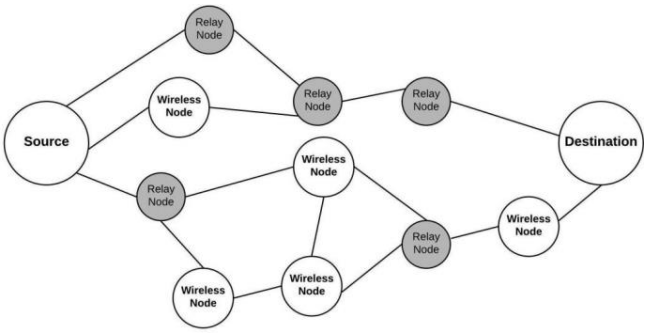

Figure 6: Example of Relay Nodes in WSN

Since data is sent from source to destination across a number of closely spaced hops rather than a direct link, which uses more power, multi-hop communication assures that communications use less power. (Broch 1998, Huang, J. 2017, Jiang 2017, Johnson 2001). The appropriate placement of relay nodes in the wireless network is a different class of problem which is not considered in our research work here for this thesis. Figure 7 shows a multi-hop communication between source and destination.

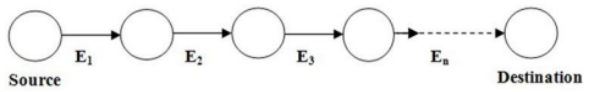

Figure 7: Multi-hop communication between source and destination, where E > E1 + E2 + E3 + ... + En

Apart from deploying relay nodes for saving the energy, RBSEE approach also ensures that the sensors which are not required for performing the work remains in standby mode and consume the minimal energy required for their survival. The remaining sensors in the proposed system stay dormant until the event of interest takes place, with just those necessary for fulfilling crucial requests—which call for immediate responses—becoming active. SDN controls how data is transferred within the RBSEE network. SDN dynamically updates the routing information and keeps the list of the optimal routes for the data transfer within the network. It is also responsible for managing the network congestion and traffic by dynamically updating the network configuration as and when required in order to maintain a fast, reliable and congestion free network.

CONCLUSION

Internet of things (IoT) has given us the concept wherein we can have several types of wireless connections and communication among the constituent objects of the IoT ecosystem. IoT has brought a paradigm shift and totally transformed the way we visualize and observe a network. The current era of IoT-enabled computing has seen an exponential growth with worldwide adoption and acceptance in the last few years. The world is on the verge of extreme technological advancements and if it is not done in a secured and privacy preserved manner, the result will surely be adverse. Therefore, security and privacy must be the primary concern for the development of effective management, deployment and collaboration tools & techniques for creating an effective IoT ecosystem. The RBSEE system takes care of the security and privacy concerns of the data captured and transmitted by the sensors. Furthermore, with the adoption of SDN, the network configuration and management are dynamically controlled by the RBSEE administrator which enables the system to choose appropriate route for transfer of data from one node to another within the network. The RBSEE approach provides energy efficiency by incorporating relay nodes and multi-hop communications along with selectively notifying the sensors to wake up only when the event of interest occurs.

REFERENCES

1. Ahad, M. A. & Biswas R (2018), PPS-ADS: A Framework for Privacy-Preserved and Secured Distributed System Architecture for Handling Big Data, International Journal on Advanced Science, Engineering and Information Technology, Vol. 8 (4), 1333-1342, Doi:10.18517/ijaseit.8.4.5465.

2. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of things: A survey on enabling technologies, protocols, and applications. IEEE Communications Surveys & Tutorials, 17(4), 2347-2376.

3. Ashton, K. (2009). That 'internet of things' thing. RFID journal, 22(7), 97-114.

4. Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. Computer networks, 54(15), 2787-2805

5. Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation computer systems, 25(6), 599-616.

6. Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future generation computer systems, 29(7), 1645-1660.

7. Jara, A. J., Ladid, L., & Gómez-Skarmeta, A. F. (2013). The Internet of Everything through IPv6: An Analysis of Challenges, Solutions and Opportunities. JoWua, 4(3), 97-118.

8. Kopetz, H., 2011. Internet of things. In Real-time systems (pp. 307-323). Springer, Boston, MA.

9. 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.

10. Miorandi, D., Sicari, S., De Pellegrini, F., & Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. Ad hoc networks, 10(7), 1497-1516.

11. Ray, P. P. (2016). A survey on Internet of

Things architectures. Journal of King Saud University-Computer and Information Sciences, 30(3), 291-319

12. Sethi, P., & Sarangi, S. R. (2017). Internet of things: architectures, protocols, and applications. Journal of Electrical and Computer Engineering, Vol. 2017, 1-26

13. Weber, R. H., & Weber, R. (2010). Internet of Things: Legal Perspectives, vol. 49. 1-129. Springer. [187] What is Deep Learning? Three things you need to know, by MathWorks, https://in.mathworks.com/discovery/deep-learning.html \

14. Xia, F., Yang, L. T., Wang, L., & Vinel, A. (2012). Internet of things. International Journal of Communication Systems, 25(9), 1101-1102.

15. Xuan, P., Ligon, W. B., Srimani, P. K., Ge, R., & Luo, F. (2017). Accelerating big data analytics on HPC clusters using two-level storage. Parallel Computing, 61, 18- 34.

16. Zanella, 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.

Corresponding Author

Rajendra Mahto*

Research Scholar, Kalinga University, Raipur, Chhattisgarh, India

Email: rmahto2250@gmail.com