Big Data in Wireless Sensor Network: Issues & Challenges

Ms. Sarita V. Halde[1], Prof. Sucheta T. Khot[2]

[1]P.G.Student, Department of E&TC, B.V.C.O.E.W, Pune, India
[2]Asst. Professor, Department of E&TC, B.V.C.O.E.W, Pune, India

Abstract— The development of Information and Communication Technology (ICT) is contributing to increase the volume of data. Wireless Sensor Network is having no. of applications which consists of sensor nodes. WSN is used for military applications, environment monitoring in smart homes, in industries to detect emergency conditions. Amongst them the important use of WSN for collecting data from sensor nodes. In this continuous collection of sensed data is there for few or all sensor nodes & forwarded to central or base station for further processing. This produces significant portion of Big data. This paper includes in brief about big data in WSN, its issues and challenges in data collection.

Keywords— Big data, Wireless Sensor Network (WSN), data collection, Routing Protocols.

I. INTRODUCTION

In recent years the technical advances has enabled significant amount of data growth in wireless sensor network. In general Data is nothing but collection of values and variables which are related in some sense and are different in some other sense. In WSN the collected data is used to take decisions after analyzing it. Therefore collected data in WSN is considered as rich source for decision making. The distributed sensor network is the production of big data. The network having single sensor may not give significant data, but the data sensed by millions of sensors in distributed wireless sensor network produce a big data[1].

Nowadays WSN is playing important role in number of applications. Also due to development in communication technology, it is required to deal with large amount of data generated while communication. The applications such as supervisory security systems in army, forest monitoring, home automation, vehicle to vehicle communication produces massive information. These kind data is posing challenge to the existing data collection methods which is referred as Big data.

The explosion of big data in WSN is very recent phenomenon and therefore it is becoming challenging to capture and analyze it, and getting some information out of it.

Big data term is used for large datasets having greatest complexity [8]. Usually this data is unable to store, manage and analyze with the conventional data processing methodologies. As the WSN is considered as heterogeneous, the information available with this is also having heterogeneity in it. The extracted information from the wireless network consists of data of continuous monitoring i.e. used for updating status in case of environment monitoring, also the sensed values for all heterogeneously distributed sensors in the network. By virtue of this enormous amount of data is generated at every minute in the network

Big data collection refers the activity that is going through big datasets to acquire requested data from user at the base station. Big datasets are observed in atmospheric science, natural disaster, sensor networks, in telecommunication etc.

II. RELATED WORK

According to D.Taikashi ,*et al.*, in future the expected key contributor for big data in network is the distributed WSN. The data produced due to no. of sensors deployed in a network produces significant portion of big data[1]. Z. Gaffari, *et al.* [3] presented the comparison between different routing protocols. D. Agrawal, *et al.* have put forward challenges and opportunities with big data[6]. But this paper has not explained about big data produced due to wireless sensor network.

The authors have explained about the big data and its analysis in brief. But up till now no one has explained about exact amount of big data. Wireless sensor network is significantly generating big data due to numerous sensors deployed in network [2]. The data volumes are increasing rapidly due to development in communication technology.

III. WIRELESS SENSOR NETWORK

The wireless sensor network is made up with the sensing and communication elements which give ability of monitoring and measuring the specified area to master node of the network which is at base station[9]. There are four key elements of wireless sensor network:

 - Densely distributed or deployed sensors.

- The interconnection of network.
- A central point for clustering data during multi-hop communication.
- Computing resources at base station.

Wireless sensor network can be categorized as follows:

1) Mesh based systems – This system uses multi-hop technique to make connectivity between sensor nodes .In this dynamic routing is used.

2) Star based or point to point systems – Single hop technique is used by this system to make connectivity among the sensor nodes.
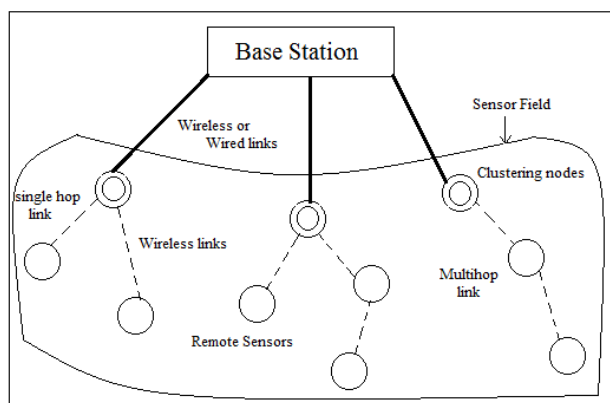

Fig.1: Typical arrangement of sensor network

Sensor nodes are distributed as shown in fig.1. In WSN every sensor node is capable for collecting data and routing this data back to the sink node and to the end user. In case of long distance communication sensor node uses multi-hopping to route data back to the end user. To collect data from each sensor node, WSN is constructed in such a way that, sensor nodes should transmit their data to the end user. In case of monitoring or supervising application, an individual sensor node may not give the correct information hence it is required to collect data from multiple sensors. The unique quality about sensor network is the cooperative communication that takes place between sensor nodes.

While constructing a WSN the points to be considered are; accuracy of data collected, low latency i.e. low network, and optimal use of sources such as power, bandwidth.

A. BIG DATA IN WIRELESS SENSOR NETWORK

Big data term is becoming popular term in field of Information and Communication technology. Since the applications of WSN are increasing enormously, WSN is considered as key contributor of Big data. As shown in figure no. of sensors deployed are responsible for producing the big data in large volume. According to the literature survey it is observed that amount of user data in the network is getting doubled in two years. Big data can be characterized with 3V's [2], which are defined as follows:

- Volume: Amount of data or high potential of information having with corresponding sensor node in the network.
- Velocity: It defines the speed at which data comes at sensor nodes.
- Variety: Heterogeneity of data arrives from no. of sensors in WSN.
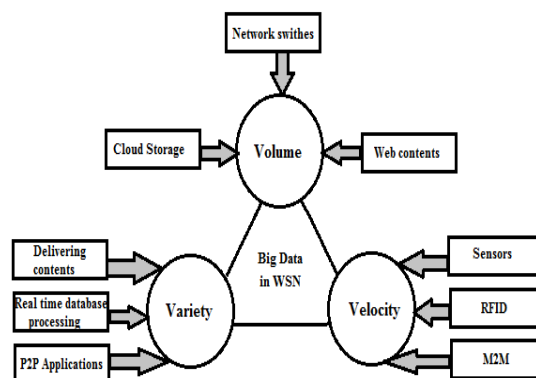
Following fig.2 shows big data characteristics.

Fig.2: Big data in WSN

B. ROUTING PROTOCOLS FOR WSN

Routing protocols are used to handle issues which occurs during data dissemination. There are three different types of routing protocols in WSN. That are Location based protocol, Hierarchical protocols, Data centric protocols [8]. These protocols uses data aggregation concept to avoid redundancies in data.

Data centric protocol uses query based approach to get data from nodes in the network i.e. Sink node from base station sends query to the particular area of WSN for acquiring data from sensor nodes deployed in that specified region. Some e.g. of this protocol are SPIN, Directed Diffusion, Gradient based Routing (GBR) COUGAR, ACQUIRE etc. [3].

The main goal of hierarchical routing is to reduce energy consumption of sensor nodes in the network by using multi hop communication among the clusters. This will help to reduce latency in data delivery. Some Hierarchical protocols are, Energy-adaptive clustering hierarchy (LEACH) Threshold-sensitive energy-efficient sensor network protocol (TEEN), Adaptive Threshold sensitive energy efficient sensor network protocol (APTEEN), Power efficient gathering in sensor information system (PEGASIS).

In Location based routing location information of sensor nodes is utilized to determine energy consumption. Location based routing is used in mobile ad hoc routing. E.g. of location based routing protocol are MECN (Minimum energy communication network) , SMECN (small MECN), GAF (Geographic adaptive fidelity), GEAR (Geographic and energy aware routing).

IV.   ISSSUES OF BIG DATA IN WSN

In WSN each individual node is surrounded with huge amount of data. The main factor regarding a node in network is its coverage area or the area of interest. Since for a particular sensor node, variation in data values in its coverage area is probably less [4]. Hence in that area the redundancy is also less. Following are some common issues observed during analysis.

i. *Computation:*

It is required to do in network processing based on locality of collected data because queries have to be directed as per the requirement by the end user. Data processing includes extracting useful information, data aggregation, fusion of data, analysis of collected data, and computation of hierarchies.

In network processing it is not sufficient to do processing at each node rather this processing should interpret all distributed events and the data associated with that events.

ii. *Data Management:*

This is critical issue to manage rapidly growing volumes and large datasets. This consists of constructing architecture for querying mechanism and for data storage, data warehousing. To manage data for proper communication data needs to be collected at centralized storage to which then queries are addressed. This arranged data at central level avoids redundant transmissions. Therefore it should be done to support distributed data querying. Multi-tiered data storage and retrieval should be done at the storages.

iii. *Sensor network organization and routing:*

Generally in case of wireless sensor network self-configuration takes place. Design of sensor network includes the hardware constraints, power constraints, operating environment, network architecture, routing topology etc. While designing a sensor network expected accuracy of information is considered.

V.   CHALLENGES IN COLLECTION OF BIG DATA IN WSN

Analysis of big data in WSN includes capturing the large data sets and processing them. During this analysis there are some challenges faced by user. Following are the challenges observed during analyzing big data in WSN:

i. *Heterogeneity:*

In densely distributed wireless sensor network, data coming from no. of sensor nodes consists of both structured as well as non-structured data. The existing database systems were designed to address only structured data which is in small volume. Therefore the heterogeneity with WSN is becoming challenging for storage and analyzing data. The difficulties are faced due to data arrived in large volumes and of mixed form i.e. of different patterns or types.

ii. *Energy consumption:*

The requirement of energy for sensor nodes in a wireless sensor network varies with the routing protocol that the sensor node used for communication. Energy is the main constrain of the network since limited lifetime of the network as the batteries are employed with each individual sensor node. During communication of sensor nodes to the base station, energy is consumed during wireless transmission.

Though data volume associated with sensor node is not large enough, lot of energy is consumed to route surrounded sensor nodes data. Due to this the densely distributed WSN has short lifetime.

Energy efficient data collection can be achieved by adopting some techniques. One way is to compress data i.e. by shrinking the data volumes energy utilization can be reduced. In addition to this other way is by reducing redundant transmissions in the network, and hence required energy for wireless transmission is also reduced.

VI.   CONCLUSION

It is observed that Wireless sensor network is the key contributor for generation of Big data. The volume of data in wireless sensor network is increasing exponentially. Big data in WSN, the issues and challenges occurred during data collection in sensor network are included in brief in this paper. Also the routing protocols are elaborated which helps to deal with data dissemination issues of large scale data in the Wireless Sensor Network. Routing protocols different categories are: Location based Hierarchical and data centric protocols. These protocols uses concept of data aggregation i.e. combining data at the time of data arrival, this mechanism reduces redundancies in data collected.

ACKNOWLEDGEMENT

REFERENCES

[1] D. Takaishi, H. Nishiyama, N. Kato and R. Miur, "Toward Energy Efficient Big Data Gathering in Densely Distributed Sensor Networks", *IEEE Trans. on Emerging topics in Computing* Vol.2, no.3, pp. 388-397, Sept. 2014.

[2] R. Bergelt, M. Vodel and W. Hardt, "Energy efficient handling of big data in embedded, wireless sensor networks" *IEEE Conference*, Sept. 2014.

[3] Z. Ghaffari, T. Jafari and H. E. Shahraki, "Comparison and Analysis Data-Centric Routing protocols in wireless sensor networks", *International*

Conference on Communication Systems and Network Technologies, pp. 351-355, 2013.

[4] Y. Demchenko, Z. Zhiming, P. Grosso, A. Wibisono, and C. de Laat, " Addressing big data challenges for scientific data infrastructure", *in Proceedings of the IEEE 4th International Conference on Cloud Computing Technology and Science, IEEE Computer Society*, pp.614-617, 2012.

[5] Q. Zia, "A Survey of Data-Centric Protocols for Wireless Sensor Network", *Journal of Computer Science Systems Biology*, Vol. 8, pp. 127-131, March 2015.

[6] D. Agrawal et al., "Challenges & Opportunities With Big Data", http: //cra.org /ccc /docs /init / Bigdatawhitepaper.pdf

[7] Fu Xiao, Chongsheng Zhang, and Zhijie Han, "Big Data in Ubiquitous Wireless Sensor Networks", *International Journal of Distributed Sensor Networks*, Vol. 2014.

[8] Jaseena K. U. and Julie M. David *"*ISSUES, CHALLENGES, AND SOLUTIONS: BIG DATA MINING", *International journal of Computer Science & Information Technology (CS & IT),* Vol. 4, pp. 131–140, 2014.

[9] K. Sohraby,D. Minoli and T. Znati (2007) , "WIRELESS SENSOR NETWORKS- Technology, Protocols, and Applications", John Wiley and sons publication.