ISSN:2231-0711

Divyadharsini S et al | IJCSET(www.ijcset.net) | 2017 | Volume 7 Issue 5, 41-43

An Enhanced Lockstep Broadcasting in Big Data

Divyadharsini S[#1],Krishneswari k[*2]

[#]*PG scholar & *Head of the Department & Tamilnadu College of Engineering*
Coimabatore ,Tamilnadu ,India

***Abstract* - Transferring of bigdata in wsn is complex due to its size and energy consumption available at node. Also to find the near by transferring node in terms of neighbour node prediction. In our proposed system we find and estimate the neighbouring node by broadcasting signal to near node . it consumes low energy as compared to other techniques to compute neighbour. After neighbour computation we randomly select cluster heads and form the cluster based on the distance and energy metrics. After the clustering of environment we choose the data to transfer in the WSN . This network now split the bigdata into chunks of number of cluster avail. Then transfer the data..Due to limited resources of the multi-sensor system, it is a challenging task to reduce the energy consumption to survive a network for a longer period. Keeping in view the challenges above, this paper presents a novel technique of using a hybrid algorithm for clustering and cluster member selection in the wireless multi-sensor system. After the selection of cluster head and member nodes, the data fusion technique is proposed that is used for partitioning and processing the data. The proposed scheme efficiently reduces the blind broadcast messages but also decreases the signal overhead due to cluster formation. Afterward, the routing technique is provided based one the layered architecture.**

***Keywords* – WSN,WMSS**

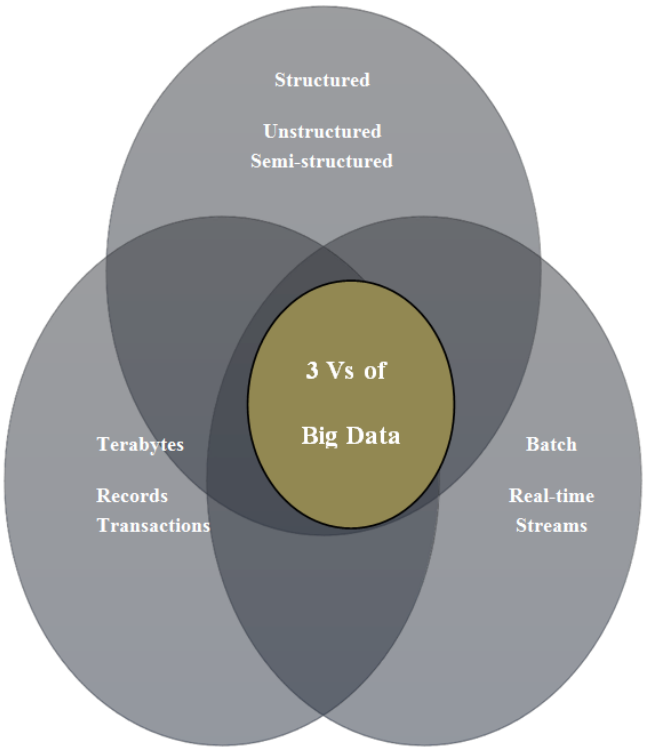
I INTRODUCTION

Big Data is a phrase also named as “veracity” is applied to data sets whose size is beyond the power of commonly used software tools to capture, control and process. The sheer size of the data, combined with complexity of search and exchange to create value from it, has led to a new class of technical knowledge and tools to adopt it. The term Big Data tends to be useful in different ways, often attributing to both the type of data being managed as well as the technology used to mine and process it. These technologies come into existence from companies such as Google, Amazon, Face book and Linked-In, where they were developed for each company’s own use in order to analyze the impressive amounts of social media data they were dealing with. Due to the nature of these companies, the highlight was on low cost scale-out specialty hardware and open source software.

The world of Big Data is progressively being defined by the 3 Vs. But now it has given defined by 4Vs, these ‘Vs’ become a equitable test as to whether a Big Data approach is the right one to maintain for a new area of analysis.

The Vs are:

Volume
Velocity
Variety



II RELATED WORK

- ***Lock Step Broadcast Tree***

There are a lot of utilization domains that widely apply broadcasting operations, such as mathematical data distributions, database transaction logs backups, the latest security patches, multimedia streaming utilizations, and data replica or virtual appliance deployment among distributed materials centers. Since the size of data becomes so enormous, the crush of broadcasting operation also becomes increasingly significant. Consider the big data broadcasting problem in a heterogeneous network where nodes may have multiple uploading capacities. The big data broadcasting problem is about how the nodes may obtain a given big data concurrently in a minimum amount of total transmission time. Assume that there are n nodes in a heterogeneous network system. And the node n is the broadcasting source that has the data item divided into m chunks of equal size, it advertise the data item to all the other nodes. The upload capacities of those nodes measured in kilobyte per second (KBps). In addition, Assume that the downloading capacity of each node is larger than or same degree to its uploading capacity. In Specifically, Focus on investigating the following questions: What is the relation between a single overlay tree with a fixed uplink rate and

the broadcast operation itself, and how to construct a single spread tree that minimizes the maximum completion time in heterogeneous networks? Introduce the novel LockStep Broadcast Tree (LSBT) to model the Big Data broadcast problem. LSBT is a broadcast tree where data chunks can be sent in a pipelined fashion with a good throughput.

Specifically, we focus on inspecting the following questions: What is the relation between a single spread tree with a fixed uplink rate and the broadcast operation itself, and how to construct a single spread tree that minimizes the maximum completion time in heterogeneous networks? Introduce the novel LockStep Broadcast Tree (LSBT) to model the Big Data broadcast problem. LSBT is a broadcast tree where data hunk can be sent in a pipelined fashion with a good throughput. The main idea is to define a basic unit of upload bandwidth, r , such that the upload link of each node is divide into separate connections each being allocated with the bandwidth r in broadcasting. In so doing, the number of upload connections is proportional to the capacity of a node. Furthermore, we also divide the broadcast data into m hunks.

These data chunks are then broadcast down the tree by the nodes in a pipeline mannerism. It is based on the LSBT model, the maximum number of rounds required to complete the broadcast of entire data chunks is O ðm þlog n Þ steps, where n is the number of nodes. In a homogeneous network encompassment in which each node has the same uploading capacity c, we show that the optimal uplink rate r of LSBT is either c=2 or c=3. For heterogeneous networks, we present an Oðn log2 nÞ algorithm to select an optimal uplink rate r and to compound an optimal LSBT. Numerical results show that the maximum completion time of our LSBT approximates to the optimum of the big data broadcast problem.

B.Wireless Sensor Networks

Normally, the wireless sensor device is a compact and tiny device that is operated by battery, which is intermittently difficult to charge in real-time scenarios. These sensors have deployed a region of interest, where they can sense, collect, and process information. In the deployed region, these devices are operated exclusive of in remote or hostile areas where human intervention is not possible. Hence, the security and survivability of a network are of an utmost important. In such region, these devices forming a collection called cluster where sensed data is transmitted to the base station (BS) my means of multi-hop intercommunication. Wireless ad-hoc and sensor network has resource constraints, such as a limited battery (energy), limited computational capabilities, less storage capacity, and lower communication bandwidth. Despite its limitations, it has various practical applications including battlefield (surveillance), health control, environmental and structural control In various applications where humans are not accessible to these devices, it is difficult to recharge or replace the batteries. Hence, limited energy in wireless ad-hoc and sensor network is a critical constraint in surviving network for a longer period. Network lifetime is one of the important features for evaluating the completion of wireless ad-hoc and sensor network.

III PROPOSED SYSTEM

Our proposed system, the datasets mentioned above are implemented using Java iterations and the Hadoop with proposed algorithm. The proposed algorithm using enhanced MapReduce in the light of the proposed algorithm is more efficient that the simple Java iteration implementation 1MB of data using proposed algorithm as well as Java iterations. From the figure, it is clearly seen that the proposed algorithm require approximate half of the second to process 1 MB of data. Apparently, Java iterations require more time to process the same amount of data. Also, ASA-APS requires more time for both cases since the size of the data is too much. Moreover, products of the ASA-WSM are processed quite efficiently since the size of the data is very less. It is concluded from the figure that size of the data plays an important role in any system. Likewise, if the size increases, the performance of Java iteration drastically reduces, whereas, the performance of the proposed system.

Advantages

- Throughput of the proposed system is increased by increasing of datasets
- All the Wireless sensor nodes gets their equal status of node
- They can easily switch over when they cant able to receive the data
- Data splitting is very easy one
- Here Hadoop is used to implementation thus the more datasets can be processed for big data processing.

IV ARCHITECTURE

In the clustering technique, a set of devices is grouped together in a topographical region. After grouping, a cluster head is selected based on some certain algorithms, in which the selected node is called a cluster head, whereas, rest of the nodes are called member nodes. Cluster head acquires data from its member nodes and aggregates it. Then it forwards the data to the bordering cluster heads at the base station via direct hop or multi-hop. Routing data in clusters are divided into two broad categories, i.e., intra-cluster (within the same cluster) and inter-cluster (within clusters) data transmission. Such outline of the cluster scale down a significant amount of energy in the network. Wireless ad-hoc and sensor network is composed of hundreds or even thousands of nodes communicating with each other. Such densely deployed nature of the network consumes more energy in alternate data with the unstable additive load and excruciating faults.

Different algorithms for selecting nodes as a cluster head and member nodes, the inter-communication and cluster head play a vital role in facilitating network in surviving for a longer period of time. It is known as flat architecture based network. In flat architecture based network, there is a uniformity in all nodes, i.e., structure and configuration of a node are homogeneous. Thus, they lack conservancy techniques that may be

supported by themselves. Apparently, in cluster architecture based network, the high energy node that is appropriately divisional cluster head acts as a gateway, which plays an important role in solving different issues . Cluster architecture based network is considered to be energy efficient network by route discovery, data aggregation, fault tolerance, and end-to-end nature . Furthermore, cluster architecture based network shows substantial advantages over flat architecture based network.

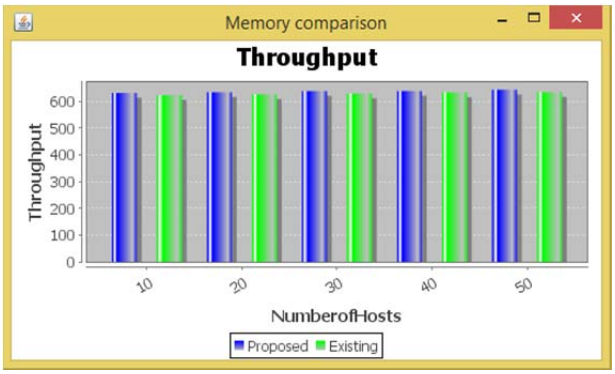


Fig 1. Throughput

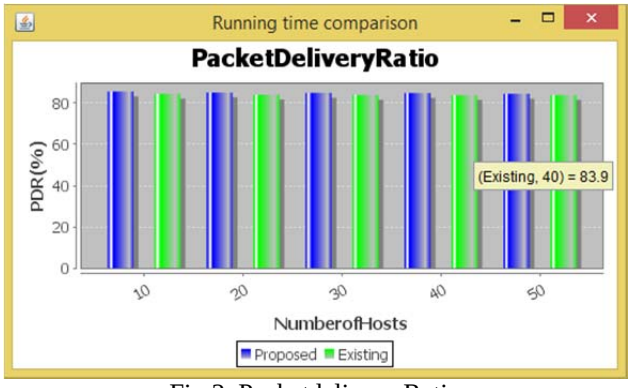


Fig 2**.** Packetdelivery Ratio

V CONCLUSION AND FUTURE ENHANCEMENT

Usually, the nodes involve in the network requires high energy power throughout the communication period. And due to the inherent properties of the network, it is really difficult to survive the network in such environment. Therefore, keeping in view the above limitations, we proposed a novel architecture for wireless ad-hoc and sensor network. The proposed architecture is based on the layered deployment of clustering technique, which is followed by the routing scheme. Initially, nodes recognize themselves to be close enough to establish direct hop communication, which is referred to layer-1. Secondly, nodes that establish multi-hop communication toward the station and are having a high density of nodes are referred to as second layer cluster. Moreover, after formation of the cluster, there is a novel algorithm to assign a member to each node, which is based on the distance as well as RSSI of the node. After selection of cluster head and assigning nodes to it, data fusion technique is employed that partitioned and process the data efficiently. Also, we have also calculated blocking probability for the cluster, which calculates the number of nodes to be attached to each cluster head. And finally, simulation results show that the proposed scheme conserve energy in all the circumstances as compared with three competing algorithms.

Scope for Future Enhancements

The System provides sophisticated environment to succeed the pattern and it involves the basic requirement. It allows promoting the software as per the user adaptability and new requirements. So the maintenance based on the changing environment and requirements can be incorporated easily. Any changes that are likely to cause failures are prevented with security and preventive measures could be taken. In future better algorithms or any technology are implement to enhance this project.

REFERENCES

[1] Deshpande, Vaibhav V., and Arvind R. Bhagat Patil. Energy efficient clustering in wireless sensor network using a cluster of cluster heads. In2013 Tenth International Conference on Wireless and Optical Communications Networks (WOCN), pp. 1-5. IEEE, 2013.

[2] K. K. Chintalapudi and L. Venkatraman, On the design of mac protocols for low-latency hard real-time discrete control applications over 802.15.4 hardware, in IPSN '08, pp. 356–367.

[3] I. Talzi, A. Hasler, S. Gruber, and C. Tschudin, Permasense: investigating permafrost with a wsn in the swiss alps, in EmNets '07, Cork, Ireland, pp. 8–12.

[4] S. Upadhyayula and S. Gupta, Spanning tree based algorithms for low latency and energy efficient data aggregation enhanced convergecast (dac) in wireless sensor networks, Ad Hoc Networks, vol. 5, no. 5, pp. 626–648, 2007.

[5] T. Moscibroda, The worst-case capacity of wireless sensor networks, in IPSN '07, Cambridge, MA, USA, pp. 1–10.

[6] Imrich Chlamtac, Shay Kutten, Tree-based broadcasting in multi-hop radio networks, IEEE Transactions on Computers C-36 (10) (1987).

[7] Imrich Chlamtac, Orly Weinstein, The wave expansion approach to broadcasting in multi-hop radio networks, IEEE Transactions on Communications 39 (3) (1991).

[8] Wendi Rabiner Heinzelman, Anantha Chandrakasan, Hari Balakrishnan, Energy-efficient communication protocol for wireless micro sensor networks, in: Proceedings of the Hawaii International Conference on System Science, January 2000.

[9] Rex Min, Anantha Chandrakasan, Energy-efficient communication for ad-hoc wireless sensor networks, in: Conference Record of the Thirty-Fifth Asilomar Conference on Signals, Systems and Computers, 2001, vol. 1, 4–7 November 2001.

[10] Sarma Upadhyayula, Valliappan Annamalai, Sandeep Gupta, A low-latency and energy-efficient algorithm for convergecast in wireless sensor networks, in: IEEE Global Communications Conference, 2003.