Efficient Energy Conservation Technique Using Sink Node Mobility for Big Data Gathering in Densely Distributed WSN

Karthik K J, Anitha K
PG Scholar, Rajarajeswari College of Engineering, Bengaluru -74
karthikkj1992@gmail.com, anithakrishna14@gmail.com

ABSTRACT- Big data has considered being promising application in the field of information and communication technology (ICT). WSNs are used to gather big data in communication technology. In most of the case the data from single sensor is not enough informative to user. To gather significant data, it is necessary to gather information from distributed sensors. Because of limited power in WSNs big data gathering is challenging task. To increase network lifetime in WSNs the paths for data transfer are chosen in such a way that the total energy consumed along the path is reduced. To support improved data aggregation and high scalability, sensor nodes are often grouped into non-overlapping, disjoint subsets called clusters. Clusters create hierarchical WSNs which incorporate efficient utilization of limited resources of sensor nodes and hence extends lifetime of network. In this paper we proposed an efficient way for data gathering using clustering based technique. In the proposed system we used k-medoids clustering technique for the optimal number of clusters and created sink node trajectory, used as efficient data collector. And also using novel approach for data encryption and decryption for providing security.

Keywords - Big data, Data aggregation, clusters, k-medoids, sink node trajectory.

I. INTRODUCTION

Wireless sensor networks (WSNs) have come up with an effective solution for many of the applications like military, healthcare, home, etc. recent advances in wireless communications and electronics have enabled the development of inexpensive, low power and multi-utility sensor system. A WSN consists of several number of sensor nodes that are distributed densely within a network area. The position of sensor nodes are need not be predetermined or engineered. This allows random distribution in inaccessible terrains or disaster relief operations. On the other hand, this also means that WSN network protocols and algorithms possess self-organizing capabilities. Another unique feature of sensor networks is the cooperation of sensor nodes. Sensor nodes are fitted with an onboard processor. Instead of sending the unprocessed data to the nodes are responsible for further action, they use their on-going abilities to carry out simple computations and move only the required and partially processed data. The described features ensure a broad range of applications for WSNs. Some of the applications are health, military and home. In military, for example, the rapid deployment, self-organization, and fault tolerance of sensor networks make them very promising sensing techniques for military operations, control, interactions, performing, intelligence, surveillance and targeting systems. Implementation of above mention and other WSNs applications require wireless ad hoc networking techniques. Protocols used in wireless ad hoc networks were not well suited the unique features and application requirements of sensor networks. The major limitations of wireless sensor networks their limited power storage capacities and prone to failures. Hence energy efficiency is the major issue in sensor network.

As information technology grows rapidly, volume of the data also increasing simultaneously. Accessing the retrieving big data is crucial for the user in many applications. Big data is a buzzword, or catch-phrase, used to describe a massive amount of both structured and unstructured data that is so huge that it's tough to process using database and software techniques. In most of these scenarios the data is too large or it moves rapidly or it exceeds present processing capacity. Big data has the potential to help companies improve operations and make faster, more intelligent decisions. Collecting large amount data from sensor nodes is the major concern in the field of ICT. Individual sensor nodes may not provide accurate information. Therefore collecting data from multiple sensor nodes is very essential.

In order to gather these data, the WSNs are constructed in such a way the sensors relay their data to the ―sink‖. However, in case of widely and densely distributed WSNs there are two problems in gathering the data sensed by billions of sensors. First, the network is divided to some sub-networks groups because of the limited wireless communication range. Second, the wireless transmission consumes the energy of sensors. Even though the size of data generated by an individual WSN is not significant, each sensor requires a lot of energy to process the data formed by surrounding sensors. Especially in dense WSNs, the life time of sensors will be very less because each sensor node compute a lot of data generated by enormous number of surrounding sensors. In order to overcome these problems, we need an energy-efficient method to gather huge volume of data from a large number of sensors in the densely distributed WSNs.

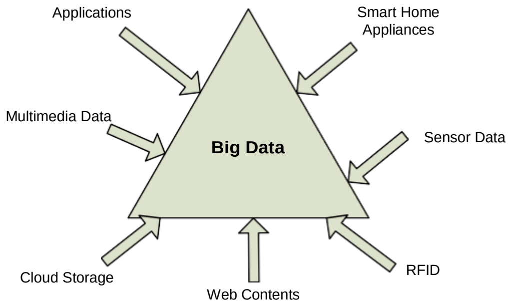

Fig 1. An Overview of Big Data Gathering

To achieve energy-efficient data gathered in densely distributed WSNs, there have been many existing methods. Clustering is one of the most widely used techniques to make WSNs energy efficient. Low energy adaptive clustering hierarchical (LEACH) clustering algorithm is introduced to maximize the life time of wireless sensor network. In [] used expectation maximization clustering technique.

In this paper we proposed an improved method to gathering large data in densely distributed sensor network using k-medoids clustering algorithm. K-medoids algorithm is used as clustering of WSNs. K-medoids is not only effective in clustering also robust in selecting cluster heads in large WSNs. Later we introduced an effective sink node routing for collecting data from each cluster heads..

II. RELATED WORK

[5]The promise of data-driven decision-making is now being recognized broadly, and there is growing enthusiasm for the notion of ―Big Data‖. Heterogeneity, scale, timeliness, complexity, and privacy problems with Big Data impede progress at all phases of the pipeline that can create value from data. The problems start right away during data acquisition, when the data tsunami requires us to make decisions, currently in an ad hoc manner, about what data to keep and what to discard, and how to store what we keep reliably with the right metadata. Much data today is not natively in structured format; for example, tweets and blogs are weakly structured pieces of text, while images and video are structured for storage and display, but not for semantic content and search: transforming such content into a structured format for later analysis is a major challenge. The value of data explodes when it can be linked with other data, thus data integration is a major creator of value. Since most data is directly generated in digital format today, we have the opportunity and the challenge both to influence the creation to facilitate later linkage and to automatically link previously created data. Data analysis, organization, retrieval, and modeling are other foundational challenges. Data analysis is a clear bottleneck in many applications, both due to lack of scalability of the underlying algorithms and due to the complexity of the data that needs to be analyzed.

[4]In this paper, we formulate a flow control optimization problem for wireless sensor networks with lifetime constraint and link interference in an asynchronous setting. Our formulations based on the network utility maximization frame work, in which a general utility function is used to characterize the network performance such as throughput. To solve the problem, we propose a fully asynchronous distributed algorithm based ondual decomposition, and theoretically prove its convergence. The proposed algorithm can achieve the maximum utility. Extensive simulations are conducted to demonstrate the efficiency of ouralgorithm and validate the analytical results.

[2]This paper presents and analyzes three-tier architecture for collecting sensor data in sparse sensor networks. Our approach exploits the presence of mobile entities (called MULEs) present in the environment. When in close range, MULEs pick up data from the sensors, buffer it, and deliver it to wired access points. This can lead to substantial power savings at the sensors as they only have to transmit over a short-range. This paper focuses on a simple analytical model for understanding performance as system parameters are scaled. Our model assumes a two-dimensional random walk for mobility and incorporates key system variables such as number of MULEs, sensors and access points. The performance metrics observed are the data success rate (the fraction of generated data that reaches the access points),latency and the required buffer capacities on the sensors and the MULEs. The modeling and simulation results can be used for further analysis and provide certain guidelines for deployment of such systems.

[3] Clustering is a standard approach for achieving efficient and scalable performance in wireless sensor networks. Traditionally, clustering algorithms aim at generating a number of disjoint clusters that satisfy some criteria. In this paper, we formulate a novel clustering problem that aims at generating overlapping multi-hop clusters. Overlapping clusters are useful in many sensor network applications, including inter-cluster routing, node localization, and time synchronization protocols. We also propose a randomized, distributed multi-hop clustering algorithm (KOCA) for solving the overlapping clustering problem. KOCA aims at generating connected overlapping clusters that cover the entire sensor network with a specific average overlapping degree. Through analysis and simulation experiments we show how to select the different values of the parameters to achieve the clustering process objectives. Moreover, the results show that KOCA produces approximately equal-sized clusters, which allows distributing the load evenly over different clusters. In addition, KOCA is scalable; the clustering formation terminates in a constant time regardless of the network size.

[8]A WSN is a specialized wireless network made up of a large number of sensors and at least one base station. The foremost difference between the WSN and the traditional

wireless networks is that sensors are extremely sensitive to energy consumption. Energy saving is the crucial issue in designing the wireless sensor networks. Since the radio transmission and reception consumes a lot of energy, one of the important issues in wireless sensor network is the inherent limited battery power within network sensor nodes. In order to maximize the lifetime of sensor nodes, it is preferable to distribute the energy dissipated throughout the wireless sensor network. The data gathering schemes should be power efficient. In our proposed work we are changing the idea related to the data gathering and transmission protocol Chiron. The main goal of our research is reduce of energy consumption and improve the lifetime of network as chain leader belonging to the certain covering angle will only transmits the gathered data to the another chain leader of the same covering angle and then we send the data of the another covering angle in sequential manner . So the data is transferred to some angle based chain leader rather than to the nearest chain leader. By this method of data gathering we found that energy consumption is reduced and lifetime is improved significantly.

III. PROBLEM STATEMENT

There are three Problems in gathering the data sensed by millions of sensors.

1. Network is divided to some sub-networks because of limited wireless communication range.
2. The wireless transmission consumes the energy of the sensors during data gathering.
3. There should be only one sink node for each sub-network, which reduces delay in data-gathering.

.

IV. METHODOLOGY

1. Network initialization.
2. Cluster formation.
3. Calculating cluster centroids.
4. Identifying sink node and evaluating trajectory.
5. Efficient data gathering using sink node.
6. Result analysis.

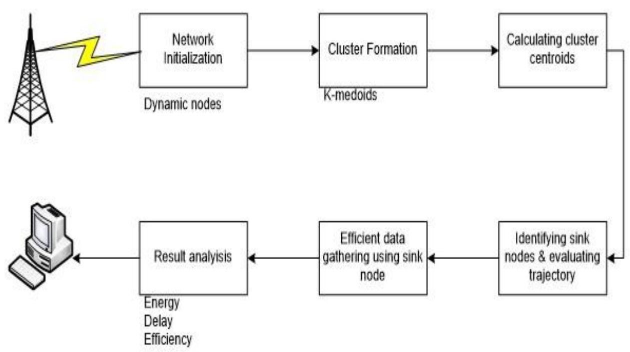

FIG 2: PROPOSED ARCHITECTURE.

V. ALGORITHM

Input: the network $G(V, E)$; k, the number of clusters to be formed; λ, the fading factor of physical channel

Conditions: a set of clusters involving all nodes
1: Pick k nodes randomly as initial cluster heads, and use M to denote the set of those k cluster heads.
2: Establish an empty set C_i for head h_i of M, and initialize
3: Establish a set
4: while any node u of V/M do
5: u chooses h_j of M as its head such that
Where d_{u,h_j} is the Euclidean distance from u to h_j.
6: $C_i \leftarrow C_i \cup u$
7: end while
8: while any C_i do
9: Select a node u from C_i as new head such that

$$\arg_u \min \left\{ \sum_{v=1, v \neq u}^{|C_i|} d_{u,v}^{\lambda} | \forall u \in C_i \right\}$$

10: $N \leftarrow N \cup u$
11: end while
12: if $N \neq M$ then
13: $M \leftarrow N$
14: $N \leftarrow \emptyset$
15: go to step 4
16: end if

VI. PERFORMANCE EVALUATION

The following graphs are compared to the existing method which are based on Expectation-Maximization technique (EM-METHOD) to the K-medoids technique. The graphs are showed better performance in energy, delay and data delivery efficiency.

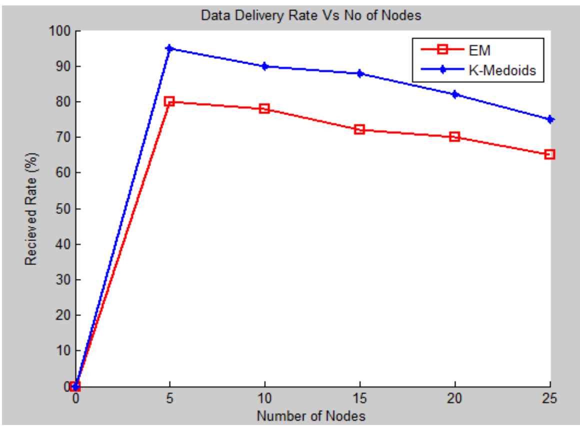

Fig 3: Data Delivery Performance

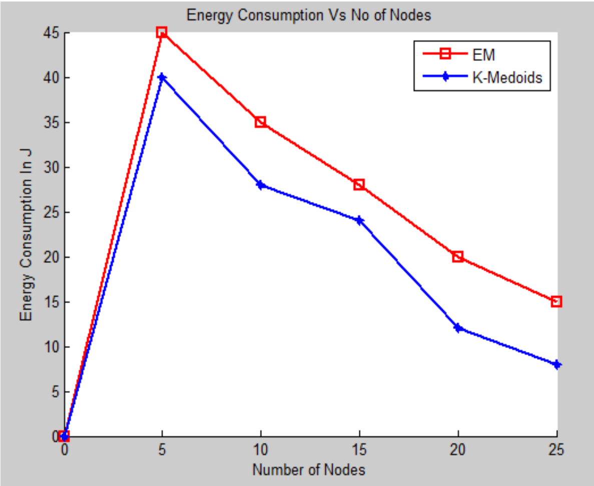

Fig 4: Energy Level Performance

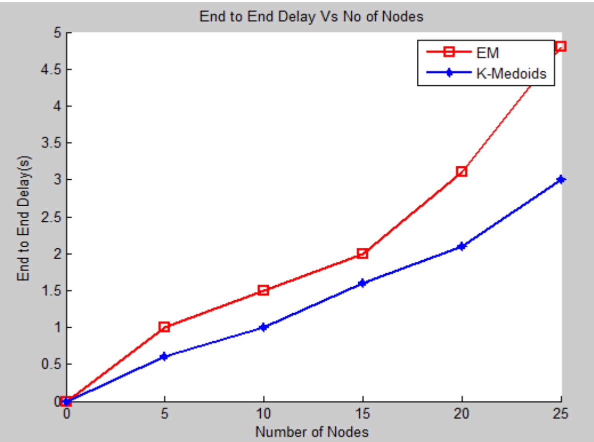

Fig 5: End to End Delay Performance

VII. CONCLUSION

The system gathers big-data energy efficiently using this proposed scheme for large-scale wireless sensor networks. This system suggests that energy efficient big data gathering in such networks is, indeed, necessary. We use optimal routing algorithm for communicating between sink node and cluster head. Once the cluster head gathers information from each node within a cluster, then that information should be passed to sink node in single-hop path. If the communication range is far than normal then multi-hop path is taken for gathering data. By using optimal routing mechanism we can gather big-data energy efficiently and also by using novel approaches it is possible to secure the data effectively from eavesdroppers..

ACKNOWLEDGEMENTS

We would like to thank the management and faculty of RRCE, CSE Dept.

REFERENCES

[1]. C. Saraydar, N. Mandayam, and D. Goodman, ―Pricing and powercontrol in a multicell wireless data network,‖ *IEEE J. Sel. Areas Commun.*, vol. 19, no. 2, pp. 1883–1892, 2001

[2] R. C. Shah, S. Roy, S. Jain, and W. Brunette, ―Data MULEs: modeling and analysis of a three-tier architecture for sparse sensor networks,‖ Ad Hoc Networks, vol. 1, no. 2-3, pp. 215 – 233, 2003

[3]. M. Youssef, A. Youssef, and M. Younis, ―Overlapping multihop clustering for wireless sensor networks,‖ IEEE Transactions on Parallel and Distributed Systems, vol. 20, no. 12, pp. 1844–1856, Dec. 2009.

[4] C. Jiming, X. Weiqiang, H. Shibo, S. Youxian, P. Thulasiraman, and S. Xuemin, ―Utility-based asynchronous flow control algorithm for wireless sensor networks,‖ IEEE Journal on Selected Areas in Communications, vol. 28, no. 7, pp. 1116–1126, Sep. 2010.

[5]. A. Divyakant, B. Philip, and et al., ―Challenges and opportunities with Big Data,‖ 2012, a community white paper developed by leading researchers across the United States.

[6] E. Begoli and J. Horey, "Design Principles for Effective Knowledge Discovery from Big Data", Software Architecture (WICSA) and European Conference on Software Architecture (ECSA) Joint Working IEEE/IFIP Conference on, Helsinki, August 2012

[7] Intel IT Center, "Peer Research: Big Data Analytics", Intel‘s IT Manager Survey on How Organizations Are Using Big Data, August 2012.

[8] Parminder Kaur,Mamta Katiyar, ―Energy-Efficient Chain-Based Data Gathering in WSN,‖ International Journal of Science, Engineering and Technology Research (IJSETR) Volume 1, Issue 6, December 2012.