ENERGY EFFICIENT DATA AGGREGATION IN REGULAR AND RANDOM WIRELESS SENSOR NETWORKS BASED ON COMB-NEEDLE MODEL

A THESIS

Submitted

in the partial fulfilment of the requirement for

the award of the degree of

DOCTOR OF PHILOSOPHY

In

COMPUTER SCIENCE AND ENGINEERING

By

M. SHANMUKHI

[Regd. No. 1103PH0644]

DIRECTORATE OF RESEARCH AND DEVELOPMENT

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD,

KUKATPALLY, HYDERABAD-500 085, TELANGANA

INDIA

DECEMBER 2016

INTRODUCTION

Wireless Sensor Network (WSN) has become ubiquitous and there has been considerable research on energy efficient data aggregation to ensure that life of the network is enhanced. Various data aggregation mechanisms are proposed by researchers for this purpose. The sensor node consumes more energy for performing data gathering, data processing, and communication to sink. The node has very limited battery power, and its battery cannot be recharged or replenished once deployed. Hence, it is very essential to design the energy efficient data aggregation mechanisms for randomly deployed sensor networks. The Comb Needle Model (CNM) is available in literature to perform data aggregation for grid networks. The basic comb needle model is proposed by X. Liu et al [1] for grid network (regular deployment). In this research work that model is extended for supporting random networks. The basic comb needle model is extended by redefining the comb and needle that supports random networks which also uses push and pull strategies for information discovery. The push-pull based comb needle model performs the process of combing for needles in a pool of sand or a haystack. Here the combing operations for needles referred as events in WSN and pool of sand is referred as WSN region.

In this thesis CNM is studied and understood in order to improve it further. Towards this end, different variants of CNM such as extended CNM, cluster-based CNM and extended cluster-based CNM are proposed and implemented. NS2 simulations are used to demonstrate the proof of concept. The following sections provide more details of the study including hypotheses and the evaluation of results against hypotheses.

1. STATE-OF-THE-ART

As a key communication rule, data aggregation plays a vital role in the field of wireless sensor networks. Several performance aspects associated to data aggregation have been widely discussed in this literature survey. For example, significant construction of data aggregation routing protocol is analyzed [2], energy-efficient algorithms in data aggregation are studied [3], and effective scheduling algorithm for reducing the delay in the data aggregation are discussed in [4] and so on.

Apart from the energy efficiency, data exactness or quality and delay are also very essential elements in the data aggregation process. The trade-off between the energy efficiency and data aggregation was initially found by the author Pham et al [5]. Further, this study was analyzed by the Zhu et al [6] for estimating the quality of service in the data aggregation to obtain the energy efficiency in the wireless sensor networks. It is provided dissemination based on the data precision for the specific task. Later on, author Tang and Xu [7] had demonstrated about the data precision in different manner for gathering the data from the various sensors to maintain the minimum energy consumption, which had improved the network life span. Hence, there are various related studies associated with improving the energy efficiency, data aggregation and network life span. Particularly, author [8] had concentrated on reducing the total error bound with improved network life span; they have proposed a new technique to meet the arbitrary data precision requirement of different data aggregation applications in terms of obtaining the energy-efficiency.

Liu et al [9] had introduced the comb needle model, which utilizes push and pull based query strategy in order to obtain the information discovery and data dissemination. The soldier forward its query to the senor nodes in the network (forms comb), this query response is obtained from the source nodes. Therefore, sensor nodes identify the events in the network. These responses are noted as the needles. The process of information collection looks like the combing for the needle in the sand. This is implemented for the grid networks.

2. MOTIVATION

In the basic Comb-Needle Model, whenever an event occurs all the sensor nodes present on the comb participate in transmitting the reply towards the base station. This results in lot of communication cost, which depletes nodes' energy. To lessen this communication cost cluster-based approach is used for data gathering and subsequently data aggregation before forwarding the result towards the base station. Cluster-based approach groups the sensor nodes in the network into certain number of clusters. For each such cluster one node is designated as Cluster Head (CH). All the nodes detecting the events communicate event occurrence to their CH. Then CH aggregates the received data and forwards it towards the base station. There are few issues involved in this clustering approach [10]. First issue is, how many clusters need to be formed that could optimize the communication Cost. Second issue is which node should act as CH. By resolving these issues can leverage the energy efficient data aggregation in WSN. This is the motivation behind the research which is aimed at exploring energy efficient data aggregation in WSN by implementing different variants of CNM.

3. NEED FOR THE RESEARCH

Nodes in WSN are resource constrained. They often work in hostile environments where they are supposed to sense data for different purposes. An important observation here is that the nodes have limited life time due to energy constraint. Therefore, it is indispensable to have energy efficient approaches with respect to all operations of WSN such as data gathering, information dissemination and so on. Optimization of WSN functionality to improve life span of the network is an open research problem to be addressed. Considering plethora of researches which already came into existence, it is very tedious to discover new areas for improvement. Nevertheless, of late CNM came into existence which paved way for Comb-Needle based methodology for energy efficient data aggregation. From the review of literature it is understood that CNM is an important research area which can have huge impact on the WSN when CNM is further improved with different variants.

4. PROBLEM FORMULATION

From the literature it is found that Comb-Needle Model achieves $O(\sqrt{n})$ performance and can balance push and pull in the best case. Nevertheless, the authors of CNM opined that the model needs to be explored with hierarchical structure of WSN for better performance besides considering optimal structures that can leverage data gathering and information dissemination. Therefore this research work is taken up as an optimization problem and different variants of CNM are proposed and implemented. The data aggregation, efficient dissemination of information, reduction of number of communications, energy efficiency, and testing CNM in random networks are all the contributors to the problem formulation. Thus the problem identified is to enhance CNM and make different variants such as extended CNM, cluster-based CNM, and extended cluster-based CNM for making CNM as a useful technique for energy efficient data aggregation in WSN.

5. SCOPE OF THE RESEARCH

The scope of the research is to explore CNM for its energy efficiency besides improving it further with different variants of CNM in order to leverage data aggregation and information dissemination in WSN. The scope includes proposal and implementation of extended CNM, cluster-based CNM, and extended cluster-based CNM and tests them against the hypotheses of the research. The performance of different variants is compared. Conclusions are made and directions are provided for future research.

6. OBJECTIVES

The aim of this research work is to have empirical study on energy efficient data aggregation in WSN by leveraging Comb-Needle model. The following objectives are identified to fulfil the aim of the study:

1. To review the present state-of-the-art on energy efficient data aggregation techniques in WSN.
2. To propose a methodology that can leverage Comb-Needle model with diversified improvements which lead to improvised energy efficiency.
3. To propose and implement an extended Comb-Needle model for reducing energy consumption while data is aggregated in random networks.
4. To propose and implement a cluster-based Comb-Needle model which improves energy efficiency further in grid networks.
5. To propose and implement an extended cluster-based Comb-Needle model to minimize data transmissions and enhance lifetime of randomly deployed WSN.
6. To evaluate the proposed concepts using performance metrics such as energy efficiency, delay, throughput, packet delivery ratio, packet drop and communication cost, and compare them.

7. METHODOLOGY

A methodology is given to guide the study of energy efficient data aggregation in WSN in a systematic way. The proposed concept throws light into different ways and means in which energy efficiency is achieved which is crucial for extended life span of WSN. From the review of literature the following hypotheses are conceived. These are basis for this research work. The following are the hypotheses tested in the research:

1. Comb-Needle Model (CNM) can balance push and pull models for information gathering and dissemination. It results in energy efficiency in large-scale sensor networks.
2. Comb Needle Model can be extended in random networks to prove that the energy consumption and overall communication costs are substantially minimized.
3. Cluster based approach with CNM for grid networks can reduce communication cost thereby improving energy efficiency and performance of data aggregation and information dissemination in WSN.
4. Cluster based approach with CNM for Randomly deployed WSN can minimize the energy consumption and performs the data aggregation very well thereby improves the throughput and life of the WSN is extended

The origin of the research started with push and pulls models for data aggregation and information dissemination in WSN. The push and pull models when employed differently exhibit pros and cons. However they can have synergic effect when combined and balanced together. Thus CNM model came into existence. The CNM model focused on leveraging the combination of pull and push models for energy efficiency and efficient gathering and dissemination of information in WSN. The CNM model is further extended to test the hypothesis such as “Comb Needle Model can be extended in random networks prove that the energy consumption and overall communication costs are substantially minimized". Afterwards the CNM model is explored with clustering besides improving it further for minimizing number of transmissions thereby reducing communication overhead and energy consumption. This results in enhanced lifetime of WSN that is the cornerstone of the research undertaken. The hypothesis "Cluster based approach with CNM can reduce communication cost thereby improving energy efficiency and performance of data aggregation and information dissemination in WSN" is tested. Performance is evaluated among CNM for grid networks, extended CNM for random networks, clustered CNM for grid networks and Extended clustered CNM for randomly deployed WSN.

8. RESEARCH CONTRIBUTIONS

This research is aimed at exploring energy efficient approaches for data aggregation in WSN. Especially it throws light into making different variants of CNM in order to leverage data gathering and information dissemination in WSN for increased life span of the network. The contributions of the research are as follows:

- CNM model is explored for realization of synergic benefits of balancing pull and push models.
- Proposed and implemented an extended CNM in random networks for optimization of CNM for making it more energy efficient model in the arena of data aggregation research in WSN.
- Proposed and implemented a cluster-based CNM in grid networks for reducing number of data transmissions, improving efficiency, decreasing overhead besides enhancing performance of the network in terms of delay, throughput and packet delivery.

- Proposed and implemented an extended cluster-based CNM in random networks for further optimization of data gathering and information dissemination to maximize lifetime of the network.
- Implemented CNM variants using NS2. Extensive simulations are made and the research hypotheses were tested. The empirical study revealed the significance of CNM and its variants in producing high performance sensor networks in terms of energy efficient data aggregation.

9. RESULTS AND DISCUSSION

The study of energy efficient data aggregation is made with WSN. Different variants of CNM proposed are explored for data aggregation efficiency and information dissemination. The following sub sections provide them in some detail. All variants of CNM are evaluated by using the metrics given in next section.

9.1 Performance Metrics

- **Packet Delivery Ratio (PDR):** It is determined as the ratio of overall packets received to the overall packets sent.
- **Throughput: It** is determined as rate of successful message delivered over a communication channel in the random networks.
- **Average Delay**: It means time difference between packets sent and packets received.
- **Energy Consumption**: It is determined as the average energy consumed on idle sleep, data processing, sensing, and data transmission.
- **Communication Cost**: It is determined as the number of packets transmitted and received for query and event notification.

- **Packet Drop**: It is determined as the ratio of number of packets dropped to the overall packets sent.

9.2 Extended Comb-Needle Model

Comb Needle Model is extended for the randomly deployed WSN, which performs efficient data aggregation. When this model is applied to random network (named as Extended CNM), the communication cost, overhead, and energy consumption are reduced significantly when compared to that without CNM. The simulation results for the proposed Extended Comb Needle Model prove that the energy consumption and overall communication costs are substantially minimized. Compared with the simple random network, the Extended Comb Needle Model performance provides better packet delivery ratio, improved throughput, minimized energy consumption and reduced delay. Thus the proposed scheme is energy efficient and it provides optimal solution. The simulation results as well as theoretical analysis prove that when Extended Comb Needle Model in random network is applied, the life of the network is extended.

10.2.1 Assumptions

The following assumptions are made in the extended CNM.

- All the sensor nodes are aware of their location.
- All sensor nodes deployed randomly in sensor region.
- Comb Needle Model is applied in random network as the nodes are randomly deployed so it is difficult to form a comb with perfectly shaped spikes as in basic Comb Needle Model, so the terms - comb (and spikes), and needle are redefined.

- Base station is in top left corner of the sensor region.
- Note: Multiple Base stations are likely to be present in the WSN region as the nodes with soldiers generate queries. Only one sink node for simplicity is considered.
- When a query (e.g., Where is the tank?) is generated, the comb is formed and routes are established.
- Once the tank is found (i.e. event is detected) the event detection information is passed vertically towards the spikes of the comb (resulting in a structure called needle). The route which was already constructed is used for notifying the base station about the event detection. As the route is repeatedly used for query propagation and event notification.

10.2.2 Results

The simulation results show that Extended Comb Needle Model in random networks and Simple random networks provides efficient means for data gathering and information dissemination.

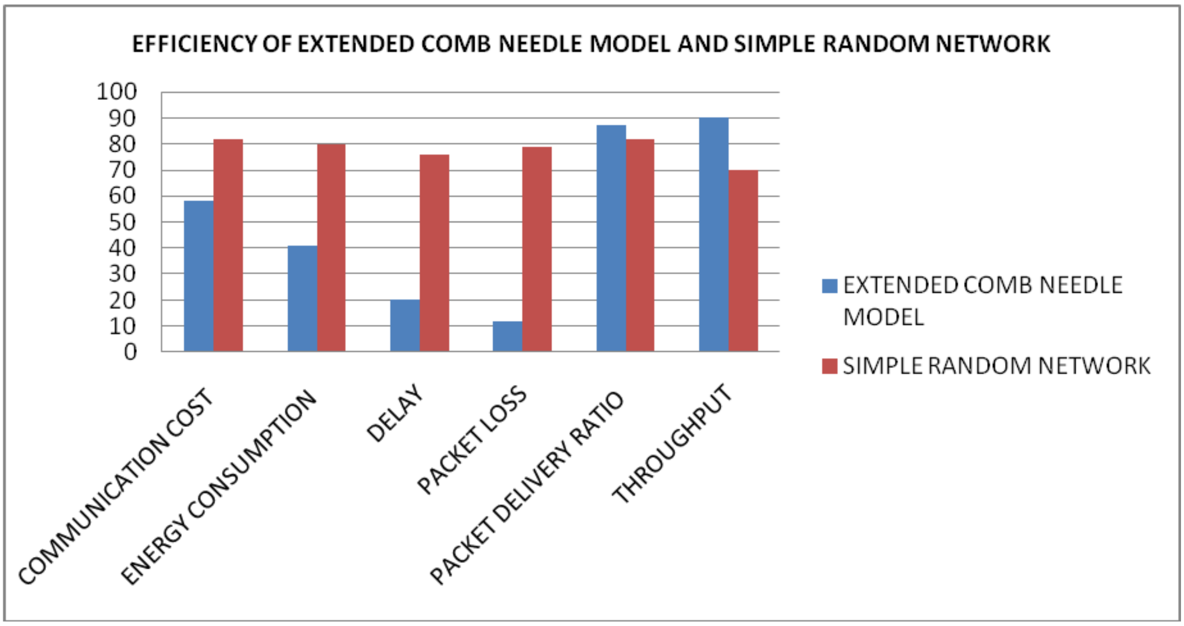

Figure 1. Performance of Extended Comb Needle Model

As shown in Figure 1, communication cost in Extended Comb Needle Model for Random Networks is 58 % and in Simple random Network is 82%. So 24% Communication Cost is decreased in the proposed model. Average Energy Consumption in Extended Comb Needle Model for Random Network is 41 % and in Simple random Network is 80%. So 39 % energy is saved in the proposed model. Delay in Extended Comb Needle Model for Random Network is 20 % and in Simple random Network is 76%. So 56% delay is reduced in the proposed model. Packet Loss is 12 % in Extended Comb Needle Model for Random Network where as 79% in Simple random Network. So 67% is reduced in the proposed model.

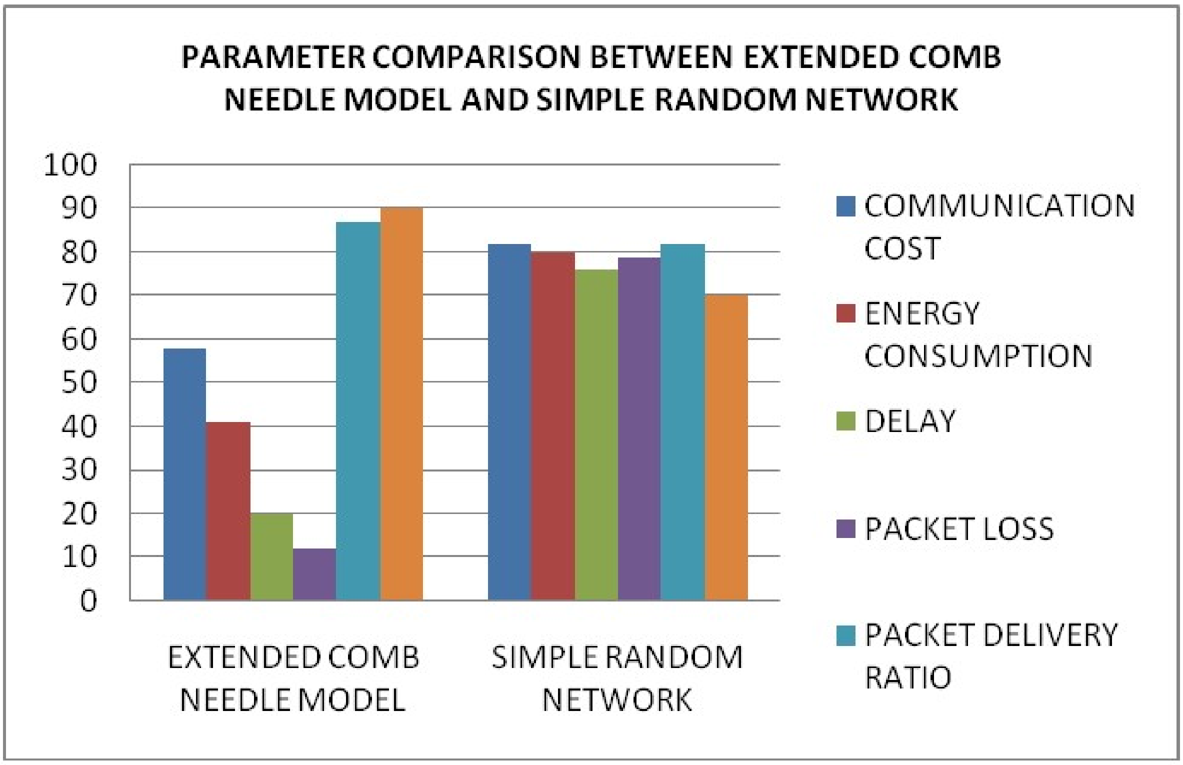

Figure 2. Performance comparison between Extended Comb Needle Model and Simple Random Network.

The results shown in Figure 2 revealed the fact that Packet Delivery Ratio is 87% in Extended Comb Needle Model and 82% in Simple Random Network. 5% is improved in the proposed model. Throughput is 90% in Extended Comb Needle Model and 70% in Simple random network. 20% is improved in the proposed model.

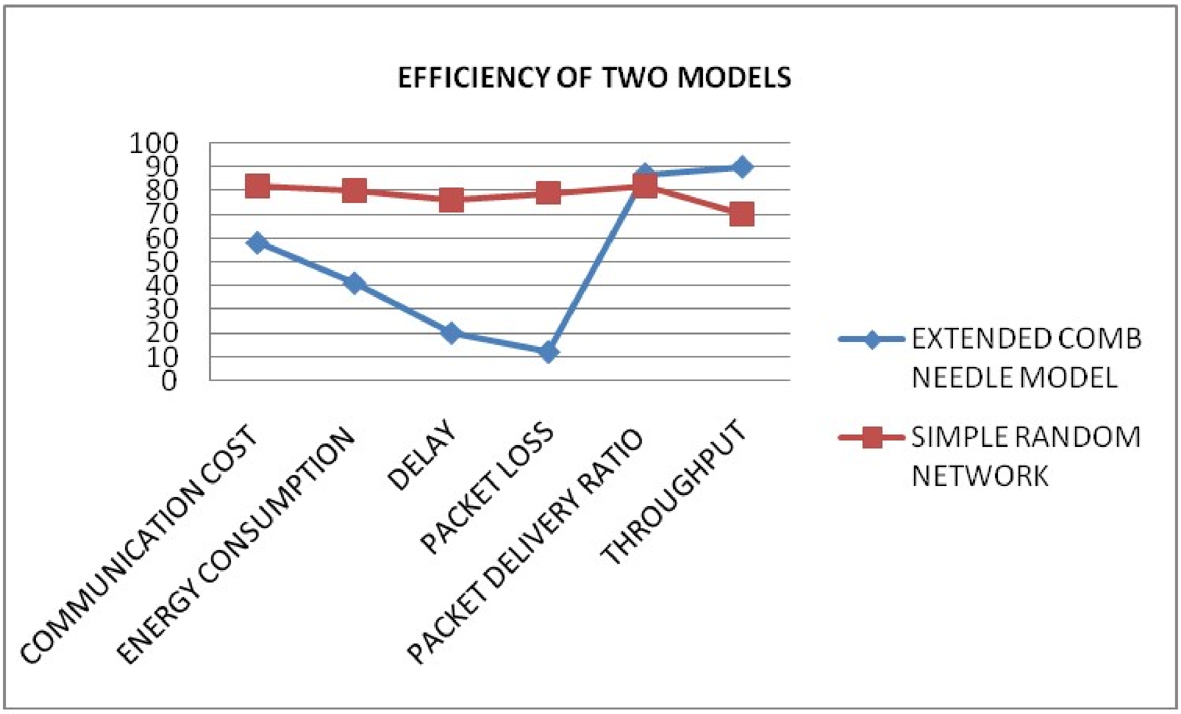

Figure 3. Performance of Two models

As shown in Figure 3, it is evident that the proposed model which is an extended variant of CNM shows very significant performance improvement over its simple random network counterpart.

9.3 Cluster Based Comb-Needle Model

The cluster based CNM deals with information discovery and aggregation in large scale wireless sensor networks applied for mission-critical applications like military reconnaissance. To support query processing based on the gathered information, an efficient and reliable information discovery mechanism is proposed for sensor networks. The basic Comb-Needle Discovery Support Model is extended by including Cluster-based data aggregation mechanism, which helps minimize the communication cost. Cluster-based approach groups the sensor nodes in the sensor network. Each node of a group will send information to its Cluster Head, which then aggregates and forwards the information to the base station (Sink). The performance of proposed Cluster-based Comb-Needle Model and the basic Comb-Needle Model are compared using the parameters, namely, energy consumption and communication cost. The experimental results through Avrora Simulator on TinyOS Platform reveal that Cluster-based Comb-Needle Model for Data Aggregation helps a sensor network in conserving its energy, and thereby extending its life time.

9.3.1 Assumptions

- ✓ Anywhere and anytime the event may occur. Any node can get query at any point of time.
- ✓ All the nodes have information about their locations.

9.3.2 Results

The simulation results show the comparison of Cluster based CNM for grid networks and Basic CNM are presented in this section. The results reveal different parameters such as Communication Cost, Energy consumption, Delay, Packet Loss, Packet Delivery Ratio and Throughput.

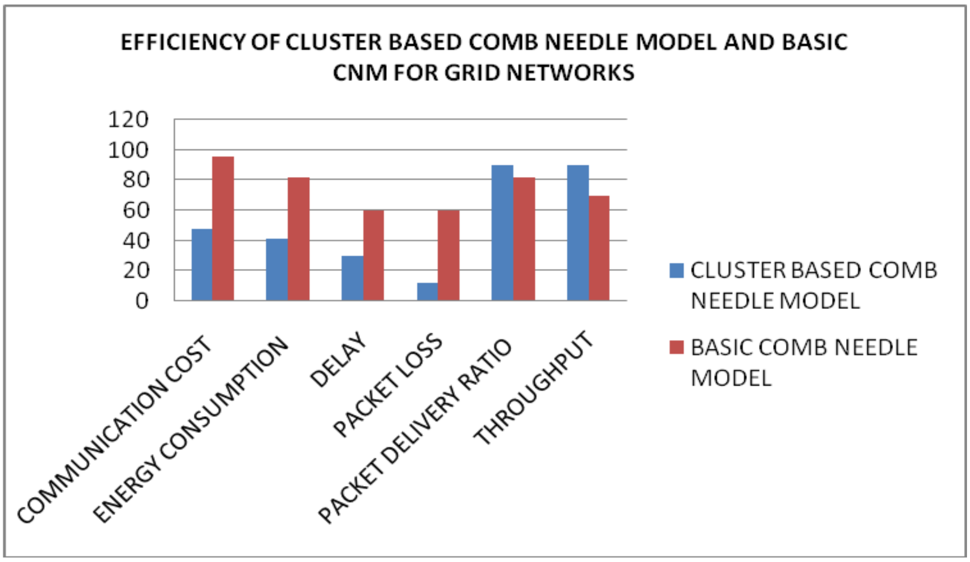

Figure 4. Performance of Cluster based CNM and Basic CNM

As shown in Figure 4, Communication Cost in Cluster based CNM for Grid networks is 48% and in Basic CNM is 96%. So 48% Communication Cost is decreased in the proposed model. Average Energy Consumption in Cluster based CNM for Grid network is 41% and in Basic CNM is 82%. So 41% energy is saved in the proposed model. Average Delay in Cluster based CNM for Grid networks is 30% and in Basic CNM is 60%. So 30% Delay is reduced in the proposed model. Packet Loss in Cluster based CNM for Grid networks is 12% and in Basic CNM is 60%. So 48% Packet Loss is reduced in the proposed model.

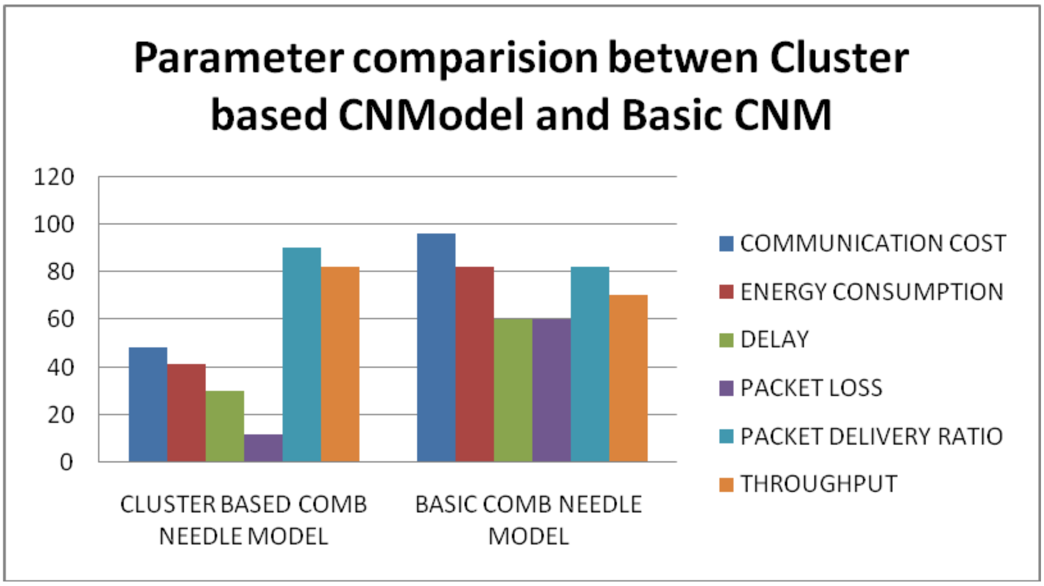

Figure 5. Performance comparison between Cluster based CNM and Basic CNM

The results shown in Figure 5 revealed the fact that Packet Delivery Ratio is 90% in Cluster based CNM and 82% in Basic CNM. 8% is improved in the proposed model. Throughput is 90% in Cluster based CNM and 70% in Basic CNM. 20% is improved in the proposed model.

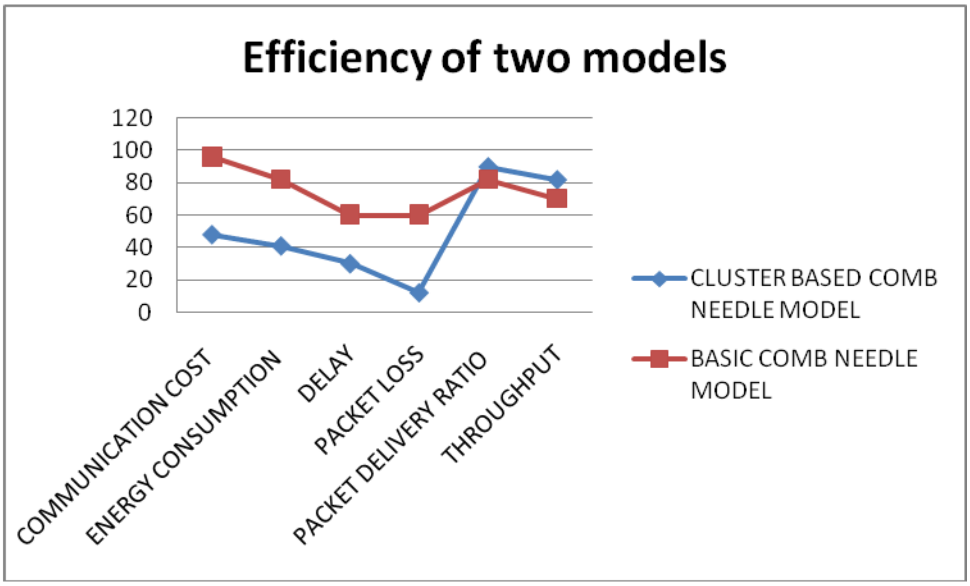

Figure 6. Efficiency of Cluster based CNM and Basic CNM

Communication Cost is reduced in Cluster-based Comb Needle model compared with basic Comb-Needle model (Shown in Figure 6). It is proved that analytically that if Cluster-based data aggregation mechanism in basic Comb Needle Model is included, it reduces the communication cost as well as energy consumption.

9.4 Extended Cluster Based Comb-Needle Model

The extended CNM for random WSNs is augmented with clustering mechanism. In addition, weighted compressive sensing method is added to the Extended CNM to minimize the number of data transmissions, and thereby extend the network life span. Weighted compressive sensing performs better data aggregation, and balances the traffic load during the transactions between sensor nodes to cluster heads, and cluster heads to sink. The CNM uses the push-pull data distribution approach. It may overload certain sensor nodes, and lead to hotspots which causes excess amount of energy loss. The CNM is extended with clustering in random network to overcome these issues and perform energy efficient processing. This paper makes the comparative analysis of the Extended CNM with clustering with that of without clustering. The performance metrics considered are energy consumption, communication cost, delay, packet loss, packet delivery ratio, and throughput. It is empirically observed that the network life span is improved substantially. The simulation is carried out using NS2.

9.4.1 Assumptions

- Wireless sensor network consists of N stationary and location-aware (using GPS or some other localization method) sensor nodes deployed randomly in the square field region of size m X n.
- Base station is in top left corner. (Multiple base stations may be available in network, each soldier is considered as base station)
- All the sensor nodes have the equal transmission power.
- All communication links are assumed to be symmetric.
- There are no transmission errors.
- Sensor node turns as an aggregator, if the size of the data is bigger than the certain limit.
- Clustering is done using ISODATA [31] scheme.

9.4.2 Algorithm (ISODATA)

Input: N_C: Initial number of clusters, M: Final (desired) number of clusters, CT: Convergence threshold,

K: No of clusters required

Output: M number of clusters along with their centroids.

begin

1. Select N_C (< M) initial cluster centroids either arbitrarily or using some criterion;

2. *while* (stopping criteria not satisfied)

    2.1. **repeat** /* k-means Algorithm */

        Assign each pattern to the closest cluster centroid;

        Re-compute the cluster centroids using the current memberships.;

    until (clustering converges) /* Convergence rate < CT */

    2.2. Choose the clusters eligible for split and/or merge;

    2.3. Update the number of clusters N_C ;

    2.4. Determine the current centroids.

End

9.4.3 Results

Simulation results show efficiency of Cluster based Comb Needle Model in random networks and Extended Comb Needle Model in random networks (shown in Figure 7). Communication cost in Cluster based Comb Needle Model for Random Network is 30 % and in Extended Comb Needle Model for Random Networks is 58%. So 28% Communication Cost is decreased in the proposed model. Average Energy Consumption in Cluster based Comb Needle Model for Random Network is 25 % and in Extended Comb Needle Model for Random Network is 41%. So 16 % energy is saved in the proposed model.

- Delay in Cluster based Comb Needle Model for Random Network is 5 % and in

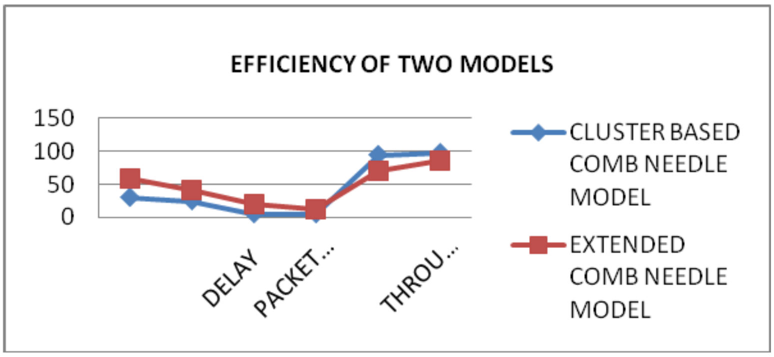

Figure 7 - Efficiency of Two models

Extended Comb Needle Model for Random Network is 20%. So 15% delay is reduced in the proposed model

- Packet Loss is 5% in Cluster based Comb Needle Model for Random Network where as 12% in Extended Comb Needle Model for Random Network. So 7% is reduced in the proposed model

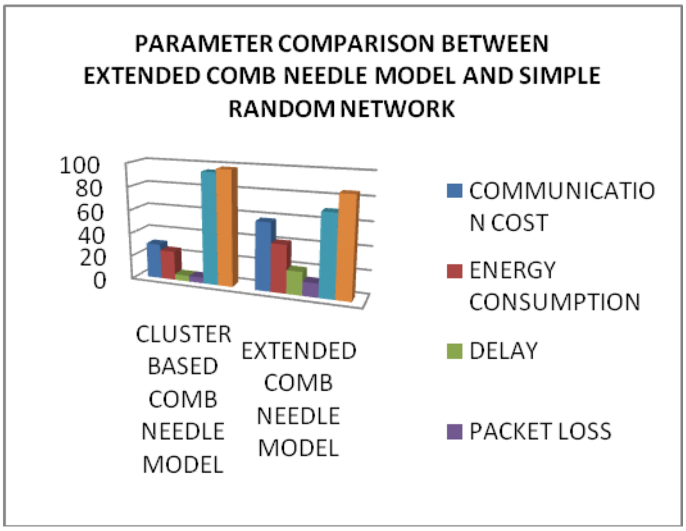

Figure 8 – Comparison of two models

As shown in Figure 8, Packet Delivery Ratio is 95% in Cluster based Comb Needle Model and 87% in Extended Comb Needle Model for Random Network. 8% is improved in the proposed model. Throughput is 98% in Cluster based Comb Needle Model and 90% in Extended Comb Needle Model for Random Network. 8% is improved in the proposed model.

10. CONCLUSIONS AND FUTURE WORK

This research work has brought forth many observations. The following are the conclusions made and directions given for future work.

10.1 Conclusions

- The Extended Comb Needle model for random networks improved and showed better performance. The empirical results in terms of packet delivery ratio, throughput, communication cost, energy consumption, and delay showed performance improvement over basic CNM.
- The Cluster-based Comb-Needle model for grid networks integrates the basic push-based and pull-based data dissemination strategies. It is shown that Cluster-based Comb Needle Model performs better than Basic Comb Needle model. The communication cost is computed for observing query dissemination and query response and found that it was minimized to a great extent.
- The extended cluster-based CNM for randomly deployed WSN improved the performance of WSN further as it made use of aggregated data. The performance metrics considered are energy consumption, communication cost, delay, packet loss, packet delivery ratio, and throughput. It is empirically observed that the network life span is improved substantially.
- Evaluations of hypotheses have been made and are proved positively with empirical evidence.

10.2 Future Work

The following are the directions for possible future work in the area of energy efficient data aggregation in WSN.

- The CNM and its variants are to be enhanced further to use them in Internet of Things (IoT) environment in order to realize the real time need for surveillance and make well informed decisions.
- Another future direction is to investigate whether the cloud has any role to play in future of WSN in the context of IoT as the cloud computing is an established phenomenon and widely used in the real world.

11. REFERENCES

[1] X. Liu, Q. Huang, and Y. Zhang. (2004). “Combs, Needles, Haystacks:Balancing Push and Pull for Discovery in Large- Scale Sensor Networks,” Proc. ACM Conf. Embedded Networked Sensor Systems (SenSys ’04), Nov. 2004.

[2] D.L. Donoho. (2006) Compressed sensing, IEEE Trans. Inform. Theory 52 (4) 1289–1306.

[3] H. Nyquist. (1928). Certain topics in telegraph transmission theory, Trans. Am. Inst. Electr. Eng. 47 (2) (1928)

[4] D . Gong and Y. Yang. (2014). "Low-latency sinr-based data gathering in wireless sensor networks." INFOCOM.

[5] Q. Le-Trung, A. Taherkordi, T. Skeie, H.N. Pham, P.E. (2009). Engelstad, Information storage, reduction and dissemination in sensor networks: a survey, in:

IEEE Conference on Consumer Communications and Networking Conference, pp. 1111–1116.

[6] L. Yuan, Y. Zhu, T. Xu. (2008). A multi-layered energy-efficient and delay reduced chain-based data gathering protocol for wireless sensor network, Mechtronic and Embedded Systems and Applications 13–18

[7] X. Tang, J. Xu. (2008). Optimizing lifetime for continuous data aggregation with precision guarantees in wireless sensor networks, IEEE/ACM Trans. Netw. 16 (4) (2008) 904–917.

[8] D. Gong, Y. Yang, Z. Pan. (2013). Energy-efficient clustering in lossy wireless sensor networks, J. Parall. Distrib. Comput. 73 (9) 1323–1336.

[9] X. Liu, Q. Huang, and Y. Zhang. (2004). "Combs, Needles, Haystacks:Balancing Push and Pull for Discovery in Large- Scale Sensor Networks," Proc. ACM Conf. Embedded Networked Sensor Systems (SenSys '04), Nov. 2004.

[10] Nandini,S.Patil,Prof.P.R.Patil,"Data aggregation in Wireless sensor network," IEEE International Conference on Computational Intelligence and Computing Research,2010.