

Energy Efficient Data Aggregation in Mobile Agent Based Wireless Sensor Network

Divya Lohani¹ · Shirshu Varma¹

Published online: 19 April 2016
© Springer Science+Business Media New York 2016

Abstract Energy efficiency have always been a priority while designing wireless sensor networks. Introduction of mobile agent technology in wireless sensor networks for collaborative signal and information processing has provided the new scope for efficient processing and aggregation of data. Mobile agent based distributed computing paradigm offers numerous benefits over the existing and commonly used client/server computing paradigm in wireless sensor networks. Mobile agent performs the task of data processing and data aggregation at the node level rather than at the sink, thus, eliminating the redundant network overhead. One of the most important issues in mobile agent based paradigm is planning of an itinerary for agent traversal. In this paper, we have proposed a dynamic mobile agent based data aggregation approach that takes into consideration energy efficiency, network lifetime, end to end delay and aggregation ration while making the decision for migration of agent in multihop sensor network. As our approach focuses on finding the most informative route by traversing comparatively less number of nodes consequently mobile agent takes less time to return to processing element, thus, exhibiting lower delay.

Keywords Wireless sensor networks · Distributed computing paradigm · Collaborative information processing · Client server · Mobile agent · Itinerary planning

1 Introduction

The recent advances and developments in the field of micro-electro mechanical system and wireless communications have paved a way for wireless sensor networks (WSN). WSN have come to be known as one of the significant research areas because of their abilities to change the way of interaction between human and physical world. In the last few years,

✉ Divya Lohani
divya_ld@yahoo.co.in

¹ Indian Institute of Information Technology, Allahabad, Allahabad, India

WSN have turned out to be a subject matter of immense interest. WSN is a simple network with little infrastructure, consisting of a large number of tiny sensor nodes with limited energy and computational capability. In WSN, sensor nodes are densely deployed in varied environmental conditions and are used for monitoring factors like temperature, humidity, pressure etc. [1, 2].

Dense and distributed systems capable of performing more complex tasks and inferences are replacing the traditional centralized architectures at a prodigious rates. A distributed sensor network (DSN) is a collection of large number of homogenous or heterogeneous sensor nodes distributed logically, spatially or geographically over a region of interest and connected through a network [3]. The sensors continuously collect data from their environments, processes it and then transmit it through the network. The information gathered from different parts of the network is then fused together to get the final inferences [4]. Since the sensors are randomly distributed in hazardous and unreliable environments, therefore, human intervention is not much required. Sensors are self-aware, self-configurable and autonomous entities collecting data and transmitting it by themselves. Some of the important applications of WSN include: military applications, remote environment monitoring, health monitoring, home automation, and habitat monitoring etc. [1].

The need for using distributed computing in WSN can be justified with the following two points: firstly, sending of all the raw sensed data to the centralized base station consumes a lot of energy and is quite impractical for larger networks. Secondly, since sensor nodes are embedded devices, so using them merely for collecting data is not really an optimum utilization of resources.

Sensor nodes are battery powered devices having constrained computational capability and restricted sensing range, so, collaboration (i.e. correlation among spatially located nodes to reduce the total data transmitted to the processing centre), is required so that the nodes can make up for one another's shortcomings to control the redundant flow of data in the network, thus, making it energy efficient. In order to facilitate implementation of collaborative information processing in WSN, distributed computing paradigms are required [5, 6].

Client server paradigm is one of the most popular and widespread model used for distributed and collaborative computing. In this model, server offers services, resources and other information for the execution of various services, and client sends request to server for execution of service. On receiving the request, server executes the service with the help of the corresponding information contained in it. However, it faces some disadvantages like single point of failure, high bandwidth consumption, huge network traffic and reduced network lifetime.

To overcome these problems, mobile agent (MA) paradigm [7, 8] was proposed for WSN. In this model, in order to perform the allotted task, MAs migrate from node to node using the resources of the nodes. MA based model has many advantages like reduction in network load, overcoming the problem of network latency, robustness and fault tolerance. But at the same time some disadvantages like network overhead, energy issues in case of low density networks are also there [9]. An MA selectively migrates among sensor nodes by moving the processing code among the target nodes, performs local processing by efficiently utilizing resources available at the local nodes rather than transmitting entire redundant data to a central processor (sink), and incrementally fuses local information on each sensor node to reach a progressively accurate global consensus [10].

In this paper, we propose MA based data aggregation approach, in which the route is decided on the fly. We have also included remaining energy of the node as an important parameter in the cost function so that the energy consumption is balanced among the

nodes. Thus, our approach is also able to balance the energy consumption among the nodes that increases the overall life time of the sensor network.

Our approach provides the flexibility in terms of deciding the impact of individual parameters included in the cost function. Thus depending upon the application's requirement one can adjust the weight factors accordingly giving the higher or lower priority to the parameters based on the scenario. Also, if required we can adjust the weight factors of the cost function for changing environments to offer more flexibility in terms of impact of individual parameters in deciding the route of MA migration.

The paper proceeds as follows: Section 2 reviews various MA based approaches. Section 3 gives the network preliminaries. Section 4 explains the proposed approach for data dissemination. Section 5 presents the simulation results and discussions. Section 6 is devoted to the conclusion of the paper.

2 Related Work

Discovering an optimal itinerary for MA traversal is a significant research issue for the effective and efficient data collection from numerous sensor nodes in mobile agent based sensor network [11]. The sequence of nodes in an itinerary has a crucial impact on the quality and accuracy of data fusion which eventually influences the major purpose of WSN, in applications like target tracking or environment monitoring. Itineraries can be categorized in two ways: firstly, whether they are static or dynamic; secondly, whether they use a single agent or multiple agents. Itineraries can be planned either statically or dynamically [11]. Static itineraries are calculated at the processing centre before MA is dispatched while dynamic itineraries are calculated on the fly after MA has been dispatched. Even though dynamic itineraries are quite flexible to changing topologies and are best suited for varying environmental conditions, calculating them result in added computational time and energy consumption as compared to their static counterparts. Itinerary planning can use either a single agent or multiple agents for the network traversal. In single agent based itinerary planning, only one MA traverses the entire network and concludes the task of data fusion; whereas in multi-agent itinerary planning, multiple MAs are dispatched, each MA is assigned a group of source nodes and has an itinerary for traversal to complete the task of data fusion.

In [12], authors have presented two approaches namely: Local Closest First (LCF) in which the MA looks up for the next node having the least distance from the current node; and Global Closest First (GCF), where the MA searches for the next node having the smallest distance from the PE. An approach quite similar to LCF has been proposed in [13] called Mobile Agent-based Directed Diffusion (MADD). MADD differs from LCF only in the selection of the first source node; rather than choosing the nearest node from the PE it starts by selecting the farthest node as the initial one. However, all these approaches take into consideration only the spatial locations of the nodes and thus are not energy efficient. A Genetic Algorithm (GA) [14] based approach does not require finding of any starting node for the algorithm to execute; instead it chooses any active node as the initial node. As, each node has to report its status for maintaining global information at the PE, so GA incurs a lot of communication overhead. Two energy efficient approaches Itinerary Energy Minimum for First-source-selection (IEMF) and Itinerary Energy Minimum Algorithm (IEMA) are presented in [15]. IEMF selects the first source node as the one whose subsequent itinerary has lowest estimated energy cost amongst other itineraries, and then it uses LCF approach to plan the remaining itinerary. IEMA iterates IEMF k times to further enhance the energy efficiency.

3 Network Preliminaries

3.1 Basic Assumptions

Following assumptions have been taken about the sensors and sensor networks in our proposed algorithm for the mobile agent migration:

- Sensor nodes are localized and their location is known in form of Cartesian coordinates.
- Each node has at least two neighbor nodes and deployment is dense.
- Clocks of all sensor nodes are locally synchronized.
- We have assumed 2D plane for the simplicity.
- Energy level at the time of deployment of each and every node is same and known.

3.2 Network Model

In our proposed methodology we have modelled the network as a graph $G(V, E)$ where V is the set of static sensor nodes, $V = \{v_1, v_2, \dots, v_n\}$, and E is the set of wireless links e_{ij} between nodes, $E = \{e_{ij} = \{vi, vj\} | vi, vj \in v, i \neq j\}$. The network consists of N sensor nodes, scattered in a rectangular area A having Gaussian distribution. The PE denoted by v_0 is the start and end points of MA migration route. It is assumed that all sensor nodes are resource constrained especially in terms of energy and bandwidth except for the PE. The sensor nodes are aware of their geographical location in the form of (x, y) coordinates and also aware of their remaining energy. Each node has maximum transmission range R , where they can recognize their neighbor nodes presence in every time intervals. Each sensor node broadcasts a beacon frame including the amount of remaining energy, the geographical location, and the number of times it was visited by MA. It is supposed that the current source node being met by MA is denoted by v_i . The candidates of next source node are denoted by v_j as one of them will be selected as the next hop among the source nodes. The MA only passes through intermediate nodes between current and next source nodes.

3.3 Energy Model

We use the same energy model as stated in [15]. The communication energy at a node is computed by:

$$e(S_{rx}, S_{tx}) = m_{rx} \cdot S_{rx} + m_{tx} \cdot S_{tx} + e_{ctrl}$$

where S_{rx} and S_{tx} are the size of received packet and transmitted packet respectively. e_{ctrl} be the energy consumed in exchange of control messages. m_{rx} and m_{tx} be the energy used in transmitting and receiving a data bit.

Let Src_{ip}^{k-1} and Src_{ip}^k be two nodes and hop distance between them is calculated by $H_{k-1}^k = \frac{d(k-1,k)}{R} \cdot S_{ma}^{k-1}$ is the size of MA at Src_{ip}^{k-1} . The energy consumed in MA traversal from one node to another node is computed as:

$$E_{k-1}^k(S_{ma}^{k-1}) = m_p \cdot S_{data} + e(0, S_{ma}^{k-1}) + H_{k-1}^k \cdot e(S_{ma}^{k-1}, S_{ma}^{k-1}) + e(S_{ma}^{k-1}, 0)$$

where, $m_p \cdot Src_{ip}^{k-1}$ be the data processing energy at Src_{ip}^{k-1} . $e(0, S_{ma}^{k-1})$ is the energy required by Src_{ip}^{k-1} to transmit MA. $H_{k-1}^k \cdot e(S_{ma}^{k-1}, S_{ma}^{k-1})$ is the energy consumed in intermediate sensor nodes between Src_{ip}^{k-1} and Src_{ip}^k . $e(S_{ma}^k, 0)$ is the energy required by Src_{ip}^k to receive MA.

So, the total itinerary energy is computed as:

$$E_{itinerary} = H(t, Src_{ip}^1) \cdot e(S_{ma}^0, S_{ma}^0) + \sum_{k=2}^n E_{k-1}^k(S_{ma}^{k-1}) + H(t, Src_{ip}^n) \cdot e(S_{ma}^n, S_{ma}^n)$$

4 Proposed Approach

Once MA is dispatched from PE it migrates to the nearest node with minimum cost calculated as per the cost function equation. After reaching at first source node, MA aggregates the data at that node. After that MA finds the next candidate for source node.

Firstly the entire one-hop neighbors will be checked using the cost function to designate the least costly candidate node. Second the difference of information gain for this candidate node with the current value of information gain is calculated. If difference is greater than the predefined threshold value, then MA will migrate to that node and perform the data aggregation. Otherwise MA will migrate to the nearest node two hops away from the current node for which data is aggregated.

Because MA has no prior knowledge regarding the network topology so that it first migrates to nearest neighbor at one hop distance that has not been selected; after that MA moves to next closest neighbor decided according to cost function. After arriving at node two hop away, the threshold for information gain is checked. If difference does not satisfies the threshold value MA will migrate to nearest node four hops away from the current node. This process of searching is continued in all 2^n hops nodes until a most informative node is found as the next source node. Upon finding the informative node MA again starts to aggregate the data and finds next node in all one hop neighbors. Finally MA will migrate to PE, when it reaches to a desired level of accuracy or visits all the desired nodes.

Decision of selecting the best candidate node for the next hope is done by cost function C_{ij} . Cost function represents the cost of migration of MA from current source node v_i to next candidate source node v_j . Cost function basically have the tradeoff between increasing benefits and reducing the losses. Motive of cost function is to decrease the energy consumption and increase information gain and network life time. Therefore the cost function tries to increase the probability of selecting the candidate node with the lowest cost among all the candidate nodes for the next hop migration for MA. Cost function C_{ij} for migrating an MA from v_i to v_j have been defined as:

$$C_{ij} = a \left(1 - \frac{I_j(x, y)}{I_{max}} \right) + b(N_{visit} + 1) \left(\frac{e_j}{e_{max}} \right) + c \frac{E_{ij}}{E_{max}} \quad 0 \leq a, b, c \leq 1, N_{visit} \geq 0$$

where,

$I_j(x, y)$: information gain of candidate node v_j for MA migration.

I_{max} : maximum information gain of the nodes.

E_{ij} : amount of energy consumed in migration of MA from v_i to v_j .

E_{max} : maximum energy required for migrating the MA from one node to another.

e_j : remaining energy level of the node v_j .

e_{max} : maximum initial energy level of the nodes at the time of deployment.

N_{visit} : number of visits of node v_j by an MA. Number of times that a node can be visited by MA has been kept limited, because frequent selection of node v_j as the source node may cause more energy loss of a single node and hence may reduce network life time.

a, b, c are the constant weight factors given to the parameters information gain, remaining energy and migration energy respectively.

Definition:
 prev: information gain at previous node
 curr: information gain at current node
 node: $\begin{cases} 1, & \text{if next source node is found} \\ 0 & \text{otherwise} \end{cases}$
 n: Number of hops
 threshold: predefined information gain required for any application

At the PE (v_0):
 Find the first source node using the cost function:
 prev=curr
 node = 1
 n=1
 jump=1

At the node $v_i (i \neq 0)$:
 If node =1 then then{
 Find next source node using above equation:
 gain=curr-prev
 if gain \geq threshold then{
 Select node as next source node
 prev=curr
 node = 1
 n=1
 jump=1
 }
 else
 {
 Find nearest neighbor that has not been selected
 node = 0
 n = n*2
 jump = n
 }
 else
 if node=0 then
 {
 jump = jump-1;
 if jump = 0 then{
 gain = curr - prev
 prev=curr
 node = 1
 n=1
 jump=1
 }
 else
 {
 Find nearest neighbor that has not been selected
 node = 0
 n = n*2
 }
 else
 {
 Find nearest neighbor that has not been selected.
 }
 }

5 Results and Discussion

In the simulation of our proposed scheme we have considered all ways of energy consumption including both communication and computational costs. First, impact weight factor ' α ' of information gain used in the cost function on the performance of our approach have been evaluated. Then we have calculated Aggregation Ratio and End-to-End Delay. And finally we have compared our scheme with the model proposed by Xu-Qi's [7] by varying number of nodes from 100 to 400. Comparison has been done based on the parameters average remaining energy and average end-to-end delay (Table 1).

The metrics used for the purpose of evaluation are as following:

- Aggregation ratio: is the precision of aggregated data by MA. If MA visits all the nodes in the network, precision will be 1.
- End-to-end delay: refers to the time taken by MA from its dispatch by PE to the time it finally returns.
- Residual energy: A node having the highest residual energy among the other nodes in its neighborhood will be chosen. In order to prolong the network lifetime, a node with lower residual energy will not be selected as the subsequent node for MA migration.
- Migration cost: It is the cost of migration of MA from current node to the subsequent node. This cost function is a tradeoff between decrease in energy consumption and increase in information and network life time.
- Information gain: If a source node senses the data all the nodes in that locality may sense the same data with nominal differences. So, instead of migrating the MA to adjacent nodes, migrating it to a far away node helps to accomplish the higher degree of information gain. Thus information gain is directly associated with the nodes distance.

Figure 1 shows the impact of hop count on migration time. It can be seen that with the increase in migration time, the hop count increases i.e. MA migrates in the network for more number of hops and its distance from the MA dispatching unit is more. For 5 hop count, migration time of MA is around 25 ms, which means that MA has been migrating in the network from 25 ms and has traversed 5 hops away from the dispatching unit. Similarly it goes for the other values also. The migration goes on until the desired information gain has been achieved or the energy of the network gets exhausted. However, in extreme cases, the migration time can lowered to 50 ms (hop count between 10 and 15) if the event has occurred in a close proximity of the dispatching unit. Further to evaluate the proposed MA itinerary planning algorithm: aggregation ratio, energy usage and delay have been used as performance metrics.

Figure 2 shows the impact of weight factor on aggregation ratio. In this experiment, we have varied the value of weight factor ' α ' in the range 0–1 and analysed its effect on the

Table 1 Simulation parameters

Parameter	Value
Terrain area	1000 m \times 1000 m
Number of nodes	100–400
Transmission range	115 m
MAC	IEEE 802.11
Simulation time	600 s

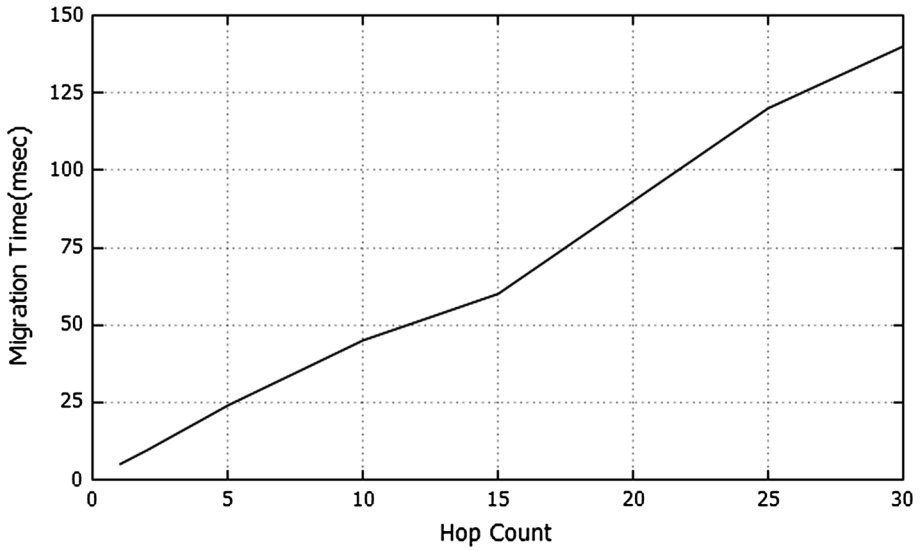


Fig. 1 Migration time versus hop count

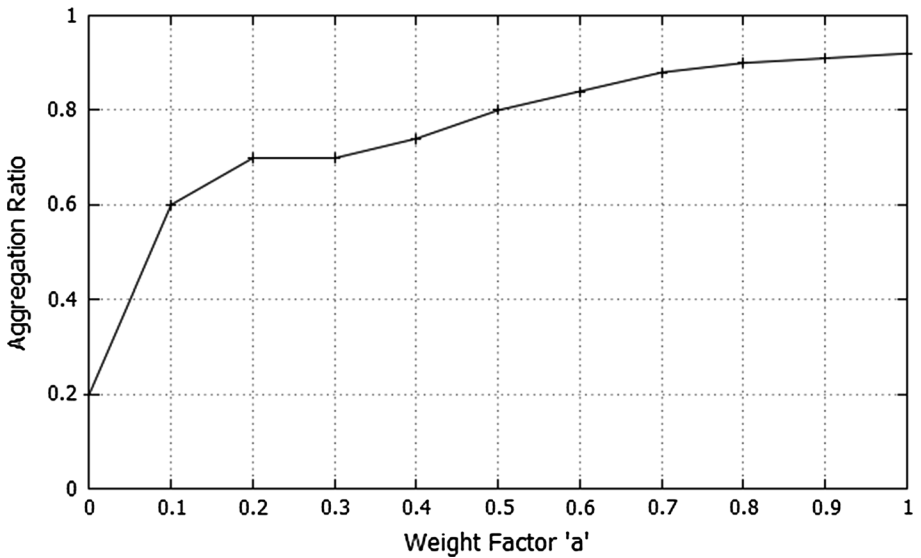


Fig. 2 Aggregation ratio versus weight factor 'a'

aggregation ratio. It can be observed from the figure that there is a direct relationship between the information gain and the weight factor 'a' used in the cost function.

Therefore on increasing the value of weight factor 'a', the effectiveness of information gain on the cost function is enhanced and consequently the data aggregation is performed with higher precision.

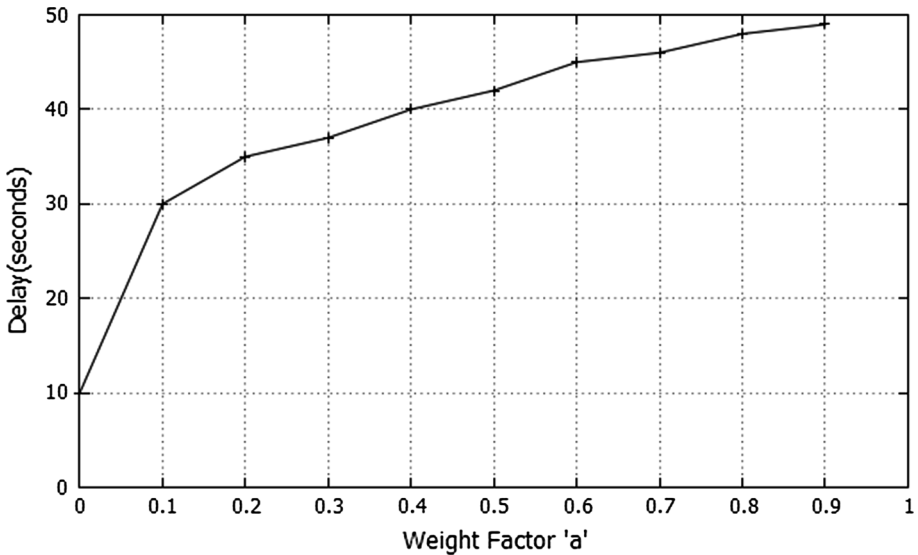


Fig. 3 Average end-to-end delay versus weight factor 'a'

Figure 3 shows the impact of weight factor on end-to-end delay. In this experiment, we vary the value of weight factor 'a' from 0 to 1 and analyse its effect on end-to-end delay. There are basically two related points: firstly information is gained by increasing weight factor 'a'; second, more information is gained by visiting farther nodes from current source node. Considering these points the MA migrates to farther nodes by increasing the value of weight factor 'a'. Consequently, end-to-end delay will be increased.

Figure 4 shows the comparative analysis of average percentage of remaining energy. In the next phase of simulation we have compared our scheme with the Xu–Qi's model [7] based on the average remaining energy percentage. In our simulation we vary the number of nodes from 100 to 400. Since our proposed algorithm selects the next node for MA migration which consumes comparatively less power for transmission of MA. Thus our algorithm balances the energy consumption among the source nodes. Also the desired accuracy is achieved by visiting limited number of source nodes, hence our algorithm can save approximately 40 % energy as compared to existing scheme.

Figure 5 shows the comparative analysis of end-to-end delay. It can be seen that on increasing the number of nodes in the network the end-to-end delay will be increased. But our proposed algorithm has lower delay as compared to existing scheme. The reason for this enhancement is that our algorithm is able to find most informative route by traversing comparatively less number of nodes consequently MA takes less time to return to PE. Our scheme is able to reduce the average end-to-end delay by 12 %.

Thus simulation results verify the practicability of the proposed algorithm. It is clear from the results shown above that our proposed scheme improves the model proposed by Xu–Qi in terms of energy consumption, end-to-end delay and aggregation precision ratio while balancing energy consumption among nodes thus increasing the overall life time of the WSN.

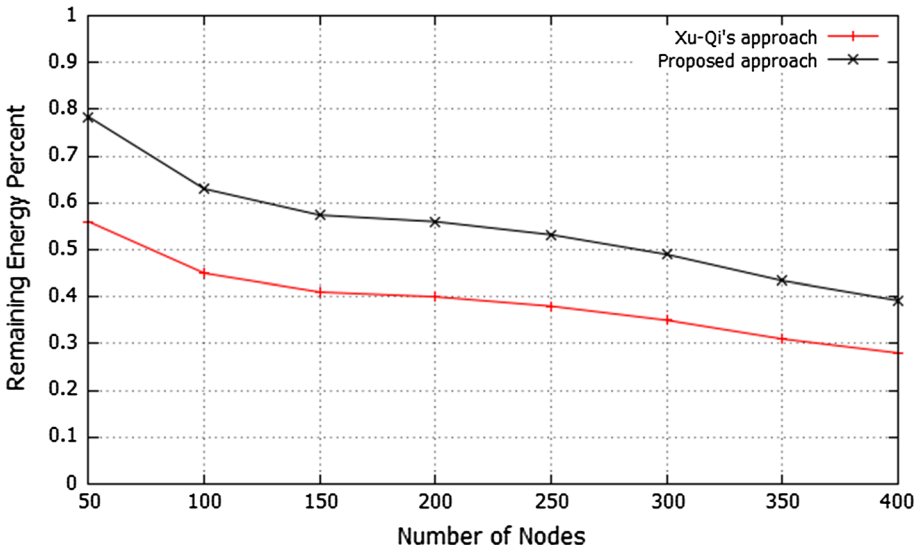


Fig. 4 Remaining energy percent

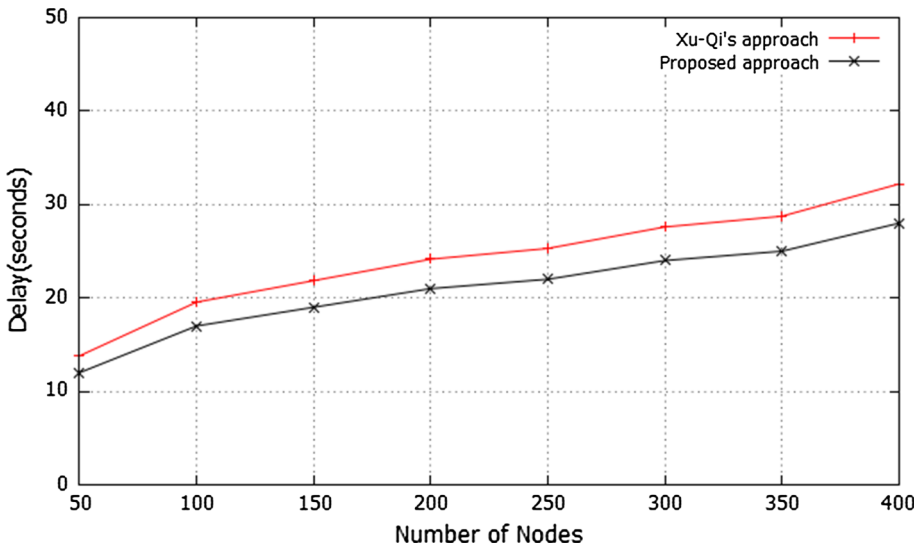


Fig. 5 End to end delay

6 Conclusion

In this paper, the authors' have presented an energy efficient data aggregation approach using MA for WSN. Unlike the previously proposed schemes where, source node selection process is based upon optimum value of cost function (trade-off between energy consumption and information gain), in our approach the cost function has been improved such

that highest possible information gain can be achieved while keeping the energy consumption minimum, thereby increasing the accuracy of aggregated data. Simulation results verify the viability of our approach. It is clear from the results that our proposed approach performs better than earlier schemes in terms of energy consumption, end-to-end delay, aggregation precision ratio and network lifetime.

Future work includes extending the proposed work to support the multiple cooperative agents to further enhance aggregation precision and reduce delay in MA migration. Also, by considering the security and reliability parameters in the cost function, data dissemination process could be made more secure and reliable.

References

1. Akyildiz, I. F., Su, W., Subramaniam, Y., & Cayirci, E. (2002). A survey on sensor networks. *IEEE Communications Magazine*, 40(8), 104–112.
2. Yick, J., Mukherjee, B., & Ghosal, D. (2008). Wireless sensor network survey. *Elsevier Computer Networks*, 52, 2292–2330.
3. Qi, H., Iyengar, S., & Chakrabarty, K. (2001). Distributed sensor networks—A review of recent research. *Journal of the Franklin Institute*, 338, 655–668.
4. Zhao, F., Shin, J., & Reich, J. (2002). Information-driven dynamic sensor collaboration. *IEEE Signal Processing Magazine*, 19(2), 61–72.
5. Xu, Y., Qi, H., & Kuruganti, P. T. (2003). Distributed computing paradigm for collaborative processing in sensor networks. In *IEEE GlobeCom* (pp. 3531–3535).
6. Xu, Y., & Qi, H. (2004). Distributed computing paradigm for collaborative signal and information processing in sensor networks. *Journal of Parallel and Distributed Computing*, 64(8), 945–959.
7. Qi, H., Xu, Y., & Wang, X. (2003). Mobile-agent-based collaborative signal and information processing in sensor networks. *Proceedings of the IEEE*, 91(8), 1172–1183.
8. Chen, M., Kwon, T., Yuan, Y., & Leung, V. C. M. (2006). Mobile agent based wireless sensor networks. *Journal of Computers*, 1, 6–10.
9. Xu, Y., & Qi, H. (2008). Mobile agent migration modelling and design for target tracking in wireless sensor networks. *Ad Hoc networks Journal*, 6(1), 1–16.
10. Biswas, P. K., Qi, H., & Xu, Y. (2008). Mobile agent based collaborative sensor fusion. *Information Fusion Journal*, 9(3), 399–411.
11. Chen, M., Gonzalez, S., & Leung, V. C. (2007). Applications and design issues of mobile agents in wireless sensor networks. *IEEE Wireless Communications Magazine (Special Issue on Wireless Sensor Networking)*, 14(6), 20–26.
12. Qi, H., & Wang, F. (2001). Optimal itinerary analysis for mobile agents in adhoc wireless sensor networks. In *Proceedings of the IEEE international conference on communications (ICC), Helsinki, Finland*.
13. Chen, M., Kwon, T., Yuan, Y., Choi Y., & Leung, V. C. (2007). Mobile agent-based directed diffusion in wireless sensor networks. *EURASIP Journal on Advances in Signal Processing*, 2007(1), 219.
14. Wu, Q., Rao, N. S. V., Barhen, J., Sitharama Iyengar, S., Vaishnavi, V. K., Qi, H., & Chakrabarty, K. (2004). On computing mobile agent routes for data fusion in distributed sensor networks. *IEEE Transactions on Knowledge and Data Engineering*, 16(6), 740–753.
15. Chen, M., Leung, V., Mao, S., Kwon, T., & Li, M. (2009). Energy-efficient itinerary planning for mobile agents in wireless sensor networks. In *Proceedings of the IEEE international conference on communications (ICC), Bresden, Germany* (pp. 1–5).



Divya Lohani is pursuing Ph.D. in the area of Wireless Sensor networks from the Indian Institute of Information Technology, Allahabad. Before this she did her Bachelors of Technology in Information Technology followed by Masters of Technology in Computer Science with specialization in Parallel Computing. Her areas of interest include, but are not confined to wireless communications, adhoc networks, sensor networks, and distributed and parallel computing.



Dr. Shirshu Varma Presently working as an Associate Professor in Indian Institute of Information Technology-Allahabad (IIIT-Allahabad). He has about 20 years of experience of teaching and research. He has published about 75 papers in international and national journals and conferences of repute and is the author of 3 book chapters. Number of citations of his papers is approximately 25. He has total 1 patent awarded and 1 patent in progress. Supervised 3 Ph.D.'s and 2 are in progress in the area of wireless sensor network and distributed sensing. He is a visiting professor of Dongseo University Korea. Visited Kuala Lumpur-Malaysia, Dongseo University Korea. Russian Academy of Sciences, Moscow, Russia, AIT Bangkok, Singapore etc. for various academic purposes. Dr. Varma has worked as a member of many national and international conferences of repute. His area of work includes: wireless sensor networks-coverage and connectivity, sensor deployment and localization, wireless sensor statistical routing etc., WI-FI, WiMAX. He has also worked as a Member of the editorial

board of various international journals. His areas of interest are intelligent sensor network, wireless sensor network-localization, geographical routing, optical wireless communication, and next generation wireless sensor networks. His group is basically focused on analysis of statistical routing, geographical routing in heterogeneous network, localization error etc. in wireless sensor networks with the help of sensor motes using various simulation tools.