

A novel approach for trade-off between computation and communication cost in wireless sensor networks

Narges Arastouie, Masoud Sabaei, Hosseinali Shirali Shahreza

Computer Engineering & Information Technology dept.

Amirkabir University of Technology

Tehran, Iran

{arastouie,sabaei,hosseinshirali}@aut.ac.ir

Abstract— Wireless sensor networks consist of nodes with severe energy constraints and thus, they must use energy saving algorithms in different layers. Despite the prior work, the computational cost in a wireless sensor network is usually neglected compared to the communication cost; it may not admit to in which large volumes of data packets are being transmitted, such as those cases in which This packets are video frames captured by video sensor nodes. In this paper, we consider the problem of minimizing the communication needed to send readings from a set of sensors to a single destination with obtaining a trade-off between computation and communication energy. We propose a centralized as well as a distributed algorithm. In the former, the sink node and in the latter, each node endeavors to reduce network energy consumption adaptively, making a trade-off between data compression at the appropriate level and sending raw data. The data compression technique used is dependent on the application type and considers spatial correlation. Comparing the simulation results of our approach with similar work, without consideration of trade-off between computation and comparison, shows significant improvement in energy saving of the network.

Keywords- sensor network, data compression, energy consumption, computation cost, data entropy, aggregation

I. INTRODUCTION

A wireless sensor network (WSN) consists of sensor nodes that are small in size and equipped with limited sense, processing and transmission capabilities. These networks are being proposed for a wide range of applications, e.g. health systems, military, home and rescue situations. The wireless transmission of data packets constitutes the most energy consuming action in a WSN, especially when there exist multi-hop paths from source nodes towards the sinks. The problem of increased energy consumption in transmitting data packets becomes more challenging when packets of large volumes need to be forwarded. Although this situation is application-dependent, it is more likely to appear as sensor networks are growing rapidly in everyday life. Examples include transmission of video frames of an object tracking application, where the objects locations need

to be reported to the sink in a predefined deadline. In spite of great improvements in the field of compression techniques adapted for video frames, sensor networks still need to trade-off between sending timely data packets and reduced volumes of data after some amount of computation. In fact, in applications with huge packets of huge volumes, it is recommended to perform computational operations on each node's data and to delay transmissions until a certain level of compression and aggregation is achieved.

Data aggregation and gathering techniques, with a great body of literature, can significantly reduce the amount of data transmitted towards the sink nodes. The numerous types of data aggregation schemes are centralized, tree-based, static-cluster-based, and dynamic-cluster-based. Each suggests a method to aggregate data, in terms of the aggregation path. Although the proposed studies suggest very effective methods to aggregate sensing data, there is no means for deciding how much a sensor node cost spends and less of the aggregation methods attend to make a trade-off between computation and communication cost.

The motivation behind the work is that the computational cost in receiving a large data packet, which identifies whether send it as it was received or wait to make one more level of compression combined with a number of other received packets, is usually neglected in current aggregation methods. While most prior work on data aggregation focuses on minimizing only the communication cost, e.g. in [2] and [3], our goal is to minimize the sum of both computation and communication costs simultaneously, and in an adaptive manner. We aim to reduce the energy consumption first by introducing a centralized approach, in which the sink node is responsible for determining the appropriate action in each node. This centralized method is then extended to a distributed algorithm, in a way that each intermediate node makes its own decision on computation and communication. In our distributed algorithm, each node compares the communication cost both with and without computation cost in the aggregation tree as shown in Fig. 1. Our contribution is to represent a comprehensive model that makes a trade-off

between communication and computation cost in order to make the best decision for appropriate level of compression, while considering spatial correlation and sending raw data. We refer to our method as the Adaptive Approach for

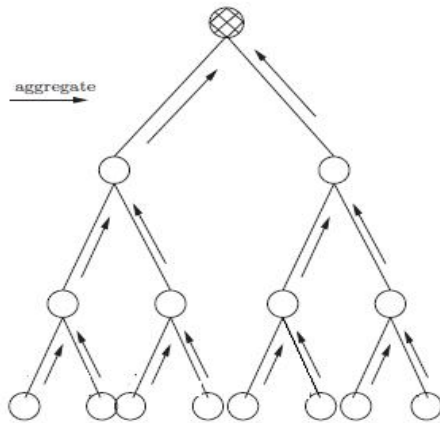


Figure 1: tree-base aggregation

Reducing Aggregation Energy Consumption (ARAE) problem.

The remainder of this paper is organized as follows: in Section II, we describe some related work. Section III illustrates our system model, while numerical results from simulations are presented in Section IV. Finally, concluding remarks are given in Section V.

II. RELATED WORK

A number of previous studies have attempted to reduce the energy consumption, using in-network processing methods. The main goal of these algorithms is to reduce the number of packets propagated within the network, which results in reduced energy consumption in the network. For example, Leone, et al. proposed the Adaptive Blind Algorithm for Energy Balanced Data Propagation in Wireless Sensors Networks [2].

The tree structure has been also exploited for trade-off between computation and communication in data gathering. For instance, Acimovic, et al. [3] developed adaptive Distributed Algorithms for Power-Efficient Data Gathering in Sensor Networks and Liang, et al. [4] has applied online data gathering for maximizing network lifetime in sensor networks. In these works, the authors assume that there is a sequence of data gathering queries, which arrive one by one. In order to respond to each query as it arrives, the system builds a routing tree for it. Within the tree, the volume of the data transmitted by each internal node depends on not only the volume of the sensed data by the node itself, but also on the volume of data received from its children. The objective is to maximize the number of answered data gathering queries until the first node in the network fails, without any knowledge of future query arrivals and generation rates.

As mentioned in the previous section, aggregation techniques have a great importance in the research field of

WASNs. This is due to the simple concept of aggregation which can be executed at each node with minimal effort. In [5] the authors are concerned with the problem of monitoring values of every sensor node over time. This can be considered as a degenerated version of an aggregation function where a node does not actually aggregate or merge its received or sensed data. However, there is a basic notion of spatial-temporal correlation between the sensor values and a periodic monitoring scheme tries to exploit it in a similar way of an aggregation scheme. The authors have mentioned that this spatial-temporal redundancy is exploited in video compression to save storage size and/or bandwidth usage. They attempted to use and modify the basic techniques of MPEG-2 for sensor network monitoring to save communication traffic. A central node calculates predictions and transmits them to all of the nodes. Each node sends its update only if it is different from the prediction received from the central node. There is a hierarchy of nodes to perform this task in a distributed manner. The scheme proposed in [5] raises the question of whether the transmission of prediction is actually more efficient in a large random network or not. Unfortunately, their algorithms are tested in small networks where every node is one hop away from the central node.

Other related endeavors can be found in [6], [7], [8], [9], [10], [11].

Our approach, using joint compression and communication, is closest to recent works presented in [12] and [13]. In [12], Yu et al. has developed a data gathering tree algorithm with tunable compression for communication, named as the Tunable Compression-based Data Gathering (TCDG) problem. To tune data compression over the gathering tree, a flow-based model was proposed where data from each source is compressed and transmitted as a data flow over the corresponding path from the source to the sink. In [13], Eswaran, et al. has developed the same idea. However, they produce a utility-based rate control model with joint congestion control for a WSN consist of multiple competing missions that have different utilities.

While we share the same motivations as in [12] and [13], we develop an adaptive approach for reducing energy consumption, considering sink in central processing where each node in distributed processing is able to make a trade-off between computation cost and communication cost to make the best decision.

III. SYSTEM MODEL

A. Network Model

This paper assumes that sensor nodes are randomly deployed in a square field. By this we mean a large number of sensor nodes are densely deployed in a two dimensional geographic space of size $A \times A$, uniformly distributed and do not move. The sensor network has the following properties:

- 1) The sensor network consists of N nodes, denoted by $s_1, s_2, \dots, s_N$.

2) Each node is demonstrated by (X, Y) in coordination space, while X and Y are random variables, uniformly distributed within the range of [0, A].

3) The wireless network is modeled as a connected weighted graph. $G = (V, E, W)$, where the vertex set V represents the set of sensor nodes and the sink, the edge (link) set E represents the wireless connection between nodes, and a weight w_{e_i} is associated with each edge $e_i \in E$.

4) w_{e_i} is the energy cost of sending a data packet of unit size over e_i . The edge weight is determined by the distance between two adjacent nodes, the radio device, and the communication environment.

5) There exists only one sink node.

6) The energy of sensor nodes cannot be recharged. It means a sensor node will die if its energy is exhausted.

7) Each node has a homogenous energy capacity.

8) Each source node generates a data packet of unit size.

9) A data aggregation tree is a subtree of G rooted at the sink, denoted as $T = \langle V', E' \rangle$, where $V' \subseteq V$ and $E' \subseteq E$.

10) There is a simplified communication mechanism with a medium access control (MAC) protocol that has no packet collision or interference in the tree.

The first three properties are typical assumptions in wireless sensor networks. These assumptions are widely adopted by most previous work.

B. Generic Cost Model of Energy

Generic cost model consist of computation and communication energy in central and distributed approach.

1) Energy consumption model

We do use the compression model, presented in [12]. Following the above rationale and for the purpose of illustration, the energy cost of compressing source information depends on size $F = \sum_{i=1}^n f_i$, where F is related to the length of messages coming from its children, and also its own information in case of being a source node, where n is the number of inputs. f_i is modeled as below:

$$f_i = \begin{cases} \beta_i & \text{if the packet is compressed} \\ S_i & \text{if the packet is not compressed} \end{cases} \quad (1)$$

The received information may contain the compressed form of formerly received packets, which are demonstrated by β_i , or non compressed data, denoted as S_i .

We define a prespecified system parameter, $\gamma \geq 0$, to abstract the energy cost of compressing one unit of data normalized by the cost of communicating one unit of data. The energy required for preparing an output of size f_o is modeled by the following function:

$$\epsilon_{compression}(f_o) = \gamma F \frac{F}{f_o} \quad (2)$$

The computation power required for compression amplifies with an increase in the compression factor (i.e. a decrease in the $\frac{\text{output size}}{\text{input size}}$ ratio) and F .

The overall energy costs, denoted as $\epsilon_{computation}(f_o)$, can then be modeled as follows:

$$\epsilon_{computation}(f_o) = \gamma F \frac{F}{f_o} + f_o \cdot w_e + \epsilon_{recieve} \quad (3)$$

Since the energy of transmitting 1 bit is typically about 500-1,000 times greater than a single 32-bit computation [14], the practical meaning behind $\gamma = 0.1$ is that around 50-100 instructions need to be executed for generating each bit in the output.

The compression is assumed to be lossless. We applied gzip for compression in different levels and the decompression cost is neglected [15]. After doing decompression on β_i , we are able to add the obtained packet to S_i , without increasing any cost.

When the compressed output message f_o , is ready for departure, the relay node sends it on link e with weight w_e . The computation cost is modeled as follow:

$$\epsilon_{computation} = \epsilon_{compression} + \epsilon_{transmit} + \epsilon_{recieve} \quad (4)$$

Written in a simple way:

$$\epsilon_{computation} = \gamma \frac{F^2}{f_o} + f_o * w_e + \epsilon_{recieve} \quad (5)$$

This model shows the output data as f_o that has a common lower bound with joint entropy of source nodes, where H_i is Joint entropy of $i \geq 1$ unit data and δ_v is the number of nodes in the sub-tree rooted at $v \in V$. This lower bound is computed as below:

$$f_o \geq H_{\delta_v} \quad (6)$$

Using a set of pratical joint entropy models from [16], we've define H_{δ_v} . We plot $\epsilon_{computation}$ in Fig 2 with $w_e=1$, $\gamma = 0.1$, $F=1$ and $f_o \in [0.1, 1]$. Obviously, where the first derivation of $\epsilon_{computation}$, the minimum of computation cost, is achieved.

$$\epsilon'_{computation}(f_o) = 0 \Rightarrow f_{opt} = F \sqrt{\frac{\gamma}{w_e}} \quad (7)$$

2) Energy communication model

The communication cost of sending entry data is modeled as bellow:

$$\epsilon_{communication} = \epsilon_{transmit} + \epsilon_{recieve} \quad (8)$$

$$\epsilon_{communication} = F' * w_e + \epsilon_{recieve}(F') \quad (9)$$

In the above formula, F' illustrates the sum of arrival packets, either compressed or not. In other words:

$$F' = \sum_{i=1}^n f_i \quad (10)$$

Where

$$f_i = \begin{cases} \beta_i & \text{if the packet is compressed} \\ S_i & \text{if the packet is not compressed} \end{cases} \quad (11)$$

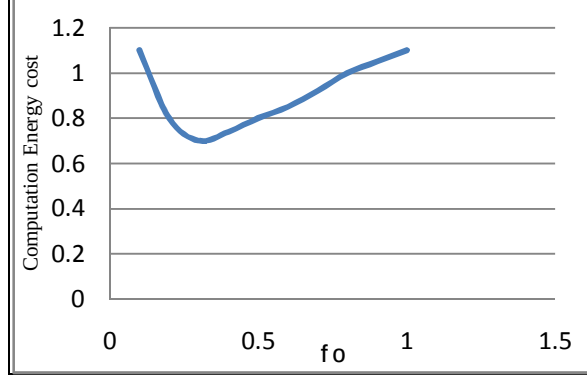


Figure 2: computation energy cost ($w_e=1$, $\gamma = 0.1$, $F=1$)

C. Problem Definition

The regarded problem is reducing the energy consumed within a network. In central assumption, sink make decision for each node to achieve better saving and in distributed sensor network each node can be able to obtaining a trade-off between computation energy and communication energy which is used to send information, does its best to reduce the total energy consumption. In other words, each node endeavors to reduce network energy consumption adaptively, regarding application type, by means of making a trade-off between data compression at an appropriate level considering spatial correlation and sending row data.

D. Proposed Method

In central assumption after the tree is constructed, sink tries to decrease the energy consumption by comparing the computation and communication cost and decision about it in each node. In the other hand, sink can be able to make a trade-off between computation and communication cost for each node which is demonstrated in equations (5),(7) and (9), therefore, every node known data compression at an appropriate level and sending row data or sending data without any computation. So decision making is modeled as below:

$\forall i \in V, \text{if } (\varepsilon_{communication_i} \geq \varepsilon_{computation_i})$
then node i compress data before sending
else node i send row data

We set $f_o = f_{opt}$, so $\varepsilon_{communication_i} \geq \varepsilon_{computation_i}$ can be modeled as follow:

$$\varepsilon_{communication_i} \geq \varepsilon_{computation_i} \Rightarrow$$

$$\begin{cases} F' * w_e + \varepsilon_{recieve}(F') \geq \gamma F \frac{F}{f_{opt}} + f_{opt} * w_e + \varepsilon_{recieve}(F) \\ f_{opt} = F \sqrt{\frac{\gamma}{w_e}} \\ \varepsilon_{recieve}(F') \geq \varepsilon_{recieve}(F) \end{cases}$$

$$\Rightarrow F' * w_e \geq 2F \sqrt{\gamma * w_e}$$

$$\Rightarrow F' * w_e \geq 2F \sqrt{\gamma * w_e}$$

$$\Rightarrow F' \geq 2f_{opt}$$

Sink follow the above semi-code:

```

cn = sink;
Decision-making(tree * cn)
{
  If ( cn != null)
  {
    Decision-making(cn -> Lchild);
    Decision-making(cn -> Rchild);
    If (F' ≥ 2fopt)
    then node i compress data before sending
    Else node send row data
  }
}

```

In distribute assumption, every relay node receives i packets and makes a decision based upon it in each data gathering round. On the other hand, each relay node is able to make a trade-off between computation and consumption energy which is demonstrated in equations (5) and (8), i.e. a decision making operation is performed by an intermediate node. So decision making is modeled as following:

$\Rightarrow$ in each node if $\varepsilon_{communication_i}$
 $\geq \varepsilon_{computation_i}$ compress data before sending
 else send row data

As it was mentioned, we are trying to minimize energy consumption. Each node tries to decrease the energy consumption. The decrease in energy consumption is modeled as follows:

$$\sum_{i \in V} \min(\varepsilon_{communication_i}, \varepsilon_{computation_i}) \quad (12)$$

IV. SIMULATION RESULTS

We assume that the sensor nodes are deployed uniformly in a 3200×3200 two-dimensional field. The sink node was fixed at bottom-left corner of the square. Each node is weighted as $c*d^2$, where d was the distance between the two incident nodes and $c = 2*10^{-5} \text{ J/m}^2$.

We evaluate distributed energy cost parameter by changing different parameters in each scenario. The description of baseline scenarios is as follow:

Scenario 1:

Parameter p is varied within [0, 0.4], when $N = 200$, $\gamma = 0.1$ and $H_i = \rho \sqrt{i} \log i$ [18]. The results are

obtained from running ARAE and TCDG [15] algorithms on SPT (Shortest Path Tree) and MST (Minimal Steiner Tree). the optimal value for ρ equals to 0.1; this is illustrated in Fig. 3.

Scenario 2:

Parameter γ is varied within $[0, 0.24]$, when $N = 200$, $p = 0.1$ and $H_i = \rho\sqrt{i}\log i$. We illustrate the results in Fig 4.

Scenario 3:

In fig 5, Parameter N changes, when $\gamma = 0.1$, $p = 0.1$ and $H_i = \rho\sqrt{i}\log i$.

In general, We observe SPT and MST when p , γ , and N increases in AREA and TCDG algorithms. When p rises, MST demonstrates better performance due to the needed joint compression for further reduction in communication cost. In addition, joint compression becomes less necessary when γ increased, while the advanced computation cost decreased a high compression ratio. Also, we are able to see the energy consumption rise by increasing the number of nodes.

V. CONCLUSION

Energy consumption is critical for wireless sensor networks. In this paper, we first present generic central cost models of energy consumption in sensor networks, then demonstrate a/the distributed energy model is demonstrated. After that we show that each node is able to make a trade-off between computation and communication cost, while in central model, sink is responsible for making this trade-off.

We finally conduct extensive experiments by simulation to evaluate the performance of the proposed method. The experimental results show that ARAE algorithm significantly outperforms TCDG algorithm.

Our future study includes developing an accurate energy model that applies the approaches described here, furthermore we will use more sophisticated joint entropy and accurate models of compression cost in special application to make a trade-off between delay and accuracy.

ACKNOWLEDGMENT

This work was supported in part by Iran Telecommunication Research Center (CTRC).

REFERENCES

- B. Krishnamachari, D. Estrin, and S. Wicker, "The Impact of 1] Data Aggregation in Wireless Sensor Networks," in *Distributed Event-Based Systems (DEBS '02)*, vol. 575578, Vienna, Austria, July 2002, p. 575–578.
- P. Leone, S.E. Nikolettseas, and J. Rolim, "An Adaptive Blind 2] Algorithm for Energy Balanced Data Propagation in Wireless Sensors Networks," in *Distributed Computing in Sensor Systems (DCOSS)*, vol. 3560, Berlin, June 2005, pp. 35–48.
- A. Goel and D. Estrin, "Simultaneous Optimization for Concave 3] Costs: Single Sink Aggregation or Single Source Buy-at-Bulk," in *Symp. Discrete Algorithms (SODA '03)*, vol. 43, 2005, pp. 5–15.
- J. Acimovic, B. Beferull-Lozano, and R. Cristescu, "Adaptive 4] Distributed Algorithms for Power-Efficient Data Gathering in Sensor Networks," in *Symp. Wireless Sensor Networks*, vol. 2, June 2005, pp. 946–951.
- Weifa Liang, Senior Member, IEEE, and Yuzhen Liu, "Online 5] Data Gathering for Maximizing Network Lifetime in Sensor Networks," *IEEE TRANSACTIONS ON MOBILE COMPUTING*, vol. 6, pp. 2–11, JANUARY 2007.
- Y. Zhu, K. Sundaresan, and R. Sivakumar, "Practical limits on 6] achievable energy improvements and useable delay tolerance in correlation aware data gathering in wireless sensor networks," in *Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, Santa Clara, CA, US, Sep. 2005.
- S. J. Baek, G. de Veciana, and X. Su, "Minimizing energy 7] consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation," *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS*, vol. 22, pp. 1130–1140, Aug 2004.
- R. Kanna and S.S. Iyengar, "Game-Theoretic Models for Reliable 8] Path-Length and Energy-Constrained Routing with Data Aggregation in Wireless Sensor Networks," in *IEEE J. Select. Areas Commun*, vol. 22, Aug. 2004, pp. 1141–1150.
- B. Hong and V.K. Prasanna, "Optimizing a Class of In-Network 9] Processing Applications in Networked Sensor Systems," in *Mobile Ad Hoc and Sensor Systems (MASS '04)*, Oct 2004, p. 154–163.
- W. Choi and S.K. Das, "A Novel Framework for Energy- 10] Conserving Data Gathering in Wireless Sensor Networks," in *The 24th Conference of the IEEE Communications Society (IEEE INFOCOM 2005)*, vol. 3, Miami, Florida, USA, Mar. 2005, p. 1395–1404.
- A. Jarry, P. Leone, O. Powell, and J. Rolim, "An Optimal Data 11] Propagation Algorithm for Maximizing the Lifespan of Sensor Networks," in *Distributed Computing in Sensor Systems (DCOSS)*, vol. 4026, June 2006, pp. 405–421.
- R. Cristescu, B. Beferull-Lozano, and M. Vetterli, "On Network 12] Correlated Data Gathering," in *INFOCOM '04*, vol. 4, Hong Kong, Mar. 2004, pp. 2571–2582.
- H. Gupta, V. Navda, S. Das, and V. Chowdhary, "Efficient 13] Gathering of Correlated Data in Sensor Networks," *ACM Transactions on Sensor Networks*, vol. 4, January 2008.
- Yang Yu, Member, IEEE, Bhaskar Krishnamachari, Member, 14] IEEE, and Viktor K. Prasanna, Fellow, IEEE, "Data Gathering with Tunable Compression in Sensor Networks," *IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS*, vol. 19, pp. 276–287, February 2008.
- S. Eswaran, M. Johnson, A. Misra, and T. Porta, "Adaptive In- 15] Network Processing for Bandwidth and Energy Constrained Mission-Oriented Multi-hop Wireless Networks," , Verlag New York Inc, 2009, pp. 87–100.
- V. Raghunathan, C. Schurgers, S. Park, and M.B. Srivastava, 16] "Energy-Aware Wireless Microsensor Networks," *IEEE Signal Processing Magazine*, vol. 19, pp. 40–50, May 2002.
- K. Barr and K. Asanovi, "Energy Aware Lossless Data 17] Compression," in *Mobile Systems, Applications, and Services (MobiSys'03)*, vol. 24, San Francisco, CA, USA, May 2003, pp. 231–244.
- D. Marco, E.J. Duarte-Melo, M. Liu, and D.L. Neuhoff, "On the 18] Many-to-One Transport Capacity of a Dense Wireless Sensor Network and the Compressibility of Its Data," in *Information Processing in Sensor Networks (IPSN '03)*, Apr 2003, pp. 1–16.
- A. Knoll, "Compression of bi-level images: compressor 19] performance report," in *INFORUM 2000 Conference*, Praha, Albertina income, May 2000, pp. 23–25.

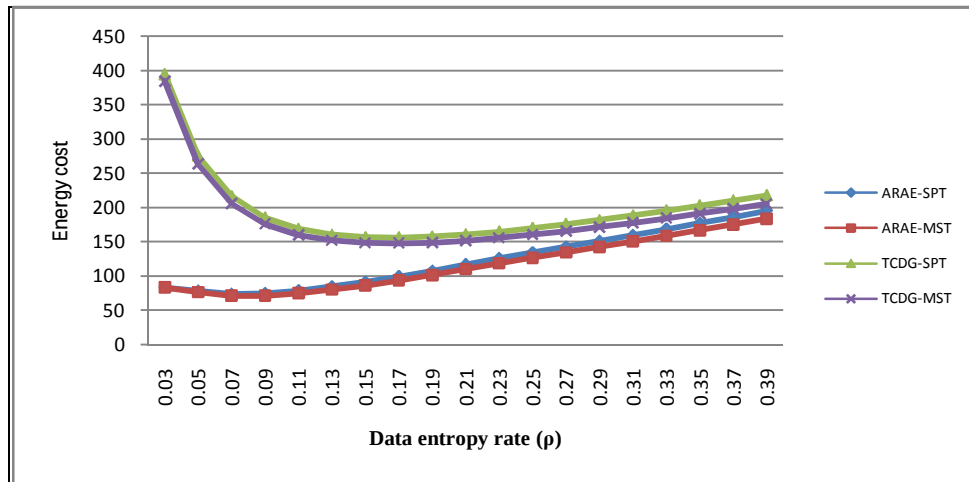


Figure 3: impact of data entropy rate (ρ) with $N=200$ and $\gamma = 0.1$

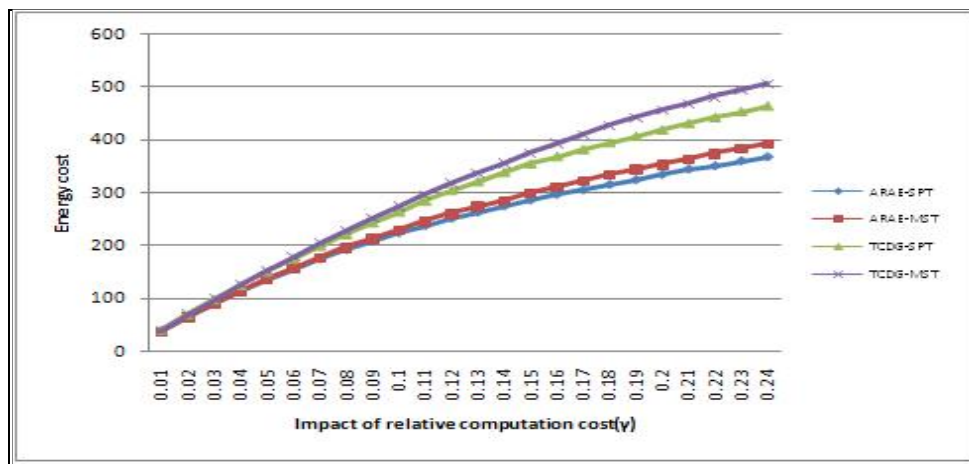


Figure 4: impact of relative computation cost (γ) with $N=200$ and $\rho = 0.1$

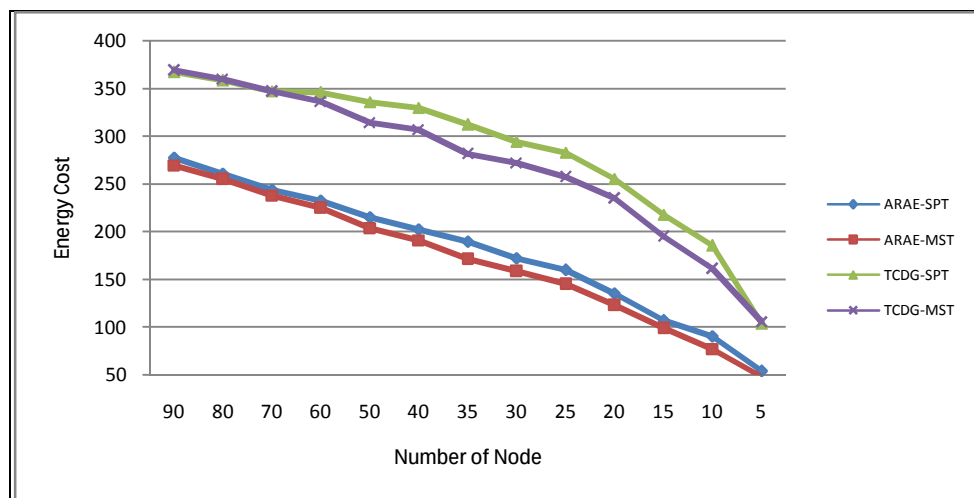


Figure 5: impact of number of source node