CSIT (June 2014) 2(2):97–107
DOI 10.1007/s40012-014-0049-1

ORIGINAL RESEARCH

Cluster head selection by randomness with data recovery in WSN

Devesh Pratap Singh · R. H. Goudar · Bhasker Pant · Sreenivasa Rao

Received: 20 October 2012 / Accepted: 6 July 2014 / Published online: 1 August 2014

Abstract Wireless sensor networks are potentially used in the field of surveillance and monitoring since last few years. In such applications, clustering play an important role in enhancement of the life span and scalability of the network. Each cluster contains a cluster head that controls the whole cluster working, in which various researcher focuses on the good selection of the cluster head that can improve the life of the WSN. Previous works on cluster head selection lacks in data recovery. In this paper we are proposing Cluster Head Selection by Randomness with Data Recovery in WSN (CHSRDR) method for selecting the cluster head inside the cluster with data recovery. The proposed method, CHSRDR, considers the heterogeneity in power and maintains a cluster of vice-heads on the basis of randomness inside the cluster; these vice-heads can work as a head in future, when the main head come to end of power. The headship circulates among the vice-heads of the cluster. We have simulated the method and got the enhancement in throughput.

Keywords WSN · Clustering · Randomness · Recovery · Lifetime

D. P. Singh
Department of Computer Science & Engineering, Graphic Era University, Dehradun 248002, India
e-mail: devesh.geu@gmail.com

R. H. Goudar (✉)
Department of Computer Network and Engineering, Visvesvaraya Technological University, Jnana Sangam, Belgaum 590018, India
e-mail: rhgoudar@gmail.com

B. Pant
Department of Information Technology, Graphic Era University, Dehradun 248002, India
e-mail: pantbhaskar2@gmail.com

S. Rao
SIT, JNTUH, Jawaharlal Nehru Technological University, Hyderabad, India
e-mail: prof.srmeda@gmail.com

1 Introduction

In past few years, Wireless sensor network becomes an interesting field of research. The wireless sensor networks are applicable in the disastrous area, battle fields and for the purpose of surveillance and environmental change monitoring [1, 2]. The wireless sensor network consists of the small battery powered sensor nodes with a sink or base station (BS). It is tough task to recharge the battery of the nodes due to the area of deployment [3]. The sensing and tracking task can be completed by the cooperation of many sensor nodes [4]. For providing the more reliability in the sensor networks, heterogeneous sensor networks are considered for main research in recent past [5–8]. Clustering can be used in the sensor network to improve the lifetime of the network and scalability [9]. Clustering also reduces the size of routing table, use of the power and the communication overhead, because there is less amount of inter cluster communication occur [9, 10]. Generally each cluster contains a leader node that is known as Cluster head. A cluster head can be selected from the nodes present in the cluster or specified by the network designer. Cluster head may be a member of cluster with many resources. The cluster head can be fix or variable; cluster head also reduces the topology overhead because we considered the connection between the cluster head only [11]. The cluster heads are to reduce the rate of consumption of the power by scheduling the tasks inside the cluster, it also perform the task of data aggregation inside the cluster [12]. There are many

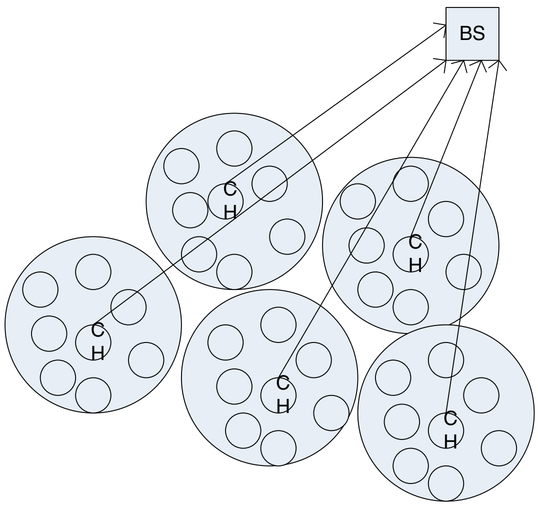


Fig. 1 Network system

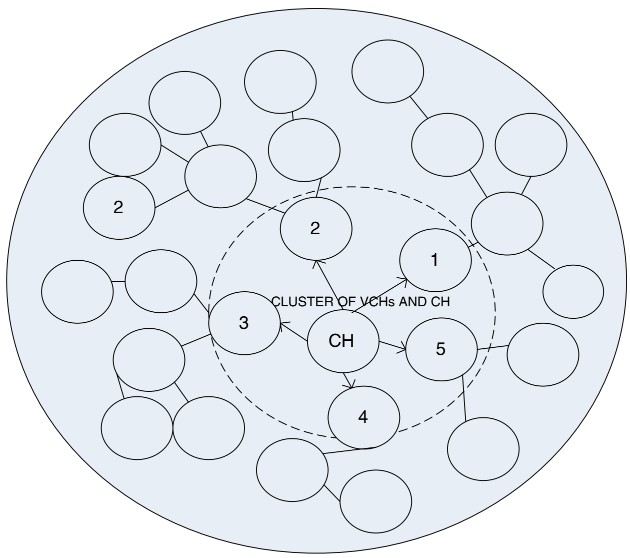


Fig. 2 A cluster with a CH and 5 VCHs

clustering protocols are proposed in the literature e.g. LEACH [13], PEGASIS [14], HEED [15], EEUC [16], and FLOC [17], these protocols do not support the adaptive multilevel clustering, in these protocols the level of clustering can not be changed until new configuration is not made. Therefore these protocols can not be used in the harsh environment of sensor networks.

In this work, CHSRDR, Cluster Head Selection by Randomness with Data Recovery in WSN is proposed. CHSRDR's objective is to provide the cluster selection inside the cluster on the basis of randomness of the network. We have considered the data recovery inside the network by using the vice-heads as the node to maintain the log of the transactions.

This paper is organized as: Sect. 2 discusses the related works, Sect. 3 explains the problem description, Sect. 4 presents the cluster selection algorithm, Sect. 5 gives the

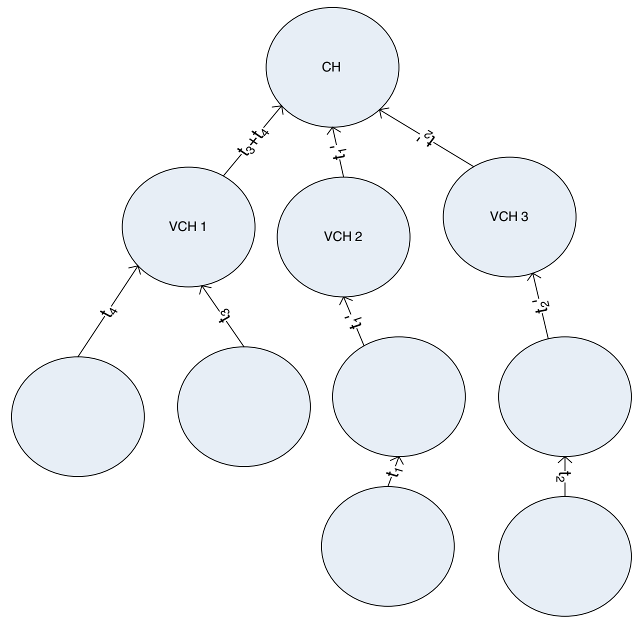


Fig. 3 Result collection process

algorithm for the working of different types on node inside the cluster, Sect. 6 presents the recovery mechanism, Sect. 7 shows the simulation results, Sect. 8 concludes with a brief summary.

2 Related work and motivation

Heinzelman et al. [18] have given a hierarchical clustering algorithm known as LEACH, in which the cluster formation occurs in distributed manner. The role of the cluster head randomly rotates among the nodes to increase the lifetime of the network. It works in two steps, first is setup phase in which cluster formation occur and in other, steady state phase the data distribution takes place. In this algorithm there is no any provision for the recovery.

Hansen et al. [19] introduced a minimum separation distance algorithm, in which they have selected the cluster head from the set of eligible nodes; eligible nodes are selected on the basis of remaining battery of the network. They have not considered the recovery of the nodes in WSN.

TEEN [20] (threshold sensitive energy efficient protocols) designed for real time applications works in hierarchical manner. In this protocol the closer sensor nodes form the cluster, this process repeats in second level until it reached the BS. After the cluster formation the cluster head sends the two threshold hard and soft, these thresholds decrease the number of transmission because when any sensor node senses the environment and get the data beyond the hard threshold, it can send the data only if the attribute change is greater than or equal to soft threshold. If the thresholds are not received by the nodes then we can not get any data from the network.

APTEEN [21] is an extension of TEEN with periodic data collections for real time applications. The problem with this protocol is the two level of formation and maintenance of clusters and how to deal with attribute based queries.

Hybrid Energy Efficient Distributed Clustering (HEED) [15] is a protocol that uses energy and communication cost to elect the cluster head. The intra-cluster and inter cluster communication cost both are used to find the communication cost. This protocol can be used to increase the lifetime of the wireless sensor network. The work done on this is good but HEED lacks recovery of data inside the network.

Distributive Energy Efficient Adaptive Clustering (DEE-AC) [22] introduces the cluster formation on the basis of spatio-temporal variations on the generation of data in different regions. A parameter 'hotness' is decided on the basis of the rate of generation of data on different regions. The cluster head is decided by using the energy and the hotness of the node.

Energy Efficient Distributed Unequal Clustering (EEDUC) [23] presents a method for creating the distributive cluster. Each node in the wireless sensor network maintains a parameter, waiting time. The waiting time is calculated through the consideration of neighborhood nodes.

Power-Efficient Gathering in Sensor Information Systems (PEGASIS) [24] is a chain based protocol, in which the node sets it its strength of signal so that it can reach only to its neighbor node. In this way a chain is formed after forming the chain a head is selected that can only send the data to BS.

Energy Efficient Hierarchical clustering (EEHC) [25] is method of randomized clustering. This method is divided into two stages Initial and extended. In Initial stage (single level clustering) the volunteers are chosen on the basis of the probability of the node. In the Extended stage the clustering is maintained in multiple levels. The cluster head at level 1 send the aggregated data to the head at level 2 and so on.

3 Problem formulation

3.1 System model

We have formulated different system models for elaborating the proposed work, such as Network System, Task Model and Result Collection System etc.

3.1.1 Network system

The network consists of many clusters $N = \{c_1, c_2, c_3, c_4 \ldots c_s\}$, s denotes the number of clusters in the network as shown in fig. 1 above. Each cluster is formed by the set of nodes as $C = \{n_1, n_2, n_3, n_4 \ldots n_m\}$, m gives the total number of nodes in cluster C. Each cluster maintains a cluster-head, $CH \in C$. The nodes in the system consist of different amount of power. The BS is placed at some distance from the network. The BS collects the information from the network. The CHs send the data collected from each node n_i, $i = 1,2,3,4 \ldots m$ present in the cluster c_j, $j = 1,2,3,4,5 \ldots s$ to BS. No other node except CH can send the data to BS.

3.1.2 Task model

The task model can be represented by the help of DAG [26]. Cluster contains one CH and some selected Vice-Cluster-Heads (VCH). The VCH are selected on the basis of randomness as shown in fig. 2 above. The VCHs are the nodes that reside nearby the CH. The Task is distributed by the CH; CH divides the task (T) in parts $(t_1, t_2, t_3 \ldots t_r)$ and assigns the tasks t_i, $i = 1,2,3 \ldots r$ to VCHs, VCHs assigns the task to other nodes, present in the lower levels. The nodes at level i assign some part of the task to itself except node at level 1 and forward the remaining part to the nodes at level $i - 1$, the nodes at level $i - 1$ handles some part itself and forward remaining part to the nodes at level $i - 2$ and so on the task is distributed towards the depth of the DAG. The tasks are assigned to the node on the basis of the performance of the node. Performance of the node depends upon the power of battery. The VCHs can work as CH in future when CH comes to end. Before getting the threshold, CH maintains a set of the VCHs and assigns the charge of CH to VCH that comes first in the set. The Remaining VCHs can work as the nodes for recovery also. Suppose the set of VCHs is $\{VCH_1, VCH_2, VCH_3, VCH_4\}$. If VCH_2 is selected for holding the charge of CH then VCH_1, VCH_3 and VCH_4 can work as the node for recovery purpose, it is possible for the VCH_1, VCH_3 and VCH_4 become CH in future.

The randomness of the node is calculates as follows:

$$Randomness = \sum_{i=1}^{m} P_i \log P_i \qquad (1)$$

The deviation can be measured as follows:

$$Deviation = \sum_{i=1}^{m} P_i \log P_i - P_i \qquad (2)$$

In our system we have placed the BS at the location (0,0) and calculated the randomness for the cluster. We have got the result that the nodes whose randomness is between −2.04383 and 2.839258 are good candidate for selecting the VCH.

3.1.3 Result collection system

Results are collected from the leaf level towards the root node of the DAGs shown in fig. 3 above. Nodes at level i − 4 send the results to their parent at level i − 3. The nodes at level i − 3 collect the results from the level i − 4 nodes. Put the data into aggregated form and forward the data towards the nodes at level i − 2. The VCHs at level one less than the root get the result and forward the data towards the root node (CH). The VCH maintains the log for the data. This log is utilized for the recovery purpose when the current CH gets finished. A VCH is selected as the CH of the cluster; this CH collects the state of the old CH from the log of the different VCHs. The data collected by the CHs is send to the BS of the network.

3.2 Types of heterogeneity

There are three types of heterogeneity in wireless sensor networks.

3.2.1 Power heterogeneity

The nodes inside the WSN contains different amount of battery backup. The lifetime of the nodes inside the network can not be taken same for each node.

3.2.2 Computational heterogeneity

Each node inside the network contains different parameters for performance. Some nodes have better computational power as compared to other nodes present in the WSN.

3.2.3 Link heterogeneity

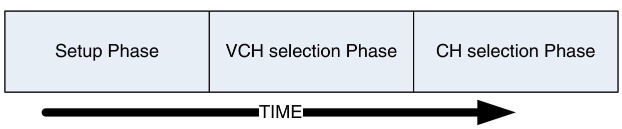
The Link heterogeneity occurs inside the network, this heterogeneity generates the effect on the wireless sensor network.

4 Cluster head selection algorithm

The cluster head selection works on there phases:

(1) Setup phase
(2) VCH selection phase
(3) CH selection phase

The sequence of cluster head selection process is as shown below.




```
Begin(Setup phase)
        BS maintains a record of nodes location and power.
        For each cluster C
        Repeat
                For i=1 to m
                Repeat
                        Calculate the randomness
                        If (-2.04383<=Randomness <=2.839258)
                        Select the node i as the CH.
                        Exit
                End
                BS sends the DB entries corresponding the cluster C to the CH.
        End
End
```

```
Begin(VCH selection phase)
        For each cluster C
        Repeat
                For each node i=1 to m
                Repeat
                        If(Node i=CH)
                                Call slect_VCH()
                                Count_CH++;
                End
                If  (Count_CH !=1) then
                Call the procedure for setup phase for the cluster C
        End
End
```

```
Begin (select_VCH)
        CH broadcast the request for getting the power and location value of
        one hop distance nodes.
        CH maintains a record for reply from the one hop distance nodes.
        For each one hop distance node
        Repeat
                Calculate the randomness of the node
                If (-2.04383<=Randomness <=2.839258) then
                        Select the node as VCH. Put in a list.
                Else
                        Node work as simple node
                End
```

```
Begin (CH selection)
        CH gets the power signal
        If (power< Threshold) then
                CH selects the first VCH (VCH1) from the list.
                Broadcast the message as VCH1 becomes CH
                CH sends the required data to VCH1
        End
        Now VCH1 works as the cluster head.
End
```

5 Algorithms for different types of nodes inside the cluster

In this section we are presenting the algorithm about the work of different types of nodes present in the considered WSN. Here, we are considering that the CH resides at the root level (level 0) of the DAG.

5.1 Algorithm for VCH

The VCH node (level 1) work as the repository of the log and it forward the packet to the parent and child nodes.

```
Start(VCH)
        Read the packet from the QUEUE
Switch(packet Type)
        Case VCH_REQ:
                Create the log entry for the node and send the reply to the
                requesting node.
        Case FLUSH_LOG:
                Append the log entries in the created log of the node.
        Case VCH_DESTR:
                Remove the entry of the log from the VCH for the sender node
        Case COMMIT:
                Remove the log entry from the VCH with the positive response
                to the sender.
        Case REP_NODE:
                Send the reply to the node if it is chosen for recovery.
```

5.2 Algorithm for node at level 2

The nodes at level 2 are working like the simple nodes that perform the assigned part of the task. These nodes are also used to forward and distribute the data.

```
Begin(Node at level 2)
    If(new task assigned )
    Begin
        Assign the part of the task to itself
        Assign the TID to it.
        Request for log entry in the parent VCH.
    End
For i=3 to DEPTH of the DAG
    REPEAT
        Read the packet from QUEUE
        Switch(Packet type)
        Case NODE_REQ:
                Assign a ID to the task and assign the partial task
                to the lower level nodes. Put this ID into the set of
                pending tasks.
        Case ABORT:
                Reassign the task to other child nodes. Readjust
                the entry of the task in the pending list.
        Case REPLY:
                Get the reply of the node from lower level.
        Case REQ_ACK:
                Send acknowledge to the sender nodes.
        Case REQ_TASK:
                Create the part of the task and assign the task to
                the lower level nodes.
```

5.3 Recovery mechanism

Recovery mechanism is used to recover the data of the node, in wireless sensor network node crash occur frequently so recovery mechanism is an important module in WSN.

6 Simulation results

We have taken the metric for evaluating the method as throughput. Before presenting the results, we are presenting the simulation scenario as: there are 20 sensors arranged in two clusters, we have taken the two clusters

Begin (recovery mechanism)

  Broadcast the crash information of the node or CH.

  Select new CH from the VCHs.

  Get the log entries from the remaining VCHs.

  Get the database items and the instruction set of the node or CH.

  Analyze the log and redo the updates

  Begin

    Redo the changes in the data items and remove the uncommitted task's values.

  End

Now the VCH work as the CH with the state of the old CH.

End

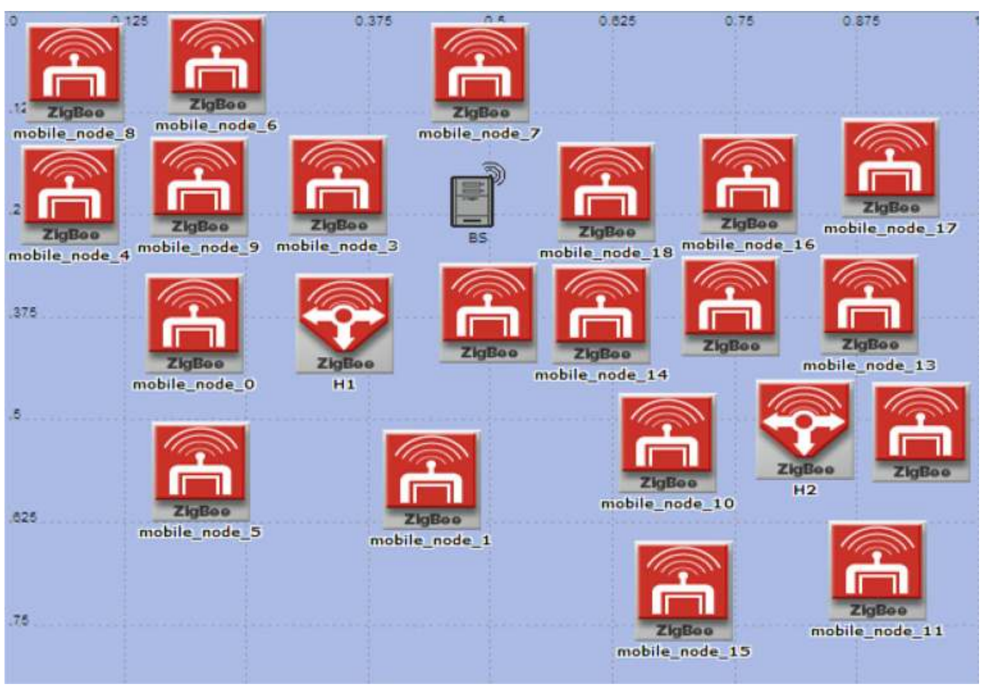


Fig. 4 Simulation scenario for CHSRDR

Table 1 Throughput of the Network with and without application of proposed method

Time (s)	Throughput(bits/s)		Enhancement in throughput (%)
	With application of proposed method	Without application of proposed method	
600	2,223.86	1,666.99	33.41
1200	2,117.739	1,336.074	58.504
10,800	1,549.497	1,042.496	48.63
18,000	2,007.304	1,531.229	31.09

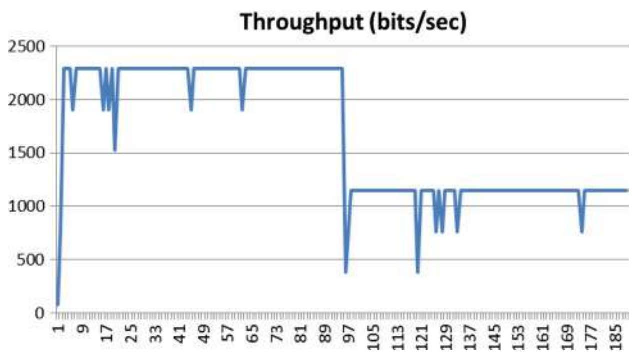


Fig. 5 Throughput without application of CHSRDR (for 600 s)

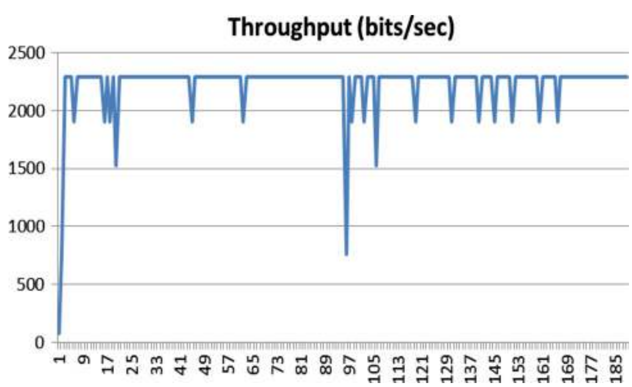


Fig. 6 Throughput with application of CHSRDR (for 600 s)

heads as H1 and H2. Figure 4 shows the simulation scenario taken by us. Table 1 shows the throughput of the network with and without application of proposed method. We have simulated the scenario for different times as 600, 1200, 10800 and 18000 s. Figures 5, 6, 7, 8, 9, 10, 11, and 12 show the throughput of the network with and without application of the proposed method for different times.

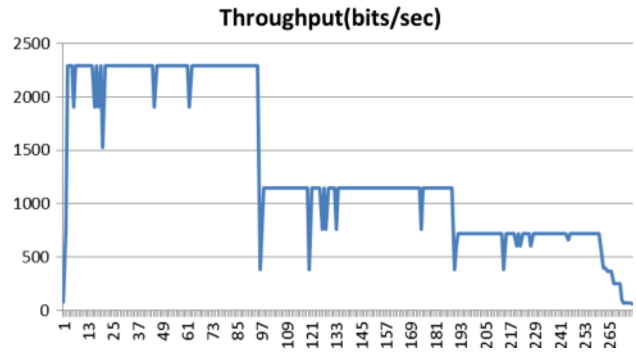


Fig. 7 Throughput without application of CHSRDR (for 1,200 s)

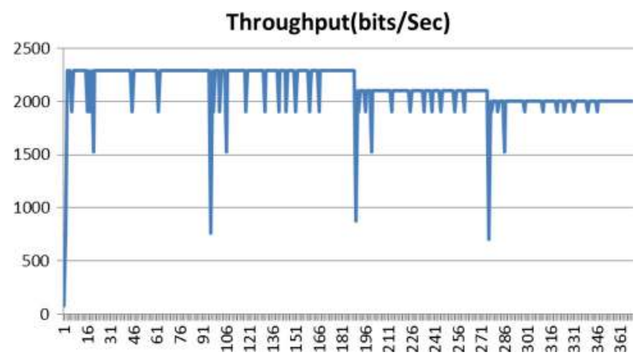


Fig. 8 Throughput with application of CHSRDR (for 1,200 s)

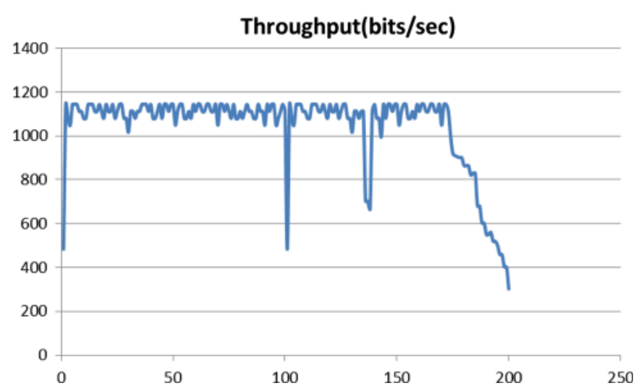


Fig. 9 Throughput without application of CHSRDR (for 10,800 s)

7 Conclusion

This paper presented CHSRDR, a method for cluster selection by randomness with data recovery in wireless sensor networks. This paper provides the efficient cluster head selection within the cluster and do not decrease the throughput when cluster head come to end. The wireless sensor networks are applied in the disastrous area, node

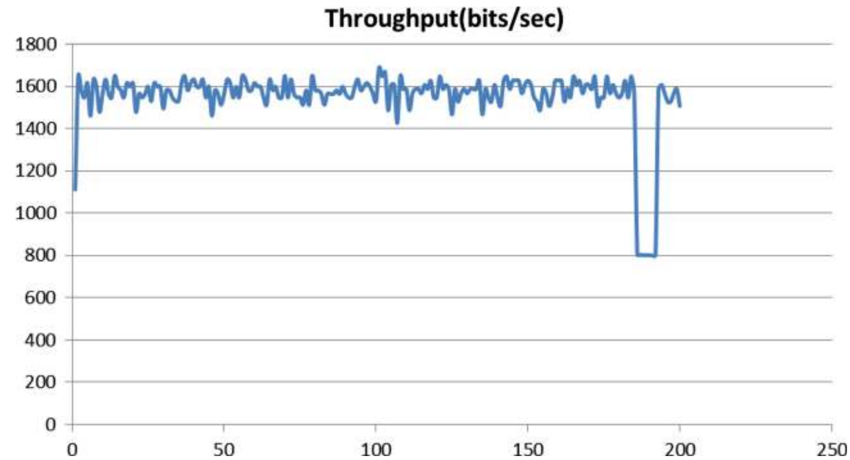


Fig. 10 Throughput with application of CHSRDR (for 10,800 s)

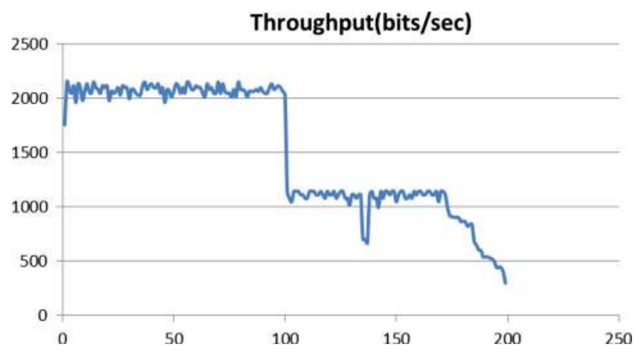


Fig. 11 Throughput without application of CHSRDR (for 18,000 s)

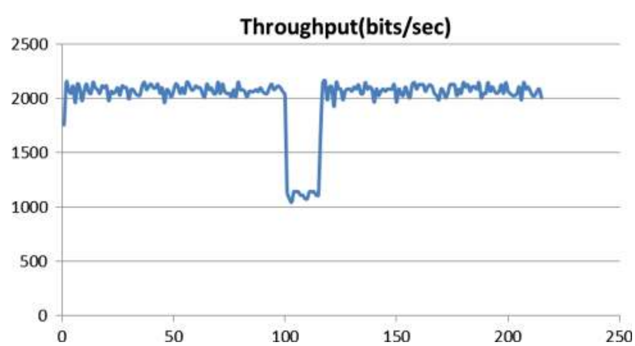


Fig. 12 Throughput with application of CHSRDR (for 18,000 s)

crash happens frequently so the data recovery is a required feature within the network. This paper also provides a data recovery mechanism. The enhancement in the throughput after the application of CHSRDR for 600 s is 33.41 %, 1,200 s is 58.504 %, 10,800 s is 48.63 % and 18,000 s is 31.09 %. Future work can be done on the efficient allocation of the task within the cluster that maintains the efficiency of the CHSRDR.

Appendix 1: Packet types with description

In this section we want to elaborate the meaning of different packet types that are used in the algorithms. Packets are the means to convey the information in the networks. The packet contains the following different fields.

(1) Packet type
(2) Source ID
(3) Destination ID
(4) Data

Packet type	Description
VCH_REQ	The node requires the VCH send the message to create the log entry for that particular node
FLUSH_LOG	The node sends the packet to append the log entries in the log of its own
VCH_DESTR	Destroy the entries of the log in VCH that are created for the sender node
COMMIT	Remove the log entries from the VCH with response in affirmation to that node
REP_NODE	VCH sends the reply to the node if it is selected for recovery
NODE_REQ	Assign the other nodes to that particular task
ABORT	Stop the task performed by the node and assign the task to other child node. Readjust the entry in the pending list
REPLY	Receive the reply from the nodes at lower level
REQ_ACK	Send the acknowledgement to the sender node
REQ_TASK	Divide the task into subtask and assign the task to sender node

References

1. Hart JK, Martinez K (2006) Environmental sensor networks: a revolution in the earth system science. Earth Sci Rev 78:177–191
2. Bokareva T, Hu W, Kanhere S, Ristic B, Gordon N, Bessell T, Rutten M, Jha S (2006) Wireless sensor networks for battlefield surveillance. In: Proceedings of the land warfare conference (LWC-2006), Brisbane, Australia
3. Romer K, Kastin O, Mattern F (2002) Middleware challenges for wireless sensor networks. ACM SIGMOBILE Mob Comput Commun Rev 6(4):59–61
4. Shorey R, Ananda A, Ooi WT (2006) Mobile, wireless, and sensor networks, 1st edn. IEEE Press/Wiley, Toulouse/New York
5. Duarte-Melo EJ, Mingyan (2002) Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks. In: Global telecommunications conference (GLOBECOM'02), vol 1. IEEE, 17–21 November 2002
6. Lu K, Qian Y, Hu J (2006) A framework for distributed key management schemes in heterogeneous wireless sensor networks. In: 25th IEEE international on Performance, computing, and communications conference (IPCCC 2006), April 2006
7. Yu L, Wang N, Zhang W, Zheng C (2007) Deploying a heterogeneous wireless sensor network. In: International conference on wireless communications, networking and mobile computing (WiCom 2007), 21–25 September 2007, pp 2588–2591
8. Wu C-H, Chung Y-C (2007) Heterogeneous wireless sensor network deployment and topology control based on irregular sensor model. In: Advances in grid and pervasive computing. Lecture notes in computer science, vol 4459. Springer, Berlin
9. Younis M, Youssef M, Arisha K (2003) Energyaware management in cluster-based sensor networks. Comput Netw 43(5):649–668
10. Akkaya K, Younis M (2005) A survey on routing protocols for wireless sensor networks. Elsevier J Ad Hoc Netw 3(3):325–349
11. Hou YT, Shi Y, Sherali HD (2005) On energy provisioning and relay node placement for wireless sensor networks. IEEE Trans Wireless Commun 4(5):2579–2590
12. Dasgupta K, Kalpakis K, Namjoshi P (2003) An efficient clustering based heuristic for data gathering and aggregation in sensor networks. In: Proceedings of the IEEE wireless communications and networking conference (WCNC, 2003), New Orleans, LA, March 2003
13. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy efficient communication protocol for wireless micro sensor networks. In: Hawaii international conference on system sciences (HICSS)
14. Lindsey S, Raghavendra C, Sivalingam KM (2002) Data gathering algorithms in sensor networks using energy metrics. IEEE Trans Parallel Distrib Syst 13(9):924–935
15. Younis O, Fahmy S (2004) HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans Mob Comput 3(4):366–379
16. Li C, Ye M, Chen G, Wu J (2005) An energy-efficient unequal clustering mechanism for wireless sensor networks. In: Proceedings of the 2nd IEEE international conference on mobile ad hoc and sensor systems (MASS'05)
17. Demirbas M, Arora A, Mittal V (2004) Floc: A fast local clustering service for wireless sensor networks. In: Workshop on dependability issues in wireless ad hoc networks and sensor networks (DIWANS/DSN)
18. Heinzelman W, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocols for wireless microsensor networks. In: Proceedings of the Hawaii international conference on systems sciences, January 2000
19. Hansen E, Neander J, Nolin M, Bojorkman M (2006) Efficient cluster formation for sensor networks. The Online Publication Documentation System (OPUS), Mälardalen University
20. Manjeshwar A, Agrawal DP (2001) TEEN: a protocol for enhanced efficiency in wireless sensor networks. In: Proceedings of the 1st international workshop on parallel and distributed computing Issues in wireless networks and mobile computing, San Francisco, CA, April 2001
21. Manjeshwar A, Agrawal DP (2002) APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks. In: Proceedings of the 2nd international workshop on parallel and distributed computing issues in wireless networks and mobile computing, Ft. Lauderdale, FL, April 2002
22. Sajjanhar U, Mitra P (2007) Distributive energy efficient adaptive clustering protocol for wireless sensor networks. In: Proceeding of international conference on mobile data management (MDM07), Mannheim, Germany
23. Yu J, Qi Y, Wang G, Guo Q, Gu X (2011) An energy-aware distributed unequal clustering protocol for wireless sensor networks. Int J Distrib Sensor Netw 2011:Article ID 202145. doi:10.1155/2011/202145
24. Lindsey S, Raghavendra CS (2002) PEGASIS: power efficient gathering in sensor information systems. In: 2002 IEEE aerospace conference, March 2002, pp 1–6
25. Bandyopadhyay S, Coyle E (2003) An energy efficient hierarchical clustering algorithm for wireless sensor networks. In: Proceedings of the 22nd annual joint conference of the IEEE computer and communications societies (INFOCOM 2003), San Francisco, CA, April 2003
26. Xie T, Qin X (2005) A new allocation scheme for parallel applications with deadline and security constraints on clusters. In: Proceedings of seventh IEEE international conference on cluster computing