RESEARCH

An efficient multi-level clustering approach for a heterogeneous wireless sensor network using link correlation

Nalamani G Praveena[1*] and Helen Prabha[2]

Abstract

The key issue in determining the lifetime of wireless sensor network (WSN) is the energy burning up of individual node. The cluster based routing improves the energy usage of WSN compared to other routing approach. In this paper, an effective multi-level cluster algorithm using link correlation is proposed for heterogeneous WSN. The level-k hierarchy with single-hop communication between nodes within a cluster is achieved using link correlation. The heterogeneous nodes are adopted as level-k cluster heads and implementing network coding on those nodes increases network lifetime significantly. Meanwhile, implementing time division multiple access (TDMA) technique within a cluster creates an organized cluster architecture improving the energy efficiency.

Keywords: Wireless sensor network (WSN); Link correlation; Network coding; TDMA

1 Introduction

The WSN is widely applied for smart environments such as battlefield surveillance, smart living, environmental monitoring, transportation and traffic monitoring [1]. Due to the recent technology advancement, the sensor nodes are available in a very low cost [2,3]. In WSN, these low-cost sensor nodes are deployed in large numbers either regularly or randomly. Each sensor node consists of a data acquisition module (sensors), data processing module and transmission module (radio transceivers) [1]. Once the network becomes active, the sensor node keeps on sensing the data and transmits it to a base station. The limited energy of sensor nodes is the major constraint in WSN [4]. The inefficient energy usage will drain the battery power faster and make the node die quickly affecting the network lifetime. In WSN, data transmission consumes more energy compared to sensing and data processing [5]. The energy usage of overall network can be improved by reducing the amount of data traffic. The WSN is classified into two types, namely homogeneous and heterogeneous networks [6]. In homogeneous networks, the sensor nodes deployed are ideal in terms of energy and hardware functionalities. In heterogeneous

networks, a certain number of nodes are deployed with more energy and additional processing capabilities to enhance the network lifetime. Since the additional battery cost is very low compared to the sensor cost, adapting the heterogeneous network does not create a huge impact on overall WSN cost.

In WSN, the nodes surrounding the sink will die quickly, due to heavy traffic flow around the sink [7]. This region is referred as bottleneck zone. To overcome this drawback, network coding will be included in those nodes, which reduces the amount of traffic using the same bandwidth. Compared to other types of routing protocols, hierarchy-based routing protocols provide better energy consumption. In hierarchy-based routing protocol, clustering algorithm is considered as the effective routing protocol [8-10]. There are many clustering-based approaches that have been evolved so far. LEACH [11] and PEGASIS [12] are some of the clustering protocols for homogeneous networks. SEP [13] and DEEC [14] are some examples of heterogeneous networks. The SEP protocol is effective for a two-level clustering approach whereas the DEEC is more suitable for a multi-level clustering hierarchy. The communications within a cluster is either single-hop or multi-hop, depending upon the application need. In a single-hop communication, the data of a node is forwarded only to a single neighbor avoiding

* Correspondence: praveenang01@gmail.com
[1]RMK College of Engineering and Technology, Tamil Nadu 601206, India
Full list of author information is available at the end of the article

redundant data transmission in the network. In a single-hop network, reliability is decreased, but the energy consumption is improved. In case of a multi-hop communication, the node transmits data to all the neighboring nodes providing higher reliability at the cost of energy [15].

In this paper, a clustering approach with a single-hop communication based upon link correlation is adopted. The heterogeneous nodes are employed along the bottleneck zone. These nodes act as a cluster head which communicate with the base station. The leaf nodes are randomly employed. At a regular interval, clusters are formed with leaf nodes and heterogeneous nodes. In this paper, section 2 deals with related work such as LEACH, SEP and DEEC in various aspects. Section 3 mainly focuses on the proposed cluster algorithm and discusses the proposed algorithm in terms of its functionality and energy model. Section 4 discusses the performance of the proposed algorithm with LEACH, SEP and DEEC.

2 Reference works

The LEACH [11] is an adaptive clustering algorithm which works well for homogeneous networks. In LEACH, the cluster head selection is based upon a desired percentage of cluster heads and the number of times it has been a cluster head so far. Each node will be given an equal chance to become a cluster head (CH) once every $1/P$ rounds. 'P' is the desired percentage of cluster heads based on our priority, and it is given by k/N. N denotes the total number of nodes in the network and k represents the expected number of CHs for this round. 'G' is the set of nodes that have not been CHs in the last $1/P$ rounds. The threshold value for a node to become a CH in the current round 'r' is given by $T(x)$. For a node 'x' to become a CH, it chooses a random number between 0 and 1. If the number is less than the threshold value, it becomes the CH for that round.

$$T(x) = \begin{cases} \dfrac{P}{1-P \times \left(r \times mod\dfrac{1}{P}\right)} & \text{if } x \in G \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

At the first round, each node has a probability of P to become a CH. The node selected as CH in the first round cannot get a chance for a cluster head in the next $1/P$ rounds, but the probability for remaining nodes to become CHs has been increased. After $(1/P - 1)$ rounds, the nodes that have not yet been CHs so far were given a chance to become CHs. Once $1/P$ rounds were over, again all nodes will be given an equal chance to become CHs. The LEACH is not suitable for heterogeneous networks. SEP [6,13] works well for heterogeneous networks, as it considers the initial energy of nodes

compared to other nodes in the network for choosing the CH. SEP improves the network stability. In a two-level heterogeneous network model, the initial energy of network depends upon normal nodes and advance nodes (heterogeneous nodes). Therefore, the total initial energy E_{total} of a two-level heterogeneous network is given by

$$E_{\text{total}} = E_{\text{nrm}} + E_{\text{adv}} \quad (2)$$

Here, E_{nrm} and E_{adv} are the initial energy of all normal and advance nodes, respectively. Let N be the total number of sensor nodes in the network. There are 'm' fractions of advance nodes and $(1-m)$ fractions of normal nodes. The advance node has 'a' times more energy than normal nodes. Then the total energy can be expanded as in (3),

$$E_{\text{total}} = N(1-m)E_0 + NmE_0(1+a) = NE_0(1+am), \quad (3)$$

where, E_0 is the initial energy of the individual normal node. The total energy of the system is increased by $(1+am)$. In SEP, the probability of choosing the advance nodes to become a CH is higher than the normal nodes. Let P_{nrm} be the probability that a normal node to become a CH, and P_{adv} is the probability that advance nodes to become a CH, which is $(1+a)$ times the normal nodes.

$$P_{\text{nrm}} = \frac{P_{\text{opt}}}{1+am} \quad (4)$$

$$P_{\text{adv}} = \frac{P_{\text{opt}}}{(1+am)} \times (1+a) \quad (5)$$

In Equations 4 and 5, P_{opt} is the optimal probability that a node to be a cluster head based on optimal construction of clusters given by

$$P_{\text{opt}} = \frac{k_{\text{opt}}}{N} \quad (6)$$

$$k_{opt} = \sqrt{\frac{N}{2\pi}} \frac{M}{d_{\text{toBS}}} = \sqrt{\frac{N}{2\pi}} \frac{2}{0.765} \quad (7)$$

where

$$d_{\text{toBS}} = 0.765 \frac{M}{2} \quad (8)$$

where k_{opt} is the optimal number of clusters, d_{toBS} is the distance between a CH to the base station in an M × M square region.

SEP does not consider the heterogeneity of nodes resulting from a node operation. Thus SEP adapts well for two-level heterogeneous networks. DEEC [14] overcomes the disadvantages of SEP and works well for multi-level heterogeneous networks. In DEEC, the initial energy and residual energy have to be considered for a node to become a cluster head. At round 'r', $\bar{E}(r)$ is the average

energy used to estimate the network lifetime and $E_i(r)$ is the residual energy of the node 'x_i'

$$\bar{E}(r) = \frac{1}{N}\sum\nolimits_{i=1}^{N} E(r) \qquad (9)$$

P_i is the average probability for node 'x_i' to become a CH and P_{opt} is the optimal probability of the network. For a node with more energy, P_i is higher than P_{opt}. In normal homogeneous networks, P_{opt} is used as a reference for P_i, as all nodes have an equal amount of energy. By using $\bar{E}(r)$ as a reference energy, the value of P_i is calculated as

$$P_i = P_{\text{opt}}\left[1 - \frac{\bar{E}(r) - E_i(r)}{\bar{E}(r)}\right] = P_{\text{opt}}\frac{E_i(r)}{\bar{E}(r)} \qquad (10)$$

The heterogeneous network contains both normal and advance with different energy levels. The average probability P_i for normal and advance nodes is given by

$$P_i = \begin{cases} \dfrac{P_{\text{opt}}E_i(r)}{(1+am)\bar{E}(r)} & \text{if } x_i \text{ is the normal node} \\[2ex] \dfrac{P_{\text{opt}}(1+am)E_i(r)}{(1+am)\bar{E}(r)} & \text{if } x_i \text{ is the advance node} \end{cases}$$

$$(11)$$

Based on these assumptions, for multi-level heterogeneous networks, the probability $P(x_i)$ of a node x_i to become a CH is expanded as

$$P(x_i) = \frac{P_{\text{opt}}N(1+a_i)}{\left(N + \sum\nolimits_{i=1}^{n} a_i\right)} \qquad (12)$$

Let x_i represent an additional energy factor for the nodes at ith level of heterogeneous networks. In multi-level heterogeneous networks, the hierarchy levels are formed based on the energy of the nodes, such that nodes in the same level have an approximately equal amount of energy. The nodes in a higher level will have more energy than lower level nodes.

3 A proposed cluster algorithm

In this proposal, using link correlation [16,17], a level-k cluster hierarchy (i.e. multi-level) is formed with a single-hop communication. The level-k clusters are the higher level in the hierarchy with {k-1, k-2, −2} denoting the hierarchy of sub-clusters in the subsequent level. The nodes in each level will act as cluster heads for its corresponding sub-level nodes and level-1 nodes act as leaf nodes. The energy level should be considered while forming a cluster hierarchy, such that nodes of a higher hierarchy level have more energy than lower level nodes.

Each node in the network is connected in a single-hop communication to its corresponding cluster head in the above hierarchy level using link correlation, whereas in level-k hierarchy multi-hop communication is adopted between cluster heads. The level-k cluster heads form the bottleneck zone of the sink which has heavy traffic flow. This results in faster depletion of its energy reducing the network life time. To overcome it, the heterogeneous nodes are adopted as level-k cluster head, since it have more energy compared to normal nodes. After establishing a level-k cluster hierarchy using link correlation as described in section 3.1, the level-k cluster head forms a TDMA time slot for its corresponding a level-(k-1) cluster head, while the level-(k-1) cluster head forms a TDMA time slot for a level-(k-2) cluster head and it is followed for all the sub-clusters. The TDMA time slot adopted between cluster heads in subsequent levels of hierarchy helps to remove the data collision and cutback the data aggregation (DA) time.

If the CH at each level randomly transmits to its next-level CH, then data loss will occur due to collision. For avoiding data collision, the CH at each level communicates with the corresponding next-level CH only in its allocated TDMA time slot. Other than the level-k CH, most of the CH incorporates DA to reduce the amount of data for transmission if needed. Data aggregation [18-20] is a suppression technique for removing redundant data that are created due to large deployment of sensor nodes. Since DA is a time-consuming process, the data collection and aggregation are always performed other than its own TDMA time slot with successive cluster heads. For a level-k hierarchy, the data packets are transmitted in a multi-hop communication between cluster heads to reach the base station. This multipath forwarding between cluster heads increases the network traffic. A random linear (RL) [21-23] network coding used in level-k cluster heads reduces the data traffic amount, thus increasing the network lifetime.

Here, a level-3 hierarchy is adopted for performance evaluation with other cluster protocols of WSN. The proposed cluster architecture for level-3 hierarchy is shown in Figure 1. The hierarchical level of a cluster head and sub-cluster head is 3 and 2, respectively. By using link correlation, a single-hop communication is established between level-3 and level-2 CH and between level-2 CH and level-1 nodes. Once the cluster is formed, a level-3 CH creates a unique TDMA time slot for each level-2 CH, and the level-2 CH will create it for level-1 nodes. The structure of TDMA time slot for level-2 is shown in Figure 2. $N_{\text{level-2}}$ represents the number of level-2 CHs. The level-2 nodes are allowed to communicate with the level-3 cluster head only in its TDMA time slot and the same is followed for level-1 nodes to level-2 CHs. The level-2 CH performs data aggregation to remove the redundant data. The level-2 node sends the aggregated data to the level-3 CH in its TDMA time slot. The level-3 CH codes the data using

Figure 1 Cluster architecture.
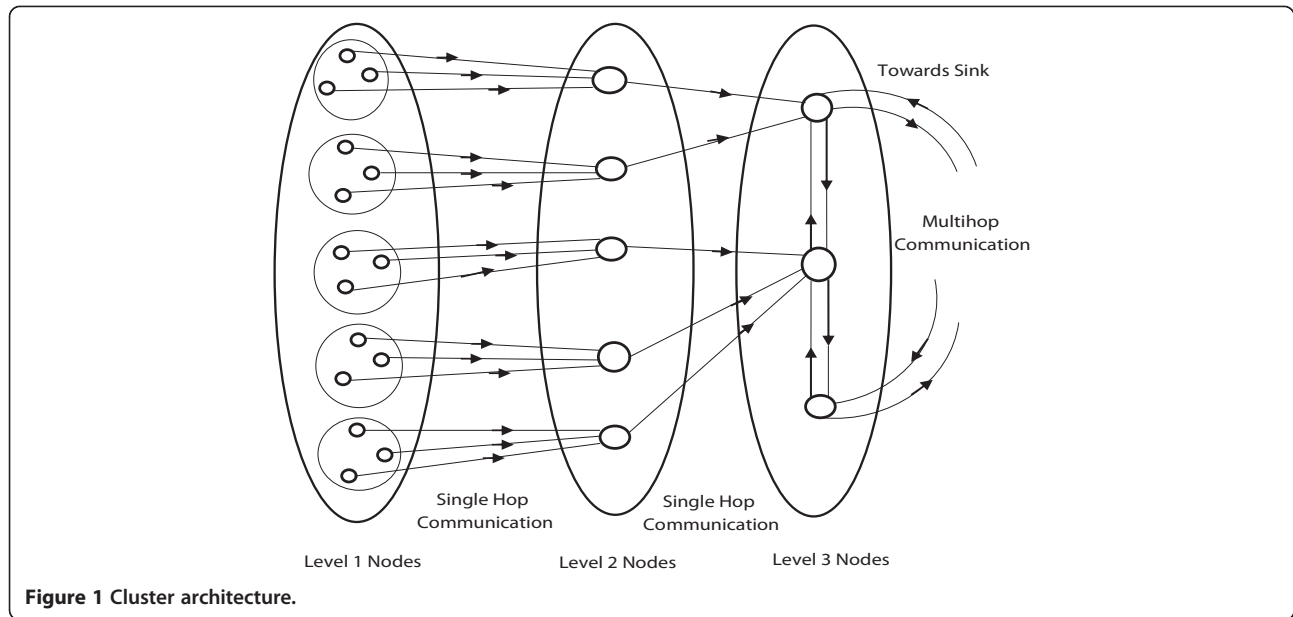
network coding and forwards it to the sink. Due to a multi-hop link between level-3 CHs, the data is forwarded in multiple paths to the sink increasing the network traffic. Network coding overcomes this drawback and provides reliable transmission to the sink.
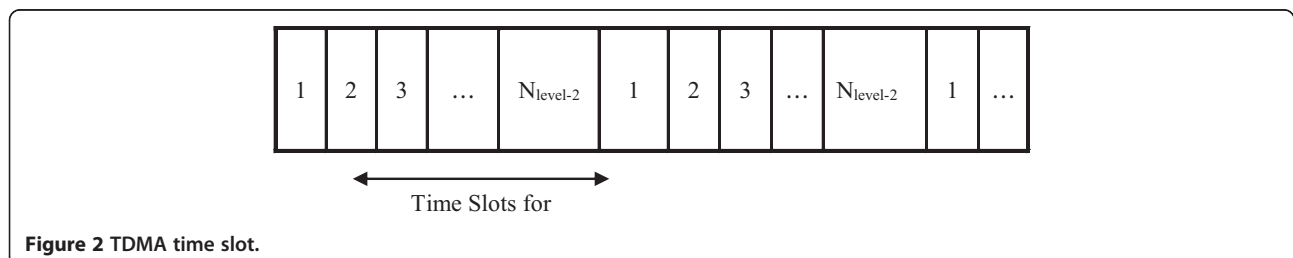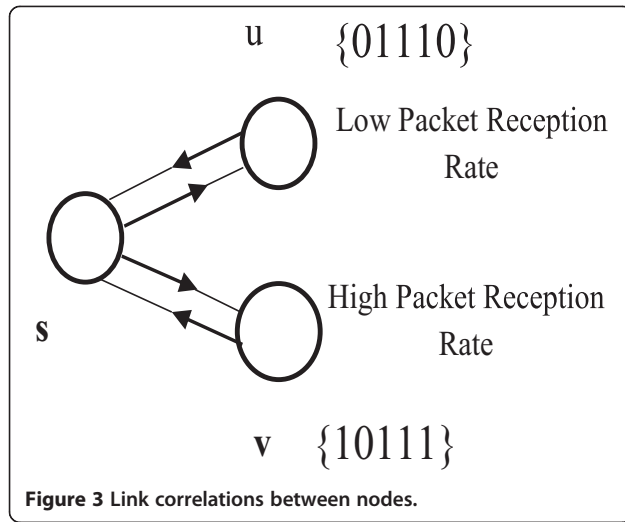
3.1 Link correlation model

In the cluster formation process, a single-hop communication is formed between all levels of clusters, based upon link correlation [16]. The key factor in shaping the link correlation is packet reception rate (PRR) [24,25] which are determined using conditional packet reception probability (CPRP). Initially, each node uses link correlation to choose its cluster head at the next hierarchy level. The nodes will transmit a sequence of 'hello' messages to the next-level nodes in short-time duration. 't' is the number of messages transmitted. The hello messages consist of a sender ID and packet number. The next-level nodes store the hello messages in a separate bitmap format from all possible nodes of a sub-hierarchy level. The bitmap formats consist of 1 s and 0 s for the message reception status as success and failure arranged upon the packet number. After receiving 't' hello messages, the next-level

nodes will transmit its bitmap as a packet in the form of node ID, sender ID, and bitmap. The nodes will receive the packet sent from next-level nodes when sender ID in the bitmap packet matches with it, for finding the link correlation using CPRP. The nodes choose its cluster head in the next hierarchy level that is having high link correlation with it. By using this link correlation technique, hierarchies of clusters are formed up to level-k. After establishing a cluster hierarchy, the stored bitmap is cleared to save memory.

Figure 3 shows a link correlation model; here the link correlation is measured between a node in a sub-hierarchy level and two nodes in next hierarchy level. Let s be the nodes in sub-hierarchy levels which act as a sender, and u and v are the next-level nodes which act as receivers. The sender s transmits five hello messages for short-time duration (i.e. t =5). X_u denotes the bitmap {01110} in receiver u for sender s, and X_v denotes the bitmap {10111} in receiver v for sender s. $P\left(\frac{u}{v}\right)$ and $P\left(\frac{v}{u}\right)$ are the CPRP for u and v, respectively.

$$X_{uv} = \sum_{i=1}^{t} x_u[i] \times x_v[i] \qquad (13)$$

1	2	3	...	$N_{level-2}$	1	2	3	...	$N_{level-2}$	1	...

Time Slots for

Figure 2 TDMA time slot.

Figure 3 Link correlations between nodes.

$$X_u = \sum_{i=1}^{t} x_u[i] \tag{14}$$

$$X_v = \sum_{i=1}^{t} x_v[i] \tag{15}$$

$$P\left(\frac{u}{v}\right) = \frac{X_{uv}}{X_v} \tag{16}$$

$$P\left(\frac{v}{u}\right) = \frac{X_{uv}}{X_u} \tag{17}$$

X_{uv} specifies the number of packets received in both receivers. $x_u[i]$ and $x_v[i]$ represent an ith bitmap entry. By using these CPRP of two receivers, the PRR for receiver u and v are calculated as

$$\text{PRR}_u = 1 - P\left(\frac{v}{u}\right) \tag{18}$$

$$\text{PRR}_v = 1 - P\left(\frac{u}{v}\right) \tag{19}$$

The receiver having a higher packet reception rate is considered as a node having more link correlations which is selected as its cluster head in the next hierarchy level.

In order to find link correlation for more number of next-level nodes, the previous link correlation procedure is extended, since the nodes that are possible to receive the 'hello' messages can be considered for finding the link correlation. The computational complexity is increased, as the link correlation process using two nodes is calculated for all combinations of nodes for finding the highly correlated node. It is overcome by avoiding the nodes having a low packet reception ratio to enter the link correlation calculation using an optimum threshold level. The threshold level is kept at (t-2), such that the nodes having a packet reception greater than this threshold level is selected. The

selected nodes are further restricted by only choosing the higher common packet reception nodes for link correlation calculation. Let $\{r_1, r_2, \ldots r_m\}$ represent the 'm' number of nodes selected for finding link correlation whose packet reception is greater than threshold level. In order to find link correlation for 'm' nodes, the nodes having a high common packet reception from all node combinations are found using Equation 20. Then for those nodes, the link correlation is calculated using Equation 14 to Equation 19. This helps to reduce the complexity involved in finding link correlation for 'm' number of nodes.

$$X_{uv} = \max_{i,j=r_1,r_2,\ldots r_m} \left(X_{ij} \right) \tag{20}$$

3.2 Network coding
Network coding is a technique used for removing redundant data transmission in a network. When packets are moved from a cluster to sink, due to a multi-hop communication among CH, redundant data will be experienced between clusters. In this proposal, an RL [21] network coding is used in all level-3 CHs to remove the redundant data transmission by discarding the previously forwarded packets from transmission. When a node receives the coded packet, it first checks the buffer; if received packet is not independent, it discards the packet. Otherwise, the network coding will code the packets with other packets in the node to generate some linear combinations with the coded packets. The level-3 CH performs RL network coding for the packets in buffer to generate 'b' linear combinations of coded data. The coded packets of random linear coding is given as

$$p_i^{'} = \sum_{i=0}^{b} c_i \times p_i \tag{21}$$

where c_i is the co-efficient over a finite field used for coding, p_i is the original packet and $p_i^{'}$ is the coded packet.

The coded packets are transmitted along with the co-efficient. To retrieve the original data, the sink will perform a decoding process after it receives the coded packet. The decoding is performed using the formula

$$p_i = \sum_{i=0}^{b} c_i \times p_i^{'} \tag{22}$$

3.3 Energy dissipation model
In WSN, the overall energy consumption of the network will depend on the energy consumed by individual node [11]. In each node, the power consumed during processing is negligible when compared to the transmission power [26]. Thus, reducing the redundant bits transmitted leads to an efficient network in terms of energy. The energy

dissipation model of the proposed cluster architecture is discussed here. Figure 4 shows the general Radio Energy Dissipation model [11] of a single sensor node. During transmission, the energy dissipated by the transmitter due to propagation loss is proportional to the distance between a transmitter and receiver d^n. 'n' is path loss exponent given by

$$n = \begin{cases} 2 & \text{for free space} \\ 4 & \text{for multipath interference} \end{cases}$$

In a direct communication between all sensor nodes and sink, the distance between them is large and it includes a multipath interference, so the propagation loss is proportional to d^4 (i.e. $d \geq d_0$). Whereas, in other cases such as multipath routing, the distance between the nodes and the sink is less, thus the propagation loss is proportional to d^2 (i.e. $d < d_0$). Based on those assumptions, the energy consumption for a transmitter and receiver to transmit 'l' bit message [2] is modeled as

$$E_{Tx}(l,d) = \begin{cases} lE_{\text{elec}} + l\varepsilon_{fs}d^2, & d < d_0 \\ lE_{\text{elec}} + l\varepsilon_{mp}d^4, & d \geq d_0 \end{cases} \quad (23)$$

$$E_{Rx}(l,d) = lE_{\text{elec}} \quad (24)$$

where $d_0 = \frac{4\pi h_r h_t}{\lambda} = 87\text{m}$

Let h_t and h_r represent a transmitter and receiver antenna height with a value of 1.5 m. 'λ' represents the signal wavelength having a value of 0.325 m. E_{elec} represents energy for transmitter electronics to transmit a bit and it is the same for the receiver also. ε_{fs} and ε_{mp} are the transmitter amplifier energy based on the amplifier chosen. In the level-k cluster hierarchy, the node at each level is allowed to communicate only with next-level nodes. The level-k CH uses a multi-hop communication for transmitting the data packets towards the sink. In both these cases, the transmission path uses only a less distance (i.e. $d < d_0$). Thus in the proposed cluster algorithm, the path loss exponent for the transmission and transmitter amplifier energy is assumed to be $n = 2$ and ε_{fs}, respectively. In this model, a level-3 cluster hierarchy for a heterogeneous network is assumed. To transmit an

'l' bit message, the energy consumption of a cluster is given by E_{cluster}

$$E_{\text{cluster}} = (n_1/k_1)E_{\text{nonCH}} + (n_2/k_2)E_{\text{CH}_1} + E_{\text{CH}_2} \quad (25)$$

In (25), the energy consumption of level-1 to level-3 nodes is given by, E_{nonCH}, E_{CH_1} and E_{CH_2}, respectively.

Here, n_1 is number of nodes in a cluster, n_2 is number of nodes in a network, k_1 is number of a sub-cluster in the cluster and k_2 is number of a cluster in the network.

Thus as level-1 nodes only transmit its data, the energy consumption for those nodes include the transmitter alone. The level-2 CHs receive the data from level-1 nodes, perform data aggregation and transmit its data to level-3 cluster heads. For level-2 nodes, the energy consumption includes all those operation. E_{DA} gives the energy of data aggregation. The energy consumption of level-3 cluster heads is derived by considering a multi-hop communication between cluster heads and energy saved by network coding. Since the network coding process dissipates negligible energy, it is not included [27]. E_{NC} represents the energy saved by using network coding.

$$E_{\text{nonCH}} = lE_{\text{elec}} + l\varepsilon_{fs}d^2 \quad (26)$$

$$E_{\text{CH}_1} = \left(\frac{n_1}{k_1}-1\right)lE_{\text{elec}} + \frac{n_1}{k_1}lE_{DA} + lE_{\text{elec}} + l\varepsilon_{fs}d^2 \quad (27)$$
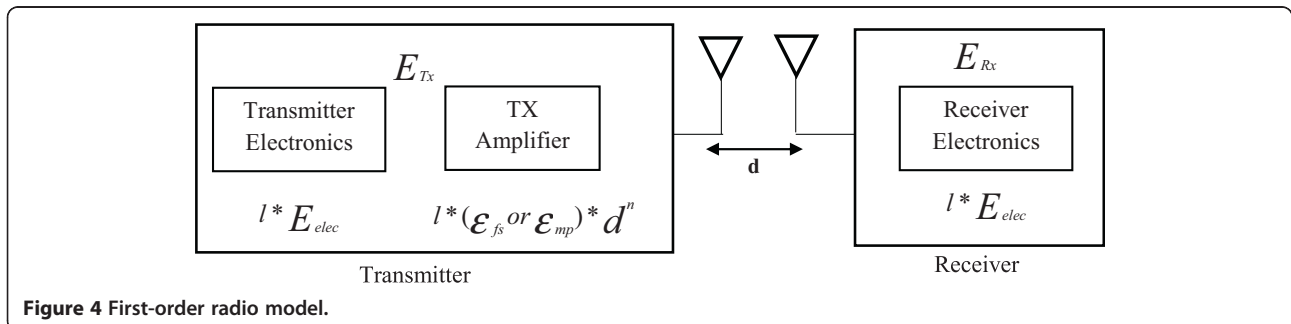
$$E_{\text{CH}_1} = \frac{n_1}{k_1}l(E_{\text{elec}} + E_{DA}) + l\varepsilon_{fs}d^2 \quad (28)$$

$$E_{\text{CH}_2} = \left(\frac{n_2}{k_2}-1\right)lE_{\text{elec}} + P_n\left[2lE_{\text{elec}} + l\varepsilon_{fs}d_m{}^2\right] - E_{\text{NC}}P_n l \quad (29)$$

$$E_{\text{NC}} = 2E_{\text{elec}}\frac{n}{n-1}\frac{D}{d_m}\frac{P_n-1}{P_n} \quad (30)$$

where $d_m = \sqrt[n]{\frac{E_{\text{elec}}}{\varepsilon_{fs}(n-1)}}$

where P_n is the number of packets that are coded into a single packet due to network coding, D is the approximate bottleneck radius and d_m is optimal hop length.

Figure 4 First-order radio model.

Table 1 Parameter of the radio model

Parameter	Value
E_{elec}	50 nJ/bit
ε_{fs}	10 pJ/bit/m^2
ε_{mp}	0.0013 pJ/bit/m^4
E_{DA}(for data aggregation)	5 nJ/bit/signal
Message size	4,000 bits

4 Performance evaluation metrics

For evaluating the performance of WSN, the following metrics are considered. They are network lifetime, message packet delivered to sink and stability for a cluster.

4.1 Simulation results

The performance of the proposed heterogeneous network model is evaluated using MATLAB. In this network model, 100 nodes are deployed in 100×100 m region. For modeling the network, the radio characteristic parameters are given in Table 1. The initial energy distribution for heterogeneous nodes is assumed to be five times (i.e. $a = 4$) of the normal node. The heterogeneous node distribution is given by a factor of $m = 0.3$ in total nodes of a network. The heterogeneous nodes are mostly distributed around the bottleneck zone. Figure 5 shows the simulation environment of the 100-node random network.

The proposed cluster algorithm is compared with SEP, LEACH and DEEC for performance evaluation. To obtain the performance improvement of the proposed algorithm upon the inclusion of a link correlation technique, TDMA

and network coding, it is compared in particularly with DEEC. The LEACH and SEP are level-2 hierarchy clusters in nature where as DEEC is a multi-level hierarchy cluster type. Since the proposed cluster architecture is of a multi-level hierarchy cluster type, here a level-3 hierarchy cluster is adopted for comparison. The LEACH and SEP are extended to a multi-level cluster for effective comparison. The DEEC clustering algorithm works similarly to that of the proposed clustering algorithm. The former algorithm organizes the cluster based on the energy level of nodes; while the latter follows the former algorithm. In addition to that, it also considers link correlation, TDMA and network coding. By comparing both algorithms, the performance improvement achieved due to link correlation and network coding is obtained. In DEEC, a multi-hop communication is established at all levels of cluster hierarchy, such that there will be more redundant data transmission which affects the network lifetime and stability. Further, due to a lack of co-ordinance between the nodes of a cluster, there will be more occurrences of data collision affecting the message delivered to the sink.

The proposed clustering algorithm overcomes those drawbacks by adopting a single-hop communication using link correlation within the clusters along TDMA and network coding at higher level CHs. The single-hop communication using link correlation along with TDMA provides organized cluster architecture with co-ordinance at all levels. This avoids the data collision increasing the message delivery to the sink and increase the network lifetime by removing redundant data transmission. Since the proposed algorithm employs a multi-hop communication between level-k CHs, implementing
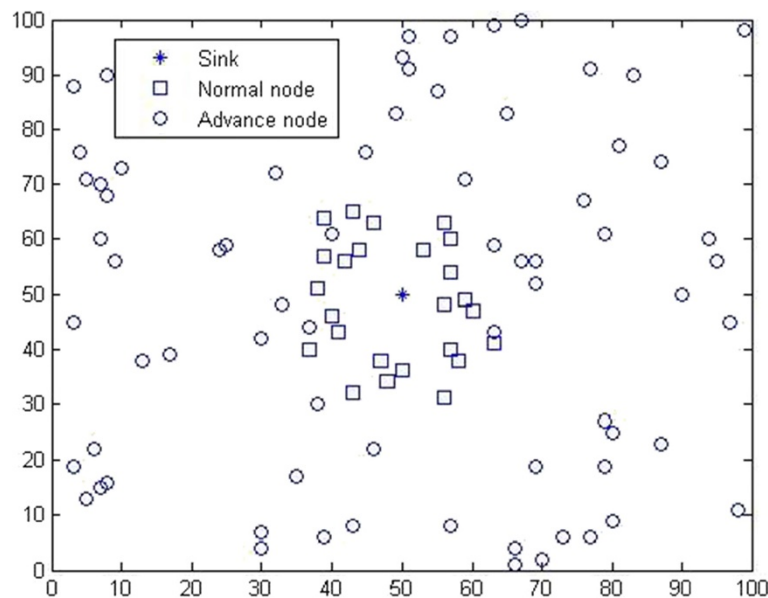
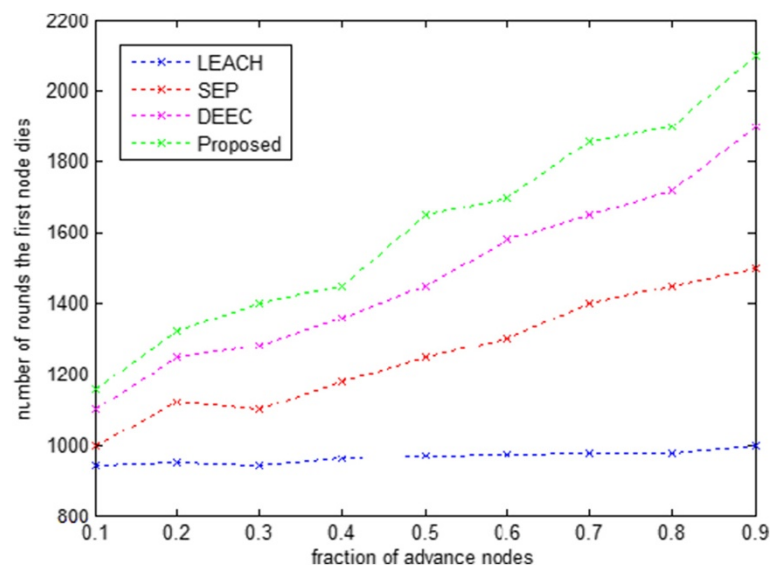**Figure 5 Simulation set-up for a 100-node random network.**
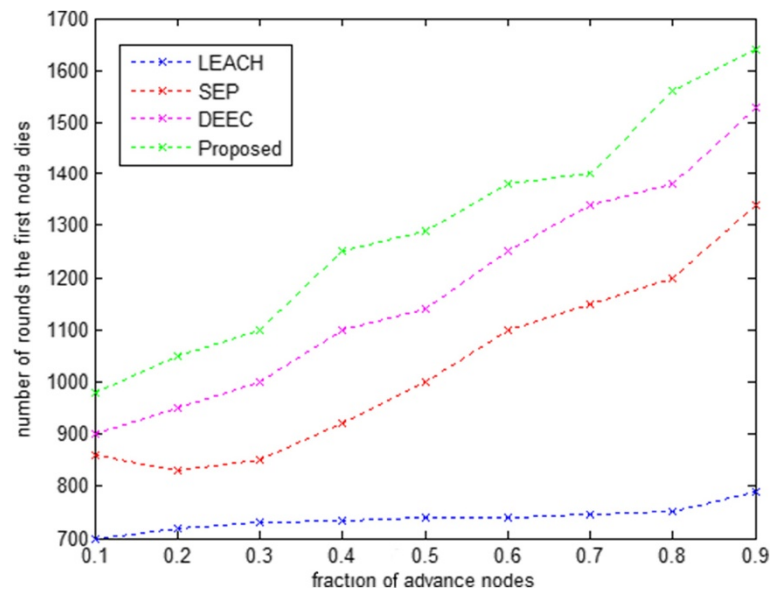
Figure 6 Comparison of stability for proposed cluster architecture over LEECH, SEP and DEEC (*a* =2). Comparison of stability for proposed cluster architecture over LEECH, SEP and DEEC with variation of number of advance nodes (assumption *a* =2).

network coding in it reduces the data traffic and maintain reliability of a network. This helps to increase the network lifetime and stability by reducing the energy depletion of nodes around the bottleneck zone. The detailed performance analysis of the proposed algorithm with DEEC, SEP and LEACH is clearly explained below.

In this model, for most of the cases, the heterogeneous nodes are selected as cluster heads. The energy consumption task such as network coding, multi-hop communication and data reception from all sub-clusters is performed by a level-3 cluster head. Thus the stability of network is increased a lot. The stable region of a proposed cluster architecture is determined by varying the fraction of advance nodes m from (0.1 to 0.9) assuming a =2 and a =4. For both cases, the LEACH does not show any difference.

Figures 6 and 7 show the comparison of a stable region of a proposed cluster architecture with other heterogeneous

Figure 7 Comparison of stability for proposed cluster architecture over LEECH, SEP and DEEC (*a* =4). Comparison of stability for proposed cluster architecture over LEECH, SEP and DEEC with variation of number of advance nodes (assumption *a* =4).
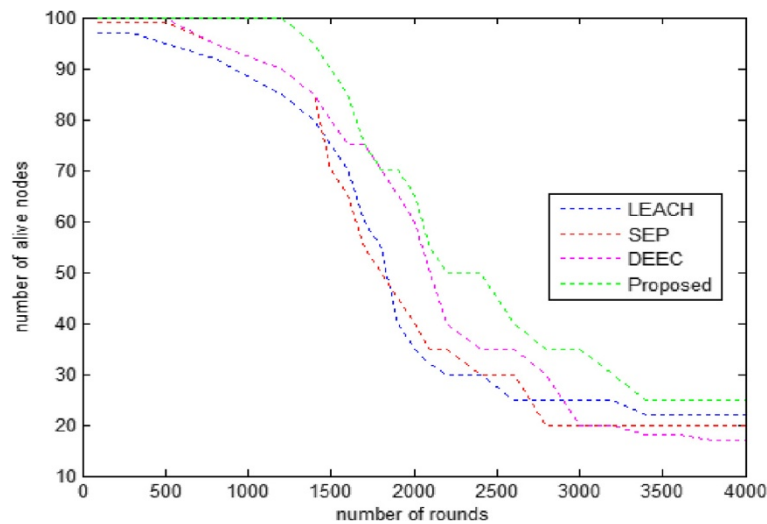
Figure 8 Comparison of network life time for proposed cluster architecture with LEECH, SEP and DEEC.

networks. The stable region of SEP and LEACH has a less number of rounds compared to other heterogeneous network.

The proposed cluster protocol achieves a 26% and 10% increase in number of rounds when the first node dies, compared to that of SEP and DEEC resulting in higher stability for a =4. Comparing to LEACH, the proposed cluster architecture achieves over a 70% increase in stability for both values of a. The proposed cluster protocol achieves a 12% increase in overall network lifetime compared to DEEC. The comparison of network lifetime is shown in Figure 8. Meanwhile, the network coding in level-3 cluster heads decreases the number of bits in a multi-hop communication and increases the reliability of

message delivery to the sink. Figure 9 shows the message delivered to the sink. Comparing to SEP and DEEC, we achieve 59% and 15% increases in message delivery to the sink, respectively.

5 Conclusions

The proposed cluster algorithm is more suitable for the multi-level heterogeneous networks. For efficient comparison, the cluster level-k chosen in this architecture is 3. Compared to LEACH, SEP and DEEC, the proposed cluster architecture is remarkable in message delivery, stability and network lifetime. The proposed cluster architecture is well organized by establishing a single-hop communication within the cluster using link correlation along
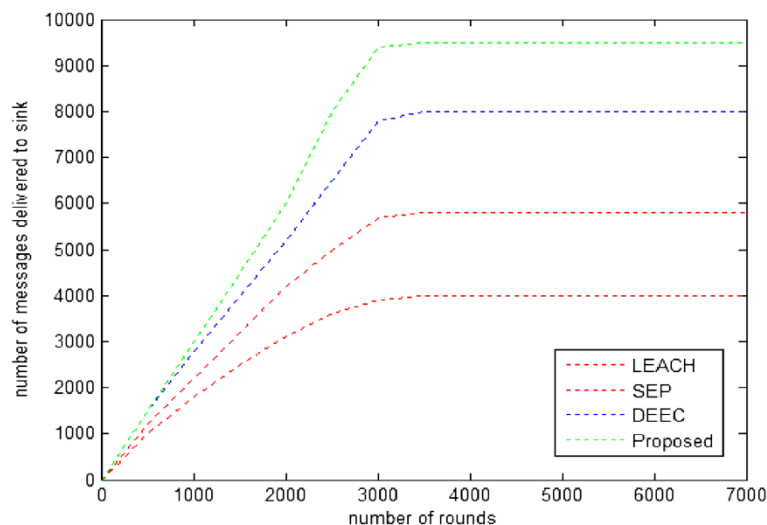
Figure 9 Comparison of message delivered for proposed cluster architecture over LEECH, SEP and DEEC.

with the TDMA time slot. On the other hand, a multi-hop communication among cluster heads is well controlled by network coding. This cluster architecture works well for medium-sized WSN. Since the heterogeneous nodes are mostly chosen as cluster heads and by incorporating the energy consuming task on those nodes increase the energy efficiency of networks.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]RMK College of Engineering and Technology, Tamil Nadu 601206, India. [2]RMD Engineering College, Tamil Nadu 601206, India.

References

1. IF Akyildiz, W Su, Y Sankara Subramaniam, E Cayirci, A survey on sensor networks. IEEE Commun. Mag. **40**(8), 102–114 (2002)
2. LP Clare, GJ Pottie, JR Agre, Self-organizing distributed sensor networks, in *SPIE Conference on Unattended Ground Sensor Technologies and Applications*, 1999, pp. 229–237
3. M Dong, K Yung, W Kaiser, Low power signal processing architectures for network microsensors, in *Proceedings 1997 International Symposium on Low Power Electronics and Design*, 1997, pp. 173–177
4. RP Mann, KR Namuduri, R Pendse, Energy-Aware Routing Protocol for Ad Hoc Wireless Sensor Networks. EURASIP J. Wirel. Commun. Netw. **5**(1900), 635–644 (2005)
5. S Tilak, NB Abu-Ghazaleh, W Heinzelman, A taxonomy of wireless micro-sensor network models. ACM SIGMOBILE Mobile Computing and Communications Review **6**(2), 28–36 (2002)
6. V Mhatre, C Rosenberg, Homogeneous vs. Heterogeneous Clustered Sensor Networks: A Comparative Study, in *Proceedings of 2004 IEEE International Conference on Communications (ICC 2004)*, 2004
7. Q Wang, T Zhang, Bottleneck zone analysis in energy-constrained wireless sensor networks. Commun. Lett., IEEE **13**(6), 423–425 (2009)
8. D Estrin, R Govindan, J Heidemann, S Kumar, Next Century Challenges: Scalable Coordination in Sensor Networks, in *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'99)*, 1999, pp. 263–270
9. S Lee, HS Lee, Analysis of network lifetime in cluster-based sensor networks. Commun. Lett., IEEE **14**(10), 900–902 (2010)
10. P Krishna, NH Vaidya, M Chatterjee, D Pradhan, A cluster-based approach for routing in dynamic networks. Comput. Comm. Rev. **27**(2), 49–65 (1997)
11. WR Heinzelman, A Chandrakasan, H Balakrishnan, Energy-efficient communication protocol for wireless microsensor networks, in *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on. IEEE*, 2000, p. 10
12. S Lindsey, CS Raghavenda, PEGASIS: Power Efficient Gathering in Sensor Information Systems, in *Proceeding of the IEEE Aerospace Conference, Big Sky, Montana*, 2002
13. G Smaragdakis, I Matta, A Bestavros, SEP: A Stable Election Protocol for Clustered Heterogeneous Wireless Sensor Networks, in *Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004)*, 2004
14. L Qing, Q Zhu, M Wang, Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. Comput. Comm. **29**(12), 2230–2237 (2006)
15. D Ganesan, R Govindan, S Shenker, D Estrin, Highly-resilient, energy-efficient multipath routing in wireless sensor networks. MCCR **5**(4), 11–25 (2001)
16. T Zhu, Z Zhong, T He, ZL Zhang, Exploring link correlation for efficient flooding in wireless sensor networks, in *Proc. NSDI*, 2010
17. K Srinivasan, P Dutta, A Tavakoli, P Levis, An empirical study of low power wireless. ACM Trans. Sens. Netw. **6**(2), (2010)
18. B Krishnamachari, D Estrin, S Wicker, The Impact of Data Aggregation in Wireless Sensor Networks, in *Proceedings of IEEE INFOCOM, New York, NY*, 2002, pp. 575–578
19. D Hall, *Mathematical Techniques in Multisensor Data Fusion* (Artech House, Boston, MA, 1992)
20. L Klein, *Sensor and Data Fusion Concepts and Applications* (SPIE Optical Engr Press, WA, 1993)
21. S Chachulski, M Jennings, S Katti, D Katabi, Trading Structure for Randomness in Wireless Opportunistic Routing, in *ACM SIGCOMM*, 2007
22. S-Y Li, RW Yeung, N Cai, Linear network coding. IEEE Trans. Inf. Theory **49**(2), 371–381 (2003)
23. OM Al-Kofahi, AE Kamal, Network coding-based protection of many-to-one wireless flows. IEEE J. Sel. Areas Commun. **27**(5), 797–813 (2009)
24. C Reis, R Mahajan, D Wetherall, J Zahorjan, Measurement Based Models of Delivery and Interference. ACM SIGCOMM Comp. Com. Rev. **36**(4), 51–62 (2006)
25. G Zhou, T He, JA Stankovic, Impact of Radio Irregularity on Wireless Sensor Networks, in *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, 2004, pp. 125–138
26. S Halawani, A Waheed Khan, Sensors lifetime enhancement techniques in wireless sensor networks. J. Comput. **2**(5), 2151–9617 (2010)
27. R Ranjan Rout, SK Ghosh, Enhancement of lifetime using duty cycle and network coding in wireless sensor networks. IEEE Trans. Wireless Comm. **12**(2), 656–667 (2013)