Journal of Network and Computer Applications 52 (2015) 116–128

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Two-tier particle swarm optimization protocol for clustering and routing in wireless sensor network

Riham S.Y. Elhabyan *, Mustapha C.E. Yagoub

School of Electrical Engineering and Computer Science, University of Ottawa, Canada

ARTICLE INFO

Article history:
Received 9 April 2014
Received in revised form
3 October 2014
Accepted 4 February 2015
Available online 14 March 2015

Keywords:
Clustering
Multi-hop
WSN
PSO
RSSI
CC2420

ABSTRACT

Many cluster-based routing techniques for Wireless Sensor Networks (WSNs) have been proposed in the literature. However, most of the proposed protocols emphasized on the Cluster Head (CH) selection ignoring how the CHs will send the aggregated data back to the Base Station (BS). Furthermore, they tend to use non-realistic parameters and assumptions. Such examples include the use of infinite transmission range and location awareness. They also used an energy model that is fundamentally flawed for modelling radio power consumption in sensor networks. In this paper, two Linear Programming (LP) formulations to the problems of clustering and routing are presented followed by two proposed algorithms for the same based on Particle Swarm Optimization (PSO). The clustering algorithm finds the optimal set of CHs that maximize the energy efficiency, cluster quality and network coverage. The routing algorithm is developed with a novel particle encoding scheme and fitness function to find the optimal routing tree that connects these CHs to the BS. These two algorithms are then combined into a two-tier protocol to provide a complete and practical clustering model. The effect of using a realistic network and energy consumption model in cluster-based communication for WSN will be investigated. Extensive simulations on 50 homogeneous and heterogeneous WSN models are evaluated and compared against well-known cluster-based sensor network protocols. The results demonstrate that the proposed protocol performs better than such protocols in terms of various performance metrics such as scalability, Packet Delivery Rate (PDR) at the CHs and delivery of total data packets to the BS.

1. Introduction

1.1. Background

Wireless Sensor Network (WSN) has emerged as a powerful technological platform with tremendous and novel applications. It has become an important technology in realizing many applications including both simple phenomena monitoring applications and heavy-duty data streaming applications such as military operations, environment monitoring and surveillance systems.

A WSN usually consists of tens to thousands of sensor nodes that communicate through wireless channels for information sharing and cooperative processing (Yu et al., 2006). Usually, the nodes are statically deployed over vast areas. However, they can also be mobile and capable of interacting with the environment.

WSN nodes also can sense the environment, communicate with neighboring nodes, and in many cases perform basic computations on the data being collected (Zungeru et al., 2012; Akkaya and Younis, 2005). These features made WSN an excellent choice for many applications like environmental monitoring, military surveillance, search and rescue, in buildings for infrastructure health monitoring, or even in bodies for patient monitoring (Yu et al., 2006).

There are some factors that affect designing and operating WSN. These factors include energy efficiency and awareness, connection maintenance, minimum resource usage limitation, low latency, network coverage and load balancing in terms of energy used by sensor nodes. Due to these unique inherent characteristics it is a challenging task to select or propose a new routing or communication algorithm for a specific WSN application (Dwivedi and Vyas, 2010).

Using clustering techniques in WSN can help solving some of those concerns, by organizing the network nodes into smaller clusters and elect a cluster head (CH). Sensor nodes in each cluster transmit their data to their respective CH and CH aggregates data and forward them to a central base station (BS) (Abbasi and Younis, 2007). The fact that only the CH is transmitting information out of the cluster helps avoid collisions between the sensors inside the cluster because they do not have to share the communication channel with the nodes in other clusters (Arboleda and Nasser, 2006).

* Corresponding author.
E-mail address: relha105@uottawa.ca (R.S.Y. Elhabyan).

http://dx.doi.org/10.1016/j.jnca.2015.02.004
1084-8045/

Once the WSN has been divided into clusters, the communication between nodes can be either intra-cluster or inter-cluster. Intra-cluster communication comprises the data exchanges between the member nodes and their respective CH. Inter-cluster communication includes transmission of the data between the CHs or between the CH and the BS.

The process by which data are forwarded efficiently between the CHs and the BS (inter-cluster communication) is an important aspect and essential feature of WSN. A simple method to accomplish this task is for each CH to exchange data directly with the BS (a single hop based approach), or allowing intermediate nodes to participate in forwarding data packets between the CH and the BS (a multihop based approach) (Zungeru et al., 2012). However, in a WSN, individual nodes have limited communication range and form an ad hoc network over a shared wireless medium. Furthermore, the BS is usually located far away from the sensing area and is often not directly reachable to all nodes due to limited communication range and signal propagation problems. A more realistic approach is to use a multihop inter-cluster communication model. For a more reliable data communication, Both data and control packets need to be routed using a multihop communication model (Saleem et al., 2011).

The objective of clustering is to search among a group of sensor nodes to find a set of nodes that can act as cluster-heads. For a given network topology, it is difficult to find the optimal set of CH nodes. For N sensor nodes, there are 2^N-1 different combination of solutions, where in each solution, a sensor node is either elected as CH or non-CH. This has been proved to be a Non-deterministic Polynomial (NP)-hard optimization problem (Agarwal and Procopiuc, 2002).

The basic function of a routing algorithm is to select a route, from the set of available routes, that is most efficient based on some specific criteria. Once the optimal set of CHs is elected in the clustering phase, the next step is to find the optimal routing tree from the CHs to the BS while minimizing the total cost of that tree. Routing is at its most basic level an optimization problem. It also has been known to be NP-hard problem (Dorigo et al., 2006). Therefore, polynomial-time algorithms are impossible to use due to their high computational complexity in real-time communications systems.

Solutions to NP-hard problems involve searches through vast spaces of possible solutions. Swarm intelligence approaches have been applied successfully to a variety of such problems.

Particle swarm optimization (PSO) is a swarm intelligence based optimization method. PSO has many advantages over other alternatives optimization techniques like Genetic Algorithms (GA) which has very high processing demands (Guo and Zhang, 2014). PSO advantages include ease of implementation on hardware or software, high-quality solutions because of its ability to escape from local optima and quick convergence (Kulkarni and Venayagamoorthy, 2011; del Valle et al., 2008). Because of its effectiveness in solving NP problems, PSO has been adopted to optimize the CH election by several centralized clustering protocols. Clustering is a repeated process; therefore, the simpler the optimization algorithm, the better the network efficiency is. This is another reason why PSO is a popular choice for WSN clustering.

1.2. Authors' contributions

In this paper, firstly, we present two Linear Programming (LP) formulations for the clustering and routing problems respectively. Then, two PSO-based protocols for the same problems are proposed.

The PSO-based clustering protocol solves the above CH selection problem by electing CHs in such way that the formed clusters maximize the energy efficiency, network coverage and data transmission reliability of the network.

Then, the PSO-based routing protocol finds the optimal routing tree that connects the elected CHs to the BS. For routing, the particles are cleverly encoded to produce complete routing tree solution. A different fitness function is used to build the trade-off between the energy efficiency and link quality of the constructed tree.

Furthermore, we develop the protocols under realistic network settings. No assumptions were made about the nodes location awareness or transmission range capabilities. The protocols were also tested using a realistic energy consumption model that is based on the characteristics of the Chipcon CC2420 radio transceiver data sheet. Extensive simulations on 90 homogeneous and heterogeneous WSN models are evaluated and compared against 7 existing protocols using several performance metrics including average energy consumption, Packet Delivery Rate (PDR), throughput, network coverage and latency. Our main contributions can be summarized as follows:

- Two LP formulations to the problems of clustering and routing respectively.
- PSO-based clustering protocol with a trade-off between energy efficiency, network coverage and data transmission reliability.
- PSO-based routing protocol with a novel particle encoding scheme for complete routing tree solution and derivation of efficient multi-objective fitness function.
- Investigate the result of using realistic network settings. No assumptions were made about location awareness.
- Investigate the effect of using a realistic energy consumption model in cluster-based communication for WSN.
- Simulation of the proposed protocol to demonstrate its performance against some of the existing protocols in both homogeneous and heterogeneous WSNs.

The remainder of this paper is organized as follows. Section 2 reviews the related work on clustering protocols and the associated drawbacks. Section 3 provides an overview of particle swarm optimization. The system model is presented in Section 4. In Section 5, we present our LP formulations for the clustering and routing problems. Section 6 gives a detailed description of the proposed protocol. In Section 7, we present the experimental results. Finally, Section 8 concludes the work and highlights a few future directions.

2. Related work

Clustering techniques have been studied extensively to improve the performance of WSN (Tyagi and Kumar, 2013; Younis et al., 2006; Abbasi and Younis, 2007). We present the review of such works based on heuristic and metaheuristic approaches.

2.1. Heuristic approaches

Low energy adaptive clustering hierarchy (LEACH) (Heinzelman et al., 2000, 2002) is one of the most popular distributed cluster-based routing algorithms in WSN that has been proven to be an effective approach to prolong the network lifetime. Each node uses a stochastic algorithm at each round to determine whether it will become a CH in this round. Nodes that have been CHs cannot become CHs again for P rounds, where P is the desired percentage of CHs. Therefore, each node has a $1/P$ probability of becoming a CH in each round. The CHs are selected without considering the residual energy or the other properties of the sensor nodes. This random mechanism of selecting the CHs does not guarantee even distribution of clusters over the network (Arboleda and Nasser, 2006).

Hybrid energy-efficient distributed Clustering (HEED) (Younis and Fahmy, 2004) is another distributed clustering protocol that is an extension of LEACH. Cluster formation is achieved with an iterative approach. CHs selection in this protocol is primarily based on the residual energy of each node. To increase energy efficiency and further prolong network lifetime, a secondary clustering parameter considers intra-cluster "communication cost" is introduced which can be a function of neighbor proximity or cluster density. The main objectives of HEED are to distribute energy consumption to prolong network lifetime, minimize energy during the CH selection phase, and reduce the control overhead of the network. The improvement over LEACH is that HEED can evenly distribute the cluster heads in the sensing area by local competition.

Energy-efficient clustering scheme (EECS) (Ye et al., 2005) is a distributed non-iterative clustering protocol. EECS extends LEACH algorithm by dynamic sizing of clusters based on cluster distance from the BS. Unlike LEACH, the CH is elected by localized competition and its no iteration property makes it differ from HEED. This competition involves candidates broadcasting their residual energy to neighboring candidates. If a given node does not find a node with more residual energy, it becomes a CH. However, the EECS protocol does not consider the structural characteristics of network topology and thus CHs are elected on the basis of residual energy. Furthermore, the set of candidate nodes in the competition is selected randomly before the competition, this may result in non-optimal CH selection.

While the above schemes assume homogeneous WSNs, several other schemes were proposed to deal with CHs selection in heterogeneous WSNs where energy heterogeneity is considered.

Kumar et al. (2009) have proposed an energy efficient heterogeneous clustered scheme for WSNs (EEHC) to study the impact of heterogeneity of nodes in terms of their energy in clustered networks. They assumed the case where a percentage of the population of sensor nodes is equipped with more energy resources than the normal sensor nodes in the network. Three types of sensor nodes equipped with different energy levels were used. Nodes under first level are known as normal nodes, second level nodes are advanced node and third level nodes are super nodes. They showed how the election process of cluster heads should be adapted appropriately to deal with heterogeneous nodes. The election probabilities of CHs are weighted by the initial energy of a node relative to that of other nodes in the network.

Like EEHC, Enhanced heterogeneous LEACH protocol for lifetime enhancement of wireless SNs (EHE-LEACH) (Tyagi et al., 2013) deals with CH election in heterogeneous networks. There are two main differences. Firstly, the authors assume two level of node heterogeneity, nodes under first level are known as normal nodes, second levels are advanced node. Secondly, a fixed distance based threshold is used by each node to choose between direct communication with the BS or cluster based communication. Sensor nodes that are near the BS send their data directly to the BS and those which are far away from the BS use cluster based communication.

Kumar (2014) has also proposed two distributed protocols, single-hop energy-efficient clustering protocol (S-EECP) and multi-hop energy-efficient clustering protocol (M-EECP) were proposed by to also deal with node heterogeneity in WSN. S-EECP uses the same weighted election probabilities concept as EEHC and the same three levels of node heterogeneity. However, they take into account the residual battery energy of nodes in calculating the weighted election probabilities of each node. They observed that in single-hop communication where data packets are directly transmitted to the BS without any relay nodes, the nodes located far away from the BS have higher energy consumption because of long range transmission, and these nodes may die out first. They solved this problem in M-EECP by using multi-hop communication to the BS. M-EECP uses a greedy approach to solve the single source shortest problem to find the shortest path from each CH to the BS. Although S-EECP outperforms EEHC in terms of energy efficiency, the assumption that each node knows all other nodes' energy level is unrealistic and impossible to obtain in such distributed setting. Furthermore, M-EECP suffers from the same problem as S-EECP and assumes that each node knows all other nodes locations.

2.2. Metaheuristic approaches

LEACH-centralized (LEACH-C) (Heinzelman et al., 2002) is a centralized version of LEACH. Unlike LEACH, where nodes self-configure themselves into clusters, LEACH-C uses the BS for cluster formation. Initially, each node sends its information (location and energy level) to the BS, which will use this information and employ a Simulated Annealing (SA) approach to find a predetermined number of CHs and configure the network into clusters. The clusters are chosen to minimize the energy required for non-CH nodes to transmit their data to their respective CHs. LEACH-C yields better results than LEACH in terms of packet delivery rate and energy consumption.

Energy balanced unequal clustering protocol (EBUC) (Jiang et al., 2010) is a centralized clustering protocol in which the authors tried to solve the hot-spot problem by creating unequal clusters using a centralized particle swarm optimization (PSO) algorithm at the BS. The clusters are created such that the ones near the BS have fewer number of nodes, and so it increases the number of clusters around the BS. For the Inter-cluster communication, the CH uses a greedy algorithm to choose a relay node based on the node's residual energy and distance to the BS.

An energy-aware clustering for WSNs using PSO algorithm (PSO-C) is a centralized clustering protocol which is implemented at the BS, was proposed in Latiff et al. (2007). It considers both energy available to nodes and physical distances between the nodes and their CHs. This protocol defines an objective function which tries to minimize both the maximum average Euclidean distance of nodes to their associated CHs and the ratio of total initial energy of all nodes to the total energy of the CH candidates. It also ensures that only nodes with sufficient energy are selected as CHs. PSO-C outperforms both LEACH and LEACH-C in terms of the network lifetime and the throughput. The authors in Abdul Latiff et al. (2007) showed that PSO-C outperforms GA and K-means-based clustering protocols as well in terms of convergence time, network lifetime and data delivery.

A genetic algorithm (GA)-based protocol that was proposed by Rahmanian et al. (2011) attempts to find appropriate CHs to minimize the total network distance. The objective function is defined as the minimization of the total distance from cluster members to their respective CHs in addition to the distance from the CHs to the BS. As PSO-C, it also ensures that only nodes with sufficient energy are selected as CHs.

An energy-aware evolutionary routing protocol (EAERP) was proposed by Khalil and Attea (2011). A centralized single-hop clustering protocol is presented where the BS runs an evolutionary-based protocol to optimize the CH election for cluster formation. The objective function is defined as the minimization of the total dissipated energy in the network, measured as the sum of the total energy dissipated from the non-CHs to send data signals to their CHs, and the total energy spent by CH nodes to aggregate the data signals and send the aggregated signals to the BS. The protocol use the energy consumption model defined by Heinzelman et al. (2002) to compute the energy dissipated during the process of data transmission and reception.

Kuila et al. (2013) have proposed a GA-based protocol to solve the problem of balancing the load of the CHs. The protocol forms clusters in such a way that the maximum load of each CH is minimized. In this protocol, the CHs are determined priori and the objective of the protocol is to find the optimal assignments of non-CHs nodes to CHs to form balanced clusters. The objective function is defined as minimizing the standard deviation of the CH load which gives even distribution of the load per cluster.

In addition to the previously mentioned problems and up to our best knowledge, all the clustering protocols that were proposed so far use the energy consumption model suggested in Heinzelman et al. (2002). This energy model is fundamentally flawed for modelling radio power consumption in sensor networks. It ignores listening energy consumption, which is known to be the largest contributor to expended energy in WSN. Moreover, most of the location-aware or link quality-based clustering protocols proposed assume that each node is equipped with self-locating hardware such as a GPS. Though this is a simple and effective solution, the resulting cost renders such a solution inefficient and unrealistic (Molina and Alba, 2011). Furthermore, several studies have shown that link quality in WSN is not correlated with distance (Srinivasan et al., 2006, 2010; Baccour et al., 2012; Srinivasan and Levis, 2006).

Table 1 summarizes the main differences between our proposed protocol and relevant related work.

3. Overview of particle swarm optimization

Particle swarm optimization (PSO) is a population based stochastic optimization technique developed by Kennedy and Eberhart (1995) and was inspired by social behavior of bird flocking or fish schooling. Until now, three successive standard PSO versions have been available namely SPSO 2006, 2007, and 2011. They all have the same principles. However, they differ slightly in their formulae.

In this paper, we use the latest standard PSO (SPSO-2011). It has been proved that SPSO-2011 has an outstanding performance and it is able to quickly converge towards the region of the global optimum (Zambrano-Bigiarini et al., 2013).

The basic PSO comprises a swarm of S potential solutions, referred to as particles, which fly through a D-dimensional problem space in search of the global optimum position that produces the best fitness of an objective function.

Initially, each particle i is randomly assigned a position x_{id} and a velocity $v_{id}(i=1,2,\ldots,S)$ where $d=(1,2,\ldots,D)$. Each particle keeps track of its personal best position $pbest_i$ and the global best position of the whole swarm *gbest*. The original PSO algorithm is shown below:

Algorithm 1. PSO algorithm.

```
for each particle do
  initialize particle
end for
while target fitness or maximum epoch is not attained do
  for each particle do
    calculate fitness
    if current fitness value better than (pbest) then
      pbest=current fitness
  end for
  set gbest to the best one among all pbest
  for each particle do
    update velocity
    update position
  end for
end while
```

3.1. Original PSO version

After finding the two best values, particle i then updates both its velocity and position iteratively using Eq. (1a) and (1b) respectively.

$$v_{id}(t+1)=w\times v_{id}(t) + c_1\times r_1\times(\mathrm{pbest}_i(t)-x_{id}(t))+c_2\times r_2\times(\mathrm{gbest}(t)-x_{id}(t)) \tag{1a}$$

$$x_{id}(t+1)=x_{id}(t)+v_{id}(t+1) \tag{1b}$$

r_1 and r_2 are random variables between $[0,1]$. c_1 and c_2 are the learning factors. w is a weight factor that controls the velocity of the particle.

3.2. Standard PSO 2011

In this standard, after finding the two best values, particle i then updates both its velocity and position iteratively using Eq. (2a) and (2b) respectively.

$$v_{id}(t+1)=w\times v_{id}(t)+x'_{id}(t)-x_{id}(t) \tag{2a}$$

$$x_{id}(t+1)=w\times v_{id}(t+1)+x'_{id}(t) \tag{2b}$$

Table 1
Main differences between our proposed protocol and relevant related work.

Clustering protocol	Clustering method	Clustering approach	Location awareness	Radio model	Network type	Clustering objective
LEACH	Distributed	Prob./Random	No	First order	Homogeneous	Rotate CH role
HEED	Distributed	Prob./Energy	No	First order	Homogeneous	Energy efficiency
EECS	Distributed	Prob./Energy	Yes	First order	Homogeneous	Load balancing
EEHC	Distributed	Prob./Energy	No	First order	Heterogeneous	Energy efficiency
EHE-LEACH	Distributed	Prob./Energy	No	First order	Heterogeneous	Energy efficiency
S-EEP	Distributed	Prob./Energy	No	First order	Heterogeneous	Energy efficiency
M-EEP	Distributed	Prob./Energy	Yes	First order	Heterogeneous	Energy efficiency
LEACH-C	Centralized	SA	Yes	First order	Homogeneous	Energy efficiency
EBUC	Centralized	PSO	Yes	First order	Homogeneous	Energy efficiency
PSO-C	Centralized	PSO	Yes	First order	Homogeneous	Energy efficiency
Rahmanian et al. (2011)	Centralized	GA	Yes	First order	Homogeneous	Energy efficiency
Kuila et al. (2013)	Centralized	GA	Yes	First order	Homogeneous	Load balancing
Khalil and Attea (2011)	Centralized	EA	Yes	First order	Homogeneous Heterogeneous	Energy efficiency
Proposed protocol	Centralized	PSO	No	CC2420	Homogeneous Heterogeneous	Energy efficiency Link quality Network coverage

$x'_{id}(t)$ is a random point in the hypersphere of center G_i and of radius $\|G_i - x_i\|$. G_i is the center of gravity of three points: the current position, a point a bit beyond the best previous position and a point a bit beyond the best previous position in the neighborhood. Detailed description of how G_i is calculated and of the all three PSO standards can be found at Clerc (2012).

4. The system model

4.1. The WSN model

For our model, we consider a two-tiered WSN with N sensor nodes, K cluster heads and one base station. Each sensor node has a unique ID and the best station ID is 0. In the cluster formation process, each sensor node belongs to only one cluster and each cluster head node acts as the cluster head of exactly one cluster.

We assume that all nodes are stationary after deployment and the locations of both the sensor nodes and the cluster heads are unknown. We consider different network densities in our experiments. Furthermore, we consider both homogeneous and heterogeneous network settings in our experiments.

4.2. The energy consumption model

In the proposed approach, a realistic energy consumption model which is based on the characteristics of the Chipcon CC2420 radio transceiver data sheet (Texas Instruments, 2013) is used. The total energy consumed by node i, E_i, is calculated as follows: (Barberis et al., 2007):

$$E_i = \sum_{state\ j} P_{state\ j} \times t_{state\ j} + \sum E_{transitions} \tag{3}$$

The index *state j* refers to the energy states of the sensor: sleep, reception, or transmission. $P_{state\ j}$ is the power consumed in each *state j*, and $t_{state\ j}$ is the time spent in the corresponding state. Moreover, the energy spent in transitions between states, $E_{transitions}$, is also added to the node's total energy consumption. The different values of $P_{state\ j}$ and $E_{transitions}$ can be found in Texas Instruments (2013).

5. LP formulations for the clustering and routing problem

In this section, we present our LP formulations for the clustering and routing problem in WSN. We have used weight sum approach (WSA) for the construction of the multi-objective fitness function in both the clustering and routing problem. This approach is computationally efficient and is straightforward to implement (Konak et al., 2006) which makes it suitable to apply in WSN.

The following subsections describe the network model used in this paper, followed by the notations used and the proposed formulations for the clustering and routing problem.

5.1. Notations used

In our formulations, we are given the following data as input:

- N: total number of sensor nodes.
- K: total number of cluster heads ($K = 5\% \times N$).
- initial$E(\mathrm{CH}_{p,k})$ is the initial energy of CH number k in particle p.
- $E(\mathrm{CH}_{p,k})$ is the remaining energy of CH number k in particle p.
- $\mathrm{RSSI}(n_i, \mathrm{CH}_{p,k})$ is the RSSI value for the link from node n_i to CH $\mathrm{CH}_{p,k}$.
- minRSSI is the worst RSSI value among all communicating pairs and is set to -97.
- $|C_{p,k}|$ is the number of members in cluster k of particle p.
- R is the total number of relay nodes in the routing tree.
- C is the total number of cluster head candidates that act as relay nodes.
- $\mathrm{RN}p, r$ is relay node number r in the route generated from particle p.
- R is the number of routes in the routing tree ($R = K$).
- rn_i is relay node number i.
- $\mathrm{wc}_1, \mathrm{wc}_2$ and wc_3 are weight coefficients that specify the contribution of each sub-objective in the main clustering objective function.
- $\mathrm{wr}_1, \mathrm{wr}_2$ and wr_3 are weight coefficients that specify the contribution of each sub-objective in the main routing objective function.

5.2. LP formulation for the clustering problem

The best CHs are selected such that they minimize the cost of the objective function. The goal of the function is to optimize the combined effect of the following properties:

5.2.1. Energy efficiency

The residual energy of a sensor node could be a criterion for selecting the best CHs since a node with a better battery life is a better candidate for the cluster management and the data aggregation. In addition to that, the consumed energy is distributed among all the sensor node. The BS uses the following function to calculate the fitness of particle p in terms of energy efficiency:

$$\mathrm{EE}_p = \sum_{k=1}^{K} \frac{\text{initial } E(\mathrm{CH}_{p,k})}{E(\mathrm{CH}_{p,k})} \tag{4}$$

5.2.2. Cluster quality

The aim of this sub-objective is to maximize the link quality between the cluster members and their respective CHs in order to maximize the Packet Reception Ratio (PRR). One fundamental indicator of link quality is RSSI, which is the strength of the received RF signal. RSSI is a register in the CC2420 transceiver that is used to measure the receiver-side link quality.

Several studies proved that RSSI can provide a quick and accurate estimate of whether a link is of very good quality (Baccour et al., 2012; Srinivasan and Levis, 2006; Srinivasan et al., 2006, 2010). In Srinivasan et al. (2006), the authors conducted empirical measurements of the packet delivery performance of various sensor platforms. They found that there is a strong correlation between RSSI and Packet Reception Ratio (PRR). Furthermore, they proved that if RSSI of a link is of -87 dBm or stronger, it is almost but not completely set to have a PRR $\geq 99\%$. Below this value, a shift in the RSSI as small as 2 dBm can change a good link to a bad one and vice versa, which means that the link is in the transitional or disconnected region (Srinivasan et al., 2010).

The link quality between cluster member n_i and CH number k in particle p can be calculated using:

$$\mathrm{LQ}_{(n_i, \mathrm{CH}_{p,k})} = \frac{\mathrm{RSSI}(n_i, \mathrm{CH}_{p,k})}{\mathrm{minRSSI}} \tag{5}$$

The higher the value of LQ, the worse is the link quality. To maximize the cluster quality in terms of link quality we need to minimize the worst cluster quality. Hence, the following sub-objective needs to be minimized:

$$CQ_p = \max_{k=1,2,\ldots,K} \frac{\sum_{\forall n_i \in C_{p_j,k}} LQ_{(n_i, CH_{p,k})}}{|C_{p,k}|} \tag{6}$$

5.2.3. Network coverage

Clustering allows sensors to efficiently coordinate their local interactions in order to achieve global goals such as scalability and more efficient resource utilization (Tubaishat and Madria, 2003). Generally, scalability refers to how well the capacity of a system to do useful work increases as the size of the system increases (Lee et al., 1998).

In order to increase the scalability of our protocol, we should increase the network coverage i.e. the number of the clustered nodes. To achieve that, the protocol tries to minimize the number of un-clustered nodes and maximize the number of clustered nodes. This can be realized by minimizing the following sub-objective:

$$NC_p = \frac{N - \sum_{k=1}^{K} |C_{p,k}|}{\sum_{k=1}^{K} |C_{p,k}|} \tag{7}$$

After calculating the previous sub-objectives, the final objective function that needs to be minimized is

$$C_p = wc_1 \times EE_p + wc_2 \times CQ_p + wc_3 \times NC_p \tag{8}$$

5.3. LP formulation for the routing problem

The optimal routing tree is selected such that it minimizes the cost of the objective function. The goal of the function is to optimize the combined effect of the following properties:

5.3.1. Energy efficiency

To achieve an energy efficient routing tree, two sub-objectives need to be met:

(1) *Save energy*: fewer sensor nodes need to be active during each round. To achieve that, the protocol needs to minimize the number of relay nodes and favor CHs as better candidates to act as relay nodes. Minimizing the following function will achieve that objective:

$$EE(a)_p = \frac{R}{C} \tag{9}$$

(2) *Balance energy consumption*: a relay node with a higher level of energy is a better candidate to include in the routing tree. The following function is used to balance the energy consumption among all the network nodes in terms of routing:##

$$EE(b)_p = \frac{\sum_{i=1}^{N} E(n_i)}{\sum_{r=1}^{R} E(RN_{p,r})} \tag{10}$$

5.3.2. Link quality

To maximize the PRR, the protocol needs to maximize the link quality between the relay nodes in the routing tree. The following function minimizes the worst link quality among all the branches in the routing tree:

$$LQ_p = \max_{b=1,2,\ldots,R} \sum_{\forall rn_i \in r} \frac{RSSI(rn_i, rn_{i+1})}{minRSSI} \tag{11}$$

After calculating the previous sub-objectives, the final objective function that needs to be minimized is

$$R_p = wr_1 \times EE(a)_p + wr_2 \times EE(b)_p + wr_3 \times LQ_p \tag{12}$$

6. The proposed protocol

In this paper, a centralized two tier PSO protocol is proposed to solve the problem of clustering and routing in WSN. The protocol is named TPSO-CR, from the initials of the words Two tier Particle Swarm Optimization for Clustering and Routing protocol.

In TPSO-CR, the network operating time is divided into rounds. Each round consists of two phases, the set-up phase and the steady-state phase. In the set-up phase, the network is configured. The BS will choose the best set of CHs and relay nodes. The set-up phase consists of the following steps:

1. *Neighbor discovery*: in this step, each sensor node in the network broadcasts a HELLO packet that includes its ID. A sensor node that receives this HELLO packet will update its neighbor table with the ID included in the packet along with the Received Signal Strength Indicator (RSSI) value in the received packet.
2. *Control data broadcasting*: TPSO-CR uses flooding method to transfer the control data to the BS. After the neighbor discovery ends by all the sensor nodes, each node broadcast the following data about itself: ID, residual energy and it neighbor table data. A node that receives this packet will rebroadcast it till it reaches the BS.
3. *Network configuration*: after a reasonable time and when the BS receives all the control packets from the network nodes, the BS starts configuring the network. For that, the BS uses a two tier PSO algorithm. The first tier is responsible for finding the optimal set of CHs and their associated cluster members. The second tier finds the optimal routes from those CHs to the BS. The clustering and routing algorithms are explained in detail in Section 6.1 and 6.2 respectively.
4. *Configuration broadcasting*: after the BS finishes the network configuration, the BS uses flooding again to transfer the configuration to all the nodes. It broadcasts a packet containing that configuration. Each node that receives that packet will modify its status to either a CH, a cluster member or a relay node. A cluster member will update its respective CH and TDMA schedule. A relay node will update its next hop to the BS.

In the steady-state phase, each non-CH node uses its TDMA schedule to transmit its data to its respective CH. When a CH receives this data, it uses its next relay node to forward the data to the BS. When a non-CH node finishes its data transmission slot, it enters the sleep state to save its energy.

The proposed TPSO-CR protocol executed at an arbitrary node u is shown below.

Algorithm 2. TPSO-CR algorithm.

```
procedure STATRUP()
  setTimer(START−ROUND, 0.0);
end procedure
procedure TIMERFIREDCALBACK (index)
  switch index do
    case START−ROUND :
      double timer = uniform(0.0, r);
      setTimer(FIND−NBRS, timer);
      setTimer(BROADCAST−INFO, r);
      if isBS setTimer(RUN−PSO, n);
      else setTimer(RUN−STEADY−PHASE, m);
      roundNumber++;
      setTimer(START−ROUND, roundLength);
    case FIND−NBRS :
      broadcast (ID);
    case BROADCAST−INFO :
      broadcast (ID, residual Energy, Neighbors'
     IDs and their RSSI);
    case RUN−PSO :
      optimalCHs=runFirstPSO (NetworkInfo);        ▹ run first tier
      optimalRoutingTree=runSecondPSO (optimalCHs,NetworkInfo); ▹ run second tier
      broadcast(configuration = optimalCHs+optimalRoutingTree);
    case RUN−STEADY−PHASE :
      if(!isCH || !isCM || !isRelayNode) then setStateSleep();
      if (isCH) then
        clusterLength = clusterMembers.size();
        setTimer(START−SLOT, clusterLength × slotLength);
      else
        if (!isRelayNode) then setStateSleep()
        setTimer(START−SLOT, myTDMATurn × slotLength);
    case START−SLOT :
      setTimer(START−SLOT, clusterLength × slotLength);
      if (isCH) then
        aggregatePackets();        ▹ aggregate packets
        processBufferedPackets();      ▹ send packets to next hop
      else                         ▹ a cluster member
        processBufferedPackets(); ▹ send packets to CH
        setTimer(END−SLOT, slotLength); ▹ go to sleep mode at end of slot
    case END−SLOT :
      if(!isCH || !isCM || !isRelayNode) then setStateSleep();
end procedure
```

6.1. The clustering algorithm

Based on the information the BS received, the BS will compute the average energy level of all nodes. Only nodes with an energy level above the average are eligible to be CH candidates for this round to ensure that only nodes with sufficient energy are selected as CHs. Next, the BS runs the first tier (the clustering algorithm) of TPSO-CR to find the best K CHs.

6.1.1. Initialization of particles

The dimension of the particle is same as the number of CH nodes (i.e., K) in the network. Let, $P_i = [X_{i,1}, X_{i,2}, X_{i,3}, \dots, X_{i,K}]$ be the ith particle of the population where each component, $X_{i,d}, 1 \le d \le K$ denotes CH number d in particle number i. Each component is initialized with a randomly generated uniformly distributed number in the range $[1, \text{networksize} - 1]$.

It should be noted that the random initialization and the velocity update (1a) gives non-integer velocity values, which are converted to the nearest integer in the implementation. In the case that a particle generates duplicate ID's after position update, it is assigned a high penalty value to ensure that the protocol generates the specified predetermined number of CHs.

Illustration 6.1: Consider a WSN with 60 sensor nodes and the number of CHs is 3 ($5\% \times 60$). Therefore, the dimension of the particle is same as the number of CHs, i.e. $K=3$.

Now, for each $X_{i,d}, 1 \le d \le 3$ of the particle P_i, a random number is generated to initialize it. Let us assume that a particle $P_i = [31.2, 20.8, 9.4]$, has been randomly generated. The second component of this particle is $X_{i,d} = 20.8$ then the 2nd elected CH $\text{ID} = \lfloor 20.8 \rfloor = 20$. Hence, the CH candidates IDs that result from this particle are 31, 20 and 9.

Now, let us consider another particle $P_j = [31.2, 31.8, 9.4]$. The CHs candidates generated are 31, 31 and 9. Since there is duplication in the generated CHs, then this particle is assigned a high penalty value to exclude it from further consideration.

6.1.2. Particles evaluation

The next step after initializing the particles is evaluating them according to some fitness function. This helps to periodically update the personal best and global best of the particles. To evaluate the particles, we use the fitness function defined before in (8).

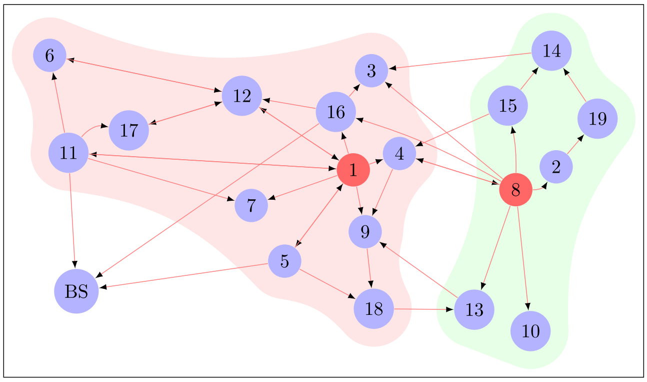


Fig. 1. A random 20 sensor node network with 2 cluster heads (1 and 8).

6.2. The routing algorithm

After finding the optimal set of CHs in the first tier, the second tier of the protocol is responsible for constructing the routing tree. The BS runs PSO again to achieve this task. The next sections give a detailed description of the PSO used in the routing tree construction phase.

6.2.1. Initialization of particles

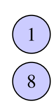
How to encode a routing tree into particle is critical for developing the second tier of TPSO-CR. Random encoding cannot be used for the following reasons:

- Random encoding results in different particle sizes due to different route lengths.
- A random sequence of edges usually does not correspond to a valid tree (that terminates on the destination node without any loop).

Node ID:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Priority:	1.0	0.6	−0.3	0.3	−0.1	0.5	−0.7	−0.1	−0.5	−0.5	−0.1	0.7	−0.3	−0.4	−0.5	0.3	0.5	−0.2	−0.1	−0.2

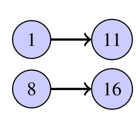
a Particle P_i encoding for network in Fig 2

Node ID:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Priority:	1.0	$-N$	−0.3	0.3	−0.1	0.5	−0.7	−0.1	−0.5	−0.5	−0.1	0.7	−0.3	−0.4	−0.5	0.3	0.5	−0.2	−0.1	−0.2
	1.0	0.6	−0.3	0.3	−0.1	0.5	−0.7	−0.1	$-N$	−0.5	−0.1	0.7	−0.3	−0.4	−0.5	0.3	0.5	−0.2	−0.1	−0.2



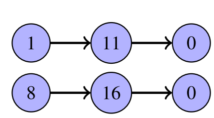
b Particle P_i after adding the CHs to the routing tree

Node ID:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Priority:	1.0	$-N$	−0.3	0.3	−0.1	0.5	−0.7	−0.1	−0.5	−0.5	−0.1	$-N$	−0.3	−0.4	−0.5	0.3	0.5	−0.2	−0.1	−0.2
	1.0	0.6	−0.3	0.3	−0.1	0.5	−0.7	−0.1	$-N$	−0.5	−0.1	0.7	−0.3	−0.4	−0.5	0.3	$-N$	−0.2	−0.1	−0.2



c Particle P_i after adding nodes 11 and 16 as relay nodes

Node ID:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Priority:	$-N$	$-N$	−0.3	0.3	−0.1	0.5	−0.7	−0.1	−0.5	−0.5	−0.1	$-N$	−0.3	−0.4	−0.5	0.3	0.5	−0.2	−0.1	−0.2
	$-N$	0.6	−0.3	0.3	−0.1	0.5	−0.7	−0.1	$-N$	−0.5	−0.1	0.7	−0.3	−0.4	−0.5	0.3	$-N$	−0.2	−0.1	−0.2



d Particle P_i after adding the BS and finishing the routing tree construction

Fig. 2. Example of priority-based encoding and decoding process for an arbitrary particle P_i: (a) Particle P_i encoding for network in Fig. 2; (b) particle P_i after adding the CHs to the routing tree; (c) particle P_i after adding nodes 11 and 16 as relay nodes; and (d) particle P_i after adding the BS and finishing the routing tree construction.

- The PSO algorithm involves arithmetic operations such as updating velocity and position which will not be applicable and will increase the number of invalid paths returned.

Mohemmed et al. (2008) have proposed an indirect priority encoding scheme to solve the problems of random encoding. This scheme has been applied successfully in many PSO based protocols like Chitra and Subbaraj (2012), Yao et al. (2011), Mohemmed and Sahoo (2007). In this scheme, the particle encodes guiding information about the solution rather than the solution itself. The guiding information used is the priorities of various nodes in the network.

In this paper, a slightly modified scheme is proposed to suit the need to find the optimal routing tree which connects all the CHs and the BS.

Particles encoding process: The dimension of the particle is same as the number of sensor nodes in the network (i.e., N). Let, $P_i=[X_{i,1},X_{i,2},X_{i,3},\ldots,X_{i,N}]$ be the ith particle of the population where each component, $X_{i,d}, 1\le d\le N$ denotes node N_d priority for selecting it as a relay node. Each component is initialized with a randomly generated uniformly distributed number in the range $[-1.0, 1.0]$.

Particles decoding process: A routing tree is built from the encoded particle in a branch growth process. Each branch is a route from a CH to the BS. For example, if there are two CHs in the network, the decoding process will generate 2 routes, one for each CH. Each route is constructed by appending relay nodes starting from the CH. At each step of the route construction, the next node with the highest priority is chosen from those which have direct links with the current node. The node that is already included in a growing path will be assigned a large negative priority value hence that node is highly unlikely to be selected again. At worst case, if a node is selected again, the concerned route can be treated as an invalid route and is assigned a high penalty value. The process continues until the BS is reached, and all the CHs are connected to the BS. A routing tree is considered invalid if it has one or more invalid branches (that does not terminate on the destination node or that have loops) and will be assigned a very high fitness value as a penalty. The best particle at the end of a run of the algorithm is that one that contains priorities that lead to the decoding procedure to select nodes forming the optimal routing tree (Fig. 1).

Illustration 6.2: Consider a WSN with 20 sensor nodes and 2 cluster heads, i.e., N_1, N_8 as shown in Fig. 2. Therefore, the dimension of the particles is same as the number of sensor nodes, i.e., N=20. Consider the directed acyclic graph $G(V,E)$ shown in Fig. 2. The edge $u\rightarrow v$ indicates that u can send to v but not necessarily vice versa.

Let us assume that a particle P_i has been randomly generated in Fig. 2(a). To find a routing tree from N_1 and N_8 to the BS, the protocol will build a route from N_1 to the BS, and another route from N_8 to the BS.

To find a branch from N_1 to the BS, a node that is connected to N_1 is identified first. As seen from Fig. 2, the nodes $[4,5,7,9,11,12,16]$ are such nodes to be considered. The priorities for them are $[-0.1, 0.5, -0.1, -0.5, 0.7, -0.3, -0.3]$ respectively. Node 11 has the highest priority and hence is used as the next relay node to N_1 and its priority is updated into a high negative value $-N$ to avoid selecting it again in the route. The possible nodes from node 11 are nodes $[0,1,6,7,17]$. The priorities of these nodes are $[1.0, -N, -0.7, -0.1, -0.2]$ respectively. Since node 0 (BS) has the highest priority than the other nodes, it is taken as the next relay node while constructing the route. Since the BS is reached, the route construction from N_1 to BS ends and results in the following route $(8,1,0)$. The same procedure is repeated for the branch from N_8 to the BS until a complete route $(8,16,0)$ is achieved. Figure 2(b–d) demonstrates this process.

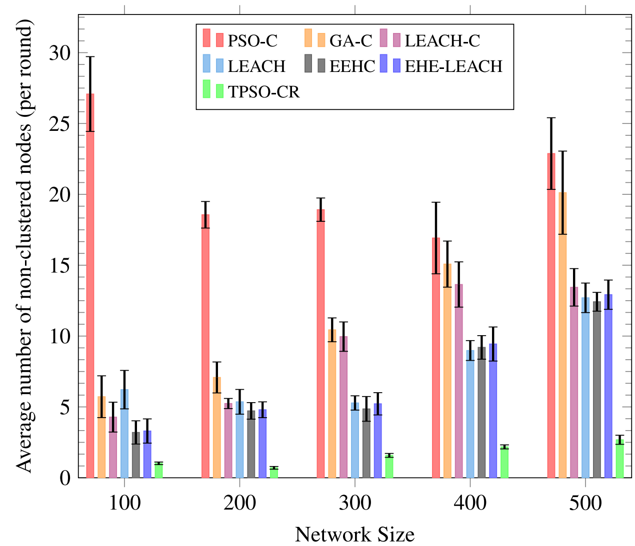


Fig. 3. Average number of non-clustered nodes per round for WSN#1.

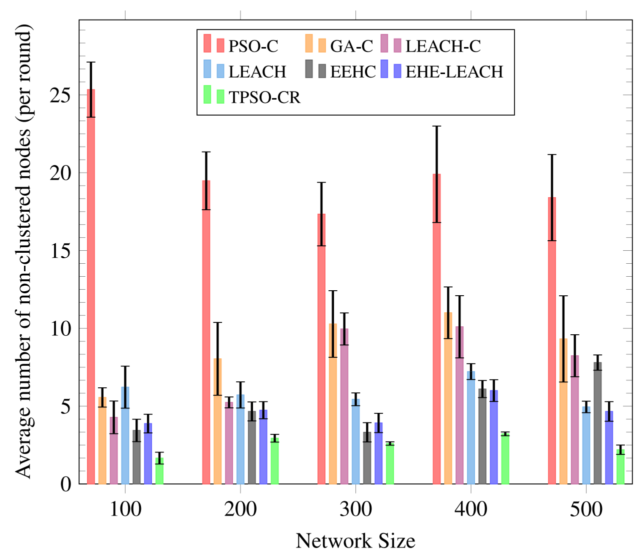


Fig. 4. Average number of non-clustered nodes per round for WSN#2.

6.2.2. *Particles evaluation*

After particles initialization, the generated routing tree that results from the decoding process is evaluated according to (12) to determine its fitness value.

6.3. *Complexity analysis*

In this subsection, we perform a complexity analysis for the proposed protocol (Algorithm 2). We consider two different types of nodes. The first one is the BS, as it executes both the clustering and the routing protocols. The second node represent any other network node (CH or non-CH).

6.3.1. *The BS node*

In the clustering algorithm, the main function is clustering the whole network for each particle. Hence, the algorithm complexity is $O(MNK)$ where M is the number of particles, N is the network size and K is the number of CH candidates. Where in the routing algorithm, the main function is decoding each particle into a routing tree which has a complexity of $O(MN^2)$.

Table 2
Mean PDR and standard deviation in WSN#1.

Protocols	100 Sensor nodes		200 Sensor nodes		300 Sensor nodes		400 Sensor nodes		500 Sensor nodes	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LEACH	0.536	0.024	0.621	0.032	0.674	0.020	0.671	0.017	0.671	0.012
EHE-LEACH	0.572	0.017	0.622	0.023	0.643	0.020	0.634	0.016	0.630	0.010
EEHC	0.589	0.032	0.661	0.037	0.680	0.006	0.669	0.009	0.670	0.015
PSO-C	0.606	0.045	0.775	0.010	0.809	0.012	0.832	0.005	0.835	0.004
GA-C	0.865	0.008	0.873	0.005	0.871	0.003	0.862	0.002	0.861	0.003
LEACH-C	0.861	0.006	0.892	0.001	0.890	0.002	0.890	0.002	0.887	0.002
TPSO-CR	**0.877**	0.008	**0.893**	0.002	**0.895**	0.001	**0.891**	0.002	**0.892**	0.001

Table 3
Mean PDR and standard deviation in WSN#2.

Protocols	100 Sensor nodes		200 Sensor nodes		300 Sensor nodes		400 Sensor nodes		500 Sensor nodes	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LEACH	0.536	0.024	0.617	0.031	0.667	0.015	0.672	0.013	0.680	0.036
EHE-LEACH	0.577	0.042	0.635	0.035	0.625	0.062	0.605	0.016	0.580	0.016
EEHC	0.588	0.029	0.651	0.016	0.577	0.058	0.669	0.020	0.639	0.031
PSO-C	0.615	0.029	0.787	0.016	0.823	0.013	0.826	0.036	0.838	0.020
GA-C	0.861	0.015	0.871	0.006	0.867	0.001	0.866	0.003	0.863	0.013
LEACH-C	0.827	0.035	0.865	0.015	0.871	0.004	0.861	0.021	0.867	0.018
TPSO-CR	**0.890**	0.008	**0.893**	0.002	**0.895**	0.001	**0.890**	0.006	**0.895**	0.010

Table 4
Mean for average consumed energy per node and standard deviation in WSN#1.

Protocols	100 Sensor nodes		200 Sensor nodes		300 Sensor nodes		400 Sensor nodes		500 Sensor nodes	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LEACH	175.19	6.545	149.22	10.11	131.07	5.712	132.37	5.500	131.74	4.400
EHE-LEACH	155.19	9.007	140.35	6.245	131.76	4.558	131.64	3.839	130.32	3.571
EEHC	158.85	9.001	137.48	10.35	131.28	1.539	131.32	3.711	130.57	4.673
PSO-C	73.855	0.042	72.208	0.061	71.593	0.028	71.303	0.085	71.102	0.027
GA-C	74.499	0.074	72.660	0.305	71.824	0.304	71.602	0.121	71.336	0.386
LEACH-C	74.549	0.003	73.060	0.004	72.559	0.005	72.308	0.013	72.161	0.006
TPSO-CR	**71.271**	0.217	**71.254**	0.175	**71.142**	0.222	**71.129**	0.115	**71.203**	0.086

Table 5
Mean for average consumed energy per node and standard deviation in WSN#2.

Protocols	100 Sensor nodes		200 Sensor nodes		300 Sensor nodes		400 Sensor nodes		500 Sensor nodes	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LEACH	175.19	6.545	150.53	10.06	133.62	3.629	130.58	7.248	128.43	5.721
EHE-LEACH	155.68	11.27	137.22	11.34	145.00	10.33	135.09	11.20	145.13	10.62
EEHC	158.98	7.263	140.80	5.602	148.40	9.181	129.70	9.201	139.55	8.120
PSO-C	73.817	0.303	72.086	0.069	71.379	0.222	73.761	0.015	73.529	0.083
GA-C	74.528	0.035	72.752	0.277	71.979	0.249	71.491	0.032	71.357	0.021
LEACH-C	74.549	0.003	73.060	0.004	72.559	0.005	72.311	0.001	72.151	0.002
TPSO-CR	**71.367**	0.133	**71.188**	0.144	**71.196**	0.168	**70.986**	0.116	**71.227**	0.192

6.3.2. *Other network nodes*

For a CH node to aggregate L buffered packets, the complexity is $O(L)$. For a non-CH node to send its buffered packets, the complexity is $O(1)$.

7. Simulations and results

In this section, the performance of TPSO-CR is investigated against the well known protocols LEACH, EHE-LEACH, EEHC, the SA-based protocol LEACH-C, the PSO-based protocol PSO-C and the GA-based protocol proposed by Rahmanian et al. (2011), which we will refer here as GA-C.

Simulations were carried on Castalia, which is based on the OMNeT++ platform and can be used to test WSN protocols in realistic wireless channel and radio models (Rastegarnia and Solouk, 2011), with a realistic node behavior. It provides a generic reliable and realistic framework for the first order validation of an algorithm before moving to implementation on a specific sensor platform (Patil and Hadalgi, 2012). The comparisons are used for the purpose of benchmarking TPSO-CR against the well known protocols cited in the literature.

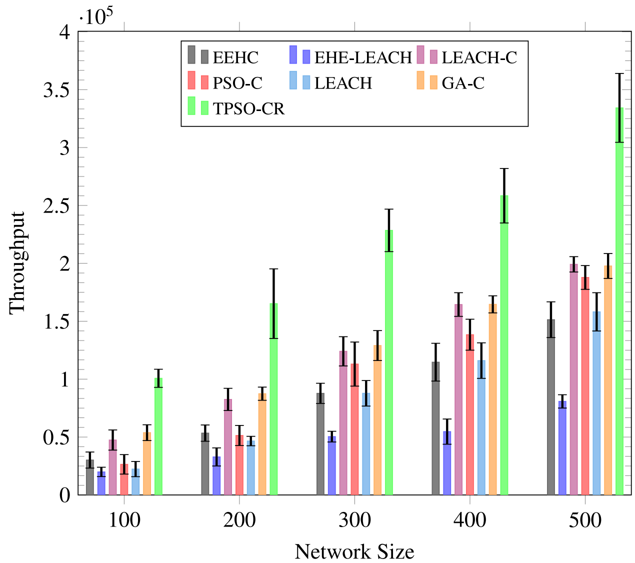


Fig. 5. Throughput for WSN#1.

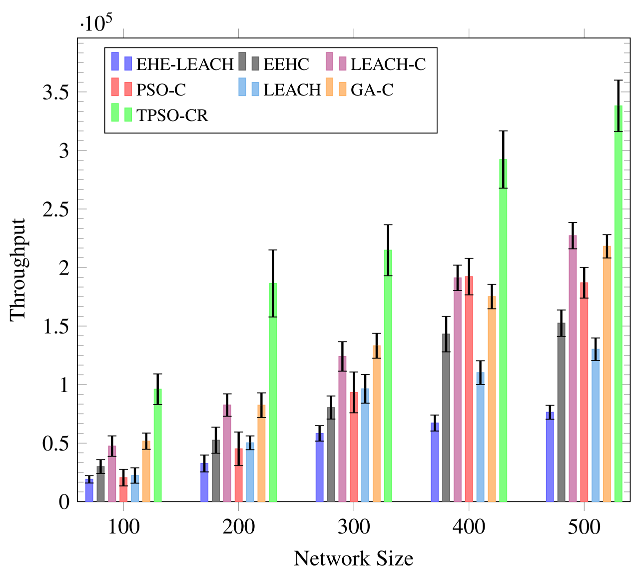


Fig. 6. Throughput for WSN#2.

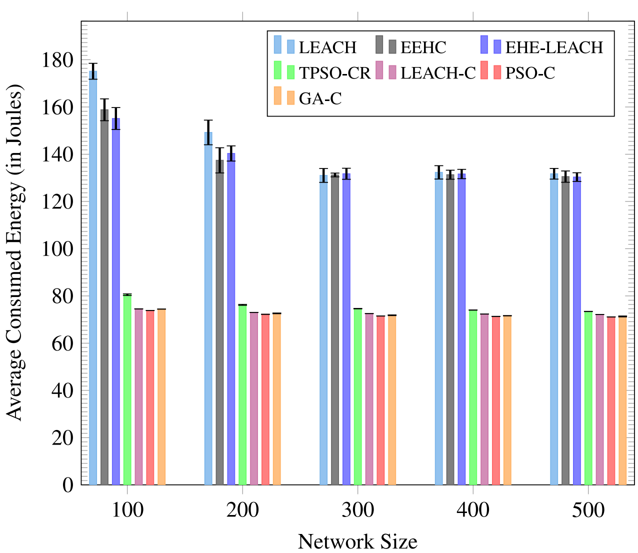


Fig. 7. Average consumed energy per node for WSN#1.

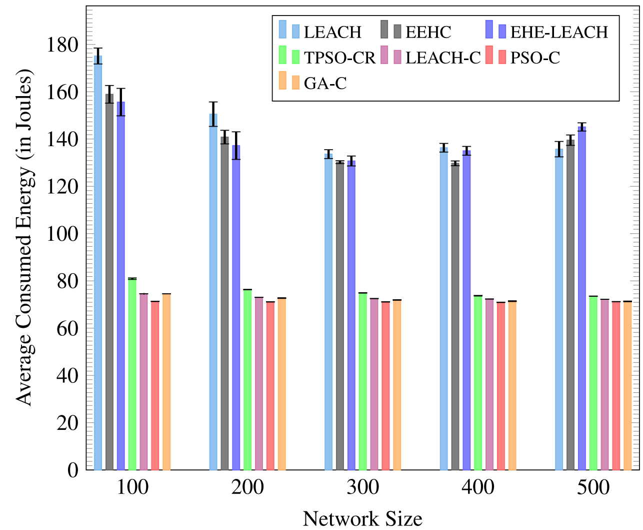


Fig. 8. Average consumed energy per node for WSN#2.

Table 6
Number of aggregated packets received at the BS.

Protocol	Number of packets
LEACH	613.5
LEACH-C	472.5
EEHC	618.5
EHE-LEACH	309.5
GA-C	778.5
PSO-C	952.5
TPSO-CR	1237.2

According to the heterogeneity of the sensors, the simulations were performed on two groups of WSNs (WSNs#1, WSNs#2), each with 25 different playground topologies. The first case assumes homogeneous sensor networks (WSNs#1) while the second experiments (WSNs#2) assume heterogeneous sensor networks with advanced nodes of 10% and super nodes of 10%.

Each WSN group consists of 5 different network sizes ranging from 100 to 500 sensor nodes. Overall, the simulation results presented herein have been averaged over 5 simulation runs for a total of 50 different networks.

The sensor nodes were deployed randomly in an area of 100 m × 100 m sensor field. The BS was located at the field's corner at position (0, 0). For the medium access control protocol, we used TMAC which is known for its energy efficiency because it adapts a variable sleep schedule that increases the battery utilization (Khatarkar and Kamble, 2013).

The percentage of CHs was set to 5% of the total nodes as was used by the competent protocols. We run the protocols for 5000 s and in order to minimize the protocol overhead, we set the round length to 500 s with a slot length of 0.4 s. Data packets were generated at a rate of 1 packet/s.

In WSNs#1, the initial energy of a standard node is set to 18,720 joules, which is the typical energy of two AA batteries (Boulis, 2011). In WSNs#2, the initial energy of a normal node is set to 6240 joules, super node initial energy is set to 12,480 joules and advanced node initial energy is set to 18,720 joules.

To execute our proposed algorithms, we considered an initial population of 50 particles and we let them evolve for 200 iterations. The values of PSO parameters are taken same default values as in Zambrano-Bigiarini et al. (2013). For the weight sum approach, we gave equal weight to each sub-objective. Hence, we set $wc_1 = wc_2 = wc_3 = 0.33$ and $wr_1 = wr_2 = wr_3 = 0.33$.

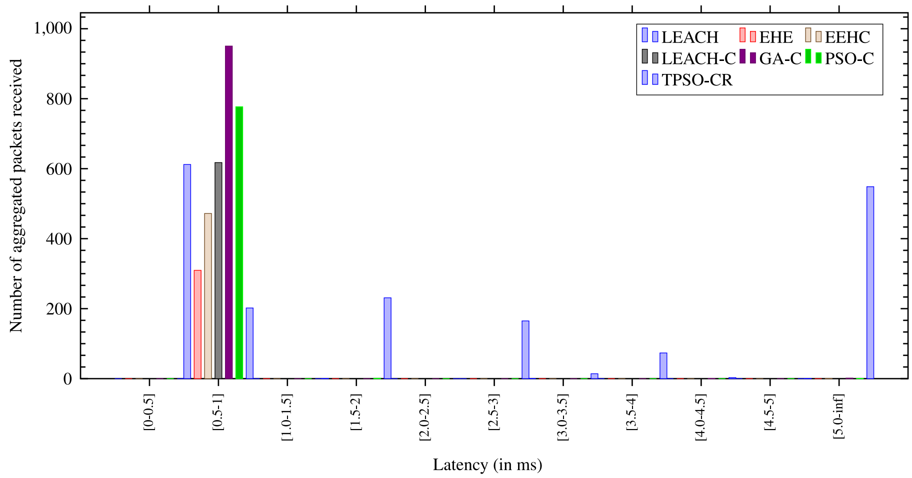


Fig. 9. Application-level latency (in ms).

First, we run the protocols for comparing the coverage quality of the proposed protocol by varying the sensor nodes from 100 to 500 on both the network scenarios, WSN#1 and WSN#2. Figures 3 and 4 show the comparison of the proposed protocol with the other protocols in terms of average number of non-clustered nodes per round in WSN#1 and WSN#2 respectively. We present the results at 99% confidence interval.

It can be observed from Figs. 3 and 4 that the proposed protocol has significantly better network coverage than the other competent protocols. This is due to the clustering phase the proposed protocol which takes care of minimizing the number of non-clustered nodes. Whereas the existing protocols do not deal with that problem.

In order to judge the cluster-based link quality of the proposed protocol, the average (mean) PDR for packets received by all the CHs for 5 runs of the protocols along with their standard deviations (SD) for both the scenarios WSN#1 and WSN#2 are calculated by varying number of sensor nodes. The results are shown in Tables 2 and 3 for WSN#1 and WSN#2 respectively. It is clear that the average PDR for the proposed protocol is maximum with nearly minimum fluctuations in the average PDR. LEACH, EHE-LEACH and EEHC protocols have more fluctuations around the average due to their probabilistic nature. Furthermore, they have much lower PDR because no link quality measure is taken in either of them.

Next, we execute the protocols to compare their energy efficiency by varying the sensor nodes from 100 to 500 on both the network scenarios, WSN#1 and WSN#2. In order to investigate the effect of using a dedicated routing tree on the proposed protocol energy efficiency, we initially executed the proposed protocol using the clustering algorithm only (no relay nodes were used). Tables 4 and 5 show the average (mean) of the average energy consumed by node (in joules) for 5 runs of the protocols along with their standard deviations (SD) for both the scenarios WSN#1 and WSN#2. It is clearly shown that TPSO-CR has the lowest average consumed energy. However, it has more fluctuations around the average.

Figures 5 and 6 show the comparison of the proposed algorithm and other protocols in terms of the network throughput in WSN#1 and WSN#2 respectively. Throughput is defined as the number of data packets successfully received at the BS. Using the number of aggregated packets delivered to the BS is not accurate, since many packets result from the aggregation process of many raw packets collected from the cluster members. In this paper, the number of the raw packets is used to calculate the throughput at the BS. We present the results at 99% confidence interval. It can be observed that the proposed protocol significantly outperforms the other protocols in terms of network throughput as shown in Figs. 5 and 6.

Now, we show the effect of using relay nodes for multi-hop data transmission on the proposed protocol energy efficiency. Figures 7 and 8 show the comparison of the proposed algorithm and other protocols in terms of the average energy consumed by node (in joules) in WSN#1 and WSN#2 respectively. We present the results at 99% confidence interval. It was noted that, in the case of sparsely deployed WSN, the average energy consumed per node in TPSO-CR is higher than LEACH-C, GA-C and PSO-C. This is mainly due to an increase in the number of active nodes during any round. This increase is caused by the adding more nodes to act as relay nodes since the number of CHs is small and their transmission range is limited. As the sensor density increases, the number of CHs that cover the same area increases. At the same time, the routing algorithm favors the inter-cluster communication between the CHs. This caused the average consumed energy for TPSO-CR to be relative to that of LEACH-C, GA-C and PSO-C in densely deployed WSN.

It was also noted that much higher energy consumption was recorded in LEACH, EHE-LEACH and EEHC. The reason behind that is the non-clustered nodes which are left unattended without any sleeping schedule. Hence, they are consuming energy even if their nodes are in the idle state. Theoretically, LEACH-C, GA-C and PSO-C cluster all the network nodes and thus give each node a sleep schedule depending on its TDMA turn to transmit. This caused both protocols to have lower energy consumption compared to that of LEACH type protocols. For TPSO-CR, any non-clustered node is set to sleep during the whole round. Although this should reduce the energy consumption for TPSO-CR compared to that of other protocols, it is not reflected in Figs. 7 and 8 because the number of non-clustered nodes is already the minimum in TPSO-CR.

To examine the application-level latency for the aggregated packets received at the BS, a heterogeneous network of 200 sensor nodes in an area of 100 m $\times$ 100 m was used. The simulation run for 1000 s (2 rounds) and was repeated using 2 different random seeds then the average was taken. Table 6 presents the number of aggregated packets received at the BS for each protocol while Fig. 9 presents the latency distribution for those packets.

Although TPSO-CR has higher throughput than the other protocols, it has higher packet latency. For example, Fig. 9 reveals that, for the competent protocols, over 90% of the received packets have a maximum latency of 1 ms. For TPSO-CR, only 16% of the packets have a maximum latency of 1 ms. The reason that TPSO-CR is experiencing higher latency is because the Medium Access Control (MAC) layer of each relay node buffers each packet before transmitting it to the next relay node.

8. Conclusions and future work

In this paper, the problem of clustering and routing in WSN was studied. A PSO inspired protocol was proposed to solve that issue. The protocol runs in two tiers: first one finds the best CHs and their associative clusters while the second tier solves the problem of the inter-cluster communication by finding the optimal routing tree. The protocol was developed and tested under realistic network and energy consumption model. Extensive simulations were conducted, and the results show that the proposed protocol can significantly improve the packet delivery rate at both the cluster heads and the BS, increase network coverage and at the same time maintain acceptable energy consumption. Furthermore, the protocol does not assume any unrealistic assumptions, for example, using GPS for location discovery.

Future research directions can be inspired from the reported results. The protocol can be extended to allow two-hop hierarchical clusters with a threshold on the link quality between any cluster member and its next hop. This may result in better clusters quality and maximize the throughput. An adaptive power control method can also be adapted to enhance the network energy efficiency.

References

Abbasi AA, Younis M. A survey on clustering algorithms for wireless sensor networks. Comput. Commun. 2007;30(14–15):2826–41.

Abdul Latiff N, Tsimenidis C, Sharif B. Performance comparison of optimization algorithms for clustering in wireless sensor networks. In: IEEE international conference on mobile adhoc and sensor systems. MASS 2007; October 2007. p. 1–4.

Agarwal PK, Procopiuc CM. Exact and approximation algorithms for clustering. Algorithmica 2002;33(2):201–26.

Akkaya K, Younis M. A survey on routing protocols for wireless sensor networks. Ad Hoc Netw. 2005;3(3):325–49.

Arboleda LM, Nasser N. Comparison of clustering algorithms and protocols for wireless sensor networks. In: IEEE canadian conference on electrical and computer engineering; May 2006. p. 1787–92.

Baccour N, Koubâa A, Mottola L, Zúñiga MA, Youssef H, Boano CA, et al. Radio link quality estimation in wireless sensor networks: a survey. ACM Trans. Sensor Netw. (TOSN) 2012;8(September(4)):34:1–33.

Barberis A, Barboni L, Valle M. Evaluating energy consumption in wireless sensor networks applications. In: 10th Euromicro conference on digital system design architectures, methods and tools, 2007. DSD 2007; August 2007. p. 455–62.

Boulis A. Castalia user's manual. ⟨https://forge.nicta.com.au/docman/view.php/301/592/Castalia+-+User+Manual.pdf⟩; 2011 [accessed 02.04.2014].

Chitra C, Subbaraj P. A nondominated sorting genetic algorithm solution for shortest path routing problem in computer networks. Expert Syst. Appl. 2012;39(1):1518–25.

Clerc M. Standard particle swarm optimisation. ⟨http://clerc.maurice.free.fr/pso/SPSO_descriptions.pdf⟩; 2012 [accessed 25.09.2014].

del Valle Y, Venayagamoorthy G, Mohagheghi S, Hernandez J-C, Harley R. Particle swarm optimization: basic concepts, variants and applications in power systems. IEEE Trans. Evol. Comput. 2008;12(April (2)):171–95.

Dorigo M, Birattari M, Stutzle T. Ant colony optimization. IEEE Comput. Intell. Mag. 2006;1(November (4)):28–39.

Dwivedi AK, Vyas OP. Network layer protocols for wireless sensor networks: existing classifications and design challenges. Int. J. Comput. Appl. 2010; 8(October (12)):30–4.

Guo W, Zhang W. A survey on intelligent routing protocols in wireless sensor networks. J. Netw. Comput. Appl. 2014;38(0):185–201.

Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd annual Hawaii international conference on system sciences, vol. 2; January 2000.

Heinzelman W, Chandrakasan A, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks. IEEE Trans. Wirel. Commun. 2002;1(October (4)):660–70.

Jiang C-J, ren Shi W, Xiang M, lun Tang X. Energy-balanced unequal clustering protocol for wireless sensor networks. J. China Univ. Posts Telecommun. 2010;17(4):94–9.

Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of IEEE international conference on neural networks, 1995, vol. 4; November/December 1995. p. 1942–8.

Khalil EA, Attea BA. Energy-aware evolutionary routing protocol for dynamic clustering of wireless sensor networks. Swarm Evol. Comput. 2011;1(4):195–203.

Khatarkar S, Kamble R. Wireless sensor network mac protocol: Smac and tmac. Ind. J. Comput. Sci. Eng. 2013;4(4):304–10.

Konak A, Coit DW, Smith AE. Multi-objective optimization using genetic algorithms: a tutorial. Reliab. Eng. Syst. Saf. 2006;91(9):992–1007.

Kuila P, Gupta SK, Jana PK. A novel evolutionary approach for load balanced clustering problem for wireless sensor networks. Swarm Evol. Comput. 2013; 12(0):48–56.

Kulkarni R, Venayagamoorthy G. Particle swarm optimization in wireless-sensor networks: a brief survey. IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev. 2011;41(March (2)):262–7.

Kumar D. Performance analysis of energy efficient clustering protocols for maximising lifetime of wireless sensor networks. IET Wirel. Sensor Syst. 2014; 4(March (1)):9–16.

Kumar D, Aseri TC, Patel R. Eehc: energy efficient heterogeneous clustered scheme for wireless sensor networks. Comput. Commun. 2009;32(4):662–7.

Latiff N, Tsimenidis C, Sharif B. Energy-aware clustering for wireless sensor networks using particle swarm optimization. In: IEEE 18th international symposium on personal, indoor and mobile radio communications (PIMR C'07); September 2007. p. 1–5.

Lee L, Nwana H, Ndumu D, de Wilde P. The stability, scalability and performance of multi-agent systems. BT Technol. J. 1998;16(3):94–103.

Mohemmed AW, Sahoo NC. Efficient computation of shortest paths in networks using particle swarm optimization and noising metaheuristics. Discrete Dyn. Nat. Soc. 2007:1–25.

Mohemmed AW, Sahoo NC, Geok TK. Solving shortest path problem using particle swarm optimization. Appl. Soft Comput. 2008;8(4):1643–53.

Molina G, Alba E. Location discovery in wireless sensor networks using metaheuristics. Appl. Soft Comput. 2011;11(1):1223–40.

Patil A, Hadalgi DPM. Evaluation of discrete event wireless sensor network simulators. Int. J. Comput. Sci. Netw. 2012.

Rahmanian A, Omranpour H, Akbari M, Raahemifar K. A novel genetic algorithm in leach-c routing protocol for sensor networks. In: 24th Canadian conference on electrical and computer engineering (CCECE), 2011; May 2011. p. 001096–100.

Rastegarnia A, Solouk V. Performance evaluation of castalia wireless sensor network simulator. In: 2011 34th international conference on telecommunications and signal processing (TSP); August 2011. p. 111–5.

Saleem M, Caro GAD, Farooq M. Swarm intelligence based routing protocol for wireless sensor networks: Survey and future directions. Inf. Sci. 2011;181(20):4597–624.

Srinivasan K, Dutta P, Tavakoli A, Levis P. Understanding the causes of packet delivery success and failure in dense wireless sensor networks. In: Proceedings of the 4th international conference on embedded networked sensor systems; November 2006. p. 419–20.

Srinivasan K, Dutta P, Tavakoli A, Levis P. An empirical study of low-power wireless. ACM Trans. Sensor Netw. 2010;6(March (2)):16:1–49.

Srinivasan K, Levis P. RSSI is under appreciated. In: Proceedings of the third workshop on embedded networked sensors (EmNets 2006); May 2006.

Texas Instruments, Chipcon CC2420 radio transceiver data sheet. http://www.ti.com/lit/ds/symlink/cc2420.pdf; 2013 [accessed 25.09.2014].

Tubaishat M, Madria S. Sensor networks: an overview. IEEE Potent. 2003;22(April (2)):20–3.

Tyagi S, Gupta S, Tanwar S, Kumar N. Ehe-leach: enhanced heterogeneous leach protocol for lifetime enhancement of wireless sns. In: 2013 international conference on advances in computing, communications and informatics (ICACCI); August 2013. p. 1485–90.

Tyagi S, Kumar N. A systematic review on clustering and routing techniques based upon {LEACH} protocol for wireless sensor networks. J. Netw. Comput. Appl. 2013;36(2):623–45.

Yao J, Yang B, Zhang M, Kong Y. Pso with predatory escaping behavior and its application on shortest path routing problems. In: 2011 3rd international workshop on intelligent systems and applications (ISA); May 2011. p. 1–4.

Ye M, Li C, Chen G, Wu J. EECS: an energy efficient clustering scheme in wireless sensor networks. In: 24th IEEE international performance, computing, and communications conference, 2005, IPCCC 2005; April 2005. p. 535–40.

Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans. Mob. Comput. 2004; 3(October (4)):366–79.

Younis O, Krunz M, Ramasubramanian S. Node clustering in wireless sensor networks: recent developments and deployment challenges. IEEE Netw. 2006;20(May (3)):20–5.

Yu Y, Prasanna VK, Krishnamachari B. Information processing and routing in wireless sensor networks. World Scientific Pub Co Inc., Singapore; 2006 [Chapter 1].

Zambrano-Bigiarini M, Clerc M, Rojas R. Standard particle swarm optimisation 2011 at cec-2013: a baseline for future pso improvements. In: IEEE congress on evolutionary computation (CEC); June 2013. p. 2337–44.

Zungeru AM, Ang L-M, Seng KP. Classical and swarm intelligence based routing protocols for wireless sensor networks: a survey and comparison. J. Netw. Comput. Appl. 2012;35(September (5)):1508–36.