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

freepaper.me FREE paper

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.

© 2015 Elsevier Ltd. All rights reserved.

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/© 2015 Elsevier Ltd. All rights reserved.

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

freepaper.me FREE paper

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

freepaper.me FREE paper

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.

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

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

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.

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

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

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

freepaper.me FREE paper