EVOLUTIONARY ALGORITHMS FOR CLUSTERING IN WIRELESS SENSOR NETWORKS

Kamal S. Hamza and Fathy A. Amir

Department of Information Technology, Cairo University,Giza, Egypt

ABSTRACT

Recent studies have shown that hierarchical routing is one of the best approaches for routing in Wireless Sensor Networks (WSN). In typical hierarchical routing techniques, the network divided into regions (a.k.a. Clusters); each cluster is then managed as a standalone network with its Cluster Head serving as the node of communication coordination and data collection. Accordingly, nodes across the network can save their energy when attempting to communicate with the other nodes far across the network. Several hierarchical routing techniques have been proposed over the last few years, despite the fact that most of these algorithms are fundamentally similar. Due to the nature and complexity of hierarchical routing, Evolutionary Optimization algorithms have been used recently to perform clustering and routing techniques. In this paper, we present, to the best of our knowledge, one of the first surveys on the use of Evolutionary Algorithms (EA) for clustering in hierarchical routing in WSNs. Categorization and classification of the various clustering techniques are presented and used to understand the existing algorithms. A comparison among these techniques is also presented in this work**.**

1. INTRODUCTION

Internet of Things (IoT) received an increasing attention over the last few years, mainly due to its role in solving social and economic challenges [1,2,3]. IoT typically refers to the interconnection of smart objects, such as sensor and Radio Frequency Identification (RFID), to deliver business value in a particular context [4,5,6]. Sensor networks play a fundamental role in the development and implementation of IoT as most of "smart things" in reality will be realized equipping objects with sensors that can collect and communicate data of a specific interest [7,2]. Most objects today either have sensors attached to them or can be extended with sensors. Most smart phones, for instance, have at least four built-in sensors that can continuously stream data to the Internet or other destinations. With such rapid advances in deploying sensors everywhere, communicating and routing the mass amount of collected data have become important; yet very challenging. The large diversity of sensor types; large number of nodes; and large distribution and coverage areas have all posed real challenges in the development of routing algorithms for WSNs.

Routing in WSNs can be generally classified into *flat* and *hierarchical* routing. In the former, all nodes in the sensor network can communicate with each other via single or multiple hops. In the latter, the network is viewed as a set of regions or clusters [8]; each cluster is seen as an independent small network. Each cluster has a Cluster Head (CH) that manages the communication among the nodes within the cluster and between the nodes of the clusters and the rest of the network. Intra-cluster communication is typically performed in two methods: single-hop or multiple-hop where nodes in a cluster communicating with the CH directly or through

another node respectively. Similarly, in Inter-cluster communication where CHs relay their data to BS either direct connection which called single-hop or through other CHs in case of multiple-hop connection used [9,10,11]. Clearly, hierarchical routing is more suitable for WSNs with large nodes, coverage areas, or diversity as it seems more natural to group and cluster nodes according to their location, size, or type. By clustering, nodes can reduce their routing tables and states, and hence, their communication overheads. As a typical result, networks can use clustering and hierarchical routing to prolong their lifetime and optimize their energy usage.

Accordingly, there has been an increasing focus on the WSN community in exploring hierarchical routing and improving the performance of the clustering and communication functions in such routing techniques. Several hierarchical routing algorithms have been proposed over the last decade with variations in complexity, performance, and efficiency. Even though the fundamental idea of hierarchical routing is the same; however, the various techniques vary in terms of the techniques used for forming and operating the clusters in the network. Cluster formation refers to the process of dividing a set of nodes into sets based on one or more objective functions. The cluster operation shows how the cluster is maintained, updated, and restructured during the operation of the network as shown in Figure 1.
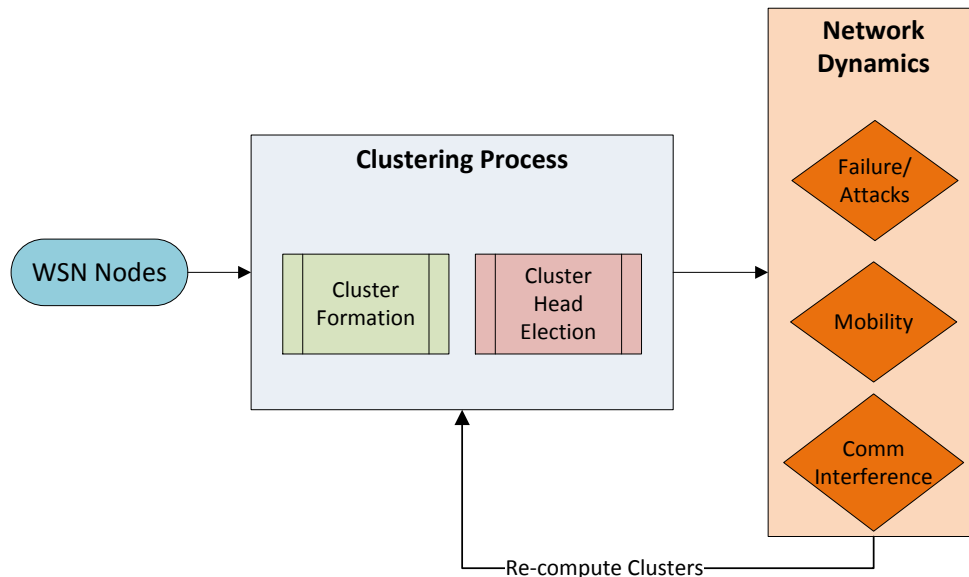

Figure 1. Wireless Sensor network clustering process

Clustering formation and operation tend to be a computationally complex problem as several factors may impact the quality and efficiency of the formed clusters at different points of times during the life-time of the network. Moreover, mobility and death of nodes result in changes in the structure of the cluster during the operation of the network. In addition, due to the variety in the applications types and needs, the objective for the clustering itself may change from one network to another. For instance, in some contexts, reduction of communication delay is favored over optimization of energy. In other applications, increasing the life-time of the network is more important than reducing delay or the probability of packet loss. With the wide diversity of applications and their needs, clustering is becoming a more challenging computing problem in the context of WSN.

In the recent few years, there have been an increasing number of researches on using Evolutionary Algorithms to develop efficient clustering for hierarchical routing. There are several

evolutionary techniques reported and used in the literature; however, they share the same underlying concept of randomly creating a set of candidate solutions and select the best among these solutions based on some quality function. The selected solutions are then used as a seed to a next iteration of selection and so on until a solution with sufficient quality is found or a previously set computational limit is reached.

Several surveys in the literature have collected and analyzed various hierarchical routing and their clustering techniques (See for example [9-14]). However, none of these surveys cover evolutionary techniques except [15], which represented a brief survey of the centralized evolutionary clustering algorithms but not the distributed ones. In this paper, we present, to the best of our knowledge, one of the first surveys on the use of Evolutionary Algorithms in developing hierarchical routing in (Centralized – Distributed) WSNs. Categorization and classification of the various techniques are presented and used to understand the existing algorithms. A comparison among these techniques is also presented in this work.

The rest of the paper is organized as follows. Section 2 introduces a synopsis of different existing evolutionary clustering algorithms and an explanation of clustering algorithms. The parameters used for comparison are presented in Section 3. In Section 4, a comparison of the various Evolutionary Algorithms is presented; Conclusion is given in Section 5.

2. EVOLUTIONARY CLUSTERING ALGORITHMS

Different techniques of Evolutionary Algorithms have developed during the last few decades. Despite their differences; they are all inspired by the same fundamental principles of biological evolution. Evolutionary Algorithms operate on a population of potential solutions applying the principle of survival of the fittest to produce better and better approximations to a solution.

In this section we review nine different existing evolutionary families of algorithms used in optimizing the clustering problem in WSN. Figure 2 shows the various types of evolutionary algorithms considered in this paper. Most of Evolutionary Algorithms focused on the cluster reformation process, while others paid attention to work on both clustering and routing processes. The simulation results of the various reviewed algorithms are also presented.
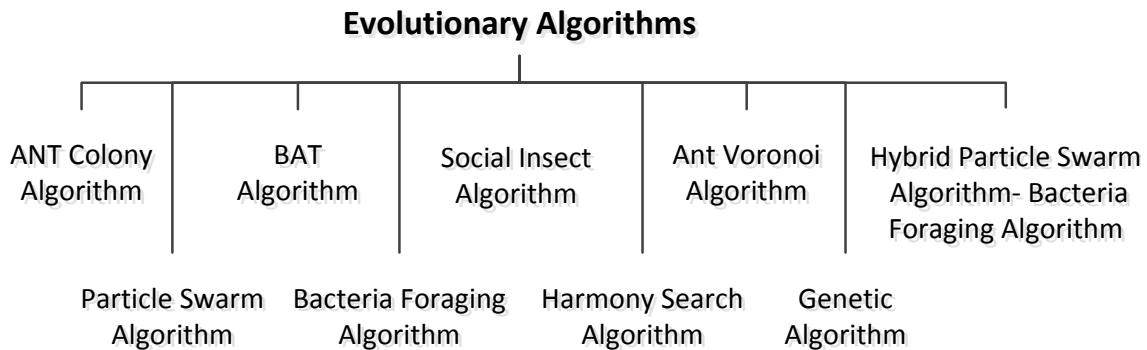
Evolutionary Algorithms

- ANT Colony Algorithm
  - Particle Swarm Algorithm
- BAT Algorithm
  - Bacteria Foraging Algorithm
- Social Insect Algorithm
  - Harmony Search Algorithm
- Ant Voronoi Algorithm
  - Genetic Algorithm
- Hybrid Particle Swarm Algorithm- Bacteria Foraging Algorithm

Figure 2. Evolutionary Algorithms for clustering in WSN

2.1 Ant colony optimization (ACO)

In [16], *an Adaptive Virtual Area Partition Clustering Routing Protocol Using Ant Colony Optimization (AVAPCR-ACO)* is proposed. AVARPCR-ACO clusters the network adaptively then builds the routing path using ant colony optimization. The flow chart of AVAPCR-ACO is shown in Figure 3.
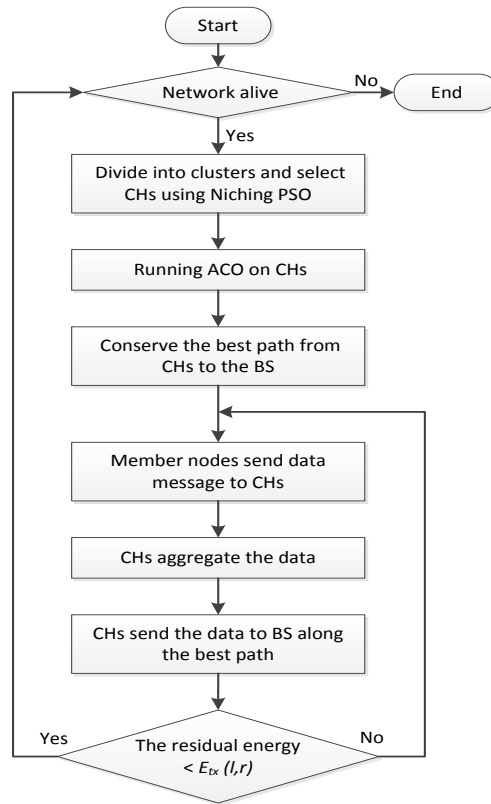
Figure 3. Flow chart of AVAPCR-ACO [16]

The network virtually partitioned into an optimum number of clusters calculated according to [17]. After partitioning, the network nodes start creating a routing path based on Ant Colony Optimization (ACO). ACO inspired by the using a chemical substance which called *Pheromone* in ants' communication. The *pheromone* carries the Ants' experience information to construct the problem solution in every irritation [18,19]. It used to build the optimum routing path between CHs. The proposed algorithm compared to LEACH-C, VAP-E and ACO-MNCC. It gives a higher network lifespan and delivers more data to the base station.

2.2 Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a well-known population based stochastic optimization technique that has been applied intensively in various networking problems. In the following, we review some of PSO-based algorithms presented in [21].

2.2.1 *Particle Swarm Optimization based Clustering (PSOBC)* **[20]: PSOBC modifies the original** PSO algorithm that is used for continuous search space by defining a new operator to make it suitable for discrete search space. Each single solution in PSO will be considered as a Particle. Particles fly through the searching area of solution looking for the global optimum position that produces the best fitness of an objective function. Each particle keeps track of personal best position p_{best} and the global best position of the wholes warm g_{best} [22]. The simulation of the algorithm applied in three different cases, each one assumed a different position of the BS which obviously affected on the number of alive nodes as showed in the simulation results. Overall, a considerable increment in network lifetime appeared as compared to LEACH and LEACH-C.

2.2.2 ***Two-Tier Particle Swarm Optimization for Clustering and Routing Protocol (TPSO-CR)***
[23]: two PSO based protocol proposed, a clustering algorithm which used to find the optimal set of CHs and a routing algorithm to find the optimal routing tree for inter-cluster communications. In this article a realistic network and novel radio model in cluster-based communication for WSN has been used and investigated.

Most of existed algorithms referred to the first radio model [13] assuming an unrealistic energy consumption model which ignores the quantity of energy consumed listening mode. Moreover, they assume that all nodes are location aware and ignore the effect of distance on link quality. Instead of this, the latest standard PSO (SPSO-2011) is used to improve performance and converge towards the region of the global optimum [12].

The simulation results show that the proposed protocol minimized the average number of non-clustered nodes in addition to improving the data rate in inter-cluster communication, increase network coverage and maintain energy consumption.

2.3 Genetic Algorithm (GA)

Genetic Algorithm (GA) based on a natural selection process that mimics biological evolution. The generic flow chart of the Genetic Algorithm is depicted in Figure 4. Generally speaking, a GA codified a typical WSN as a "chromosome." A set of chromosomes is called a "generation", and they are kept or dropped based on their fitness functions. In the following, we review GA-based algorithms presented in [24].

2.3.1 *Genetic Centralize Dynamic Clustering (GCDC)* [25]: GCDC uses GA to find the optimal combinations of CH selection parameters as: residual energy and distance of inter-clustering and intra-clustering communication. The dynamic clustering changes the CHs over the time, equalizing the energy consumption across all nodes and, thus, extending the network lifetime.

The proposed algorithm improved the network coverage due to determining the optimal number of CHs and their position, unlike LEACH algorithm which consider fixed percentage of CHs without identifying their positions. That can lead to overlapped clusters. The simulation results show an improvement in a live nodes percentage and residual energy compared to LEACH.

2.3.2 *Genetic Algorithm Clustering and Routing (GACR)* [26]: is a genetic based algorithm for both clustering and routing. They are based on the distance of intra-cluster and inter cluster communication respectively, and the residual energy of CHs.
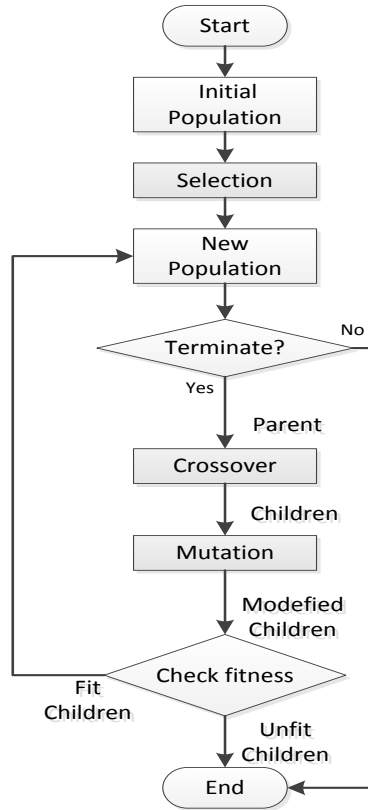
Figure 4. Genetic Algorithm flow chart.

The simulation results of the proposed algorithm have shown an improvement in energy consumption, number of active nodes, first gateway die and the number of dead gateway per round compared to GA based clustering algorithm, LDC and PSO based algorithm.

They have also shown the capability of the proposed routing algorithm in building a trade-off between hop counts and the total distance covered in a round, this performance compared with MHRM and GAR.

One of the weak points in this article that it is not considering the dynamic situation in which any sensor node or gateway may fail to work.

2.4 Social Insect

The similarity between WSN and the social structures of social insects are chosen as the sources of inspiration in this study [27], because of their massive number of simple individuals and decentralized control mechanisms. Unlike ACO which used in general purpose, it used to locate an optimum point in search spaces by releasing agents to perform random search and leave trails of artificial *pheromones* into the search space to lead to some optimum points, the proposed algorithm designed specifically for the clustering in WSN as exchanging the artificial *pheromones* among individuals in a multi-hop manner [28].

To balance the load on CH nodes caused by the long distance communications between them and BS, the proposed algorithm reduced the intra-cluster communication distance by arranging the nodes of a cluster in multiple layers. These layers separated carefully to allow CHs having almost the same number of the direct connection nodes.

A higher residual power node gets the priority in the random selection process of CHs in the network. The performance of the proposed algorithm evaluated against three existing algorithms LEACH, PEGASIS and PEDAP. The simulation results showed that:
- The Social Insect based algorithm reduced the energy consumption of the network, thus giving a much longer network lifetime.
- The using of multi-layers in arranging the network reduced the number of nodes which connect directly to the CH with respect to two-hop networks. This has improved network coverage in terms of avoiding the heavy load on some CHs led to early expiration.
- The time taken in the data collection is shortened due to the possibility of interleaving provided by the optimum distribution of the nodes in the multiple layer formation.

2.5 Bacteria Foraging Algorithm (BFA)

Bacteria Foraging Algorithm (BFA) [29]: is a novel algorithm for optimal sensor node deployment leading to optimal clustering of WSN nodes is presented. Using a bacteria foraging algorithm for optimizing the locations are adjusted such that all the nodes in the network moves to vertices of regular hexagons connected with each other. This leads to complete coverage of the area and all the nodes are equidistant.

Most of WSN created by improper deployment where nodes deployed randomly in the network field, this produce uncovered portion of the network.

The purpose of proposed algorithm is to bring the nodes on vertices of a regular hexagon and the cluster head placed in the center.

Steps of the proposed algorithm are:
1. Any node X with location (x1,y1) begins the algorithm.
2. It creates a virtual regular hexagon around itself by calculating coordinates of its vertices {v11, v12, v13, v14, v15, v16} and broadcast them to all the nodes in its radius r.
3. This is the objective function of the bacteria foraging algorithm that all nodes try to reach to the vertices of the hexagons.
4. Any six nodes decided by X moves to the vertices performing the swimming step of Chemotaxis.
5. Remaining nodes move out of the range *r* up to a random distance from hexagon performing tumbling step to change the cluster.
6. The message will be a directed signal in all six directions separated 60 degrees with each other from the center.
7. The vertices of the node X are moved into a queue since they have moved to their final location now they will be new center nodes and will make a cluster around them.
8. When all the nodes are in the queue, it means all the nodes have reached to some vertex then the chemotaxis converges.
9. If two nodes reach on the same vertex, then they reproduce a new node with double battery power, but single communication link as per BFO.
10. ID of one of the nodes is eliminated so to decrease extra communication without covering an extra area in the system.
11. Convergence of the algorithm will remove the redundancy of the nodes due to random deployment with their optimal placement positions this will also lead to usage of these nodes as extra battery power or to be located in the regions which are left without any node.
12. This unplaced area will be treated in the second pass of the algorithm when all the nodes have reached up to some vertices.
13. So new nodes will be created using reproduction with new locations.

The algorithm proved that it is optimal in the context of power utilization and minimum number of nodes to cover complete area. The needed movement of nodes to reach their optimal position is very less so it's not conflict with the power saving strategy in WSN. This method is very useful in case, deploying network nodes manually in the area.

2.6 Harmony Search Algorithm (HSA)

HSA is a music based optimization method [30] which inspired from the searching concept of musical instruments for the best harmony by polishing the pitches [31]. The main objective of deploying this algorithm is to minimize the intra-cluster distance and optimizing the energy consumption of the network. The selection of the CHs according to their residual energy of nodes –attached with data sent by each node to the BS – will be repeated in each round of data exchange.

HSA used to optimize the selection of CHs by finding optimal k CHs. BS sends information of CHs and which cluster to join to the rest of the nodes. After the forming of the clusters, nodes send their data to BS through their CHs.

The optimal cluster distribution of the nodes as a result of using HSA, showed a marked improvement in the average energy consumption, especially in large area networks (large-scale) because the optimal distribution of clusters in the network. It can also avoid the shortcoming of premature convergence of GA method and also the inability of PSO to maintain the desired levels of population diversity and the balance between local and global searches.

2.7 BAT Algorithm (BA)

2.7.1 *Bat algorithm (BA)* [32]: is a promising algorithm that looking forward to minimize the total communication distance and energy consumption basing on the loudness parameter of bats taking in consideration that the closest CHs to the BS are dying faster than other nodes in the network due to the heavy rely-rate occurring what called hotspot problem. To overcome this problem, network nodes clustered in different sizes known as "unequal clustering method". The author assumed that PSO and Harmony Search are the special cases of the Bat Algorithm. While the last one uses a combination of major advantages of PSO, GA and HSA.

The process of BA is depicted as follows:
- **Step 1**. Initialize the bat population, the pulse rates, the loudness, and define the pulse frequency.
- **Step 2**. Update the velocities to update the location of the bats, and decide whether detonate the random walk process.
- **Step 3.** Rank the bats, according to their fitness value, find the current near best solution found so far, and then update the loudness and the emission rate.
- **Step 4**. Check the termination condition to decide whether to go back to step 2 or end the process and output the result.

The simulation has implemented into two cases. First case, without taking in consideration hotspot problem where the BA-WSN compared to PSO-TVIW and PSO-TVAC and showed more convergence and accuracy. A second case, taking the hotspot problem into consideration where BA-WSNHS compared to BA-WSN and the first one showed better optimization than other methods because it is the most convergence among them.

2.7.2 The well-known (LEACH) algorithm has been optimized with BAT algorithm [33] to be used for clusters formation and CH selection. Pipelining is used for packet scheduling which either increases clock speed (sampling speed) or reduces power consumption at the same speed in a digital signal processor (DSP) system.

Due to BAT routing techniques can be adopted to work in highly challenging environments, it can improve the routing efficiency of the network.

The BAT algorithm inspired by the echolocation behaves that used by bats in tracking prey and developed three main rules:

(a) Bats use echolocation to sense distance, and they 'know' the difference between food/prey and background barriers in some way.
(b) Bats fly randomly with velocity v_i at position x_i with frequency f_{min} varying wavelength λ and loudness A_0 when searching for the prey. They adjust wavelength (or frequency) of emitted pulses automatically.
(c) They adjust the rate of pulse emission *are $\in$ [0:1]* depending on the proximity of the target; though loudness varies in many ways, it is assumed that loudness varies from a large (positive) A_0 to a minimum constant value A_{min}.

Simulation results showed that the proposed cluster based pipelining with BAT optimization performs better in terms of throughput, delay, Data drop and retransmission attempts in comparison with the existing algorithms.

2.8 Hybrid Bacteria Foraging using Particle Swarm Optimization

BFPSO LEACH-C [34]: is the applying of Hybrid Bacteria Foraging using Particle Swarm Optimization to the well-known clustering algorithm LEACH-C for the purpose of improving the lifetime of the network by increasing the number of alive nodes for a longer period of time and reducing the energy consumption.

The earlier proposal applied BFA to form the k-optimal to the LEACH-C protocol showed up a drawback in its need of iterative loops many times to track the global solution and consumes very long time for reaching the global position in the tumble behavior of the each bacterium which affect negatively on network lifetime. Hybridization the bacteria foraging particle swarm optimization algorithm to obtain the p_{best} and g_{best} location generated by PSO and these values utilized by each bacterium in tumble behavior inspired by the bacteria foraging to improve the network lifetime. The simulation results compared the proposed algorithm to the LEACH-C and BFA LEACH-C and showed the improvement in network life time and the reduction of the power consumption.

2.9 ANT Voronoi (VAS)

VAS [35] is a routing algorithm based on a functional combination of two popular approach i.e. Voronoi diagram and Ant System. The objective of it is to improve the network energy efficiency and lifetime.

Initially, the CH broadcasts a packet that contains the information about its position and Voronoi control packet VCP which contains the equations of the lines that bound the cluster area and the co-ordinates of the respective CH. Number of equations in the VCP depend upon the number of neighbor CHs. Nodes may receive VCPs from more than one CH. Based upon the VCP, all the nodes will identify which CH to follow.

The VAS algorithm described as follows:
1. Network Initialization, Node Initialization, Neighbor Initialization
2. Cluster formation:
   (a) CH run the Voronoi algorithm with neighbor CHs.
   (b) The information about the edges of the cluster is stored in the Voronoi Control Packet (VCP)
3. Run Voronoi Ant System on CH
   (d) Generate position information of the CH
   (e) Generate VAS packet that contains both 3(a) and 2(b)
   (f) Broadcast VAS with destination address as broadcast address
4. Every node that receives (VAS) packet:

    (a) Store position information of the CH in the routing table
    (b) Check the destination address as the broadcast address, if so broadcast it further
    (c) After the broadcast, update seen table
    (d) VCP is used to check its membership in a cluster
5. Repeat steps 2-4 till all the nodes are covered.

The performance of the VAS has been compared to AODF and OSPFv2 with varying data send rate, network area size, number of sensor nodes and number of CHs. Showed in the following tables (1: 4):

Table 1: In case of Varying the Data Send Rate

Algorithm	average delay	average jitter	data delivery ratio	Energy consumed	overhead in of number of bytes	overhead in number of packets
VAS	Better	Better	Better	Better	Well	Well
AODF	Bad	Bad	Bad	Well	Poor	Poor
OSPFv2	Worse	Worse	Worse	Well	Well	Well

Table 2: In case of Varying the Network Area Size

Algorithm	average delay	average jitter	data delivery ratio	Energy consumed	overhead in of number of bytes	overhead in number of packets
VAS	Poor	Poor	Poor	Better	Well	Better
AODF	Poor	Poor	Poor	Bad	Well	Well
OSPFv2	Poor	Poor	Poor	Worse	Worse	Worse

Table 3: In case of Varying the Number of Nodes

Algorithm	average delay	average jitter	data delivery ratio	Energy consumed	overhead in of number of bytes	overhead in number of packets
VAS	Better	Better	Better	Better	Better	Better
AODF	Well	Well	Well	Well	Worse	Worse
OSPFv2	Worse	Worse	Worse	Well	Bad	Bad

Table 4: In case of Varying the Number of CHs

Algorithm	average delay	average jitter	data delivery ratio	Energy consumed	overhead in of number of bytes	overhead in number of packets
VAS	Well	Well	Better	Better	Better	Better
AODF	Well	Well	Well	Well	Well	Well
OSPFv2	Well	Well	Worse	Well	Well	Well

3. PARAMETERS FOR COMPARISON OF EVOLUTIONARY ALGORITHMS

Different Evolutionary Algorithms reported in the literature for clustering in WSN have different assumptions about the network structure and parameters. This diversity can be attributed mainly to the lack of benchmarks and standards in WSN. Accordingly, in order to better compare and understand the differences, in this section, we present a set of WSN parameters. The parameters are explained and used in the following sections to compare the various Evolutionary Algorithms.

We classify the parameters into three main sets; namely, optimization characteristics, clustering characteristics, and network characteristics.

Optimization characteristics
- *Optimization Approach:* Many different Evolutionary Algorithms featured in the last few decades. They all exploited different nature, evolution technique looking for optimum solutions. Some recent published articles employed some these techniques used to optimize WSN clustering algorithms.

The Evolutionary Algorithms compared in this paper are: Ant colony optimization, Particle Swarm Optimization, Genetic Algorithm, Social Insect, Bacteria Foraging Algorithm, Harmony Search Algorithm, BAT Algorithm and Voronoi ANT System.

- *Optimization objective:* The limited resources of WSN node due to the need of minimizing their size and the fact that they are commonly deployed in isolated areas, energy efficiency remains the main challenge to extend the network lifetime. It can be represented in terms such as reducing the average energy consumption, increasing alive nodes over time and average fitness value. Many traditional algorithms applied to overcome this main issue. Few algorithms paid attention to other major issues such as: throughput, delay, coverage, robustness, link quality, data delivery ratio, overhead and etc. . .

Clustering characteristics [29,21]
- *Clustering Method:* If the algorithm is centralized the control message for cluster formation and CH selection will be received from the BS based on information gathered from all nodes to elect the appropriate number of clusters and the most efficient nodes to act as CHs every iteration but in case distributed algorithm these decisions will be taken by the nodes themselves. Hybrid scheme can be performed by a combination of centralized and distributed.

Network characteristics
- *Radio Model:* The mathematical calculation for each algorithm needs to assume a Radio Model which simulating the energy consumed by nodes in the network. The First order of Radio Model established by [3].However, most of the proposed algorithms depending on it in building their calculations.
  - *First Order Radio Model:* In this model, both the free space and multi-path fading channels are used depending on the distance (d) between the transmitter and receiver. When the distance is less than a threshold value (d_o), then the free space (fs) model is used, otherwise, the multipath (mp) model is used. Let E_{elec}, ε_{fs} and ε_{mp} be the energy required by the electronics circuit and by the amplifier in free space and multipath respectively. Then the energy required by the radio to transmit an l-bit message over a distance d is given as follows:

$$E_T(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2 & for \quad d < d_o \\ lE_{elec} + l\varepsilon_{mp}d^4 & for \quad d \geq d_o \end{cases} \qquad (1)$$

The threshold d_o is calculated as:

$$d_0 = \sqrt{E_{fn}/E_{mp}} \qquad (2)$$

The energy required by the radio to receive an l-bit message is given by:

$$E_R(l) = lE_{elec} \qquad (3)$$

The E_{elec} depends on several factors such as digital coding, modulation, filtering, and spreading of the signal, whereas the amplified energy, $\varepsilon_{fs}d^2/\varepsilon_{mp}d^4$, depends on the distance between the transmitter and the receiver and also on the acceptable bit-error

rate. Note that this is a simplified model which ignores listening energy consumption that known to be the largest contributor to expend energy in WSN.

- *CC2420 Radio Model:* one of the promised Evolutionary Algorithms discussed in this article that used a realistic energy consumption model which is based on the characteristics of the Chipcon CC2420 radio transceiver data sheet [36]. The total energy consumed by node *i*, E_i, is calculated as follows:

$$E_i = \sum_{state\,j} P_{state\,j} \times t_{state\,j} + \sum E_{transmitions} \qquad (4)$$

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, $E_{transmitions}$, is also added to the node's total energy consumption. The different values of $P_{state\,j}$ and $E_{transmitions}$ can be found in [8].

- *Mobility*: Some of the Evolutionary Algorithms assumed that the nodes are able to move from one position to another in order to optimize the distribution of nodes to contribute to the optimal clustering solution. On the other hand, most networks deployed in a random distribution of stationary nodes where the optimization tries to overcome the coverage problem without moving nodes.

- *Location awareness:* Some networks consist of nodes supplied with GPS chip or any other geographical location awareness method. This valuable information is used in different phases and decision in the WSN. But as a cost and size this is inefficient and unrealistic node characteristic.

- *Node capabilities:* In some Networks nodes are not all the same, heterogeneity can be in the energy, computation capability (data aggregation) or link communication (transmit power (range), bandwidth, location awareness, etc. . . Otherwise network can be built from Homogeneous nodes where all have the same capabilities. This is can be suitable in a traditional WSN. However, nowadays networks tend more to heterogeneity due to the complex interactions between multiple tasks different networks.

- *Number of nodes*: This parameter represents the total number of nodes in the network. Networks with large number of nodes (typically few hundreds) are becoming very common due to the emerging applications. Small networks are becoming increasing limited in terms of practical use.

4. COMPARISON OF EVOLUTIONARY ALGORITHMS
In this section we represent a comparison between most recent existed Evolutionary Algorithms based on different nine optimization approaches. The summary of the comparison is given in Table 5 below. From Table 5, we can deduce the following main observations:

- Most studies run the simulation on a few number of nodes. None of them applied for clustering in a large-scale WSN (thousands of nodes).
- Some studies compared to other Evolutionary Algorithms such as in [30,34,32]. And showed the superiority of some Evolutionary Algorithms than others, such as HSA, Hybrid PSO-BFA and Two-Tier PSO. This illustrates the need of applying some

modification on well-known Evolutionary Algorithms for expanding the optimization to the desired level.

- One study assumed a different radio model (CC2420 Radio Model) not the first order. Where it accuses the old model that it is unrealistic.
- Due to the complex simulation of the heterogeneous WSN, most algorithms assumed that all nodes are homogeneous (same parameters).
- None of the above algorithms assumed a hybrid method of clustering (combine of distributed and centralized).

5. CONCLUSION

In this paper, we present a survey of using Evolutionary Algorithms in clusters for hierarchical routing in WSNs. Nine different evolutionary techniques are identified and corresponding techniques from the literature are investigated and summarized.

Key findings include the following (summarize observations from Section 4). It can be seen that more research is needed in including security aspects as a factor in forming the clusters in WSNs.

Experimentation with large-scale WSN is still needed as most studies reported range from 10 to 500 nodes. In typical emerging WSNs, thousands of sensor nodes are expected in a single network.

No studies deal with special clustering based on sensor types. In particular, emerging and future applications of IoT and sensors include very heterogeneous sensors and devices in the same network. Variations in the energy models, transmission capability, etc. make conventional clustering techniques that assume homogeneous nodes with the similar capabilities inaccurate or inefficient.

Table 5. Comparison between most recent existed Evolutionary Algorithms for clustering in WSN

Algorithm	Optimization approach	Clustering method	Location awareness	Radio model	Nodes Capabilities	Clustering Objectives	Mobility	Optimization objective		No. of nodes
								Clustering	Routing	
AVAPCR-ACO [16]	Ant colony optimization (ACO)	Distributed	No	First order [3]	Homogeneous	Load balance Reduce delay Network lifetime Energy efficiency	No	Yes	Yes	100
PSOBC [20]	Particle Swarm Optimization (PSO)	Centralized	Yes	First order [3]	Homogeneous	Optimize intra cluster and inter cluster communication energy	No	Yes	No	100
TPSO-CR [23]		Centralized	No	CC2420 radio [36]	Homogeneous and Heterogeneous	Energy efficiency Link quality Network coverage	No	Yes	Yes	100-500
GCDC [25]	Genetic Algorithm (GA)	Centralized	Yes	First order [3]	Homogeneous	Reduces the energy depletion rate Network coverage	No	Yes	No	100, 1000
GACR [26]		Distributed	No	First order [3]	Homogeneous	Reduces the energy consumption	No	Yes	Yes	200-600
Proposed Algorithm [27]	Social Insect	Distributed	No	First order [3]	Homogeneous	Network lifetime Network coverage Data collection Time Robustness	No	Yes	No	300-500

EVOLUTIONARY ALGORITHMS FOR CLUSTERING IN WIRELESS SENSOR NETWORKS

Algorithm	Optimization approach	Clustering method	Location awareness	Radio model	Nodes Capabilities	Clustering Objectives	Mobility	Optimization objective		No. of nodes
								Clustering	Routing	
BFA [29]	Bacteria Foraging Algorithm (BFA)	Distributed	Yes	--	Homogeneous	Network coverage	Yes	Yes	No	-
HSA [30]	Harmony Search Algorithm (HSA)	Centralized	Yes	First order [3]	Homogeneous	Minimize the intra-cluster distance Reduce power consumption Extend lifetime	No	Yes	No	100
BA-WSN [32]	BAT Algorithm	Centralized	Yes	First order [3]	Homogeneous	More convergence and accurate	Yes	Yes	No	100
LEACH- BAT optimized [33]		Distributed	Yes	--	Heterogeneous	End – to - end delay and retransmission attempts	Yes	Yes	Yes	-
BFPSO LEACH-C [34]	Hybrid Bacteria Foraging – Particle Swarm Algorithm	Distributed	Yes	--	Homogeneous	Reduce the energy consumption Extend lifetime	No	Yes	No	20-100
VAS [35]	Voronoi ANT System (VAS)	Distributed	Yes	--	Homogeneous	Energy Efficiency Overall performance	No	No	Yes	10: 100

REFERENCES

1. O. Bello, O.Jumira, S.Zeadally, "Communication issues in the Internet of Things (IoT)", Journal of Network and Computer Applications, 2012.
2. D. Singh, G. Tripathi, A.J. Jara, "A survey of Internet-of-Things: Future vision, architecture, challenges and services," in Internet of Things (WF-IoT), 2014 IEEE World Forum on , vol., no., pp.287-292, 6-8 March 2014.
3. W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan: "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," IEEE Trans. Wireless Commun. 1(4), pp.660–770, 2002.
4. Jayavardhana GubbiAuthor Vitae, Rajkumar Buyya, Slaven MarusicAuthor Vitae, Marimuthu Palaniswami, Internet of Things (IoT): A vision, architectural elements, and future directions, Journal of Future Generation Computer Systems, Vol. 29, No. 7,pp.1645–1660, 2013.
5. Aditya Gaur, Bryan Scotney, Gerard Parr, Sally McClean, Smart City Architecture and its Applications Based on IoT, Procedia Computer Science, Volume 52, pp. 1089-1094, 2015.
6. Laura Belli, Simone Cirani, Luca Davoli, Andrea Gorrieri, Mirko Mancin, Marco Picone, Gianluigi Ferrari, "Design and Deployment of an IoT Application-Oriented Testbed", Computer, vol.48, no. 9, pp. 32-40, Sept. 2015.
7. D. Miorandi, S. Sicari , F. De Pellegrini , I. Chlamtac, "IoT: Vision, applications and research challenges", Ad Hoc Net. J., 1497-1516, 2012.
8. A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," ACM Computing Surveys (CSUR), vol. 31, no. 3, pp. 264–323, Sep.1999.
9. O. Boyinbode, Hanh Le, A. Mbogho, M. Takizawa, R. Poliah, "A Survey on Clustering Algorithms for Wireless Sensor Networks," in Network-Based Information Systems (NBiS), 2010 13th International Conference on , vol., no., pp.358-364, 14-16 Sept. 2010
10. Congfeng Jiang; Daomin Yuan; Yinghui Zhao, "Towards Clustering Algorithms in Wireless Sensor Networks-A Survey," in Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE , vol., no., pp.1-6, 5-8 April 2009
11. Sudeep Tanwar, Neeraj Kumar, Joel J.P.C. Rodrigues, A systematic review on heterogeneous routing protocols for wireless sensor network, Journal of Network and Computer Applications, Volume 53, July 2015, Pages 39-56.
12. Santar Pal Singh, S.C. Sharma, A Survey on Cluster Based Routing Protocols in Wireless Sensor Networks, Procedia Computer Science, Volume 45, 2015, Pages 687-69.
13. X. Liu, "A Survey on Clustering Routing Protocols in Wireless Sensor Networks," Sensors, vol. 12, pp.11113-11153, 2012.
14. C. Li, H. Zhang, B. Hao, and J. Li, "A Survey on Routing Protocols for Large-Scale Wireless Sensor Networks," Sensors, vol. 11, no. 4, pp. 3498-3526, 2011.
15. Kamal S. Hamza and Fathy Amir. 2016. Centralized Clustering Evolutionary Algorithms for Wireless Sensor Networks. Proceedings of the 10th International Conference on Informatics and Systems (INFOS '16). ACM, New York, NY, USA, 273-277.
16. D. Ma et al., "An adaptive virtual area partition clustering routing protocol using ant colony optimization for wireless sensor networks," in Advances in Wireless Sensor Networks, ser. Communications in Computer and Information Science, L. Sun et al., Eds. Springer Berlin Heidelberg, 2014, vol. 418, pp. 23–30.
17. M. Zambrano-Bigiarini, M. Clerc, R. Rojas, "Standard particle swarm optimization 2011 at cec-2013: a base line for future pso improvements," In: IEEE congresson evolutionary computation (CEC); p.2337–44. June2013
18. D.E. Jackson, M. Holcombe, F.F.W. Ratnieks, " Trail geometry gives polarity to ant foraging network," Nature 432(7019), pp. 907–909 (2004).
19. M. Dorigo, V. Maniezzo, A. Colorni, "Ant System: Optimization by a Colony of Cooperating Agents," IEEE Trans. Systems, Man, and Cybernetics-Part B 26(1), 29–41 (1996).

20. R. Yadav et al., "A discrete particle swarm optimization based clustering algorithm for wireless sensor networks," in Emerging ICT for Bridging the Future - Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2, ser. Advances in Intelligent Systems and Computing, S.C. Satapathy et al., Eds. Springer International Publishing, 2015, vol. 338, pp. 137–144.
21. R.C. Eberhart, J. Kennedy, "A new optimizer using particle swarm theory," In: Proceedings of the Sixth International Symposium on Micromachine and Human Science, Nagoya, Japan, pp. 39–43 (1995).
22. J. Kennedy, R.C. Eberhart, "Particle swarm optimization," In: Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ, pp. 1942–1948 (1995).
23. Riham S.Y. Elhabyan, Mustapha C.E. Yagoub, Two-tier particle swarm optimization protocol for clustering and routing in wireless sensor network, Journal of Network and Computer Applications, Volume 52, June 2015, Pages 116-128, ISSN 1084-8045, http://dx.doi.org/10.1016/j.jnca.2015.02.004.
24. E. D. Goldberg, "Genetic algorithms: Search optimization and machine learning," Massachusetts: Addison Wesley. (2007).
25. M. Kheireddine et al., "Genetic centralized dynamic clustering in wireless sensor networks," in Computer Science and Its Applications, ser. IFIP Advances in Information and Communication Technology, A. Amine et al., Eds. Springer International Publishing, 2015, vol. 456, pp. 503–511.
26. S. Gupta and P. Jana, "Energy efficient clustering and routing algorithms for wireless sensor networks: Ga based approach," Wireless Personal Communications, vol. 83, no. 3, pp. 2403–2423, 2015.
27. Chi-Tsun Cheng; Tse, C.K.; Lau, F.C.M., "A Clustering Algorithm for Wireless Sensor Networks Based on Social Insect Colonies," in Sensors Journal, IEEE , vol.11, no.3, pp.711-721, March 2011
28. J. B. Free, 'The Social Organization of Honeybees," London, U.K.: Edward Arnold (Publishers) Limited, 1977.
29. P. Nagchoudhury et al., "Optimal sensor nodes deployment method using bacteria foraging algorithm in wireless sensor networks," in Emerging ICT for Bridging the Future - Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2, ser. Advances in Intelligent Systems and Computing, S.C. Satapathy et al., Eds. Springer International Publishing, 2015, vol. 338, pp. 221–228.
30. D. C. Hoang, P. Yadav, R. Kumar, S. K. Panda, "A Robust Harmony Search Algorithm Based Clustering Protocol for Wireless Sensor Networks," in Communications Workshops (ICC), 2010 IEEE International Conference on , vol., no., pp.1-5, 23-27 May 2010
31. Z.W. Geem, J.H. Kim, G.V. Loganathan, "A new heuristic optimization algorithm: harmony search," Simulation Vol. 76 (2), pp. 60-68, 2001.
32. T.T. Nguyen et al., "Unequal clustering formation based on bat algorithm for wireless sensor networks," in Knowledge and Systems Engineering, ser. Advances in Intelligent Systems and Computing, V.H. Nguyen et al., Eds. Springer International Publishing, 2015, vol. 326, pp. 667–678.
33. K. Seelam et al., "An improved bat-optimized cluster-based routing for wireless sensor networks," in Intelligent Computing and Applications, ser. Advances in Intelligent Systems and Computing, D. Mandal et al., Eds. Springer India, 2015, vol. 343, pp. 115–126.
34. B. Pitchaimanickam, S. Radhakrishnan,"A hybrid bacteria foraging using Particle Swarm Optimization algorithm for clustering in wireless sensor networks," in Science Engineering and Management Research (ICSEMR), 2014 International Conference on , vol., no., pp.1-6, 27-29 Nov. 2014.
35. N. Gautam et al., "An ant voronoi based clustering approach for wireless sensor networks," in Ad Hoc Networks, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, M.H. Sherif et al., Eds. Springer International Publishing, 2014, vol. 129, pp. 32–46.
36. Texas Instruments, Chipcon CC2420 radio transceiver data sheet. (http://www.ti.com/lit/ds/symlink/cc2420.pdf); 2013 [accessed 16.11.2015].