A Novel Clustering Algorithm for Homogenous and Large-Scale Wireless Sensor Networks: Based on Sensor Nodes Deployment Location Coordinates

Hossein Jadidoleslamy

Information Technology Engineering Group, Department of Information Technology, Communications and Security, Malekashtar University of Technology (MUT), Tehran, Iran

Abstract
Wireless sensor networks (WSNs) are new generation of wireless networks which have many potential applications and unique challenges. Some of their most important problems are their hard organization and management, limited resources, dynamic topology and scalability. One significant technique against to these problems is clustering. Clustering leads to more scalability, energy efficiency, prolonged network lifetime and easy management of large-scale WSNs. As a result, this paper is focused on WSNs' clustering and it be represented a new centralized clustering algorithm for homogenous and large-scale WSNs. The proposed clustering algorithm is based on calculating distances between sensor nodes, Distance Average (DA) between each node and other nodes, finding distance average range, dividing it into K sub-range and selecting associated nodes of each sub-range. The proposed algorithm using of the WSN's sensor nodes deployment location coordinates. Also, all of these calculations will be done by Sink through using of mathematical and statistical formula, in centralized. This work enables us to verify the purpose and capabilities of the WSNs and clustering techniques. Also, it would enable WSNs' designers and managers to design, organize and manage WSNs, more significantly.

Keywords:
Wireless Sensor Network (WSN), Clustering, Sensor Nodes Coordinates, Distances, Distance Average, Centralized, Homogenous, Large-Scale.

1. Introduction

Wireless Sensor Networks (WSNs) are homogeneous or heterogeneous systems consist of many small devices, called sensor nodes, that monitoring different environments in cooperative; i.e. sensor nodes cooperate to each-others and combine their local data to reach a global view of the operational environment [1, 2, 6, 23]. In WSNs there are two other components, called "aggregation points" (i.e. Cluster-Heads deployment locations) and "base station" (i.e. the Sink deployment location), which have more powerful resources and capabilities than usual sensor nodes. As shown in Figure1, Cluster-Heads (CHs) collect information from their nearby sensor nodes, aggregate and forward them to the Sink to process gathered data [3, 4, 5, 10, 24]. Some of most important posed problems in WSNs are their hard organization and management, limited resources, dynamic topology and low scalability. In attending to these WSNs' constraints, clustering is a vital and complex requirement for these networks. Clustering leads to more scalability, energy efficiency, prolonged network lifetime and manageability in large-scale WSNs. As a result, this paper is focused on WSNs' clustering; the main purpose of this paper is presenting a centralized and light-weight clustering algorithm for homogenous and large-scale WSNs. The proposed clustering algorithm is including of following steps:

- Assumption: count of clusters = K; count of all sensor nodes = N;
- Calculating distances between sensor nodes to each-others (each node to other nodes);
- Calculating distance average of each sensor node to other nodes;
- Specifying range of distances' average: DARange = [minimum of distance average, maximum of distance average];
- Calculating variations' range of distances;
- Calculating average of variations rate of distances' ranges;
- Determining ranges of distance average;
- Determining sensor nodes of each range ⇨ it is proportional to determining members of each cluster;

The proposed algorithm is based on the WSN's sensor nodes deployment location coordinates. All of these calculations will be done by the Sink through using of mathematical and statistical relations, in centralized. This work enables us to verify the different aspects of WSNs and their clustering techniques; also it is proposing a new clustering algorithm which it would enable WSNs' designers and managers to design, organize and manage WSNs, more significantly. Also, by using of the proposed algorithm they will be able to solving the energy limitation and scalability problems of WSNs. The rest of this paper is organized as following: Section 2 expressed an overview of WSNs; Section 3 considering WSNs' clustering; section

Manuscript received February 5, 2014
Manuscript revised February 20, 2014

4 expressed the proposed clustering algorithm; section 5 is represented properties of the proposed clustering algorithm; section 6 represented case study; section 7 expressed reached results and conclusion; and finally future works, are drawn in section 8.

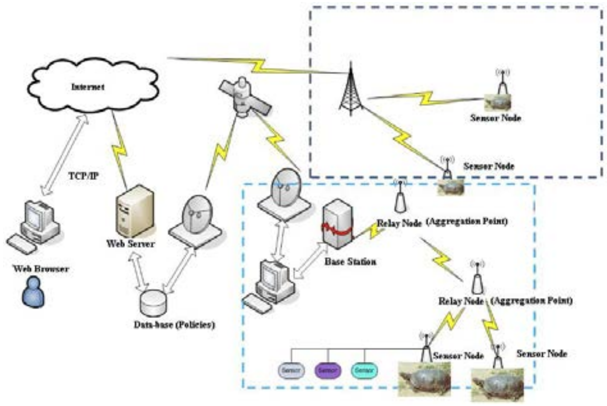

Fig. 1 WSNs' communication architecture

2. An Overview of WSNs

Wireless Sensor Network (WSN) is a wireless computer network with following major features:

- Infrastructure-less [1, 7, 8];
- No public address, often [1, 9, 34, 35] (data-centric);
- Consists of many tiny sensor nodes [2, 16, 17, 18] (small size, low-cost and low-power);
- Nodes distribution along with high-density in operational environment;
- Insecure radio links [3, 13, 14, 15, 26];
- Different communication models: hierarchical/distributed WSNs; or homogenous/heterogeneous WSNs [2, 22, 23, 25, 27];
- Limited resources of sensors [2, 28, 29, 32] (radio range, bandwidth, energy, memory and processing capabilities);
- Main application domains including of monitoring and tracking [4, 19, 20, 33] (as shown in Figure3);

In continue of this section, it is represented sensor node's architecture (by using of Figure2) and an outline of different aspects of WSNs, such as their characteristics, applications, communication architectures and vulnerabilities (according to Figure3).

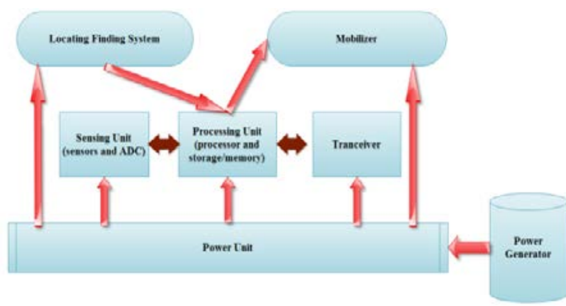

Fig. 2 Sensor node's architecture

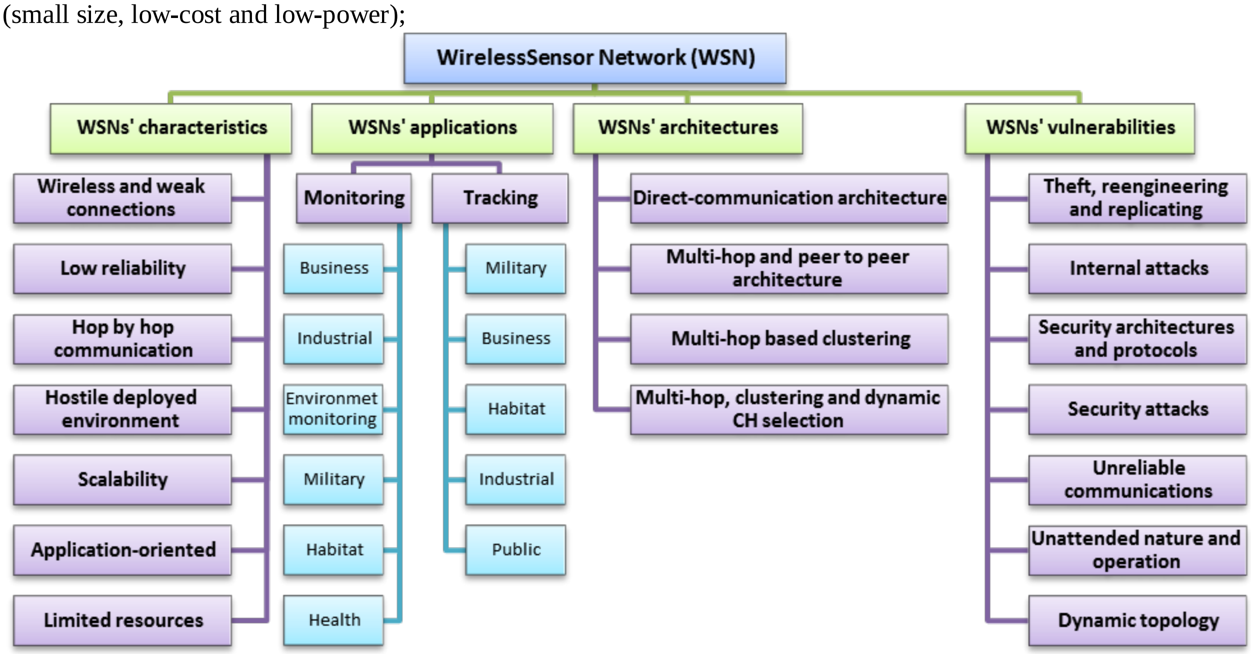

Fig. 3 WSNs' different aspects

3. Clustering on WSNs

Clustering means dividing the WSN's sensor nodes in virtual group (called cluster) according to some rules and then, sensor nodes belonging in a group can execute different functions than other nodes [9, 11, 17]. As Figure4 is shown, clustering is involving grouping nodes into clusters and selecting a Cluster-Head (CH) [8, 18, 19, 29] as following,

- Members of a cluster can communicate with their CH directly or multi-hop (hop by hop);
- CH can forward the aggregated data to the Sink through other CHs or directly [9];

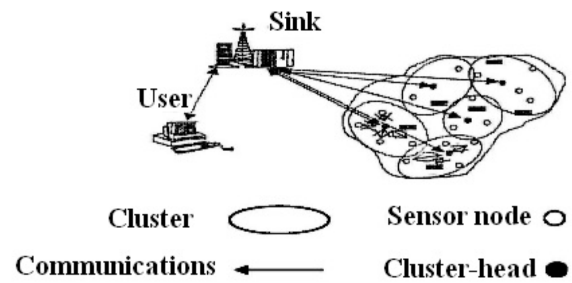

Fig. 4 Clustering overview

According to the Figure4, in high-level approach, clustering algorithms have three main phases, including of [9, 12, 21]: clusters formation phase, selecting CHs and stable-state phase. Also, required time interval for phase1 and phase2 is too less than phase3 [9, 30, 31]. In continue of this section, following figure (Figure5) represents goals and necessity of WSNs' clustering algorithms.

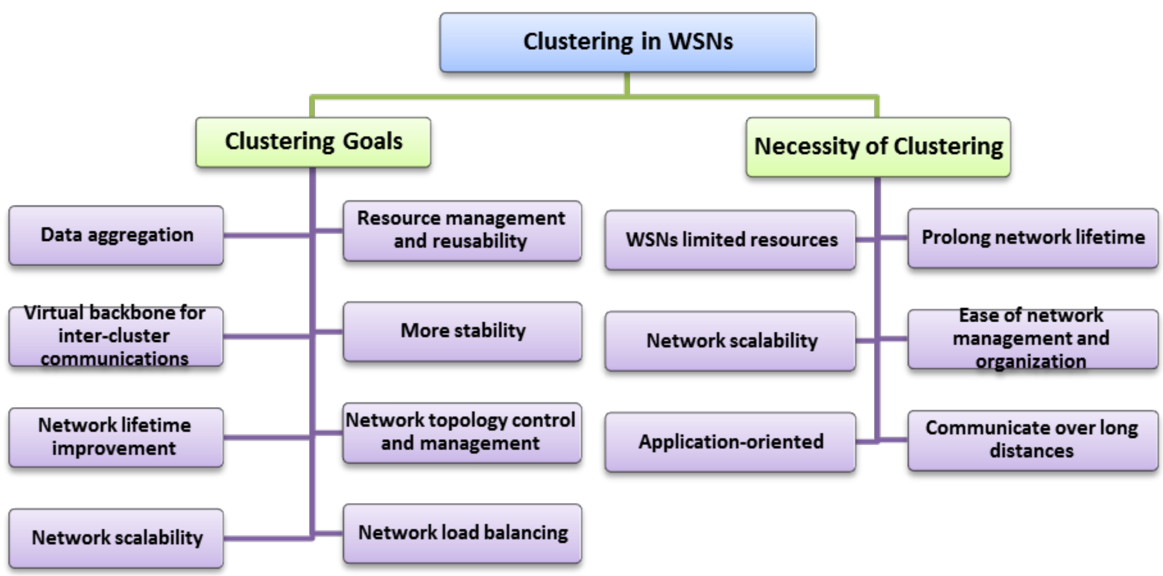

Fig. 5 Goals and necessity of clustering in WSNs

4. The Proposed Clustering Algorithm: Based on Coordinates of Deployment Location of Sensor Nodes, Distances between Sensor Nodes and Their Distance Average Calculations

The proposed clustering algorithm using of this idea: "sensor nodes which have similar distance average to other nodes, will be putted into a cluster". It is based on following calculations:

- Main assumption: count of clusters = K; count of sensor nodes = N;
- Calculating distances between sensor nodes to each-others (each node to other nodes);
- Calculating distance average of each sensor node to other nodes;
- Specifying range of distances' average: DARange = [minimum of distance average, maximum of distance average];
- Calculating variations' range of distances;
- Calculating average of variations rate of distances' ranges;
- Determining ranges of distance average;

- Determining sensor nodes of each range ⇨ it is proportional to determining members of each cluster;

The proposed method using of the WSN's sensor nodes coordinates. Also, all of these will be calculated by Sink through using of mathematical and statistical relations, in centralized.

1.1. Algorithm of The Proposed Clustering Algorithm

The proposed clustering algorithm operates as following steps.

1.1.1. Step1

Selecting radix node and allocating coordinates to other nodes;

1.1.2. Step2

Calculating distances between sensor nodes to each-others (each node to other nodes). In other words, calculating distances between node1 and other nodes, distances between node2 and other nodes and etc.

- D(i,j) = D(j,i) = Distance between node(i) and node(j) $= \sqrt{(\text{node}[j].\text{length} - \text{node}[i].\text{length})^2 + (\text{node}[j].\text{width} - \text{node}[i].\text{width})^2}$;

1.1.3. Step3

Calculating distance average between each node and other nodes; as following:

- Distance Average = DA = $\frac{(\text{sum of distances between a node and other nodes})}{N}$;
- Distance average between node i and other nodes = $DAi = \frac{\sum_{j=1}^{N} D(i,j)}{N} = \frac{D(i,1)+D(i,2)+\dots+D(i,N)}{N}$;

1.1.4. Step4

Specifying range of distances' average;

- DARange = [minimum of distance average, maximum of distance average] = [MinDA, MaxDA];
- Minimum of distance average = MinDA
- Maximum of Distance average = MaxDA

Calculating variations' range of distances as following:

- Distances' variations range = DVR = MaxDA – MinDA;

Calculating average of variations rate of distances' ranges,

- C = Dividing the distances' variation range to K = $\frac{\text{distances' variations range}}{K} = \frac{DVR}{K}$;

1.1.5. Step5

Determining ranges of distance average, as following:

- Range 1: [MinDA, MinDA+C); ⇨ Cluster 1
- Range 2: [MinDA+C, MinDA+2C); ⇨ Cluster 2
- …
- Range K: [MinDA+(K-1)C, MinDA+K.C] = [MinDA+(K-1)C, MaxDA]; ⇨ Cluster K

1.1.6. Step6

Determining sensor nodes of each range based on previous steps ⇨ it is proportional to determining members of each cluster (for example, sensor nodes of range1, they are members of cluster1).

1.2. Statistical Analysis Criteria of The Proposed Clustering Algorithm

This section is representing most important statistical criteria of evaluating the proposed clustering algorithm; as following:

- Drawing frequency table of clusters' nodes;
- Drawing diagrams such as column diagram to showing frequency of each cluster and their comparison;
- Calculating each cluster nodes' frequency, average of nodes' frequency, variance, standard deviation and coefficient of variation:
    - Frequency = count of existent nodes into each cluster or range;
    - Average of nodes' frequency = Average = $\frac{\text{sum of frequencies}}{\text{count of clusters}} = \frac{\text{sum of all nodes}}{K}$;
    - X = count of nodes of each cluster or range;
    - Variance (X) = Var (X) = σ^2 (X) = $\frac{\sum_{i=1}^{K}(Xi - Average)^2}{K}$;
    - Standard Deviation = SD = σ ;
    - Coefficient of Variation = CV = $\frac{\text{Standard Deviation}}{\text{Average}} = \frac{SD}{\text{Average}} = \frac{\sigma}{\text{Average}}$;

1.3. Pseudocode of The Proposed Clustering Algorithm

// Public assumptions
node[] nodescoordinates = new node [N]; // A matrix including of coordinates of all sensor nodes

int N; // Count of all sensor nodes including of $\{S_1, S_2, S_3, \dots, S_N\}$

int K; // Count of Cluster-Heads (CHs) or clusters, as: K<<N; including of $\{CH_1, CH_2, \dots, CH_k\}$ which they organize clusters $\{C_1, C_2, \dots, C_k\}$

int n_{max}; // Maximum count of nodes into different clusters

node[][] clustersnodescoordinates = new node [K][n_{max}]; // A matrix including of coordinates of sensor nodes of each cluster, separately

int [] clusterscountnodes = new int[K]; // An array including of different clusters' nodes count, i.e. $\{n_1, n_2, n_3, \dots, n_k\}$

Class node

{

float length;

float width;

Public static void node()

{

length = this.length;

width = this.width;

}

}

// Step1

point radix-node = new point();

radix-node.length = 1;

radix-node.width = 1;

// All nodes coordinates allocation

for (int i = 1; i ≤ N; i++)

{

Calculate and Allocate (nodescoordinates[i].length);

Calculate and Allocate (nodescoordinates[i].width);

}

// Step2

float [][] distancenodesmatrix = new float [N][N]; // A matrix consist of existent distances between nodes;

// distancenodesmatrix[i][j] = distance between node[i] and node[j]= distance (node[i], node[j]) = $\sqrt{\begin{matrix}(\mathbf{node[j].length - node[i].length})^2 + \\ (\mathbf{node[j].width - node[i].width})^2\end{matrix}}$;

for (int i=1; i≤N; i++)
for (int j=1; j≤N; j++)
distancenodesmatrix[i][j] = Math.SQRT
(
$\begin{pmatrix}\text{nodescoordinates[j].length} - \\ \text{nodescoordinates[i].length}\end{pmatrix}^2 +$
$(\text{nodescoordinates[j].width} - \text{nodescoordinates[i].width})^2$
) ;

// Step3

float[] DA = new float [N]; // An array is including of average of distances between nodes

for (int i=1; i≤N; i++)
float sumdistances = 0;
{
for (int j=1; j≤N; j++)
sumdistances+= distancenodesmatrix[i][j];
DA[i] = $\frac{\text{sumdistances}}{N}$;
}

// Step4

float MinDA, MaxDA;

```
MinDA = DA[1];
MaxDA = DA[1];
for (int i=2; i<=N;i++)
{
if (MinDA > DA[i])
MnDA = DA[i];
if (MaxDA<DA[i])
MaxDA = DA[i];
}

// Float DARange = [MinDA, MaxDA];
float DVR = MaxDA – MinDA;
float C = DVR/K ;

// Step5
// Determining ranges of distance average

float[][] DARangesmatrix = new float [K][2];
DARangesmatrix[1][1] = MinDA;
DARangesmatrix [1][2] = MinDA+C;
for (int i=2; i<=K;i++)
{
DARangesmatrix[i][1] = DARangesmatrix[i-1][2];
DARangesmatrix[i][2] = DARangesmatrix[i][1]+C;
}

// Step6
// Determining sensor nodes of each range ⇨ it is proportional to members of associated cluster
// For each range/cluster, for example range1

for (int i=1; i<=N; i++)
if (DA[i]>= DARangesmatrix[1][1] and DA[i]<=DARangesmatrix[1][2])
Node[i] is belonging to the cluster1 or range1;
```

5. Different Properties of The Proposed Clustering Algorithm

Following tables are representing some of most important properties of the proposed clustering algorithm, including of: special characteristics, characteristics of the operational environment (WSN), necessities of using the proposed method (Table1), its goals and applications (Table2), the proposed algorithm's advantages and weaknesses (Table3).

Table 1: The proposed clustering algorithm's especial characteristics, its necessity and the operational environment's characteristics

No.	Properties	Description
1	**Special characteristics of the proposed clustering algorithm**	• High accuracy clustering; due to using of mathematical and statistical formulas; • Using of quantitative variables which they are measurable by statistical and mathematical methods; • The proposed approach can be used in both static and dynamic WSNs (especially static); • Centralized-nature (in management and calculations); • Execution time complexity: : O (N^2); • High computational overhead and execution time complexity; • Using of deterministic-nature criteria and method (not probabilistic); • Sensor nodes' clustering method: calculating distances between sensor nodes to each-others, calculating average-distances, finding nodes which have similar distance average and then, selecting those nodes as associated cluster's members;
2	**Characteristics of the operational environment (WSN)**	• Homogenous; • Centralized management and calculations by sink; • Communicational architecture: clustering based single-hop and multi-hop communications; • Information transfer model: combinational (time-driven, event-driven or query-driven);
3	**Necessities of using the proposed clustering algorithm**	• Centrality and load balancing ⇨ increasing the WSN lifetime; • Reducing signal interference between CHs ⇨ the WSN resources' maintenance ⇨ the WSN lifetime prolonged; • More regularity, hierarchical organizing and ease of WSNs management; • High flexibility: supporting, usable and customizable for different operational environment;

Table 2: Goals and applications of the proposed clustering algorithm

No.	Properties	Description
1	**Goals of the proposed clustering algorithm**	• Data aggregation⇨ gathered and transmitted data redundancy reduction; • Resource management and reusability; • Virtual backbone for inter-cluster communication; • More stability; • Network lifetime improvement; • Network topology control and management; • Network scalability; • Network load balancing;
2	**Applications of the proposed clustering algorithm**	• Evaluating clustering algorithms in more efficiently and precisely; • Data gathering protocols; • Data aggregation protocols; • Hierarchical routing protocols; • WSNs' organization and centralized management; • Balanced and distributed energy consumption;

Table 3: The proposed clustering algorithm's advantages and its weaknesses

No.	Properties	Description
1	**Advantages of the proposed clustering algorithm**	• WSNs' scalability; • Load distribution in whole the WSN, in fairness; • Usually centralized calculations and management by Sink; • Supporting hierarchical and flat communication architectures; • Usually using of mathematics criteria ⇨ high accuracy; • Possibility of data aggregation ⇨ reducing volume of transmitted data; • Deterministic-nature (not probabilistic);
2	**Disadvantages of the proposed clustering algorithm**	• Low fault tolerant: by attention to high dependency to Sink for calculations, if the Sink be disabled, the proposed approach will be disrupted; • High execution time complexity and sophisticated computations in some steps ⇨ Necessity of using power and strong Sink; • Imposing almost high computational overhead; • Imposing delay to WSNs to clustering nodes, at first; • High dependency to coordinates of sensor nodes' deployment location;

2. Result: Case study

General assumptions are:

- Count of nodes: N = 9;
- Operational environment dimensions = 6 m * 4 m;
- Cluster radius = CR = 2 m;
- Sensor nodes are distributed in operational environment, randomly; supposing they are deployed on following coordinates (according to Figure6):

Node1 coordinates: (3,4); Node2 coordinates: (4,4); Node3 coordinates: (5,1);
Node4 coordinates: (4,3); Node5 coordinates: (3,3); Node6 coordinates: (2,4);
Node7 coordinates: (2,3); Node8 coordinates: (4,1); Node9 coordinates: (6,2);

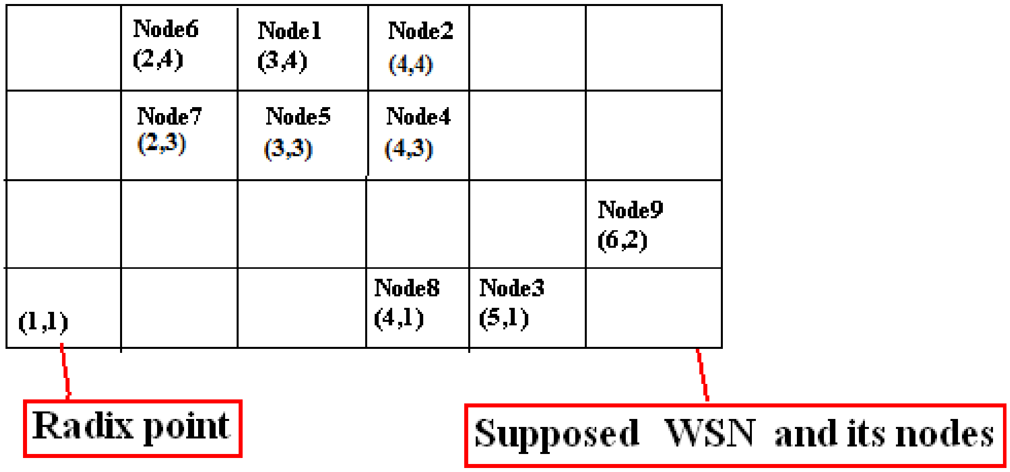

Fig. 6 The operational environment and its sensor nodes distribution

2.1. Case study (1)

- Count of clusters: K = 2;
- Distances between sensor nodes and their distance average (according to the Table4)

Table 4: Distances between sensor nodes and their distance average for case study (1)

Distance	Node1	Node2	Node3	Node4	Node5	Node6	Node7	Node8	Node9	Distance Average
Node1	0	1	$\sqrt{13}$	$\sqrt{2}$	1	1	$\sqrt{2}$	$\sqrt{10}$	$\sqrt{13}$	**1.80**
Node2	1	0	$\sqrt{10}$	1	$\sqrt{2}$	2	$\sqrt{5}$	3	$\sqrt{8}$	**1.85**
Node3	$\sqrt{13}$	$\sqrt{10}$	0	$\sqrt{5}$	$\sqrt{8}$	$\sqrt{18}$	$\sqrt{13}$	1	$\sqrt{2}$	**2.45**
Node4	$\sqrt{2}$	1	$\sqrt{5}$	0	1	$\sqrt{5}$	2	2	$\sqrt{5}$	**1.57**
Node5	1	$\sqrt{2}$	$\sqrt{8}$	1	0	$\sqrt{2}$	1	$\sqrt{5}$	$\sqrt{10}$	**1.56**
Node6	1	2	$\sqrt{18}$	$\sqrt{5}$	$\sqrt{2}$	0	1	$\sqrt{13}$	$\sqrt{20}$	**1.92**
Node7	$\sqrt{2}$	$\sqrt{5}$	$\sqrt{13}$	2	1	1	0	$\sqrt{8}$	$\sqrt{17}$	**2.02**
Node8	$\sqrt{10}$	3	1	2	$\sqrt{5}$	$\sqrt{13}$	$\sqrt{8}$	0	$\sqrt{5}$	**2.23**
Node9	$\sqrt{13}$	$\sqrt{8}$	$\sqrt{2}$	$\sqrt{5}$	$\sqrt{10}$	$\sqrt{20}$	$\sqrt{17}$	$\sqrt{5}$	0	**2.53**

- MinDA = 1.56; MaxDA = 2.53; DARange = [1.56, 2.53];
- Distances' variation range = 2.53 – 1.56 = 0.97
- $C = \frac{0.97}{2} = 0.48$
- Ranges of distance average are as following:
  - Range 1: [1.56, 2.04);
  - Range 2: [2.04, 2.53];
- Clustering (as shown in Figure7):
  - Members of cluster1 = {Node1, Node2, Node4, Node5, Node6, Node7}
  - Members of cluster2 = {Node3, node8, Node9}

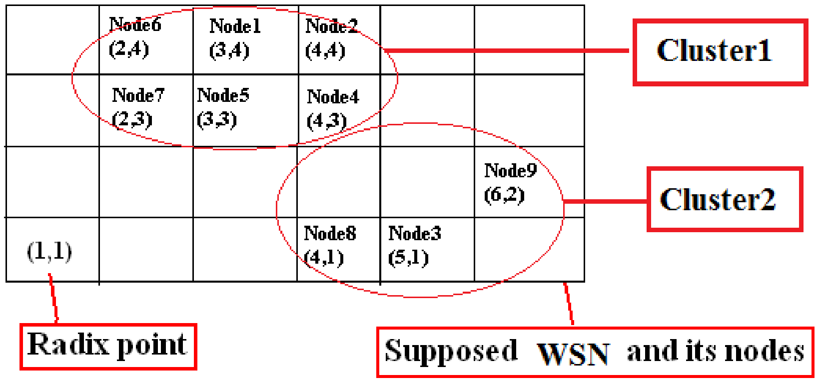

Fig. 7 Clustering the WSN by using of the proposed clustering algorithm for case study (1)

2.1.1. Statistical analysis

Table 5: Frequency table of case study (1)

Cluster	DARange	Frequency of nodes	Percentage
Cluster1	[1.56, 2.04)	6	66.66 %
Cluster2	[2.04, 2.53]	3	33.34 %

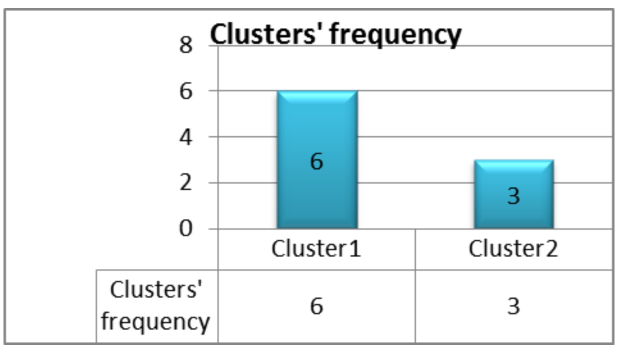

Fig. 8 Column diagram of case study (1)

- Average = $\frac{\text{sum of frequencies}}{\text{count of clusters}} = \frac{\text{sum of all nodes}}{K} = \frac{9}{2} = 4.50$;
- X = count of nodes of each clusters or ranges;
- Var (X) = σ^2 (X) = $\frac{\sum_{i=1}^{K}(Xi - Average)^2}{K} = \frac{(6-4.50)^2 + (3-4.50)^2}{2} = \frac{2.25+2.25}{2} = \frac{4.50}{2} = 2.25$;
- SD = $\sigma = \sqrt{2.25} = 1.50$;
- CV = $\frac{SD}{Average} = \frac{\sigma}{Average} = \frac{1.50}{4.50} = 0.33$;

As a result, the WSN is not uniform, the sensor nodes dispersion is not fairness and balanced; because, variation and SD are almost high. In other words, it is indicating rate of nodes' dispersion in different clusters is almost high; also, nodes' dispersion rather than average is much (i.e. nodes' distribution is not balanced and all clusters are not equal to each-others in term of count of nodes).

2.2. Case study (2)

- Count of clusters: K = 3;
- Distances between sensor nodes and their distance average (according to the Table6)

Table 6: Distances between sensor nodes and their distance average for case study (2)

Distance	Node1	Node2	Node3	Node4	Node5	Node6	Node7	Node8	Node9	Distance Average
Node1	0	1	$\sqrt{13}$	$\sqrt{2}$	1	1	$\sqrt{2}$	$\sqrt{10}$	$\sqrt{13}$	**1.80**
Node2	1	0	$\sqrt{10}$	1	$\sqrt{2}$	2	$\sqrt{5}$	3	$\sqrt{8}$	**1.85**
Node3	$\sqrt{13}$	$\sqrt{10}$	0	$\sqrt{5}$	$\sqrt{8}$	$\sqrt{18}$	$\sqrt{13}$	1	$\sqrt{2}$	**2.45**
Node4	$\sqrt{2}$	1	$\sqrt{5}$	0	1	$\sqrt{5}$	2	2	$\sqrt{5}$	**1.57**
Node5	1	$\sqrt{2}$	$\sqrt{8}$	1	0	$\sqrt{2}$	1	$\sqrt{5}$	$\sqrt{10}$	**1.56**
Node6	1	2	$\sqrt{18}$	$\sqrt{5}$	$\sqrt{2}$	0	1	$\sqrt{13}$	$\sqrt{20}$	**1.92**
Node7	$\sqrt{2}$	$\sqrt{5}$	$\sqrt{13}$	2	1	1	0	$\sqrt{8}$	$\sqrt{17}$	**2.02**
Node8	$\sqrt{10}$	3	1	2	$\sqrt{5}$	$\sqrt{13}$	$\sqrt{8}$	0	$\sqrt{5}$	**2.23**

Node9	$\sqrt{13}$	$\sqrt{8}$	$\sqrt{2}$	$\sqrt{5}$	$\sqrt{10}$	$\sqrt{20}$	$\sqrt{17}$	$\sqrt{5}$	0	**2.53**

- MinDA = 1.56; MaxDA = 2.53; DARange = [1.56, 2.53];
- Distances' variation range = 2.53 – 1.56 = 0.97
- C = $\frac{0.97}{3}$ = 0.32
- Ranges of distance average are as following:
  - Range 1: [1.56, 1.88);
  - Range 2: [1.88, 2.20);
  - Range 3: [2.20, 2.53];
- Clustering (as shown in Figure9):
  - Members of cluster1 = {Node1, Node2, Node4, Node5}
  - Members of cluster2 = {Node6, Node7}
  - Members of cluster3 = {Node3, Node8, Node9}

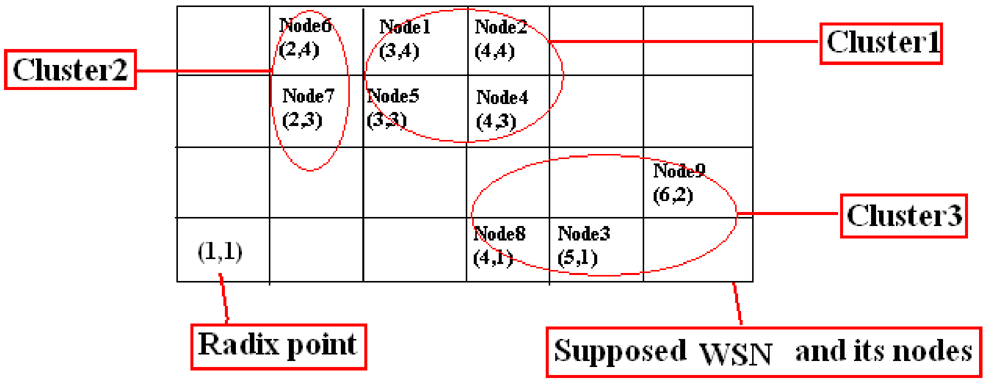

Fig. 9 Clustering the WSN by using of the proposed clustering algorithm for case study (2)

2.2.1. Statistical analysis

Table 7: Frequency table of case study (2)

Cluster	DARange	Frequency of nodes	Percentage
Cluster1	[1.56, 1.88)	4	44.44 %
Cluster2	[1.88, 2.20)	2	22.22 %
Cluster3	[2.20, 2.53]	3	33.34 %

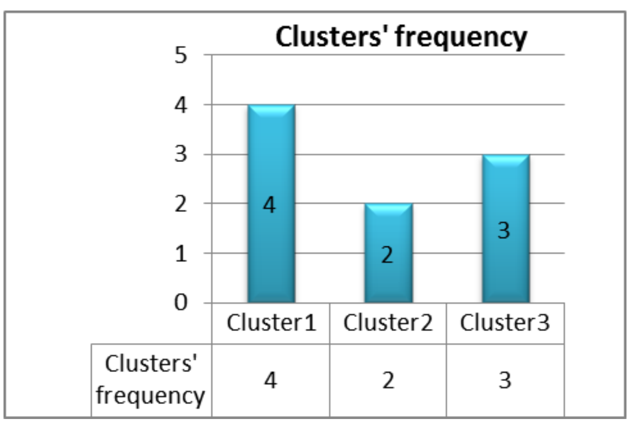

Fig. 10 Column diagram of case study (2)

- Average = $\frac{\text{sum of frequencies}}{\text{count of clusters}} = \frac{\text{sum of all nodes}}{K} = \frac{9}{3}$ = 3;
- X = count of nodes of each clusters or ranges;
- Var (X) = σ^2 (X) = $\frac{\sum_{i=1}^{K}(Xi - Average)^2}{K}$ = $\frac{(4-3)^2+(2-3)^2+(3-3)^2}{3} = \frac{1+1+0}{3}$ = 0.67;
- SD = σ = $\sqrt{0.67}$ = 0.82;
- CV = $\frac{SD}{Average} = \frac{\sigma}{Average} = \frac{0.82}{3}$ = 0.27;

As a result, the WSN is more uniform than case study (1), the nodes dispersion is more fairness and balanced than case study (1); because, variation and SD are less than case study (1). In other words, it is indicating rate of nodes' dispersion in different clusters is almost appropriate and nodes' dispersion rather than average is low (i.e. nodes' distribution is almost balanced and all clusters are almost equal to each-others in term of count of nodes).

6. Conclusion

This paper discussed on WSNs and their clustering algorithms. Also, it be represented a new centralized clustering algorithm for homogenous and large-scale WSNs along with its algorithm, pseudocode and properties such as its goals, advantages and weaknesses. In summarized, the proposed clustering algorithm is including of following steps:

- Main assumption: count of clusters = K; count of sensor nodes = N;
- Calculating distance average of each node to other nodes;
- Specifying range of distances' average: DARange = [minimum of distance average, maximum of distance average];
- Calculating variations' range of distances;
- Calculating average of variations rate of distances' ranges;
- Determining ranges of distance averages;
- Determining sensor nodes of each range ⇨ it is proportional to determining members of each cluster;

The proposed method using of the WSN's sensor nodes coordinates. Also, all of these calculations will be done by Sink through using of mathematical and statistical formulas, in centralized. Some of most important findings of this paper are as following:

- One of most important feature which leads to more uniform and balanced distribution of the energy consumption is clustering nodes ⇨ Like the proposed clustering algorithm;
- Execution time complexity of the proposed method is O(N^2) ⇨ The clustering algorithm has almost high computational overhead;
- The proposed algorithm supporting from static and dynamic clustering (especially static) ⇨ Leads to better load balanced, less energy consumption, balance energy distribution, network lifetime prolonged and high scalability;
- The proposed method supporting from multi-hop intra-cluster and inter-cluster communications; also, it usable on large-scale WSNs ⇨ Preserving energy efficiency independent of the network growth;
- The proposed approach using of qualitative and quantitative properties and variables (especially discrete quantitative variables) ⇨ Measurable by mathematical and statistical functions ⇨ High precise;
- If the operational environment be non-uniform, it is also possible to using of this clustering algorithm, but it should increase the count of clusters (K) ⇨ leads to more balanced nodes distribution (according to the case study (1) and case study (2));
- Clustering algorithms are different in many parameters such as their CH selection criteria, clusters formation methods and criteria, communication types, applications and architectures;
- Clustering has three main sections, including of clusters formation, CHs selection and stable-state phase. Required time for CH selection is too less than stable-state phase; i.e. Time (CH selection phase) << Time (stable-state phase);
- Calculating distances between sensor nodes to each-others (each node to other nodes);
- Some of most important challenges in clustering and CH selection of WSNs are:
  - How to divide the WSN into some of separated clusters;
  - Node cannot be member of two clusters (overlapping);
  - CHs communicates to the Sink directly or via other CHs (multi-hop);
- The proposed method is especially applicable for non-uniform WSNs which the sensor nodes are not distributed uniformly in whole of the WSN surface. In other words, sensor nodes' density is non-uniform;
- There is no universal clustering and CH-selection algorithm which fits to different operational environments, because clustering design decisions and CH selection methods are depending on application environments.

7. Future Works

There are several additional issues should be further studied in future research. Some of the most challenging and proposed topics of these issues are including of:

- Development of a generic method for finding the optimal number of clusters in order maximize the energy efficiency;
- Developing a method for estimation of the optimal frequency of CH re-selection to gain better energy distribution and efficiency;
- Discussing on clusters overlapping and boundary nodes;
- Presenting some energy efficient CHs selection methods;
- Proposing some new dynamic clustering techniques for WSNs;
- Proposing a new dynamic CH selection method based on distance between CH to (5% of) nodes which have least energy on that cluster;
- Designing an algorithm for CHs balanced distribution in whole of the WSN surface;
- Proposing new proper CH selection criteria;
- Combining proposed and existent criteria and then, representing hybrid clustering and CH selection algorithms;

References

[1] H. Jadidoleslamy; A High-level Architecture for Intrusion Detection on Heterogeneous Wireless Sensor Networks: Hierarchical, Scalable and Dynamic Reconfigurable; International Journal of Wireless Sensor Network (WSN); Vol. 3. No. 7, pp. 241-261; July, 2011.

[2] H. Jadidoleslamy; A Distributed and Hierarchical Intrusion Detection Architecture for Wireless Sensor Networks; International Journal of Network Security and its Applications (IJNSA); Vol. 3, No. 5, pp. 131-154; September, 2011.

[3] H. Jadidoleslamy; Designing an Agent-based Intrusion Detection System for Heterogeneous Wireless Sensor Networks: Robust, Fault Tolerant and Dynamic Reconfigurable; International Journal of Communications, Network and System Sciences (IJCNS); Vol. 4, No. 8, pp. 523-543; August, 2011.

[4] S. Mohammadi, R. A. Ebrahimi and H. Jadidoleslamy; A Comparison of Routing Attacks on Wireless Sensor Networks; International Journal of Information Assurance and Security (JIAS); ISSN 1554-1010 Volume 6, pp. 195-215; 2011.

[5] S. Mohammadi, R. A. Ebrahimi and H. Jadidoleslamy; A Comparison of Link Layer Attacks on Wireless Sensor Networks; Journal of Information Security (JIS); Vol. 2, No. 2, pp. 69-84; April, 2011.

[6] S. Mohammadi and H. Jadidoleslamy; A Comparison of Physical Attacks on Wireless Sensor Networks; International Journal of Peer to Peer Networks (IJP2P); Vol. 2, No. 2, pp. 24-42; April, 2011.

[7] S. Mohammadi and H. Jadidoleslamy; A Comparison of Transport and Application Layers Attacks on Wireless Sensor Networks; International Journal of Information Assurance and Security (JIAS); ISSN 1554-1010 Volume 6, pp. 331-345; April, 2011.

[8] H. Jadidoleslamy; A Comparison of Common Dynamic Cluster-Head Selection Clustering Techniques on Wireless Sensor Networks; International Journal of Wireless Sensor Network (WSN); Not published yet (Primary accepted); 2013.

[9] H. Jadidoleslamy; An Introduction to Various Basic Concepts of Clustering Techniques on Wireless Sensor Networks; International Journal of Computational Engineering Research (IJCER); Not published yet (Primary accepted); 2013.

[10] J. Yick, B. Mukherjee and D. Ghosal; Wireless Sensor Network Survey; Elsevier's Computer Networks Journal 52 (2292-2330); Department of Computer Science, University of California; 2008.

[11] A. A. Abbasi, M. Younis; A survey on clustering algorithms for wireless sensor networks; ComputerCommunications, vol. 30, pp. 2826–2841; June, 2007.

[12] O. Younis, M. Krunz, and S. Ramasubramanian; Node Clustering in Wireless Sensor Networks: Recent Developments and Deployment Challenges; IEEE Network; May/June, 2006.

[13] H. Taheri, P. Neamatollahi, M. Naghibzadeh, M. H. YaghmaeeMoghadam; Improving on HEED Protocol of Wireless Sensor Networks using Non Probablistic Approach and Fuzzy Logic (HEED-NPF); IST'10; Dec, 2010.

[14] G. Ran, H. Zhang, S. Gong; Improving on LEACH protocol of Wireless Sensor Networks Using Fuzzy Logic; Journal of Information &Comutational Science, pp. 767-775; 2010.

[15] M. C. M. Thein, T. Thein; An Energy Efficient Cluster-Head Selection for Wireless Sensor Networks; International Conference on Intelligent Systems, Modelingand Simulation, pp. 287-291; 2010.

[16] S. Bandyopadhyay and E. J. Coyle; An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks; IEEE INFOCOM; April, 2003.

[17] M. Liliana, C. Arboleda, and N. Nidal; Comparison of clustering algorithms and protocols for wireless sensor networks; in Proceedings of IEEE CCECE/CCGEI Conference, Ottawa, Ontario, Canada, pp. 1787–1792; May, 2006.

[18] V. Loscri, G. Morabito, and S. Marano; A two-level hierarchy for low-energy adaptive clustering hierarchy; in Proceedings of IEEE VTC Conference 2005, Vol. 3, pp. 1809–1813; 2005.

[19] A. Youssef, M. Younis, M. Youssef, and A. Agrawala; Distributed formation of overlapping multi-hop clusters in wireless sensor networks; in Proceedings of the 49th Annual IEEE Global Communication Conference (Globecom06), San Francisco, CA; November, 2006.

[20] M. Ye, C. Li, G. Chen, and J. Wu; EECS: An energy efficient clustering scheme in wireless sensor networks; in Proceedings of IEEE International Performance Computing and Communications Conference (IPCCC'05), pp. 535–540, April 7–9, Phoenix, AZ; 2005.

[21] Y. Jin, L. Wang, Y. Kim, and X. Yang; EEMC: An energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks; Computer Networks Journal, 52, 542–562; 2008.

[22] G. Li and T. Znati; RECA: A ring-structured energy-efficient clustering architecture for robust communication in wireless sensor networks; International Journal Sensor Networks, 2(1/2), 34–43; 2007.

[23] C. Liu, K. Wu, and J. Pei; A dynamic clustering and scheduling approach to energy saving in data collection from wireless sensor networks; in Proceedings of the 2nd Annual IEEE Conference on Sensor and Ad Hoc Communications and Networks(SECON'05), pp. 374–385; September, 2005.

[24] C. Wen and W. Sethares; Automatic decentralized clustering for WSNs; EURASIP Journal on Wireless Communications and Networking, 5(5), 686–697; 2005.

[25] K. Yanagihara, J. Taketsugu, K. Fukui, S. Fukunaga, S. Hara, and K.I. Kitayama; EACLE: Energy-aware clustering scheme with transmission power control for sensor networks; Wireless Personal Communications, 40(3), 401–415; 2007.

[26] S. Yi, J. Heo, Y. Cho, and J. Hong; PEACH: Power-efficient and adaptive clustering hierarchy protocol for WSNs; Computer Networks, 30, 2842–2852; 2007.

[27] P. Ding, J. Holliday, and A. Celik; Distributed energy efficient hierarchical clustering for wireless sensor networks; in Proceedings of the IEEE International Conference on Distributed Computing in Sensor Systems(DCOSS05), Marina Del Rey, CA; June, 2005.

[28] R. Virrankoski and A. Savvides; TASC: Topology adaptive clustering for sensor networks; in Proceedings of the Second IEEE International Conference on Mobile Ad-Hoc and Sensor Systems, (MASS 2005), Washington, DC; November, 2005.
[29] J. Kamimura, N. Wakamiya, and M. Murata; A distributed clustering method for energy-efficient data gathering in sensor networks; International Journal on Wireless and Mobile Computing, 1(2), 113–120; 2006.
[30] N.M. Abdul Latiff, C.C. Tsimenidis, and B.S. Sharif; Energy-aware clustering for wireless sensor networks using particle swarm optimization; in IEEE Intl. Symposium PIMRC'07, pp. 1–5, Athens, Greece; September, 2007.
[31] S. Yoon and C. Shahabi; Exploiting spatial correlation towards an energy efficient Clustered Aggregation technique (CAG); in Proceedings of IEEE International Conference on Communications (ICC'05), pp. 82–98; May, 2005.
[32] X. Zhang, H. Wang, and A. Khokhar; An energy-efficient data collection protocol for mobile sensor networks; in IEEE 64th VTC,Montreal, Quebec, Canada, pp. 1–5; 2006.
[33] K. Dasgupta, K. Kalpakis and P. Namjoshi; An efficient clustering– based heuristic for data gathering and aggregation in sensor networks; in: Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC); 2003.
[34] C. Duan, and H. Fan; A Distributed Energy Balance Clustering Protocol for Heterogeneous Wireless Sensor Networks; International Conference of Wireless Communications, Networking and Mobile Computing (WiCom), pp. 2469-2473; 2007.
[35] Z. Haibo, W. Yuanming and X. Guangzhong; EDFM: a stable election Protocol based on energy dissipation forecast method for clustered heterogeneous wireless sensor networks; Proceedings of the 5th International Conference on Wireless communications, networking and mobile computing, pp. 3544-3547; 2007.

Author Biography

H. Jadidoleslamy is a PhD student in Information Technology (IT)-Information Security at the University of MalekAshtar in Tehran, Iran. He received his Bachelor Degree in Information Technology (IT) engineering from the University of Sistan and Balouchestan (USB), Zahedan, Iran, in September 2009. He also has been received his Master of Science degree from the University of Guilan, Rasht, Iran, in March 2011. His research interests is including Computer Networks (especially Wireless Sensor Network), Information Security (by focusing on Intrusion Detection System), and E-Commerce. He may be reached at tanha.hossein@gmail.com or jadidoleslamy@gmail.com.