

A Secure, Hierarchical and Clustered Multipath Routing Protocol for Homogenous Wireless Sensor Networks: Based on the Numerical Taxonomy Technique

Hossein Jadidoleslami

Information Technology Engineering Group, Department of Technical and Engineering, University of Zabol (UOZ), Zabol, Sistan and Baluchestan, Iran

Abstract

Wireless Sensor Networks (WSNs) have many potential applications and unique challenges. Some problems of WSNs are: severe resources' constraints, low reliability and fault tolerant, low throughput, low scalability, low Quality of Service (QoS) and insecure operational environments. One significant solution against mentioned problems is hierarchical and clustering-based multipath routing. But, existent algorithms have many weaknesses such as: high overhead, security vulnerabilities, address-centric, low-scalability, permanent usage of optimal paths and severe resources' consumption. As a result, this paper is proposed an energy-aware, congestion-aware, location-based, data-centric, scalable, hierarchical and clustering-based multipath routing algorithm based on Numerical Taxonomy technique for homogenous WSNs. Finally, performance of the proposed algorithm has been compared with performance of LEACH routing algorithm; results of simulations and statistical-mathematical analysis are showing the proposed algorithm has been improved in terms of parameters like balanced resources' consumption such as energy and bandwidth, throughput, reliability and fault tolerant, accuracy, QoS such as average rate of packet delivery and WSNs' lifetime.

Keywords:

Wireless Sensor Network (WSN), Multipath Routing, Hierarchical and Clustered Routing, Trust-based Data Aggregation, Numerical Taxonomy Technique.

1. Introduction

Wireless Sensor Networks (WSNs) are consisting of many tiny sensor nodes. In hierarchical WSNs there are two other components, called Cluster-Heads (CHs) and Sink; they usually have more powerful resources and capabilities than usual sensor nodes. In Hierarchical WSNs, CHs gather data from their nearby sensor nodes, aggregating them and then, forwarding to the Sink [1, 2, 3]. Figure1 is representing different properties of WSNs. Some common problems of WSNs are: limited resources, low reliability and fault tolerant, low throughput, low scalability, low QoS and insecure operational environments. Hierarchical and clustering-based multipath routing is a significant solution to solving these problems. Multipath routing is discovering and using of multiple paths between source and destination nodes for data transmission; Figure2 is showing different properties of

multipath routing [4–28]. Existent multipath routing algorithms have many weaknesses such as: imposing high overhead to WSNs, security vulnerabilities, address-centric, low-scalability, permanent usage of optimal paths and severe resources' consumption. In other direction, WSNs have constraints in using of multipath routing such as severe resources' constraints, dynamic topology, high density (interference occurrence) and asymmetric and unbalanced traffic pattern. As a result, this paper will be discussed on multipath routing in WSNs. The main purpose of this paper is proposing an energy-aware, congestion-aware, location-based, data-centric, scalable, hierarchical and clustering-based multipath routing algorithm based on Numerical Taxonomy technique for homogenous WSNs. It leads to balanced and efficient resources' consumption such as energy and bandwidth and increasing bandwidth, throughput, reliability, fault tolerant and scalability; also, it results in improving security, Quality of Service (QoS) and WSNs' lifetime. Main phases of the proposed algorithm are as following:

- Clustering the WSN: centralized and location-based clusters' formation
- Multiple CHs selection based on Numerical Taxonomy technique
- Balanced traffic distribution between elected CHs (data segmentation)
- Secure and trust-based data aggregation
- Aggregated data transmission
- Routes maintenance and rediscovery

Finally, performance of the proposed algorithm has been compared with performance of LEACH routing algorithm; results of simulations and statistical-mathematical analysis are showing the proposed algorithm has been improved in terms of different criteria like balanced resources' consumption such as energy and bandwidth, throughput, reliability and fault tolerant, accuracy, QoS parameters like average rate of packet delivery and WSNs' lifetime. Rest of this paper has been organized as following: Section2 is represented proposed hierarchical and clustering-based multipath routing algorithm, in details; Section3 is

discussing on different properties of the proposed algorithm, including of: its assumptions, advantages, disadvantages and its especial and major characteristics; section4 is presented considered algorithm for comparison; i.e. it describes

LEACH routing algorithm; Section5 is including of results, analysis and evaluations (quantitative and qualitative comparison); and finally, Section6 is expressed conclusion and future directions.

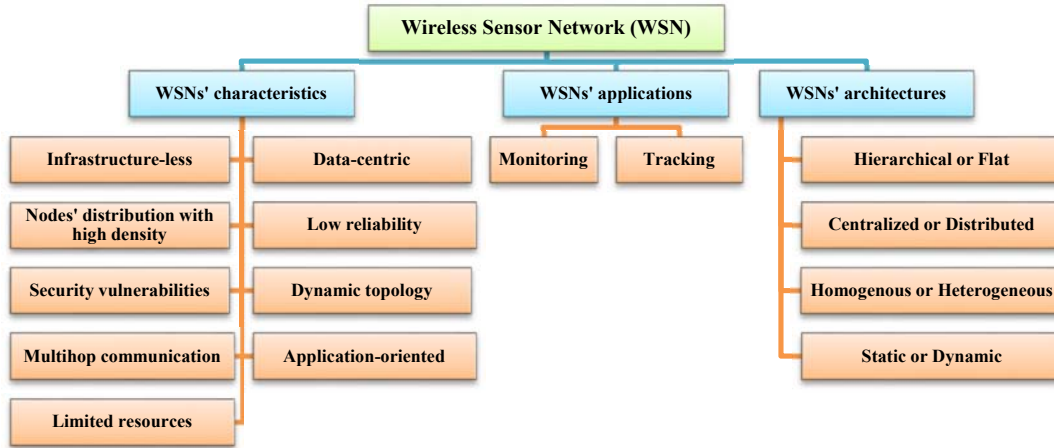


Fig. 1 An overview on WSNs: major properties, applications and architectures

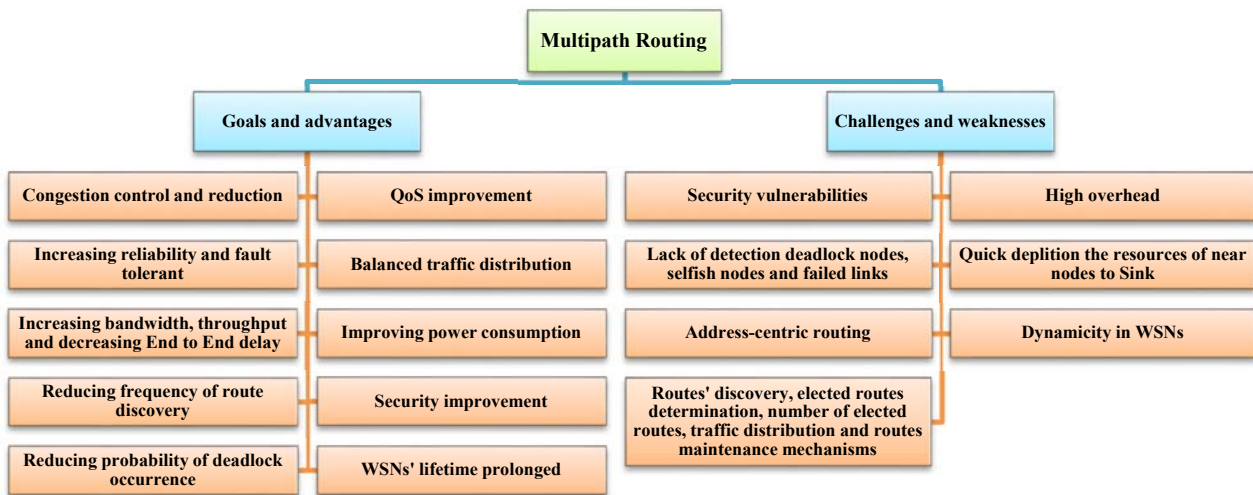


Fig. 2 Different properties of multipath routing: goals and challenges

1. Proposed Multipath Routing Algorithm

Main idea of the proposed algorithm is: time is divided into number of super-rounds; each super-round is including of a few time intervals. In each super-round, it clustering the WSN and selecting multiple Cluster-Heads (CHs) for each cluster based on Numerical Taxonomy technique. Now, notifying the CHs of each cluster to its members; then, determining elected CHs by each one of the cluster's members; elected CHs can be used simultaneously, during each super-round together (full-multipath routing) or elected CHs can be used separately, in each time interval; i.e. in each time interval, it using of one of elected CHs for data transmission to the Sink (partial-multipath routing). Now,

distributing ready data between elected CHs; finally, trust-based data aggregation and transferring aggregated data. In this protocol, following phases are repeated periodically, in each super-round.

1.1 Phase1: Clustering the WSN: Centralized and Location-based Clusters' Formation

Clustering means dividing and grouping the WSN's sensor nodes into virtual groups, called clusters, according to some rules and then, selecting a Cluster-Head (CH) for each cluster [29–35]. Idea of the proposed clustering algorithm is: "sensor nodes which have similar average distance to other nodes, they will be putted into a cluster". In the proposed algorithm, Sink is clustering the WSN by using of mathematical-

statistical relations and based on sensor nodes' deployment location coordinates, according to following steps:

- Selecting radix node and allocating coordinates to other nodes.
- Calculating Average Distance (AD) between each node and other nodes, as following:

$$AD = \frac{(\text{sum of distances between the node and other nodes})}{N};$$

$$AD_i = \frac{\sum_{j=1}^N D(i,j)}{N} = \frac{D(i,1)+D(i,2)+\dots+D(i,N)}{N} = \frac{\sum_{j=1}^N (\sqrt{(node_{j,length} - node_{i,length})^2 + (node_{j,width} - node_{i,width})^2})}{N}$$

N: number of sensor nodes

D(i,j): distance between node(i) and node(j)

AD_i: average distance between node(i) and other nodes

- Specifying Range of AD (RAD), variations' range of distances (VRD) and average variations rate of distances' ranges (C) as following:

RAD = [Minimum of AD, Maximum of AD] = [MinAD, MaxAD];

VRD = MaxAD – MinAD;

$$C = \frac{VRD}{K};$$

K: number of clusters

- Determining ranges of AD, as following:

Range₁: [MinAD, MinAD+C] ⇒ Cluster 1

Range₂: [MinAD+C, MinAD+2C] ⇒ Cluster 2

...

Range_K: [MinAD+(K-1) × C, MinAD+K × C] = [MinAD+(K-1) × C, MaxAD] ⇒ Cluster K

- Determining sensor nodes corresponding to each range; it is proportional to members of associated clusters (for example, sensor nodes of range₁, they are members of cluster1).

1.2 Phase2: Multi-criteria Multiple CHs' Selection for Each Cluster based on Numerical Taxonomy Technique

Numerical Taxonomy technique is a Multi-Criteria Decision Making (MCDM) method. It can be used in clustering process for CH selection; it can rank members of the cluster and then, selecting appropriate cluster's members as CHs [36–41]. So, purpose of this phase is proposing an algorithm to ranking the cluster's members and determining multiple CHs through evaluating their capability for data aggregation and transmission. For each cluster, the proposed multiple CHs selection algorithm is including of following major steps:

1) Determining number of required CHs for the cluster

It is based on criteria such as: volume of ready data for transmission, number of the cluster's members or nodes' density of the cluster, available resources such as energy and

bandwidth, amount of existent congestion such as length of waiting queue, size of empty buffer and delay.

2) Ranking the cluster's members and selecting multiple CHs

It is ranking the cluster's members based on the CHs' selection criteria and determining multiple CHs for the cluster by using of the Numerical Taxonomy technique; then, notifying elected CHs to other members of the cluster.

Proposed criteria for ranking the cluster's members to determining multiple elected CHs are as following:

- Remainder Energy (RE: +): if the remainder energy be less than a predefined threshold, its value is assumed zero (RE = 0).
- Distance to the Gravity Center (DGC: -)
- Distance to the Sink (DS: -)
- Average Distance (AD: -): average distance between the node and other cluster's members.
- Congestion (Con: -): it is calculable through parameters such as length of weighting queue and average delay in the queue.
- Size of empty buffer (Buff: +)
- Available bandwidth (BW: +)
- Deadlock (DL): it is showing if the node is a deadlock node or not; possible values for this criterion are: {0, 1}; it has no weight.
- Density around the node (Den: +): it is equal to number of its neighbors.

Note: '+' and '-' signs are indicating positive and negative criteria.

Note: If (DL = 0) or (RE = 0) or (Buff = 0) or (BW = 0) ⇒ Rank_{member} = 0

Algorithm of multi-criteria multiple CHs' selection based on Numerical Taxonomy technique is as following:

In first super-round, for each cluster, the proposed algorithm selects a few CHs, randomly; then, since second super-round, current CHs of each cluster select new multiple CHs for corresponding clusters by using of Numerical Taxonomy technique. For this purpose, there are following approaches:

- Every one of existent CHs selects one or more replacement CHs instead of itself.
- One of existent CHs selects all required multiple CHs for the cluster.

Now, every one of current CHs is following these steps to ranking its cluster's members and determining replacement CHs instead of itself. Steps of Numerical Taxonomy technique are as following:

Goal: Multi-criteria multiple CHs' selection for the cluster.

Step1: Defining criteria of ranking the cluster's members.

Step2: Cluster's Members Specifications Matrix (CMSM) formation: every one of the cluster's members sends its Specifications Notification Packet (SNP) to the current CH.

Then, current CH is constructing the Cluster's Members Specifications Matrix (CMSM) by using of received SNP control packets from its cluster's members. This matrix is including of values of every one of the cluster's members per above criteria. Number of its rows is equal to the cluster's members and number of its columns is equal to criteria. So, it is a $(n \times m)$ matrix, as:

- n : rows of the CMSM matrix: cluster's members;
- m : columns of the CMSM matrix: criteria;
- X_{ij} : score of node i with respect to criterion j ;

$$\text{CMSM} = \begin{bmatrix} X_{1,1} & \cdots & X_{1,m} \\ \vdots & \ddots & \vdots \\ X_{n,1} & \cdots & X_{n,m} \end{bmatrix}$$

- Calculating sum, average (\bar{X}) and Standard Deviation (SD) of each criterion, as following:

$$\text{Sum} = \sum_{j=1}^n X_j$$

$$\bar{X} = \frac{\sum_{j=1}^n X_j}{n} = \frac{\text{sum}}{n}$$

$$\text{SD} = \sqrt{\frac{\sum_{j=1}^n (X_j - \bar{X})^2}{n}}$$

Step3: Data homogenization and standard matrix formation in order of removing different measurement units and replacing a same scale; so, it leads to a scale-less matrix. In other words, constructing the normalized CMSM matrix by current CH through mapping values of CMSM's elements to the range $[0, 1]$; i.e.:

$$Y_j = \frac{X_j - \bar{X}}{\text{SD}}$$

$$B = \begin{bmatrix} Y_{1,1} & \cdots & Y_{1,m} \\ \vdots & \ddots & \vdots \\ Y_{n,1} & \cdots & Y_{n,m} \end{bmatrix}$$

In this step:

- Ideal value of each criterion = largest value of that criterion;
- SD of each criterion or column = 1;
- Average of each criterion or column = 0;

Step4: constructing differences' matrix or composite distances' matrix: calculating differences or composite distances between sensor nodes into a symmetric matrix;

$$D_{ab} = \sqrt{\sum_{j=1}^m (Y_{aj} - Y_{bj})^2}$$

$$C = \begin{bmatrix} D_{1,1} & \cdots & D_{1,m} \\ \vdots & \ddots & \vdots \\ D_{n,1} & \cdots & D_{n,m} \end{bmatrix}$$

$$\bar{D} = \frac{\sum_{j=1}^n D_j}{n}$$

$$\text{SD}_D = \sqrt{\frac{\sum_{j=1}^n (D_j - \bar{D})^2}{n}}$$

- Determining shortest distances into each row of the symmetric matrix;

Step5: Calculating upper limit and lower limit of distances to determining homogenous cluster's members; i.e. nodes which are into range of upper and lower limits, they are called

homogenous nodes. Nodes which are outside of this range, they are representing lack of similarity in posed criteria; so, they should be removed.

$$\bar{D} = \frac{\sum_{j=1}^n D_j}{n}$$

$$\text{SD}_D = \sqrt{\frac{\sum_{j=1}^n (D_j - \bar{D})^2}{n}}$$

$$D^- = \bar{D} - 2 \text{SD}_D$$

$$D^+ = \bar{D} + 2 \text{SD}_D$$

Step6: Determining ideal value of each criterion from the standard matrix: after removing the non-similar values, ideal value of each criterion is its largest value.

D_o = Ideal value = maximum value in matrix B

Step7: Calculating model of development (C_{io}), as following:

For each criterion: $(\text{Ideal value} - \text{standard value})^2 =$

$$(D_o - D_j)^2 \Rightarrow C_{io} = \sqrt{\sum_{j=1}^n (D_o - D_j)^2}$$

Whatever this parameter is smaller, that node has appropriate and desired criteria more than others.

Step8: Calculating degree of development (F_i) of nodes which they have the desired criteria, as following:

$$F_i = \frac{C_{io}}{C_o} = \frac{C_{io}}{C_{io} + 2 \text{SD } C_{io}}$$

$$C_o = C_{io} + 2 \text{SD } C_{io}$$

$$\bar{C}_{io} = \frac{\sum_{j=1}^n C_{io}}{n}$$

$$\text{SD}_{C_{io}} = \sqrt{\frac{\sum_{j=1}^n (C_{io} - \bar{C}_{io})^2}{n}}$$

Value of F_i is between 0 and 1 ($0 \leq F_i \leq 1$); whatever it is closer to zero, that node has those criteria more; i.e. that node is more developed and it has higher priority. So the cluster's members will be sorting in ascending order; after that, multiple CHs for the cluster will be selected. In other words, ranking the cluster's members based on ascending order of F_i values; then, selecting the cluster's members with F_i closer to zero as multiple elected CHs; i.e. cluster's members who their F_i is less, they are better to selecting as CHs.

Therefore, current CH knows capability of its cluster's members for data aggregation and transmission. Finally, it is storing calculated scores of the cluster's members, ranking them and selecting multiple elected CHs for the cluster.

Note: If count of the cluster's members be greater than a predefined threshold, it is better to breaking that cluster to some new smaller clusters (with less count of sensor nodes).

Note: Another approach is ranking the cluster's members by Sink and selecting multiple CHs; it is a centralized clustering (clusters' formation and multiple CHs selection) approach, including of following steps:

- Sink is constructing a decision making matrix for each cluster by using of the received CMSM matrix from corresponding CH.
- Normalizing above decision making matrix (mapping its values to range $[0, 1]$).

- Weighting to the normalized decision making matrix.
- Calculating score of every one of the cluster's members by using of the proposed ranking formula and existent information into the weighted normalized decision making matrix.
- Selecting the appropriate cluster's members as multiple elected CHs for the cluster.

1.3 Phase3: Balanced Traffic Distribution between Multiple Elected CHs: Data Segmentation

For this purpose, there are following approaches:

- Every one of the cluster's members selects one CH in each time interval/super-round and transferring its data through elected CH during that time interval/super-round.
- Every one of the cluster's members segments its data packets, selecting more than one CH and then, transferring its data through multiple CHs in each time interval/super-round; i.e. every one of the cluster's members selects several CHs in each time interval/super-round. For this purpose, like the previous phase, every one of the cluster's members is scoring and ranking CHs of the cluster. Then, it is distributing its data between elected CHs, in balanced; this distribution can be based on elected CHs' scores and number of elected CHs.

Besides, for determining criteria and proposing algorithm of ranking the cluster's CHs for balanced traffic distribution between them there are following approaches:

- **Approach (1):** normalizing and using of the previous phase's scores.
- **Approach (2):** weighting the criteria of previous phase, again and then, using of them as criteria of scoring the cluster's CHs for balanced traffic distribution; i.e. based on these criteria, every one of the cluster's members is scoring the cluster's CHs, selecting elected CHs and distributing its traffic between them, in balanced.
- **Approach (3):** defining new criteria and new formula for this purpose and then, weighting to the criteria and scoring the cluster's CHs; finally, determining elected CHs and distributing traffic between them, in balanced.
- **Approach (4):** every one of the cluster's members allocates some percentage of its data to every one of the cluster's CHs, randomly.

In this phase, radio antenna of the cluster's members is off and they usually do not sending and receiving data; it leads to reducing energy consumption.

Note: Percentage of allocated traffic to every one of elected CHs = $\frac{\text{Score of the CH}}{\text{Total scores of elected CHs}} \times 100$

Note: If volume of ready data be less than a predefined threshold, multipath routing and traffic distribution is not required. In this case, whole data will be forwarded through the optimal CH (single-path routing).

Note: Average number of elected CHs (C) by every one of the cluster's members for data transmission to the Sink is:

$$C = \frac{VRD}{ABWL} \times \frac{1}{LR}$$

ABWL: average bandwidth of each link (bps)

VRD: volume of ready data (bit)

LR: length of each time interval (second)

1.4 Phase4: Secure and Trust-based Data Aggregation

This phase is proposing a trust model for secure data aggregation. Each CH evaluates consistency and reliability of arrived data from its cluster's members and calculating their data' trust values; then, it is aggregating arrived data based on their trust values (trust is a number between 0 and 1). For data trust calculation, the proposed algorithm is assumed primary trust value of all the cluster's members is 0.5; every one of the cluster's members is reported its sensed data and result of its measurement (R) to the corresponding CH; then, the CH is aggregating incoming data according to following steps:

- Receiving data (measurement's result: R) from its cluster's members and calculating their average (R_{avg}).
- Calculating Absolute Deviation (AD) of reported value by every one of the cluster's members than R_{avg} and Average of Absolute Deviation (AAD), as following:

$$AD = |R_{avg} - R|$$

$AAD = \frac{\sum_{i=1}^{CCM} AD}{CCM}$, CCM: count of the cluster's members which reported data;

- Calculating reliability and consistency of reported data by the cluster's members as following:

$$C_o = \text{Maximum} \left\{ 1 - \frac{AD}{AD_{thr}}, 0 \right\}$$

$AD_{thr} = \alpha \times AAD$ = threshold of difference = if a cluster's member has been deviation equal or greater than AD_{thr} than the average value, its trust value is zero (0).

$\alpha > 0$: Tolerance factor of deviation;

- Now, trust value of every one of the cluster's members will be calculated as following:
 - Regardless of previous trust value: $T = C_o$
 - By attention to previous trust value: $T = W \times T_{Previous} + (1-W) \times T_{Current} = W \times T_{Previous} + (1-W) \times C_o$

W: weight of previous trust's value

- Determining threshold of acceptable trust value (T_{min}); the cluster's members which their trust's value is less than T_{min} , they are malicious, compromised or selfish nodes (M); i.e.:

A = set of all data reporting members

N = set of normal reporting members which they have acceptable trust's value

M = set of malicious reporting members = $\{i \mid i \in A, T_i < T_{min}\} \Rightarrow W_{Ri} = 0;$

So, malicious reporting members which they have no minimum acceptable trust's value, they do not participate in data aggregation process.

- Weight of reported data by every one of the reporting cluster's members (W_{Ri}) for secure and trust-based data aggregation is equal to its trust's value ($W_{Ri} = T_i$). So, aggregated value of the considered parameter is (R_{agg}):

$$R_{agg} = \frac{\sum_{i=1}^{|N|} W_{Ri} \times R_i}{|N|} = \frac{\sum_{i=1}^{|N|} T_i \times R_i}{|N|}, \quad |N|: \text{ number of normal reporting members.}$$

1.5 Phase5: Aggregated Data Transmission

In this phase, each CH transfers its aggregated data to the Sink.

1.6 Phase6: Routes Maintenance and Rediscovery: Repeating above Phases

The proposed algorithm is a recursive algorithm; i.e. in each super-round, above phases will be repeated until the WSN be alive.

2. Different Properties of the Proposed Algorithm

According to Table1, this section is representing some of most important properties of the proposed algorithm, including of: its assumptions, advantages, disadvantages and its especial characteristics.

Table 1: Different properties of the proposed algorithm

No.	Properties	Description
1	Assumptions and considerations of the proposed algorithm	<ul style="list-style-type: none"> • Sink is aware from deployment location coordinates of the WSN's nodes. • Periodically execution and repeating (dynamic-nature). • The WSN is: static/dynamic, constant/mobile nodes, homogenous (especially static with constant nodes; then, clusters' formation is only doing one time). • CHs operate as aggregator, router and relay. • Collision-prevention by using of CSMA-MAC method. • Interference-prevention through using of CDMA technique.
2	Strengths and advantages of the proposed algorithm	<ul style="list-style-type: none"> • No need to maintaining multiple paths, reducing the overhead of route maintenance and reducing the memory consumption. • High scalability. • Security improvement and secure data transmission through: <ul style="list-style-type: none"> ○ Preventing from unauthorized access to data by traffic distribution and data packets' segmentation. ○ Preventing from false or forged data injection through detecting malicious or compromised nodes by trust-based data aggregation. ○ Reducing attacks like selective forwarding and black-hole through using of different multiple CHs for data transmission in different time intervals. • Capability of data aggregation leads to reducing volume of transmitted data and consumed energy for data transmission. • Improving QoS and WSNs' lifetime. • High flexibility and accuracy: possibility of modifying proposed algorithms and criteria of different phases proportional to the WSN's properties and routing goals. • Increasing the reliability and fault tolerant; also, no need to use of error notification control packet; due to existence backup routes (multiple CHs) ⇒ reducing volume of control traffic on error prompt emission. • Balanced and distributed energy consumption: energy-aware CHs selection, dynamic rotation of CHs role between the cluster's members and balanced and non-uniform load distribution between them. • Simplicity and practicality.
3	Weaknesses and disadvantages of the proposed algorithm	<ul style="list-style-type: none"> • Security weaknesses: vulnerability against attacks like Sybil attack. • Necessity of synchronization in CHs determination and selection. • Energy and time wastage, especially in cluster's formation and CHs selection phases (imposing delay to the WSN's operations). • High computational overhead.
4	Special and major properties of the proposed algorithm	<ul style="list-style-type: none"> • Routing nature: data-centric, energy-aware, congestion-aware, QoS-aware, multipath, multi-criteria, unreactive, location-based, semi-centralized, scalable, hierarchical and clustering-based. • Congestion control and reduction: attention to the traffic status and congestion of nodes, links and routes in CHs selection and traffic distribution processes; it leads to reducing rate of data packet loss. • Guarantee the WSN's stability through high accuracy, congestion control and balanced load distribution. • High accuracy leads to imposing much overhead to the network, complex computations and consumed more resources such as energy and bandwidth (increasing size of packets' header); but it results in decreasing packet loss rate. • Possibility of selecting optimal path (supporting single-path routing). • Topology-independent. • Deterministic-nature criteria and methods (not probabilistic).

3. Considered Algorithm for Comparison: LEACH Routing Algorithm

LEACH is abbreviation of Low Energy Adaptive Clustering Hierarchy routing protocol. It is a hierarchical and clustering-based routing protocol for WSNs. LEACH is a self-organized protocol along with dynamic categorization, which using of a random method to distributing energy consumption between all members of the cluster. In this protocol, time is dividing to slots, called round; then, following steps are repeated periodically, in each round. It is including of [29, 30, 42]:

- CHs selection; dynamic rotation of CH's role between the cluster's members in different rounds, randomly.
- Clusters formation
- TDMA scheduling program creation by CHs
- Data transmission

4. Results, Analysis and Evaluations: Quantitative and Qualitative Comparison Between the Proposed Algorithm and LEACH Algorithm

This section is expressed reached results of simulations and statistical-mathematical analysis. As following tables (Table2–Table6) and figures (Figure3–Figure15) are showing, the proposed algorithm and LEACH algorithm are simulated and analyzed by statistical-mathematical techniques. Then, the proposed algorithm's performance is compared with performance of LEACH algorithm in terms of different evaluation and validation criteria.

4.1 Quantitative Evaluation and Comparison

According to the following tables (Table2, Table3 and Table4), this section is analyzed and evaluated the proposed algorithm in terms of different positive and negative statistical-quantitative parameters and then, comparing it with LEACH algorithm.

Table 2: Quantitative comparison of the proposed algorithm with LEACH algorithm: proposed criteria and formulas

No.	Proposed Criteria	Proposed Formulas	Proposed Algorithm	LEACH
1	Energy Consumption Model for data transmission: ECM [43]	$E_T(K, d) = E_{elec} \times K + \epsilon_{amp} \times K \times d^2;$ $E_R(K) = E_{elec} \times K$	$2eK + \epsilon K d^2 + e K' + \epsilon K' d^4$	$3eK + \epsilon K (d^2 + d^4)$
2	Average of Packet Loss Rate: APLR	$APLR = (\frac{1}{DLN} + \frac{1}{BLD} + Con) \times \frac{1}{BW} \times \frac{1}{Buff}$	$\frac{0.2 + Con_1}{BW_1 \times Buff}$	$\frac{1.2 + Con_2}{BW_2 \times Buff}$
3	Average of End to End Delay: AEED	$AEED = \frac{L \times (Con+P)+L}{BW \times V}$	$\frac{L \times (Con_1 + P_1) + L}{BW_1 \times V}$	$\frac{L \times (Con_2 + P_2) + L}{BW_2 \times V}$
4	Average of Route Discovery Delay: ARDD	$ARDD = \frac{L \times (Con+P+TL)+L}{BW \times V}$	$\frac{L \times (Con_1 + P_1 + TL_1) + L}{BW_1 \times V}$	$\frac{L \times (Con_2 + P_2 + TL_2) + L}{BW_2 \times V}$
5	Average of Routes Setup Time: RST	Average delay of setting up a route between source and destination nodes	$t_{Clusters' formation} + t_{CHs selection} + t_{CHs notification} + t_{Membership notification} + t_{Balanced traffic distribution}$	$t_{CH selection} + t_{CH notification} + t_{Membership notification}$
6	Memory Consumption: MC	Volume of stored information on the CHs	mV_1	mV_2
7	Routing Overhead in route discovery process: RO	Volume of transmitted data for route discovery	Clusters formation + Multiple CHs selection + $m_1 \times Packet_{CH}$ + $(n-m_1) \times Packet_{Membership}$	$CHs selection + m_2 \times Packet_{CH} + (n-m_2) \times Packet_{Membership}$
8	Fault Tolerant: FT	Average number of backup routes between source and destination	$(\text{average number of CHs in each cluster}) - 1$	0
9	Throughput: T	Volume of transmitted data per time unit or per each time interval	$C \times BW$ (bps)	BW (bps)

Table 3: Quantitative comparison of the proposed algorithm with LEACH algorithm: assumptions and results

No.	Proposed Criteria	Considerations and assumptions	Result of Comparison
1	ECM	$K' < K$	$ECM_{\text{Proposed algorithm}} < ECM_{\text{LEACH}}$
2	APLR	$Con_1 < Con_2, BW_1 > BW_2, \frac{Con_1}{BW_1} < \frac{Con_2}{BW_2}$	$APLR_{\text{Proposed algorithm}} < APLR_{\text{LEACH}}$
3	AEED	$P_1 > P_2, BW_1 > BW_2, Con_1 < Con_2, \frac{Con_1}{BW_1} < \frac{Con_2}{BW_2}$	If: $\frac{(Con_1+P_1)}{BW_1} > \frac{(Con_2+P_2)}{BW_2} \Rightarrow AEED_{\text{Proposed algorithm}} > AEED_{\text{LEACH}}$ Else: $AEED_{\text{Proposed algorithm}} < AEED_{\text{LEACH}}$
4	ARDD	$P_1 > P_2, TL_1 > TL_2, BW_1 > BW_2, Con_1 < Con_2$	If: $\frac{(Con_1+P_1+TL_1)}{BW_1} > \frac{(Con_2+P_2+TL_2)}{BW_2} \Rightarrow ARDD_{\text{Proposed algorithm}} > ARDD_{\text{LEACH}}$ Else: $ARDD_{\text{Proposed algorithm}} < ARDD_{\text{LEACH}}$
5	RST	-----	$RST_{\text{Proposed algorithm}} > RST_{\text{LEACH}}$
6	MC	m: average count of each cluster's members; V_i : size of each record of routing table; $V_1 > V_2$	$MC_{\text{Proposed algorithm}} \gg MC_{\text{LEACH}}$
7	RO	m: average count of CHs; $m_1 > m_2$	$RO_{\text{Proposed algorithm}} > RO_{\text{LEACH}}$
8	FT	(average number of CHs in each cluster) ≥ 1 ; Proposed algorithm: Energy-aware CH selection and deterministic-nature; LEACH: Random CHs selection and probabilistic-nature;	$FT_{\text{Proposed algorithm}} \geq FT_{\text{LEACH}}$
9	T	C: average number of CHs in each cluster; $C \geq 1$	$T_{\text{Proposed algorithm}} \gg T_{\text{LEACH}}$

Table 4: Words and associated abbreviations in formulas and variables of quantitative comparison

No.	Word	Abbreviation
1	Consumed energy for transferring K bit data	$E_T(K, d)$
2	Consumed energy for receiving K bit data	$E_R(K)$
3	Distance between source and destination	d
4	Volume of ready data (bit)	K
5	Volume of aggregated data (bit)	K'
6	Required energy for electronics transmission; $E_{elec} = e = 50$ nj/bit	E_{elec}
7	Signal relay when data transmission to acquiring acceptable signal to noise ratio; $\epsilon_{amp} = \epsilon = 100$ pj/bit/m	ϵ_{amp}
8	Capability of deadlock nodes and failure links detection: most attention = 10, no attention = 1	DLN
9	Amount of existent Congestion	Con
10	Available bandwidth	BW
11	Size of free buffer	Buff
12	Capability of balanced load distribution: most score = 10 and lowest score = 1	BLD
13	Average length of each route to the Sink	L
14	Speed of transmission media (for radio waves is 3×10^8 m/s)	V
15	Complexity and volume of intermediate computations and processes	P
16	Traffic overhead and volume of control packets which exchanged for routes discovery	TL
17	Delay of operation i	t_i
18	Size of packet: CH notification	Packet _{CH}
19	Size of packet: Membership notification	Packet _{Membership}

4.2 Qualitative Evaluation and Comparison

According to the following tables (Table5 and Table6) and figures (Figure3–Figure15), this section is analyzed and evaluated the proposed algorithm in terms of different positive and negative statistical-qualitative parameters and then, comparing it with LEACH algorithm.

2	Scalability and distributed nature	Much	Much
3	Throughput	Much	Low
4	Reliability and fault tolerant	Much	Very low
5	Accuracy	Very much	Very low
6	Execution speed	Moderate	Very much

Note: Numerical equivalents are: {very low: 1, low: 3, moderate: 5, much: 7, very much: 9}

Table 5: Statistical-qualitative comparison of the proposed algorithm with LEACH algorithm: in terms of positive criteria

No.	Criteria	Proposed algorithm	LEACH
1	Balanced energy consumption	Much	Low

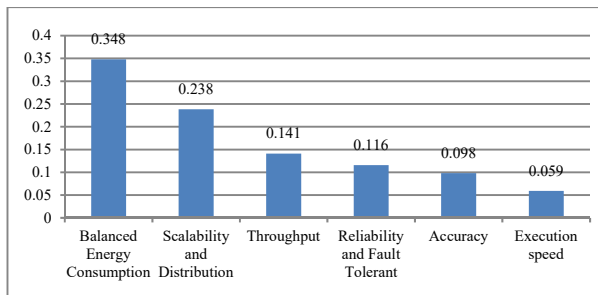


Fig. 3 Weights of positive statistical-qualitative criteria

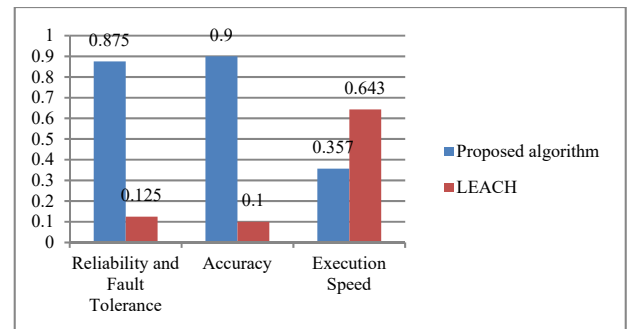


Fig. 5 Comparison between performance of the proposed algorithm and performance of LEACH algorithm in terms of reliability and fault tolerant, accuracy and execution speed

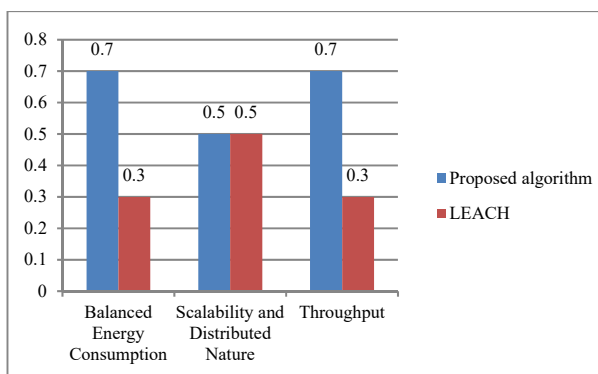


Fig. 4 Comparison between performance of the proposed algorithm and performance of LEACH algorithm in terms of balanced energy consumption, scalability and throughput

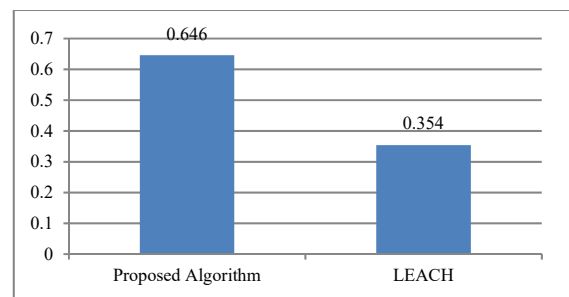


Fig. 6 Comparison between total performance of the proposed algorithm and total performance of LEACH algorithm in terms of all positive statistical-qualitative criteria

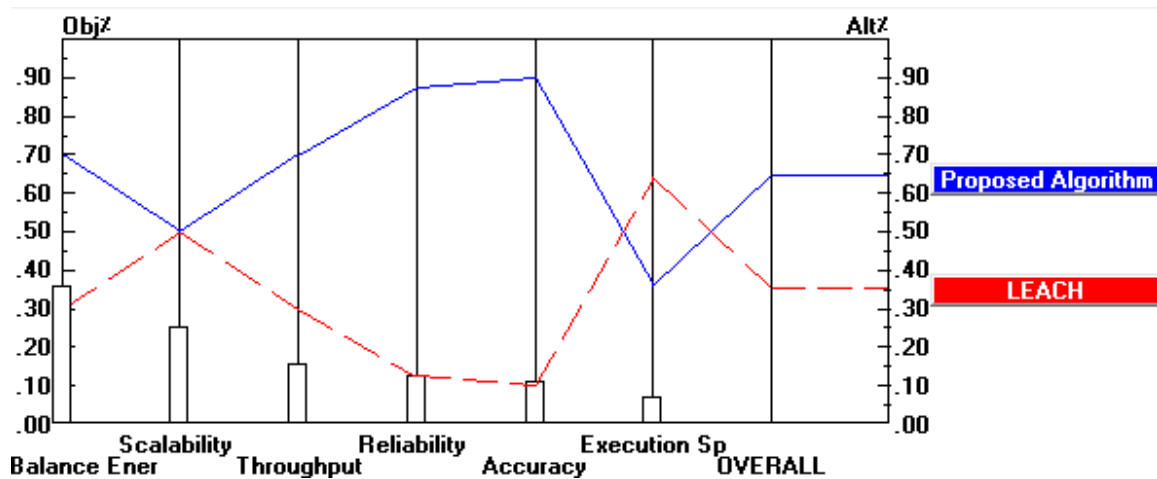


Fig. 7 Sensitivity analysis of the algorithms' performance: comparison between performance of the proposed algorithm and performance of LEACH algorithm in terms of positive statistical-qualitative criteria

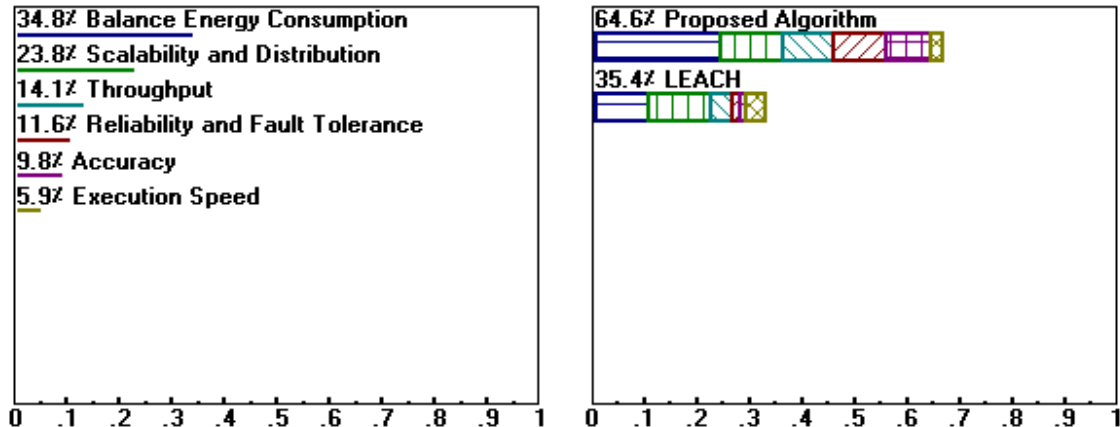


Fig. 8 Sensitivity analysis of the algorithms' dynamicity in terms of positive statistical-qualitative criteria

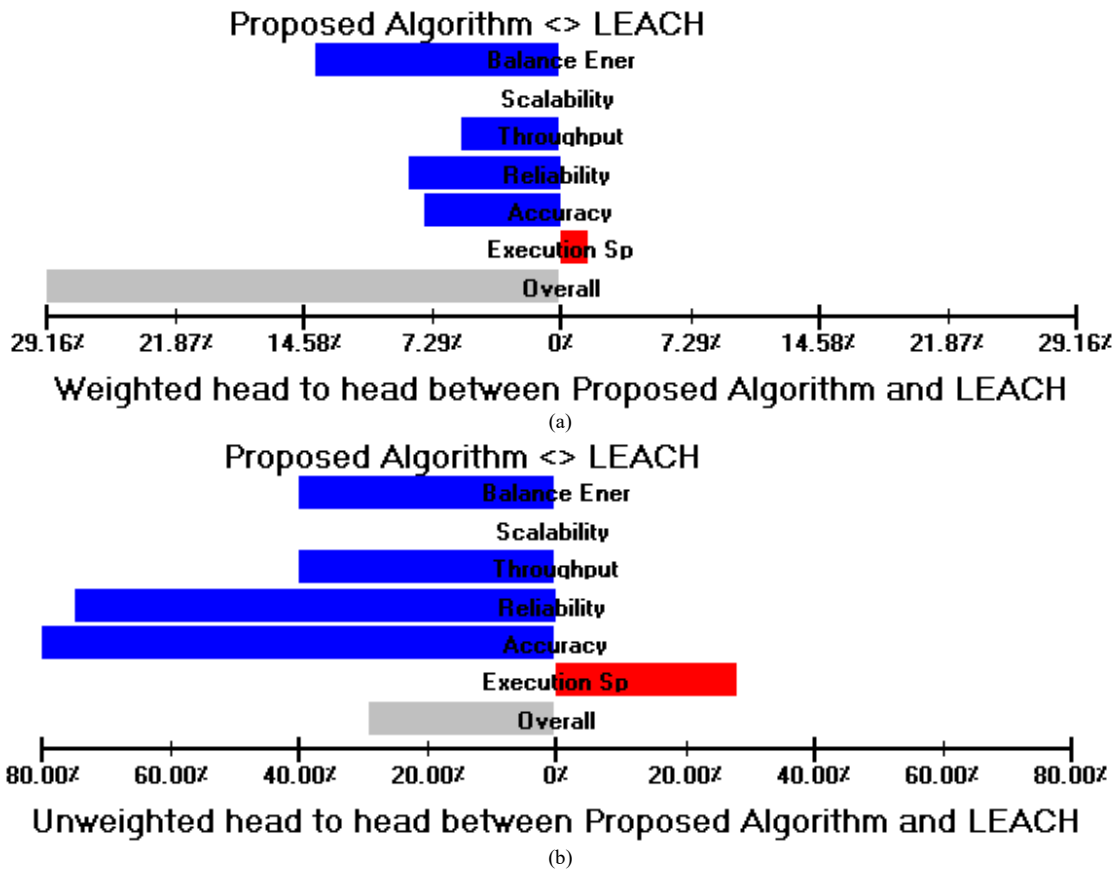


Fig. 9 (a) Weighted and (b) Unweighted head to head sensitivity analysis between proposed algorithm and LEACH algorithm in terms of positive statistical-qualitative criteria

4	Computational complexity	Moderate	Very low
---	--------------------------	----------	----------

Table 6: Statistical-qualitative comparison of the proposed algorithm with the LEACH algorithm: in terms of negative criteria

No.	Criteria	Proposed algorithm	LEACH
1	Packet loss rate	Very low	Moderate
2	Route discovery time (taken time for clustering)	Much	Low
3	Execution time	Moderate	Low

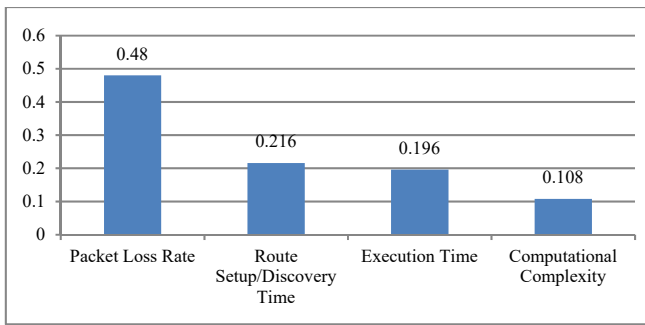


Fig. 10 Weights of negative statistical-qualitative criteria

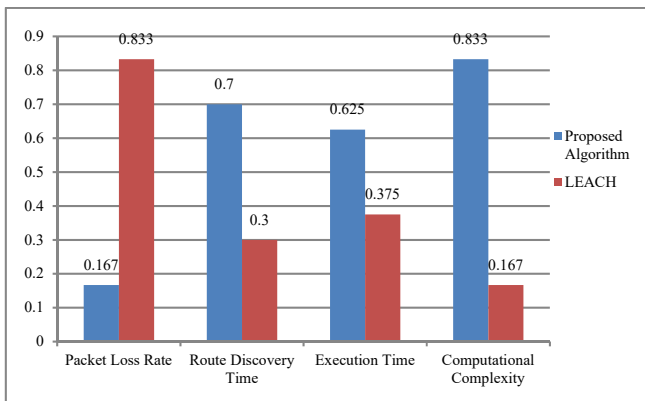


Fig. 11 Comparison between performance of the proposed algorithm and performance of LEACH algorithm in terms of packet loss rate, route discovery time, execution time and computational complexity

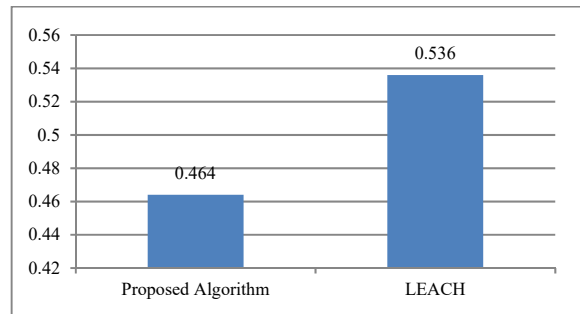


Fig. 12 Comparison between total performance of the proposed algorithm and total performance of LEACH algorithm in terms of all negative statistical-qualitative criteria

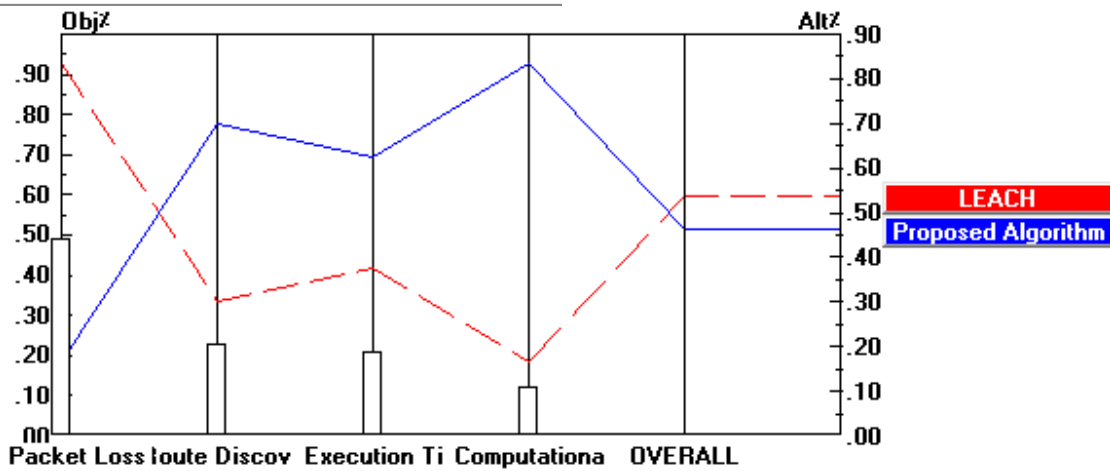


Fig. 13 Sensitivity analysis of the algorithms' performance: comparison between performance of the proposed algorithm and performance of LEACH algorithm in terms of negative statistical-qualitative criteria

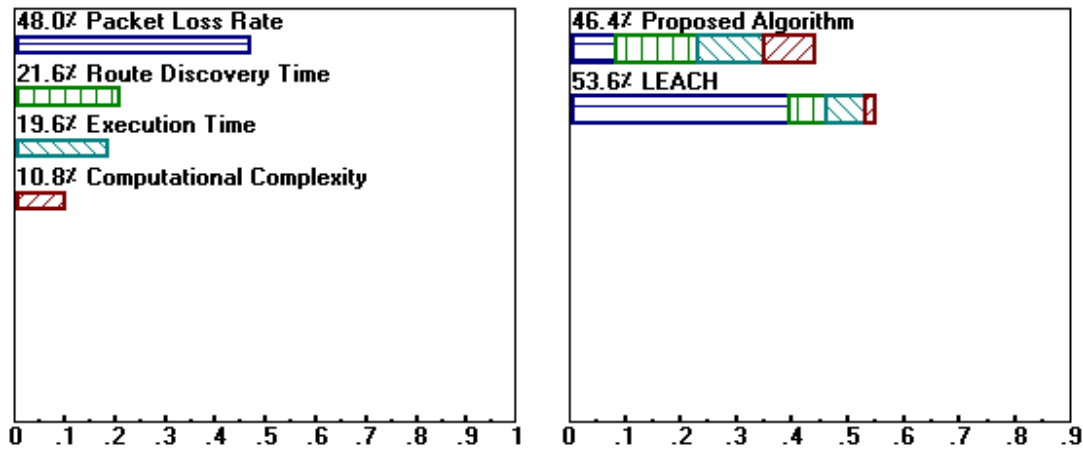


Fig. 14 Sensitivity analysis of the algorithms' dynamicity in terms of negative statistical-qualitative criteria

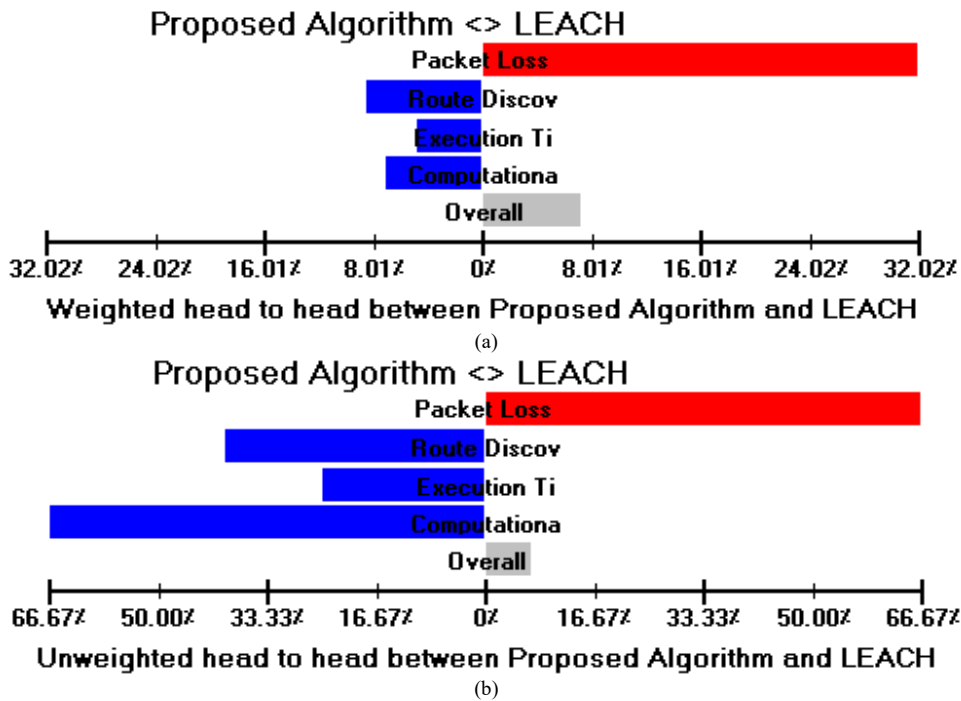


Fig. 15 (a) Weighted and (b) Unweighted head to head sensitivity analysis between proposed algorithm and LEACH algorithm in terms of negative statistical-qualitative criteria

5. Conclusion and Future Directions

Nowadays, clustering and multipath routing are high interests topics in WSNs. But, existent algorithms have many weaknesses and challenges such as: imposing high overhead to WSNs, security vulnerabilities, address-centric, low-scalability, permanent usage of optimal paths and severe resources' consumption; then, they are not appropriate for WSNs. Therefore, this paper is proposed an energy-aware, congestion-aware, location-based, data-centric, scalable, hierarchical and clustering-based multipath routing

algorithm based on Numerical Taxonomy technique for homogenous WSNs. Figure16 and Figure17 are showing main steps of the proposed multipath routing algorithm and different steps of the proposed multiple CHs selection algorithm based on Numerical Taxonomy technique.

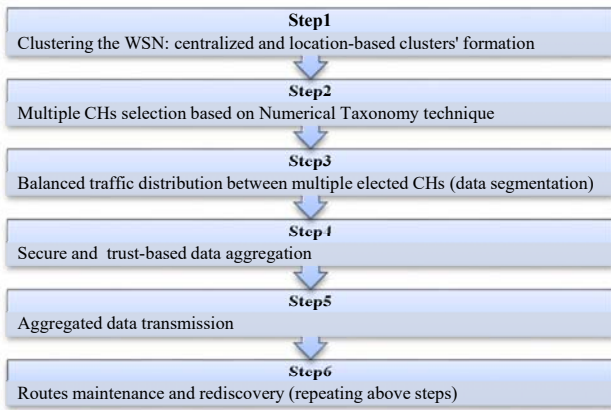


Fig. 16 Main steps of the proposed multipath routing algorithm

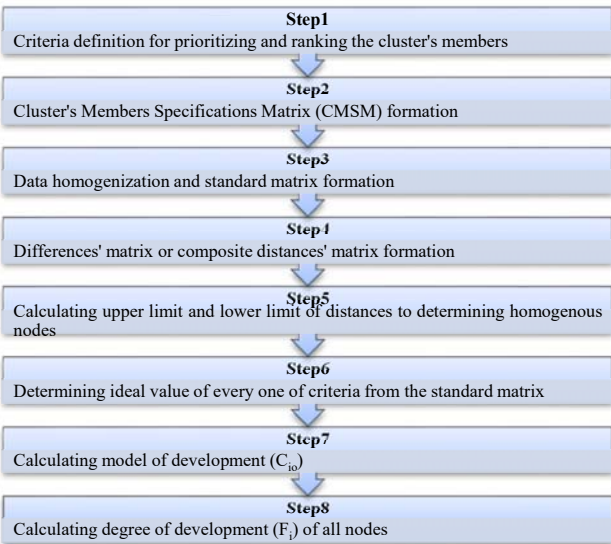


Fig. 17 Different steps of the proposed multiple CHs selection algorithm based on Numerical Taxonomy technique

Some major properties of the proposed algorithm are:

- Data-centric, energy-aware, congestion-aware, location-based, hierarchical and clustering-based routing
- Topology-independent
- High scalability and semi-centralized nature
- Improving QoS and WSNs' lifetime
- High flexibility and accuracy
- Balanced and uniform distribution of energy consumption
- Balanced and non-uniform distribution of traffic load
- Simplicity and practicality

According to the following tables (Table7, Table8 and Table9) and figures (Figure18 and Figure19), performance of the proposed algorithm has been compared with performance of LEACH algorithm; results of simulations and statistical-mathematical analysis are representing the proposed algorithm has been improved than LEACH algorithm in terms of parameters such as balanced resources' consumption such as energy and bandwidth, throughput, reliability and fault tolerant, accuracy, QoS parameters like average rate of packet delivery and WSNs' lifetime.

Table 7: Comparison between the proposed algorithm and LEACH algorithm in terms of quantitative criteria

No.	Proposed Criteria	Result of Comparison
1	ECM	$ECM_{Proposed\ algorithm} < ECM_{LEACH}$
2	APLR	$APLR_{Proposed\ algorithm} < APLR_{LEACH}$
3	AEED	If: $\frac{(Con_1+P_1)}{BW_1} > \frac{(Con_2+P_2)}{BW_2} \Rightarrow AEED_{Proposed\ algorithm} > AEED_{LEACH}$ Else: $AEED_{Proposed\ algorithm} < AEED_{LEACH}$
4	ARDD	If: $\frac{(Con_1+P_1+TL_1)}{BW_1} > \frac{(Con_2+P_2+TL_2)}{BW_2} \Rightarrow ARDD_{Proposed\ algorithm} > ARDD_{LEACH}$ Else: $ARDD_{Proposed\ algorithm} < ARDD_{LEACH}$
5	RST	$RST_{Proposed\ algorithm} > RST_{LEACH}$
6	MC	$MC_{Proposed\ algorithm} \gg MC_{LEACH}$
7	RO	$RO_{Proposed\ algorithm} > RO_{LEACH}$
8	FT	$FT_{Proposed\ algorithm} \geq FT_{LEACH}$
9	T	$T_{Proposed\ algorithm} \gg T_{LEACH}$

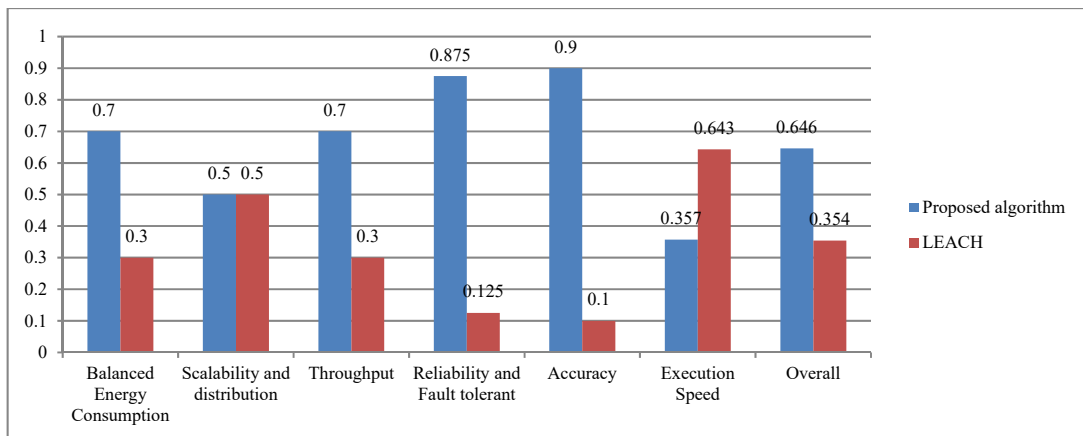


Fig. 18 Statistical-qualitative comparison of the proposed algorithm with LEACH algorithm: in terms of positive statistical-qualitative criteria

Table 8: Improvement percentage of the proposed algorithm than LEACH algorithm in terms of positive statistical-qualitative criteria

No.	Positive criteria	LEACH (in percentage)
1	Balanced Energy Consumption	57.14
2	Scalability and Distributed nature	0
3	Throughput	57.14
4	Reliability and Fault Tolerant	85.71
5	Accuracy	88.89
6	Execution Speed	-44.48
	Total Improvement Percentage (positive criteria)	45.20

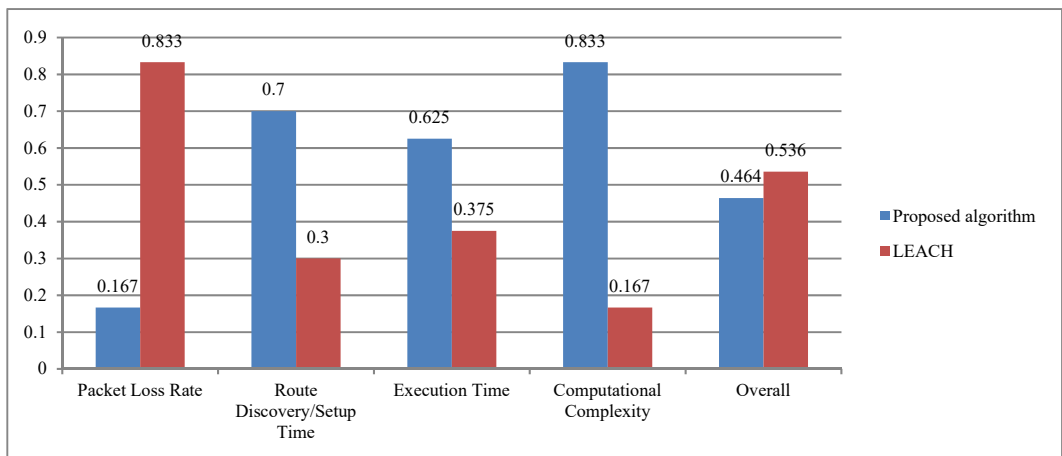


Fig. 19 Statistical-qualitative comparison of the proposed algorithm with LEACH algorithm: in terms of negative statistical-qualitative criteria

Table 9: Improvement percentage of the proposed algorithm than LEACH algorithm in terms of negative statistical-qualitative criteria

No.	Negative criteria	LEACH (in percentage)
1	Packet Loss Rate	79.95
2	Route Setup/Discovery Time	-57.14
3	Execution Time	-40
4	Computational Complexity	-79.95
	Total Improvement Percentage (negative criteria)	13.43

There are several additional issues should be further studied in future researches. Some of most challenging proposed topics of these issues are as following:

- Proposing a method for detecting deadlock nodes, failed links and selfish nodes.
- Security analysis of the proposed algorithm, finding its security vulnerabilities and improving them.
- Presenting a secure and trust-based multipath routing algorithm for WSNs.
- Determining length of each super-round for executing and repeating the proposed algorithm.
- Suggesting a method for balanced and appropriate multiple CHs distribution in the cluster.
- Presenting a light-weight and energy-efficient clustering algorithm for WSNs.
- Discussing on finding optimal number of clusters, CHs and also, estimation of optimal frequency of CHs re-selection to gain better energy-efficiency.
- Discussing on data segmentation, data aggregation and data reassembly techniques in traffic distribution process.

References

- [1] J. Yick, B. Mukherjee and D. Ghosal; Wireless Sensor Network Survey; Elsevier's Computer Networks Journal, 52, (2292-2330); 2008.
- [2] H. Jadidoleslami; A Comprehensive Comparison of Attacks in Wireless Sensor Networks; International Journal of Computer Communications and Networks (IJCCN); Vol. 4, No. 1, February, 2014.
- [3] H. Jadidoleslami; A Novel Clustering Algorithm for Homogenous and Large-Scale Wireless Sensor Networks: Based on Sensor Nodes Deployment Location Coordinates; International Journal of Computer Science and Network Security (IJCSNS); Vol. 14, No. 2, February, 2014.
- [4] K. Sha, J. Gehlot and R. Greve; Multipath Routing Techniques In Wireless Sensor Networks: A Survey; Wireless Personal Communications: An International Journal, Vol. 70, Iss. 2, pp. 807-829; 2013.
- [5] M. Radi, B. Dezfouli, K. Abu Bakar and M. Lee; Multipath Routing in Wireless Sensor Networks: Survey and Research Challenges; MDPI journals, Sensors Journal, Vol. 12, Iss. 1; pp. 650-685; 2012.
- [6] A. Jayashree, G. S. Biradar and V. D. Mytri; Review of Multipath Routing Protocols in Wireless Multimedia Sensor Network: A Survey; International Journal of Scientific & Engineering Research, Vol. 3, Iss. 7; 2012.
- [7] Y. Chen, E. Chan and S. Han; Energy efficient multipath routing in large scale sensor networks with multiple sink nodes; In Advanced Parallel Processing Technologies, Vol. 37, pp. 390-399; 2005.
- [8] Ch. Ahn, J. Shin and E. Huh; Enhanced Multipath Routing Protocol Using Congestion Metric in Wireless Ad Hoc Networks; International Federation for Information Processing, pp. 1089-1097; 2006.
- [9] X. Huang and Y. Fang; Multi-constrained QoS Multipath Routing in Wireless Sensor Networks; Journal of Wireless Networks, Vol. 14, Iss. 4, pp. 465-478; 2007.
- [10] Y. M. Lu and V. W. S. Wong; An energy-efficient multipath routing protocol for wireless sensor networks; International Journal of Communication Systems, Vol. 20, No. 7, pp.747-766; 2007.
- [11] R. Vidhyapriya and P. T. Vanathi; Energy Efficient Adaptive Multipath Routing for Wireless Sensor Networks; IAENG International Journal of Computer Science; 2007.
- [12] J. Y. Teo, Y. Ha and C. Tham; Interference-Minimized Multipath Routing with Congestion Control in Wireless Sensor Network for High-Rate Streaming; Journal of IEEE transaction of mobile computing, Vol. 7, No. 9, pp. 1124-1137; 2008.
- [13] S. Han, Z. Zhong and H. Li; Coding-aware multi-path routing in multi-hop wireless networks; IEEE International Conference on Performance, Computing and Communications Conference (IPCCC), pp. 93-100; 2008.
- [14] A. Bagula and K. Mazandu; Energy Constrained Multipath Routing in Wireless Sensor Networks; In Proceeding of the 5th International Conference on Ubiquitous Intelligence and Computing, pp. 453-467; 2008.
- [15] M. Maimour; Maximally Radio-Disjoint Multipath Routing for Wireless Multimedia Sensor Networks; In Proceedings of the 4th ACM Workshop on Wireless Multimedia Networking and Performance Modeling, pp. 26-31; 2008.
- [16] K. Saleem, N. Faisal and S. Hafizah, S. Kamilah and R. A. Rashid; A self-optimized multipath routing protocol for wireless sensor networks; International Journal of Recent Trends in Engineering, Vol. 2, No. 1, pp. 93-97; 2009.
- [17] Z. Wang, E. Bulut and B. K. Szymanski; Energy efficient collision aware multipath routing for wireless sensor networks; In Proceedings of the IEEE international conference on Communications; 2009.
- [18] S. Li, R. K. Neelisetti and C. Liu; Efficient multi path protocol for wireless sensor networks; International Journal of Wireless and Mobile Networks, Vol. 2; 2010.
- [19] X. Wang, C. Che and L. Li; Reliable multi-path routing protocol in wireless sensor networks; In Proceedings of the International Conference on Parallel and Distributed Computing, Applications and Technologies, pp. 289-294; 2010.
- [20] J. Yang, M. Xu, W. Zhao and B. Xu; A multipath routing protocol based on clustering and ant colony optimization for wireless sensor networks; Sensors, Vol. 10, Iss. 5; 2010.
- [21] B. J. Othman and B. Yahya; Energy Efficient and QoS Based Routing Protocol for Wireless Sensor Networks; Journal of Parallel and Distributed Computing; Vol. 70, pp. 849-857; 2010.
- [22] I. T. Almalkawi, M. G. Zapata and J. N. Al-Karaki; A secure cluster-based multipath routing protocol for wmsns; Sensors, Vol. 11, Iss. 4; 2011.
- [23] S. Pratheema, K. G. Srinivasagan and J. Naskath; Minimizing end-to-end delay using multipath routing in wireless sensor networks; International Journal of Computer Applications, Vol. 21, Iss. 5, pp. 20-26; 2011.
- [24] M. Cherian and T. R. Gopalakrishnan Nair; Multipath routing with novel packet scheduling approach in wireless sensor networks; International Journal of Computer Theory and Engineering, Vol. 3; 2011.
- [25] S. Soundararajan and R. S. Bhuvaneshwaran; Adaptive Multi-Path Routing for Load Balancing in Mobile Ad Hoc Networks, Journal of Computer Science, pp. 648-655; 2012.

- [26] S. Soundararajan and R. S. Bhuvaneshwaran; Multipath Load Balancing & Rate Based Congestion Control for Mobile Ad Hoc Networks (MANET); Second IEEE International Conference on Digital Information and Communication Technology and its Applications (DICTAP); 2012.
- [27] R. Vinod Kumar and R. S. D. Wahida Banu; E2AODV Protocol for Load Balancing in Ad-Hoc Networks; Journal of Computer Science, Vol. 8, Iss. 7, pp. 1198-1204; 2012.
- [28] N. Nasser and Y. Chen; SEEM: Secure and energy-efficient multipath routing protocol for wireless sensor networks; Computer Communications, Special issue on security on wireless ad hoc and sensor networks, Vol. 30, Iss. 11; 2007.
- [29] H. Jadidoleslami; An Introduction to Various Basic Concepts of Clustering Techniques on Wireless Sensor Networks; International journal of Mobile Network Communications & Telematics; Vol. 3, No.1, pp. 1-17; 2013.
- [30] A. A. Abbasi, M. Younis; A survey on clustering algorithms for wireless sensor networks; Computer Communications, Vol. 30, pp. 2826-2841; June, 2007.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] 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; September, 2007.
- [35] 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.
- [36] S. Opricovic and G. H. Tzeng, Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS, European Journal of Operational Research 156 (2), pp. 445-455, 2004.
- [37] G. L. Fu, C. Yang and G. H. Tzeng, A multi-criteria analysis on the strategies to open Taiwan's mobile virtual network operators services, International Journal of Information Technology & Decision Making 6 (1), pp. 85-112, 2007.
- [38] S. Opricovic and G. H. Tzeng, Multi-criteria planning of post-earthquake sustainable reconstruction, Computer-Aided Civil and Infrastructure Engineering 17 (3), pp. 211-220, 2002.
- [39] Y. Shi, Y. Peng, G. Kou and Z. Chen, Classifying credit card accounts for business intelligence and decision making: a multiple-criteria quadratic programming approach, International Journal of Information Technology & Decision Making 4 (4), pp. 581-599, 2005.
- [40] G. H. Tzeng, M. H. Teng, J. J. Chen and S. Opricovic, Multi-criteria selection for a restaurant location in Taipei, International Journal of Hospitality Management 21 (2), pp. 171-187, 2002.
- [41] G. H. Tzeng, C. W. Lin and S. Opricovic, Multi-criteria analysis of alternative-fuel buses for public transportation, Energy Policy, 33 (1), pp. 1373-1383, 2005.
- [42] K. Akkaya and M. Youni; A survey on routing protocols for wireless sensor networks; ScienceDirect, Vol. 3, No. 3, pp. 325-349; 2005.
- [43] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan; Energy-Efficient Communication Protocol for Wireless Micro-sensor Networks; Proceedings of the 33rd IEEE Hawaii International Conference on System Sciences; 2000.



Hossein Jadidoleslami is a faculty member at the University of Zabol (UOZ), Department of Computer Engineering, Zabol, Sistan and Baluchestan, Iran. He received his Ph.D. degree in Computer Science-Information Technology Engineering (CS-IT) from the Malekashtar University of Technology (MUT), Tehran, Iran, in March 2016. He received his B.Sc. degree in Information Technology (IT) Engineering from the University of Sistan and Baluchestan (USB), Zahedan, Iran, in September 2009. He also has been received his M.Sc. degree in Information Technology (IT) Engineering from the University of Guilan, Rasht, Iran, in March 2011. His research interests are including of Computer Networks (Wireless Sensor Network), Network Security (Trust Management Systems), Information Security (Intrusion Detection System) and E-Commerce. He may be reached at tanha.hossein@gmail.com or jadidoleslami@gmail.com.