Reputable Path Recognition in Wireless Sensor Network using Intelligent Water Drop

Heena Rathore
Indian Institute of Technology, Jodhpur

Manish Goyal
Infosys, Bangalore

ABSTRACT
Applying trust in Wireless Sensor network (WSN) is an emerging area where researchers are engrossed in developing novel design archetype to address the security issues. Security plays an important role in WSN where trustworthy sensor node turns into a untrustworthy one because of hardware and software faults. Aim of the paper is to propose an optimized path coverage algorithm with the help of an emerging biologically inspired technique, i.e., intelligent water drop (IWD). The proposed work is guaranteed to have most reputable path leading to most trustworthy nodes in WSN. This approach uses dynamic parameters for finding all the optimal paths using basic properties of natural water drops. It proposes how different test cases can be considered as an IWD moving on the edges of the network for finding the optimal paths. The algorithm guarantees complete path coverage.

Keywords
Biologically inspired, Trust, WSN, IWD.

1. INTRODUCTION
Exploring the symbiotic nature of biological systems can result in valuable knowledge for computer networks [2]. Biologically inspired techniques seem promising because of their intrinsic appealing characteristics as self-healing, self-adapting, and self-evolving [3]. They are apter because the techniques are more robust at the time of catastrophic errors when optimized systems fail to deliver optimum performance. While closely looking at nature, we can derive inspiration from everything we see [4]. May it be from animals, birds, waterfall, etc. Birds’ flock together synchronizing their pattern, bees perform nectar search with a well-defined procedure, and ants find their food source with the stigmergy. All these social insects along with numerous examples have been studied thoroughly and are used in deriving inspiration in various fields.

Wireless Sensor Network is a network of remote stations /nodes sending sensing information to the gateway. The nodes have the capability to sense various environmental parameters such as temperature, pressure, humidity, etc. However, there are times when the sensor nodes knowingly/ unknowingly send bad information to the gateway. This is because the nodes are remotely located and anybody can hamper the device. For this reason, there is a requirement to look at the security issues of WSN. It is better to find most trustworthy/benevolent nodes in the network, on which the system can rely at its best. The objective of the paper is to present a novel approach for finding the most reputable path leading to trustworthy nodes in the network. Section II gives the related work in the field of trust. Section III is the novel approach that can be used in WSN for recognition of reputable path. Section IV summarizes the paper and presents the scope of future plans.

2. BACKGROUND
In WSN, nodes work selflessly to achieve a common goal like measuring the temperature parameter using thermocouple sensor nodes. Also, the architecture of WSN can be either as flat configuration or clustered configuration as shown in Figure 1.

Nodes can either send the information directly to the gateway or to the cluster heads /relay nodes which in turn sends it to the gateway. Hence the topology can be considered as a tree structure having a source node which senses the information, processes it and sends it the respective destination node.

Sensor nodes precisely have two main functionalities besides processing data, i.e., sensing and communicating data. Both at the time of sensing as well as communicating data, node can misbehave. While sensing the data, the sensor node can misinform by becoming a faulty node because of noise, environmental harsh conditions, or may be because of interference. And while communicating the data, the sensor node can become malicious and become selfish to exhaust nearby node energy as compared to its own. It can be malicious while forwarding or routing packets or it can self exclude data by not sending the data to the base station. This scenario makes it ardent to apply trust mechanisms to WSN to find the trustworthy nodes of the network.

There are various mechanisms used to find the trustworthiness of the network. Momani et al. present a survey of trust models in different network domains. Several techniques such as ratings, weightings, probability, Bayesian network approach, game theory approach, swarm intelligence, neural network and fuzzy logic are used in assessing trust factor of sensor nodes [16]. Swarm intelligence is the bio-inspired technique which uses ant colony optimization, where ants build paths which fulfill certain conditions in a graph. These ants leave some pheromone traces that help next ants to discover and follow those routes [1]. The resultant is that we find the most reputable path. Besides this there are other techniques which were developed, like, ATRM, QDV, ATSN, RFSN, CORE, DBRTS [5], [6], [7], [8], [9], [10].
3. BIO-INSPIRED, INTELLIGENT WATERDROP TECHNIQUE

3.1 Assumptions

There are two status of WSN, either it can be static one, where the nodes are strict to a place or it is a dynamic one where nodes are mobile. There are nodes which remain idle/asleep most of the time or there are nodes which are highly active providing some service. We have assumed a topology where there is high-sensor activity, where there are nodes asking for services and there are nodes providing those services.

We will also assume that every node will only know its neighbors, that is, those nodes within its wireless range. Our model is aimed to help a node requesting a certain service to the network to discover the most trustworthy route leading to a node providing the right requested service. A node (equally a path) can be considered untrustworthy either because it intentionally provides a fraudulent service or because it provides a wrong one due to hardware failures or performance deterioration.

3.2 Intelligent Water Drop

IWD algorithm has been studied extensively which is nothing but a idealized upon studying natural water drops in rivers [13]. Natural water drops have two properties that are soil and velocity. Every water drop flows with some velocity and carries an amount of soil with itself. The velocity of the water drop helps in removing some soil from the riverbeds, and this soil is added to the soil of the water drop. During the transition, the water drops gains speed; due to which a water drop with more velocity can erode more soil. One more interesting feature of the water drop is that it chooses an easy path when there are many paths available at the same time. Here, easy path means the path with less soil. In this way constructively, the water drop chooses its optimal path. Using these remarkable properties IWDs are created. These drops possess two important properties of natural water drops that are:

1. the soil it carries, denoted by soil (IWD)
2. the velocity that it possesses, denoted by velocity (IWD).

Both these values can change when an IWD flows from one location to another in its environment (given problem statement). The IWDs’ velocity is non-linearly proportional to the inverse of the soil on the path between the two locations. The IWDs’ soil is increased by removing some soil from the path, and the amount of soil added is non-linearly proportional to the inverse of the time taken to move between the locations. This time is proportional to the velocity and inversely proportional to the distance between the two locations. In addition, the path choosing property is implemented by calculating an uniform random distribution function, such that the probability of choosing a path is inversely proportional to the soil of the available paths. This paper briefs about IWDs, which is used to find optimized path of WSN to find the most trustworthy nodes in the network. The following section explains how the properties of water drops are applied to the desired problem and the process of achieving the reputable paths.

3.3 Proposed Work

We presume to know the topology of the network at the start. Initially, the important factors affecting the working of the algorithm are explained followed by a stepwise explanation of the algorithm.

1. Fitness function

A fitness function has been designed for the calculation of soil on all the links of the graph [13]. The basic idea behind using the fitness function is to give importance to more deep and complex paths. A well-constructed fitness function may increase the chances of finding a solution and reaching higher coverage, which is always desirable. For the proposed algorithm, the fitness function can be calculated as:

\[
\text{soil} (i, j) = a \times \text{subgraph} (i) + b \times \text{condition} (j)
\]

where soil (i,j) is the soil on the edge between the two nodes i and j. Condition (j) of a node j can be either 0 or 1, depending on whether a particular node has a decision element or not. If it has a decision statement then it is assigned a value of 1 else 0. Subgraph (j) is the subgraph value, i.e., number of nodes below that particular node (j) in a given graph. The IWD starts from the initial node of the graph and after that the probability of each path is calculated using the above mentioned fitness function. The path with the highest probability is considered as the desired path for the water drop. The configuration parameter ‘a’ and ‘b’ is assigned a constant value of 2 and 1 respectively. The reason behind assigning these values is that, more weight should be assigned to nodes with a larger subgraph value rather than nodes with smaller subgraph value. Subgraph value has been given greater importance in the designing of the fitness function, because by traversing the part of the code with greater depth the probability of achieving complete code coverage is high. This fitness function was so chosen, because during code coverage there are chances of errors in either decision nodes or more dense and complex paths. In order to assign weight to links based on their complexity, this fitness function is chosen.

2. List of visited paths

A list of visited paths known as ‘Ve’ is maintained to record the paths traversed by the IWD. In order to keep track of the nodes visited, a check is performed on the node, and if all its children are visited then it is deleted. This operation is performed recursively. This is done in order to avoid the IWD from traversing the path which has been already traversed.

3. Constant values

Various configuration parameters \(a, b, c\), \(a, b, c\), \(p\) are used in the calculation and updating of soil and velocity respectively during implementation of the IWD algorithm [11, 12].

3.4 Algorithm

Parameters which are used are:

- Probability \((i, j)\): probability of the IWD to move from node i to node j.
- Time \((i, j)\): time taken by the IWD to move from node i to node j.
- \(\Delta\text{Soil} (i, j, \text{IWD})\): soil eroded by the IWD on the link between node i and node j.
- \(\text{Vel} (i, j, \text{IWD})\): velocity of the IWD when it moves from node i to node j.

Step wise implementation of the algorithm is [15]:
Algorithm 1: Bio-inspired Intelligent Water drop algorithm

**STEP 1: INITIALIZATION OF PARAMETERS**

a. Velocity updating parameters are $a_v$, $b_v$, $c_v$. set as: $a_v = 1$, $b_v = 1$, $c_v = 1$.

b. Soil updating parameters are $a_s$, $b_s$, $c_s$ set as: $a_s = 1$, $b_s = 1$, $c_s = 1$.

c. The local soil updating parameter is $\rho$. Here $\rho = 0.1$ is chosen for better experimental results. Value of $\rho$ can range from 0 to 1, depending upon the amount of soil on the link (nodes).

d. The initial soil on each edge of the graph is computed using the fitness function and the configuration parameters

e. The initial velocity of each IWD is set to 100.

f. Every IWD has a visited path list $V_c$ (IWD) which is initially empty i.e. $V_c$ (IWD) = {}.

**STEP 2: UPDATE OF PARAMETERS VALUES**

Repeat below steps until topology does not gets empty:

2.1. Calculation of probability:

For the IWD receding at node $i$, the next node $j$ (which is not in visited node list $V_c$ (IWD)) is chosen on the basis of probability $p_i$ (IWD)$. The probability for IWD to move from the current node $i$ to node $j$ is calculated using:

$$\text{Probability} (i, j) = \frac{\text{Soil} (i, j)}{\sum_{k \neq V_c} \text{Soil} (i, k)}$$

Where: $\text{Soil} (i, j)$ is the soil between the two nodes, $\sum_{k \neq V_c} \text{Soil} (i, k)$ denotes the summation of the soils of all the paths which can be traversed from the current node $i$. In the original algorithm, the IWD moves to the location where the soil is less, but in our case we traverse the IWD to the path with more soil.

2.3 Calculation of time taken:

After calculating the probability, IWD chooses the next node which it has to move on. After selecting the respective node, the time taken to move from node $i$ to node $j$ is calculated as:

$$\text{Time} (i, j) = \frac{\text{subgraph} (i) - \text{subgraph} (j)}{\text{vel} (iwd)}$$

Where $\text{subgraph} (i)$ – $\text{subgraph} (j)$ correspond to the distance between the nodes and $\text{vel}(iwd)$ is the original velocity of the IWD.

2.4 Calculation of Soil of the IWD:

Soil carried by IWD is computed using 2.3:

$$\Delta \text{Soil} (i, j, iwd) = \frac{a_s}{(b_s + c_s \times \text{time} (i, j))}$$

Where: $a_s$, $b_s$, $c_s$ are the positive parameters as specified earlier.

2.5 Calculation of Velocity of the IWD:

The velocity of the IWD after it has moved from node $i$ to $j$ is calculated as:

$$\text{vel} (i,j,iwd) = \text{vel}(iwd) + (a_v + (b_v + c_v \times \text{soil}(i,j)))$$

Where: soil ($i,j$) is the soil on the path before the IWD traversed the required path. $a_v$, $b_v$, $c_v$ are the positive parameters as specified earlier and vel($iwd$) is the previous velocity of the IWD.

2.6 Soil updation of the link:

Since, the IWD carries some amount of soil with it; the soil on the link is reduced. Therefore, the soil on the link is updated as:

$$\text{Soil} (i,j) = (1 - \rho) \text{soil} (i,j) - \rho \times \Delta \text{soil} (i,j,iwd),$$

Where: $\rho$ is a positive parameter, $\Delta \text{soil} (i,j,iwd)$ is the soil carried by the IWD while moving from the node $i$ and $j$.

2.7 Current node is updated and the same procedure from step 2.2 is repeated until it does not reach leaf node, or one complete path is not reached. After it reaches the leaf node then that node is deleted, and its parent is also checked. If all its leaves have been deleted, then it is also deleted.

2.8 After completing one possible path by the water drop, the visited path list is updated and the IWD again starts with a new iteration where all parameters are initialized again from step 1 onwards. If the topology becomes empty, that means all the paths have been traversed, and we have to exit the loop else all the steps from 2.1 are executed.

In this way, all the possible paths are determined with their corresponding weights. The path with highest weight should be the most trustworthy path. Next section shows the demonstration of the above mentioned algorithm on an illustrated problem.
The complete flowchart of the algorithm is:

![Flowchart for the IWD algorithm](image)

**Fig 2: Flowchart for the IWD algorithm**

Next we describe the above algorithm with the help of example:

**Fig 3: Topology of the sensor network**

In this topology, A can be considered as the source node from where data is transmitted. Sub-graph of all the nodes is calculated first. Sub-graph value of a node is the number of nodes below that particular node. Node A has all the nodes, i.e., B, C, D, and so on connected to it. This is done using the depth first traversal algorithm [14]. Hence the value of the sub-graph of all nodes is:

- A = 17, B=7, D=4, E= 2, F=4, G=1, H=2, M=2
- C, I, J, K, L, N, O, P, Q and R have a sub graph value of 0.

Now, the decision condition for the node is checked. A, B, D, E, F, G, H and M are decision nodes so they get 1 and rest of them get a value of 0. The various iterations of the IWD are shown below.

**Iteration 1:** The IWD starts from starting node A as shown in Figure 4.

![Path selection by IWD from A](image)

**Fig 4: Path selection by IWD from A**

Note: Visited path (Vc)= A>B

As the water-drop, have four paths to travel, i.e., via B, C, D and E. So, in order to find out which node to be traversed next, the probability of two nodes is computed using the respective link soild and on the basis of the next node for the 1st path is selected.

The calculation of the various parameters is:

- Soil (A, B) = 2*7+1*1 = 15
- Soil (A, C) = 2*0+1*0 = 1 (as sub-graph(j)-cond(j)==0)
- Soil (A, D) = 2*4+1*1 = 9
- Soil (A, E) = 2*2+1*1 = 5
- P (A, B) = Soil (A, B)/ \{Soil (A, B) + Soil (A, B) + Soil (A, B) + Soil (A, B)\} = 15/30
- P (A, C) = 1/30
- P (A, D) = 9/30
- P (A, E) = 5/30

The node with a higher probability is chosen first. In this case probability of node B is higher than that of node C, D and E so node B is selected. The probability is calculated using the Step 2.2 of the algorithm mentioned above. The time taken by the IWD to move from A–B is calculated using Step 2.3 of the algorithm mentioned above:

- Time (A, B) = max \{(17–7, 1)/Velocity (Vwld)\} = 10/100 = 0.1 sec

The soil carried by the IWD and its velocity is computed which is then subtracted from the soil of the link A-B to compute the remaining soil on the link using steps 2.3 and 2.4:

- ∆Soil (A, B) = 1/(1 + 1*0.1) = 0.9
- ∆Velocity (A, B) = 100 + 1/1 + 1*15 = 100.0625
- Updated Soil (A, B) = 0.9*15 - 0.1*0.9 = 13.41
- Soil (iwd) = 0.9

So the path will be: A – B.

Now, after reaching node B the water-drop has two available nodes for selecting the next node, which is shown in Figure 5.

![Path selection by IWD from B](image)

**Fig 5: Path selection by IWD from B**

For pages other than the first page, start at the top of the page, and continue in double-column format. The two columns on the last page should be as close to equal length as possible.
\[ \Delta \text{Velocity} \left( B, F \right) = 100.0625 + \frac{1}{1+1*9} = 100.1625 \]

\[
\text{Updated Soil} \left( B, F \right) = 0.9*0 - 0.1*0.970 = 8.003
\]

\[
\text{Soil (iwd)} = 0.9 + 0.970 = 1.87
\]

So the path will be: \( A \rightarrow B \rightarrow F \).

Similarly, after reaching node F the water-drop has to decide the next node to be traversed as in Figure 6.

The calculations for the various parameters are shown below:

\[
\text{Soil} \left( F, L \right) = 2*0 + 1*0 = 1
\]

\[
\text{Soil} \left( F, M \right) = 2*2 + 1*1 = 5
\]

\[
P \left( F, L \right) = \frac{\text{Soil} \left( F, L \right)}{\text{Soil} \left( F, L \right) + \text{Soil} \left( F, M \right)} = \frac{1}{6} = 0.167
\]

\[
P \left( F, M \right) = \frac{5}{6} = 0.83
\]

As the probability of node M is greater so the water drop will flow to node M.

\[
\text{Time} \left( F, M \right) = \frac{4 - 2}{100.069} = 0.0199 \text{ sec}
\]

\[
\Delta \text{Soil} \left( F, M \right) = \frac{1}{1 + 1*0.0199} = 0.98
\]

\[
\Delta \text{Velocity} \left( F, M \right) = 100.069 + \frac{1}{1+1*8.003} = 100.180
\]

\[
\text{Updated Soil} \left( F, M \right) = 0.9*8.003 - 0.1*0.98 = 7.105
\]

\[
\text{Soil (iwd)} = 2.85 + 0.98 = 3.83
\]

So the path will be: \( A \rightarrow B \rightarrow F \).
Now,
Time (M, R) = {−2.0}/100.365= 0.0199 sec
∆Soil (M, R) = 1/ (1 +1*0.98) = 0.98
∆Velocity (M, R) = 100.365 + 1/ 1*1= 100.865
Updated Soil (M, R) = 0.9*1-0.1*0.98 = 0.802

So the path will be: A→B→F→M→R

In a similar manner, we apply the algorithm and compute various paths. The path with the maximum weight indicates the most trustworthy path.
Path 1: A → B → F → M → Q (Total weight of path is: 30)
Path 2: A → B → F → M → R
(Total weight of path is: 26.617)
Path 3: A → B → F → L (Total weight of path is: 22.9476)
Path 4: A → B → G → N (Total weight of path is: 17.9881)
Path 5: A → D → H → O (Total weight of path is: 15)
Path 6: A → D → H → P (Total weight of path is: 9.6476)
Path 7: A → D → I (Total weight of path is: 7.99)
Path 8: A → E → J (Total weight of path is: 6)
Path 9: A → E → K (Total weight of path is: 5.21)
Path 10: A → C (Total weight of path is: 3.88)

This data can be shown in Table 1.

<table>
<thead>
<tr>
<th>Path No</th>
<th>Path</th>
<th>Weight</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A → B → F → M → Q</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>A → B → F → M → R</td>
<td>26.617</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>A → B → F → L</td>
<td>22.9476</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>A → B → G → N</td>
<td>17.9881</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>A → D → H → O</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>A → D → H → P</td>
<td>9.6476</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>A → D → I</td>
<td>7.99</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>A → E → J</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>A → E → K</td>
<td>5.21</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>A → C</td>
<td>3.88</td>
<td>10</td>
</tr>
</tbody>
</table>

The proposed algorithm uses IWD concept to generate all the possible independent paths with their priority, which is depicted in the path strength Table 1. The weight of each path is calculated by adding the soil on all the links of the path. It shows that the path with highest weight has the highest priority. The beauty of the algorithm is that it traverses all the paths as well as the nodes and computes the longest and the best path. Weight can be said as the trust value of that path and ∆Soil (i, j) is the trust of a particular communication between node i and j.

4. Conclusions
Managing trust in Wireless Sensor Networks (WSNs) is a crucial issue and necessary steps should be taken for security purpose in an efficient, accurate and robust way. Providing this management would notably increase the security in such a sentient environment, supporting thus its development and deployment.

In this paper, we have proposed a new method which is biologically inspired and is used to find the most reputable path leading to find the trustworthy nodes in the network. The topology here used is a tree topology, however we can apply this algorithm to any of the network graph structure where there are nodes asking for the services known as source nodes and there are nodes providing certain services as destination. Specifically we have explained how to measure the quality of the path. This in turns computes the trust value of a particular link.

5. REFERENCES


