

A Prediction Algorithm based on Markov Chains for finding the Minimum Cost Path in a Mobile Wireless Sensor Network

Germán A. Montoya
 System and Computing Engineering Department
 Universidad de Los Andes
 Bogotá, Colombia
 ga.montoya44@uniandes.edu.co

Yezid Donoso
 System and Computing Engineering Department
 Universidad de Los Andes
 Bogotá, Colombia
 ydonoso@uniandes.edu.co

Abstract—In this paper we propose the usage of a prediction technique based on Markov Chains to predict nodes positions with the aim of obtain short paths at minimum energy consumption. Specifically, the valuable information from the mobility prediction method is provided to our distributed routing algorithm in order to take the best network decisions considering future states of network resources. In this sense, in each network node, the mobility method employed is based on a Markov model to forecast future RSSI states of neighboring nodes for determining if they farther or closer within the next steps. The approach is evaluated considering different algorithms such as: Distance algorithm, Distance Away algorithm and Random algorithm.

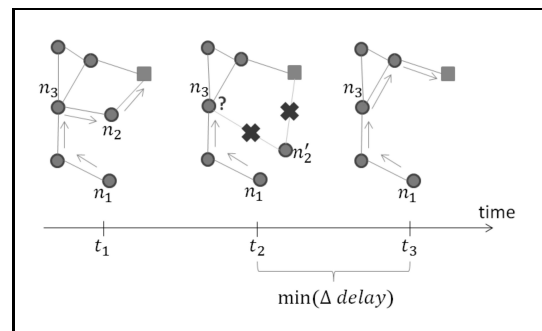
Keywords-Markov Model; RSSI; MWSN;

I. INTRODUCTION

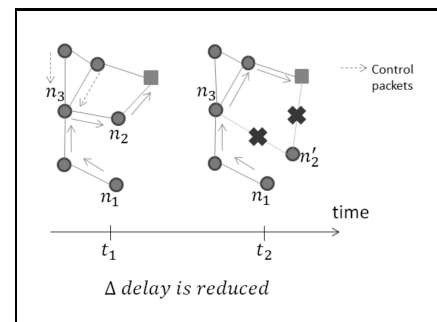
The advances of WSN have allowed attaching the sensors to an entity such as objects, animals or humans, to monitor a physical variable presented in its environment. However, the sensors are equipped with limited batteries whereby it is required to implement energy efficient routing techniques to extend the lifetime of the sensors as much as possible [1][2]. In addition, communication disruptions caused by mobility in wireless sensor networks introduce undesired delays which affect the network performance in delay sensitive applications, such as health monitoring applications. Due to these applications deal with health states, illness and continuous medical supervision, a base station should no experiment delays from the information collected by the sensors [3][4].

Given the scenario described above, a possible solution would consist to implement energy efficient routing techniques considering the sensor position to know the mobility level of the network. Based on this information, it is possible to determine the nodes that considerably affect the communication performance of the network. Some of these solutions propose the usage of sensors equipped with GPS devices, called GPS non-free approaches. However, these GPS non-free solutions have in most of cases drawbacks such as high implementation costs, delays for acquiring position information and non-accurate position information [5]. In addition, these types of solutions require an extra chip for the GPS [6], whereby more

energy consumption is experimented. For these reasons, our work will not take into account sensors equipped with GPS devices. Thus, in order to be aware of the network mobility we are going to use RSSI measurements, which indicates an approximated distance between two nodes.



(a) Problem



(b) Solution

Fig. 1: Problem Definition

The figure 1.a) presents the problem we want to solve. Suppose we have a MWSN where at time t_1 there is a communication path between the source sensor node n_1 and the base station. However, at time t_2 , the node n_2 moves away from the node n_3 , causing a communication disruption for

carrying the information from n_1 to the base station. Once n_3 has realized of this problem, at time t_3 , n_3 has to perform routing corrections in order to reestablish the communication path between n_1 and the base station. The communication reestablishment between n_1 and the base station can be perfectly performed using routing techniques, but at the expense of introduce an undesired delay in this communication path. In some applications these delays can be omitted because do not affect the purpose itself of the application, but in other ones, such as delay sensitive applications like health monitoring, this disadvantage might mean a very low network performance.

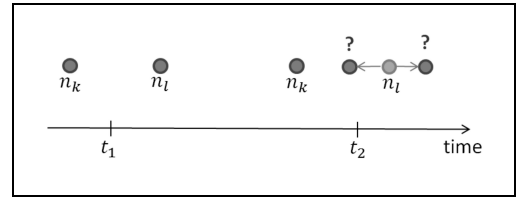
Given the problem above, our proposal consists to use a predicting technique which is described in the figure 1.b) [7][8]. It consists of the same situation showed in the figure 1.a), but in this case, at time t_1 , the node n_3 receive information that indicates the node n_2 will rapidly be away from its communication range, at time t_2 . Given this information, n_3 , at time t_1 , is also analyzing a possible candidate which could replace n_2 , in the case n_2 fails in a future time. If, indeed, at time t_2 the node n_2 fails because it has moved away from n_3 , this node at time t_2 can promptly reestablish the communication path between n_1 and the base station, reducing the delay described in the figure 1.a).

In order to solve the problem presented above, we propose to use a predicting method based on a Markov Model for estimating future RSSI states for a node with the aim of minimizing the delay experimented in the network. In this sense, our approach will be evaluated considering a Gauss-Markov mobility mode [9] where the mobility nodes can be considered predictable in order to test our prediction algorithm. Our work pretends to show an increasing network performance in terms of end-to-end delay and energy consumption against different algorithms such as: Distance algorithm, Distance Away algorithm and Random algorithm, which will be described in detail in the next sections.

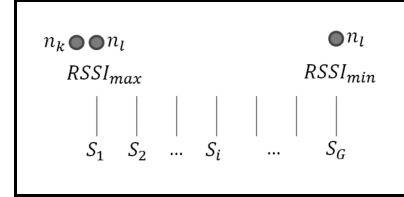
II. PROPOSED SOLUTION

In order to solve the problem presented above, we propose to use a predicting method based on a Markov Model for estimating future RSSI states for a node with the aim of minimizing the delay experimented in the network. For this purpose, a detailed explanation, supported along the following figures, will be presented.

In relation to the figure 2.a), suppose we have a network compound of two nodes: n_k and n_l , where n_l is a neighboring node of n_k . There are two times, t_1 and t_2 , at which our little network is evolving in time. At time t_1 the node n_l is located at certain distance from n_k . However, at time t_2 we want to predict if n_l will be farther or closer (or at the same distance in t_1) from n_k .



(a) Possible movement of n_l .



(b) RSSI States.

Fig. 2: Defining Markov States.

Respect to the figure 2.b), there is a minimum and maximum distance at which n_l can be located in order to establish a communication link with n_k . At the minimum distance, n_l will have a maximum RSSI, $RSSI_{max}$, and, at the maximum distance, n_l will have a minimum RSSI, $RSSI_{min}$. At t_2 , n_l could be located at any distance between $RSSI_{min}$ and $RSSI_{max}$. Our goal consists to estimate the location between $RSSI_{min}$ and $RSSI_{max}$ at which n_l will be in a future time (in this case, t_2). Theoretically, there are infinite locations between $RSSI_{min}$ and $RSSI_{max}$, but for our model we assume discrete locations equitably spaced. These possible locations, at which n_l could be, are called states. In this sense, at a future time t_2 , n_l could be at S_1, S_2, S_r or S_G , where G is the maximum number of states. The initial probability of n_l for being at any state S_i is $1/G$, which is called *Initial Probability Distribution of set S* (π), can be expressed as follows:

$$\pi = \{P_{s_1}, P_{s_2}, \dots, P_{s_G}\} \quad (1)$$

According to the figure 3.a), suppose we want to know the probability to go from the the state S_2 to the state S_4 , which is calculated with the following expression:

$$P_{24} = \frac{N(S_2, S_4)}{\sum_{j=1}^G N(S_2, S_j)} \quad (2)$$

Where $N(S_i, S_j)$ is the number of times that the state S_i follows state S_j .

This expression can be extensible for the rest of probabilities, as it is indicated in the following expression:

$$P_{ij} = \frac{N(S_i, S_j)}{\sum_{j=1}^G N(S_i, S_j)} \quad (3)$$

In this sense, we have the probability to go from any state S_i to any state S_j . These probabilities can be expressed in a matrix, which is called *Transition Matrix*:

III. IMPLEMENTATION

We have designed a Mobile Wireless Sensor Network Simulator in MATLAB, which has the following basic network components:

- *Destination node*: it is the final node that will receive a data message. In our simulations this node will always be the last network node.
- *Source node*: This node will have a data message, which must arrive to the destination node. In our simulations this node will always be the first network node.
- *Connected node*: If a message arrive to this node, this node knows the path to achieve the destination node.

In order to test the Prediction technique above, our simulator is compound of the following main processes:

- *Forwarding node selection*: When a node has a data message, this process consists to select properly a neighbour node as a forwarding node, which is selected according to the following priorities:
 - If among the neighbour nodes there is the destination node, then, the forwarding node is the destination node.
 - If among the neighbour nodes there is not the destination node, but there is a connected node, then, the forwarding node is the connected node.
 - If among the neighbour nodes there is not a destination node neither a connected node, then, the forwarding node is a node given by the Prediction method.
- *Sink refreshing*: This process consists to determine which nodes will be connected nodes at each certain period. This refreshing process is required due to network mobility, since it causes that connected nodes established in a previous state period, they will not possibly be connected nodes in the next period.
- *Loop detection*: It is important that a message can achieve the destination node, whereby it is necessary avoiding the message fall into a loop.

- *Prediction at each k-state*: At each network state the *Transition Matrix* (T) is calculated for all network nodes, except the destination node. Remember that this Transition Matrix stores the probability of each node to be at certain distance level respect to their neighbour nodes.
- *Prediction for selecting a forwarding node*: As we say before, if among the neighbour nodes there is not a destination node neither a connected node, then, the forwarding node is a node given by the Prediction method. This forwarding node is selected based on the information given by the Transition Matrix.

In order to test the Prediction algorithm performance, we have designed more algorithms with the aim of doing comparisons and obtain valuable information. Next, there is a description of each algorithm respect to its forwarding node process selection:

- *Distance Algorithm*: Considering there is not a destination or a connected node among the neighbour nodes of

$$T = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1G} \\ P_{21} & P_{22} & \dots & P_{2G} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ P_{G1} & P_{G2} & \dots & P_{GG} \end{bmatrix} \quad (4)$$

In relation to the figure 3.b), suppose that in a current time t_1 , n_l is at state S_3 and we want to estimate the future state of n_l at a future time t_p . For this purpose, we can apply the following expressions:

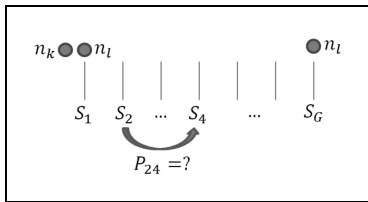
$$\pi_p = \pi * T^p \quad (5)$$

$$S_p = \max\{\pi_p\} \quad (6)$$

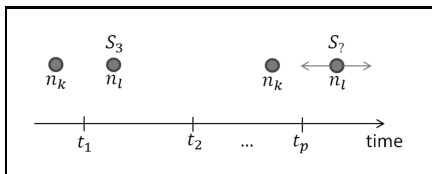
$$S_p = \max\{P_{s_1}, P_{s_2}, \dots, P_{s_G}\} \quad (7)$$

According to the expression 7, n_k can finally obtain the most probable future state at which n_l will be at a time t_p , and use this information for routing decisions in order to reduce the delay caused for probable communication disruptions in the future.

The present approach will be evaluated considering a Gauss-Markov mobility mode [9] where the mobility nodes can be considered predictable in order to test our prediction algorithm. Our work pretend to show an increasing network performance in terms of end-to-end delay and energy consumption against an approach without using a mobility prediction method and other approaches existent. Additionally, we will compare our algorithm results against a mathematical model optimization which minimizes energy consumption considering delay and network resources constraints.



(a) Probability to go from state S_2 to S_4 .



(b) Predicting the future state of n_l .

Fig. 3: Defining Markov States.

a current node, the forwarding node is the node with the shortest distance to the current node. The current node is the one that currently has a message that must arrive to the destination node.

- *Distance Away Algorithm*: Considering there is not a destination or a connected node among the neighbour nodes, the forwarding node is the node with the longest distance to the current node.
- *Prediction Algorithm*: Considering there is not a destination or a connected node among the neighbour nodes, the forwarding node is the node with the best probability to be near to the current node.
- *Prediction Away Algorithm*: Considering there is not a destination or a connected node among the neighbour nodes, the forwarding node is the node with the best probability to be far away to the current node.
- *Random Algorithm*: Considering there is not a destination or a connected node among the neighbour nodes, the forwarding node is a random node.

IV. RESULTS

In this section we will present the main results for the different algorithms showed in the previous sections. The metrics used for showing these results are: *Energy Consumption* and *Hops*. The *Energy Consumption* metric indicates the energy wasted by all the network nodes until the destination node is found. The *Hops* metric indicates the amount of hops needed to find the destination node. These two metrics are showed versus the number of network nodes. In addition, in order to obtain valuable statistical results, the performance evaluation of each algorithm and the mathematical model was made for 10000 tests for each network size. The next figures show the performance of the different algorithms and the mathematical model proposed for finding the minimum cost path in the network.

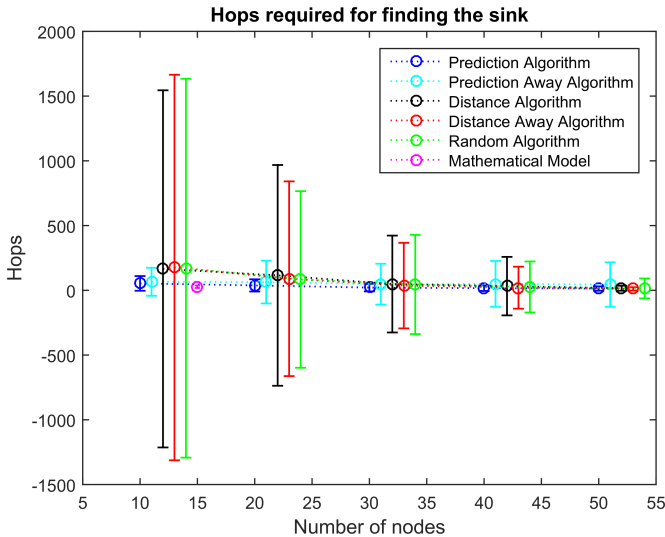


Fig. 4: Hops Performance along the Network Size.

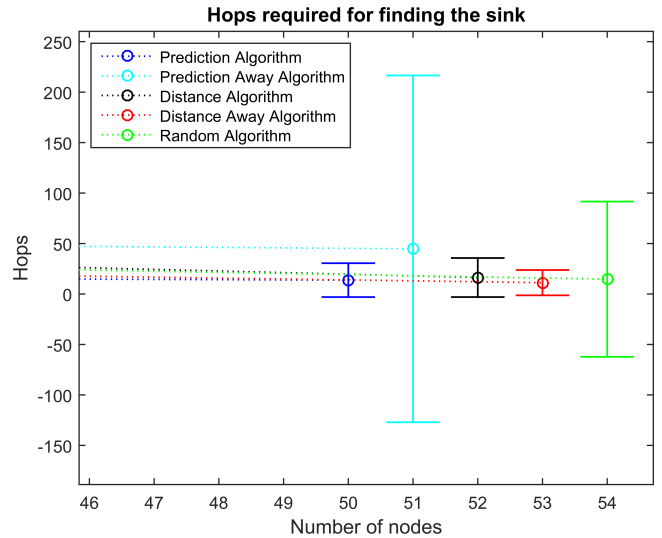


Fig. 5: Hops for 50 nodes.

The figure 4 shows the hops performance by the different algorithms along the network size. From this figure, as the size decreases the performance of Distance, Distance Away and Random algorithms gets worse because they require more hops to find the destination node. By contrast, the performance of Prediction and Prediction Away algorithms is better than the other ones because it requires less hops to find the destination node. This can be explained by the usage of prediction techniques, which offer more reliable paths. The following figures are focused in each network size.

The figure 5 shows the algorithms hops performance for 50 nodes. The number of hops for each algorithm is presented in the table I. The best performance is obtained by the Distance algorithm, while the Prediction algorithm is second best because the network size is big (50 nodes), allowing more path alternatives for the Distance Away algorithm to find faster the destination node.

TABLE I: Hops for 50 nodes.

	Hops	Ranking
Prediction Algorithm	13.72	2
Prediction Away Algorithm	44.71	5
Distance Algorithm	16.30	4
Distance Away Algorithm	11.20	1
Random Algorithm	14.66	3

The figure 6 shows the algorithms hops performance for 40 nodes. The number of hops for each algorithm is presented in the table II. The best performance is obtained by the Prediction algorithm because the number of nodes begins to decrease compared with the 50 nodes scenario, generating less path alternatives for the others algorithms and, then, thanks to the reliable feature given by the prediction technique, this algorithm can achieve faster the destination node.

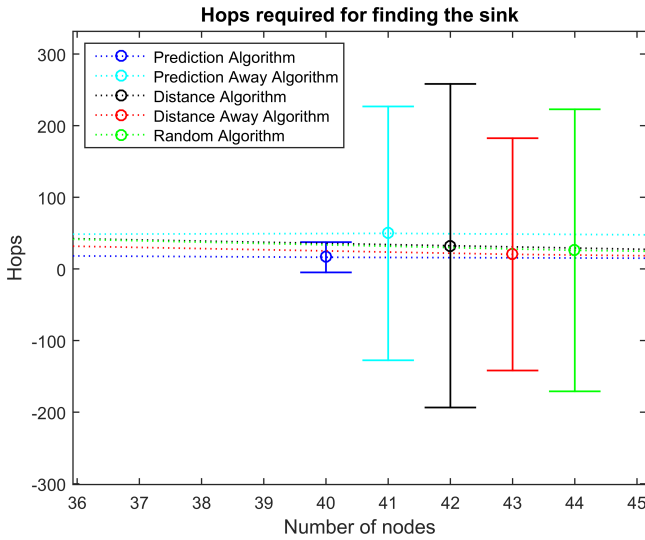


Fig. 6: Hops for 40 nodes.

TABLE II: Hops for 40 nodes.

	Hops	Ranking
Prediction Algorithm	16.24	1
Prediction Away Algorithm	49.61	5
Distance Algorithm	32.38	4
Distance Away Algorithm	20.31	2
Random Algorithm	25.95	3

The figure 7 shows the algorithms hops performance for 30 nodes. The number of hops for each algorithm is presented in the table III. The best performance is obtained again by the Prediction algorithm for the same reason as the previous figure. The less size of the network, the less path alternatives will have the rest of algorithms.

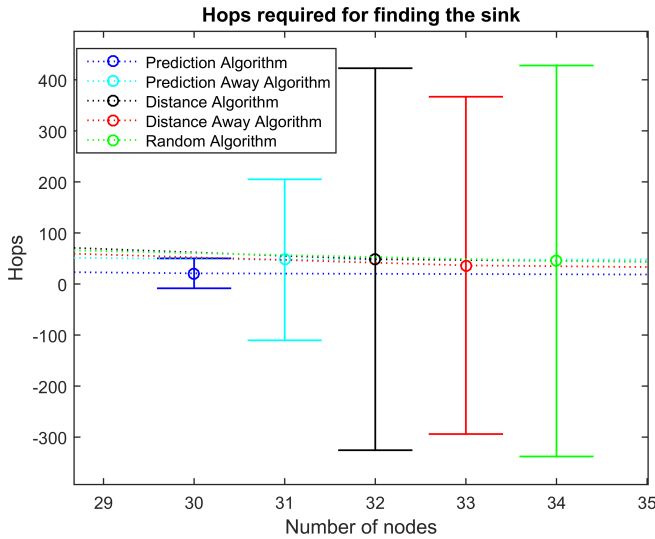


Fig. 7: Hops for 30 nodes.

TABLE III: Hops for 30 nodes.

	Hops	Ranking
Prediction Algorithm	20.86	1
Prediction Away Algorithm	47.38	4
Distance Algorithm	48.53	5
Distance Away Algorithm	36.42	2
Random Algorithm	45.05	3

The figure 8 shows the algorithms hops performance for 20 nodes. The number of hops for each algorithm is presented in the table IV.

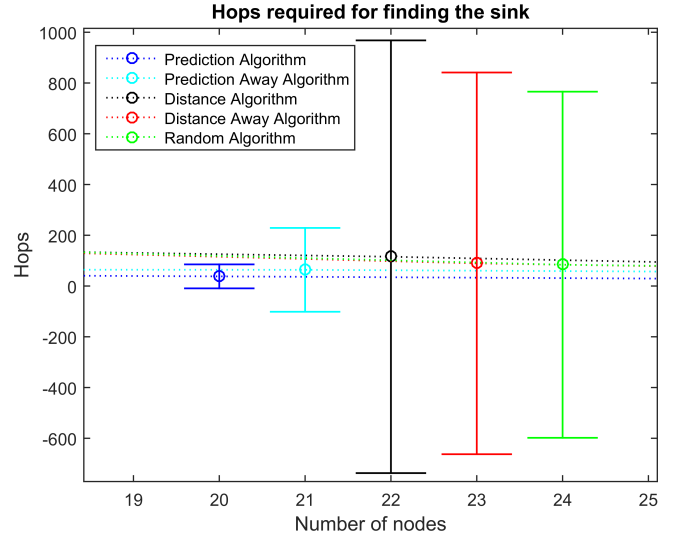


Fig. 8: Hops for 20 nodes.

TABLE IV: Hops for 20 nodes.

	Hops	Ranking
Prediction Algorithm	37.99	1
Prediction Away Algorithm	63.75	2
Distance Algorithm	115.38	5
Distance Away Algorithm	89.30	4
Random Algorithm	83.85	3

The figure 9 shows the algorithms hops performance for 10 nodes. The number of hops for each algorithm is presented in the table V. Here we can notice the large difference in terms of hops of using prediction techniques compared with not-using prediction techniques. This means that if our network has few nodes and, as a consequence, it is more difficult to find a path to the destination node, our prediction algorithm is capable of obtain a large advantage respect the others algorithms for finding the destination node. This advantage is represented in the hop different respect to the second algorithm in the ranking, which is 25.75 hops of difference. This comparison is among the Prediction and Prediction Away algorithms. However, if the comparison is done between the Prediction algorithm and the best algorithm that does not use prediction techniques (the Random Algorithm), the advantage of using the Prediction algorithm is even higher (45.85 hops). This indicates using prediction techniques are suitable when finding

paths is a critical task, that is, when the network is compound of few nodes. Notice that in addition there are presented the results for the mathematical model, which obviously presents the best performance, showing a hops performance difference of 25.81 respect to the Prediction algorithm. Notice that the mathematical model tests where suitable in terms of time execution and memory usage for a maximum of 15 nodes. For instance, a test of 30 nodes or even 20 nodes, the mathematical model solution unfortunately never ended. For this reason, only solutions for 10 nodes is provided to be compared with the prediction routing algorithm and no-prediction routing algorithms.

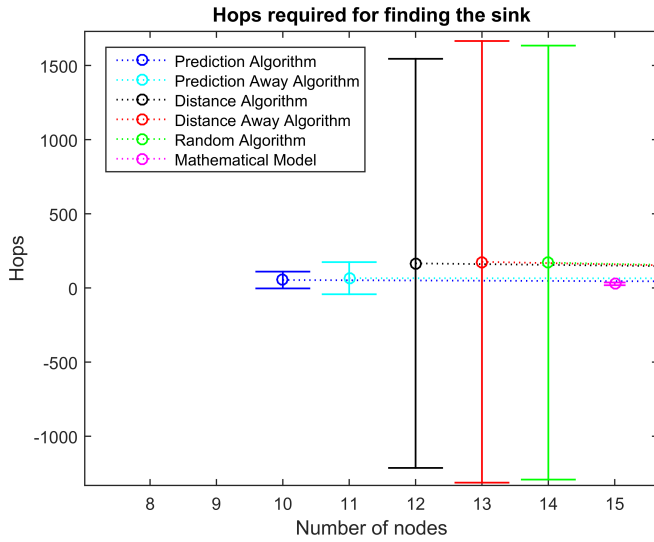


Fig. 9: Hops for 10 nodes.

TABLE V: Hops for 10 nodes.

	Hops	Ranking
Prediction Algorithm	53.38	1
Prediction Away Algorithm	65.75	2
Distance Algorithm	165.62	3
Distance Away Algorithm	176.07	5
Random Algorithm	170.92	4
Mathematical Model	27.57	

The figure 10 shows the algorithms energy consumption performance for different network size. This figure indicates that the Prediction Algorithm is suitable for low size network, showing that it starts to be efficient in terms of energy consumption from almost 30 nodes to 10 nodes.

V. CONCLUSIONS

We proposed the usage of a prediction technique in the context of a mobile wireless sensor network with the aim of the shortest path possible from a source node to a destination node. Employing this technique allowed to building the most reliable path for finding the destination node and at the same time it allowed to obtain the shortest path to the destination node. In other words, the reliability offered by the prediction

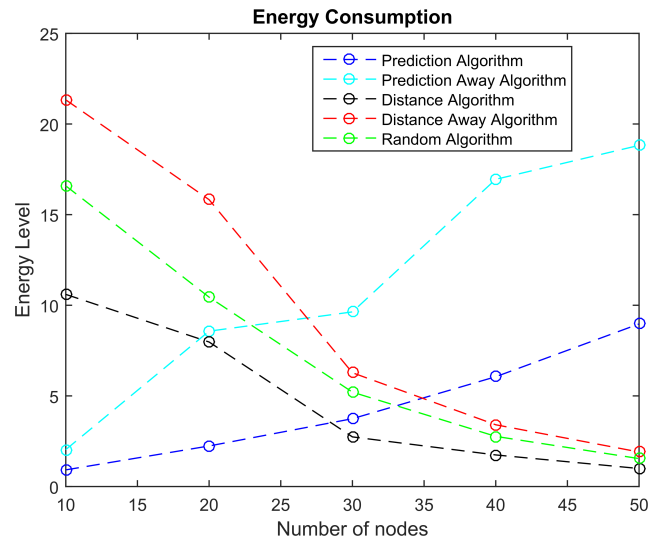


Fig. 10: Energy Consumption of the Network.

technique allowed to select the most stable forwarding nodes in terms of their network connectivity. In this sense, through the prediction technique it was less likely that a data message would be in isolated network zones, and then, there was a higher probability for reaching the destination node by the data message. For this reason, when the number of network nodes was scarce, 10 or 20 nodes, the prediction algorithm performance was too high in comparison with the rest of the algorithms, obtaining 45.85 and 112.24 hops of difference with the second best no-prediction algorithm for the 10 and 20 nodes of network size. The impact of this finding is very interesting. Suppose a cattle application where the network nodes (20 nodes) changes each 100 milliseconds. This means that if we use the prediction algorithm, a data message will reach the destination node 11.22 seconds faster than the second best no-prediction algorithm. This time, 11.22 seconds, could be a significant advantage in delay sensitive applications where the timeliness is an imperative factor.

In terms of energy consumption, a prediction technique is suitable for scarce networks (10, 20 or 30 nodes) because the energy consumption was the less than the rest of algorithms. This energy performance besides to the hops performance make the prediction algorithm totally suitable for scarce networks, that is, MWSN applications where the number of nodes is not too high and it is required data messages arrive to the destination nodes as soon as possible.

REFERENCES

- [1] I. F. Akyildiz and M. C. Vuran, *Wireless Sensor Networks*, 2010.
- [2] J. Zheng and A. Jamalipour, *Wireless Sensor Networks: A Networking Perspective*, 2009.
- [3] A. A. Ahmed, An enhanced real-time routing protocol with load distribution for mobile wireless sensor networks. *Computer Networks*, vol.57, 2013.
- [4] A. A. Ahmed, *Real-Time Wireless Sensor Networks*. University of Virginia, 2007.

- [5] B. Buchli, F. Sutton and J. Beutel. GPS-Equipped Wireless Sensor Network Node for High-Accuracy Positioning Applications. *Wireless Sensor Networks Lecture Notes in Computer Science*, Vol. 7158, pp 179-195, Springer, 2012
- [6] S. Li, X. Ma, X. Wang, M. Tan. Energy-efficient multipath routing in wireless sensor network considering wireless interference. *Journal of Control Theory and Applications*. Vol. 9, Issue 1, pp. 127-132. 2011.
- [7] G. M. de Arajo, J. Kaiser and L. B.Becker . Genetic Machine Learning Approach for Link Quality Prediction in Mobile Wireless Sensor Networks. *Cooperative Robots and Sensor Networks*. 2014.
- [8] G. M. de Arajo, J. Kaiser and L. B.Becker. An Optimized Markov Model to Predict Link Quality in Mobile Wireless Sensor Networks. 2012.
- [9] J. A. Torkestani Young. Mobility prediction in mobile wireless networks. *Journal of Network and Computer Applications*. Vol.35, 2012.
- [10] A Community Resource for Archiving Wireless Data At Dartmouth. <http://crawdad.org/>.