On Base Station Localization in Wireless Sensor Networks

H. E. KIZILOZ

Abstract— **Wireless sensor networks (WSN) has been a prominent topic for the past decade. WSN consist of multiple sensor nodes, which collect and convey data to the base station(s). Sensor nodes are expected to run on batteries, and it makes energy the scarce resource for sensor nodes. The energy expenditure of a sensor node mainly depends on data transmission, which is exponentially affected by transmission distance. Consequently, if sensor nodes forward their data to the base station directly, distant sensor nodes will exhaust quickly. On contrary, minimization of transmission distance for each sensor node, i.e., each node transmits its data to the closest sensor node on its path to the base station, depletes the energy of sensor nodes that are closer to the base station fast. As a result, the flow balance in the network must be optimized. In this study, we investigate the effect of optimization of the base station location along with flow balance optimization. For this purpose, we compare five different localization methods on different topologies; three statically located linear programming approaches, a dynamically located nonlinear programming approach and a heuristic-based hybrid approach. Experimental results indicate that lifetime improvement of up to 42% is possible in selected scenarios.**

Index Terms—**lifetime optimization, linear programming, wireless sensor networks.**

I. INTRODUCTION

WIRELESS SENSOR NETWORKS (WSN) has gained much attention recently. The emergence of cheap devices and the functionalities it provides make WSN an interesting field of topic. Even though military and/or emergency applications were the main causes in its appearance, it has become a common concept in academia, commercial market, and even among hobbyists [1]. Anyone can deploy their WSN at their homes, offices, etc. as of today [2]. Moreover, insights on the Internet of Things and urban computing raise interest in WSN [3].

A typical WSN consists of multiple sensor nodes and at least one sink called the base station. A sensor node is responsible for two actions: i) sensing, which is collecting data using its sensor. Any type of sensors may be used for data collection, e.g. humidity, air temperature, pressure, acoustic, magnetic, etc. ii) conveying its data to the base station. For this purpose,

sensor nodes may directly communicate with the base station, or they may choose to use other sensor nodes as a relay.

A possible drawback of WSN is that the number of sensors in a WSN topology could increase up to hundreds or thousands, and all sensor nodes are assumed to work on a battery, hence, they have limited and nonrenewable energy. Since energy is the scarce resource for a sensor node, the energy expenditure of the nodes must be optimized. Even though energy dissipation for sensing may be application or sensor type-specific; in general, a sensor node uses most of its energy for data transmission. As a result, minimization of the amount of energy used for data transmission helps to maximize the lifetime of the WSN.

Energy usage for data transmission is highly bound to the distance between communicating parties. Amount of energy required to transmit one-bit data increases exponentially as distance increases. However, selecting the closest sensor node as the relay causes the sensor nodes that are closer to the base station to exhaust all of its energy very quickly. Since we define the lifetime as the period until the first sensor node depletes its whole energy, the lifetime would be short in such a case. As a result, flows between sensor nodes must be optimized in such a way that it does not exhaust the nodes that are either close to or far from the base station.

Given the sensor node and base station locations, optimizing the flows between parties has been a resolved issue [4]. Since distance is the key factor for energy expenditure, there have been several studies for optimizing sensor node locations in literature [5, 6, 7]. The mentioned studies proposed different methods for sensor node deployment that increases lifetime and ensures coverage.

In this study, we look at the problem from a different angle. The main purpose of this study is to investigate the effects of using different base station localization methods over the lifetime in the WSN domain. For this purpose, we utilize the model proposed in [4] and compare five base station localization methods, in terms of their effect on the lifetime of the network and their execution times, with empirical results.

The rest of the paper is organized as follows. The literature review is given in Section 2. The model is described in Section 3, and experimental results are given in Section 4. We conclude the paper with concluding remarks in Section 5.

HAKAN EZGI KIZILOZ, is with Department of Computer Engineering University of Turkish Aeronautical Association, Ankara, Turkey, (e-mail: hakanezgi@etu.edu.tr).

https://orcid.org/0000-0002-2306-6008

II. RELATED WORK

In this section, related work about base station location optimization in literature is given.

Oyman and Ersoy [8] proposed an iterative search for finding the optimal number and locations of base stations. After the sensor nodes are deployed, their model initiates with one base station. At each iteration, they apply the k-means clustering algorithm to locate the base stations at the center of each cluster and estimate a lifetime value accordingly. Iterations continue to increase the number of the base station until the estimated lifetime reaches a predetermined threshold value. Kim *et al.* [9] proposed a mixed-integer linear programming model for optimal multi-sink positioning and routing. In their model, sensor nodes are located into a topology first, and then some of them are selected as base stations during the optimization phase. Experiment results indicate improvement in lifetime and fairness when compared to multi-sink aware Minimum Depth Tree model.

In 2010, Güney *et al.* [10] proposed mixed-integer linear programming models for optimal base station location and routing in heterogeneous WSN for two separate objectives: minimizing energy expenditure, and minimizing financial cost. Also, they provided such techniques to find lower bounds. Later in 2012, the authors proposed a hybrid approach utilizing mixed-integer linear programming and Tabu Search to search for the optimal base station location [11]. In both models, sensor nodes and the base station(s) can only be located at predefined locations. Random deployment of the sensor nodes is also investigated in the latter study [11]. Different heuristic approaches also exist. Tripathi *et al.* [12] proposed a heuristic approach for finding the optimal base station location in two-tiered WSN. In their approach, they begin from the centroid of the nodes, assign link weights with respect to distances, and deploy the base station to the weighted average of the sensor nodes. Experimental results indicate that their proposed algorithm improves lifetime, yet the authors state that it does not find the theoretically optimal location. Fouad *et al.* [13] proposed a Particle Swarm Optimization (PSO) based heuristic method for finding the optimal base station location. According to experimental results, their method improves the performance of both topology construction and maintenance phases, and hence improves lifetime. Pan *et al.* [14] proposed a compact PSO method to find the optimal base station location. This method approximates the genetic algorithm and conventional PSO in terms of obtained results, but it requires less resources.

The work by Tung and Binh [15] is the most similar to this study. They use the same mathematical model, and they also compare five different base station localization methods. Their proposed method, integrated greedy method, improves the lifetime by *10%*. Similarly, a recent study by Saha *et al.* [16] shows that their probabilistic localization method performs well.

The mobility of the base stations has attracted interest recently. Cayirpunar *et al.* [17] analyzed possible mobility patterns for multiple base stations, whereas Cicek *et al.* [18] provided mathematical formulations for flying base stations.

III. MODEL

In this section, we describe the employed energy model and base station localization methods separately, and in the specified order, in the following subsections.

A. Energy Model

In this study, we utilize the linear programming (LP) based energy model described in [4]. The model takes positions of the sensor nodes and the base station as input and optimizes the flow balance to maximize the lifetime of the network accordingly.

The energy model used in this study is as follows:

Maximize t

Subject to:

$$f_{ij} \geq 0, \forall (i,j) \in A \tag{1}$$

$$f_{ij} = 0, \forall (i,j) \in A, if\ i\ is\ B\ or\ i\ =\ j\ or\ d_{ij} \geq d_{max} \tag{2}$$

$$\sum_{j \in V \cup B} f_{ij} = \sum_{j \in V \cup B} f_{ji} + s_i t, \forall i \in V \tag{3}$$

$$E_{rx} \sum_{j \in V} f_{ji} + \sum_{j \in V} f_{ij} E_{tx,ij} + f_{iB} E_{tx,iB} \leq energy_i, \forall i \in V \tag{4}$$

In this model, the objective is to maximize lifetime, t, with respect to the constraints given with equations (1) through (4). We would like to mention that, B represents the base station, whereas V and A represent all sensor nodes and all edges (links) in the network, respectively. Moreover, data transmission (flow) from node i to node j is shown with f_{ij}. Among the constraints, Eq. (1) indicates that all flows must be non-negative. Flow in the network can either be positive, indicating that flow between sensor nodes exists; or zero, indicating that there is no flow between sensor nodes. In Eq. (2), we introduce additional constraints on flow. According to these additional constraints: i) base station cannot transmit data, ii) a sensor node cannot transmit data to itself, and iii) a sensor node cannot transmit data to another sensor node that is located further than d_{max} meters, which is *82.92 m*. Flow balance at each sensor node is guaranteed with Eq. (3), that is, each node must transmit all the data they receive and produce. In this constraint, s_i represents the amount of data produced by each sensor node at a unit time interval (1 bps). Finally, Eq. (4) is a constraint for energy. The total amount of energy used by each sensor node cannot exceed its initial energy. In this model, sensor nodes are only responsible for collecting and conveying data. Complex tasks such as data aggregation or compression is not a concern. Also, the energy used for sensing and/or sleeping is considered as negligible. As a result, they only use energy for transmitting and receiving information. Energy used for receiving one bit of information is constant, $E_{rx} = \rho$, and equal to the electronics energy (*50nJ*). On the other hand, sensor nodes spend a variable amount of energy to transmit one bit of information, which is calculated with respect to the distance of the receiving node or

base station: $E_{tx,ij} = \rho + \varepsilon\, d_{ij}^{\alpha}$. In this formula, ε represents the amplifier energy (*100pJ*), d_{ij} represents the Euclidean distance between sensor nodes i and j, and α represents path loss exponent. The $E_{tx,iB}$ value in Eq. (4) is a special case of $E_{tx,ij}$, where the receiver is not a sensor node but the base station. Finally, the path loss exponent, α, is considered low ($\alpha=2$) on a free space propagation topology without any obstacles, but it increases ($\alpha=4$) as propagation gets harder due to physical obstacles, such as rock, building, etc.

B. Base Station Localization Methods

In this study, we compare five different base station localization methods. Three of these methods employ static approaches, meaning that the base station is localized initially, and the network's lifetime is maximized with the given energy model. Another method dynamically searches for the best location for the base station, that maximizes the lifetime. The remaining approach is hybrid, and it utilizes a heuristic search algorithm, namely Particle Swarm Optimization (PSO). All the methods are described in detail as follows.

center: This is a static approach. The base station is simply located at the center of the area of interest, e.g. at the center of a circular topology. This approach is common in the literature.

mean: This is, again, a static approach. The base station is located at the mean values of *x*- and *y*-coordinates of the sensor nodes. Intuitively, this method seems to be a simple yet effective improvement to localization at the center.

random: Third and the final static approach. The base station is located at a random location inside the area of interest. A possible scenario for this model could be deploying both sensor nodes and the base station into a danger zone by throwing them from an airplane. In this scenario, the base station could be designed as a more powerful computing device with large enough energy, and a GSM connector.

dynamic: This method introduces two new variables to the model: *x*- and *y*-coordinates (position) of the base station. Rather than statically locating the base station, this method searches for the optimal location for the base station that maximizes the network's lifetime. The introduction of these two variables and their effects on the energy model (dynamically calculating the Euclidean distance between each sensor node and the base station) make it more complex, hence, the energy model becomes a nonlinear programming (NLP) model. Unlike an LP model, an NLP model may get stuck in a local optimum and cannot ensure global optima. Moreover, an NLP model may require more time to solve than an LP model. Nonetheless, it may still be valuable in this domain and it is worth testing it as a possible method.

hybrid: This hybrid approach utilizes PSO, a heuristic search algorithm. PSO algorithm consists of several particles searching for the best solution, i.e. *x*- and *y*-coordinates (position) of the base station that maximizes the fitness value, the lifetime. Each particle represents a possible solution. Initially, the particles are initialized randomly. Then, the fitness value for each particle is calculated. The fitness value is calculated as similar to the static approaches: the particle provides the position of the base station, and the LP energy model maximizes the lifetime. Once fitness calculation for each particle is complete, the particle having the best fitness value is

determined and all particles begin searching towards the best particle. Fitness evaluation, the best particle's determination, and updating each particle's new search location is called a generation in PSO. The algorithm continues until a termination criterion, i.e. completing a predefined number of generations, is met. Finally, the overall best particle through all generations is reported as the solution of the algorithm.

To eliminate any bias, all base station localization methods are tested on equal terms; that is, all sensor nodes are initially deployed randomly in a circular topology, then the base station is localized separately for each method on this network, and finally, the energy model maximized the network's lifetime. As a result, any deviation in the lifetime solely depends on the location of the base station, which is only affected by the localization method.

IV. EXPERIMENTAL RESULTS

We use GAMS [19] to solve our proposed model. We use the energy model described in the previous section, with each sensor node having an initial energy of *2W*. However, we would like to note that, initial energy and energy dissipation for transmitting and/or receiving data is hardware and application-specific; hence may vary in different hardware or applications. As a result, obtained lifetime results should not be evaluated as exact timespan in seconds, but should be considered as a ratio. Hence, we give normalized result values, i.e., all lifetime values are divided by the maximum lifetime value of the same context.

The hybrid method employs the PSO algorithm, which requires the determination of algorithm-specific and execution parameters. Algorithm specific parameters are obtained from [13], as follows: inertia weight w is *0.8*, cognitive coefficient c_1 and social coefficient c_2 are *2*, and the random numbers r_1 and r_2 are selected in the range *[0,1]*. The two execution parameters, population size and the number of generations, are selected as *5* and *12*, respectively.

A sample topology depicting base station locations in all five localization methods are given in Fig. 1.
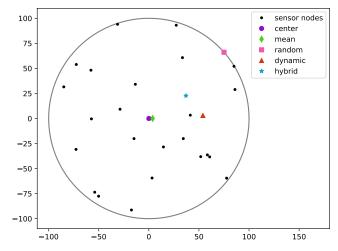

Fig.1. A sample topology with 25 sensor nodes, depicting base station locations for all localization methods.

In our experiments, we randomly deployed sensor nodes inside a circular topology having a radius of *100 m*. We used varying numbers of sensors (*25*, *50*, *100*, *250* and *500*) to

simulate sparse and dense topologies. Also, we chose two different values for path loss exponent, $\alpha = 2$ and $\alpha = 4$, to simulate free space and multi-path propagation, respectively. Finally, to eliminate possible bias due to randomness, we share the average results of *1000* different sensor node deployments when the number of sensors is *25*, *50* or *100*. Topologies having *250* or *500* sensor nodes are averaged over a smaller number of deployments (*100* and *30*, respectively) due to their larger execution time (see Table 1).

Experimental results for the two path loss exponents, $\alpha = 2$ and $\alpha = 4$, are given in Fig. 2 and Fig. 3, respectively. In both figures, normalized lifetime values are plotted as the number of nodes increases. From the figures, it is clear that random deployment of the base station is a bad choice, for all cases. On free space propagation, where $\alpha = 2$, base station localization is more important for sparse topologies than it is for dense topologies. The dynamic method improves the lifetime for up to *7.1%* in a sparse topology (where the number of sensor nodes is *25*). On the other hand, as the network gets denser, the lifetime of all methods (except for the random method) approximates to the maximum. The hybrid method, contrary to its good lifetime improvement performance, is not a viable option due to its huge execution time. In topologies where multi-path fading is a concern, $\alpha = 4$, the huge execution time of the hybrid method may be negligible due to its fruitful solutions. In sparse topologies with multi-path fading, the hybrid method improves lifetime by *27.5%*, *27.0%* and *4.4%* as compared to the center, mean and dynamic methods, respectively. Lifetime improvement becomes even more critical as the topology gets denser. The hybrid method improves lifetime by *42.6%* as compared to the center and mean methods, and by *15.1%* as compared to the dynamic method when the network consists of *500* sensor nodes.
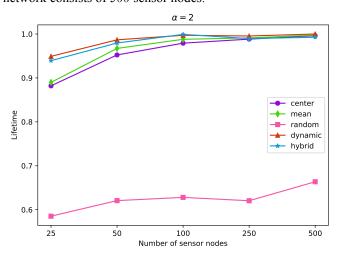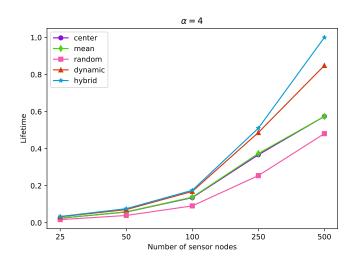

Fig.3. Normalized lifetime values with varying numbers of sensor nodes on multi-path propagation topology (α=4).

We also provide average execution time for solving a problem in each model in Table 1. All static approaches require the same and minimum amount of the time, in all cases. The hybrid method requires roughly *60* times more amount of time as compared to the static methods, in all cases. This outcome is natural due to its implementation and the selected execution parameters (population size and the number of generations). In its comparison to the dynamic method, it requires more time in all cases, yet, the difference is negligible for dense topologies with multi-path fading. The dynamic method requires a similar amount of time to those of static methods for sparse and medium-sized topologies (up to *100* nodes). For dense topologies, however, its execution time increases massively.

Our findings on the lifetime and the execution times are summarized as follows.

• The base station should not be located randomly, if possible.

• On free space propagation, where $\alpha = 2$, the dynamic method can be executed in less than 5 seconds for sparse and medium-sized topologies, and it improves lifetime. For dense topologies, however, the base station can simply be located at the mean or the center.

• On multi-path propagation, where $\alpha = 4$, the hybrid method dominates all other methods in terms of improving lifetime. For sparse topologies, the dynamic method may still be a viable option considering its low execution time (under 5 seconds). For dense topologies, however, the hybrid method performs better and its usage is crucial.

Fig.2. Normalized lifetime values with varying numbers of sensor nodes on free-space propagation topology (α=2).

TABLE I

AVERAGE EXECUTION TIMES OF SOLVING A PROBLEM FOR EACH TYPE OF TOPOLOGY AND LOCALIZATION METHOD (IN SECONDS).

	$\alpha = 2$					$\alpha = 4$				
# of nodes	25	50	100	250	500	25	50	100	250	500
center	0.55	0.57	0.75	3.92	40.14	0.52	0.57	0.66	3.38	37.59
mean	0.54	0.56	0.75	3.94	40.89	0.53	0.56	0.66	3.27	34.60
random	0.54	0.55	0.75	3.96	41.79	0.52	0.56	0.67	3.42	32.91
dynamic	0.59	0.81	3.55	91.53	992.60	0.56	0.87	4.12	185.30	2114.65
hybrid	30.55	32.13	46.45	235.68	2517.39	29.96	31.54	44.11	197.82	2120.15

V. CONCLUSION

Energy is the scarce resource of sensor nodes in wireless sensor networks (WSN), and it directly affects the lifetime of the network. The main factor in energy depletion is data transmission, which is highly dependent on transmission distance. Linear programming models have been employed so far to balance flows in a given WSN topology, to prevent premature energy depletion of a specific sensor node. These models work well, however, most of these studies assign a static location for the base station, e.g. the center. In this study, we analyze the effects of employing different base station localization methods on lifetime.

We evaluate five localization methods on different types of topologies: sparse or dense topologies on a free space channel or multi-path propagation. We compare the obtained results in terms of their effects on the lifetime and their execution time. Experimental results show that the dynamic method provides viable solutions in sparse and medium-sized topologies. It improves lifetime in a considerable amount of time (in under 5 seconds). In dense topologies, however, it becomes incompetent. Dense topologies having free space channel should employ the center or mean methods, whereas, the use of the hybrid method is crucial in dense topologies with multi-path propagation.

The main advantages and disadvantages of these methods are as follows. Among these methods, the random and center methods require no additional information and calculation of the base station location is imminent. The remaining methods require the knowledge of the sensor node locations, which may be hard or even impossible to acquire in a danger zone, e.g. natural disaster site. Moreover, the calculation time may be considered as a disadvantage, especially for the dynamic and hybrid methods. The most outstanding advantage is lifetime improvement, which is up to *42%* in specific scenarios. Additionally, these methodologies are directly related, hence applicable to Internet of Things (IoT) studies, as WSN is a specialized form of it. Briefly, WSN may be considered as a more controlled and compact version of IoT, as IoT may consist of heterogeneous and geographically distributed devices.

Possible future work for this study could be making a similar analysis for networks having multiple base stations. Also, the effects of these methods on topologies combining different propagation channels could be investigated, e.g. $\alpha = 2$ for communication between some nodes, and $\alpha = 4$ for others.

REFERENCES

[1] M. A. Gray and P. N. Scherer, "Web services framework for wireless sensor networks," in SERVICE COMPUTATION 2014: The Sixth International Conferences on Advanced Service Computing, IARIA, 2014, pp. 15–23.
[2] S. Newberry and G. S. Gupta, "Wireless sensor based home automation as an educational springboard," in Sensors Applications Symposium (SAS), 2015 IEEE. IEEE, 2015, pp. 1–6.
[3] B. Rashid and M. H. Rehmani, "Applications of wireless sensor networks for urban areas: A survey," Journal of Network and Computer Applications, vol. 60, pp. 192–219, 2016.
[4] Z. Cheng, M. Perillo, and W. B. Heinzelman, "General network lifetime and cost models for evaluating sensor network deployment strategies," IEEE Transactions on mobile computing, vol. 7, no. 4, pp. 484–497, 2008.
[5] H. Subir, G. Amrita, S. Sanjib, D. Avishek, and D. Sipra, "A lifetime enhancing node deployment strategy in WSN," in International Conference on Future Generation Information Technology. Springer, 2009, pp. 295–307.
[6] X. Liu, "Sensor deployment of wireless sensor networks based on ant colony optimization with three classes of ant transitions," IEEE Communications Letters, vol. 16, no. 10, pp. 1604–1607, 2012.
[7] S. Singh, S. Chand, R. Kumar, and B. Kumar, "Optimal sensor deployment for WSNs in grid environment," Electronics Letters, vol. 49, no. 16, pp. 1040–1041, 2013.
[8] E. I. Oyman and C. Ersoy, "Multiple sink network design problem in large scale wireless sensor networks," in Communications, 2004 IEEE International Conference on, vol. 6. IEEE, 2004, pp. 3663–3667.
[9] H. Kim, Y. Seok, N. Choi, Y. Choi, and T. Kwon, "Optimal multi-sink positioning and energy-efficient routing in wireless sensor networks," in International Conference on Information Networking. Springer, 2005, pp. 264–274.
[10] E. Güney, N. Aras, İ. K. Altınel, and C. Ersoy, "Efficient integer programming formulations for optimum sink location and routing in heterogeneous wireless sensor networks," Computer Networks, vol. 54, no. 11, pp. 1805–1822, 2010.
[11] E. Güney, N. Aras, İ. K. Altınel, and C. Ersoy, "Efficient solution techniques for the integrated coverage, sink location and routing problem in wireless sensor networks," Computers & Operations Research, vol. 39, no. 7, pp. 1530–1539, 2012.
[12] R. K. Tripathi, Y. N. Singh, and N. K. Verma, "Two-tiered wireless sensor networks - base station optimal positioning case study," IET Wireless Sensor Systems, vol. 2, no. 4, pp. 351–360, 2012.
[13] M. M. Fouad, V. Snasel, and A. E. Hassanien, "Energy-aware sink node localization algorithm for wireless sensor networks," International Journal of Distributed Sensor Networks, vol. 11, no. 7, 2015.
[14] J. S. Pan, T. K. Dao, and T. S. Pan, "Compact particle swarm optimization for optimal location of base station in wireless sensor network," in International Conference on Genetic and Evolutionary Computing. Springer, 2016, pp. 54-62.
[15] N. T. Tung and H. T. T. Binh, "Base station location-aware optimization model of the lifetime of wireless sensor networks," Mobile Networks and Applications, vol. 21, no. 1, pp. 10-17, 2016.
[16] K. Saha, J. Aich, S. Chakraborty, and S. Bose, "Probabilistic Sink Placement Strategy in Wireless Sensor Network," in Contemporary Advances in Innovative and Applicable Information Technology. Springer, 2019, pp. 169-175.
[17] O. Cayirpunar, B. Tavli, E. Kadioglu-Urtis, and S. Uludag, "Optimal mobility patterns of multiple base stations for wireless sensor network lifetime maximization," IEEE Sensors Journal, vol. 17, no. 21, pp. 7177-7188, 2017.
[18] C. T. Cicek, H. Gultekin, B. Tavli, and H. Yanikomeroglu, "UAV Base Station Location Optimization for Next Generation Wireless Networks: Overview and Future Research Directions," in 2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS). IEEE, 2019, pp. 1-6.
[19] G. D. Corporation, "General Algebraic Modeling System (GAMS) Release 24.2.1," Washington, DC, USA, 2013. [Online]. Available: http://www.gams.com/

BIOGRAPHIES

HAKAN EZGI KIZILOZ is currently an assistant professor at the University of Turkish Aeronautical Association. After receiving his BSc degree from the Mathematics department of TOBB University of Economics and Technology in 2008, he received his MSc and PhD degrees from the Computer Engineering department of the same university in 2010 and 2016, respectively. His research interests include usable security and optimization.