OPTIMAL PLACEMENT OF RELAY STATIONS IN WIRELESS SENSOR NETWORKS

A Paper
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Kiran Kumar Vallabh

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Computer Science

March 2013

Fargo, North Dakota

North Dakota State University

Graduate School

Title

Optimal Placement of Relay Stations in Wireless

Sensor Networks

By

Kiran Kumar Vallabh

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Kendall E. Nygard

Chair

Dr. Simone Ludwig

Dr. Darshi De Saram

Approved by Department Chair:

4/2/2013

Date

Dr. Brian M. Slator

Signature

ABSTRACT

Wireless sensor networks (WSNs) are a collection of nodes organized into a cooperative network with sensing, processing and transmitting capabilities. WSNs are becoming an increasingly prominent technology that can be used in diverse application areas. In WSNs, cooperative relay stations are projected as one of the most cost effective solutions to meet the demanding requirement of capacity enhancement.

In this paper, major concerns of the wireless sensor networks addressed are optimizing the number of relay stations required for covering the desired percentage of sensor nodes by optimal placement of relay stations and optimal assignment of the sensors to the relay stations. The joint problem of relay station placement and coverage is formulated into a mixed integer program which is solvable by commercial GAMS software with Xpress-MP Solver. Sensitivity analysis is carried out, along with a case study to demonstrate the performance gain of the model.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my deepest gratitude to my adviser, Dr. Kendall E. Nygard, who was abundantly helpful and offered invaluable assistance, support and guidance throughout this paper. I would also like to thank Dr. Simone Ludwig, Dr. Darshi De Saram for being a part of my graduate supervisory committee, and without whose knowledge and support this study would not have been successful.

Special thanks to the faculty and staff of the Computer Science Department for their support throughout my master's program. I would also like to thank my friend Vinay Gonela, without whom I would not have come this far.

Finally, I am grateful to my parents, whose love and guidance are with me in whatever I pursue.

On a different note, many people have been a part of my graduate education, and I am grateful to all of them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Wireless Sensor Networks (WSNs) are ad hoc multi-hop systems containing sensor nodes connected by wireless links organized into a cooperative network. The flourishing research on WSNs is driven by the advances in Micro-Electro Mechanical Systems (MEMS) technology, Complementary Metal Oxide Semiconductor (CMOS) logic, and wireless networking [11, 12]. These advances have made it possible to deploy sensors in a particular geographical area and sense the environment by calculating physical parameters such as heat, sound, temperature, light etc. [1]. A sensor network consists of a group of sensor nodes, which are able to sense, process and transmit data. Sensor networks have many potential applications which include monitoring battlefields and detecting enemies, sense the environmental changes in oceans, forests, and atmosphere, providing security to the buildings and also monitoring human body, which is immensely useful in bio-medical applications [1].

WSNs are used to create macro-scale effects from micro-devices through synchronized activities of various sensors; thus connectivity is a hugely critical issue in WSN architecture design [10]. Wireless links are principally determined by transmit powers of sensors, and higher transmit power produces affluent connectivity. Though, in the context of untethered nodes, a primary design constraint is the limited energy resources [10]. Communications is a significant energy consumer, due to ground reflections from short antenna heights, so for slightly longer distances to be reached, the sensor needs to dispatch much greater transmit power [10]. The second reason for the restraining higher transmission power is the higher interference to on-going traffic. The higher the power a sensor transmits, the more the number of direct neighbors the sensor has, and the higher the negative impact the sensor has on the network throughput. The third reason is the lifespan of the network [14], which is determined by the lifespan of sensors as a whole. Wireless sensors are battery powered; either battery replenishment is prohibited by monetary considerations or it is impossible to recharge or replace a battery in a WSN. The fourth reason, but not the last, is the heat emitted by higher power, transmission may hamper the sensing function (i.e. Temperature sensors) [10].

On the other hand, these observations do not mean that the lesser the transmit power, the better. Very low transmit power may result in disconnected topology, and thus network breakdown. It may cause shooting up of the hop count for message broadcasting, thus ascending error rate and deteriorating throughput. Economically deploying a sensor network with remarkably low transmit power may be prohibitive, as the number of sensors needed may be significantly increased twice or thrice. For every wireless network to remain active, it has to maintain a threshold level of coverage for the area where sensor network is deployed ensuring balance between high coverage and longer network lifetime [5]. Most of the applications do not require 100% of the network coverage and only partial coverage is necessary to maintain the network to be active. These partial coverage requirements of the wireless networks also improves network lifetime further [1].

Consequently, a sensor network designer has to seek a compromise among performance, network lifetime, and cost. A number of problems have been formulated to study this compromise. Many of them focus on topology control by minimizing the maximum transmit power [15, 16] or minimizing total transmit power [17, 18] to maintain global topology. Optimal node placement is a hugely challenging problem that has been proven to be NP-Hard for most of the formulations of sensor deployment [2, 3].

To meet the growing demand and stringent design requirements for coverage extension, throughput and capacity enhancement, deploying relay stations has been considered as a promising solution to Point-to-Multi-Point (PMP) networks [20].

A network operator always wishes the most economical solution with the minimum deployment expenditure to provide a satisfactory service [19]. The Relay Station (RS) location for sensor networks in the network planning phase is critical and will address fundamental impacts on the subsequent service provisioning scenario [4].

This paper proposes and describes a mixed integer programming model; that can be used to find the optimal number of relay stations and positions for placing them in a wireless sensor network, also optimal assignment of sensors to the relay stations. The main aim of this mixed integer programming model is to propose the positions and number for relay stations, so that using a minimal number of relay stations, the desired amount of coverage for wireless sensors can be achieved. The relay stations are responsible for relaying data between the sensors and the base station. The relay stations considered in the study have sufficient power and are not directly connected through cable. The links between the relay stations and sensors are assumed relatively static, and a deterministic Time Division Multiple Access (TDMA)/ Code Division Multiple Access (CDMA) scheme can be utilized as the communication technique. This paper considers optimizing the number of relay stations and the relay station placement problem only. The model built will take the characteristics like transmission range of relay station, the desired percentage of coverage of sensors by the relay stations and the positions of the sensors as inputs and returns the number of relay stations and their positions. The remainder of the paper is organized as follows. Section 2 describes the related work done in this area. In Section 3, the problem statement and the mixed integer programming model formulation is described. In Section 4, the results are shown in a graphical format along with the tables supporting the graphs. In Section 5, a case study for Maine State was designed and conducted. Section 6 concludes the paper and also talks about the future directions this research work can take.

2. BACKGROUND STUDY

In [6], the authors mentioned that the technology of wireless networks can be useful for various applications like environmental monitoring, infrastructure management, public safety, health care, home and office security, transportation, military surveillance etc. A sensor network is a collection of sensors deployed in a location to perform specific tasks. The sensors were primarily used to sense or detect or to track a target or monitor a particular location. As the technology of the sensors evolved, there were several breakthroughs and achievements in this area enabling sensors to be useful in many more situations.

In [4], the authors addressed the task of relay station placement and relay time allocation in IEEE 802.16j Mobile Multi-hop Relay (MMR) networks. By incorporating advanced cooperative relaying technologies as Decode-Forward (D-F) and Compress-Forward (C-F), the authors aimed at finding the optimal location of a single relay station and the resource allocation for all the subscriber stations (sensors). The authors also conducted numerical analysis through some case studies, and demonstrated the performance gain by using the approach proposed in the paper for relay station placement and relay time allocation. The authors considered a practical deployment scenario, where each subscriber station imposes some amount of traffic demand during a specific time window. In a metropolitan area, the load on a particular subscriber station may vary based on the time of the day. The authors formulated the single relay station placement problem in multi subscriber station model in order to yield the optimal deployment, and resource allocation for each individual relay station for a given set of subscriber stations.

The major challenge in designing WSNs is the support of the functional, such as data latency, and the non-functional, such as data integrity and requirements while coping with the computation, energy and communication constraints [7]. Careful node placement can be a highly effective optimization means for achieving the desired design goals. In [7] the authors reported the research on optimized node placement in wireless sensor networks. The authors classified the placement strategies into static and dynamic depending on whether the optimization is executed at the time of deployment or while the network is in operation, respectively. For many, wireless sensor networks will consist of hundreds of nodes that operate on small batteries. Wireless sensor networks should be carefully managed in order to meet applications requirements while conserving energy. The aim of the authors is to assist the application designers in identifying alternative resolutions and select suitable approaches.

In [8] the authors studied the capacity enhancement problem by means of relay stations placement to achieve an efficient and scalable design in broadband wireless access networks. The authors developed an optimization framework to maximize the capacity as well as to meet the minimal traffic demand by each subscriber station. The problem of joint relay station placement and bandwidth allocation is formulated into a mixed integer nonlinear program. To avoid exponential computation time, the authors proposed a heuristic to solve the formulated problem efficiently. The authors conducted numerical analysis through case studies and demonstrated the performance gain of cooperative relaying and the comparison between the proposed algorithms against the optimal solutions. With the relay stations, the quality of wireless channels can be significantly improved not only by replacing one long distance low-rate link with multiple short-distance high rate links, but also due to the ability of circumventing any obstacles between subscriber stations and base station that may hinder the channel quality. In [9], the impact of relay station placement in IEEE 802.16j network performance is analyzed. A throughput maximization relay station placement problem is mathematically formulated as a binary integer programming problem. The authors proposed an efficient near-optimal placement solution to find the sub-optimal solution to the problem with large input size. The throughput performance shows that, with the strategy the authors proposed, the network capacity can be tremendously enhanced.

3. PROBLEM STATEMENT AND MODEL FORMULATION

This chapter discusses about the structure of the problem, formulating and building the mixed integer programming model.

3.1. Problem Statement

In wireless sensor networks, a large number of sensors with limited energy supply are responsible for relaying the sensed data hop by hop to the base station. The sensors neighboring the base station deplete their energy much faster than distant nodes, because they carry heavy traffic, which results in premature ending of the network lifetime. Commissioning relay stations can alleviate this problem by relaying data between the sensors and the base station.

Figure 1 presents the generalized structure of the problem studied in this paper. Let us consider an irregular shaped area, where i be the locations of the deployed sensors and j be the potential locations of the relay stations. Let d_{ij} be the distance between sensor at location { i } and relay station at location { j }. Let D_{Max} be the maximum transmission range of the relay station. Given such a structure, a mixed integer programming (MIP) model is proposed to determine the:

1. Optimal number of relay stations
2. Optimal locations of the relay stations
3. Optimal assignment of the sensors to the relay stations

Sensitivity analysis is carried out by varying the transmission ranges of relay stations, number of sensor nodes and percentage coverage of sensors to provide deep insights to the decision makers. In addition, a case study for Maine State is conducted to accommodate a profound understanding of the proposed model.

3.2. Assumptions

Following are some of the assumptions that are made in the modeling:

1. Sensors are assumed to be static. This means that the sensors do not change their positions in the WSN.
2. The potential locations for establishing the relay stations are pre-determined.

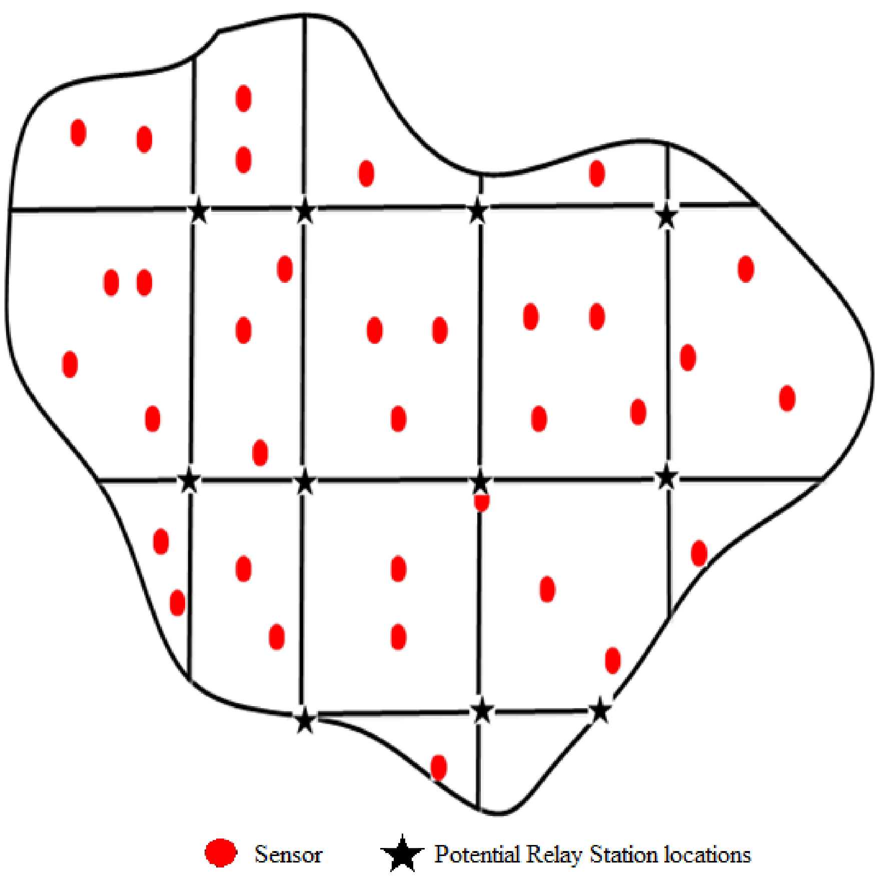


Figure 1. Structure of the problem

3.3. Model Formulation

This section describes the parameters, decision variables, objective function and the constraints of the mixed integer programming model.

3.3.1. Parameters

i Index for sensor locations where $i = \{1, 2, 3 \ldots I\}$

I Total number of sensors

j Index for relay station locations where $j = \{1, 2, 3 \ldots J\}$

J Total number of relay stations

d_{ij} Distance between sensor at location $\{i\}$ and relay station at location $\{j\}$

D_{Max} Maximum allowable transmission range of relay station

P Percentage of coverage and an element between 0 and 1

3.3.2. Free variables

Z Number of relay stations

3.3.3. Binary variables

Y_j 1, if relay station at location $\{j\}$ is open

0, else

X_{ij} 1, if sensor at location $\{i\}$ is assigned to relay station at location $\{j\}$

0, else

3.3.4. Objective function

The objective is to minimize the number of relay stations.

$$Min\ \ Z = \sum_{j} Y_j$$

Subject to,

Each sensor can only be assigned to a relay station if distance between the sensor at location { i } and relay station at location { j } is less than the maximum allowable distance (transmission range of the relay station).

$$d_{ij} X_{ij} \leq D_{Max} Y_j \ \forall i, \forall j$$

Each sensor can be assigned to no more than one relay station.

$$\sum_{j} X_{ij} \leq 1 \forall i$$

Eq. (4) presents the percentage of coverage area.

$$\sum_{i} \sum_{j} X_{ij} \geq P * I, \ \ where\ \ P \in [0,1]$$

3.4. Model Building

The mixed integer programming model is coded in General Algebraic Modeling System (GAMS) and solved using Xpress-MP Solver. The data for the distances between the sensors and relay station locations was generated randomly using the RAND () function in Microsoft Excel.

3.4.1. Algebraic modeling language

Algebraic modeling language is a high level computer programming language for describing and solving high complexity problems for large scale optimization problems. One advantage of algebraic modeling language is the similarity of its syntax to the mathematical notion of optimization problems, and this allows for a very concise and readable definition of problems in the domain of optimization.

3.4.2. General algebraic modeling system

[23] GAMS was the first algebraic modeling language. It is a high-level modeling system for mathematical programming and optimization. It consists of a language compiler and a stable of integrated high-performance solvers. GAMS is tailored for complex, large scale modeling applications, and allows you to build large maintainable models that can be adapted quickly to new situations. GAMS is specifically designed for modeling linear, nonlinear and mixed integer optimization problems. The system models problems in a highly compact and natural way.

GAMS language is formally similar to commonly used programming languages. Models are described in concise algebraic statements which are easy for both humans and machines to read. GAMS automatically generates each constraint equation, and lets the user make exceptions in cases where generality is not desired. Statements in models can be reused without having to change the algebra when other instances of the same or related problems arise. GAMS handles dynamic models involving time sequences, lags and leads and treatment of temporal endpoints.

Figure 2 shows the Dashboard of GAMS Integrated Development Environment.

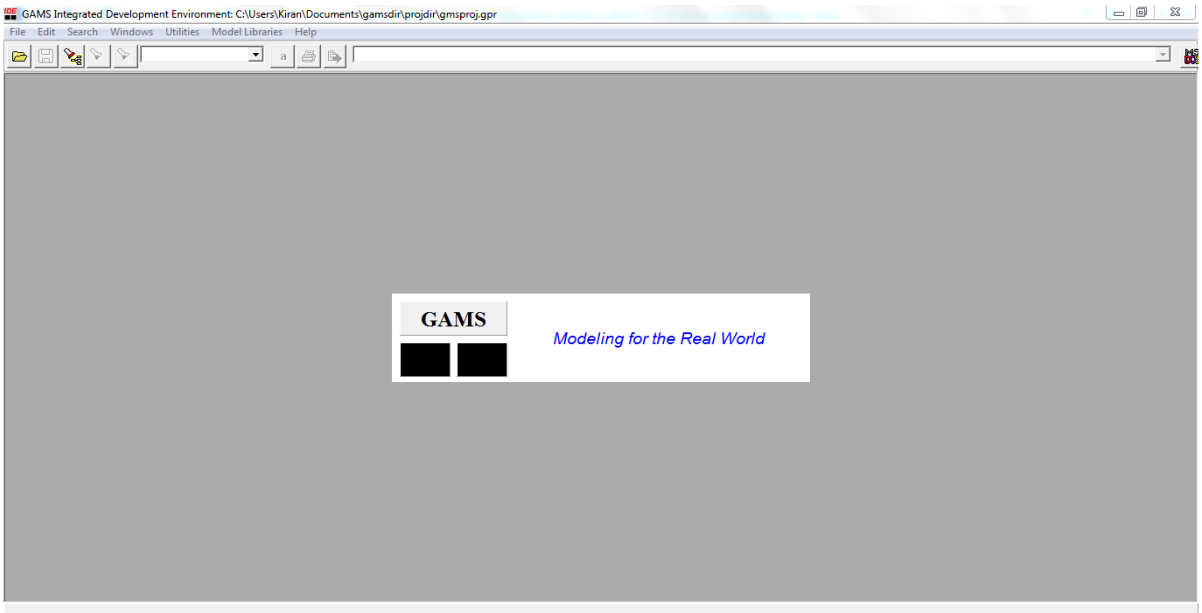


Figure 2. GAMS IDE dashboard

Figure 3 shows the Compiler of GAMS Integrated Development Environment.

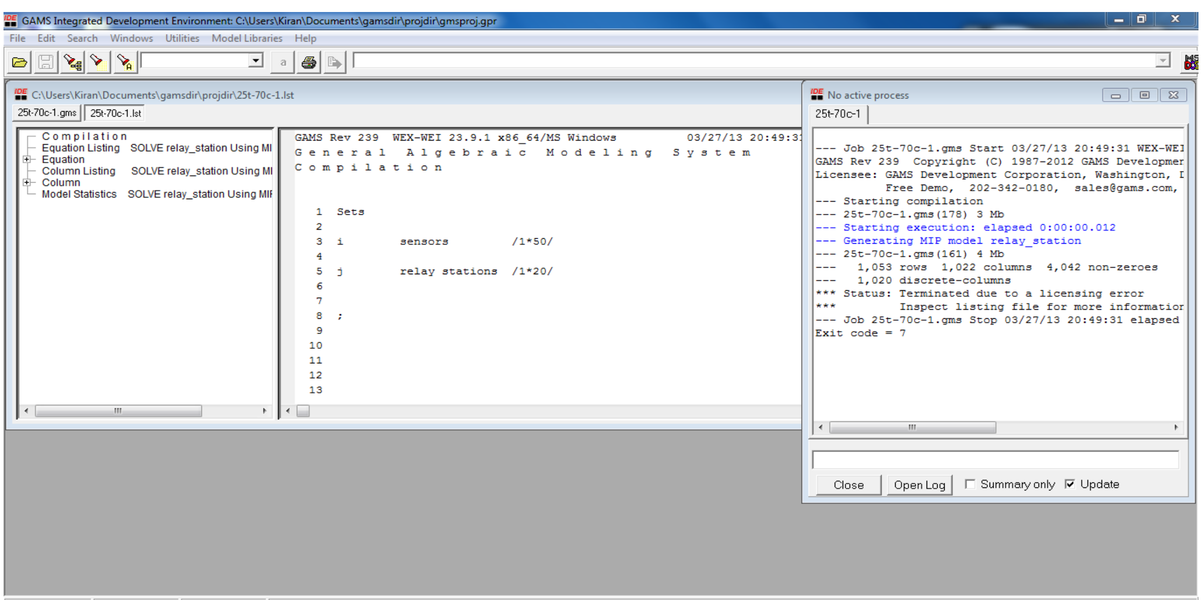


Figure 3. GAMS IDE compiler

3.4.3. Xpress-MP solver

[24] Xpress-MP Solver is a versatile, high performance optimization system. The system integrates a powerful simplex-based LP solver, a MIP module with cut generation for integer programming problems and a barrier module implementing an interior point algorithm for very large LP problems. Xpress-MP Solver runs only in conjunction with the GAMS modeling system.

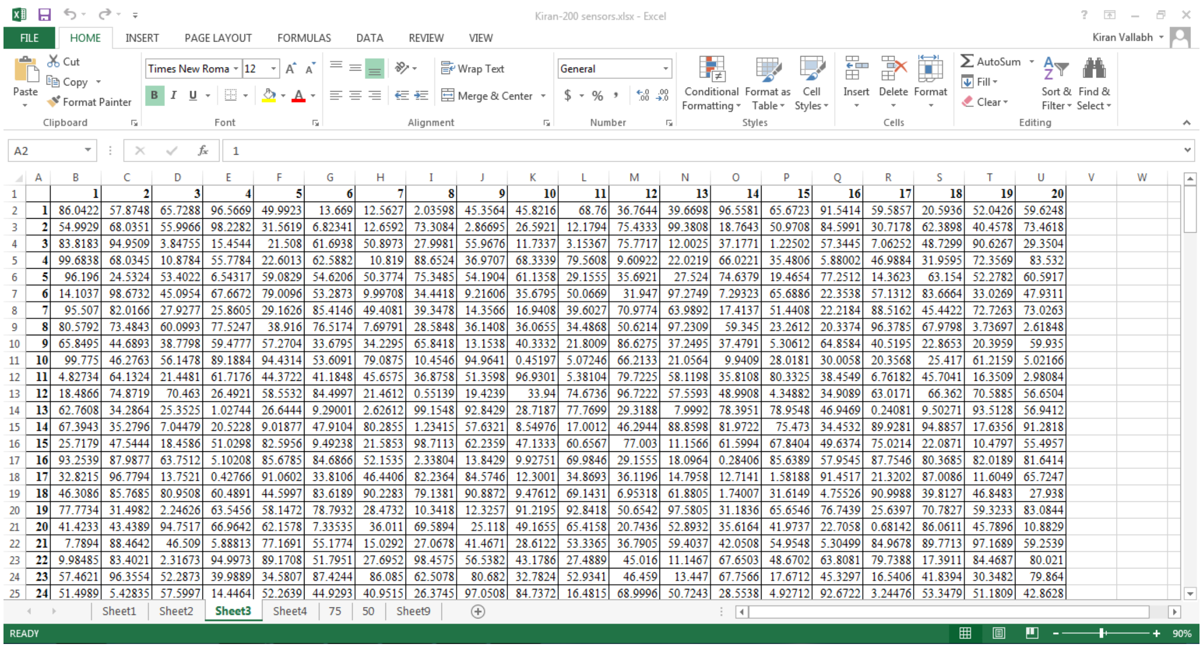
Figure 4 shows the NEOS Xpress-MP Solver page.

Figure 4. NEOS Xpress-MP solver

3.4.4. Microsoft Excel

[25] Microsoft Excel is a spreadsheet application developed by Microsoft. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for applications.

Figure 5 shows the Microsoft Excel spreadsheet with random data generated using the RAND () function.



	1	2	3	4	5	6	7	8	9	10
1	86.0422	57.8748	65.7288	96.5669	49.9923	13.669	12.5627	2.03598	45.3564	45.8216
2	54.9929	68.0351	55.9966	98.2282	31.5619	6.82341	12.6592	73.3084	2.86695	26.5921
3	83.8183	94.9509	3.84755	15.4544	21.508	61.6938	50.8973	27.9981	55.9676	11.7337
4	99.6838	68.0345	10.8784	55.7784	22.6013	62.5882	10.819	88.6524	36.9707	68.3339
5	96.196	24.5324	53.4022	6.54317	59.0829	54.6206	50.3774	75.3485	54.1904	61.1358
6	14.1037	98.6732	45.0954	67.6672	79.0096	53.2873	9.99708	34.4418	9.21606	35.6795
7	95.507	82.0166	27.9277	25.8605	29.1626	85.4146	49.4081	39.3478	14.3566	16.9408
8	80.5792	73.4843	60.0993	77.5247	38.916	76.5174	7.69791	28.5848	36.1408	36.0655
9	65.8495	44.6893	38.7798	59.4777	57.2704	33.6795	34.2295	65.8418	13.1538	40.3332
10	99.775	46.2763	56.1478	89.1884	94.4314	53.6091	79.0875	10.4546	94.9641	0.45197
11	4.82734	64.1324	21.4481	61.7176	44.3722	41.1848	45.6575	36.8758	51.3598	96.9301
12	18.4866	74.8719	70.463	26.4921	58.5532	84.4997	21.4612	0.55139	19.4239	33.94
13	62.7608	34.2864	25.3525	1.02744	26.6444	9.29001	2.62612	99.1548	92.8429	28.7187
14	67.3943	35.2796	7.04479	20.5228	9.01877	47.9104	80.2855	1.23415	57.6321	8.54976
15	25.7179	47.5444	18.4586	51.0298	82.5956	9.49238	21.5853	98.7113	62.2359	47.1333
16	93.2539	87.9877	63.7512	5.10208	85.6785	84.6866	52.1535	2.33804	13.8429	9.92751
17	32.8215	96.7794	13.7521	0.42766	91.0602	33.8106	46.4406	82.2364	84.5746	12.3001
18	46.3086	85.7685	80.9508	60.4891	44.5997	83.6189	90.2283	79.1381	90.8872	9.47612
19	77.7734	31.4982	2.24626	63.5456	58.1472	78.7932	28.4732	10.3418	12.3257	91.2195
20	41.4233	43.4389	94.7517	66.9642	62.1578	7.33535	36.011	69.5894	25.118	49.1655
21	7.7894	88.4642	46.509	5.88813	77.1691	55.1774	15.0292	27.0678	41.4671	28.6122
22	9.98485	83.4021	2.31673	94.9973	89.1708	51.7951	27.6952	98.4575	56.5382	43.1786
23	57.4621	96.3554	52.2873	39.9889	34.5807	87.4244	86.085	62.5078	80.682	32.7824
24	51.4989	5.42835	57.5997	14.4464	52.2639	44.9293	40.9515	26.3745	97.0508	84.7372

	11	12	13	14	15	16	17	18	19	20
1	68.76	36.7644	39.6698	96.5581	65.6723	91.5414	59.5857	20.5936	52.0426	59.6248
2	12.1794	75.4333	99.3808	18.7643	50.9708	84.5991	30.7178	62.3898	40.4578	73.4618
3	3.15367	75.7717	12.0025	37.1771	1.22502	57.3445	7.06252	48.7299	90.6267	29.3504
4	79.5608	9.60922	22.0219	66.0221	35.4806	5.88002	46.9884	31.9595	72.3569	83.532
5	29.1555	35.6921	27.524	74.6379	19.4654	77.2512	14.3623	63.154	52.2782	60.5917
6	50.0669	31.947	97.2749	7.29323	65.6886	22.3538	57.1312	83.6664	33.0269	47.9311
7	39.6027	70.9774	63.9892	17.4137	51.4408	22.2184	88.5162	45.4422	72.7263	73.0263
8	34.4868	50.6214	97.2309	59.345	23.2612	20.3374	96.3785	67.9798	3.73697	2.61848
9	21.8009	86.6275	37.2495	37.4791	5.30612	64.8584	40.5195	22.8653	20.3959	59.935
10	5.07246	66.2133	21.0564	9.9409	28.0181	30.0058	20.3568	25.417	61.2159	5.02166
11	5.38104	79.7225	58.1198	35.8108	80.3325	38.4549	6.76182	45.7041	16.3509	2.98084
12	74.6736	96.7222	57.5593	48.9908	4.34882	34.9089	63.0171	66.362	70.5885	56.6504
13	77.7699	29.3188	7.9992	78.3951	78.9548	46.9469	0.24081	9.50271	93.5128	56.9412
14	17.0012	46.2944	88.8598	81.9722	75.473	34.4532	89.9281	94.8857	17.6356	91.2818
15	60.6567	77.003	11.1566	61.5994	67.8404	49.6374	75.0214	22.0871	10.4797	55.4957
16	69.9846	29.1555	18.0964	0.28406	85.6389	57.9545	87.7546	80.3685	82.0189	81.6414
17	34.8693	36.1196	14.7958	12.7141	1.58188	91.4517	21.3202	87.0086	11.6049	65.7247
18	69.1431	6.95318	61.8805	1.74007	31.6149	4.75526	90.9988	39.8127	46.8483	27.938
19	92.8418	50.6542	97.5805	31.1836	65.6546	76.7439	25.6397	70.7827	59.3233	83.0844
20	65.4158	20.7436	52.8932	35.6164	41.9737	22.7058	0.68142	86.0611	45.7896	10.8829
21	53.3365	36.7905	59.4037	42.0508	54.9548	5.30499	84.9678	89.7713	97.1689	59.2539
22	27.4889	45.016	11.1467	67.6503	48.6702	63.8081	79.7388	17.3911	84.4687	80.021
23	52.9341	46.459	13.447	67.7566	17.6712	45.3297	16.5406	41.8394	30.3482	79.864
24	16.4815	68.9996	50.7243	28.5538	4.92712	92.6722	3.24476	53.3479	51.1809	42.8628

Figure 5. Microsoft Excel spreadsheet

4. RESULTS

This section describes various test cases that are performed with the mixed integer programming model developed. The different parameters for carrying out the test cases were the number of sensors, number of potential relay stations, and transmission range of relay station and desired percentage of coverage.

In the test cases, for each set of parameters, the number of relay stations opened and the execution time of the mixed integer programming model for that situation are captured. After running the model several times with different data sets and after gathering all the results; the mean, standard deviation and variance are calculated. Then one of the parameters is varied, and then the mean, standard deviation and variance are calculated, and this process is repeated by varying another parameter and the results are calculated. We determine various results, and the performance of the model is compared based on the number of relay stations and execution time.

4.1. Initial Study

The initial study is a preliminary analysis conducted to determine whether there is substantial evidence that the mixed integer programming model developed yields the desired output.

4.1.1. 100% coverage of sensor nodes

The data for distance between the sensor location and relay station location is randomly sampled in Microsoft Excel by using the RAND () function with the distance range from 0 to 100 miles. 50 sensor nodes and 20 potential relay station locations are considered and modeled for 100% coverage. The model is coded in GAMS and solved with Xpress-MP Solver.

Table 1 shows the number of relay stations opened when the transmission range is varied between 25, 50 and 75 miles.

Table 1. No. of relay stations opened for 50 sensors, 100% coverage

50 sensors, 100% coverage			
	25 mile transmission range	**50 mile transmission range**	**75 mile transmission range**
No. of relay stations	6	3	2

Figure 6 suggests that as the transmission range increases, the number of relay stations opened decrease.

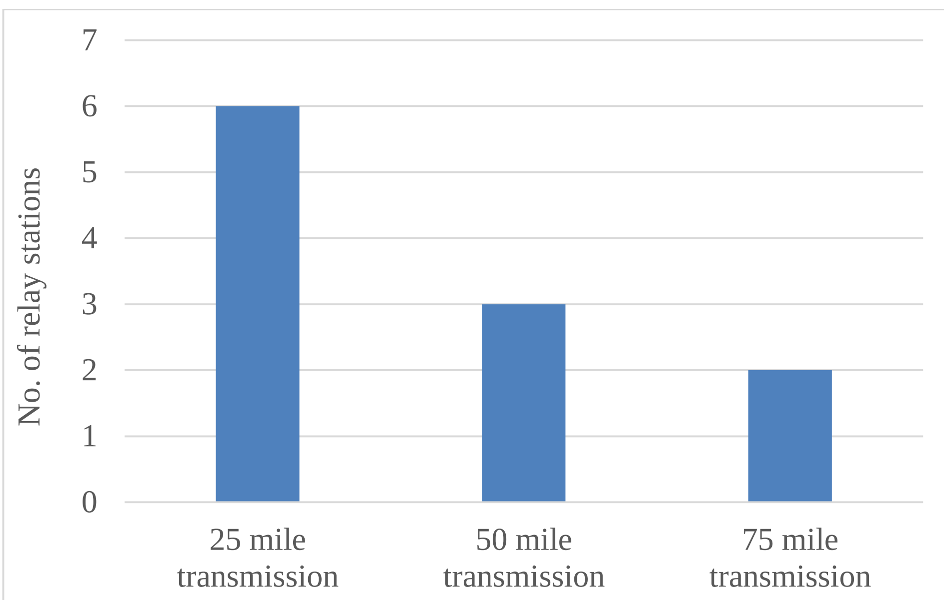


Figure 6. No. of relay stations opened for 50 sensors, 100% coverage

4.1.2. 90% coverage of sensor nodes

50 Sensor nodes and 20 potential relay station locations are considered and modeled for 90% coverage. The model is solved in commercial GAMS software with Xpress-MP Solver.

Table 2 shows the number of relay stations opened when the transmission range is varied between 25, 50 and 75 miles.

Table 2. No. of relay stations opened for 50 sensors, 90% coverage

50 sensors, 90% coverage			
	25 mile transmission range	**50 mile transmission range**	**75 mile transmission range**
No. of relay stations	5	2	2

Figure 7 suggests that as the transmission range increases, the number of relay stations opened decrease.

Figure 7. No. of relay stations opened for 50 sensors, 90% coverage

4.1.3. 70% coverage of sensor nodes

50 sensor nodes and 20 potential relay station locations are considered and modeled for 70% coverage. The model is solved in commercial GAMS software with Xpress-MP Solver.

Table 3 show the number of relay stations required when the transmission range is varied between 25, 50 and 75 miles.

Table 3. No. of relay stations opened for 50 sensors, 70% coverage

50 sensors, 70% coverage			
	25 mile transmission range	**50 mile transmission range**	**75 mile transmission range**
No. of relay stations	3	2	1

Figure 8 suggests that as the transmission range increases, the number of relay stations opened decrease.

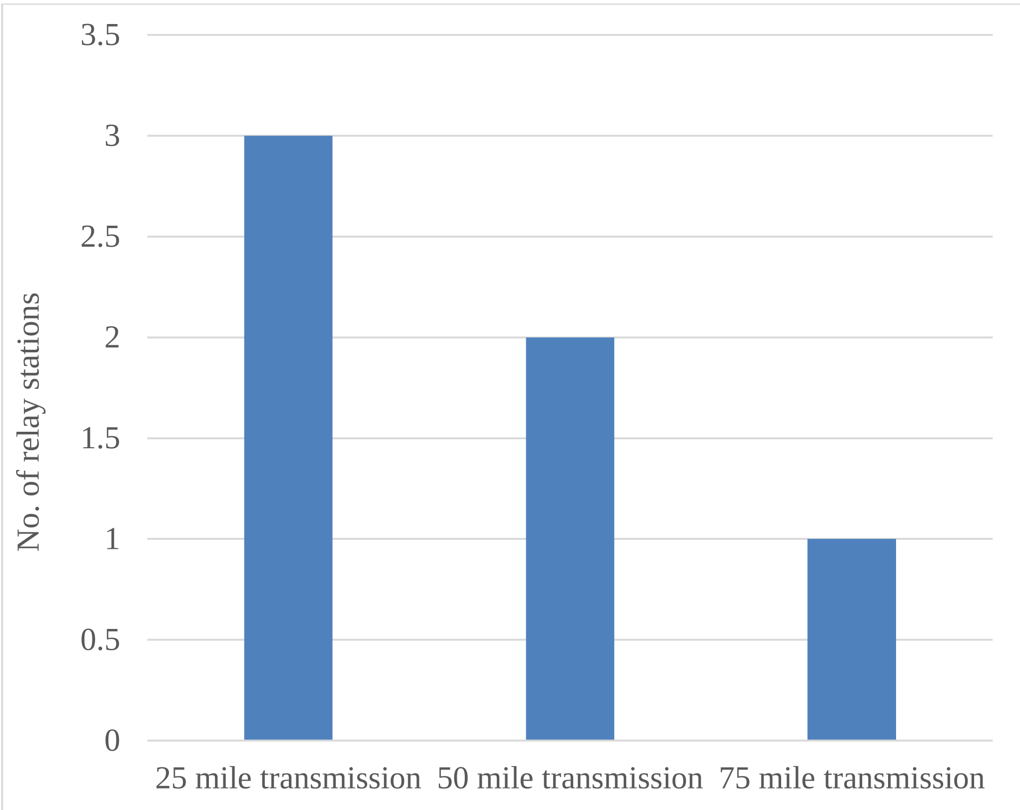


Figure 8. No. of relay stations opened for 50 sensors, 70% coverage

4.2. Sensitivity Analysis

Sensitivity Analysis deals with finding out the amount by which we can change the input data for the output of our mixed integer programming model to remain comparatively unchanged. This helps us in determining the sensitivity of the data we supply for the problem. If a small change in the input doesn't affect its optimal solution as much, we can conclude that the model is robust and is less sensitive to the changes in the input data.

Sensitivity analysis can be useful for a range of purposes, including:

- Testing the robustness of the results of a model in the presence of uncertainty.
- Increased understanding of the relationships between input and output variables in a model.
- Searching for errors in the model by encountering unexpected relationships between inputs and outputs.
- Model simplification by fixing model inputs that have no effect on the output, or identifying and removing redundant parts of the model structure.
- Enhancing communication from modelers to decision makers (e.g. by making recommendations more credible, understandable, compelling or persuasive).

Sensitivity analysis is carried out for the mixed integer programming model with the help of following test cases.

4.2.1. Test cases

A test case is a specific executable test that examines all the aspects including inputs and outputs of a system and then provides a detailed description of the steps that should be taken, the results that should be achieved, and other elements that should be identified.

The various parameters that are chosen for the test cases are number of sensor nodes, number of potential relay stations, transmission range of relay station, and percentage of coverage for the sensors. Initially to test the model with a small number of sensors, 50 sensor nodes and 20 potential relay stations are considered. Next, the number of sensors are increased by four times to 200. Later, with the sensitivity analysis, the number of sensor nodes are increased by a large numbers until 1500. The transmission ranges for the relay stations considered are 25, 50 and 75 miles and the percentage of coverage of sensor nodes is varied between 70, 90 and 100%.

4.2.1.1. Test case 1

The constant values of the parameters are number of sensor nodes = 50, number of potential relay stations = 20, transmission range of the relay station = 25 miles. The percentage of coverage of the sensors is varied between 70%, 90% and 100%.

The model is run 10 times with different data sets for 70% coverage, and similarly for the other coverage percentages.

The results in table 4 shows the mean, standard deviation and variance of the number of relay stations for the three different coverage percentages.

Table 4. Statistics for no. of relay stations for test case 1

50 sensors, 25 mile transmission range, 10 runs			
Number of relay stations			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	3	0	0.00000084
90% coverage	4.9	0.316227766	0.1
100% coverage	6.7	0.674948558	0.455555556

Figure 9 shows the graphical representation of mean and standard deviation of the no. of relay stations for test case 1.

The results in Table 5 shows the mean, standard deviation and variance of the execution time (sec) of the model for the three different coverage percentages.

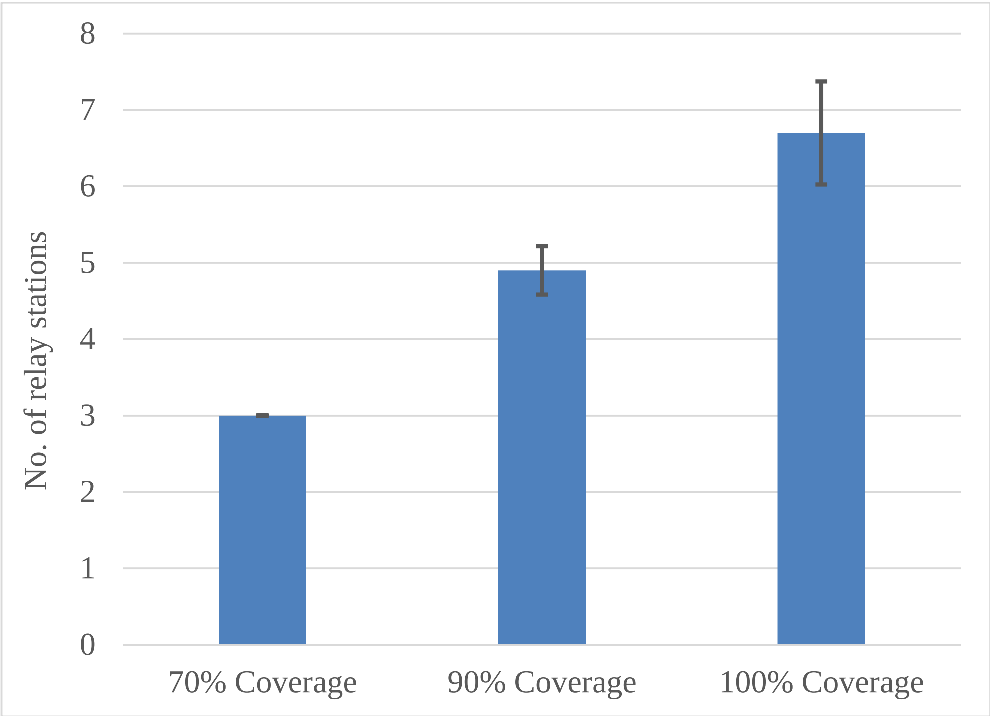


Figure 9. Mean and standard deviation of no. of relay stations for test case 1

Table 5. Statistics for execution time (sec) for test case 1

50 sensors, 25 mile transmission range, 10 runs			
Execution time (sec)			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	0.0044	0.000916515	0.00000084
90% coverage	0.0043	0.001187434	0.00000141
100% coverage	0.0062	0.002299758	5.28889E-06

Figure 10 shows the graphical representation of mean and standard deviation of the execution time (sec) for test case 1.

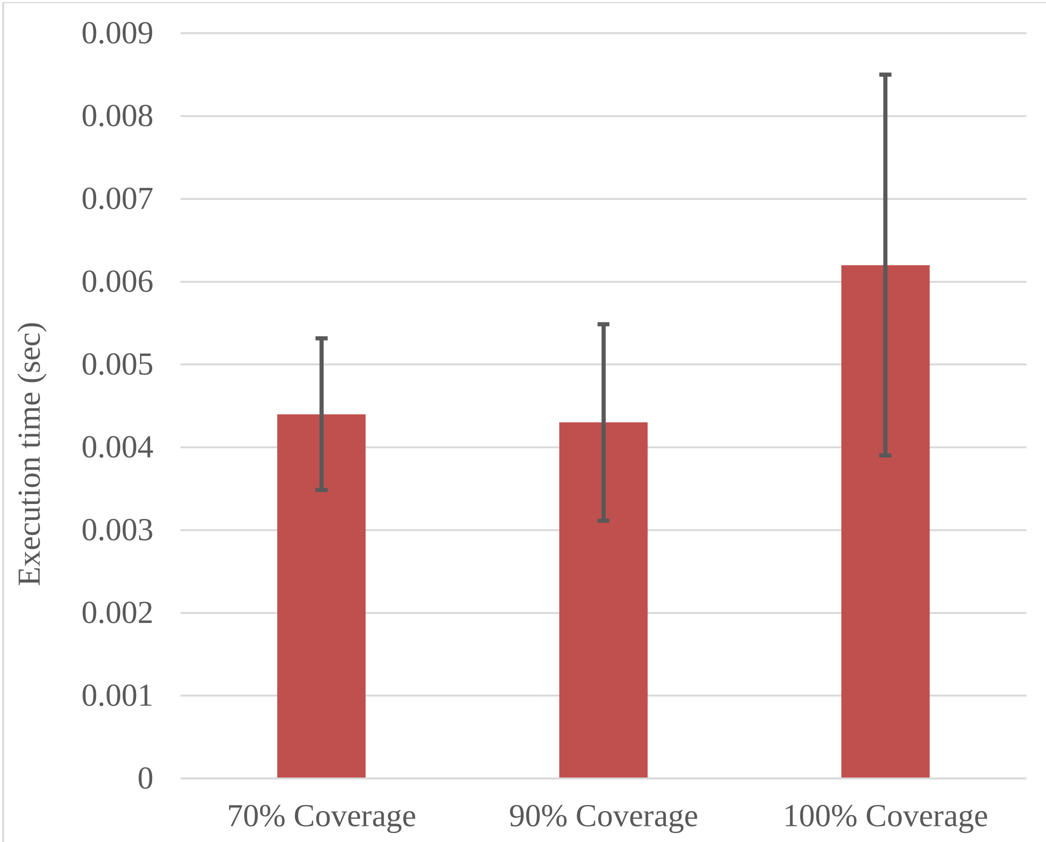


Figure 10. Mean and standard deviation of execution time (sec) for test case 1

4.2.1.2. Test case 2

The constant values of the parameters are number of sensor nodes = 50, number of potential relay stations = 20, transmission range of the relay station = 50 miles. The percentage of coverage of the sensors is varied between 70%, 90% and 100%.

The model is run 10 times with different data sets for 70% coverage, and similarly for the other coverage percentages.

The results in Table 6 shows the mean, standard deviation and variance of the number of relay stations for the three different coverage percentages.

Table 6. Statistics for no. of relay stations for test case 2

50 sensors, 50 mile transmission range, 10 runs			
Number of relay stations			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	1.9	0.316227766	0.1
90% coverage	2.4	0.489897949	0.266666667
100% coverage	3.4	0.516397779	0.266666667

Figure 11 shows the graphical representation of mean and standard deviation of the no. of relay stations for test case 2.

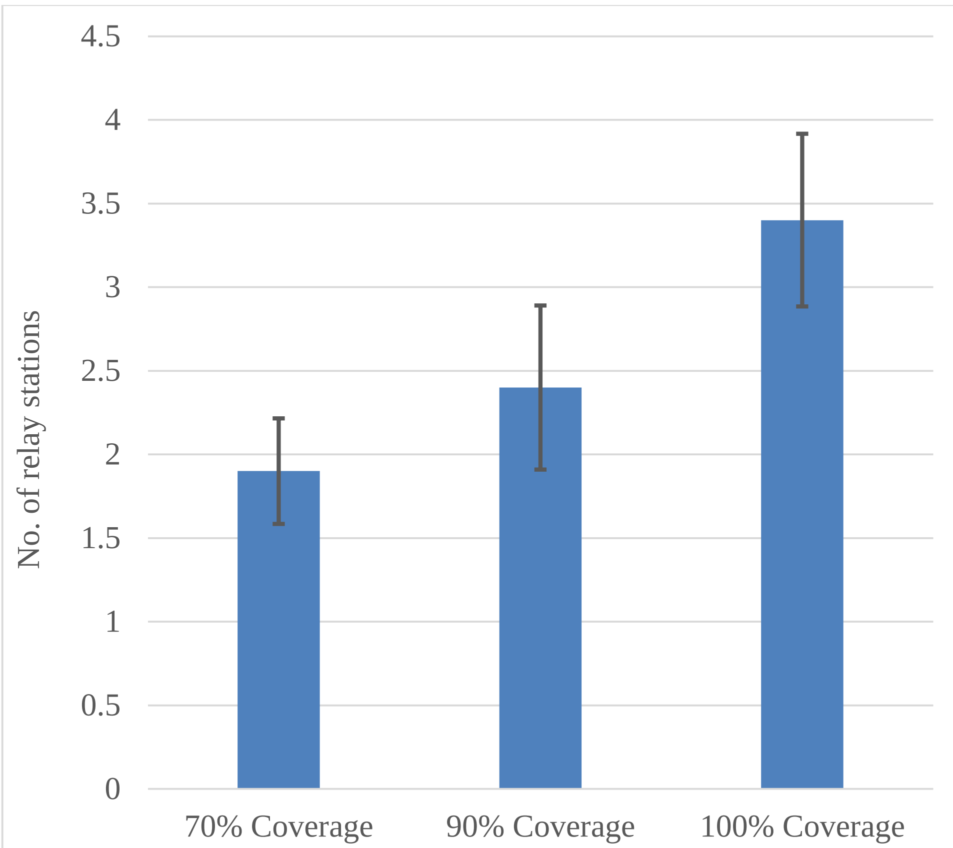


Figure 11. Mean and standard deviation of no. of relay stations for test case 2

The results in Table 7 shows the mean, standard deviation and variance of the execution time (sec) of the model for the three different coverage percentages.

Table 7. Statistics for execution time (sec) for test case 2

50 sensors, 50 mile transmission range, 10 runs			
Execution time (sec)			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	0.0057	0.002359378	0.002359378
90% coverage	0.0058	0.002441311	6.62222E-06
100% coverage	0.0087	0.003377869	1.26778E-05

Figure 12 shows the graphical representation of mean and standard deviation of the execution time (sec) for test case 2.

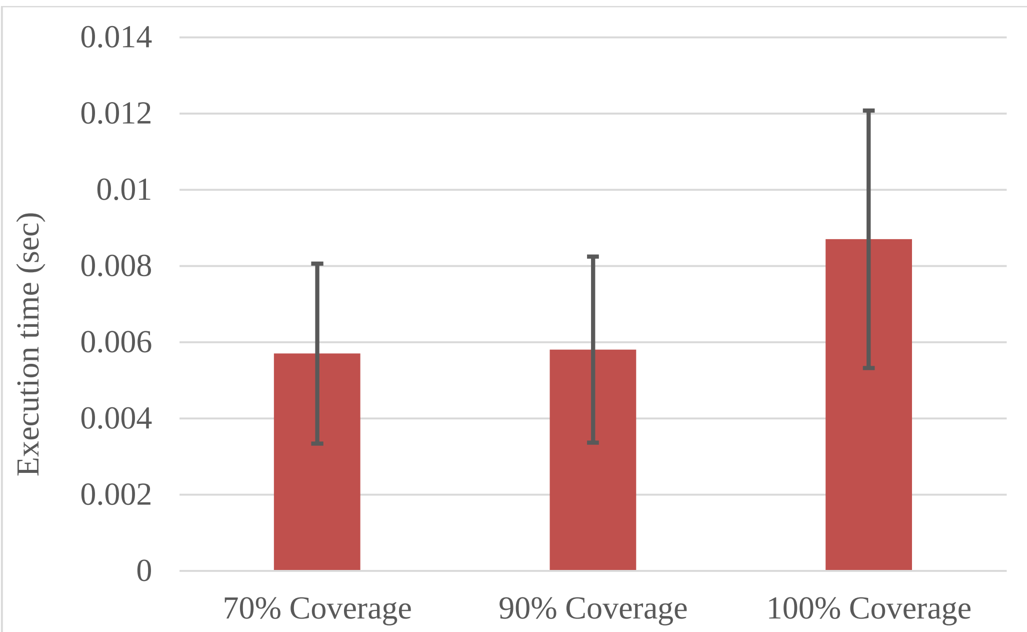


Figure 12. Mean and standard deviation of execution time (sec) for test case 2

4.2.1.3. Test case 3

The constant values of the parameters are number of sensor nodes = 50, number of potential relay stations = 20, transmission range of the relay station = 75 miles. The percentage of coverage of the sensors is varied between 70%, 90% and 100%.

The model is run 10 times with different data sets for 70% coverage, and similarly for the other coverage percentages.

The results in Table 8 show the mean, standard deviation and variance of the number of relay stations for the three different coverage percentages.

Table 8. Statistics for no. of relay stations for test case 3

50 sensors, 75 mile transmission range, 10 runs			
Number of relay stations			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	1	0	0
90% coverage	2	0	0
100% coverage	2	0	0

Figure 13 shows the graphical representation of mean and standard deviation of the no. of relay stations for test case 3.

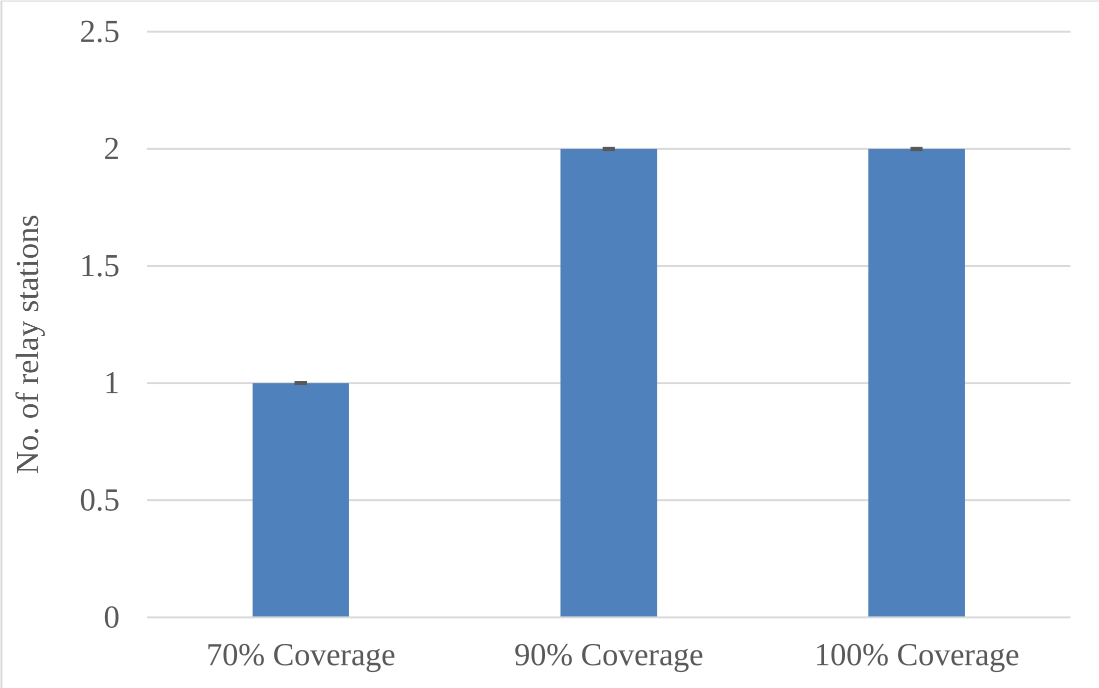


Figure 13. Mean and standard deviation of no. of relay stations for test case 3

The results in Table 9 shows the mean, standard deviation and variance of the execution time (sec) of the model for the three different coverage percentages.

Table 9. Statistics for execution time (sec) for test case 3

50 sensor, 75 mile transmission range, 10 runs			
Execution time (sec)			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	0.00971	0.003654054	1.33521E-05
90% coverage	0.008	0.000942809	8.88889E-07
100% coverage	0.0051	0.002330951	5.43333E-06

Figure 14 shows the graphical representation of mean and standard deviation of the execution time (sec) for test case 3.

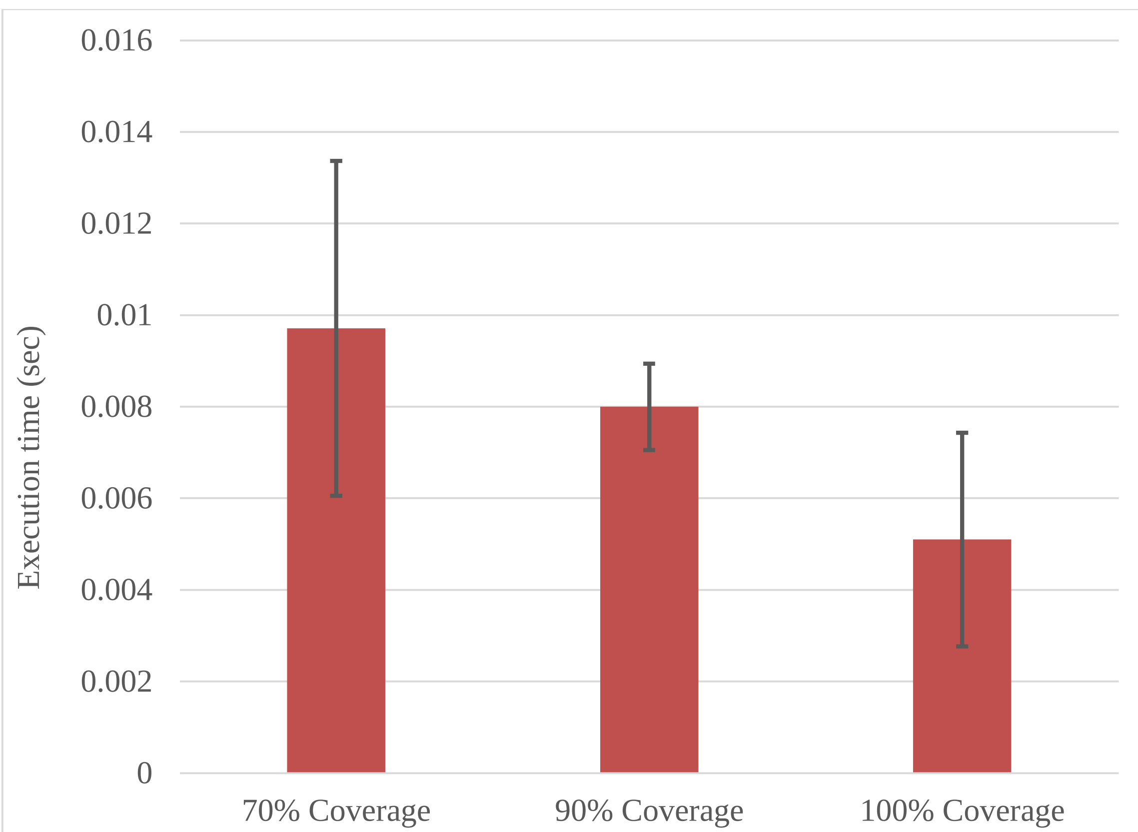


Figure 14. Mean and standard deviation of execution time (sec) for test case 3

4.2.1.4. Discussion of sections 4.2.1.1, 4.2.1.2 and 4.2.1.3

It can be inferred from the figures that, as the coverage percentages increases, the mean number of relay stations required increase; for all the transmission ranges. The number of relay stations opened are reduced by a maximum of 85% for test case 1, 90.5% for test case 2 and 95% for test case 3.

The mean of execution time (sec) increases with the coverage percentages, but for the 75 mile transmission range, the mean execution time decreases.

4.2.1.5. Test case 4

The constant values of the parameters are number of sensor nodes = 200, number of potential relay stations = 20, transmission range of the relay station = 25 miles. The percentage of coverage of the sensors is varied between 70%, 90% and 100%.

The model is run 10 times with different data sets for 70% coverage, and similarly for the other coverage percentages.

The results in Table 10 shows the mean, standard deviation and variance of the number of relay stations for the three different coverage percentages.

Table 10. Statistics for no. of relay stations for test case 4

200 sensors, 25 mile transmission range, 10 runs			
Number of relay stations			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	4	0	0
90% coverage	6.2	0.4	0.177777778
100% coverage	11	1.264911064	1.777777778

Figure 15 shows the graphical representation of mean and standard deviation of the no. of relay stations for test case 4.

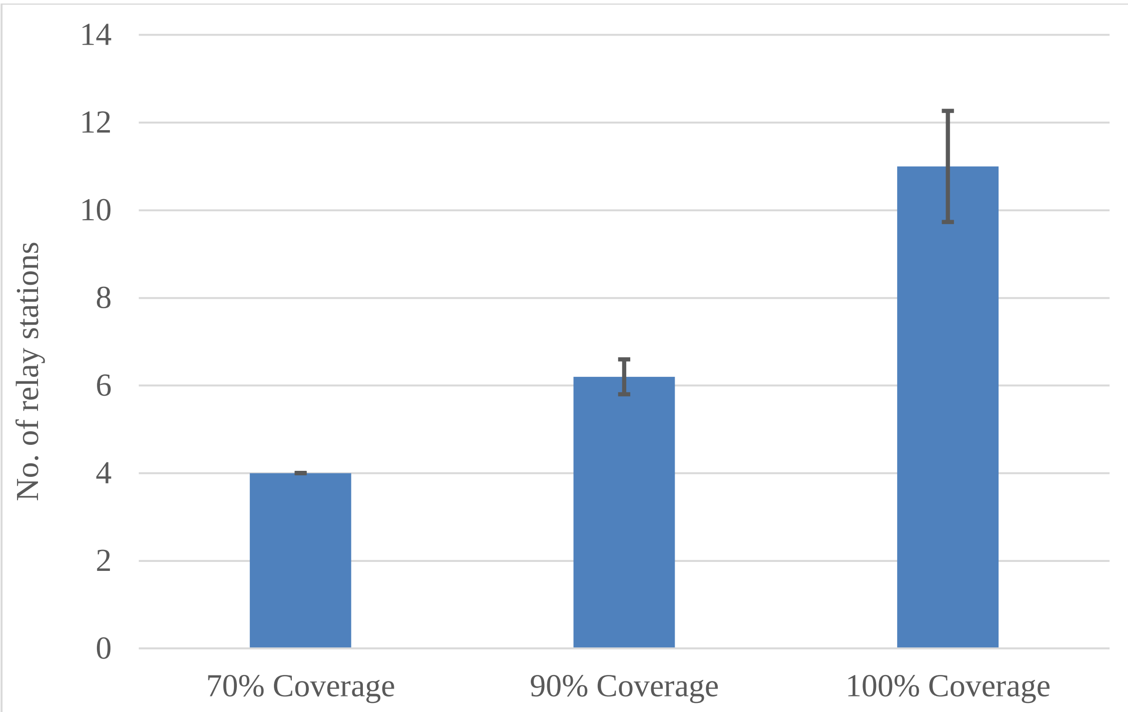


Figure 15. Mean and standard deviation of no. of relay stations for test case 4

The results in Table 11 shows the mean, standard deviation and variance of the execution time (sec) of the model for the three different coverage percentages.

Table 11. Statistics for execution time (sec) for test case 4

200 sensors, 25 mile transmission range, 10 runs			
Execution time (sec)			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	0.0279	0.015687256	0.000273433
90% coverage	0.0426	0.019314243	0.000414489
100% coverage	0.0318	0.015961203	0.000283067

Figure 16 shows the graphical representation of mean and standard deviation of the execution time (sec) for test case 4.

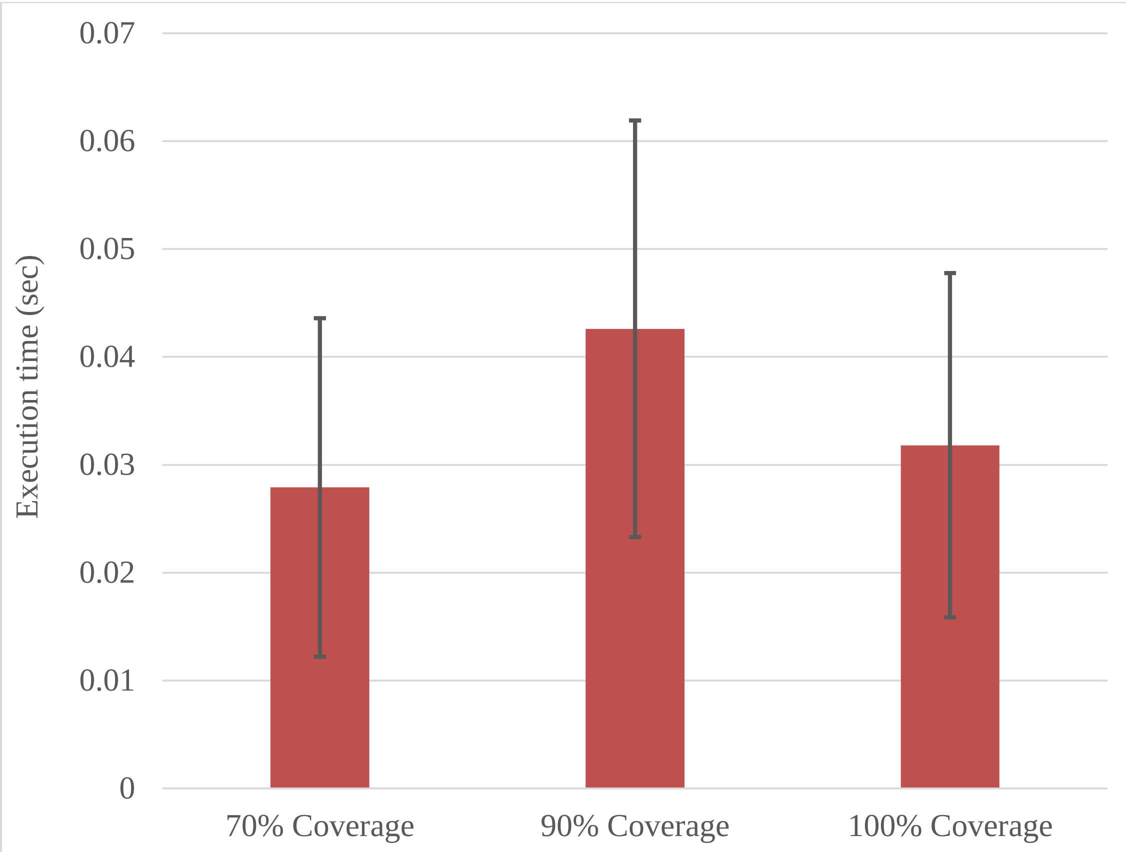


Figure 16. Mean and standard deviation of execution time (sec) for test case 4

4.2.1.6. Test case 5

The constant values of the parameters are number of sensor nodes = 200, number of potential relay stations = 20, transmission range of the relay station = 50 miles. The percentage of coverage of the sensors is varied between 70%, 90% and 100%.

The model is run 10 times with different data sets for 70% coverage, and similarly for the other coverage percentages.

The results in Table 12 shows the mean, standard deviation and variance of the number of relay stations for the three different coverage percentages.

Table 12. Statistics for no. of relay stations for test case 5

200 sensors, 50 mile transmission range, 10 runs			
Number of relay stations			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	2	0	0
90% coverage	3	0	0
100% coverage	5	0	0

Figure 17 shows the graphical representation of mean and standard deviation of the no. of relay stations for test case 5.

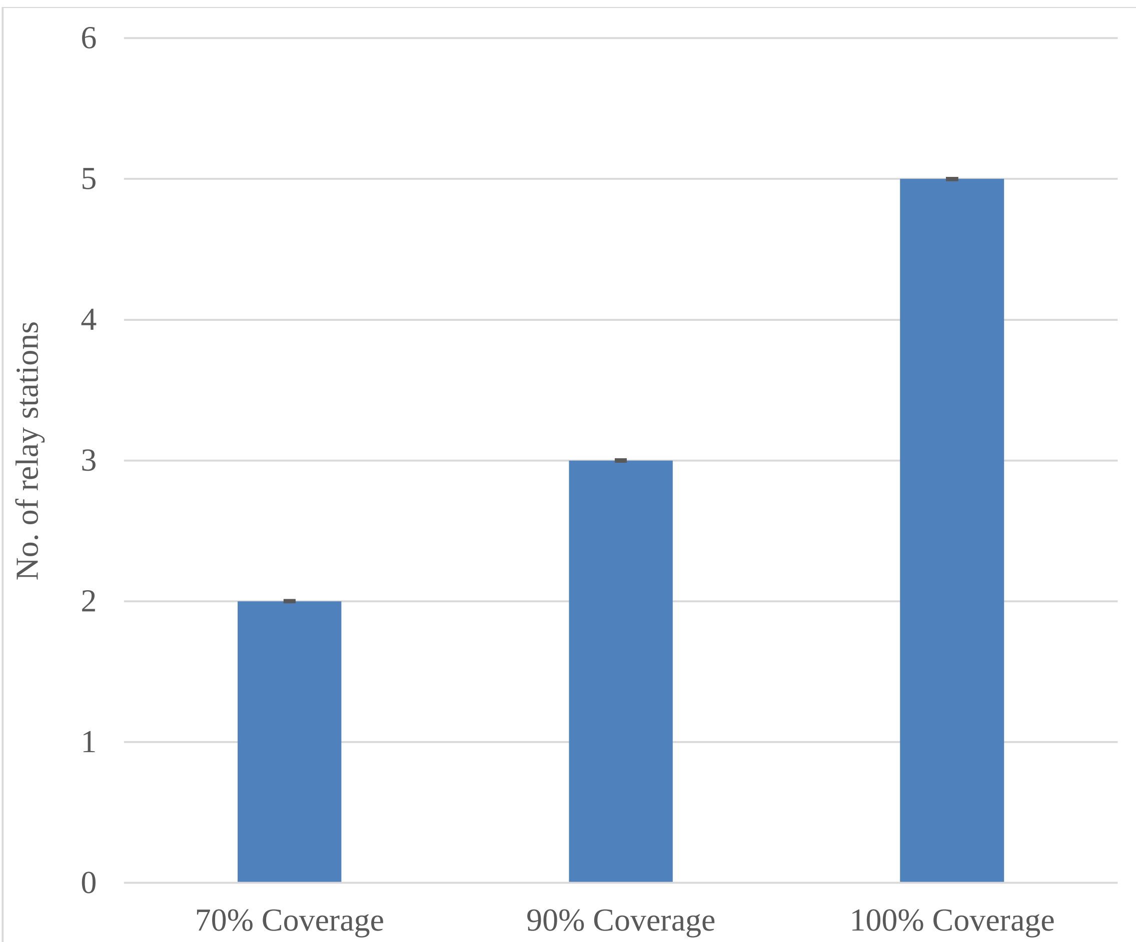


Figure 17. Mean and standard deviation of no. of relay stations for test case 5

The results in Table 13 shows the mean, standard deviation and variance of the execution time (sec) of the model for the three different coverage percentages.

Table 13. Statistics for execution time (sec) for test case 5

200 sensors, 50 mile transmission range, 10 runs			
Execution time (sec)			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	0.0278	0.014132704	0.000199733
90% coverage	0.0364	0.014361987	0.00026267
100% coverage	0.0294	0.013696715	0.0001876

Figure 18 shows the graphical representation of mean and standard deviation of the execution time (sec) for test case 5.

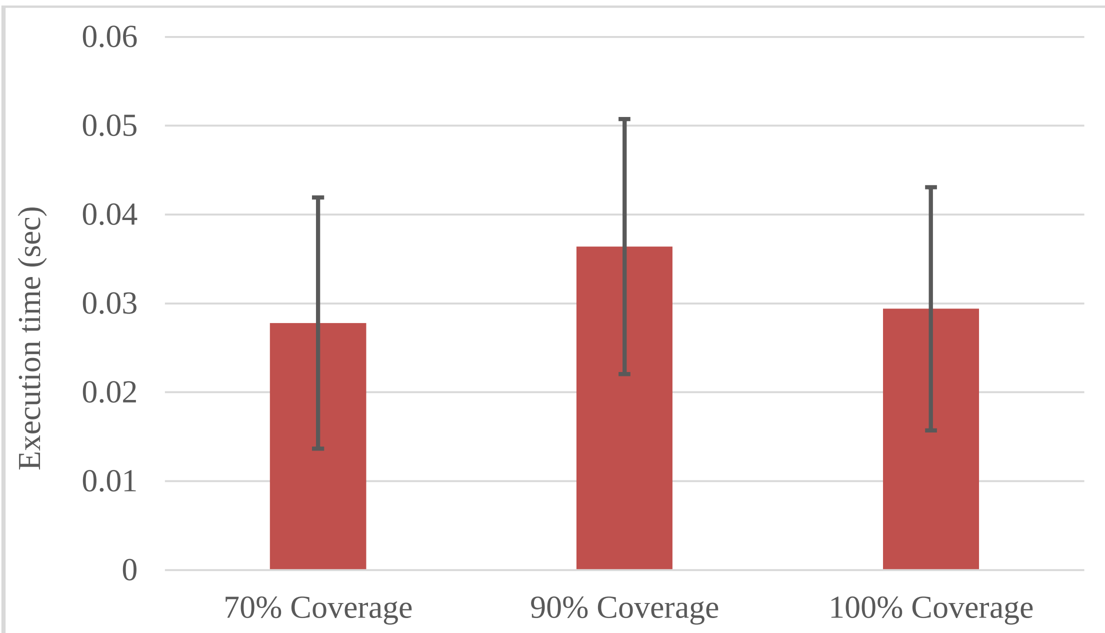


Figure 18. Mean and standard deviation of execution time (sec) for test case 5

4.2.1.7. Test case 6

The constant values of the parameters are number of sensor nodes = 200, number of potential relay stations = 20, transmission range of the relay station = 75 miles. The percentage of coverage of the sensors is varied between 70%, 90% and 100%.

The model is run 10 times with different data sets for 70% coverage, and similarly for the other coverage percentages.

The results in Table 14 shows the mean, standard deviation and variance of the number of relay stations for the three different coverage percentages.

Table 14. Statistics for no. of relay stations for test case 6

200 sensors, 75 mile transmission range, 10 runs			
Number of relay stations			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	1	0	0
90% coverage	2	0	0
100% coverage	3	0	0

Figure 19 shows the graphical representation of mean and standard deviation of the no. of relay stations for test case 6.

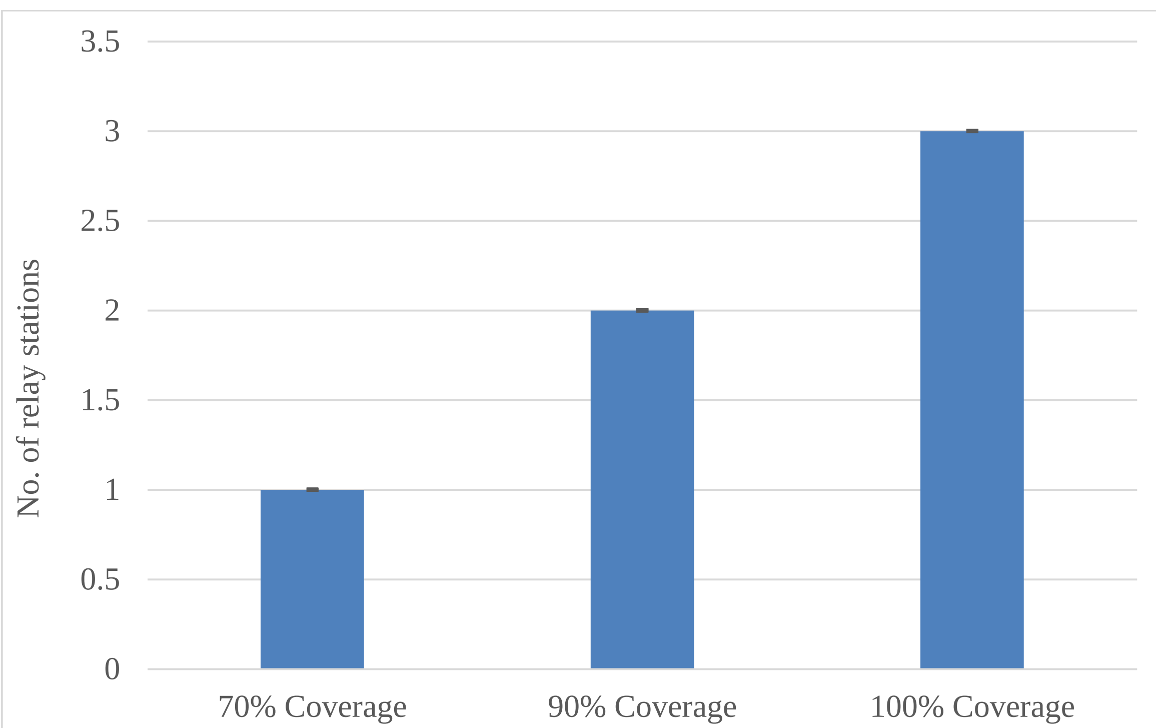


Figure 19. Mean and standard deviation of no. of relay stations for test case 6

The results in Table 15 shows the mean, standard deviation and variance of the execution time (sec) of the model for the three different coverage percentages.

Table 15. Statistics for execution time (sec) for test case 6

200 sensors, 75 mile transmission range, 10 runs			
Execution time (sec)			
Case	**Mean**	**Standard Deviation**	**Variance**
70% coverage	0.0292	0.012612163	0.000159067
90% coverage	0.0269	0.010082438	0.000101656
100% coverage	0.0295	0.019822826	0.000392944

Figure 20 shows the graphical representation of mean and standard deviation of the execution time (sec) for test case 6.

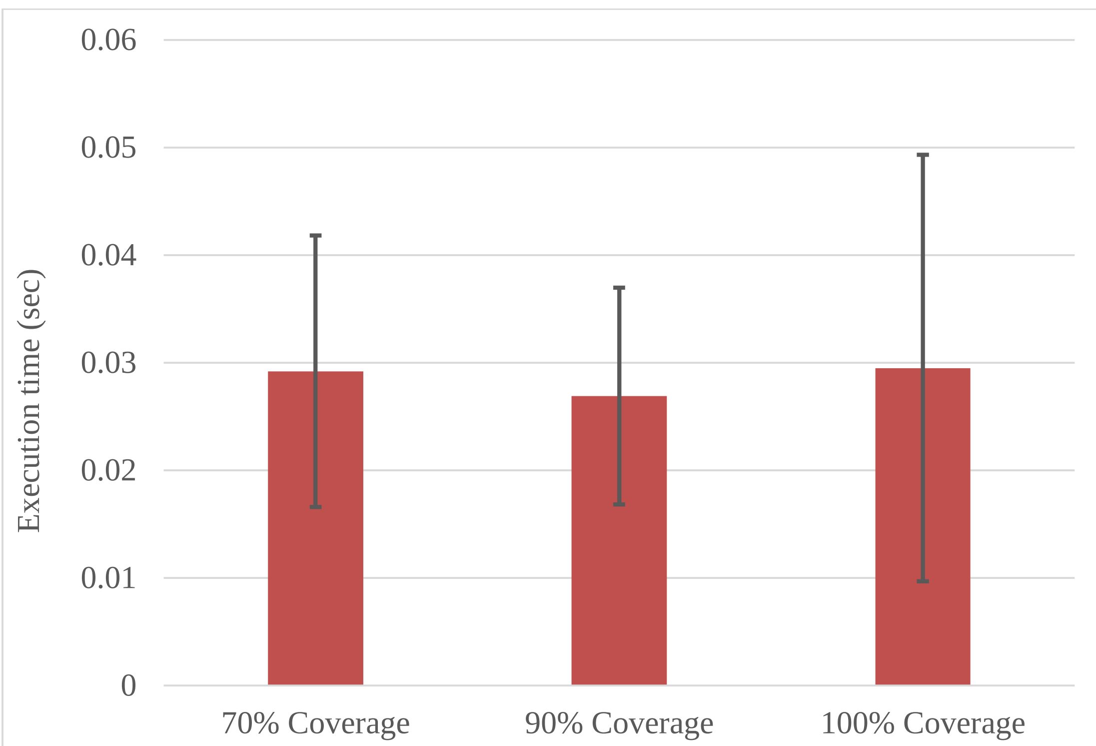


Figure 20. Mean and standard deviation of execution time (sec) for test case 6

4.2.1.8. Discussion of sections 4.2.1.5, 4.2.1.6 and 4.2.1.7

The figures indicate that as the coverage percentages increase, the mean number of relay stations required also increase; across all the transmission ranges. The number of relay stations opened are reduced by a maximum of 80% for test case 4, 90% for test case 5 and 95% for test case 6.

For 25 and 50 mile transmission ranges, the mean of the execution time (sec) starts at one point for the 70% coverage, increases for the 90% coverage and falls slightly for the 100% coverage. For the 75 mile transmission range, the mean starts at a point for 70% coverage, dips slightly for 90% coverage and rises for the 100% coverage.

4.2.1.9. Test case 7

4.2.1.9.1. 50 mile transmission range

The constant values of the parameters are number of potential relay stations = 20, transmission range of the relay station = 50 miles, percentage of coverage of the sensors = 100%. The number of sensors considered is 500, 700, 1000, 1250 and 1500.

The results in Table 16 shows the number of relay stations opened for test case 7, 50 mile transmission range.

Table 16. No. of relay stations for test case 7, 50 mile transmission range

50 mile transmission range, 100% coverage					
No. of sensors	**500**	**700**	**1000**	**1250**	**1500**
No. of relay stations	6	7	7	7	8

The results in Table 17 shows the execution time (sec) for test case 7, 50 mile transmission range.

Table 17. Execution time (sec) for test case 7, 50 mile transmission range

50 mile transmission range, 100% coverage					
No. of sensors	**500**	**700**	**1000**	**1250**	**1500**
Execution time (sec)	0.02	0.029	0.059	0.082	0.098

4.2.1.9.2. 75 mile transmission range

The constant values of the parameters are number of potential relay stations = 20, transmission range of the relay station = 75 miles, percentage of coverage of the sensors = 100%. The number of sensors considered are 500, 700 and 1000.

The results in Table 18 shows the number of relay stations opened for test case 7, 75 mile transmission range.

Table 18. No. of relay stations for test case 7, 75 mile transmission range

75 mile transmission range, 100% coverage			
No. of sensors	**500**	**700**	**1000**
No. of relay stations	4	4	4

The results in Table 19 shows the execution time (sec) for test case 7, 75 mile transmission range.

Table 19. Execution time (sec) for test case 7, 75 mile transmission range

75 mile transmission range, 100% coverage			
No. of sensors	**500**	**700**	**1000**
Execution time (sec)	0.019	0.027	0.085

4.2.2. Comparison of test cases

Quantitative test case comparison metrics suggest the amount of similarity between any test case pair, capturing what is being tested on the target binaries. We capture key aspects of MIP model including the number of relay stations opened and execution time. These are the aspects or features which drive our comparison.

4.2.2.1. Comparison 1

Comparison is made between 25, 50 and 75 mile transmission ranges for 50 sensors and 100% coverage.

Results in Table 20 shows the number of relay stations opened for comparison 1.

Table 20. No. of relay stations for comparison 1

50 sensors, 100% coverage			
	25 mile transmission range	**50 mile transmission range**	**75 mile transmission range**
No. of relay stations	6.7	3.4	2

Figure 21 shows the graphical representation of number of relay stations opened for comparison 1.

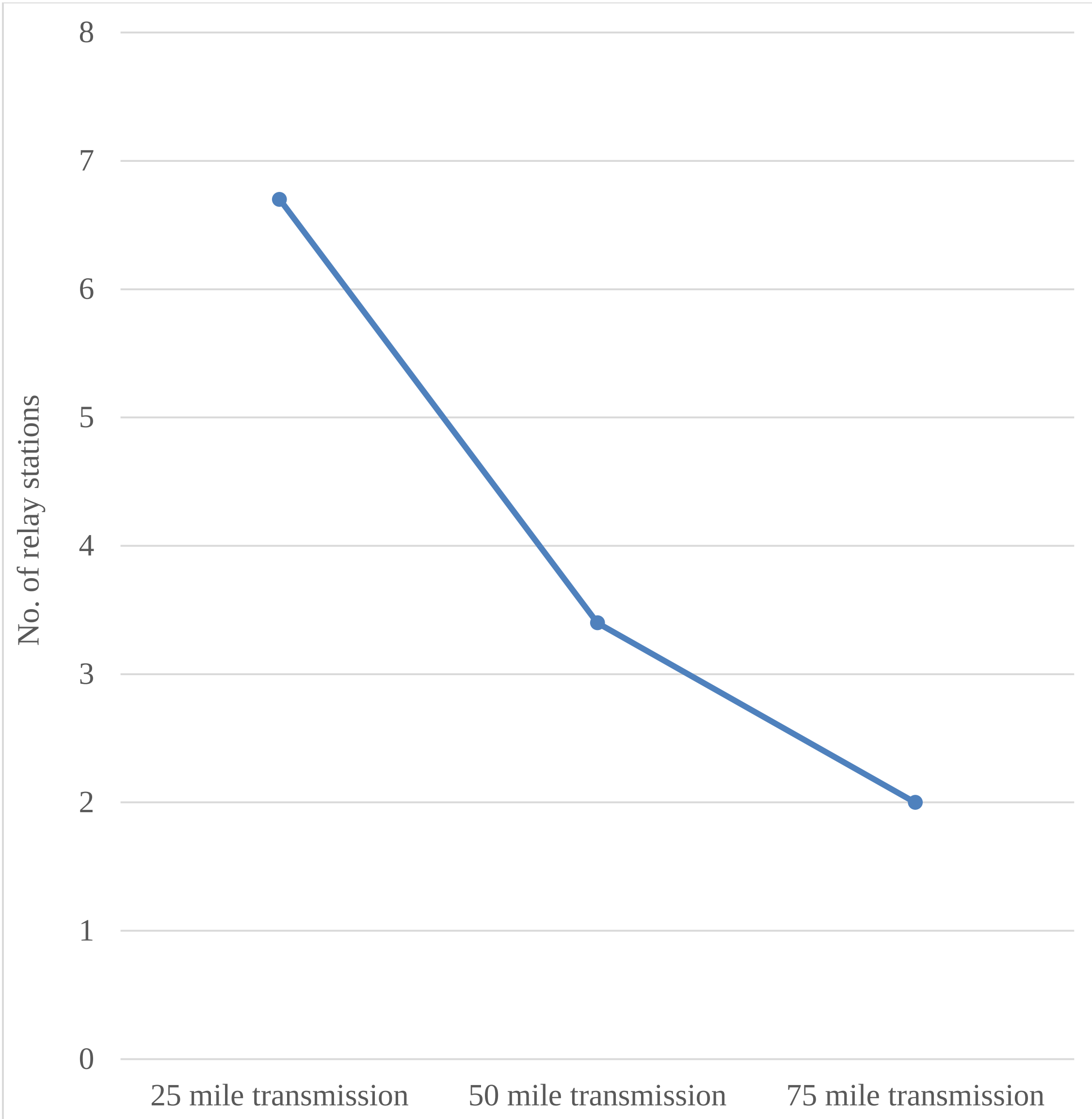


Figure 21. No. of relay stations for comparison 1

Results in Table 21 shows the execution time (sec) for comparison 1.

Table 21. Execution time (sec) for comparison 1

	50 sensors, 100% coverage		
	25 mile transmission range	**50 mile transmission range**	**75 mile transmission range**
Execution time (sec)	0.0062	0.0087	0.0051

Figure 22 shows the graphical representation of execution time (sec) for comparison 1.

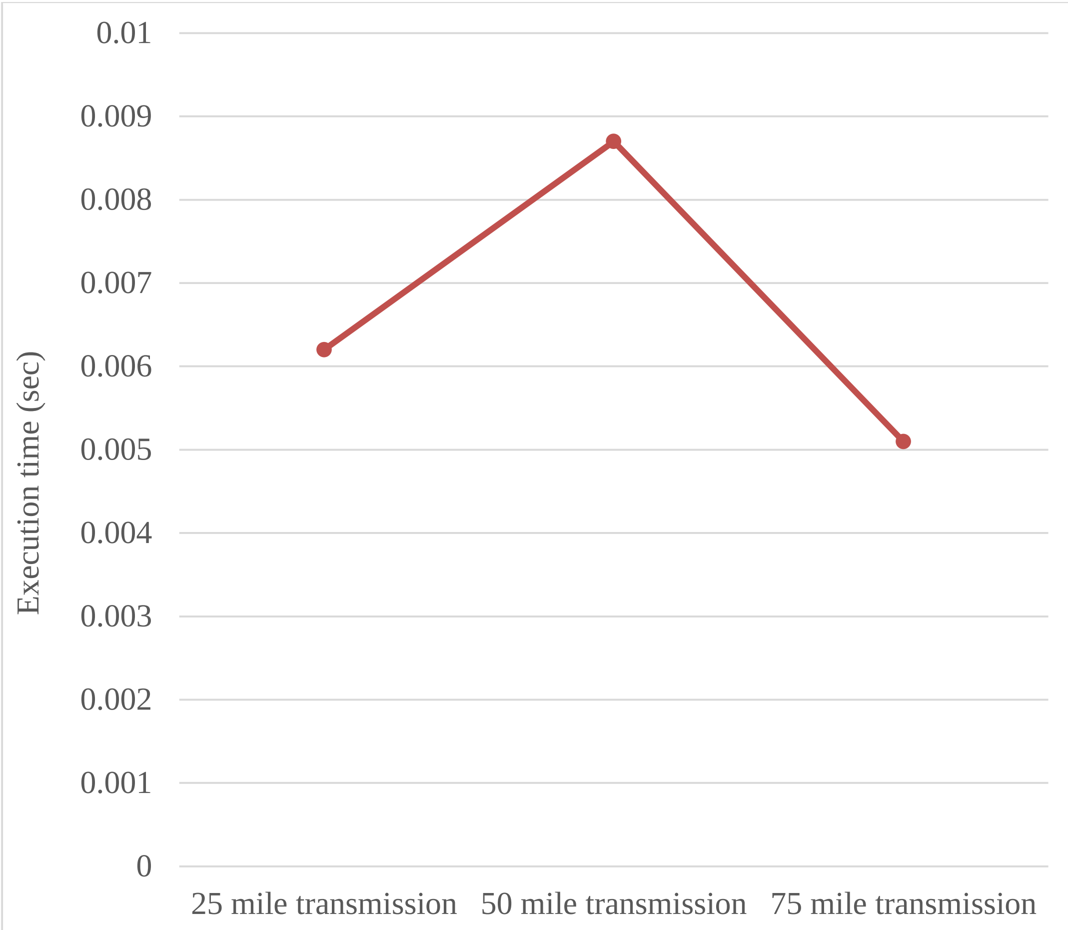


Figure 22. Execution time (sec) for comparison 1

4.2.2.2. Comparison 2

Comparison is made between 50, 200, 500, 700, 1000, 1250 and 1500 sensors for 50 mile transmission range and 100% coverage.

Results in Table 22 shows the number of relay stations opened for comparison 2.

Table 22. No. of relay stations for comparison 2

50 mile transmission range, 100% coverage							
No. of sensors	**50**	**200**	**500**	**700**	**1000**	**1250**	**1500**
No. of relay stations	3.4	5	6	7	7	7	8

Figure 23 shows the graphical representation of number of relay stations opened for comparison 2.

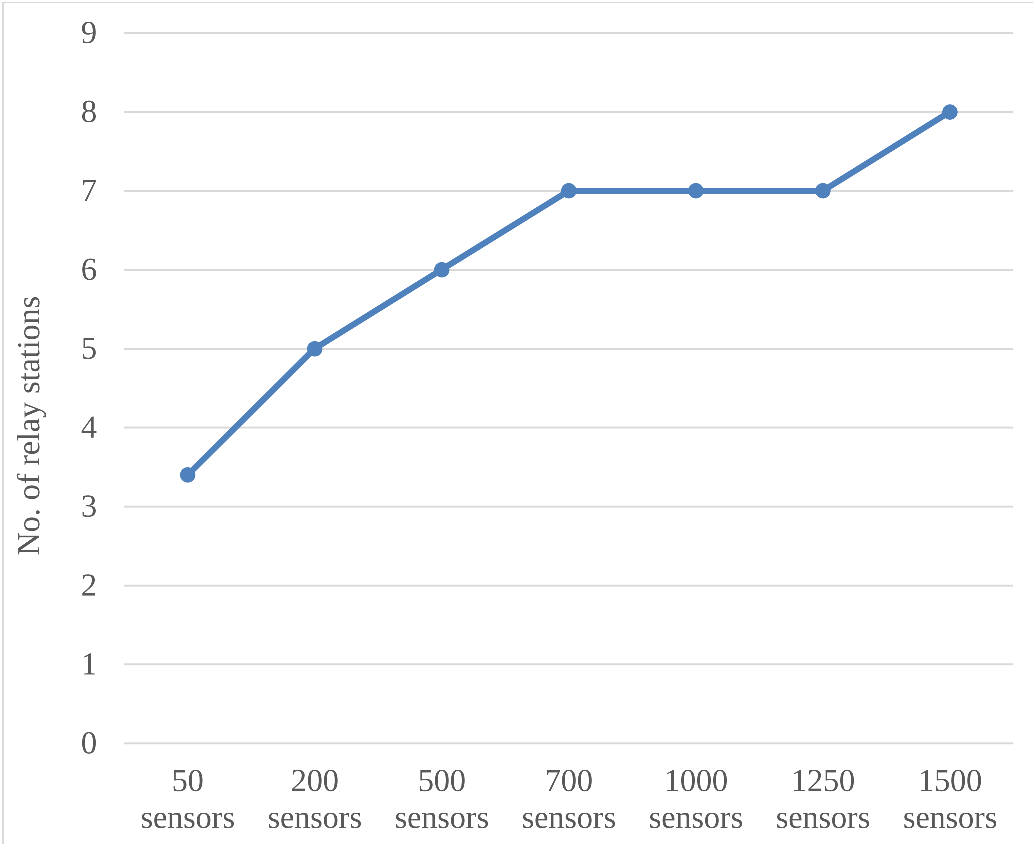


Figure 23. No. of relay stations for comparison 2

Results in Table 23 shows the execution time (sec) for comparison 2.

Table 23. Execution time (sec) for comparison 2

50 mile transmission range, 100% coverage							
No. of sensors	**50**	**200**	**500**	**700**	**1000**	**1250**	**1500**
Execution time (sec)	0.0087	0.0294	0.02	0.029	0.059	0.082	0.098

Figure 24 shows the graphical representation of execution time (sec) for comparison 2.

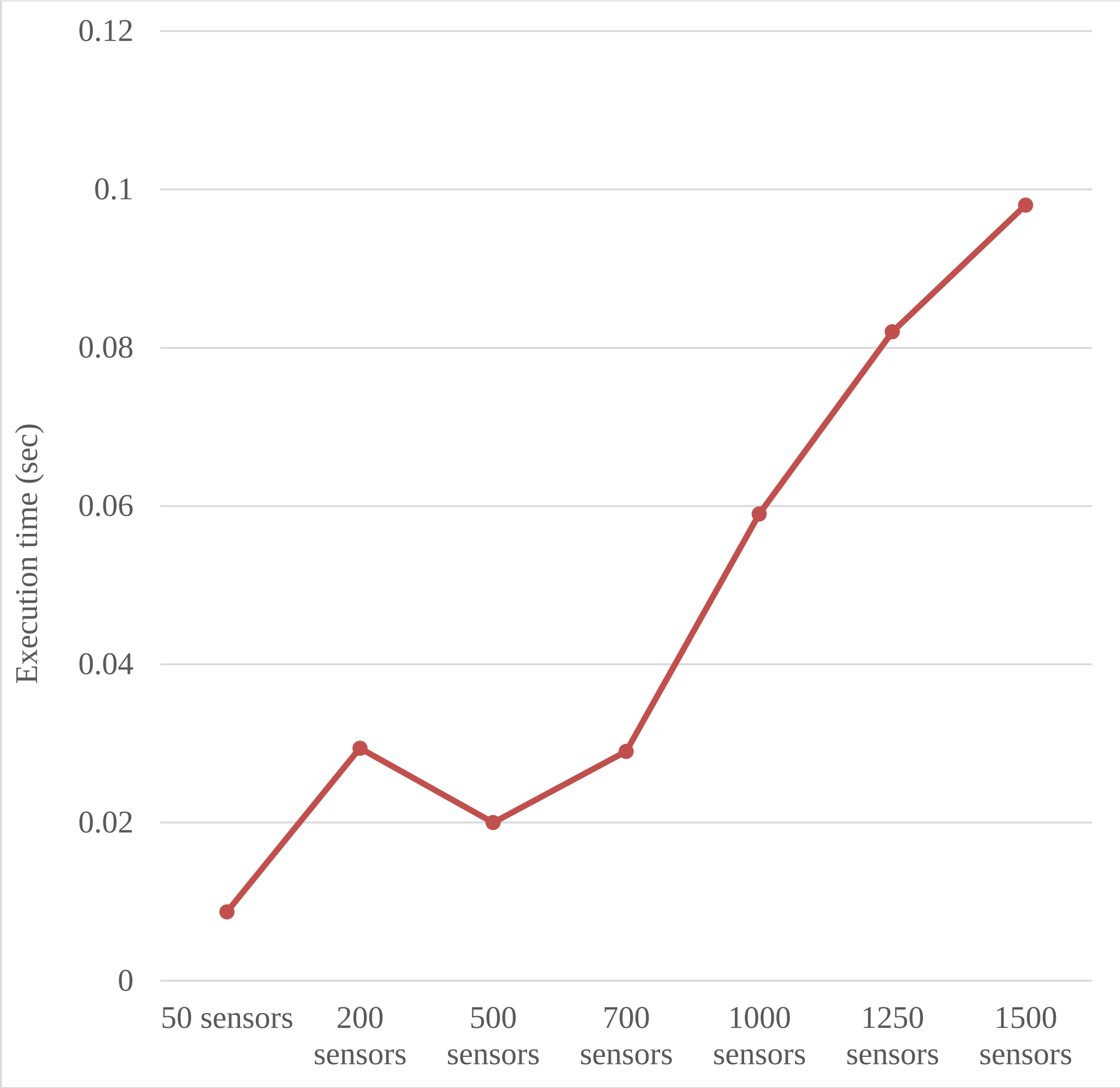


Figure 24. Execution time (sec) for comparison 2

4.2.2.3. Discussion of sections 4.2.2.1 and 4.2.2.2

From 4.2.2.1, as the transmission range increases, the number of relay stations required decrease but the execution time increase for the first two transmission ranges but decrease for the last one.

From 4.2.2.2, we infer that as the number of sensors increase, both the number of relay stations required and the execution time also increase.

5. CASE STUDY: MAINE STATE

The case study on Maine State was designed and conducted to optimize the number of relay stations required and place them optimally across the state for a given number of sensors, transmission range of relay station and desired percentage of coverage. There are a total of eleven major cities in Maine State, and we consider these cities as potential locations for the relay stations. So, 100 sensors and 11 potential relay stations are considered for this study; several runs are carried out by varying the transmission range of the relay station and the coverage percentage.

Figure 25 shows the potential relay station locations for Maine State.

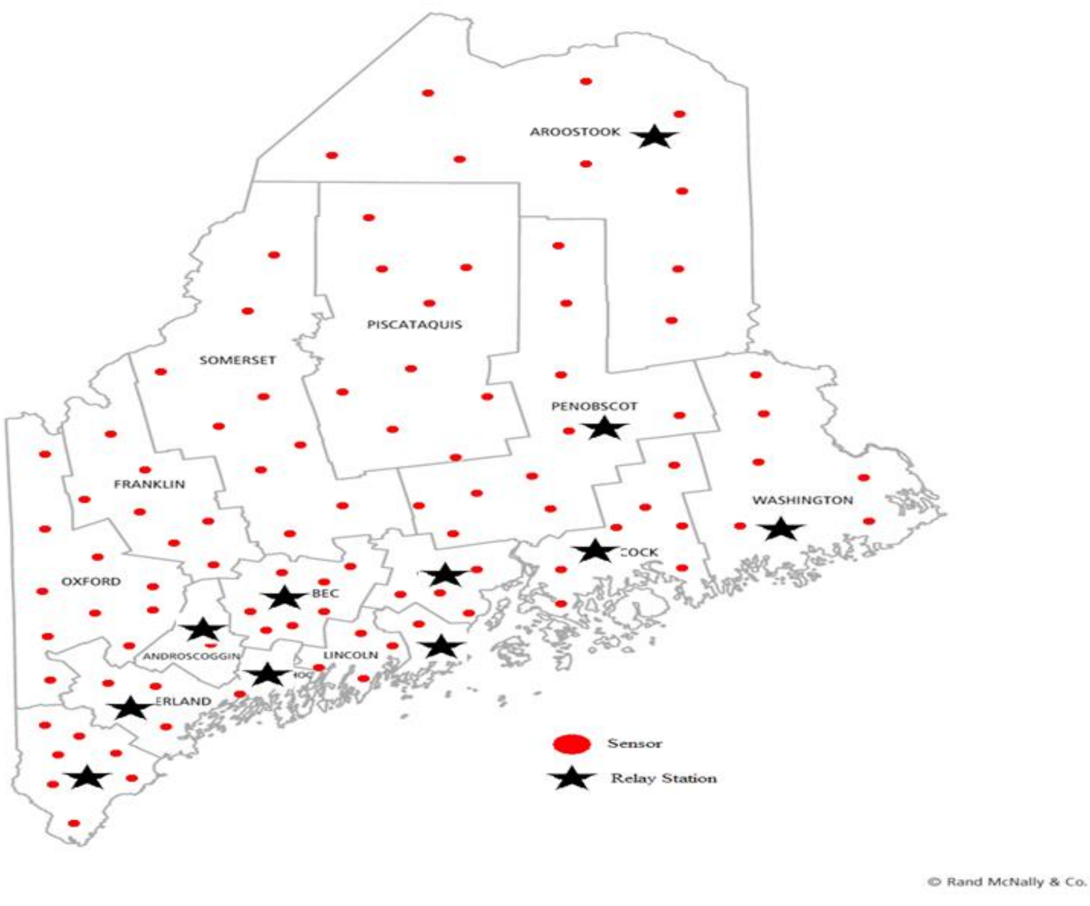


Figure 25. Potential relay station locations for Maine State

5.1. Run 1

Run 1 is carried out with the following parameters; transmission range = 90 miles, coverage percentage = 100%.

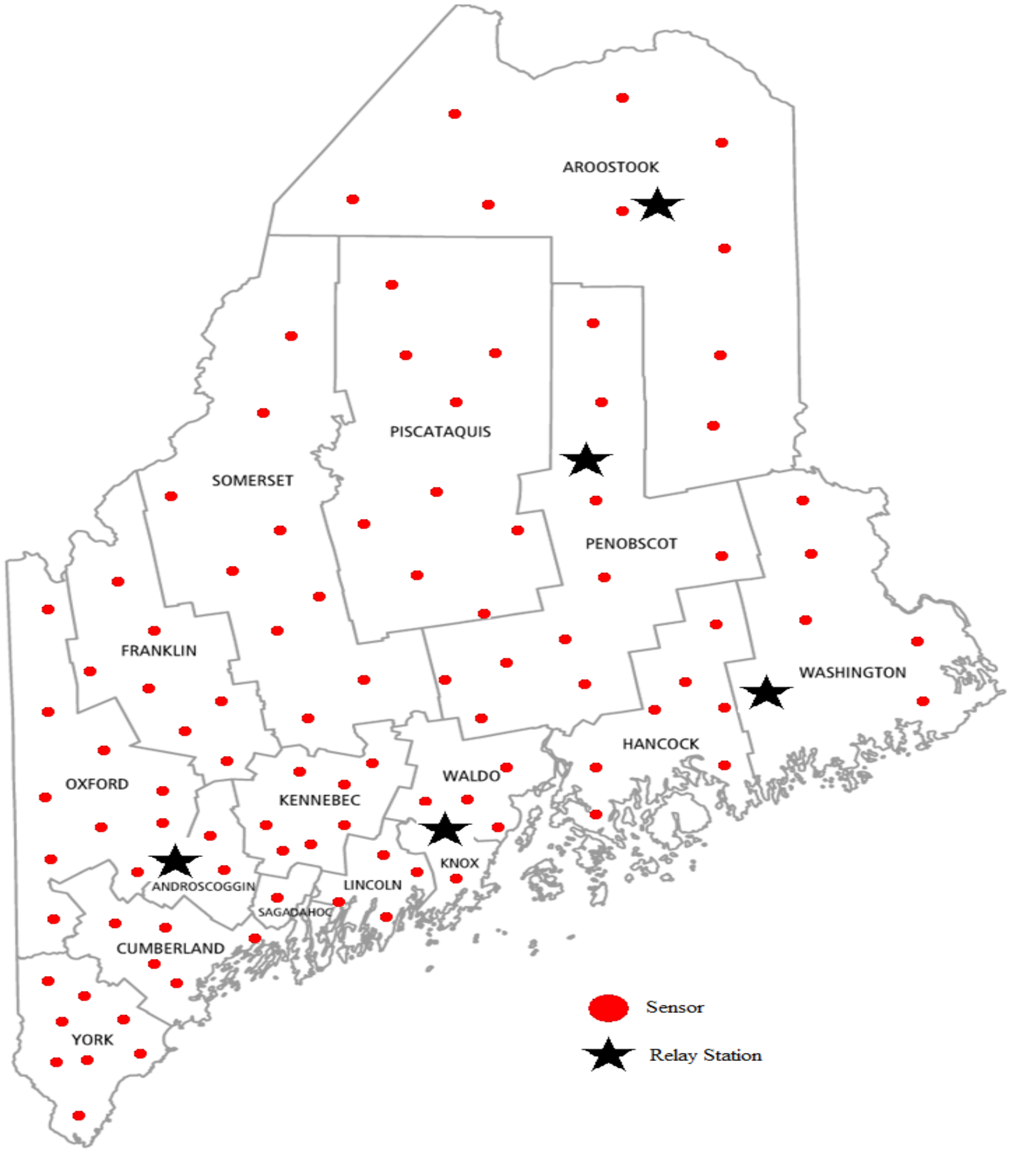


Figure 26. Optimal relay station locations for Maine State, run 1

Figure 26 shows that the model opened a total of 5 optimal relay station locations to meet the required specifications.

5.2. Run 2

Run 2 is carried out with the following parameters, transmission range = 100 miles, coverage percentage = 100%.

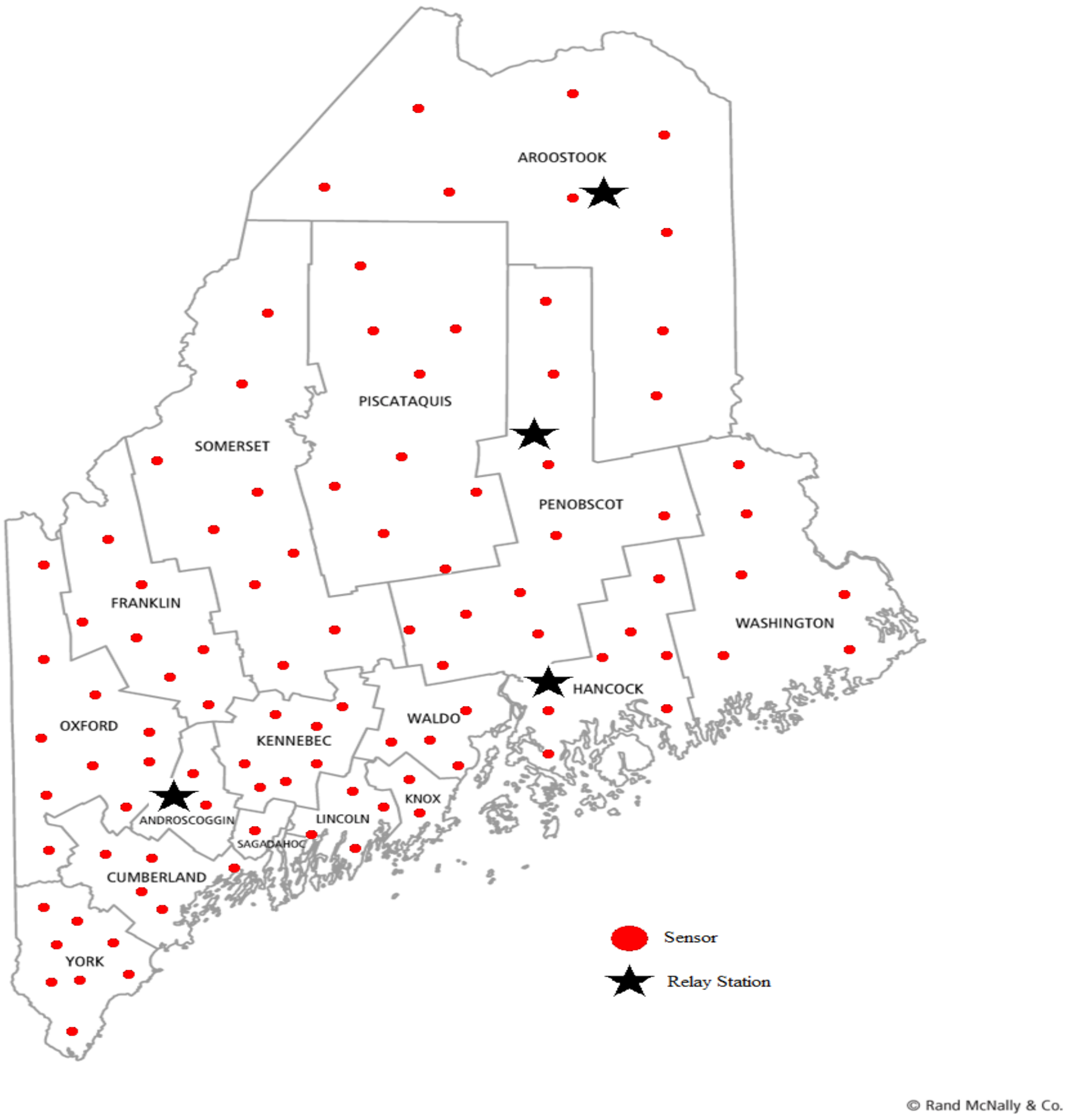


Figure 27. Optimal relay station locations for Maine State, run 2

Figure 27 shows that the model opened a total of 4 optimal relay station locations to meet the required specifications.

5.3. Run 3

Run 3 is carried out with the following parameters; transmission range = 90 miles, coverage percentage = 90%.

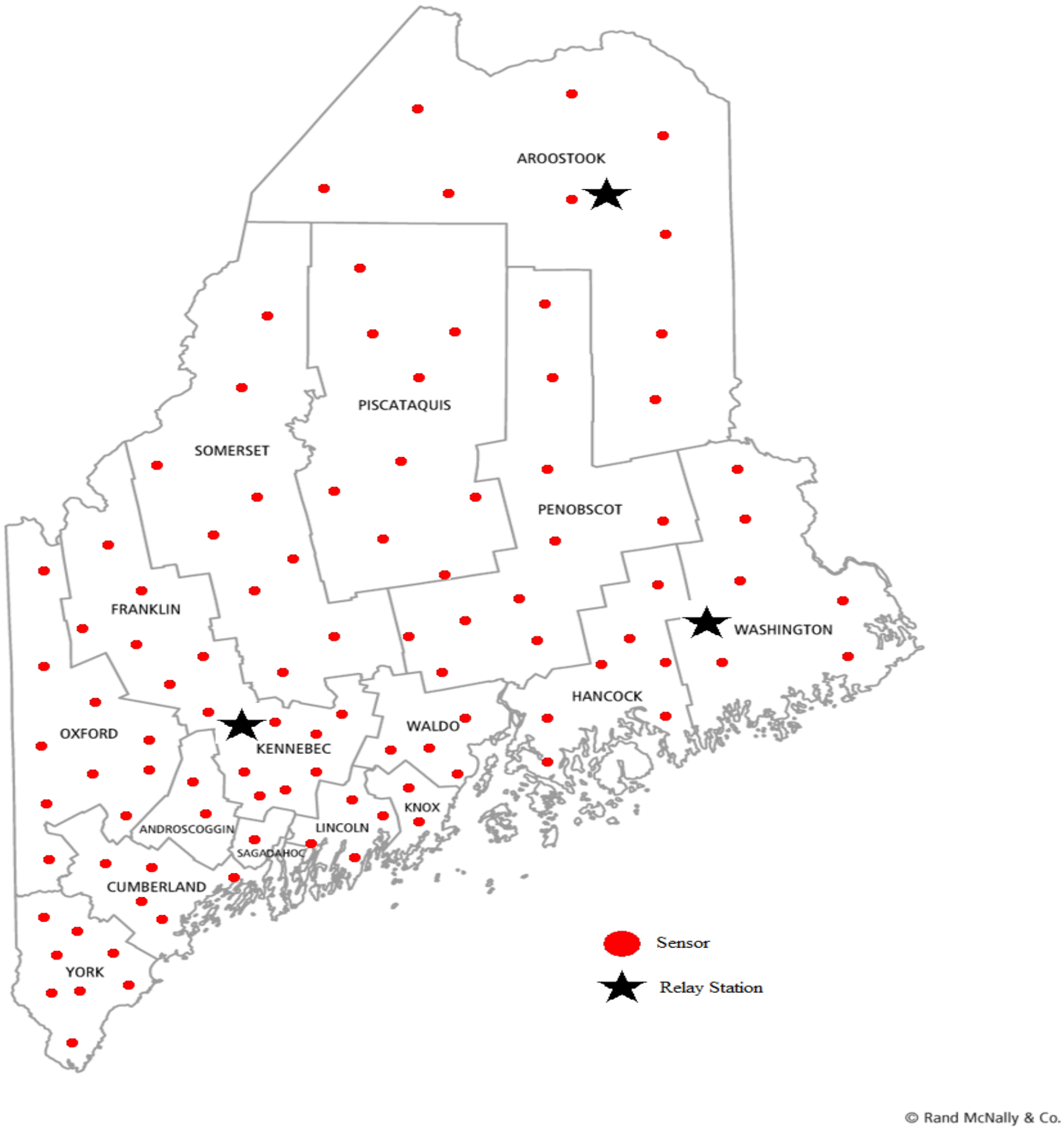


Figure 28. Optimal relay station locations for Maine State, run 3

Figure 28 shows that the model opened a total of 3 optimal relay station locations to meet the required specifications.

5.4. Run 4

Run 4 is carried out with the following parameters; transmission range = 100 miles, coverage percentage = 90%.

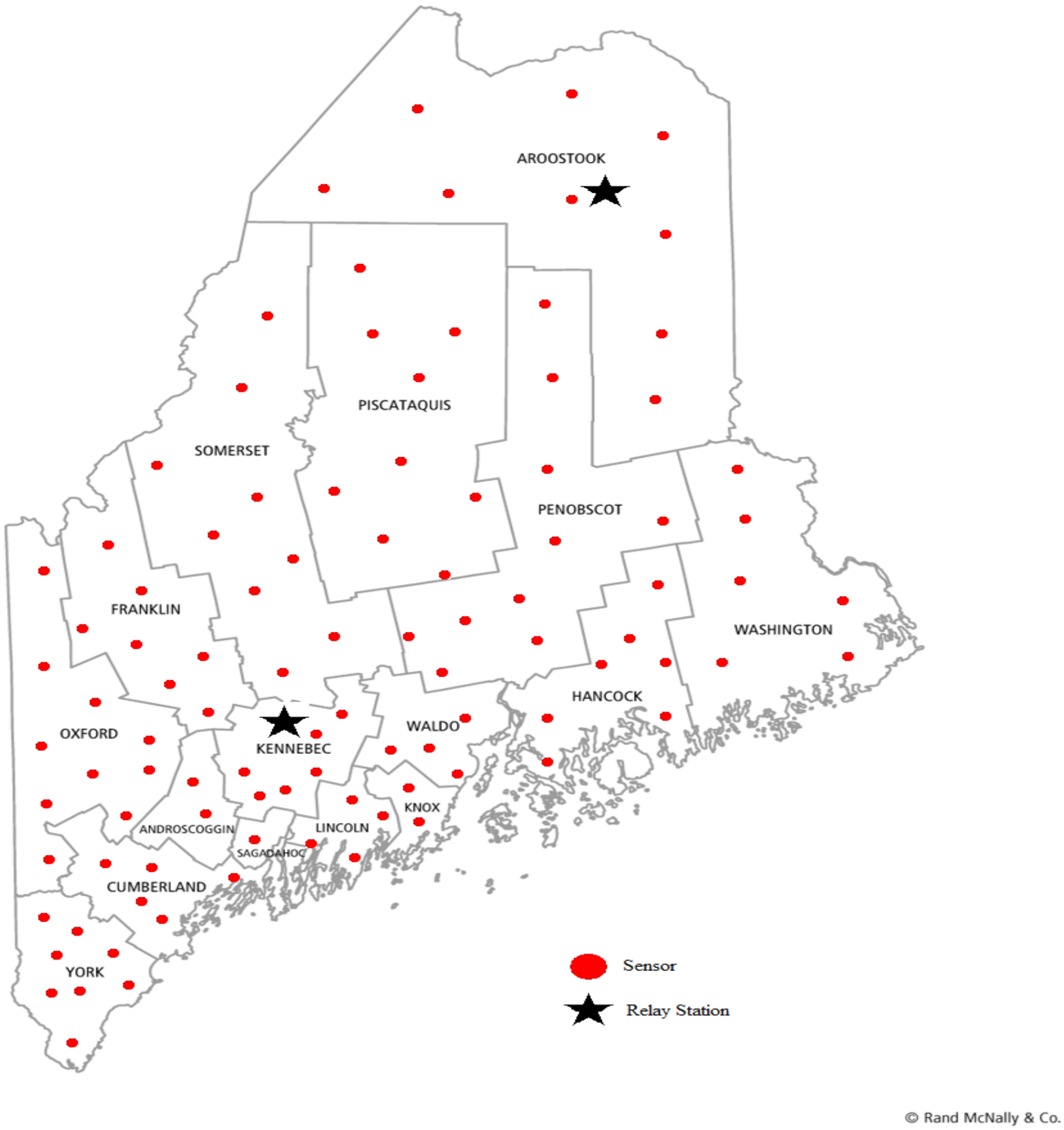


Figure 29. Optimal relay station locations for Maine State, run 4

Figure 29 shows that the model opened a total of 2 optimal relay station locations to meet the required specifications.

From the Maine State case study, we can infer that, for varying transmission ranges and coverage percentages, the model optimizes the number of relay stations required, suggests their placement and at the same time satisfying all the requirements.

In run 1, the number of relay stations opened are reduced by around 55%.

In run 2, the number of relay stations opened are reduced by around 64%.

In run 3, the number of relay stations opened are reduced by around 73%.

In run 4, the number of relay stations opened are reduced by around 82%.

6. CONCLUSIONS AND FUTURE WORK

The current research focuses on determining the optimal number of relay stations and their optimal locations in wireless sensor networks. A mixed integer programming model is developed to determine the optimal number of relay stations, their optimal locations, and optimal assignment of the sensors to the relay stations. The model is coded in the GAMS and is solved by using Xpress-MP Solver. A number of test cases have been developed along with a case study of Maine State.

Some of the key insights are:

1. The number of relay stations opened are reduced between 80% and 95%.
2. As the transmission range of the relay station increase, the number of relay stations opened decrease.
3. As the number of sensors to be covered decrease, the number of relay stations opened decrease.
4. As the scalability of the sensors increase, the number of relay stations opened increase.
5. As the scalability of the sensors increase, the execution time of the MIP model also increase.
6. For all the cases (excepting a few); as the percentage of sensors to be covered increase, the execution time of the MIP model also increase.

The future research includes, but not limited to:

1. Developing optimal relay station locations under robust conditions. Since, sensors have the potential to change their locations; a stochastic model can be developed to determine the optimal relay station locations under uncertain sensor positions.

2. The relay station problem can be combined with the energy saving sensor scheduling model. This has the potential to depict the true optimal locations of the relay stations.
3. Cost for the relay stations can be incorporated into the model based on the transmission range.

7. REFERENCES

[1] Tarik Yarbidi, “Sleep scheduling for energy conservation in wireless sensor networks with partial coverage”, Master's thesis, Bilkent University, Ankara. 2006.

[2] Chia-Pang Chen, Cheng-Long Chuang, Tzu-Shiang Lin, Chia-Yen Lee, Joe-Air Jiang, “A coverage-guaranteed algorithm to improve network lifetime of wireless sensor networks”, Proc. Eurosensors XXIV, 2010.

[3] Fan Ye, Gary Zhong, Songwu Lu, Lixia Zhang, “PEAS: A Robust Energy Conserving Protocol for Long-lived Sensor Networks”, Proc. the 23rd International Conference on Distributed Computing Systems, 2003.

[4] Bin Lin, Pin-Han Ho, Liang-Liang Xie, and Xuemin (Sherman) Shen, “Optimal Relay Station Placement in IEEE 802.16j Networks”, Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing, 2007.

[5] Fangyang Shen, Chunlei Liu, and Jun Zhang, "A Distributed Coverage-Aware Sleep Scheduling Algorithm for Wireless Sensor Networks”, Proceedings of the 2009 Sixth International Conference on Information Technology: New Generations, 2009.

[6] Linnyer Beatrys Ruiz, José Marcos Silva Nogueira, Antônio Alfredo Ferreira Loureiro, "MANNA: A Management Architecture for Wireless Sensor Networks”, Communications Magazine, IEEE, 2003.

[7] Mohamed Younis and Kemal Akkaya, “Strategies and techniques for node placement in wireless sensor networks: A survey”, Ad Hoc Networks, 2008.

[8] Bin Lin, Pin-Han Ho, Liang-Liang Xie, Xuemin (Sherman) Shen and Ja´nos Tapolcai, "Optimal Relay Station Placement in Broadband Wireless Access Networks", IEEE Transactions on Mobile Computing, 2010.

[9] Hsiao-Chen Lu, Wanjiun Liao and Frank Yeong-Sung Lin, "Relay Station Placement Strategy in IEEE 802.16j WiMAX Networks", IEEE Transactions on Communications, 2011.

[10] Xiuzhen Cheng, Ding-Zhu Du, Lusheng Wang and Baogang Xu, "Relay Sensor Placement in Wireless Sensor Networks", Wireless Networks, 2008.

[11] Jason Hill, Robert Szewczyk, Alec Woo, Seth Hollar, David Culler, Kristofer Pister, "System architecture directions for networked sensors", Proceedings of the Ninth International Cnference on Architectural Support for Programming Languages and Operating Systems, 2000.

[12] Gregory J. Pottie, "Wireless sensor networks", Information Theory Workshop, 1998.

[13] Gregory J. Pottie and William J. Kaiser, "Wireless sensor networks", Communications of the ACM, 2000.

[14] Vijay Raghunathan, Curt Schurgers, Sung Park, and Mani B. Srivastava, "Energy Aware Wireless Sensor Networks", IEEE Signal Processing, 2002.

[15] Errol L. Lloyd, Rui Liu, Madhav V. Marathe, Ram Ramanathan, Sekharipuram S. Ravi "Algorithmic Aspects of Topology Control Problems for Ad hoc Networks", MOBIHOC, 2002.

[16] Ram Ramanathan, Regina Rosales-Hain, "Topology control of multihop wireless networks using transmit power adjustment", INFOCOM, 2000.

[17] Xiozhen cheng, Bhagirath Narahari, Rahul Simha, Maggie Xiaoyan Cheng and Dan Liu, "Strong minimum energy topology in wireless sensor networks: NP-completeness and heuristics", IEEE Transactions on Mobile Computing, 2003.

[18] Roger Wattenhofer, Li Li, Paramvir Bahl, Yi-Min Wang, "Distributed Topology Control for Power Efficient Operation in Multihop Wireless Ad Hoc Networks", INFOCOM, 2001.

[19] Yongchul Kim and Mihail L. Sichitiu, "Optimal Placement of Transparent Relay Stations in 802.16j Mobile Multihop Relay Networks", IEICE Transactions on Communications, 2011.

[20] Shahid Mumtaz, Lee Tham Tu, Rasool Sadeghi, Atilio Gameiro, "Performance Evaluation of Fixed and Mobile Relay in WiMAX System", Journal of Digital Information Management, 2010.

[21] RandMcNally <http://www.randmcnally.com/mileage-calculator.do> [accessed 2.7.13].

[22] RandMcNally <http://www.randmcnally.com/> [accessed 2.7.13].

[23] GAMS <http://www.gams.com/docs/intro.htm> [accessed 1.9.13].

[24] NEOS Xpress-MP <http://www.neos-server.org/neos/solvers/milp:XpressMP/GAMS.html> [accessed 1.9.13].

[25] Microsoft Excel <http://office.microsoft.com/en-us/excel/> [accessed 1.9.13].