

CHAPTER 4

AVI ANTENNAS SITE LOCATION OPTIMIZATION

4.1 - Introduction

In many cities the Travel Speed and Travel Time Database is an integral component of the Advanced Traveler Information System (ATIS) program. The data for this database are collected from highways and arterials instrumented with Automated Vehicle Identification (AVI) equipment. Potential benefits of the AVI technology are in the fields of measuring traffic parameters in real time, traffic operations and control, reduction of traffic congestion at transportation facilities (toll plaza facilities), transportation planning (O-D studies), fleet management, information and control, electronic toll collection, national security, vehicle weighting and classification, etc. The basic output of the AVI technologies is accurate traffic information which can help in making appropriate decisions. AVI technologies are able to identify equipped vehicles when they pass certain points on the transportation network links without any action taken by the drivers.

The users can obtain the ATIS data via In-Vehicle Navigation Device (IVN) and Traveler Information Kiosks. The quality of AVI coverage and the total costs of the AVI system primarily depend on the number and chosen locations of the AVI antennas site. The decisions about the number and locations of the AVI antennas site are usually made after evaluating available funding and desired areas of AVI coverage. In other words, these decisions are primarily based on subjective judgment of the experts. On the other hand, the problem of determining the locations of the AVI antennas site is combinatorial by its nature. The quality of a location highly depends on the measure of effectiveness being used. Every possible deployment of the AVI readers in the space is characterized by the appropriate number of “readings”. When determining the number and locations of the AVI antennas in the space, the O-D matrix must be taken into account, as well as the set of shortest paths between all pairs of nodes. The basic idea behind this statement is that the readers must be located in such a way that they are able to “catch” the greatest number of trips in the network.

Let us introduce the following notation:

N - the set of the nodes in the network

A - the set of the links in the network

$|N|$ - the total number of nodes in the network

$|A|$ - the total number of links in the network

T - the total number of trips in the network

I - the set of all O-D pairs in the network (the set of all shortest paths in the network)

$|I|$ - the total number of origin-destination pairs in the network

R - the set of the readers in the network

$|R|$ - the total number of readers in the network

It is assumed in this model that the total number of trips in the network T is known. In this model the total number of trips in the network T is treated as a deterministic quantity. In figure 3.1 we showed one possible deployment of the readers in the network. This deployment is characterized by the appropriate total number of achieved “readings” in the network, as well as by the total number of covered origin-destination pairs by AVI readers. Let us explain these terms into more details. One reading means that the same vehicle is recorded on two different locations in the network (Figure 4.2).

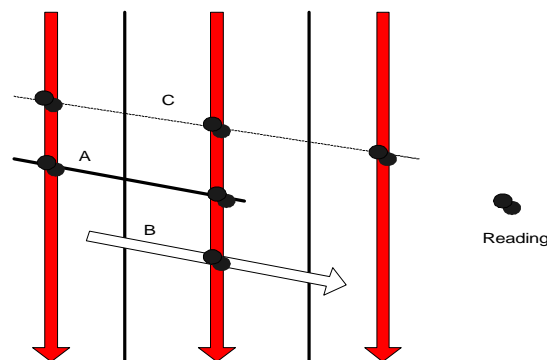
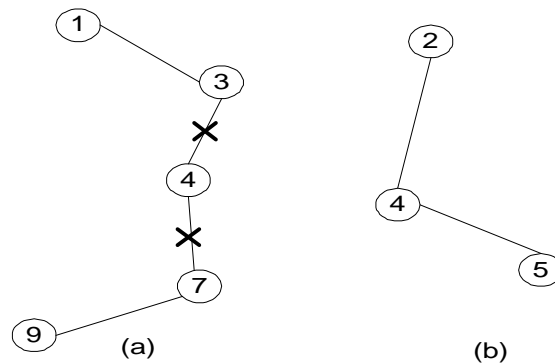


Figure 4.2 – Graphical Representation of the AVI Readings

In Figure 4.2 by heavy lines are denoted locations containing AVI readers. As we can see from the Figure 4.2, vehicle A produced one reading, vehicle B produced no readings, while vehicle C produced two readings. In Figure 4.3(a) one shortest path connecting one particular O-D pair is shown. Existing at least two AVI readers along this path will enable receiving some information about travel time along the part of this path. The more AVI readers along some path we have the more accurate information about travel times along that path is. Figure 4.3(b) shows shortest path between some other O-D pair with no AVI readers along the path. In the case of path shown in Figure 4.3(a) we will say that the path is covered by the readers, while in the case of path shown in Figure 4.3(b) we can say that the path is not covered.



**Figure 4.3 – (a) Shortest Path “Covered” by the Readers
(b) Shortest Path Not “Covered” by the Readers**

Let us denote by \mathbf{Re} the total number of achieved readings in the network. Obviously, some different deployment of readers in the network would cause different total number of readings. Let us also denote by \mathbf{Iod} - the total number of covered origin–destination pairs by readings. Different readers deployments correspond to the different total number of covered origin destination pairs. It is very logical that the questions of the following type could appear: (a) Where should we locate $|R|$ readers to maximize (or minimize) defined objective function? (b) What is the smallest number of readers which will provide to us certain standard of performance and where these readers should be located? It was assumed in this model that the total number of AVI readers $|R|$ is known. The potential locations of the readers are also known (It is assumed

that the readers could be located anywhere at any link in the network). The problem considered in this chapter can be formulated in the following way: For the given number of the AVI readers find the optimal readers locations along the network links in order to maximize (minimize) defined objective function.

4.3 - Proposed Solution to the Problem

When determining the best locations for the AVI equipment we will use simplified assumption that all drivers use the shortest path between origin and destination. Let us denote by f_i - the flow along the shortest path i . Let us again consider our transportation network (Figure 4.1). The length of any link (i, j) is denoted by $l(i, j)$. Using “classical” Floyd’s algorithm we can find the shortest paths between all pairs of nodes.

It is also assumed in this model that the origin-destination matrix is known. The O-D matrix could be obtained from the planning model (or by fields measurements) and is the basic, rough indicator of the traffic pattern in the considered network. The elements of the O-D matrix represent the total number of monthly, weekly, daily or hourly trips between particular node pair. Every element in the O-D matrix is treated in this paper as deterministic quantity. As it is well known, one link can be part of a large number of different paths (Figure 4.4).

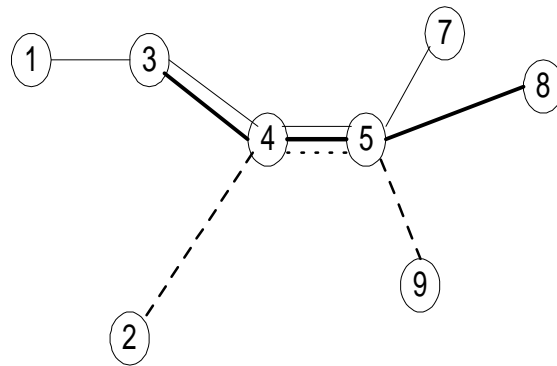


Figure 4.4 - Link Which Is Part of A Large Number of Different Paths

Let us introduce binary variables d_{ia} that are defined as follows:

$$\begin{aligned}
\mathbf{d}_{ia} = & \begin{cases} 1 & \text{if link } a \in A \text{ is part of the shortest path } i \in I \\ 0 & \text{otherwise} \end{cases} \quad (4.1)
\end{aligned}$$

The shortest paths are known for all node pairs in the network. The link considered could be part to none of them, it can belong to one shortest path, it can belong to two different shortest paths, to three different shortest paths, etc,... In this research the simplified assumption was made that all users will use the shortest path during their trip. Using this assumption, and taking the appropriate flow values from the O-D matrix, it is possible to calculate the flows along all shortest paths.

Based on Figure 4.3, we conclude that the following relation must be satisfied:

$$\sum_{i \in I} \mathbf{d}_{ia} f_i = f_a \quad \forall a \in A \quad (4.2)$$

In other words, the sum of all flows on different shortest paths which include the considered link must be equal to the flow on the link.

Let us introduce the following binary variables:

$$\begin{aligned}
\mathbf{x}_{ra} = & \begin{cases} 1 & \text{if reader } r \in R \text{ is located on the link } a \in A \\ 0 & \text{otherwise} \end{cases} \quad (4.3)
\end{aligned}$$

As we already mentioned, existing at least two AVI readers along the path will enable receiving some information about travel time along the part of this path. The total number of readings \mathbf{Re}_i along the i -th path equals:

$$\mathbf{Re}_i = d_i \max \left\{ 0, \left[\sum_{a=1}^{|A|} \sum_{r=1}^{|R|} \mathbf{d}_{ia} \mathbf{x}_{ra} - 1 \right] \right\} \quad (4.4)$$

Where:

d_i - the total number of trips along i-th path

The total number of readings in the network Re equals:

$$Re = \sum_{i=1}^{|I|} Re_i \quad (4.5)$$

Let us introduce the following binary variables:

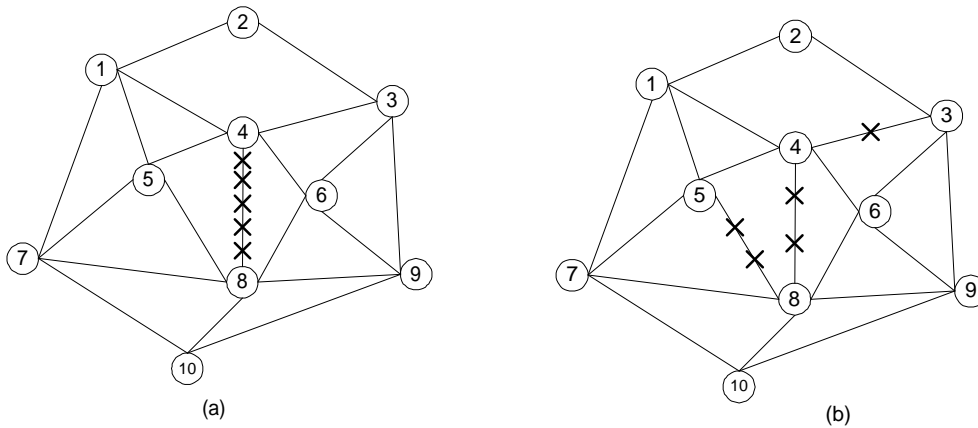
$$y_i = \begin{cases} 1 & \text{when } \sum_{a=1}^{|A|} \sum_{r=1}^{|R|} d_{ia} x_{ra} > I \\ 0 & \text{otherwise} \end{cases} \quad (4.6)$$

Binary variable y_i is equal to 1 when it is possible to obtain readings along i-th path. This binary variable equals 0 when there is maximum 1 reader along the path. In this case we do not have any information about travel time. The total number of different O-D pairs covered by readers is:

$$I_{OD} = \sum_{i=1}^{|I|} y_i \quad (4.7)$$

Let us return to our basic question: Where should we locate $|R|$ readers to maximize (or minimize) defined objective function? To answer this question properly, let us first try to define the objective function. It is naturally to try to make as much as possible readings with the available readers. In other words, it is logical to try to maximize the total number of readings with the available number of readers. Let us assume that the total number of readers which are placed in certain positions could produce the total number of readings which is equal to 1,000 and that the all readings correspond only to one pair of the nodes (Figure 4.4(a)). Let us change

the locations of the readers and let us assume that the new locations could produce 500 readings which are readings between 15 different pairs of nodes (Figure 4.5(b)).



**Figure 4.5 (a) Total of 1,000 Readings Between One Pair of Nodes;
(b) Total of 500 Readings Between 15 Different Pairs of Nodes**

Which locations are better? Obviously, the second locations will produce smaller total number of readings, but at the same time the “coverage” of the network will be much better than in the previous case. In other words, when determining the quality of certain locations it is necessary to take into account the total number of readings, as well as the total number of “covered” pairs of nodes by readings. It was decided in this model to introduce the following performance index as a measure of the quality of readers locations:

$$F = w_1 \frac{Re}{T} + w_2 \frac{Iod}{|I|} \quad (4.8)$$

Where:

$$w_1 + w_2 = 1$$

w_1 - the weight (importance) that we are assigning to the maximization of the total number of readings

w_2 - the weight (importance) that we are assigning to the maximization of the total number of origin-destination pairs that are covered by readings

The highest the performance index values, the better the locations considered. As we can see, the performance index is composed of two components. First component reflects our desire to maximize the total number of readings, while the second component reflects our desire to maximize the total number of the origin-destination pairs covered by readings. When $w_1 = 1$, we are taking care exclusively about maximization of the total number of readings. In the case when $w_2 = 1$, we are taking care exclusively about maximizing the total number of covered origin-destination pairs. The model for determining the AVI locations is thus given by:

Maximize

$$F = w_1 \sum_{i=1}^{|I|} \frac{d_i \max \left\{ 0, \left[\sum_{a=1}^{|A|} \sum_{r=1}^{|R|} d_{ia} x_{ra} - 1 \right] \right\}}{T} + w_2 \frac{\sum_{i=1}^{|I|} y_i}{|I|} \quad (4.9)$$

Subject to

$$\sum_{a=1}^{|A|} x_{ra} = 1 \quad r = 1, 2, \dots, |R| \quad (4.10)$$

$$0 \leq \sum_{r=1}^{|R|} x_{ra} \leq |R| \quad a = 1, 2, \dots, |A| \quad (4.11)$$

$$x_{ra} = 0 \quad \text{or} \quad 1 \quad (4.12)$$

Relation (4.9) express our wish to maximize defined objective function. Relation (4.10) explains the fact the every reader must be located on some link in the network. Any link could be without reader, with one reader, with two readers, ..., or with all $|R|$ readers. This is explained by the relation (4.11).

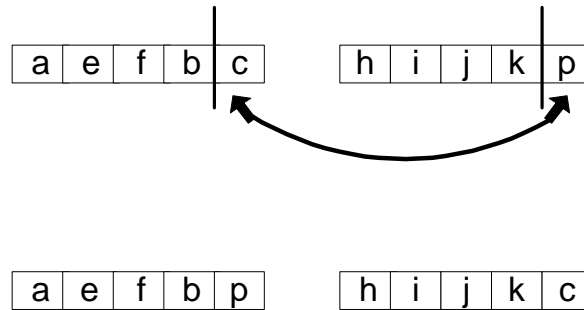
4.4 - Searching the Best Locations of the Automated Vehicle Identification Equipment Using Genetic Algorithms

Finding the best locations of the automated vehicle identification equipment is difficult combinatorial optimization problem. Genetic algorithms (Holland (1975), Goldberg (1989)) represent search techniques based on the mechanics of nature selection used in solving complex combinatorial optimization problems. In the case of genetic algorithms, as opposed to traditional search techniques, the search is run in parallel from a population of solutions. In our case one solution is one particular placement of the automated vehicle identification equipment in the network. In the first step, various solutions are generated. In this project we generated these first generation of the solutions randomly. In the next step, the evaluation of these solutions that is, the estimation of the defined performance index is made. Some of the “good” solutions yielding a better “fitness” (objective function value) are further considered. The remaining solutions are eliminated from consideration. The chosen solutions undergo the phases of reproduction, crossover and mutation. After that, a new generation of solutions is produced to be followed by a new one, and so on. Each new generation is expected to be “better “ than the previous one. The production of new generations is stopped when a prespecified stopping condition is satisfied. The final solution of the considered problem is the best solution generated during the search. In the case of genetic algorithms an encoded parameter set is used. Most frequently, binary coding is used. The set of decision variables for a given problem is encoded into a bit string (chromosome, individual). Genetic algorithms are procedure where the strings with better fitness values are more likely to be selected for mating. Let us denote by F_i the value of the objective function (fitness) of string i . The probability p_i for string i (i -th placement of the automated vehicle identification equipment in the network) to be selected for mating is equal to the ratio of F_i to the sum of all strings’ objective function values in the population:

$$p_i = \frac{F_i}{\sum_j F_j}$$

This type of reproduction, that is selection for mating represents a proportional selection known as the “Roulette wheel selection”. Crossover operator was used to combine the genetic material.

At the beginning pairs of strings (parents) were randomly chosen from a set of previously selected strings. Later, for each selected pair the location for crossover is randomly chosen. Each pair of parents creates two offsprings (Figure 4.6)



**Figure 4.6 – A Single-point Crossover Operator (a) Two Parents
(b) Randomly Chosen Location is Before the Last Bit (c) Two Offsprings**

After completing crossover, the genetic operator mutation is used. In the case of binary coding, mutation of a certain number of genes refers to the change in value from 1 to 0 or vice versa.. The probability of mutation is very small (of order of magnitude 1/1000). The purpose of mutation is to prevent an irretrievable loss of the genetic material at some point along the link.

Searching the best locations of the automated vehicle identification equipment was based in this research on the following genetic algorithms:

Step 1:

Form the initial population $P(0)$ consisting of n string (n possible placement of the automated vehicle identification equipment). Make an evaluation of the fitness of each string.

Step2:

Select n parents from the current population. (selection probability is proportional to the fitness).

Step 3:

Randomly select a pair of parents for mating. Create two offsprings by exchanging strings with the one –point crossover.. To each of the sreated offsprings apply mutation. Apply crossover and mutation operators until n offsprings (new population) are created.

Step 4:

Substitute the old population of strings with the new population. Evaluate the fitness of all members in the new population.

Step 5:

If the number of generations (populations) is smaller than the maximal prespecified number of generations, go back to Step 2. Otherwise, stop the algorithm. For the final solution chose the best string discovered during the search.

4.5 - Numerical Example

The proposed model is accompanied by a computer program written in C++. The developed algorithm was tested on the example of the transportation network shown in Figure 4.7. The network consists of 10 nodes and 20 links.

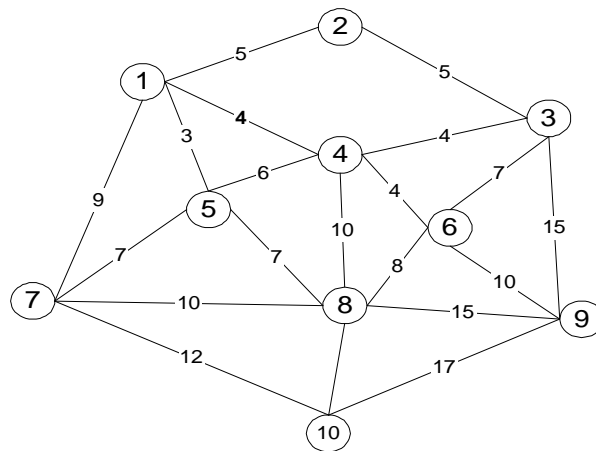


Figure 4.7 – Transportation Network Used for Testing

The corresponding O-D matrix is assumed in the following:

To node	1	2	3	4	5	6	7	8	9	10
From node 1	-	500	200	700	800	300	200	300	200	500
2	400	-	500	600	300	800	100	100	200	700
3	150	500	-	100	200	300	400	100	800	700
4	650	600	200	-	100	200	300	400	200	100
5	900	200	300	100	-	200	200	100	100	200
6	350	700	300	300	250	-	100	200	100	100
7	300	200	400	300	200	150	-	300	400	100
8	400	100	200	400	100	100	300	-	100	100
9	200	200	700	100	100	100	400	100	-	200
10	450	600	600	100	300	100	100	100	200	-

In this paper we entered the genetic algorithm search with the following parameters:

- The total number of strings (possible placement of the automated vehicle identification equipment) in one population is 100
- The total number of population is 100
- The probability of the mutation is 1/1000

Figure 4.8 shows the alterations in the objective function value depending on the number of generations (these results refer to the case when $w_1=0.8, w_2=0.2$).As we can see from the figure 4.8 objective function values converge relatively fast (there are minor changes in the objective function value between generation 20 and generation 100).

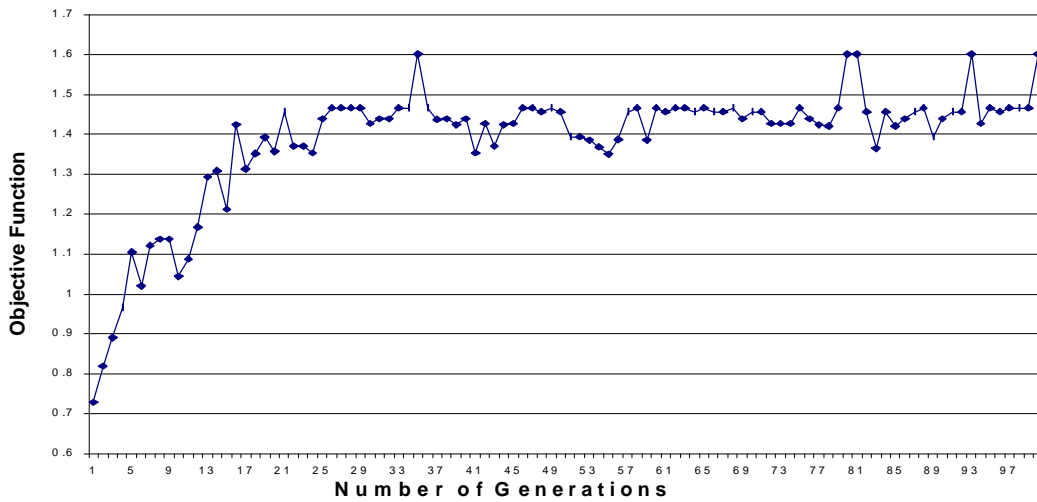


Figure 4.8 – Alterations in the objective function value depending on the number of generation.

Figure 4.9 shows the corresponding alterations in the total number of readings depending on the number of generations.

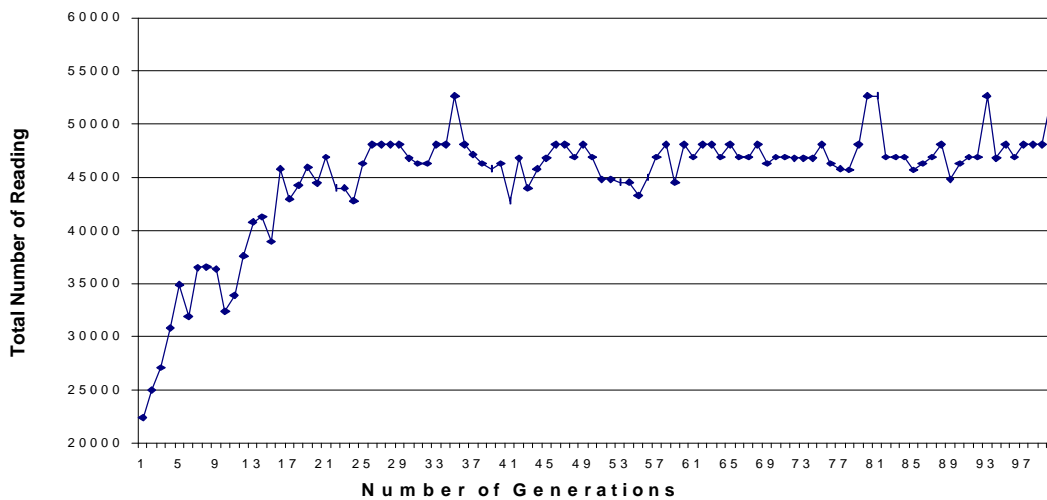


Figure 4.9 - Alterations in the total number of readings depending on the number of generations

An analysis of the influence of the weight (importance) of individual criteria on the results obtained was also made. Figure 4.10 shows changes in the total number of readings, as well as total number of “covered” O-D pairs as a function of weight w_1 (the importance we give to maximize the total number of readings in the network).

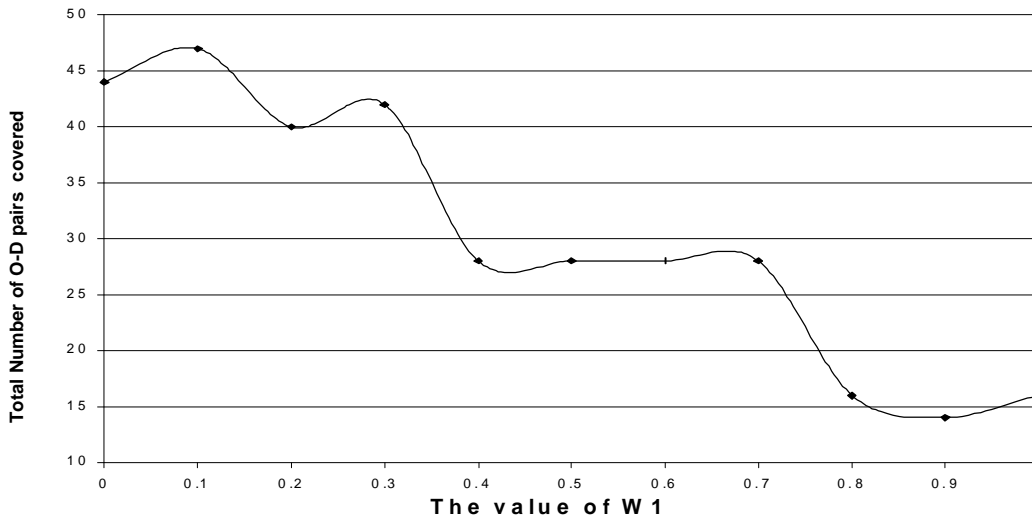


Figure 4.10 - Changes in the total number of readings and the total number of “covered” O-D pairs as a function of weight w_1

As can be seen in Figure 4.10 the chosen criteria weights have an extremely great influence on the values of the criteria. There is clearly noted general tendency for the total number of readings to increase and for the total number of “covered” O-D pairs to decrease with a rise in the value of weight w_1 . Varying the weight values enables the generation of a large number of different deployment of the readers in the network which facilitates the analyst understanding of the problem and the choice of the final solution. Finally, an analysis of the influence of the number of readers in the network is also done. Figure 4.11 shows changes in the total number of readings as a function of the total number of readers.

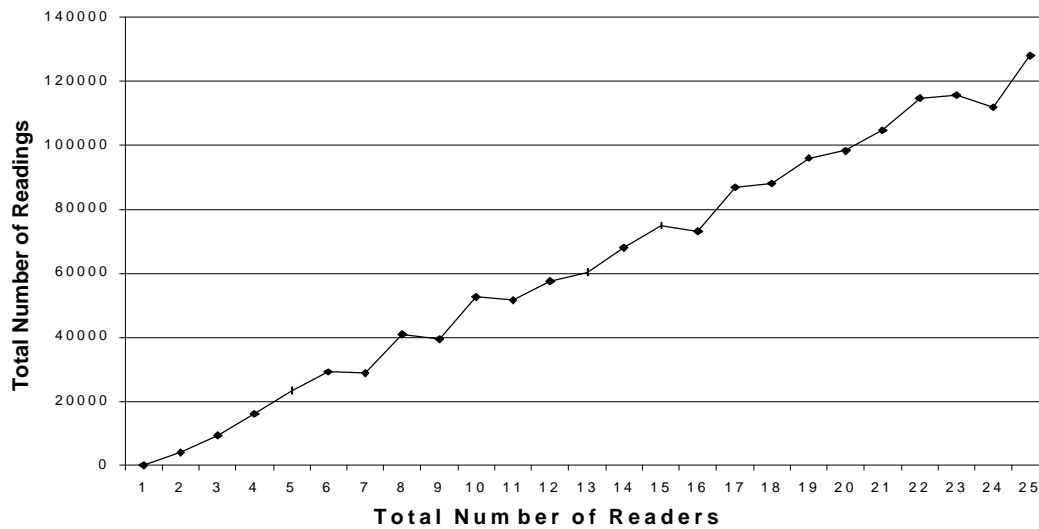


Figure 4.11 - Alterations in the total number of readings depending on the number of readers in the network

4.6 - Summary

This chapter develops a model to solve the locations of the automated vehicle identification site problem in the large-scale transportation networks. Bearing in mind the combinatorial nature of the problem, its possible large dimensions, and the desire to solve problem based on more than one criteria, the proposed algorithm is heuristic.

The developed model was tested on a relatively small numerical example. The value of O-D matrix was assumed and fixed. The achieved CPU times in this stage of the research are promising. Testing of the model developed on the greater transportation networks is one of the directions for the future research. The development of models based on other Metaheuristic approaches (Simulated Annealing, Taboo Search) is certainly one of the extremely important directions for future research. Testing of the model with different O-D matrix, such as peak hour O-D matrix, is also one of the important directions for further research.

The optimized AVI antennas site locations will maximize the “coverage” of the network and thus produce more representative traffic data information. The accuracy and reliability of the overall traffic data collection system and travel time forecasting also will increase because of this. The travel time forecasting is presented in *chapter 5*.