

Low complexity heuristics for multi-objective sensor placement in traffic networks

Marina Sofokleous, Yiolanda Englezou, Aristotelis Savva, Stelios Timotheou and Christos G. Panayiotou

Abstract—Monitoring the traffic state of the road network is very important for a plethora of reasons, such as the prevention of traffic congestion and the development of estimation and control policies. In order to efficiently obtain high-quality information on traffic, the sensors must be installed at optimal locations in the road network under study. This problem is known as the Network Sensor Location Problem (NSLP). In this work, a multi-objective NSLP is proposed for the installation of a pre-defined number of sensors to maximise (i) the covered traffic flow volume and (ii) the minimum distance between candidate links for sensor installation, while taking into account pre-installed sensors. We reformulate the problem into a single-objective mixed-integer linear program (MILP) that yields the optimal sensor locations. In addition, we propose four low-complexity heuristics for the solution of the problem. The performance of the proposed algorithms is evaluated for the traffic network of the Republic of Cyprus under real-life conditions and traffic data. Results show that the four low-complexity approaches yield a different trade-off between execution speed and solution quality.

I. INTRODUCTION

Traffic monitoring has been one of the major tools used for transportation planning, decision-making and the implementation of various control strategies [1]. Traffic monitoring tasks include traffic state estimation (TSE) [2], origin-destination (OD) matrix estimation [3], travel-time [4], and queue estimation [5] and traffic state and demand prediction [6]. Traffic counts, i.e. the number of vehicles passing through a specific road in specific temporal windows, have been long used for traffic monitoring purposes. Stationary sensors (e.g. loop detectors, cameras) have been widely used to collect measurements of the flow at observed links (e.g. links equipped with loop detectors), as well as to provide information for inference at unobserved links. These sensors are used to obtain traffic counts with their position being of high significance as it affects both the quality and the quantity of the acquired information.

The Network Sensor Location Problem (NSLP) has received significant attention for many decades and numerous

studies in the literature have been proposed to provide efficient solutions subject to an objective related to the problem at hand [7]. Typical objectives used in the literature include the accurate estimation of OD flows, the efficient traffic state estimation/observation, and the maximum coverage of traffic flow [8]. Typically the problem is formulated as a mixed-integer mathematical program that aims to determine the minimal set of locations from which the flows of the whole network are obtained, utilising error-free sensor measurements. In terms of OD estimation, several works have been proposed using different mathematical models [9], or data from heterogeneous sources while eliminating the need of data which might be difficult to obtain [10]. Other works focused on OD flows and travel time estimation from an information theoretical point of view and under several possible cost functions [11]. A number of papers considered the NSLP for efficient traffic state estimation. A recent study, proposed an optimisation problem aiming to minimise the error in the traffic density estimation using information of the traffic flow and turning proportions [12]. Furthermore, an Ant Colony Optimisation technique was proposed to maximise the number of captured traffic flows [13]. Incorporating pre-existing sensors while maximising the obtained traffic flow volumes was studied in [14], using a modified version of the Maximum Flow Coverage (MFC) method [7].

The NSLP still remains a challenging problem due to (i) the difficulty in obtaining the required data, such as turning proportions and information on route choice in real networks [10], (ii) the practical difficulty in solving the NSLP for real-life, large-scale road networks [9], (iii) the incorporation of pre-existing sensors into the solution [14], and (iv) the partial observability of the problem, due to the sparsity of existing and installed sensors [15]. Heterogeneous sources of data [10] and budgetary constraints [16] also impose additional challenges in solving the NSLP. Despite the extensive literature on the NSLP, there is a major gap in the application of the proposed methodologies in real-life large-scale networks. Some research works presented numerical tests or case studies using simulated data or small networks [13]. Others presented case studies using sections of real road networks of moderate size (e.g. 300-500 road links) [17]. Two case studies on real-life networks focused on the NSLP for OD matrix estimation; a case study of the A3 Naples–Salerno motorway with 1,414 road links [9] and of the road network of the Province of Benevento in Italy with 1,800 road links [18]. More recently, a case study was published for the road network of the city of Thessaloniki in Greece with a total of 137,852 road links from which 3,073 links were selected to be included in the optimisation procedure for traffic volume observation and estimation [14],

This work is supported by the European Union (i. ERC, URANUS, No. 101088124, and ii. Horizon 2020 Teaming, KIOS CoE, No. 739551), and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation, and Digital Strategy. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

M. Sofokleous, Y. Englezou, S. Timotheou and C. G. Panayiotou are with the KIOS Research and Innovation Center of Excellence, and the Department of Electrical and Computer Engineering, University of Cyprus, {sofokleous.marina, englezou.yiolanda, timotheou.stelios, christosp}@ucy.ac.cy.

A. Savva is with the Public Works Department of the Ministry of Transport, Communications & Works, Nicosia, Cyprus, asavva@pwd.mcw.gov.cy.

however, pre-existing sensors where not taken into account.

In this context, this work aims to define a methodology for the installation of a pre-defined number of loop detectors while taking into account pre-existing sensors. The aim of the proposed approach is to simultaneously maximise the flow coverage of the network under study while maximising the minimum distance of candidate installation links. The main contributions of this work are as follows.

- Building upon the NSLP literature, we formulate a multi-objective optimisation problem to maximise (i) the covered traffic flow and (ii) the minimum distance between candidate links for sensor installation.
- We reformulate the problem to a single-objective Mixed Integer Linear Program (MILP). While MILP solvers can optimally solve combinatorial problems, such as the NSLP, their usage is often avoided.
- We propose four low-complexity heuristics for the solution of the considered problem that yield a different trade-off between execution speed and solution quality.
- The proposed approach is applied to the traffic network of the Republic of Cyprus, using real-life data, which consists of 40 pre-specified sensors and 74 new sensors to be installed¹. The developed solution approaches are extensively compared in terms of their execution speed and solution quality.

Following, Section II formulates the problem and Section III presents the MILP formulation and the proposed heuristic solution approaches. Section V discusses the available traffic data and the main results obtained in the real-life network of the Republic of Cyprus. Finally, Section VI discusses the main contributions.

II. PROBLEM STATEMENT

Assume a road traffic network consisting of the set of links, \mathcal{L} , where $L = |\mathcal{L}|$ is the total number of links, and the set of nodes \mathcal{N} , with $N = |\mathcal{N}|$ the total number of nodes. In addition, we define the set of road links that are already equipped with static sensors, \mathcal{M} , with $M = |\mathcal{M}|$ the total number of pre-installed sensors and $\mathcal{P} = \mathcal{L} \setminus \mathcal{M}$, i.e. \mathcal{P} contains all links that do not have a pre-installed sensor and can be chosen for new sensor installation.

The main goal of NSLP is to find the optimal location of K sensors to simultaneously maximise the traffic flow covered and the spread of the sensors into the traffic network. Towards this direction, we define a bi-objective optimization problem as follows:

$$\text{maximize } \{\Psi_1(x), \Psi_2(x)\} \quad (1a)$$

$$\text{subject to: } \sum_{i \in \mathcal{P}} x_i = K, \quad (1b)$$

$$x_i = 1, \forall i \in \mathcal{M}, x_i \in \{0, 1\}, \forall i \in \mathcal{L}. \quad (1c)$$

In formulation (1), x_i is a binary decision variable such that $x_i = 0$ if link $i \in \mathcal{L}$ is selected as a sensor installation, otherwise $x_i = 1$. Objective $\Psi_1(x)$, considers the maximisation of traffic flow covered by the sensors given by $\Psi_1(x) = \sum_{i \in \mathcal{L}} f_i x_i$, where f_i is the traffic flow volume, given in

Passenger Car Units (PCUs), at each link $i \in \mathcal{L}$. Objective $\Psi_2(x)$, considers the maximisation of the minimum distance between the pre-existing and newly installed sensors, i.e. $\Psi_2(x) = \min_{i, j \in \mathcal{L}} d_{ij}(x)$, where

$$d_{ij}(x) = \begin{cases} D_{ij} & \text{if } x_i = 1 \text{ and } x_j = 1, \forall i, j \in \mathcal{L}, \\ \infty, & \text{otherwise,} \end{cases} \quad (2)$$

where D_{ij} is the Euclidean distance between links i and j . Note that the geometry of the road network is considered during the calculation of the distances, hence, vehicle directions are also considered. Therefore, the calculated distances refer to the distance that a vehicle is required to traverse from the upstream point of link i to the downstream point of link j , $\forall i, j \in \mathcal{L}$. Constraints (1b) and (1c) indicate the need to install K new sensors, identify the pre-installed sensors, and define the binary decision variables, respectively.

Formulation (1) is transformed into a single-objective problem by considering the second objective $\Psi_2(x)$ as a constraint, according to the ϵ -Constraint method [19]. This yields the following optimisation problem:

$$\text{maximize } \sum_{i \in \mathcal{L}} f_i x_i \quad (3a)$$

$$\text{subject to: } \min_{i \in \mathcal{P}, j \in \mathcal{L}} d_{ij}(x) \geq \epsilon, \quad (3b)$$

$$\sum_{i \in \mathcal{P}} x_i = K, \quad (3c)$$

$$x_i = 1, \forall i \in \mathcal{M}, x_i \in \{0, 1\}, \forall i \in \mathcal{L}. \quad (3d)$$

The Pareto front of problem (3) is challenging to solve due to the definition of $d_{ij}(x)$ given in Equation (2).

III. OPTIMAL SOLUTION APPROACH

To overcome the complexity of Problem (3) due to the definition of $d_{ij}(x)$, we formulate a new optimisation problem:

$$\text{maximize } \sum_{i \in \mathcal{L}} f_i x_i, \quad (4a)$$

$$\text{subject to: } \sum_{i \in \mathcal{L}} x_i = K, \quad (4b)$$

$$x_j \leq 1 - x_i, \forall j \in \mathcal{N}_i^\epsilon, i \in \mathcal{P} \quad (4c)$$

$$x_i = 1, \forall i \in \mathcal{M}, x_i \in \{0, 1\}, \forall i \in \mathcal{L}. \quad (4d)$$

In (4), $\mathcal{N}_i^\epsilon = \{j \in \mathcal{L} : D_{ij} < \epsilon\}$ is the set of links where no sensors can be installed in the neighbourhood of link i , if a sensor is installed on the particular link, due to the distance constraint. As a result, constraint (4c) ensures that if a sensor is installed on link i , and $j \in \mathcal{N}_i^\epsilon$ is a neighbouring link within the distance threshold ϵ , then no sensor can be installed on link j .

Problem (4) is a Mixed Integer Program (MIP), a variation of the set covering problem, which is NP-hard to solve in the general case. The optimal solution of Problem (4) can be obtained using MILP solvers which rely on branch-and-bound algorithms [20]. While MILP software can optimally solve combinatorial problems, their usage is often avoided for two reasons. First, MILP solvers cannot provide guarantees such that execution times can vary by

¹The case study was conducted in collaboration with the Public Works Department of the Ministry of Transport, Communications and Works.

several orders of magnitude, even for optimization problem instances of the same size. Second, commercial solvers, such as Gurobi [21], are very powerful in solving combinatorial optimization problems and orders of magnitude faster than non-commercial solvers [22], but they are expensive. Hence, for computational and budgetary purposes, we propose the use of four low-complexity heuristics that yield near-optimal solutions at low computational cost.

IV. LOW-COMPLEXITY HEURISTICS

A. Greedy Heuristics

The greedy algorithm is known as one of the best polynomial time algorithm for maximum coverage problems. A greedy algorithm is any algorithm that follows the problem-solving heuristic of making the locally optimal choice at each stage in a greedy fashion. Specifically, for the NSLP the algorithm works by iteratively selecting one link $i \in \mathcal{L}$ at each iteration under some rule until K links are selected.

First we formulate the *Deterministic Greedy Heuristic* (DGH) approach, in which the optimisation procedure selects one link i at each iteration, such that $f_i = \max(f_j), \forall j \in \mathcal{P}$, i.e. the link that has the maximum traffic flow. The proposed algorithm is given as follows.

- Set $\mathcal{B} \equiv \mathcal{P}$.
- Repeat until K links are selected:
 - 1) Select link $i \in \mathcal{B}$ that has the highest traffic flow, $f_i = \max(f_j), \forall j \in \mathcal{B}$.
 - 2) Remove the chosen link i and the neighbouring links using the adjacency matrix from set \mathcal{P} such that $\mathcal{B} = \mathcal{P} \setminus (\{i\} \cup \{j\}_{j \in \mathcal{N}_i^\varepsilon})$, i.e. remove link i and all neighbouring links that lie within a certain distance threshold.
- Return the K total chosen links and the objective function value ψ^{DGH} .

The second proposed low-complexity algorithm is the *Random Greedy Heuristic* (RGH). The RGH algorithm is run for a fixed number of trials S such that the best solution (higher objective function value) over the total number of trials is obtained. In each trial, a weighted random selection procedure is followed and the optimisation procedure is repeated until K links are selected. The proposed algorithm is given as follows.

- For $s = 1, \dots, S$:
 - Set $\mathcal{B} \equiv \mathcal{P}$.
 - Repeat until K links are selected:
 - 1) Calculate the probability of selection given as

$$w_i^{(s)} = \begin{cases} f_i \setminus \sum_{i \in \mathcal{L}} f_i, & \text{if } i \in \mathcal{B}, \\ 0, & \text{if } i \in \mathcal{L} \setminus \mathcal{B} \end{cases}$$

- 2) Sample $u \sim \text{Uniform}(0, 1)$.
- 3) Select the first link $i \in \mathcal{B}$ that satisfies $u < \sum_{j=1}^i w_j^{(s)}$.
- 4) Remove the chosen link i and the neighbouring links from set \mathcal{P} such that $\mathcal{B} = \mathcal{P} \setminus (\{i\} \cup \{j\}_{j \in \mathcal{N}_i^\varepsilon})$ using the adjacency matrix, i.e. remove link i and all neighbouring links that lie within a certain distance threshold.

- Return the K total chosen links and the objective function value $\psi^{s, RGH}$.
- Select the solution with the highest objective function value, $\psi^{RGH} = \max_{s=1, \dots, S} \psi^{s, RGH}$.

The greedy approaches can be combined to yield the *Combined Greedy Heuristic* (CGH) approach that takes the best (maximum) solution between the DGH and RGH, as described above.

B. Linear Programming (LP) Heuristics

To avoid the combinatorial nature of Problem (4), we relax the binary variables to take continuous values, yielding:

$$\text{maximise } \sum_{i \in \mathcal{L}} f_i x_i, \quad (5a)$$

$$\text{subject to: } \sum_{i \in \mathcal{L}} x_i = K, \quad (5b)$$

$$x_j \leq 1 - x_i, \forall j \in \mathcal{N}_i^\varepsilon, i \in \mathcal{P} \quad (5c)$$

$$x_i = 1, \forall i \in \mathcal{M}, \quad (5d)$$

$$x_i \in [0, 1], \forall i \in \mathcal{L}. \quad (5e)$$

Problem (5) is a Linear Program (LP) that can be fast and reliably solved using standard LP solvers. Problem (5) is the basis for the development of two heuristics, namely the *Deterministic Linear Programming* (DLP) and the *Random Linear Programming* (RLP).

First we introduce the *Deterministic Linear Programming* (DLP) approach which upon solving Problem (5) selects the link i with the maximum $x_i f_i$ value, i.e. $x_i f_i = \max(x_j f_j), \forall j \in \mathcal{P}$; the procedure is terminated once K links are selected. The proposed algorithm is given as follows.

- Set $\mathcal{B} \equiv \mathcal{P}$.
- Repeat until K links are selected:
 - 1) Solve Problem (5) using standard LP solvers.
 - 2) Select link $i \in \mathcal{B}$ that has the highest value $x_i f_i$, $i = \text{argmax}_{j \in \mathcal{B}} (x_j f_j)$.
 - 3) Set $x_i = 1$.
 - 4) Remove chosen link i and the neighbouring links using the adjacency matrix from set \mathcal{P} such that $\mathcal{B} = \mathcal{P} \setminus (\{i\} \cup \{j\}_{j \in \mathcal{N}_i^\varepsilon})$.
- Return the K total chosen links and the objective function value ψ^{DLP} .

The second approach is a *Random Linear Programming* (RLP) algorithm which selects a link i in each iteration using weighted random selection. The weights are given by the normalised value of $x_i f_i, \forall i \in \mathcal{L}$ as:

$$w_i = \frac{x_i f_i}{\sum_{i \in \mathcal{L}} x_i f_i} \quad (6)$$

The RLP approach runs for a fixed number of trials Q and obtain the best solution (higher objective function value) over the total number of trials. The difference of this approach with DLP is that each link i is selected using the weights $\max(x_i f_i) \setminus \sum_{i \in \mathcal{P}} x_i f_i, \forall i \in \mathcal{P}$. The proposed algorithm is given as follows.

- For $q = 1, \dots, Q$:
 - Set $\mathcal{B} \equiv \mathcal{P}$.
 - Repeat until K links are selected:

- 1) Solve Problem (5) using standard LP solvers.
 - 2) Calculate the probability of selection given as $w_i^{(q)} = f_i x_i / \sum_{i \in \mathcal{L}} x_i f_i, \forall i \in \mathcal{B}$.
 - 3) Sample $u \sim \text{Uniform}(0, 1)$.
 - 4) Select link $i \in \mathcal{B}$ that satisfies $u < \sum_{j=1}^i w_j^{(q)}$.
 - 5) Set $x_i = 1$.
 - 6) Remove the chosen link i and the neighbouring links from set \mathcal{P} such that $\mathcal{B} = \mathcal{P} \setminus (\{i\} \cup \{j\}_{j \in \mathcal{N}_i^\varepsilon})$ using the adjacency matrix.
- Return the K total chosen links and the objective function value $\psi^{q, RLP}$.
 - Select the solution with the highest objective function value, $\psi^{RLP} = \max_{q=1, \dots, Q} \psi^{q, RLP}$.

Note that the RLP approach converges to the optimal solution with a significantly smaller number of trials compared to the RGH approach, i.e. $Q \ll S$.

V. PERFORMANCE EVALUATION

To evaluate the performance of the proposed approach we examine a real-life network, the traffic network of the Republic of Cyprus, using real-life data. The network is monitored by the Public Works Department (PWD) of the Cyprus Ministry of Transport, Communications and Works and consists of $L = 2383$ links. We define as $K = 74$ the total number of static sensors that are available for installation, while $M = 40$ sensors are already installed. In addition, traffic flow data are available for a subset of the links. However, different sources of information such as previous annual traffic census and pre-existing traffic sensors are also utilised.

To avoid the installation of sensors only to the larger cities of the network that contain by default the larger traffic flow volume, we perform a clustering procedure. The traffic network of the Republic of Cyprus is clustered into R distinct regions, with \mathcal{R} the set of all clustered $R = 6$ distinct regions, distinguishing in this way the five city centres (Nicosia, Limassol, Larnaca, Paphos, Famagusta) and the area around Troodos, the larger mountain range of the Republic of Cyprus which consists of urban centres communities and has a relatively low traffic flow volume. Then, we calculate the proportion of sensors to be installed per cluster $r \in \mathcal{R}$, to ensure that sensors are evenly spread in the network, while taking into account regions with higher traffic flow. The proportion of traffic flow volume at the links of each cluster is calculated using $t_r = \sum_{j \in \mathcal{C}_r} (100 f_j / \sum_{u \in \mathcal{L}} f_u)$, $r \in \mathcal{R}$, where $\mathcal{C}_r \subset \mathcal{L}$ is the subset of links in cluster r . Hence, the number of available sensors to be installed per cluster, K_r , is given by $K_r = K/100t_r, \forall r \in \mathcal{R}$.

A. Available Traffic data

Annual traffic census were conducted in 2006, 2007, 2008, 2010, 2012 and 2014 and the information obtained is available either directly from the Public Work Department (PWD) of the Cyprus Ministry of Transport, Communication and Works, or through the data.gov.cy website [23]. The information provided includes the Annual Average Daily Traffic (AADT), i.e. the total volume of vehicle traffic on a link for a number of days and the Passenger Car Units

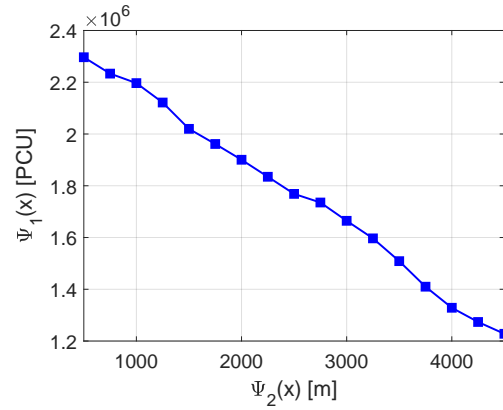


Fig. 1: Pareto front of multi-objective Problem (1).

(PCU values) per year of reference at each link. The latest traffic census covered a large proportion of the road network, obtaining traffic data from 1016 links of the road network. Data obtained from previously installed loop detectors are also available for 2015, 2019, 2021 and 2022.

While this case study was conducted, there were only $M = 40$ pre-existing installed detectors, and therefore, only 40 of the total 2383 links were covered, resulting in limited information on the traffic flows across road network, in addition to the traffic census data. The information provided from the installed sensors includes time, speed, and direction of each vehicle passing through the detector.

The proportion of traffic flow volume in each of the six clusters is 29.89%, 23.65%, 24.76%, 9.18%, 8.76%, and 3.75%. Hence, the estimated number of available sensors for each cluster was found to be 22, 18, 18, 7, 6, and 3.

B. Results

Initially we solve Problem (3) for a range of ε values that result in feasible solutions. We select $\varepsilon = \{500, 750, 1000, \dots, 4250, 4500\}$ m, and present the Pareto front of the multi-objective problem (1), given in Figure 1. As shown in the figure the relationship between the two objective functions $\Psi_1(x)$ and $\Psi_2(x)$ is almost linear. For this reason for the remainder of this section we set $\varepsilon = 3000$ m to compare the OS described in Section III with the four proposed low complexity heuristics, DGH, RGH, CGH, RLP and DLP as described in Section IV.

The optimal link selection obtained using the OS is shown in Figure 2, where blue triangles denote the selected link locations at which the new sensors will be installed, while the green squares and dark blue bullets represent two types of pre-existing traffic sensors, WIM detectors and loop detectors, respectively. The selected optimal locations are very similar for all solution algorithms and are spread around the traffic network. In Nicosia city center which is the largest city in Cyprus and has the maximum traffic flow, no links have been selected for installation due to the distance constraint as it already has a large number of existing sensors compared to the other regions. As shown, multiple sensors are placed on highways that connect the six regions.

Given the similarity of the solutions obtained, we compare objective function values obtained from each algorithm

Objective function value						
Regions	OS	DGH	RGH ($S = 1000$)	CGH	DLP	RLP ($Q = 20$)
Paphos	148373.15	135663.60	131676.91	135663.60	138751.29	148373.15
Famagusta	89859.13	89859.13	88008.87	89859.13	89859.13	89859.13
Nicosia	374416.76	313015.49	366299.18	366299.18	313015.49	367611.08
Limassol	448574.78	448574.78	430142.27	448574.78	448574.78	448574.78
Larnaca	393646.13	393646.13	377298.52	393646.13	393646.13	393646.13
Troodos	53816.27	53816.27	53816.27	53816.27	53816.27	53816.27

Relative Percentage Optimality Gap						
Regions	DGH	RGH	CGH	DLP	RLP	
Paphos	8.57	11.25	8.57	6.48	0.00	
Famagusta	0.00	2.06	0.00	0.00	0.00	
Nicosia	16.40	2.17	2.17	16.40	1.82	
Limassol	0.00	4.11	0.00	0.00	0.00	
Larnaca	0.00	4.16	0.00	0.00	0.00	
Troodos	0.00	0.00	0.00	0.00	0.00	
Average	4.91	4.075	1.38	4.71	0.45	

TABLE I: The objective function values obtained using each algorithm and the RPOG compared to the optimal objective value obtained through the OS approach.

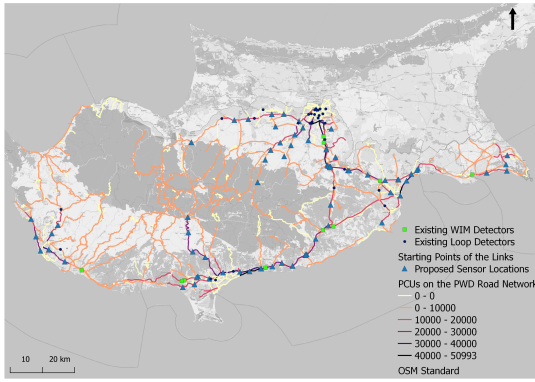


Fig. 2: Optimal sensor locations obtained by the OS approach. The proposed sensor locations are indicated by blue triangles while green squares and dark blue bullets represent pre-existing sensors.

(DGH, RGH, CGH, DLP and RLP). In addition, we calculate the Relative Percentage Optimality Gap (RPOG) given by

$$RPOG = \frac{\psi^{OS} - \psi^{alg}}{\psi^{OS}} 100\% \quad (7)$$

where ψ^{OS} is the optimal objective function value obtained from the OS approach and ψ^{alg} is the objective value obtained from each algorithm examined, i.e. $alg = \{DGH, RGH, CGH, DLP, RLP\}$. We summarise the obtained results in Table I and present the objective function and RPOG values for each solution algorithm. To obtain the results presented in Table I we set $S = 1000$ for the RGH algorithm and $Q = 20$ for the RLP algorithm.

As shown in Table I, the RLP approach resulted in the highest objective value with the smallest RPOG. The deterministic approaches, i.e. DGH and DLP, are found to behave similarly, with the DLP resulting in 0.21% higher objective value. Lastly, the RGH algorithm with 1000 trials also resulted in similar solutions as the two deterministic approaches, with the objective value being only 0.66% higher than in the DLP. As expected the CGH approach significantly reduces the high RPOG resulted from the RGH approach, as it is a combination of the two greedy heuristics

Execution Times (in seconds)						
Regions	OS	DGH	RGH	CGH	DLP	RLP
Paphos	0.0372	0.0036	0.0432	0.0468	0.0538	0.9080
Famagusta	0.0080	0.0015	0.0187	0.0202	0.0348	0.5209
Nicosia	0.0155	0.0012	0.0660	0.0672	0.2657	3.8551
Limassol	0.0075	0.0003	0.0527	0.0530	0.0766	1.3937
Larnaca	0.0111	0.0008	0.0558	0.0566	0.0964	1.9403
Troodos	0.0043	0.0002	0.0078	0.0080	0.0086	0.1421
SUM	0.0836	0.0076	0.2442	0.0252	0.5359	8.7601

TABLE II: Execution times of the different algorithms.

and selects the maximum value between the two approaches (see Section IV). It is evident that the DGH, RGH, and DLP algorithms resulted in similar results with insignificant differences in their objective values. The deterministic algorithms (DGH and DLP) either have zero or high RPOG, i.e. they either result in the optimal solution or a significantly lower objective value.

Next, we compare all solution algorithms in terms of their computational efficiency. As shown in Table I, almost all problem formulations require less than one second to obtain a solution with the RLP algorithm yielding the highest total execution time, which is still only around nine seconds. As can be seen from the results, the CGH approach yields the best trade-off between execution time and solution quality as it yields an optimality gap of less than 2% and an execution time of less than 0.2 seconds. Both random approaches result in higher objective values in comparison to the deterministic approaches. Therefore, we further investigate the impact on the number of trials. Figure 3 illustrates the RPOG values of the solutions of the two random algorithms, RLP and RGH, as the number of trials increases. For each different number of trials used, we run the algorithms 30 times. The variance of the RPOG is then taken as the range of 80% of the 30 repetitions for each number of trials.

Figure 3 shows that both approaches yield more accurate results as the number of trials is increased, i.e. the objective values obtained is much closer to the optimal objective value. However the RGH approach starts at 25% RPOG at $S = 10$ and falls to around 1% at $S = 100000$. The RLP approach starts at 1% RPOG at $Q = 10$ and falls to around 0.1% at around $Q = 60$. Both random approaches result in efficient results at low computational cost with the RLP approach

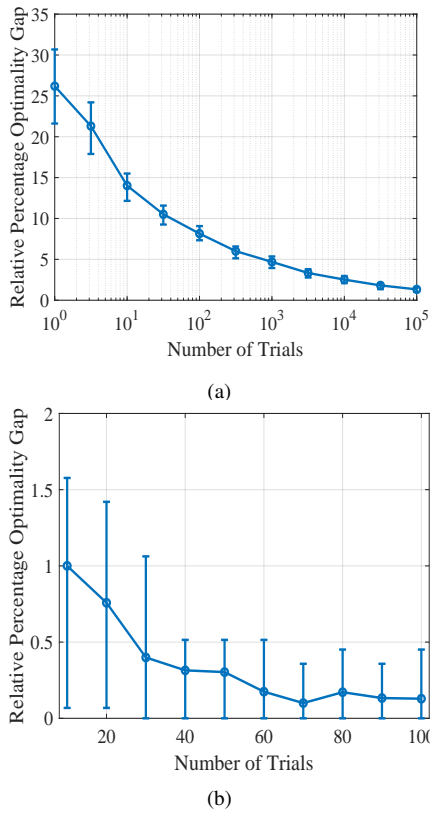


Fig. 3: The RPOG with respect to the number of trials of the (a) RGH algorithm and (b) the RLP algorithm.

converging to the optimal solution in a smaller number of trails compared to the RGH.

VI. CONCLUSIONS

In this work we have introduced a multi-objective Network Sensor Location Problem (NSLP) aiming to maximise the covered traffic flow and the minimum distance between candidate links for sensor installation. To facilitate the solution of this problem we have reformulated the problem into a single-objective combinatorial optimization which can be optimally solved using commercial MILP. To avoid the usage of commercial MILP solvers, we have proposed four low-complexity heuristics for the solution of the single-objective NSLP. We evaluate the quality of the proposed solution approach in a real-life network, the traffic network of the Republic of Cyprus. The obtained results indicate that the proposed heuristics can successfully obtain optimal sensor locations given pre-existing sensors in the problem and the use of limited data such as the traffic flow volume in PCU on links. All proposed solution approaches yield a different trade-off between execution speed and solution quality.

Future plans include the collection of data from the proposed sensor locations in the network of the Republic of Cyprus, in order to validate the efficiency of this approach. In addition, we aim to extend this approach to incorporate heterogeneous sources of data.

REFERENCES

- [1] S. Timotheou, C. Panayiotou, and M. Polycarpou, "Transportation systems: Monitoring, control, and security," *Studies in Computational Intelligence*, vol. 565, pp. 125–166, 2015.
- [2] V. Kyriacou, Y. Englezou, C. G. Panayiotou, and S. Timotheou, "Bayesian traffic state estimation using extended floating car data," *IEEE Transactions of Intelligent Transportation Systems*, vol. 24, pp. 1518 – 1532, 2022.
- [3] Y. Englezou, S. Timotheou, and C. G. Panayiotou, "Estimating the origin-destination matrix using link count observations and traffic flow dynamics," in *Proceedings of the 2021 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2022.
- [4] Z. Ma, H. Koutsopoulos, L. Ferreira, and M. Mesbah, "Estimation of trip travel time distribution using a generalized Markov chain approach," *Transportation Research Part C: Emerging Technologies*, vol. 74, pp. 1–21, 2017.
- [5] S. Carrese, E. Cipriani, U. Crisalli, A. Gemma, and L. Mannini, "Bluetooth traffic data for urban travel time forecast," *Transportation Research Procedia*, vol. 52, pp. 236–243, 2021.
- [6] W. Li, J. Wang, R. Fan, Y. Zhang, Q. Guo, C. Siddique, and Z. J. Ban, "Short-term traffic state prediction from latent structures: Accuracy vs. efficiency," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 72–90, 2020.
- [7] M. Gentili and P. B. Mirchandani, "Locating sensors on traffic networks: Models, challenges and research opportunities," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 227–255, 2012.
- [8] C. Antoniou, J. Barceló, M. Breen, M. Bulles, J. Casas, E. Cipriani, B. Ciuffo, T. Djukic, S. Hoogendoorn, V. Marzano, L. Montero, M. Nigro, J. Perarnau, V. Punzo, T. Toledo, and H. Van Lint, "Towards a generic benchmarking platform for origin-destination flows estimation/updating algorithms: Design, demonstration and validation," *Transportation Research Part C*, vol. 66, pp. 79–98, 2016.
- [9] A. Ehler, M. G. Bell, and S. Grosso, "The optimisation of traffic count locations in road networks," *Transportation Research Part B: Methodological*, vol. 40, no. 6, pp. 460–479, 2006.
- [10] S.-R. Hu and H.-T. Liou, "A generalized sensor location model for the estimation of network origin-destination matrices," *Transportation Research Part C: Emerging Technologies*, vol. 40, pp. 93–110, 2014.
- [11] T. Xing, X. Zhou, and J. Taylor, "Designing heterogeneous sensor networks for estimating and predicting path travel time dynamics: An information-theoretic modeling approach," *Transportation Research Part B: Methodological*, vol. 57, pp. 66–70, 2013.
- [12] M. Rodriguez-Vega, C. C. de Wit, and H. Fourati, "Dynamic density and flow reconstruction in large-scale urban networks using heterogeneous data sources," *Transportation Research Part C: Emerging Technologies*, vol. 137, p. 103569, 2022.
- [13] Y. Liu, N. Zhu, S. Ma, and N. Jia, "Traffic sensor location approach for flow inference," *IET Intelligent Transport Systems*, vol. 9, pp. 184–192, 2015.
- [14] E. Mitsakis, E. Chrysohoou, J. M. S. Grau, P. Iordanopoulos, and G. Aifadopoulos, "The sensor location problem: Methodological approach and application," *Transport*, vol. 32, pp. 113–119, 1 2017.
- [15] F. Viti, M. Rinaldi, F. Corman, and C. M. J. Tampère, "Assessing partial observability in network sensor location problems," *Transportation Research Part B: Methodological*, vol. 70, pp. 65–89, 2014.
- [16] E. Lovisari, C. C. de Wit, and A. Y. Kibangu, "Density/flow reconstruction via heterogeneous sources and optimal sensor placement in road networks," *Transportation Research Part C: Emerging Technologies*, vol. 69, pp. 451–476, 2016.
- [17] M. Shao, C. Xie, and L. Sun, "Optimization of network sensor location for full link flow observability considering sensor measurement error," *Transportation Research Part C: Emerging Technologies*, vol. 133, p. 103460, 2021.
- [18] F. Simonelli, V. Marzano, A. Papola, and I. Vitiello, "A network sensor location procedure accounting for o-d matrix estimate variability," *Transportation Research Part B*, vol. 46, pp. 1624–1638, 2012.
- [19] D. Kalyanmoy, *Multi-Objective Optimization using Evolutionary Algorithms*, 2001.
- [20] A. H. Land and A. G. Doig, "An automatic method of solving discrete programming problems," *Econometrica*, vol. 28, pp. 497–520, 1960.
- [21] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2022. [Online]. Available: <https://www.gurobi.com>
- [22] H. Mittelmann, "The MIPLIB2017 Benchmark Instances," 2022. [Online]. Available: <http://plato.asu.edu/ftp/milp.html>
- [23] Public Works Department, "Annual Traffic Census 2014-October 2015," 2018, accessed: 22-02-2022. [Online]. Available: <https://data.gov.cy>