Wireless Network Optimization for Massive V2I Data Collection using Multiobjective Harmony Search Heuristics

Iker Sobron[1], Borja Alonso[1], Manuel Velez[1], and Javier Del Ser[1,2,3]

[1] University of the Basque Country UPV/EHU, 48013 Bilbao, Spain,
balonso024@ikasle.ehu.eus,
{iker.sobron,manuel.velez,javier.delser}@ehu.eus
[2] TECNALIA, E-48160 Derio, Spain,
javier.delser@tecnalia.com
[3] Basque Center for Applied Mathematics (BCAM), 48009 Bilbao, Spain

Abstract. This paper proposes to improve the efficiency of the deployment of wireless network infrastructure for massive data collection from vehicles over regional areas. The increase in the devices that are carried by vehicles makes it especially interesting being able to gain access to that data. From a decisional point of view, this collection strategy requires defining a wireless Vehicular-to-Infrastructure (V2I) network that jointly optimizes the level of service and overall CAPEX/OPEX costs of its deployment. Unfortunately, it can be intuitively noted that both optimization objectives are connecting with one another: adding more equipment will certainly increase the level of service (i.e. coverage) of the network, but costs of the deployment will rise accordingly. A decision making tool blending together both objectives and inferring therefrom a set of Pareto-optimal deployments would be of utmost utility for stakeholders in their process of provisioning budgetary resources for the deployment. This work will explore the extent to which a multi-objective Harmony Search algorithm can be used to compute the aforementioned Pareto-optimal set of deployment by operating on two different optimization variables: the geographical position on which wireless receivers are to be deployed and their type, which determines not only their coverage range but also their bandwidth and cost. In particular we will utilize a non-dominated sorting strategy criterion to select the harmonies (solution vectors) evolved by Harmony Search heuristics.

Keywords: Vehicular Networks; Cost-efficient deployment; Harmony Search

1 Introduction

The latest technological advances in the design of wireless sensors have lately given rise to lower fabrication costs, which have in turn favored their widespread integration in vehicles. Indeed the automotive sector is a rich substrate for

mobility-based applications leveraging the geo-located information seamlessly collected, processed and stored by vehicles. These data, which is usually sensed locally at each vehicle, would certainly open new horizons if further processing is done to aggregate, fuse and infer knowledge from the data collected by all vehicles [1, 2].

To this end communications infrastructure must be deployed over the area to compile all the information captured by vehicles so that it can be subsequently processed, stored and managed by the applications at hand [3]. The scope of this research work is framed around the search of solutions that improve the efficiency of the Vehicular-to-Infrastructure (V2I) wireless network for the collection of massive data from vehicles over regional areas. From a decisional viewpoint this collection strategy requires defining a wireless V2I network that jointly optimizes the level of service and overall CAPEX/OPEX costs of its deployment. Unfortunately, it can be intuitively noted that both optimization objectives are conflicting with one another: adding more equipment will certainly increase the level of service (i.e. coverage) of the network, but costs of the deployment will rise accordingly. A decision making tool blending together both objectives and inferring therefrom a set of Pareto-optimal deployments would be of utmost utility for stakeholders in their process of provisioning budgetary resources for the deployment.

This work will explore to which extent a multi-objective Harmony Search (HS) algorithm can be used to compute the aforementioned Pareto-optimal set of deployment by operating on two different optimization variables: the geographical position on which wireless receivers are to be deployed and their type, which determines not only their coverage range but also their bandwidth and cost. In particular we will utilize a non-dominated sorting strategy criterion to select the harmonies (solution vectors) evolved by the Harmony Search heuristic. Experiments with the designed solver will be run over a massive case study comprising more than $7 \cdot 10^5$ vehicles deployed over the region of the Basque Country [4], as well as real target locations for the wireless receivers.

2 Vehicular and Infrastructure Databases

As a starting point we have created a geolocated vehicles database from the Basque Country region in Spain that allows simulating a realistic wireless network scenario of the entire region. For this purpose we will exploit information on the main cities such as the quantity of vehicles on each geographical location. In order to distribute the vehicles over the region, we assume that most of the traffic will be located around the most populated spots of the region. As a result we define vehicle distribution areas with a maximum radius, r_{max}, around the main populated cities in the region. The circular area of each village is divided in several concentric rings with inner and outer radius r_{int} and r_{ext}, respectively. The ratio of distributed vehicles within a m-th ring is thus given by

$$P_m = \frac{\int_{r_{int,m}}^{r_{ext,m}} e^{-cr} dr}{\int_0^{r_{max}} e^{-cr} dr}, \tag{1}$$

where r the radius and c the coefficient that modifies the gradient according to the size of the city. Note that P_m is the ratio of vehicles deployed on roads of the m-ring area with respect to its total registered in a given village. Thus, the sum of all the concentric rings around the village equals 1. The vehicles are distributed over the existing motorways, which will be classified in three categories according to the traffic. Therefore the most crowded type of motorway will have more chance to have vehicles assigned than the least crowded ones. Due to the fact that the quantity of running vehicles is variable depending on the time of the day, we can apply a correcting factor F which controls such a variability. The number of vehicles V_m to distribute in a random manner on the motorways of the m-th ring will be

$$V_m = FP_mTV, \tag{2}$$

where TV the total of vehicles recorded on each village.

It will be also necessary to create a database of wireless communications infrastructures to which the vehicles can be connected to send the data collected by their sensors. Bearing in mind that low-power wide-area systems (LPWA) such as Sigfox are focused on low-rate communications over wide areas, we assume that base stations (BSs) could be located covering wide areas. For this rationale we have built two databases of already deployed wireless networks by the Spanish operator for Internet of Things (IoT) – Cellnex Telecom – [5] and the Basque operator of TDT, Itelazpi [6]. We assume that IoT BSs can be located at the same point of the TDT repeaters since the operators may employ part of the already deployed infrastructure for wireless transmission-reception purposes.

3   Deployment Cost Profiles

For the definition of the deployment costs we consider two criteria: power and throughput of the BSs. An increase of the transmission power and the throughput corresponds to a greater range and capacity, respectively. Since we assume that infrastructure is already deployed we do not consider a location-based optimization process. The power level will be measured over the coverage radius that every transmitter reaches and the throughput will be measured on the amount of users/cars that can be simultaneously connected at the same transmitter. As a result, the cost function depends on these two parameters according to

$$C = \alpha BW + \beta \pi R^2, \tag{3}$$

where C is the cost, BW the capacity measured in number of cars/users, R the range radius and weights α and β permit to adjust the cost. We have defined the series of profiles listed in Table 1, which combine different capacities and coverages for a given transmitter. We have chosen $\alpha = 0.5$ and $\beta = 10$. The capacity values are 1000, 2000 and 5000 vehicles, and ranges are given by 5, 10 and 30 km of radius. We will also consider a profile where both values are zero, incurring no cost. This profile will emulate the case when the BS at hand is off (not activated).

Table 1: BS profiles considered in this work.

Profile type	Radius R (km)	Cars per cell (BW)	Cost (C)
1	0	0	0.00
2	5	1000	1285.40
3	5	2000	1785.40
4	5	5000	3285.40
5	10	1000	3641.59
6	10	2000	4141.59
7	10	5000	5641.59
8	30	1000	28774.33
9	30	2000	29274.33
10	30	5000	30774.33

4   Computing the Deployment Coverage and Cost

After establishing the optimization criteria we define the algorithm that cal-
culates the deployment coverage level for the users on roads according to the
combination of profiles for the deployed network. Following the previously de-
fined vehicle distribution, each profile combination – that determines the range
radius and capacity for each BS within the network – yields the total amount
of vehicles that will be covered at certain time. To compute this value we use
the flowchart depicted in Fig. 1. Using the combination of profiles as an input,
the algorithm calculates the coverage level and the corresponding cost to each
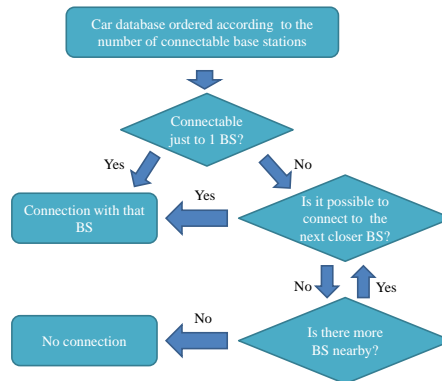specific case.

Fig. 1: Algorithmic scheme used to compute the deployment coverage ratio ac-
cording to the BS profiles.

In the region of the Basque Country 413 BSs are available as shown in Fig.
2, each capable of operating under any of the profiles defined in Table 1. Firstly,

we arrange the vehicles of the database according to the amount of BSs that it will be able to be connected to from the lowest to the highest number of BSs. Therefore, vehicles located inside the coverage area of a unique BS will be the first ones to be connected. This is due to the fact that those vehicles will be only able to connect to one BS (in this case, there is no overlapped coverage among BSs). In this manner vehicles, which can be connected to more than one BS, do not limit the access to the ones that can be only connected to a single BS. We start, car by car, with the decision criterion shown in Fig. 1. Vehicles that fall only in range of one BS will be connected if it is possible, because BSs are limited by the capacity defined in their profiles. As can be seen in Fig. 1, vehicles located within the range of two or more BSs try to connect first with their nearest BS. If this connection is not possible (due to e.g. lack of bandwidth resources of the selected profile for the BS at hand), they will attempt at connecting to their second closest BS. If the vehicle cannot connect because BSs have reached their maximum number of vehicles, the vehicle has no coverage. When the algorithm ends, we shall compute the deployment coverage level and the overall cost of the deployed profiles.
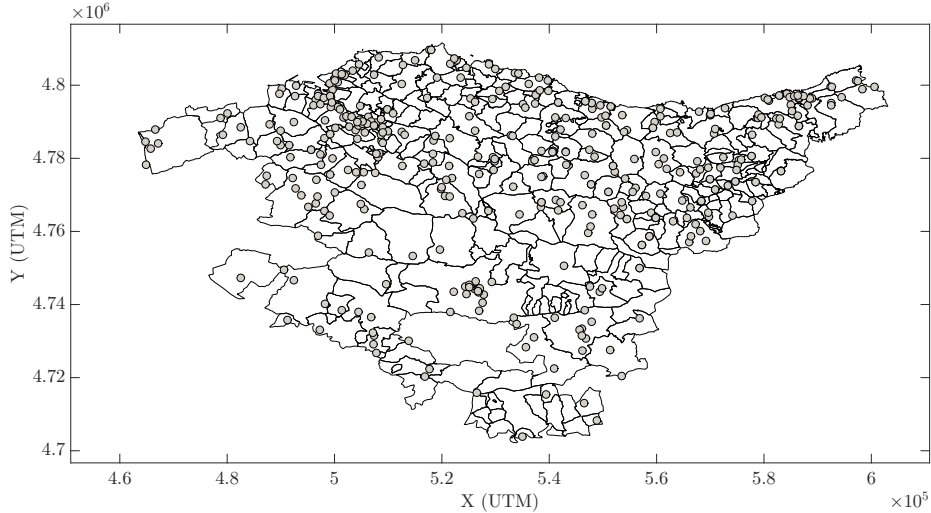

Fig. 2: Geo-location of the deployed BSs in the Basque Country in Spain.

5 Review of Multi-objective Harmony Search Algorithm

The bi-objective problem addressed in this paper – i.e. the simultaneous maximization of the coverage ratio of the deployed receivers and the minimization of their cost – will be tackled by resorting to a multi-objective version of the HS algorithm, a heuristic solver proposed in [7] that emulates the music composition process observed in jazz bands when their members jointly improvise harmonies driven by an aesthetic measure of musical quality. In essence HS is a

population-based optimization algorithm that resembles other techniques from Swarm Intelligence and Evolutionary Computation, with subtle yet relevant differences in the operators that iteratively refine the solutions contained in the K-sized population of the algorithm (commonly referred to as Harmony Memory). Such operators iteratively enhance the candidate solutions or harmonies by permuting and mutating intelligently their constituent variables or notes until a convergence criterion set beforehand is met (e.g. a maximum number of iterations $\mathcal{I}$ are completed). The outperforming behavior of HS has been evinced in several application scenarios [8], with scarce albeit interesting practices regarding the optimization of cost-efficient deployments [9–11].

The HS solver proposed in this work utilizes three improvisation operators that are applied to the harmonies within the harmony memory at every iteration:

– *Harmony Memory Considering Rate* HMCR $\in [0,1]$, which establishes the probability that the improvised value for a note is drawn from the values taken by the same note in the other $K-1$ harmonies in the harmony memory.
– *Pitch Adjusting Rate* PAR $\in [0,1]$, which tunes the probability that the value of a note is replaced with any of its neighboring values in its alphabet. In order to establish a proper relation of vicinity BS profiles are sorted in increasing order of their bandwidth, in such a way that they are equally ordered in terms of cost.
– *Random Selection Rate* RSR $\in [0,1]$, which is similar to PAR: it sets the probability that the new value for a given note is taken uniformly at random (i.e. without any neighborhood consideration) from its alphabet.

The above operators are applied to each note of the candidate solutions, after which the fitness functions of the newly improvised solutions (namely, coverage ratio and deployment cost) are evaluated. Based on the values of the objective functions for both the new harmonies and those from the previous iteration a Pareto dominance ranking and crowding distance criterion is applied to filter out the prevailing population for the next iteration. In short each solution is ranked according to its dominance level (namely, `rank` $= 1$ for the best, non-dominated subset of solutions, `rank` $= 2$ for the second best front, etc). After ranking all solutions the sum of distances to the closest harmony along each metric (crowding distance) permits to prioritize solutions within a certain dominance level: those solutions featuring large crowding distances are preferential.

6   Results

In Fig. 3, we can observe the Pareto efficiency curve achieved by HS for the defined profile combinations. The HS algorithm has been configured with $K = 30$, HMCR $= 0.5$, PAR $= 0.15$, RSR $= 0.1$, $\mathcal{I} = 200$ iterations. The overall cost has been normalized by the most expensive scenario: the overall number of transmitters at the most expensive cost (i.e. 1.2710e+07). The amount of 10^6 vehicles have been distributed according to the radial distribution in (1) using $c = 0.01$ and $r_{max} = 25km$ around the 30 most populated villages in the region.

The temporal factor F has been chosen as 0.7 in the rush hour. We have also distributed in a uniform random manner 10% of the running vehicles on roads out of the 30 circular areas.

Solutions A, B and C have been highlighted in Fig. 3 with 205, 150 and 70 transmitters switched off out of the overall 413, respectively. One can observe that an acceptable 55% of coverage level can be achieved with the minimum Pareto efficiency value (208 transmitters on). This result arises from the effect of a radial traffic distribution model centered in cities, most of cars will be distributed close to the city where a higher number of transmitters are usually deployed. As a result, those BSs can provide coverage to more amount of users compared with the rural BSs which are deployed far away from urban areas.
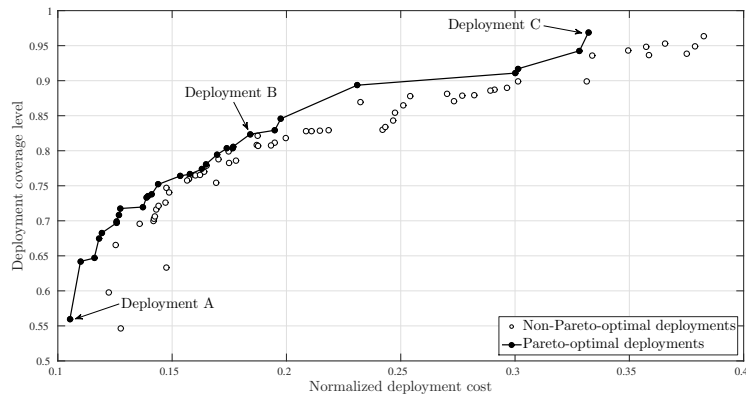

Fig. 3: Pareto efficiency curve of V2I deployment coverage level for the Basque region in Spain.

In Fig. 4 the deployment profiles of the marked scenarios are depicted. In Fig. 4.a it can be observed the high amount of BS in low profiles (type 1 (205 BSs) and 2) and that a few BSs are set with profiles with high cost (type 9 an 10). This deployment provides a coverage level of 55%. In Fig. 4.b we show an intermediate point where the trade-off between cost and coverage level is interesting: 82% of users on roads are covered at rush hour with a reduced cost since 150 transmitters can be off and rest of profiles are mainly low-cost. Finally, we can see the C deployment in Fig. 4.c where the highest coverage level can be achieved. One can notice that deployments at the rightmost part of the Pareto curve with higher cost can not achieve that coverage level. These scenarios have shown the importance of deployment analysis for operator in terms of cost and coverage since adequate deployments with lower cost can obtain relevant results in terms of coverage.

7   Conclusions

This paper have presented a multi-objective HS-based strategy to improve the usage of already deployed wireless network infrastructure for massive data col-
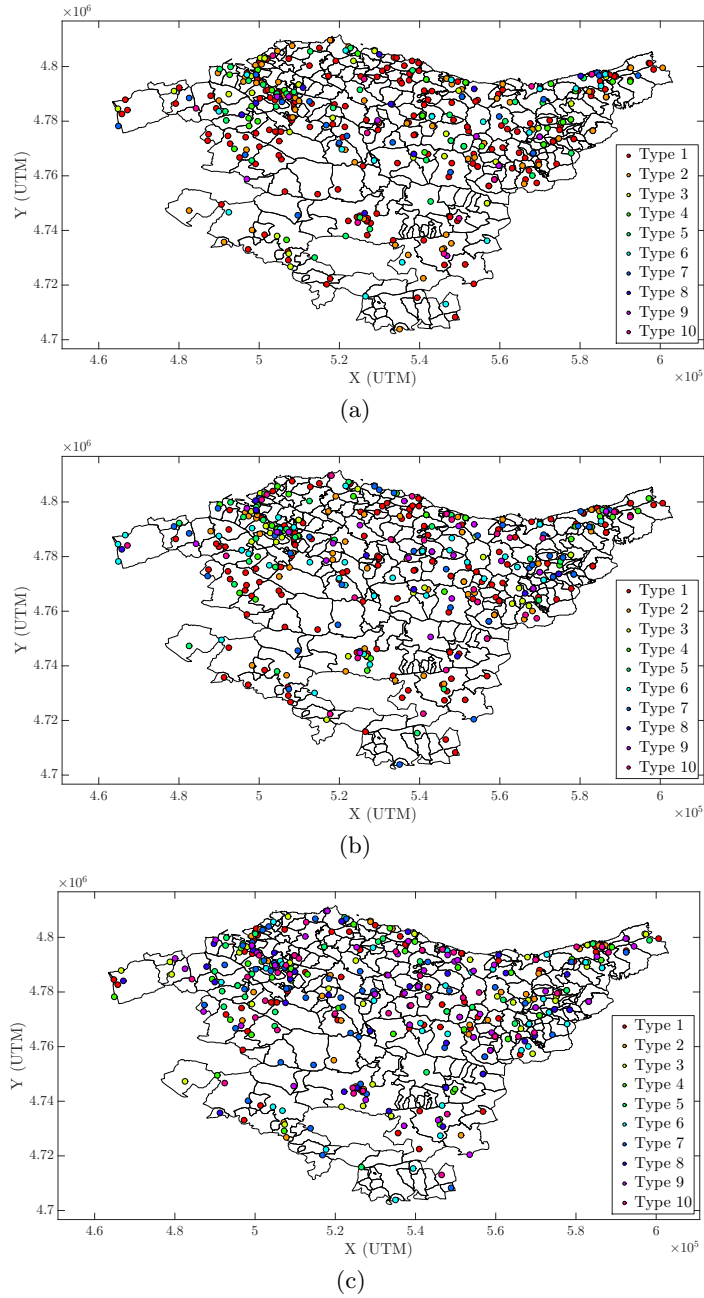
(a)

(b)

(c)

Fig. 4: Deployment profiles of the A, B and C Pareto solutions.

lection from vehicles over regional areas. This HS strategy is based on a non-dominated sorting criterion in terms of two conflicting fitness metrics: the level of service (number of connected cars) and maintenance cost of the network. Realis-

tic results have been obtained by leveraging open data from the Basque Country region in Spain. The set of Pareto-optimal solutions have demonstrated the utility of this tool for local operators in their process of providing cost-efficient resources for the deployment of a wireless network aimed at collecting data from vehicles. An important reduction of the budget can be achieved by carefully selecting the operation profiles of the already deployed wireless infrastructure.

8   Acknowledgments

References

1. Papadimitratos, P., Fortelle, A. D. L. Evenssen, K., Brignolo, R. Cosenza, S.: Vehicular Communication Systems: Enabling Technologies, Applications, and Future Outlook on Intelligent Transportation. IEEE Communications Magazine 47(11): 84-95 (2009)
2. Reininger, M., Miller, S., Zhuang, Y. Cappos, J.: A First Look at Vehicle Data Collection via Smartphone Sensors. IEEE Sensors Applications Symposium (SAS), 1-6 (2015)
3. Paulin, T., Ruehrup, S., Fuxjaeger, P., Paier, A.: Improving V2I Edge Communication by Performance Maps. IFIP Wireless and Mobile Networking Conference (WMNC) 200-207 (2015)
4. OTEUS: Basque Transport Observatory. http://www.garraioak.ejgv.euskadi.eus/r41-4833/es/
5. Cellnex Telecom. http://www.cellnextelecom.com/
6. Itelazpi: Basque Public Telecommunications Company. http://www.itelazpi.eus/
7. Geem, Z. W., Kim, J.-H., Loganathan, G. V.: A New Heuristic Optimization Algorithm: Harmony Search. Simulation 76(2), 60–68 (2001)
8. Manjarres, D., Landa-Torres, I., Gil-Lopez, S., Del Ser, J., Bilbao, M. N., Salcedo-Sanz, S., Geem, Z. W.: A Survey on Applications of the Harmony Search Algorithm. Engineering Applications of Artificial Intelligence 26(8), 1818–1831 (2013)
9. Bilbao, M. N., Del Ser, J., Salcedo-Sanz, S., Gil-Lopez, S. Portilla-Figueras, J. A.: A Bi-objective Harmony Search Approach for Deploying Cost-Effective Multi-hop Communications over Large-Area Wildfires. International Joint Conference. SOCO'14-CISIS'14-ICEUTE'14, 93-103 (2014)
10. Bilbao, M. N., Gil-Lopez, S., Del Ser, J., Salcedo-Sanz, S., Sanchez-Ponte, M., Arana-Castro, A.: Novel Hybrid Heuristics for an Extension of the Dynamic Relay Deployment Problem over Disaster Areas. TOP 22:997-1016 (2014)
11. Bilbao, M. N., Del Ser, J., Salcedo-Sanz, S., Casanova-Mateo, C.: On the Application of Multi-Objective Harmony Search Heuristics to the Predictive Deployment of Firefighting Aircrafts: a Realistic Case Study. International Journal of Bio-Inspired Computation (IJBIC) 7(5):270-284 (2015)