Nature-Inspired Metaheuristics for optimizing Information Dissemination in Vehicular Networks

Antonio D. Masegosa
University of Deusto, Bilbao, Spain
Ikerbasque, Basque Foundation for Science, Bilbao, Spain
ad.masegosa@deusto.es

Eneko Osaba
TECNALIA Research and Innovation, Derio, Spain
eneko.osaba@tecnalia.com

Juan S. Angarita-Zapata
University of Deusto, Bilbao, Spain
js.angarita@deusto.es

Ibai Laña
TECNALIA Research and Innovation, Derio, Spain
ibai.lana@tecnalia.com

Javier Del Ser
University of the Basque Country (UPV/EHU), Bilbao, Spain
javier.delser@tecnalia.com

ABSTRACT

Connected vehicles are revolutionizing the way in which transport and mobility are conceived. The main technology behind is the so-called Vehicular Ad-Hoc Networks (VANETs), which are communication networks that connect vehicles and facilitate various services. Usually, these services require a centralized architecture where the main server collects and disseminates information from/to vehicles. In this paper, we focus on improving the downlink information dissemination in VANETs with this centralized architecture. With this aim, we model the problem as a Vertex Covering optimization problem and we propose four new nature-inspired methods to solve it: Bat Algorithm (BA), Firefly Algorithm (FA), Particle Swarm Optimization (PSO), and Cuckoo Search (CS). The new methods are tested over data from a real scenario. Results show that these metaheuristics, especially BA, FA and PSO, can be considered as powerful solvers for this optimization problem.

CCS CONCEPTS

- Applied computing → Transportation;
- Theory of computation → Random search heuristics; Evolutionary algorithms; Packing and covering problems;

KEYWORDS

Nature-inspired Metaheuristics, Vertex Covering, Vehicular Communications, VANETs, Intelligent Transportation Systems

ACM Reference Format:

Antonio D. Masegosa, Eneko Osaba, Juan S. Angarita-Zapata, Ibai Laña, and Javier Del Ser. 2019. Nature-Inspired Metaheuristics for optimizing Information Dissemination in Vehicular Networks. In Genetic and Evolutionary Computation Conference Companion (GECCO ’19 Companion), July 13–17, 2019, Prague, Czech Republic. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3319619.3326847

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

GECCO ’19 Companion, July 13–17, 2019, Prague, Czech Republic © 2019 Association for Computing Machinery. ACM ISBN 978-1-4503-6748-6/19/07...$15.00 https://doi.org/10.1145/3319619.3326847

1 INTRODUCTION

Transport and mobility are immersed in a major transformation nowadays thanks to the confluence of several innovations: autonomous driving, shared mobility, electric vehicle, and connected vehicle. Regarding this last aspect, the main technology behind is the so-called Vehicular Ad-Hoc Networks (VANETs), which are communication networks in which the nodes are vehicles [21]. Depending on which element is the transmitter and the receptor in each extreme of the communication, either Infrastructure (e.g Road Side Units, a central server, etc.) or a Vehicle, VANETs can be categorized into Infrastructure-to-Infrastructure (I2I), Infrastructure-to-Vehicle (I2V), Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V). These communication configurations can enable different services related to security, leisure and entertainment, traffic management and driving assistance [33].

Many of the aforementioned services require a centralized architecture in which a central server collects and disseminates data from/to vehicles [5] (e.g. to acquire information about the instant position, speed and heading of all vehicles on a region or to warn vehicles closed to a potentially dangerous situation). Within this scenario, in this paper, we focus on the so-called downlink communication in which a central server disseminates information to vehicles. Different aspects make this task challenging in VANETs [21]. One of the most promising approaches proposed so far to address this challenge is the use of Virtual Infrastructure (VI) [11, 38], wherein vehicles are also used as infrastructure to increase the area of application and to reduce the deployment cost. The VI is composed of three phases: 1) some vehicles are selected as cluster-head (CH) to cover the broadcast area, 2) CH vehicles receive information from a central server through I2V, and 3) CH vehicles disseminate this information to nearby vehicles using V2V communications.

The process of deciding how many and what vehicles should be used as VI in each instant plays a pivotal role in this type of systems, since it can avoid the necessity of fixed infrastructure (as Road Side Units), reduce the network overload, or affect the communication quality [38]. Many of the services envisioned for VANETs require the dissemination of some messages to all or as much as possible vehicles in a specific geographical area. In this context, the priority of the VI selection phase must be to choose the minimum number of CH vehicles that can broadcast the information to the maximum
number of vehicles in the target area, taking into account all the requirements imposed by the corresponding standard [5].

Based on the communication architecture for VANETs presented in [11] and called NAVI (Neighbor-Aware Virtual Infrastructure), in a previous work, Masegosa et al. [23] proposed a new approach to address the VI selection process, modelling it as the optimization of a Covering Location Problem (CLP) [12, 14]. Concretely, they modelled those vehicles that must receive the information as demand nodes, and those vehicles that play the role of VI as potential facility locations. Using a Genetic Algorithm (GA, [13]) as optimization method, they were able to outperform an ad-hoc state-of-the-art methodology designed for the same purpose.

In this paper, our objective is to continue deepening the research line proposed in [23] through the development of better optimization methods based on nature-inspired metaheuristics. Specifically, we propose and compare four nature-inspired metaheuristics to improve the performance of this approach for information dissemination in VANETs. The algorithms considered for such purpose are: Bat Algorithm (BA, [43]), Firefly Algorithm (FA, [42]), Particle Swarm Optimization (PSO, [18]), and Cuckoo Search (CS, [44]). The main rationale behind the use of these nature-inspired metaheuristics relies on their reputed efficiency for solving covering location problems [9, 15, 35, 39] and other combinatorial optimization problems [26, 28, 30, 31] along recent years. The good acceptance of these previous studies has led us to the hypothesis that these methods may be promising also for the problem approached in this work. Furthermore, we also use a new modelling framework, shifting from a CLP to a Vertex Cover Problem (VCP), which is more appropriate to the features of the underlying optimization problem. To assess and compare the performance of the new methods and the new modelling approach, we use a real scenario consisting of 45 vehicles moving around in the downtown area of the city of Malaga in Spain. Apart from this, the proposed optimization algorithms are compared against two baseline algorithms, the GA proposed in [23] and the state-of-the-art approach for information dissemination presented in [11].

The rest of this paper is structured as follows. Section 2 gives background about information dissemination in VANETs and VCPs. The new modelling approach is presented and described in Section 3. Then, Section 5 is devoted to detail the experimental framework used to test our proposal. After that, in Section 6 we analyze the results of the comparison of the proposed metaheuristics and the two baseline methods. Finally, Section 7 presents the main conclusions of the work as well as the future research lines.

2 RELATED WORK

2.1 Information Dissemination in VANETs

Due to the interest of deploying cooperative services and applications for vehicles, information dissemination in VANETs has been extensively researched [5, 32]. The aim of a VANET is to be an infrastructure-less self-organizing traffic information system. Therefore, the vast majority of proposed methods for information dissemination are based on a decentralized architecture wherein the organization of the network is managed by vehicles creating dynamic clusters of vehicles, which is enabled by short-range communication technologies.

The creation and management of the cluster imply the periodic communication of status information [16], which in high density scenarios can lead to the exchange of a high number of messages to organize the network, an in turn to the consumption of many resources for network management. Other approaches try to reduce this overhead deploying infrastructure nodes at preferential locations, but they usually lack flexibility and the dissemination covered region is fixed [32]. Few works in the literature propose to use vehicles as mobile infrastructure. Camara et al. [8] present the virtual RSU (vRSU) concept where nodes receive and cache messages from other vRSU, or access points that are located in areas with no coverage from conventional RSUs.

Recently, heterogeneous architectures have been proposed to exploit both the wide range low latency communications of cellular technologies - Long-Term Evolution (LTE) communication (commercially known as 4G network) - and the low cost of IEEE 802.11p - similar to common WiFi networks. Some works use the heterogeneous architecture to improve the efficiency of clustering [11, 36]. In [36], Remy et al. use a heterogeneous centralized architecture to reduce the clustering overhead. An interesting proposal can be found in [11] where D’Orey et al. presented the approach Neighbor-Aware Virtual Infrastructure (NAVI), for information dissemination using vehicles as mobile Virtual Infrastructure (VI) that were selected from a central entity called GeoServer. NAVI presented many advantages as an appropriate uplink performance, reduced use of fixed infrastructure and better use of individual technologies. More recently, in [23], A.D. Masegosa et al. using this VANET architecture, they modelled the selection of vehicles used as VI by means of a CLP whose objective consists on maximizing the covering (vehicles that receive the message), while minimizing the number of vehicles used as VI. They also proposed a Genetic Algorithm to solve the optimization problem, and they show that this new methodology outperformed the VI selection used in [11].

As mentioned in the introduction, in this paper, we aim at extending the research done in [23]. Therefore, we will use the same VANET architecture based on NAVI that will be further described in Section 3.1. However, we will model the selection of the vehicles used as VI, as a VCP instead of a CLP. In the next section, we define CLPs formally and review some literature about them.

2.2 Vertex Cover Problems

Given an undirected graph $G = (V, E)$, the Vertex Cover Problem (VCP) [7] consists on finding a subset $C \subseteq V$ such that each edge in $E$ has at least one end point in $C$, and the cardinality of $C$ is limited by a constant $k (|C| < k)$. Some applications of the VCP can be found in dynamic detection of potential data races in multi-threaded programs [25], network monitoring [4], or telecommunications [24].

In the next part of this section, we describe and formulate two of the most important generalizations of the VCP: the Minimum Vertex Cover and the Maximum Partial Vertex Cover.

**The Minimum Vertex Cover Problem (MVCp).** The MVCp is an extension of the VCP and consists of finding the minimum subset $C$ that covers all the edges in $E$ [19]. It is formulated as:

- $V$ - set of vertices in the undirected graph $G$
- $E$ - set of edges in the undirected graph $G (E \subseteq V \times V)$. An edge between vertices $i$ and $j$ is denoted as the duple $(i,j) \in E$
Nature-Inspired Metaheuristics for Information Dissemination.

- $C$ - the cover set, $C \subseteq V$
- $N_i$ - $\{ j \in V | (i, j) \in E \}$, the set of vertices adjacent to $i$.
- $x_j$ - boolean variable set to 1 if a vertex $j \in C$, 0 otherwise.

$$\text{Minimize: } Z = \sum_{j \in V} x_j \quad (1)$$

$$\text{Subject to: }$$
$$\sum_{j \in N_i} x_j \geq 1 \quad \forall i \in V \quad (2)$$
$$x_j = \{0, 1\} \quad \forall j \in V \quad (3)$$

The objective function given in Equation 1 minimizes the number of vertices in the cover set $C$. Equation 2 ensures that each vertex $i$ is covered by at least one vertex in $C$, and finally, Equation 3 enforces the binary restriction on the decision variable.

The Weighted Maximum Partial Vertex Cover (WMVC). Unlike the MVC, in the WMVC, the cardinality of the cover set $C$ is bounded at a value equal or lower than $p$ [1]. The objective of the MVC consists on finding a subset $C$, such that $|C| \leq p$, which maximizes the number of edges covered, that is, the number of edges in $G$ with at least one endpoint in $C$. The mathematical formulation of the MVC is as follows:

- $V$ - set of vertices in the undirected graph $G$
- $E$ - set of edges in the undirected graph $G$ ($E \subseteq V \times V$). An edge between vertices $i$ and $j$ is denoted as the duple $(i, j) \in E$
- $w_i$ - weight associated to vertex $i$
- $C$ - the cover set, $C \subseteq V$
- $N_i$ - $\{ j \in V | (i, j) \in E \}$, the set of vertices adjacent to $i$.
- $x_j$ - boolean variable set to 1 if a vertex $j \in C$, 0 otherwise.
- $y_i$ - represent the coverage of vertex $i$. Its value is 1 if vertex $i$ is covered ($\exists j \in V | x_j = 1 \land j \in N_i$), and 0 otherwise.
- $p$ - the maximum allowed cardinality for the cover set $C$.

$$\text{Maximize } Z = \sum_{i \in V} w_i y_i \quad (4)$$

$$\text{Subject to: }$$
$$\sum_{j \in N_i} x_j \geq y_i \quad \forall i \in V \quad (5)$$
$$\sum_{j \in V} x_j \leq p \quad (6)$$
$$x_j = \{0, 1\} \quad \forall j \in V, \quad y_i = \{0, 1\} \quad \forall i \in V \quad (7)$$

Equation 4 refers to the objective function that maximizes the weighted sum of the vertex covered. Equation 5 is analogous to Equation 2. Equation 6 limits the cardinality of $C$ to $p$. Finally, Equation 7 restricts variables $x_i$ and $y_i$ to binary values.

3 APPROACH FOR INFORMATION DISSEMINATION IN VANETS

This section is devoted to describing the approach for information dissemination used in this paper. As mentioned above, this approach was first proposed in [23] that in turn was based on NAVI’s network architecture [11], but with a procedure for selecting vehicles as VI based on the optimization of a CLP through a Genetic Algorithm. As we also pointed out in Section 1 and explained in Section 2, in this paper we reformulate the optimization model as a VCP instead of a CLP because it is more consistent with the real underlying optimization problem. The main reason behind this shift is the nature of the communications that do not depend only on a distant radius but in more complex factors. Given that VANETs can be seen as a graph where vertices correspond to vehicles and edges link vehicles that can communicate between each other, VCP fits better in this scenarios than CLP, although they are analogous problems. To facilitate the understanding of modelling process used, we first introduce NAVI’s network architecture; then, we give the formal description of the VCP-based model for selecting the vehicles used as VI; and finally, the novel nature-inspired metaheuristics presented in this work to solve the optimization problem are presented.

3.1 NAVI’s Network Architecture

For the sake of simplicity and due to limited space, in this section, we will only give a quick overview of the NAVI’s network architecture. The interested reader may refer to [23] or [11] for further details. Having said this, NAVI’s architecture was designed for an efficient collection and dissemination of information in VANETs through the selection of mobile infrastructure nodes (vehicles) in a scenario with multiple technologies [11]. Figure 1 shows a scheme of the general architecture of the system. The system comprises of a heterogeneous network architecture consisting of short-range communication networks, as IEEE 802.11p, and long-range communication networks, as LTE. Vehicles can be categorized into three main classes in terms of network equipment: i) short-range communication only, ii) long-range communication only and iii) short and long-range communication. The working of the system can be divided into three main stages: a) data collection, b) virtual infrastructure selection and c) data dissemination strategy execution.

In the data collection phase, the periodic broadcast single-hop CAMs are used to establish when two vehicles can communicate between them. Concretely, if a vehicle A receives a CAM message from another vehicle B, it means that they can transfer information among them. In this case, we establish that B is a neighbour of A. Each vehicle has its own Neighbourhood Table (NT) with an entry for each neighbour. The NTs are aggregated and transmitted periodically to the Geoserver by a subset of nodes in the Region of
Interest (ROI). The VI selection phase updates the VI periodically from the information provided by the aggregated NTs, with a prefixed period length $t$. The selection of the VI is made centrally by GeoServer. The GeoServer launches this process each time it receives a dissemination request from the service provider. In the last phase, data dissemination strategy execution, the vehicles selected in the previous step perform the corresponding action: broadcast, relaying, and store-and-forward.

### 3.2 Virtual Infrastructure Selection Model based on Vertex Covering

In this section, we describe the reformulation of the model proposed for the VI selection process. Its objective consists of the efficient transfer of information from a central entity to the vehicles using a proper VI configuration. Concretely, it aims to balance the network infrastructure resources and move on a given scenario considering a proper VI configuration. Concretely, it aims to balance the networks that usually have a pay-per-use system.

The formulation used to model this optimization process is based on the VCP. The model can be divided into three main components: vehicles, scenario and network systems. Vehicles have access to network infrastructure resources and move on a given scenario $S$, which is partitioned in zones $s$ such that $s \in S$. The properties of the scenario impact vehicle mobility and the communication reliability between vehicles, but we assume that during the period between the creation of the NTs and the dissemination of information, they remain unaltered. Taking into account that the length of this period is usually lower than 200ms, this assumption can be considered realistic. The vehicles located in the scenario at the instant at which a dissemination request is received by the GeoServer it is represented by the set $V$. During the observation period, additional vehicles may join or leave the scenario depending on demand or routes, but as mentioned before, given the short length of the period, these changes are irrelevant. The objective of the underlying optimization process is to find the minimum subset $C \subseteq V$ that maximizes the number of zones covered in the scenario $S$. More formally, the model is defined as follows:

- $S$ - set of zones to cover in the region. Each zone $s \in S$ has at least one vehicle located in it.
- $V$ - set of vehicles that must be covered.
- $V^{VI}$ - subset of vehicles from $V$ ($V^{VI} \subseteq V$) that can be potentially used as VI, that is, those equipped with short and long range communication capabilities.
- $CAM_{ij}$ - boolean variable whose value is 1 if vehicle $i \in V$ has received a CAM message from vehicle $j \in V$.
- $NV_{i}$ - $\{ j \in V^{VI} | CAM_{ij} = 1 \}$, the set of neighbours of vehicle $i \in V$ (adjacent vertices).
- $Z_{ij}$ - zone $s \in S$ in which vehicle $i \in V$ is located.
- $V_{s}$ - $\{ i \in V | Z_{ij} = s \}$, set of vehicles located in zone $s \in S$.
- $N_{s}$ - $\{ i \in V^{VI} | (i \in V_{s}) \lor (3j \in V_{s} \text{ s.t. } j \in NV_{i}) \}$, the set of vehicles that can cover zone $s$ (those vehicle from $V^{VI}$ that can communicate with at least one vehicle in zone $s$).
- $x_{j}$ - boolean variable set to 1 if vehicle $j \in V^{VI}$ is selected as VI, 0 otherwise.
- $y_{s}$ - represent the coverage of zone $s \in S$. Its value is 1 if zone $s$ is covered and 0 otherwise.
- $p$ - maximum number of vehicles that can be used as VI.

**Objective Function**

$$Maximize \quad Z = (1 - F_{1}) \cdot F_{2} \quad (8)$$

**Expression $F_{1}$**

$$F_{1} = \frac{1}{|V^{VI}|} \sum_{j \in V^{VI}} x_{j} \quad (9)$$

**Expression $F_{2}$**

$$F_{2} = \frac{1}{|S|} \sum_{s \in S} y_{s} + \frac{1}{|S||V|} (-1 + \sum_{s \in S} y_{s}|V^{s}|) \quad (10)$$

**Subject to:**

$$\sum_{j \in N_{s}} x_{j} \geq y_{s} \quad \forall s \in S \quad (11)$$

$$\sum_{j \in V^{VI}} x_{j} \leq p \quad (12)$$

Equation 8 corresponds to the objective function which combines the two objectives, maximizing the zones covered while minimizing the number of vehicles used as VI. Equation 9 defines the proportion of vehicles considered as VI. In Equation 10, the first term of the expression corresponds to the ratio of zones covered, whereas the second one to the proportion of vehicles covered divided by $|S|$. Note that the values of this second term range in the interval $[0, \frac{1}{|S|}]$.

In this way, a solution is considered better if it covers more zones or at the same number of zones if it covers more vehicles. The constraint formulated in Equation 11 establishes that a zone $s$ is covered only when at least one vehicle in the VI has, as neighbour, one of the vehicles located at $s$. Equation 12 ensures that the number of vehicles used as VI is at most $p$. Finally, Equation 13 restricts variables $x_{j}$ and $y_{s}$ to binary values.

To avoid dealing with feasible and infeasible solutions, we used the penalization scheme show in Equation 14, where $p' = \sum_{j \in V^{VI}} x_{j}$ and $c$ is a parameter weights the magnitude of the penalization.

$$Z(x) = \begin{cases} (1 - F_{1}(x)) \cdot F_{2}(x) & \text{if } p' \leq p \\ (1 - F_{1}(x)) \cdot F_{2}(x) - c(p' - p) & \text{otherwise} \end{cases} \quad (14)$$

### 4 PROPOSED NATURE-INSPIRED SOLVERS

In this section, we describe the four nature-inspired metaheuristics that we have developed in order to solve the optimization problem detailed in 3.2, and therefore, to select the vehicles used as VI. Before the description of each deemed method, some important design issues are detailed in what follows, related to the codification strategy, the metric used for measuring the difference between two individuals, and the modelled movements operators.

Being one of the crucial aspects in the heuristic development, it is interesting to highlight that the binary representation has been adopted as encoding strategy. Thus, each potential solution is represented as a vector $x = [x_{1}, x_{2}, \ldots, x_{n}]$ of $n$ binary values $\{0, 1\}$, where $n$ depicts the cardinality of the subset $V^{VI}$, that is, $n = |V^{VI}|$ and it refers to the number of vehicles that can be potentially used as VI. Additionally, $x_{j}$ is analogous to the same symbol defined in...
where \( \text{VI} \), or not (e.g. \( x_i = 1 \) if vehicle \( j \in \mathcal{V}_I \) is selected as VI, 0 otherwise).

Given that we are dealing with boolean problems, we introduced an appropriate distance metric and two movement operators. Regarding the distance metric employed for assessing the similarity between two solutions, the well-known Hamming Distance (\( H(\cdot, \cdot) \)) has been used. Several previous studies in the recent literature have utilized this same function for similar purposes [27], verifying its adequacy for solving combinatorial optimization problems, such as the one addressed in the present work. Specifically, \( H(\cdot, \cdot) \) is computed as the number of non-corresponding elements between two solutions. For instance, and assuming two individuals \( x = [1, 0, 1, 0, 1, 1, 1, 0, 1] \) and \( x' = [1, 0, 1, 1, 1, 1, 0, 0, 0] \), their Hamming Distance \( D_H(x, x') \) would equal 3. It should be highlighted that this distance metric is the basis of the movement strategies inherent to each of the proposed techniques.

Lastly, two movements operators have been considered for evolving the solutions along the search process. The first one is the commonly used Uniform Crossover (\( \text{UC}_p(x, y) \)) [40], in which two different individuals share their genes probabilistically for producing a new hybrid solution. For this study, a probability \( p=50\% \) has been used for each gene, which means that every gene \( c_i \) has the same probability of being inherited by \( x \) or \( y \). The second operator used is the swapping \( \text{SW}(x, c) \) operator, in which \( c \) genes are randomly chosen from solution \( x \) for reversing their value, from 1 to 0, or from 0 to 1. In this specific research, \( c=2 \) value has been fixed.

Below we defined the four nature-inspired metaheuristics proposed for the optimization problem described in Section 3.2.

**Bat Algorithm (BA).** The BA, introduced by Xin-She Yang in [43], is a nature-inspired method based on the echolocation behaviour of microbats, which can find their prey and discriminate different kinds of insects even in complete darkness. Despite it was firstly proposed for solving continuous optimization problems, it has been adapted several times before [30, 37, 46] for solving discrete problems. Following the same philosophy, an adaptation has been developed several times before [30, 37, 46] for solving discrete problems. Following the same philosophy, an adaptation has been introduced by Yang and Deb in 2009 [44] for solving continuous optimization problems. One of its main advantages is its easy parameterization, which along with its efficiency has lead the CS to have a great success recently [3, 20, 22]. In this work, the well-known adaptation proposed in [31] for the Traveling Salesman Problem has been embraced as a base, using similar mechanisms and parameters. Specifically, for the cuckoo movement, the same \( \text{UC}_p(x, y) \) and \( \text{SW}(x, c) \) operators have been employed. Additionally, the movement of each cuckoo is conducted using the same logic depicted in Expression (15), considering \( D_H(x, \cdot) \) as similarity function, and best individual as reference.

**Cuckoo Search (CS).** The CS was introduced by Yang and Deb in 2009 [44] for solving continuous optimization problems. One of its main advantages is its easy parameterization, which along with its efficiency has lead the CS to have a great success recently [3, 20, 22]. In this work, the well-known adaptation proposed in [31] for the Traveling Salesman Problem has been embraced as a base, using similar mechanisms and parameters. Specifically, for the cuckoo movement, the same \( \text{UC}_p(x, y) \) and \( \text{SW}(x, c) \) operators have been employed. Additionally, the movement of each cuckoo is conducted using the same logic depicted in Expression (15), considering \( D_H(x, \cdot) \) as similarity function, and best individual as reference.

**Particle Swarm Optimization (PSO).** PSO has been adapted to discrete problems multiple times before [6, 45]. Taking as a base some of these interesting works, each particle in our developed PSO also represents a possible solution for the problem at hand. Additionally, velocity parameter \( v_i \) has been considered in the same way as done in BA case. Furthermore, both inclination mechanism and \( \text{UC}_p(x, y) \), \( \text{SW}(x, c) \) movement strategies have also been deemed for the PSO, as well as the movement criterion shown in Expression (15). Finally, \( D_H(x, \cdot) \) has been used as similarity measurement.

5 EXPERIMENTAL FRAMEWORK

This part of the chapter is devoted to describing the experimental framework used in this study. Concretely, in Section 5.1 we give the details of the realistic simulation scenario employed in the experimentation. Then, in Section 5.2 we describe the two methods taken as a baseline to compare it versus the four nature-inspired metaheuristics proposed in this paper. After that, we list the performance measures used in the experimentation and finally, Section 5.4 points out the details of the implementations carried out.

5.1 VANET simulation scenario

The simulation scenario used is the same employed in [11]. The VANET simulation was done using GPS traces publicly available in NS-2 format1, which were taken as input in the network simulation platform. The tool SUMO (Simulation of Urban Mobility) was then used to generate the mobility traces using realistic input data, including road network, vehicles routes or traffic lights among

---

1http://neo.lcc.uma.es/staff/jamal/vanet/?q=node/11
Table 1: Main Simulation Parameters

| Type            | Parameter            | Value          |
|-----------------|----------------------|----------------|
| Neighbor        | CAM Frequency        | 1 Hz           |
| Information     | Neighbor Table Timeout| 5 s            |
|                 | Server update Frequency | 1 Hz        |
| Dissemination   | Frequency            | 1 Hz           |
|                 | Request dissemination area | 0.44 km² |
| Scenario        | Type                 | Urban          |
|                 | Number of Vehicles   | 45             |
|                 | Simulation Duration  | 180 s          |
|                 | Vehicle Speed        | 10-50 km/h     |
|                 | Vehicle Density      | 113 veh/km²    |
|                 | Maximum VI size      | [2,4,6,8,10] nodes |
| 802.11p         | Bit Rate             | 6 Mbps         |
| Network         | Bandwidth            | 10 MHz         |
|                 | Frequency band       | 5.9 GHz        |
|                 | Maximum Tx Power     | [16, 21, 23] dBm |
| LTE             | eNodeB Tx Power      | 30 dBm         |
| Network         | UE Tx Power          | 10 dBm         |
|                 | Propagation Model    | Friis Tx Eq    |

The simulation period is 180s and the maximum vehicle speed is 50km/h. Table 1 contains the details of the simulation parameters. For further details, the interested reader is referred to [41]. Finally, we have considered three different maximum transmission powers for the short-range network (IEEE 802.11p), 16dBm, 21dBm, and 23dBm, and five different maximum VI sizes (2, 4, 6, 8 and 10 nodes). The transmission power controls the communication range of the vehicles and therefore the neighbourhood awareness levels: a higher transmission power implies a higher communication range that, in turn, entails a higher number of neighbours per vehicle, but also a higher consumption of energy. In the VCP models, the transmission power can be seen as a factor that modifies the number of vertices adjacent to each vertex. Regarding the maximum VI size, represented by the parameter p in the model described in Section 3.2, it limits the resource consumption in terms of LTE connections.

5.2 Baseline algorithms

The methods used to compare the performance of the four nature-inspired methods proposed in this paper are, on the one hand, the algorithm proposed in the original paper of NAVI architecture [11] and on the other hand, the Genetic Algorithm proposed by Masegosa et al in [23]. We will refer to these methods as to which we will refer to as NAVI_Alg and NAVI_GA, respectively.

Regarding NAVI_Alg, the authors applied a Min-Max formulation to model the optimization problem and an ad-hoc greedy algorithm as a solver. The objective consisted of maximizing the number of zones covered while minimizing the number of vehicles used as VI, using at most p vehicles. The general idea of the greedy algorithm is that, at each step, it selects as VI that vehicle that covers the maximum number of zones that has not been covered yet. The method stops when the VI size is equal to p or when all zones are covered. The interested reader is referred to [11] for further details. As for NAVI_GA, it is a generational GA with binary codification and elitism. Regarding the genetic operators, the method used tournament selection with size q, uniform crossover and uniform mutation. Further details can be found in [23].

5.3 Performance measures

The performance metrics used to compare the algorithms considered are given below:

- **Covered Area (Maximize)**: percentage of zones covered, which corresponds with those regions that have at least one vehicle that would receive the dissemination message.
- **VI size (Minimize)**: number of vehicles selected as VI. In this way, we measure the resource consumption of the solution.

5.4 Implementation details

To finish with the description of the experimental framework, we give here the details of the implementation done and the parameter settings used for all the four proposed nature-inspired metaheuristics. First of all, the population size has been fixed in 100 for each approach. For the FA, γ=0.95. Furthermore, for BA α=β=0.98, A0=1.0 and r=0.1. Besides that, for CS p=0.2. Finally, PSO have been configured as described in Section 4. For the development and parameterization of these methods, the guidelines given in [27, 29–31] have been followed. Regarding GA_NAVI, the population size was also 100, the tournament size was set to 5, the crossover rate to 0.5 and the mutation rate to 0.015. The stopping criteria for FA, BA, CS, PSO and GA_NAVI was 100 iterations/generations. In the objective function, the penalization coefficient c was empirically set to 0.2.

As for the simulation scenario, the region of interest was divided into 100 rectangular zones, all of them with the same width and height. The number of vehicles to be covered was 45, and we assume that all of them are equipped with both short and long range communication capabilities, that is, all of them can be used as VI. The positions of the vehicles were sampled every second. Given that the simulation time for the data was 180s, in this way, we have a total of 180 instances of the problem. Taking into account that we considered three different transmission powers and five different maximum VI size, this experimentation counts with a total of 180×3×5=2700 instance configurations. For each of these instance configurations, our proposal was run 10 times and the mean covered area and VI size were registered.

The implementation was done in Java 8, and the experiments were run on a computer with Ubuntu 16.08, 42GB RAM and 2 CPUs Intel Xeon Silver 4114 2,2GHz 13.75MB Cache 10 Cores. The results of the NAVI method were provided by its authors.

6 RESULT ANALYSIS

The objective of the experimentation done in this paper is two-fold:

- **Assess the performance of the proposed nature-inspired metaheuristics in a real scenario.** Concretely, our aim is to test on a real scenario the four methods presented in this paper to get insights about their performance and check under what circumstances one method is preferred over the others.
- **Compare the performance of the proposed nature-inspired metaheuristics versus two baseline algorithms.** To assess the
Beginning with the covered area, we can observe in Figure 2 that, as expected, for a specific MTxP, the higher the MVIS the higher the percentage of zones covered, and vice versa, for a specific MVIS, the higher the MTxP the higher the covering. A similar rule can be established for the complexity of the problem which is negatively correlated with both MTxP and MVIS. The lower the MTxP or the MVIS, the worse the performance of the methods. Comparing the four proposed methods among them, we can see that the difference in performances varies with the MVIS, but not with the MTxP. When MVIS is equal to two, for the three MTxP values considered, CS is the best performing method, whereas BA, FA and PSO obtain pretty similar results. However, when MVIS is equal to four and MTxP to 16 dBm, BA works slightly worse than the other three ones that show very similar performance, again. In the rest of cases, no differences in the percentage of covered zones can be appreciated. If the comparison is made w.r.t the baseline algorithms, the four proposed methods outperform NAVI_Alg and obtain better or equal results than NAVI_GA, which proves the good performance of the nature-inspired metaheuristics proposed in this paper.

Regarding the VI size, displayed in Figure 3, it is interesting to see that the relative performance of BA, CS, FA and PSO varies significantly w.r.t when the covered area was considered. In this case, we can observe that CS is actually the worse method, event for "easy" instances, as those with MVIS set to two. The second worse performing method was the PSO algorithm, especially if we look at the results with ten as MVIS, but very close to the best performing methods, which are BA and FA. The comparison w.r.t the baseline algorithms is quite similar to what we see above when BA, FA and PSO are considered, that is, they outperform NAVI_Alg and they obtain better results than NAVI_GA in almost all cases. The exception here is the CS algorithm, which performs even worse than NAVI when the MVIS has a value of two, and similar or worse than NAVI_GA for 21 and 23 dBm as MTxP. The only cases when we can observe a better performance of the CS w.r.t NAVI_GA are in those scenarios with 16 dBm as MTxP and 8 and 10 as MVIS values, respectively.

To assess whether the differences in performance observed among the different analyzed methods are significant or not, we made use of non-parametric statistical tests. Two statistical tests have been applied, following the guidelines proposed in [10], for the two measures considered: the covered area and the VI size. First, the Friedman’s test for multiple comparisons has been applied to check whether there are significant differences among the studied methods. The samples correspond with the mean performance of the algorithm for a pair MTxP-MVIS, that is, for each method, there are 15 samples. Given that the p-value returned by this test for both measures was 0.0, the null hypothesis can be rejected in all cases. The mean ranking returned by this test for the covered area and the VI size is displayed in Table 2. According to this ranking, BA and FA are the best performing methods for the two considered measures, respectively. Holm post-hoc test has also been applied using the best rank method in each case (highlighted in bold). In order to highlight significant differences, those p-values lower than 0.05 are marked with an asterisk (*). For the covered zones, BA is significantly better than the two baseline algorithms. Regarding the VI size, in addition to the two baseline algorithms, FA improving significantly CS. To sum up, this analysis confirms the better
The solutions was measured in terms of covered area and number in the downtown of the city of Malaga, Spain. Besides, we consid-
vised this optimization problem: BA, CS, FA y PSO. Furthermore, presented and compare four new nature-inspired metaheuristics to

7 CONCLUSIONS

In this work, we have aimed at improving information dissemination in VANETs deepening into the approach presented by Masegosa et al. in [23], which is based on the optimization of a Covering Location Problem by means of a Genetic Algorithm. Concretely, we have presented and compare four new nature-inspired metaheuristics to address this optimization problem: BA, CS, FA y PSO. Furthermore, we have shifted the model to Vertex Cover because its abstraction is more close to the real optimization problem addressed.

The methods presented here were tested on a real scenario consisting of 45 vehicles moving on a rectangular area of 600m × 700m in the downtown of the city of Malaga, Spain. Besides, we considered three different maximum transmission powers for the 802.11p network and five different maximum VI sizes. The performance of the solutions was measured in terms of covered area and number of vehicles used as VI (VI size). The objective of the experimentation done over this real scenario was two-fold: on the one hand, to analyze and compare the performance of the four proposed metaheuristics w.r.t. the baseline methods, particularly BA and FA.

![Figure 3: Boxplot for the VI size results. The Y axis represents the covered area, the X axis the MVIS, the panels the MTxP and the series the methods compared.](image)

Table 2: Ranking provided by Friedman’s non-parametric test for both Covered Area and VI Size

| Method    | Ranking Covered Area | Ranking VI Size |
|-----------|----------------------|-----------------|
| NAVI_Alg  | 5.33*                | 5.20*           |
| NAVI_GA   | 4.47*                | 4.13*           |
| BA        | 2.53                 | 2.17            |
| CS        | 2.83                 | 4.47*           |
| FA        | 2.67                 | 2.03            |
| PSO       | 3.17                 | 3.00            |

ACKNOWLEDGMENTS

E. Osaba, I. Laña and J. Del Ser would like to thank the Basque Government for its support through the EMAITEK program. E. Osaba, I. Laña and J. Del Ser acknowledge the financial support from the project SCOTT: Secure Connected Trustable Things (ECSEL Joint Undertaking, ref. 737422).

J. S. Angarita-Zapata thanks the funding received from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 665959.

A. D. Masegosa thanks the LOGISTAR project, which received funding from European Union’s Horizon 2020 research and innovation programme under grant agreement No. 769142.

REFERENCES

[1] Nicola Apollonio and Bruno Simeone. 2014. The maximum vertex coverage problem on bipartite graphs. *Discrete Applied Mathematics* 165 (2014), 37 – 48.
