The final publication is available at

https://doi.org/10.1109/TMC.2019.2928803

Data Transmissions using Hub Nodes
in Vehicular Social Networks

Anna Maria Vegni, *Senior Member, IEEE*, Camilo Souza, Valeria Loscrí, *Senior Member, IEEE*, Enrique
Hernández-Orallo, and Pietro Manzoni, *Senior Member, IEEE*

Abstract—**Vehicular Social Networks (VSNs) consist of groups of individuals (*i.e.*, people) who may share common interests, preferences and needs in the context of temporal spatial proximity on roads. In this environment, the impact of human social factors, such as mobility, willingness to cooperate and personal preferences, on vehicular connectivity is taken under consideration, thus extending the concept of Vehicular Ad-hoc Networks. In VSNs, vehicles are classified based on their social degree, a vehicle considered to be a "social" one if it accesses the vehicular social network and posts messages with a frequency higher than a given threshold. Therefore, to speed up the data dissemination process within a vehicular social network, a packet should be forwarded to those vehicles showing high social activity.**

In a previous paper, we introduced a new probabilistic-based broadcasting scheme called SCARF (SoCial-Aware Reliable Forwarding Technique for Vehicular Communications), and we analytically demonstrated its effectiveness in packet transmission reduction while guaranteeing network dissemination. In this paper, we assess SCARF in more realistic scenarios with real traffic traces, and we compare it with other similar techniques. We show that SCARF outperforms other approaches in terms of delivery ratio, while guaranteeing acceptable time delay values and average number of forwardings.

Index Terms—**Vehicular social networks, social degree, forwarding probability, delivery delay, real traffic scenarios**

I. INTRODUCTION

Vehicular Social Networks (VSNs) are emerging as a novel communication paradigm that exploits opportunistic encounters among vehicles for mobile social networking and collaborative content dissemination [1]. In a VSN, vehicles become members of a mobile social network, which is formed *on-the-fly* among neighbouring vehicles with common interests (*i.e.*, content-based VSN), or moving to the same location (*i.e.*, position-based VSN), or having relationship bindings (*i.e.*, relationship-based VSN). Dynamic links formed in a VSN are due to the opportunistic nature of vehicular networks that make them time limited, as well as the fact that users'

A. M. Vegni (*corresponding author*) is with COMLAB Telecommunication lab at the Department of Engineering, Roma Tre University, Rome, Italy, e-mail: annamaria.vegni@uniroma3.it

V. Loscrí is with INRIA Lille-Nord Europe, Lille, France, e-mail: valeria.loscri@inria.fr

C. Souza is with Federal University of Amazonas, Institute of Computing Manaus, Brazil, e-mail: camilo.souza@icomp.ufam.edu.br

E. Hernández-Orallo and P. Manzoni are with the Computer Engineering Department, Universität Politècnica de València, València, Spain, e-mail: {ehernandez, pmanzoni}@disca.upv.es

access is allowed only in given positions and for a limited time interval [1].

A VSN is mainly composed of two fundamental parts, namely: (*i*) a vehicular ad hoc network that represents the physical layer, and (*ii*) a social network framework running on top of said physical vehicular network. The synergy between these two layers allows content dissemination protocols and routing algorithms to be designed in VSNs, which exploit not only the mobility patterns of human beings and vehicles, but also the social properties that exist mainly among humans and also among vehicles (*i.e.*, Social Internet of Vehicles).

In order to better understand the behaviour of social-based vehicular networks, many researchers make use of the main tools of Social Network Analysis (SNA) [2], which takes into account social relationships in terms of nodes (*i.e.*, individuals) and ties (*i.e.*, relationships among nodes), to identify the important components of a social network [3]. For example, Xia *et al.* [4] describe BEEINFO, an interest-based forwarding mechanism that uses an artificial bee colony for socially-aware networking, that is, a routing mechanism that classifies communities into specified categories on the basis of personal interests. Data dissemination is based on the information on node density, such that the higher the density, the more nodes the community has. This information provides a guideline for better selecting next-hop forwarders. As another example, SocialCast [5] is based on data similarity metrics only and considers the interests shared among nodes. Messages are forwarded by a given node that represents a good message carrier for a given interest. Finally, hybrid approaches consider both social and data similarity metrics, like SCORP (Social-aware Content-based Opportunistic Routing Protocol) [6], which considers the type of content and the social relationship among nodes so that data forwarding occurs according to the social degree of a node with respect to others interested in a given message.

Derived from classical data dissemination protocols, such as those based on Delay Tolerant Networks for example, centrality metrics are used to denote how "important" a node is inside a network. An "important" node plays a crucial role in a vehicular social network and performs relevant tasks like data forwarding or decision making (*e.g.*, a node in contact with several neighboring nodes will be able to forward a data message to a huge number of next-hop forwarders due to its high encounter probability). For instance, Bubble Rap [7] is based on social metrics only and combines node centrality with the notion of community to make forwarding decisions.

The centrality metric allows *hub nodes* to be identified in the

communities. A "hub" is a node with a number of links (*i.e.*, a node degree) greater than the average in the whole network. Hubs represent a unique feature of scale-free networks, which are ruled by the power-law degree distribution.

In a previous paper, we introduced SCARF (SoCial-Aware Reliable Forwarding Technique for Vehicular Communications) [8], a data dissemination technique based on social similarity metrics. SCARF is a next-hop forwarding selection protocol for VSNs that, in addition to being based on the typical physical parameters used in vehicular ad-hoc networks, also considers the social features that may exist among nodes. In particular, SCARF selects the next-hop vehicle only if its "sociality features" indicate it as eligible. In such context, a forwarder hub vehicle (*i.e.*, a node with a very large network degree) will likely disseminate messages to a larger number of neighbors as compared to other vehicles with a lower network degree. How to define a node as a "social vehicle" is an open issue.

Unlike traditional forwarding techniques used in traditional opportunistic networks, SCARF is most suitable for social-aware vehicular scenarios.

In this paper, we further investigate the SCARF technique by assessing it with both synthetic and real traffic scenarios. A comparison has also been carried out with existing related protocols used for vehicular networking that do not necessary deal with social features. As one of the main results, SCARF presents its effectiveness in providing a higher delivery ratio and lower delay, especially when the number of vehicles increases in the network, meaning that it is able to select as nexthop forwarders those vehicles that are "socially active" and are connected with a larger number of nodes. As a result, SCARF maintains a lower number of forwardings, meaning that it is able to find those vehicles that have a higher social degree, thus allowing for a very effective data dissemination mechanism in VSNs. Finally, we observe that SCARF is not only able to forward messages to social vehicles, but is also selective in choosing the most likely hub vehicle within a transmission range.

To summarize, unlike [8], in this paper we present the system model for VSN, together with a novel network architecture based on edge computing. We also provide a definition of hub node and social degree concepts, in addition to an analytical performance evaluation of SCARF. Finally, extended simulation results of SCARF have been carried out. The following contributions are addressed in this paper:

1) The SCARF technique, already introduced in [8], has been largely assessed in both synthetic and real trace scenarios and has been compared with existing routing protocols for vehicular networks. Its effectiveness, due to selection of hub nodes as next-hop forwarders, is proven in terms of (*i*) high delivery ratio, (*ii*) low delivery delay, and (*iii*) reduced number of forwardings;

2) A VSN system model describes how a vehicular social network is built and how a vehicle can be elected as a social node. The activity degree provides information about the "social addiction" of a vehicle (*i.e.*, how social a vehicle is in a VSN), mainly depending on how many messages it uses to send when connected to a vehicular

social network and its willingness to communicate. As expected, we observe a Pareto distribution of social vehicles, meaning that most nodes present a low social degree, while only few vehicles have a high social degree;

3) Data from each vehicle (*i.e.*, the activity degree) are collected through an edge computing-based architecture, where a central node (*i.e.*, a Road Side Unit) gathers such information to be processed and shared within the network.

This paper is organised as follows. Section II presents an overview of the main techniques for message forwarding based on vehicle social behaviour. In Section III, we introduce the system model that describes how a VSN is formed and provides definitions of *hub node* and *social degree*, and then in Subsection III-A we describe the edge computing-based network architecture in order to represent a real-world based solution for gathering the needed information in the proposed SCARF forwarding approach. Section IV describes the mathematical model of SCARF and derives the average number of forwarders occurring in a transmission range. In Section V, we report and discuss the simulation results related to SCARF technique carried out both in real traffic and synthetic trace scenarios and compare it to other existing techniques. Finally, conclusions are drawn at the end of this paper.

II. RELATED WORKS

The study of social features in contexts that differ from well-known OSNs is on the rise. In the context of information networks, identifying hubs allows data to be effectively spread on the network. Many methods are based on the topology of the network and ignore other important properties, such as the nodes' features and the way they process information [9]. For instance, Ilyas *et al.* [10] consider a user who is connected to well-connected users to have a more central status and so is more likely to be eligible as central node.

Identifying hub nodes in mobile social networks –and particularly, in VSNs– may be very challenging as connectivity links –and social ties– exist in limited time intervals and spatial positions [1] due to the vehicles' mobility patterns. Additionally, the dynamics of vehicle mobility, as stated in [11], have been difficult to characterise and model, so in order to perform realistic evaluations, we should rely on the use of realistic traces. In [12], da Cunha *et al.* study social metrics in order to characterise vehicle mobility in a realistic dataset. Mobility in VSNs can be modeled as a social network following the laws of the degree distribution and the short distance among nodes. It was also possible to verify that the use of social metrics can help in improving the performance of protocols, network infrastructure, and to propose new services to assist the drivers.

How to define which social metric efficiently models the activities of members in social networks remains a challenge. Traditional social metrics (*i.e.*, closeness, betweenness, and bridging) are application-oriented, usually exploited in the design of routing protocols in VSNs while achieving higher delivery ratio, and shorter end-to-end delays. In [13], a node is

selected as next hop not only according to traditional greedy approaches (*i.e.*, the closest node to a given destination), but also considering social factors like centrality. Similarly, in [14] Bradai *et al.* select rebroadcaster vehicles based on their strategic location in the network and their capacity for reaching other vehicles by using a new centrality metric called dissemination capacity. In [15], Stagkopoulou *et al.* use a specific social inspired metric, namely Probabilistic Control Centrality, in order to identify potential vehicles for message forwarding and the coverage of a wide range of a vehicular network. Recently, in [16], we proposed two data dissemination techniques for VSNs based on graph metrics, like betweenness and the local clustering coefficient, in order to select central nodes as next-hop forwarders. Finally, in [17] Lu *et al.* investigate network performance in a social-proximity vehicular network, where each vehicle transmits to a destination vehicle sharing the same social interest.

One characteristic common to all the previous works is the focus on selecting the best vehicles for rebroadcasting data messages according to social metrics, like the clustering coefficient and the node degree. More generally, it is possible to define a *social degree* that allows the most appropriate vehicle for data forwarding to be selected in order to maximise network performance. A tentative definition of social degree has been proposed in [18] as a score that evaluates the probability that within a given time period, a vehicle can establish a reliable social connection with a neighbouring vehicle. In [19], Neto *et al.* consider the Waze dataset and investigate the posting behaviour of users in vehicular social networks. They observe the influence of speed and delay on the posting activity of vehicles so that most contributions to the vehicular social network are triggered during rush hours, mainly on weekdays. Another work [20] studies the relation between human geographic movement, temporal dynamics and the ties of the social network. The authors conclude that short-range travel is less impacted by the social network structure while a person going a long distance is more likely to travel near an existing friend. Relevant studies have shown the existence of regular driving patterns in time and in space in the vehicular context [21]. This feature is extremely interesting when considered in combination with social networking. Indeed, a social network can be associated to a social structure composed of basic units (*e.g.*, individuals or nodes) sharing some common interests or ideas. The identification of regular driving patterns on the road emerges from people that normally take the same road every day. From these premises, it seems straightforward to think about vehicular social networks as "social networks on the roads" [22], [23].

Vehicular social networks present intrinsic challenges and issues as identified in [24], that derive from both of the networks, *i.e.*, the social and the vehicular ones. In particular, the unreliable connectivity of wireless links may impact the quality of social interactions and efficiency data transmission [25], while effective information dissemination mechanisms are necessary [26]. This aspect is becoming more urgent with the advent of new potential applications and services that could highly benefit from social ties. There is an increasing need to showcase vehicular social applications and services

enabling proximity-based social interactions among vehicles during daily travel [27], [28]. The major trend to make social networks available among vehicles is Mobile Social Software (MoSoSo) [29]. MoSoSo is a class of mobile applications related to the concept of mobile Internet, with the aim of supporting social interactions among interconnected mobile users for data sharing.

Commercial applications of VSNs have recently started emerging in the daily routine of vehicles. As an example, the Twittermobile car developed by Ford [30] allows Twitter messages to be sent containing information ranging from the driver's mood to real-time traffic warnings. Similarly, traffic related voice tweets can be shared through NaviTweet [31], such that the driver's preferences can be incorporated into the navigator's route calculation. RoadSpeak [22] is a voice chat system used by daily driving commuters or a group of people who are on a commuter bus or train. Finally, another interesting application is vehicle-to-vehicle (V2V) charging/discharging systems for electric vehicles (EVs). When vehicles with excess energy can be considered to be alternative charging points for other EVs that need to be recharged, an efficient social network among these vehicles would have a huge impact in terms of charging effectiveness [32]. In general, reducing end-to-end delay and achieving a higher delivery ratio is a challenge in VSNs [33], [34]. Appropriate mechanisms have to be conceived in order to deal with the intermittent connectivity. In [35], the authors focus on this issue by considering buses that periodically run along fixed routes. Specifically, they propose a routing strategy based on an ant colony optimisation approach. This type of approach relies on the knowledge of the bus schedules. For other types of vehicles, even though regularly patterns have been envisaged, they are not subject to regular schedules. In this sense, the identification of specific social vehicular ties as we propose in this paper could be more generally applicable.

To summarize, VSNs enhance commuter communication along the roads for different applications, such as (*i*) safety, (*ii*) convenience, (*iii*) comfort, and (*iv*) entertainment-based applications [36]. As an example, commuters spend more time during rush hours in comparison to normal traffic conditions; in such a case, entertainment-based applications of VSNs facilitate commuters's access to gaming, as well as to sharing/downloading music, video, and photos. Regarding safety-based applications, social-aware dissemination protocols allow messages to be forwarded to social nodes. SocVe (Social-aware Vehicular DTN protocol) [37] has been proposed for safety applications, such as emergency support services. The goal of SocVe is to transmit safety messages to one or more predefined important node in the network. The forwarding decision is based on several social metrics such centrality, tie strength, etc.

In this paper, we present a data forwarding technique in VSNs, namely SCARF, based on the selection of hub nodes. It aims to identify social nodes based on the concept of the social activity of a node. Social nodes are expected to best disseminate messages across the VSN. SCARF can be used for different VSN applications, from security to entertainment-based applications. For instance, for safety-based applica-
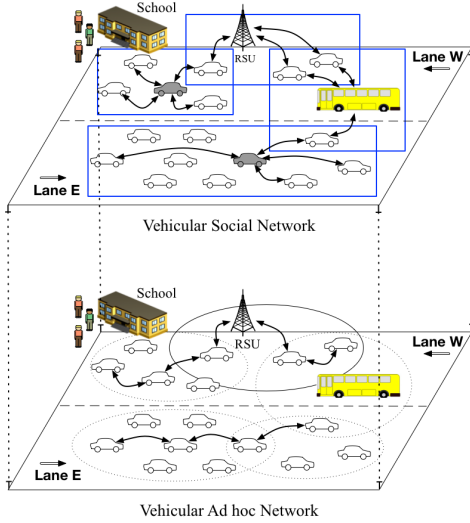
Fig. 1: Mapping between VANET and VSN. In VANET, next-hop vehicles are selected within a transmission range (*dotted circles*) according to a greedy algorithm. In VSNs, vehicles are distinguished according to social features (*i.e.*, social degree). Vehicles are members of a VSN (*blue rectangles*), if they share common interests. Social nodes (*grey vehicles*) are preferably selected as next-hop forwarders in order to reach more nodes.
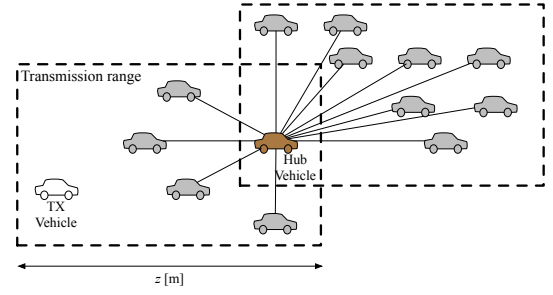

Fig. 2: Schematic of a VSN comprised of two clusters (*dotted areas*). The hub vehicle (*brown vehicle*) has a higher network degree and will be likely to disseminate an information packet to a large number of neighboring vehicles. From a social networking perspective, a hub vehicle is a social node with high activity degree.

tions, forwarding messages to social nodes (*e.g.*, in such a case, police cars, fire trucks, ambulances, and so on) can enhance the message dissemination within the overall network. SCARF stems in part from traditional forwarding probabilistic techniques applied to vehicular communications, and then it depends on physical parameters like the inter-vehicle distance and the vehicular density, so that the farther the forwarder from the source vehicle is, the higher the forwarding probability is. Intuitively, the farthest vehicle within range of the transmitter should be the most likely next-hop forwarder of the packet as this will yield the highest forward progress. However, the social features of vehicles are also taken into account, meaning that for a fixed distance, not all the vehicles are potentially eligible as next-hop forwarder.

Unlike related works on social networking applied to vehicular networks, we do not rely on well-known graph theory metrics of node centrality, but rather we introduce the activity degree of a node in a VSN based on the average number of messages (*i.e.*, posts) sent by a node when connected to the VSN. Taking into account such information, we can identify the most likely social (hub) node within a transmission range, which can be then selected as next-hop forwarder. Selecting a hub node is expected to reach a higher number of neighbors, thus providing a reliable data dissemination.

Notice that the content of a post message depends on the specific application of VSNs. For instance, the Ford Tweeting car is a socially-aware cloud-based application, which enables drivers to be connected to other neighboring drivers and share vehicle and route information. In this specific context, the content of a message should be the vehicle and route information (*e.g.*, speed, direction, position, fuel consumption,

traffic conditions, etc.).

III. VSN SYSTEM MODEL

Vehicular social networks arise from well-known VANETs, where social ties are formed *on-the-fly* according to the vehicles' pattern mobility, routines and shared interests [1]. In VANETs a node is selected as a next-hop forwarder based on some greedy algorithm. Physical features like channel quality, transmission range, inter-vehicular distances and so on, are taken into account, as are mobility models.

Fig. 1 depicts the mapping between a VANET and the corresponding VSN. As the vehicular network is based on V2V and vehicle-to-infrastructure (V2I) communication links following the *store-carry-and-forward* scheme, in the vehicular social network, vehicles distinguish the importance of a vehicle in the network according to social features. Identifying a social vehicle improves network performance as it is most likely to disseminate a message to many other neighbors. Graph metrics like centrality, betweenness and clustering coefficient may aide in the detection of social (central) nodes. In Fig. 1 the grey vehicles are social nodes because their degrees are high (*i.e.*, high number of neighbors).

A VSN is formed according to specific criteria, namely: position, content and relationship [1]. For example, vehicles moving towards a school are most likely to share common interests related to the school context. A VSN is therefore most likely to be built among such vehicles (see the blue rectangle near the school). Similar considerations apply to other VSNs depicted in Fig. 1 (*e.g.*, the bus and the interested neighboring vehicles form a VSN related to public transportation).

In each VSN a central (social) node exists. However, how to identify a social vehicle is an open challenge that we aim to address in this paper. Let us consider a VSN is built based on information content shared among drivers with shared interests (*e.g.*, the VSN of vehicles driving towards the stadium to attend a football match). Neighboring vehicles can share common interests (*e.g.*, the availability of parking slots, ticket sale, match bets, etc.) thanks to social ties built based on the common interest (*i.e.*, they are fans of a football team and are physically going to the football match).

TABLE I: Parameters used in the VSN system model.

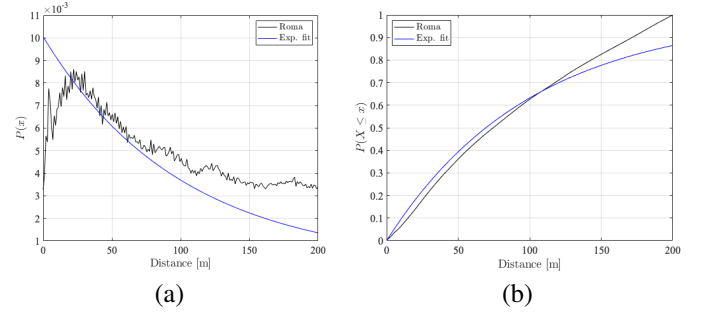| Parameter	Description
z	Transmission range [m]
d	Distance [m]
ρ	Vehicular density [veh/m]
a_i	Activity degree of the i-th vehicle, in the range $[0, 1]$
c_i	Shape factor of the i-th vehicle
S_i	Social degree of the i-th vehicle, in the range $[0, 1]$
α	Exponent of the power law distribution, in the range $[2, 3]$
N_z	Number of potential forwarders
V_{avg}	Average number of social vehicles
$a_{\min}$	Minimum value of the activity degree

Fig. 3: (a) PDF and (b) CDF of the inter-vehicle distances distribution of the Rome traffic, including an exponential fit with $\lambda = 1/\rho = 0.01$.

In such scenario, a new node can access the VSN and become member based on common characteristics. As an example, and without a loss of generality, in Fig. 2 we consider a VSN made of two clusters formed according to well-known techniques as in [38]. We focus on the cluster where a source vehicle (*i.e.*, Tx vehicle) needs to forward a packet among its N_z neighbours within the transmission range z [m]. The VSN can be modeled through a planar network topology for a generic multi-hop broadcast protocol *i.e.*, it is a relative static two-dimensional wireless network where vehicles have almost-zero relative speed[1]. A free-space propagation model and a fixed transmit power are assumed so that each vehicle has a fixed transmission range z [m]. We are aware that a two-dimensional network with free-space propagation model and a fixed transmit power could be considered not completely realistic. In previous works of some of the authors *e.g.*, [40], this aspect was thoroughly evaluated. Leveraging on such results, we consider that, specifically for the scope of this work, this model is a reasonable approximation that will not impact on the results obtained. We are aware, in any case, of the possible limitations of the approach. As a reference, Table I collects all the parameters used in the following VSN system model.

Our model assumes that inter-vehicle distances (or inter-vehicle headways) are Poisson (exponentially) distributed. An excellent and recent analysis of this topic is in [41], including several references in which empirical data were used to confirm that the exponential model is generally a good fit for the distributions of inter-vehicle distances. Furthermore, several papers have used exponential distribution for modelling the inter-vehicle distances [42], [43], [44]. In order to confirm this fact, we have performed an extensive study of the inter-vehicle distance on the Rome taxi traces [45]. We obtained the Probability Density Function (PDF) and the Cumulative Distribution Function (CDF) as depicted in Fig. 3, by considering the Euclidean distance in meters between two vehicles in movement using a period of 1 s, a distance resolution of 1 m, and the maximum distance was 200 m. In Fig. 3 we also included an exponential distribution with mean $\lambda = 0.01$, showing a reasonably good fit, and thus confirming the suitability of the exponential inter-vehicle distribution for our model.

Consequently, the nodes' positions can be generated according to a Poisson Point Process of parameter ρ, where ρ is the vehicle linear spatial density [veh/m]. It follows that the inter-vehicle spacing is exponentially distributed with mean $1/\rho = \lambda$. More formally, let X_i be a random variable denoting the space between the i-th and the $(i + 1)$-th vehicle. We assume that X_i are independent and identically distributed (*i.i.d.*) with the CDF denoted by $F_X(x)$. Finally, each vehicle is able to determine the relative distance between itself and a transmitter.

Now, let us consider the i-th vehicle is at distance d [m] from the Tx vehicle (*i.e.*, $d \leq z$). The forwarding probability of this vehicle should be proportional to (*i*) the probability that there is no vehicle in the interval of length $z - d$, and also (*ii*) to its social degree. The probability that the i-th vehicle will reforward a packet (*i.e.*, $p_{f,i}$) depends on the probability that there is no other vehicle in the interval of length $(z - d)$, which can be expressed as

$$p_{f,i} = 1 - F_X(z - d), \qquad (1)$$

where $F_X(z - d)$ is the CDF with $x = z - d$.

In Eq. (1), the dependence of the forwarding probability on distance is expressed. Other features should be also taken into account, such as the quality of channel over which the packets will be transmitted. These aspects might be enough for traditional VANETs, as each node is potentially eligible as a next-hop forwarder based solely on physical features (*i.e.*, distance and channel quality). However, as we are dealing with vehicular social networks, we also need to consider the social features of the i-th vehicle (*i.e.*, the reliability of forwarded packets depends on the willingness of the i-th vehicle to forward them to its neighbors). For instance, if the channel quality is bad (*e.g.*, low SNR, low data rate, high delays, etc.), then the forwarding probability should be low in order to avoid unsuccessful packet transmissions (*i.e.*, packet loss).

On the other hand, not all the neighboring nodes are eligible as next-hop forwarders, but specific features distinguish them from among *social* and *unsocial* vehicles. As expected, a social vehicle is likely to forward packets to a higher number of neighbors as compared to unsocial (*poor social*) nodes, which can reach a limited (*low*) number of nodes. For example, if a central node (*i.e.*, a node with the highest number of

---

[1]Notice that the assumption of static nodes is not restricting because from the perspective of a single transmitted packet, the transmission time is very short, so the network appears as static [39].

neighbors) is elected as a next-hop forwarder, packets will be disseminated more efficiently than they would be to a not central node with a lower number of neighbors. Choosing a "popular" vehicle as a next hop forwarder increases the chance of delivering the message to a broader group.

Notice that in this work we are not specifically investigating the necessary mechanisms for motivating a vehicle to collaborate in forwarding messages. We nevertheless believe that the area of Serious Games, or more in general of "gamification", should be the path forward. Serious Games are games that not only focus on pure diversion but that also take advantage of the potential of game technology and game mechanics for learning and training [46], [47]. The general objective here is to reward proper user behavior through virtual and/or real-world incentives. Rewards could be related to the ITS world, like discounts on parking, minutes of access to Congestion Charge Zones, or the like.

Based on these considerations, we assume that there are "collaborative" nodes and we can provide the following formal definitions *i.e.*:

Definition 1 – Hub node. *A node v is central if the number of direct connections to its neighbors (i.e., node degree $\delta(v)$) is the highest in the network. The node degree $\delta(v)$ is defined as:*

$$\delta(v) = \sum_{j \in V, j \neq v} \ell_{vj}, \tag{2}$$

where $\ell_{v,j}$ is the link from node v to its neighbors j (with $j \in V$, $j \neq v$, being V the set of edges of the graph associated to the given network). A hub node is a central node.

Definition 2 – Social degree. *The social degree of a node is defined as the probability that it is socially active in a VSN, assuming it is interested in the content produced and interchanged among members in that specific network, i.e.,*

$$S_i = \Pr\left\{a_i(t) \geq A\right\}, \tag{3}$$

where A is a threshold that determines if the i-th vehicle has an high activity degree (i.e., if $a_i \geq A$), with $a_i(t)$ as the activity degree of the i-th vehicle (in the range $[0,1]$) at time t.

Notice that the exact nature (*i.e.*, text, voice, images, or a combination thereof), form of production (*e.g.*, voice-to-text device or more simply by a passenger of the vehicle), and the semantic of the content produced and interchanged among nodes clearly depend on the application that is making use of this dynamically created network. It can range from simple traffic status indications with possibly some suggestions for alternatives routes, to discussions related to a topic of interest *e.g.*, a show that various people in vehicles are going to attend.

Corollary 1. *A hub node is likely to have a high social degree, but a node with a high social degree is not necessarily a hub node.*

Notice that in the case where multiple vehicles have the same value of social degree, like in the case of vehicles traveling on a freeway with the same velocity, a selection technique can be used based on the vehicle ID, like choosing
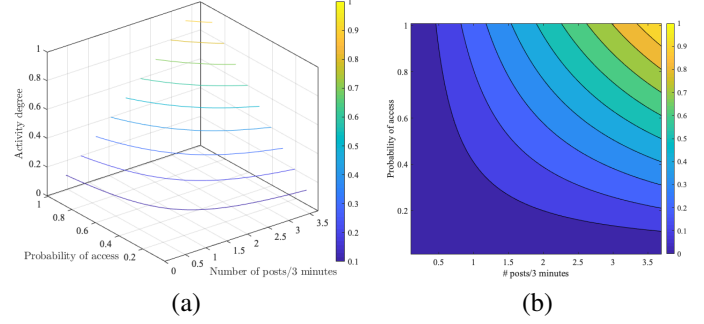

(a)   (b)

Fig. 4: (a) Contour plot of activity degree versus the number of posts over 3 minutes *i.e.*, $[0, 3.6]$, and the probability of accessing the vehicular social network *i.e.*, $[0, 1]$. (b) Top-view of the activity degree.

the one with the lowest ID. In this context, it makes sense to "clusterize" the nodes that have a relative low velocity. If there are "too quick" vehicles, their relative speeds will be too high and they will not be considered part of the cluster since the dynamics of the cluster are slower than the node dynamics.

From *Definition 2*, the activity degree of the i-th vehicle *i.e.*, $a_i(t)$, depends on the sampling time t, the number of posts[2] sent by the i-th vehicle in the VSN *i.e.*, $n_{p_i}(t)$, and the probability that the i-th vehicle accesses the VSN, *i.e.*, $p_{a_i}(t)$, *i.e.*,

$$a_i(t) = n_{p_i}(t) \cdot p_{a_i}(t)\big|_{t \in [t_{start}, t_{stop}]}, \tag{4}$$

where t_{start} and t_{stop} are the start time and end time of the mobility pattern of the i-th vehicle. Eq. (4) is within the range $[0, 1]$ when normalized w.r.t. the maximum number of posts *i.e.*, $\max_t(n_{p_i})$. Fig. 4(a) depicts the activity degree versus the probability of access and the number of posts sent by a vehicle during its activity time. In accordance with [4], we consider a maximum rate of one post every 50 seconds (*i.e.*, 3.6 posts over 3 minutes). We notice that for high probability of access and high number of posts, the normalized activity degree is expected to be maximum *i.e.*, 1. On the other hand, we have the lowest activity degree when both the probability of accessing the network and the number of posts is low. Fig. 4 (b) depicts the areas with different values of the activity degree.

In Eq. (3) the activity degree of vehicles can be assumed to be power-law distributed. In general, in a social network, few nodes present a high activity degree, while most have a lower activity degree. It follows that the threshold A in Eq. (3) can be computed according to the $80/20$ rule of power-law distribution [48]. Specifically, the fraction of vehicles N_v with an activity degree higher than A can be written as

$$N_v = \left(\frac{A}{a_{\min}}\right)^{1-\alpha}, \tag{5}$$

and the percentage N_a of the activity degree of N_v vehicles is

$$N_a = \left(\frac{A}{a_{\min}}\right)^{2-\alpha}. \tag{6}$$

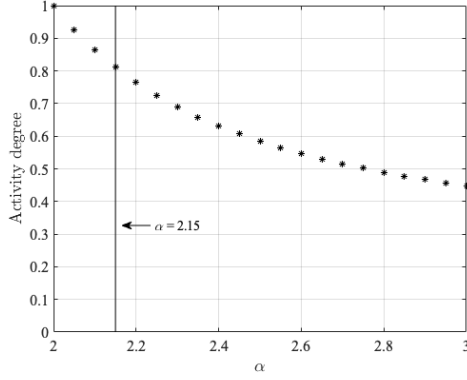
[2] A post is a message produced in a VSN.

Fig. 5: Activity degree vs. α. The threshold $A = 0.8$ is set for $\alpha = 2.15$.

Fig. 6: Edge computing-based network architecture.

Since the exponential of power-law distribution of scale-free networks is in the range $[2, 3]$ [48], and assuming $N_v = 20\%$, we can choose $\alpha = 2.15$, which provides an activity degree of ≈ 0.8. This means that the 20% of vehicles have an activity degree equal to ≈ 0.8, for $\alpha = 2.15$. For increasing α, the activity degree decreases, as depicted in Fig. 5.

Leveraging on the previous considerations, in Eq. (1) we include the shape factor c_i and the social degree S_i of the i-th vehicle, which take into account the channel quality and social features, respectively. Eq. (1) becomes

$$p_{f,i}^{SCARF} = \left[1 - F_X(z - d)\right]^{1/(c_i \cdot S_i)}, \tag{7}$$

which is the SCARF forwarding probability assignment function with $c_i \geq 1$ and $S_i = (0, 1]$. The higher the value of c_i (S_i), the higher the forwarding probability $p_{f,i}$. Finally, as we assumed the inter-vehicle spacing is exponentially distributed, Eq. (7) becomes [8]

$$p_{f,i}^{SCARF} = \exp\left[-\rho \frac{z - d}{c_i \cdot S_i}\right]. \tag{8}$$

A. Edge-computing Architecture for VSNs

One of the main goals of Vehicle Social Networks, as all Smart Cities or more generally IoT, is to provide solutions capable of handling possible scalability issues. Edge computing is essentially the process of decentralising computer services and moving them closer to the source of data. This approach has a significant impact on scalability as it can drastically reduce the volume of data moved and the distance it travels. For this reason, in this paper we introduce an edge computing architecture to support the data exchange and gathering of relevant information, such as the activity degree of vehicles. Specifically, we treat each vehicle as a "content island" *i.e.*, a node that can interconnect groups of integrated "things" to exchange and process data among other content islands. We are extending the concept of "node" represented here by a vehicle, using the idea of edge computing.

Following the results in [49], we consider a content island to be a publish/subscribe-based system where a set of "things", capable of performing processing and sensing, interchange messages inside the island and with other neighboring islands,
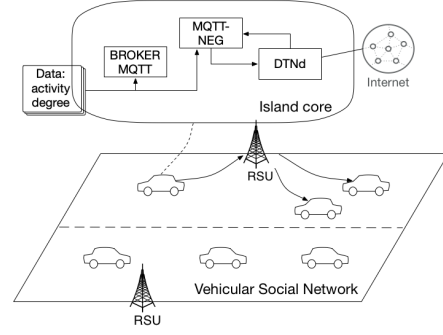
talking MQTT. The island core allows the flow of messages to be handled directly inside the island and among other islands. Fig. 6 depicts the concept of a vehicle as a content island.

We rely on an edge-based architecture in order to leverage a content-centric network paradigm where the content is cached close to the mobile user so as to improve the access delay and avoid wasting resources. Similar approaches have been used in [50], [51], [52], [53]. Information about the social degree of each vehicle is spread within a cluster through a V2I link to a Road Side Unit (RSU). Each vehicle shares its information about activity degree, which is updated in a given time interval (*e.g.*, every 3 minutes in a 30-minutes journey). The RSU can then relay the received information on to other vehicles in the network.

In our architecture, a message can be *(i)* a request for processing or data analysis, *(ii)* an action trigger, or *(iii)* a collection of data. Messages are interchanged among nodes that belong to either the same island or to other islands. A request for data analysis or processing indicates the need for data intensive processing over a data-set. In our scenario, the activity degree of the i-th node is computed. An action trigger can be a request to, if certain conditions hold, elect a node as a social node or hub. Finally, a collection of data is a set of values obtained from a node, such as the activity degree values from short and long-period records.

The island core is built around: *(i)* an MQTT broker, *(ii)* a DTN daemon, and *(iii)* the MQTT near-user edge gateway (MQTT-NEG). Messages can be *i.e.*, *(i)* local and *(ii)* global. Global messages are both distributed inside the island and forwarded to other islands by the DTN daemon. Inside the island, connectivity is provided using any common standard technology, like Wi-Fi. Inter-island connectivity is where the VSN and the SCARF forwarding technique come into play, opening up the possibility of distributing information, creating "conversations" among the members of specific groups of islands, and deploying applications that allow users to cooperate.

IV. SCARF TECHNIQUE

SCARF is a message rebroadcasting technique that selects the forwarding nodes based on their social degree; we refer to them as "social vehicles" [8]. Similarly to what happens in traditional on-line social networks, SCARF classifies vehicles into different categories based on their social degree. A social vehicle shows an appropriate social behaviour with respect
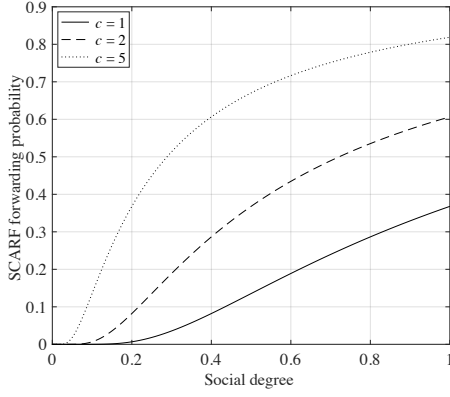
Fig. 7: SCARF forwarding probability for a vehicle at distance $d = 150$ m within the source vehicle's transmission range $z = 200$ [m], and for different values of the shape factor c.
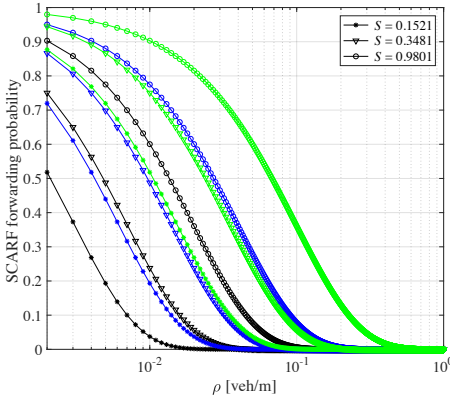

Fig. 8: SCARF forwarding probability for a vehicle within the source vehicle's transmission range z [m], versus the vehicular density ρ [veh/m] and for selected values of the social degree and of the shape factor, *i.e.*, $c = [1, 2, 5]$ for *black*, *blue* and *green lines*, respectively.

to a given VSN. With the term "appropriate", we refer to a social behaviour that makes the vehicle eligible to be a "social vehicle" in a specific VSN. For instance, a vehicle that accesses a VSN every day and has a high social activity degree (*e.g.*, it posts several comments in the VSN) can be awarded as "social vehicle". A hub vehicle, *i.e.*, a vehicle having the highest network degree, is eligible as social vehicle since it is connected with the highest number of neighbours in a VSN.

Mathematically, the social behaviour of vehicle i is expressed in terms of its social degree S_i, ranging from 0, where the vehicle has no social behaviour and so is not eligible, to 1, where the vehicle has a high social behaviour and is therefore a potential next-hop vehicle.

According to the SCARF technique, a vehicle that receives a packet should forward it only to the "social vehicle i", that is to say, the one among those within its transmission range that has the highest social-based forwarding probability, computed according to the probability assignment function in Eq. (8). Fig. 7 depicts the SCARF forwarding probability for the i-th vehicle at distance $d = 150$ m within the source vehicle's

transmission range (*i.e.*, $z = 200$ m, which can be regarded as a possible range of an IEEE $802.11p$ device) versus the social degree S_i and for different values of c_i (*i.e.*, $c_i = [1, 2, 5]$). The shape factor implicitly takes into account information about the quality of channel, acting as the signal-to-noise ratio associated to the communication channel of the i-th vehicle. The higher the c_i value, the higher the quality of the channel of the i-th vehicle and therefore the higher its forwarding probability. Notice that different values of c_i depend on the radio technology used to transmit data and on the effective transmission range of the source vehicle. Formally, c_i can be defined as:

$$c_i = \frac{z + z_{e,i}}{z - z_{e,i}}, \quad \text{with } z_{e,i} \neq z \tag{9}$$

where $z_{e,i}$ [m] is the effective transmission range of the i-th vehicle (*i.e.*, $z_{e,i} < z$). For instance, in the case of IEEE $802.11p$, the nominal transmission range is 200 m but the effective one is expected to be lower (*e.g.*, 100 m) if there is a fading and noisy channel; so in this case, the shaping factor will be $c_i = 3$. The higher the c_i value, the higher $z_{e,i}$, meaning it is approaching the nominal transmission range z. The shaping factor can be further modeled as a random variable C_I, defined to take any value in the interval $I = [1, B] = \{c \in \mathbb{N} : 1 \leq c \leq B\}$ with equal likelihood and associated probability density function as

$$f_C(c) = \begin{cases} \dfrac{1}{T}, & 1 \leqslant c \leqslant B \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

with T as the interval width (*i.e.*, $B - 1$).

Another important property of SCARF forwarding probability assignment function is that it should be able to adapt to the vehicular density changes. In Fig. 8, the SCARF probabilities are shown as functions of the vehicle density for a vehicle at a fixed distance from the transmitter (*i.e.*, 150 m in a 200 meters transmission range). It can be observed that the SCARF forwarding probability assignment function is adaptive to the vehicle density. As the vehicle density increases, the forwarding probability decreases.

The dependence of SCARF forwarding probability behavior on (*i*) distance, (*ii*) social degree, and (*iii*) shape factor, is depicted in Fig. 9. For increasing distances, the forwarding probability increases, reaching the maximum value when the social degree is also at its maximum. When the shape factor increases (*i.e.*, good channel quality), the SCARF forwarding probability increases. In this case, even if the i-th vehicle has a low social degree, its probability of forwarding is expected to be non null. In contrast, if the shape factor is low (*i.e.*, poor channel quality) but its social degree is high, then the forwarding probability associated to the i-th vehicle will be maximum, meaning that it will try to forward the message.

The main feature of SCARF is its applicability to vehicular social networks, as it distinguishes and selects only social vehicles likely to be the most suitable next-hop forwarders for VSN applications for the purpose of disseminating data information in a fast and reliable manner. Indeed, a social vehicle showing a high social degree is expected to be collaborative and willing to forward data messages in the VSN. We therefore assume
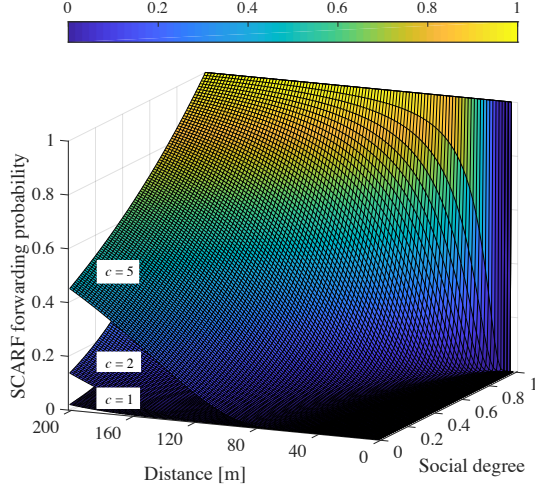
Fig. 9: SCARF forwarding probability versus the distance d [m] (with $d \leq z$) and the social degree. In this case, we assumed $\rho = 0.02$ [veh/m] and different values of the shape factor, *i.e.*, $c_i = [1, 2, 5]$.
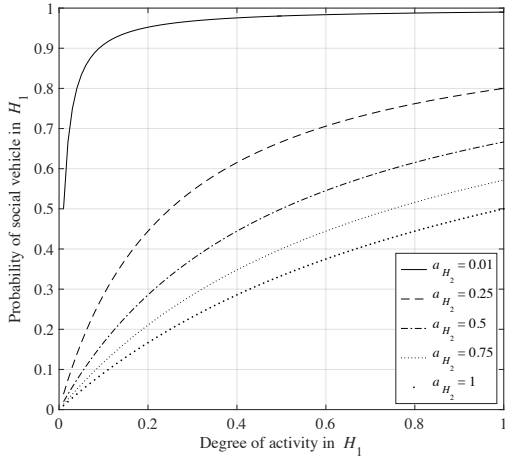

Fig. 10: Probability of a social vehicle versus the degree of activity under the hypothesis H_1, for different values of the degree of activity in H_2.

the social degree of the i-th vehicle also provides information about its willingness to collaborate and forward messages. In this way, SCARF avoids to select selfish vehicles [54] that can affect network performance (*e.g.*, causing a lower packet delivery ratio).

So far, we have considered one given vehicle to be connected in one given VSN (*i.e.*, VSN$_1$). However, other networks could be present and a vehicle might be then connected to more than one VSN (*e.g.*, VSN$_2$). We can assume the following two hypotheses:

$$H_1 : \text{the } i\text{-th vehicle is connected to } VSN_1$$
$$H_2 : \text{the } i\text{-th vehicle is connected to } VSN_2 \tag{11}$$

Where $\Pr(H_1)$ and $\Pr(H_2)$ are the probability that the i-th vehicle is under hypothesis H_1 and H_2, respectively, we have $\Pr(H_1) + \Pr(H_2) = 1$. Fig. 10 depicts the probability of

having a social vehicle to which to forward a packet. The vehicle is then considered to be socially next-hop eligible if its probability in Eq. (3) is higher than a given threshold (*e.g.*, > 0.7). This occurs for different values of the social activity parameter in H_1 and H_2. For instance, according to Fig. 10 assuming $a_{H_1} = 0.2$ as the activity degree in H_1, only the vehicle showing the lowest degree of activity in H_2 (*i.e.*, $a_{H_2} = 0.01$) is considered a social vehicle since its social degree[3] is higher than 0.7. Literally speaking, $a_{H_2} = 0.01$ means that the activity degree in VSN$_2$ is very low and the vehicle is likely not social in VSN$_2$. As a consequence, for $a_{H_2} = 0.01$ it is expected that the social activity of the vehicle in VSN$_1$ will be higher and therefore, the probability that it is a social vehicle in VSN$_1$ will be higher.

A. Analytical Performance Evaluation

We now turn to the assessment of the performance of the SCARF approach. We consider the scenario where an information packet propagates from a source vehicle to a destination one. Let us denote N_z as the number of potential forwarder vehicles in the transmission range $(0, z)$ [m] of a source vehicle. Each of these N_z vehicles can be elected as a social next-hop forwarder, according to the probability in Eq. (8). The variable associated to the event that a vehicle is social, and then rebroadcasts, or not, is a Bernoulli random one, which can be expressed as:

$$V_i = \begin{cases} 1 & \text{if the } i\text{-th vehicle rebroadcasts} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Since N_z is Poisson distributed with parameter ρz, we can define the average number of social vehicles in $(0, z)$ that rebroadcast the packet as:

$$V_{avg} = \sum_{i=1}^{N_z} V_i = \sum_{n=1}^{\infty} \mathbb{E}\left[\sum_{i=1}^{N_z} V_i \middle| N_z = n\right] \Pr\{N_z = n\} =$$
$$= \sum_{n=1}^{\infty} \sum_{i=1}^{n} \mathbb{E}\left[V_i \middle| N_z = n\right] \Pr\{N_z = n\} =$$
$$= \sum_{n=1}^{\infty} \frac{(\rho z)^n}{n!} \exp(-\rho z) \sum_{i=1}^{n} \mathbb{E}[V_i], \tag{13}$$

where we applied the law of total probability. Furthermore, since V_i is a Bernoulli random variable, we have

$$\mathbb{E}[V_i] = \Pr\{V_i = 1\}, \tag{14}$$

where $\Pr\{V_i = 1\}$ is the probability that the i-th vehicle is socially eligible and rebroadcasts. We assume that the i-th vehicle is within the transmission range of a source vehicle and is therefore potentially selectable as a vehicle forwarder. As previously stated, the forwarding probability depends mainly on social features (*i.e.*, the degree of social activity of the i-th vehicle). So we can define

$$\Pr\{V_i = 1 | S_i \geq a\} = \exp\left(-\frac{\rho(z - \tau)}{c_i \cdot a}\right), \tag{15}$$

[3]Notice that from Eq. (3) the social degree represents a probability, and so the terms "probability of social vehicle" and "social degree" of the i-th vehicle are used interchangeably in this paper.
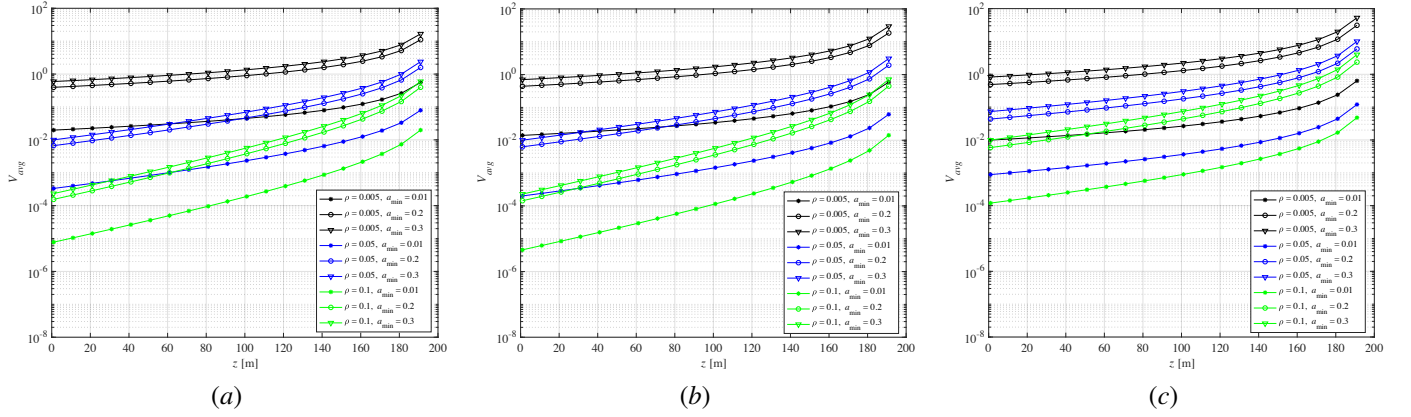
Fig. 11: Average number of forwarder vehicles inside a transmission range, for different values of ρ and $a_{\min}$, with c_i as a continuous uniform random variable in the range $[1, 4]$, and (a) $\alpha = 2$, (b) $\alpha = 2.15$, and (c) $\alpha = 2.3$.

as the probability that the i-th vehicle rebroadcasts, which depends on whether it has a social degree at least equal to a (with $a \leq 1$). In Eq. (15), we have assumed that the i-th vehicle is at distance τ [m] with $\tau \leq z$.

By applying the total probability theorem, we can expand Eq. (15) as:

$$\Pr\{V_i = 1\} = \int_0^1 \Pr\{V_i = 1 | S_i \geqslant a\} f_{S_i}(a)\, da, \quad (16)$$

where $f_{S_i}(\cdot)$ is the PDF of the social parameter S_i, which we assume as a power-law probability distribution [12] expressed as

$$f_{S_i}(a) = \frac{\alpha - 1}{a_{\min}} \cdot \left(\frac{a}{a_{\min}}\right)^{-\alpha}, \quad (17)$$

where $a_{\min}$ represents the minimum value of the activity degree a, and α is the exponent of the power law distribution, which is in the range $[2, 3]$ for small-world networks [48].

Substituting Eq. (17) into (16), we get:

$$\Pr\{V_i = 1\} = \int_0^1 \exp\left[-\rho\left(\frac{z - \tau}{c_i \cdot a}\right)\right] \cdot \frac{\alpha - 1}{a_{\min}} \cdot \left(\frac{a}{a_{\min}}\right)^{-\alpha} da$$

$$= \frac{\alpha - 1}{a_{\min}} \int_0^1 \exp\left[-\rho\left(\frac{z - \tau}{c_i \cdot a}\right)\right] \cdot \left(\frac{a}{a_{\min}}\right)^{-\alpha} da. \quad (18)$$

No closed-form solution for Eq. (18) exists, but the solutions vary for different values of $a_{\min}$, which represents the minimum value of a i.e., the minimum activity degree experienced by a vehicle. In the computation of the average number of forwarders, we set $a_{\min} = [0.01, 0.2, 0.3]$, which represent the cases of activity degree of 1%, 20%, and 30%, respectively. The following solutions hold i.e.,

$$\Pr\{V_i = 1\} = 100 e^{-4.605\alpha} (\alpha - 1) \left(\frac{c_i}{\rho(z - \tau)}\right)^{(\alpha - 1)} \cdot$$

$$\cdot \Gamma\left[\alpha - 1, \frac{\rho(z - \tau)}{c_i}\right], \quad for\ a_{\min} = 0.01,$$

$$(19)$$

$$\Pr\{V_i = 1\} = 5 e^{-1.609\alpha} (\alpha - 1) \left(\frac{c_i}{\rho(z - \tau)}\right)^{(\alpha - 1)} \cdot$$

$$\cdot \Gamma\left[\alpha - 1, \frac{\rho(z - \tau)}{c_i}\right], \quad for\ a_{\min} = 0.2,$$

$$(20)$$

and

$$\Pr\{V_i = 1\} = 3.33 e^{-1.203\alpha} (\alpha - 1) \left(\frac{c_i}{\rho(z - \tau)}\right)^{(\alpha - 1)} \cdot$$

$$\cdot \Gamma\left[\alpha - 1, \frac{\rho(z - \tau)}{c_i}\right], \quad for\ a_{\min} = 0.3$$

$$(21)$$

where $\Gamma\left[\alpha - 1, \frac{\rho(z - \tau)}{c}\right]$ is the upper incomplete gamma function, defined as:

$$\Gamma\left[\alpha - 1, \frac{\rho(z - \tau)}{c_i}\right] = \int_{\rho(z - \tau)/c_i}^{\infty} t^{\alpha - 2} e^{-t} dt. \quad (22)$$

Finally, substituting specific Eqs. (19)–(21) into Eq. (14) for given values of $a_{\min}$, we can compute the average number of forwarder vehicles.

Fig. 11 depicts the average number of forwarders within a transmission range (i.e., $z = 200$ m) for different values of ρ [veh/m], $a_{\min} = [0.01, 0.2, 0.3]$, and α. Specifically, we consider $\alpha = [2, 2.15, 2.3]$, which corresponds to different values of A, i.e., $A = [1, 0.81, 0.68]$, as shown in Fig. 5. We observe that, as in accordance with the forwarder probability of a vehicle that increases for long distances inside the transmission range, the average number of potential forwarders increases for higher values of the distance. Furthermore, the average number of forwarders increases for higher values of $a_{\min}$. Indeed, $a_{\min}$ represents the minimum value of a experienced by the vehicles inside the transmission range. The higher the $a_{\min}$, the higher the probability of social vehicle (see Fig. 10). We also observe that the higher ρ, the lower the average number of forwarders. Finally, different values of α affect the number of forwarders. For increasing α, the threshold A decreases, meaning that it is more likely to find a socially-eligible vehicle, and therefore the average number of forwarders increases (see Fig. 11(c)).
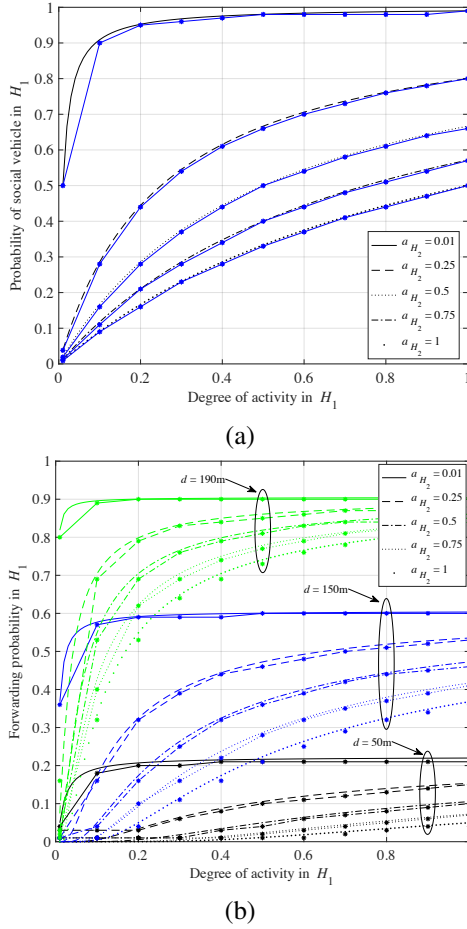
(a)

(b)

Fig. 12: (a) Analytical (*black lines*) and simulated (*blue lines*) results of the probability of social vehicle. (b) Analytical (*no marker lines*) and simulated (*lines with star marker*) results of the probability of forwarding of a vehicle versus the degree of activity under the hypothesis H_1, for different distances and a_{H_2} values.

Fig. 13: Example of traffic scenario with 40 nodes placed along the map of Helsinki city.

V. SIMULATION RESULTS

In order to assess the effectiveness of the SCARF approach, in this section we carry out the comparison between (*i*) the analytical trend of the probability of social vehicle and the probability of forwarding, and (*ii*) the simulation results obtained in real traffic scenarios through The One simulator [55].

We first observe that the probability of a social vehicle in a simulated scenario reflects the trend of Eq. (3), by using values for a_{H_2} equals to 0.01, 0.25, 0.5, 0.75 and 1. Fig. 12(a) plots the comparison between the results of mathematical and simulated results. As we can see, the results are very similar with small differences mainly due to the Java round procedure of The One simulator implementation.

Then, we evaluate the probability of social-based forwarding, as expressed in Eq. (8), versus the degree of activity under the hypothesis H_1 for different values of inter-vehicle distance and degree of activity in H_2, and assuming $z = 200$ [m], $d_i = [50, 150, 190]$ [m], $c_i = 2$, and $\rho = 0.02$ [veh/m]. The computation of S_i is done by setting $A = 0.8$ and considering
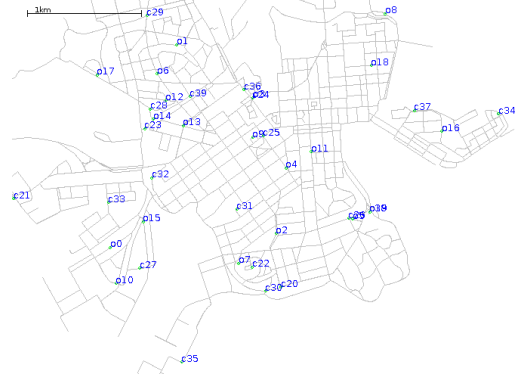
the activity degree as depicted in Fig. 4. Fig. 12(b) plots the comparison between the analytical and the simulated results for the probability of forwarding. Again, we can see that the results are very similar, and the differences between them are well justified by the Java round procedures.

In order to evaluate the performance of the SCARF protocol in a real traffic environment, we implemented a network scenario by using The One Simulator [55]. Comparison results have been made with respect to other related dissemination protocols, such as (*i*) MaxProp [56], (*ii*) Epidemic [57], and (*iii*) Prophet [58]. We also consider SCORP routing protocol [6], which exploits the concept of social proximity and content interests on each node before replicating the message to other nodes in order to improve the network's ability to send messages.

To simulate the node mobility, we use two scenarios with (*i*) real traces and (*ii*) synthetic traces, respectively. The data used in the first scenario –namely, *Rome taxi*– was acquired by an extensive measurement campaign in the city of Rome, in which 370 taxi cabs reported their positions at every 7 seconds for a period of six months [45]. On the other hand, the second scenario –namely, *Helsinki*– comprised of synthetic traces, is based on the one presented in [55], where nodes are placed on the map of Helsinki city (see Fig. 13). We considered a variable number of vehicles (*i.e.*, 20 and 40), in order to assess the performance for different vehicular densities. During the simulated 12 h time period, the vehicles move on the map roads at an average speed of 50 km/h, between random locations, and with random pause times between 5 and 15 minutes. In addition, we divided the scenario into 3 areas, each representing a VSN as the most visited areas during the simulations. Each vehicle randomly generates a number of posts (*i.e.*, in order to grant the participation of all the vehicles in the network, every vehicle generates at least one message per day). Of course, the VSN in which a vehicle can generate a message depends on its location at that moment.

Precision of the results and a manageable set of simulation experiments is a critical compromise to be found. The parameters taken in this paper are based on the experience obtained in previous works (*e.g.*, [59], [60]), where we observed that simulations with a higher number of nodes (*e.g.*, > 1000) is
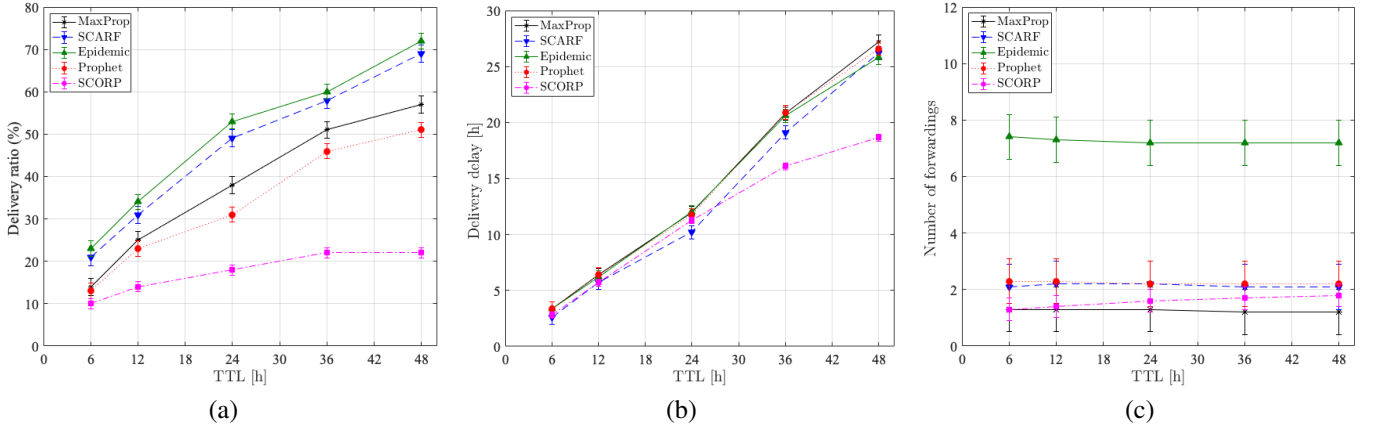
Fig. 14: Synthetic trace scenario with 20 vehicles. Simulation results for (a) delivery ratio, (b) delivery delay, and (c) average cost, versus TTL.
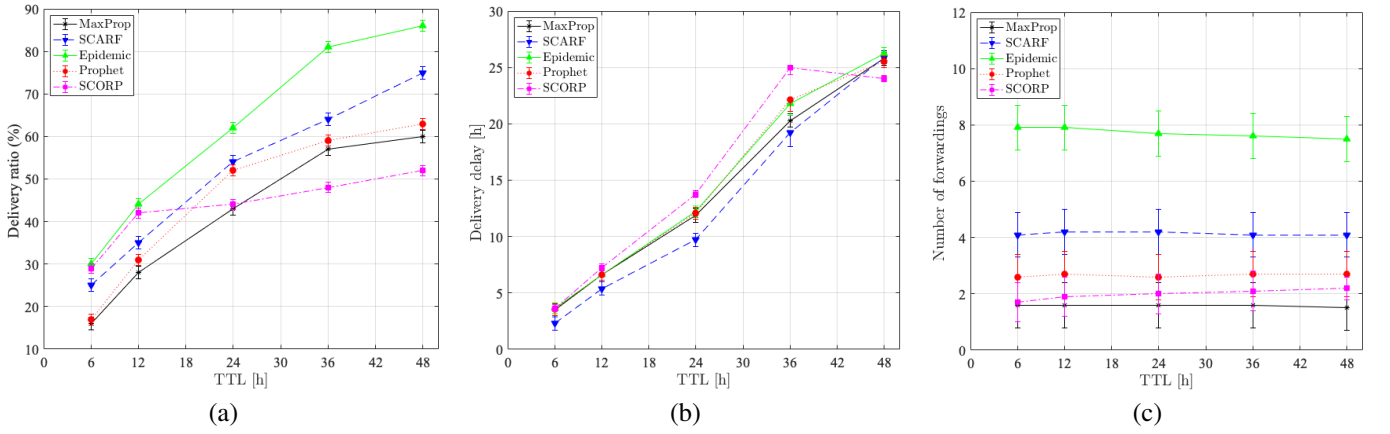

Fig. 15: Synthetic trace scenario with 40 vehicles. Simulation results for (a) delivery ratio, (b) delivery delay, and (c) average cost.

TABLE II: Main parameters used in the simulations.

Message generation rate	1000 msg per simulation
Message size	$[1, 2]$ MB
TTL	$[6, 12, 24, 36, 48]$ h
Network Interface	IEEE 802.11p
Transmission range	200 m

computationally intractable. Moreover, we must remember that our architecture by itself has integrated an edge based solution (see Subsection III-A) that allows the impact of the growth of the network nodes to be minimised. Table II shows the main parameters used in the simulated scenarios, both for real and synthetic traces.

A. Network performance assessment

Network performances are expressed in terms of (*i*) delivery ratio *i.e.*, percentage of how many messages are delivered to the receiver, (*ii*) delivery delay [h] *i.e.*, time needed for a message to be delivered to the receiver, and (*iii*) number of forwardings *i.e.*, number of hops that connect the source vehicle to the final destination. Fig. 14 presents the results in

the Helsinki scenario with 20 vehicles. As we can see from Fig. 14(a), the SCARF algorithm achieves high performance for the metrics of delivery ratio, reaching 68% for the highest Time To Live (TTL), outperformed only by the Epidemic technique –as expected– which reaches 72% of delivery ratio when TTL = 48 h. Fig. 14(b) depicts the delivery delay, which is minimized with SCARF with respect to other approaches for values of TTL less than 24 h, and thus, providing a fast and reliable transmission. On the other hand, for increased TTL, SCORP presents better performance, followed by SCARF. Finally, in Fig. 14(c) we observe lower values in the number of forwardings with SCARF (*i.e.*, around 2), meaning that a low number of hops is necessary to reach a destination. On the other hand, as expected, the Epidemic technique exhibits higher number of forwardings. It is worth mentioning that the SCARF average cost (*i.e.*, the number of forwardings) is very similar to the result achieved by the MaxProp algorithm, thus clearly showing that SCARF technique can contribute to increasing network performance without increasing the overhead.

Similar behaviours are obtained in Fig. 15, which presents the results in the Helsinki scenario with 40 vehicles. Once
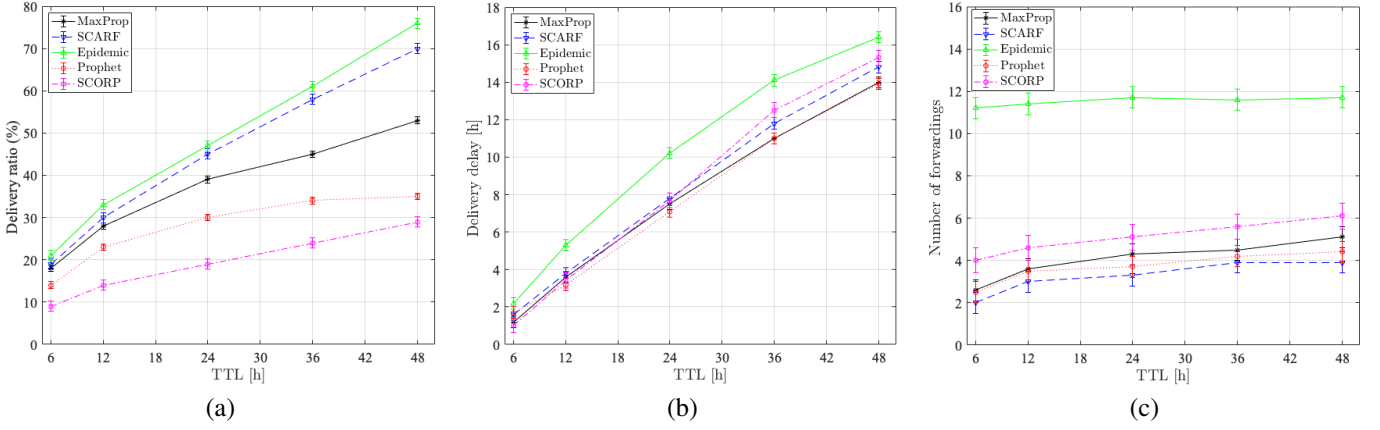
Fig. 16: Real trace scenario with 370 taxis from Roma city [45]. Simulation results for (a) delivery ratio, (b) delivery delay, and (c) average cost.
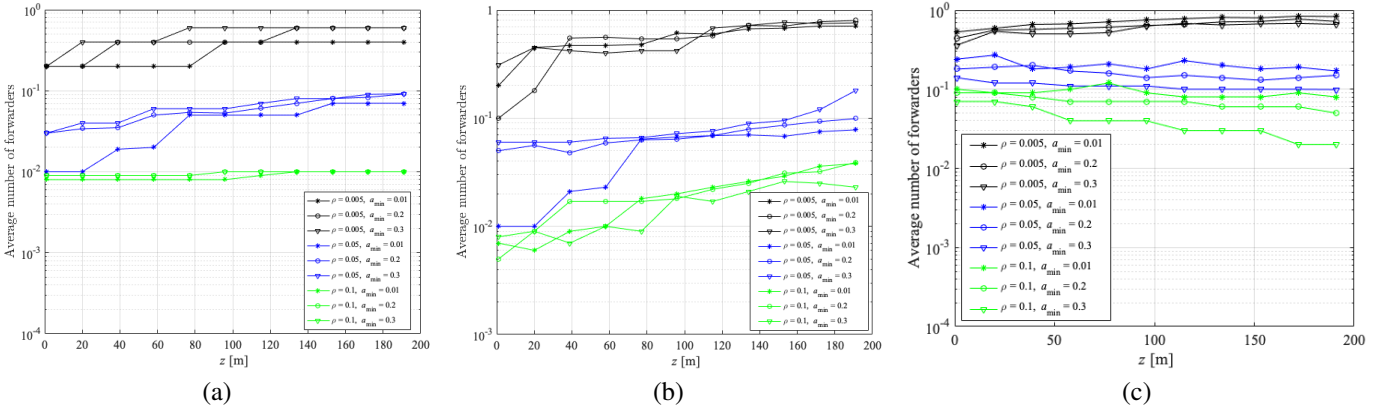

Fig. 17: Simulation results of the average number of forwarders within a source node transmission range, for (a) 20-vehicle synthetic trace (*Helsinki scenario*), (b) 40-vehicle synthetic trace (*Helsinki scenario*), and (c) 370-taxi real trace scenario (*Rome taxi scenario*).

again, SCARF outperforms the MaxProp algorithm for the metrics delivery ratio and delay. For this increasing number of vehicles, SCARF shows lower delays with a small increase of forwardings (*i.e.*, ≈ 4). It is also worth mentioning that SCARF achieves better performance in this scenario as compared to the previous one with 20 vehicles. This might clearly represent that the greater the number of nodes in the network, the bigger the number of possible relays to forward a message (see Fig. 15(c)). It follows that SCARF can increase the message delivery probability, and therefore, the delivery ratio. In this scenario, we observe that the delivery delay for SCARF is very similar to that obtained in the scenario with 20 vehicles, while the number of forwardings is duplicated. As a conclusion, when the number of nodes increases, SCARF presents better performance while still guaranteeing the lowest delivery delay.

Fig. 16 depicts the network performances in the case of a real trace scenario (*i.e.*, namely Rome Taxi), as described in [45]. In Fig. 16(a), we observe a high delivery ratio for increased TTLs for SCARF, approaching the Epidemic technique. On the other hand, for the delivery delay, in Fig. 16(b), we notice the best performance come from Prophet and MaxProp, followed by SCARF with slightly a higher delay. However, as compared to the results from the scenario with 40 vehicles in Fig. 15, where the delivery delay reaches ≈ 25 h, here the delivery delay gives back lower values (*i.e.*, ≈ 16 h as maximum value). Finally, in Fig. 16(c), we observe the number of forwardings is the lowest with SCARF, confirming its lower overhead.

This latter behaviour lies in contrast with the results from the synthetic traces, where we noticed low delays with an increasing number of forwardings with SCARF. On the other end, in the real traffic scenario, SCARF still shows low delays but with a limited number of forwardings. This can be justified by the nature of vehicles, as taxis cross the whole city and perform random trajectories, without fixed time and trip duration, while common people (*i.e.*, vehicles in the synthetic traces) use their vehicles to perform predetermined trajectories according to their routines.

Finally, in Fig. 17, we present the average number of forwarders within a source node transmission range (*i.e.*, $z = 200$ m) and for different values of ρ and $a_{\min}$. Common
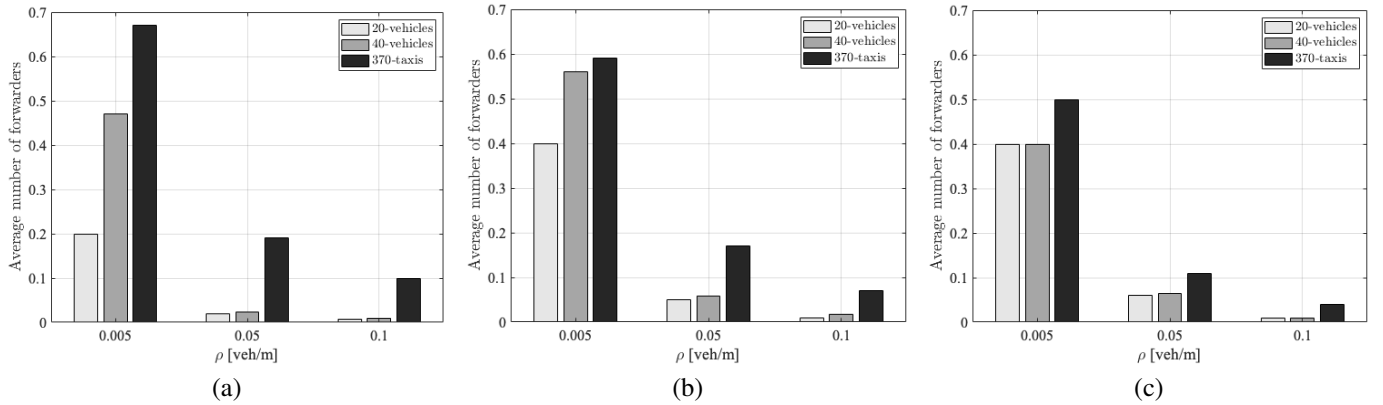
Fig. 18: Simulation results of the average number of forwarders for a fixed distance from a source node, in case of (a) $a_{\min} = 0.01$, (b) $a_{\min} = 0.2$, and (c) $a_{\min} = 0.3$ and different vehicular scenarios.

aspects to all the three figures are that (*i*) better performances are experienced when the vehicular density is low (*black lines*), and (*ii*) the average number of forwarders is ≈ 1. This aspect highlights SCARF's ability to select as a next-hop forwarder that vehicle showing a given social activity. Specifically, in Fig. 17(a), we notice that the average number of selected forwarders is higher when $a_{\min} = 0.3$. Moreover, in the case of $\rho = 0.1$ (*green lines*), there is no particular difference among different values of $a_{\min}$. Additionally, we can distinguish a slight increasing slope for increasing inter-vehicular distance. Similar considerations can be made when observing Fig. 17(b), in the case of a scenario with 40 vehicles. Different behaviour is experienced in Fig. 17(c) for the 370-taxi real trace scenario. We observe better performance for lowest values of $a_{\min}$ (*i.e.*, $a_{\min} = 0.01$), and the trend appears to be decreasing in the case of higher vehicular density (*green* and *blue lines*). Once again, this particular behaviour can be explained by the specific nature of the vehicles in this scenario.

From the last results, in Fig. 18 we show the average number of forwarders within a transmission range, computed for a fixed distance from the source node (*i.e.*, $d \approx 60$ m) and different values of $a_{\min}$, versus the vehicular density ρ, in case of different vehicular scenarios. As reflected also in Fig. 17, the average number of forwarders decreases for increasing vehicular density values. It allows to state that in case of high density scenarios, the number of transmissions will not be overloaded, thus saving the traffic load within the network.

Finally, a few considerations regarding the computational cost of SCARF. From Eq. 13, we observe that the average number of social vehicles V_{avg} depends on the number of potential forwarders N_z within a transmission range. It follows that SCARF's computational cost is expressed as $O(N_z)$.

VI. Conclusions

In this paper we evaluated the performance of a protocol we designed called SCARF in realistic scenarios with traffic traces and compared it to other similar techniques. SCARF is a reliable, socially-aware forwarding technique for vehicular communications that relies on the concept of a social hub node, that is, a node with social addiction to sharing messages

within a vehicular social network. Selecting a hub node as a next-hop forwarder allows more nodes to be reached, thus allowing for a more reliable data dissemination technique.

The next-hop selection occurs through the computation of a node social degree. An edge-based network architecture has been proposed to cache the information needed for the graph metric close to the mobile users by decreasing the delay of access and avoiding resource waste. Indeed, this stored content is effectively exploited to "categorise" the nodes from a social point of view.

Unlike forwarding techniques used in traditional opportunistic networks, SCARF is well suited to socially-aware vehicular scenarios. As a result, we showed that SCARF outperforms other approaches in terms of delivery ratio, while guaranteeing acceptable time delay values. We moreover analytically demonstrated its effectiveness in packet transmission reduction, that is, reducing the overhead, while guaranteeing network dissemination.

References

[1] A. M. Vegni and V. Loscrí, "A survey on vehicular social networks," *IEEE Comm. Surveys & Tutorials*, vol. 17, no. 4, pp. 2397–2419, Fourthquarter 2015.

[2] A. Srivastava, Anuradha, and D. Gupta, "Social network analysis: Hardly easy," in *Proc. of Intl. Conf. on Optimization, Reliabilty, and Information Technology (ICROIT)*, 2014, pp. 128–135.

[3] K. Wei, X. Liang, and K. Xu, "A survey of social-aware routing protocols in delay tolerant networks: Applications, taxonomy and design-related issues," *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 556–578, First 2014.

[4] F. Xia, L. Liu, J. Li, A. M. Ahmed, L. T. Yang, and J. Ma, "Beeinfo: Interest-based forwarding using artificial bee colony for socially aware networking," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 3, pp. 1188–1200, March 2015.

[5] P. Costa, C. Mascolo, M. Musolesi, and G. P. Picco, "Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 748–760, June 2008.

[6] W. Moreira, P. Mendes, and S. Sargento, "Social-aware opportunistic routing protocol based on users interactions and interests," *Ad Hoc Networks*, vol. 129, pp. 100–115, 2014.

[7] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: Social-based forwarding in delay-tolerant networks," *Mobile Computing, IEEE Transactions on*, vol. 10, no. 11, pp. 1576–1589, 2011.

[8] A. M. Vegni, V. Loscrí, and R. Petrolo, "Scarf: A social-aware reliable forwarding technique for vehicular communications," in *Proceedings of the 3rd Workshop on Experiences with the Design and Implementation of Smart Objects*, ser. SMARTOBJECTS '17. Snowbird, Utah, USA: ACM, 2017, pp. 1–6. [Online]. Available: http://doi.acm.org/10.1145/3127502.3127512

[9] A. Guille, H. Hacid, C. Favre, and D. A. Zighed, "Information diffusion in online social networks: A survey," *SIGMOD Rec.*, vol. 42, no. 2, pp. 17–28, Jul. 2013. [Online]. Available: http://doi.acm.org/10.1145/2503792.2503797

[10] M. U. Ilyas, M. Z. Shafiq, A. X. Liu, and H. Radha, "A distributed algorithm for identifying information hubs in social networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 629–640, September 2013.

[11] M. Fiore and J. Härri, "The networking shape of vehicular mobility," in *Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '08. New York, NY, USA: ACM, 2008, pp. 261–272.

[12] F. D. da Cunha, A. C. Viana, R. A. F. Mini, and A. A. F. Loureiro, "How effective is to look at a vehicular network under a social perception?" in *Proc. of IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, October 2013, pp. 154–159.

[13] X. Gu, L. Tang, and J. Han, "A social-aware routing protocol based on fuzzy logic in vehicular ad hoc networks," in *2014 International Workshop on High Mobility Wireless Communications*, Nov 2014, pp. 12–16.

[14] A. Bradai and T. Ahmed, "Reviv: Selective rebroadcast mechanism for video streaming over vanet," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014, pp. 1–6.

[15] A. Stagkopoulou, P. Basaras, and D. Katsaros, "A social-based approach for message dissemination in vehicular ad hoc networks," in *Ad Hoc Networks*, N. Mitton, A. Gallais, M. E. Kantarci, and S. Papavassiliou, Eds. Cham: Springer International Publishing, 2014, pp. 27–38.

[16] A. M. Vegni, V. Loscrí, and P. Manzoni, "Data forwarding techniques based on graph theory metrics in vehicular social networks," in *Proc. of IEEE 29th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2018, pp. 1771–1775.

[17] N. Lu, T. H. Luan, M. Wang, X. Shen, and F. Bai, "Bounds of asymptotic performance limits of social-proximity vehicular networks," *IEEE/ACM Transactions on Networking*, vol. 22, no. 3, pp. 812–825, June 2014.

[18] T. H. Luan, X. S. Shen, F. Bai, and L. Sun, "Feel bored? join verse! engineering vehicular proximity social networks," *IEEE Transaction on Vehicular Technology*, vol. 64, no. 3, pp. 1120–1131, March 2015.

[19] V. R. Neto, D. S. V. Medeiros, and M. E. M. Campista, "Analysis of mobile user behavior in vehicular social networks," in *Proc. of 2016 7th International Conference on the Network of the Future (NOF)*, Nov 2016, pp. 1–5.

[20] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '11. New York, NY, USA: ACM, 2011, pp. 1082–1090. [Online]. Available: http://doi.acm.org/10.1145/2020408.2020579

[21] N. Abbani, T. Tarhini, H. Artail, and W. El-Hajj, "Managing social networks in vehicular networks using trust rules," in *Proceedings of IEEE Symposium on Wireless Technology and Applications (ISWTA)*, Sept. 2011.

[22] S. Smaldone, L. Han, P. Shankar, and L. Iftode, "Enabling voice chat on roadways using vehicular social networks," in *Proceedings of the 1st Workshop on Social Network Systems, SocialNets*, 2008.

[23] ——, "Roadspeak: Enabling voice chat on roadways using vehicular social networks," in *Proceedings of the 1st Workshop on Social Network Systems*, ser. SocialNets '08. New York, NY, USA: ACM, 2008, pp. 43–48. [Online]. Available: http://doi.acm.org/10.1145/1435497.1435505

[24] S. Maaroufi, "A dynamic messaging architecture for vehicular social networks," in *PhD Thesis, Ecole polytechniuqe de Montreal*, 2016.

[25] Z. Ning, F. Xia, N. Ullah, X. Kong, and X. Hu, "Vehicular social networks: Enabling smart mobility," *IEEE Communication Magazine*, vol. 55, no. 5, pp. 49–55, May 2017.

[26] Q. Yang, B. Zhu, and S. Wu, "An architecture of cloud-assisted information dissemination in vehicular networks," *IEEE Access*, vol. 4, pp. 2764–2770, May 2016.

[27] X. Hu, V. C. Leung, K. G. Li, E. Kong, H. Zhang, N. S. Surendrakumar, and P. TalebiFard, "Social drive: A crowdsourcing-based vehicular social networking system for green transportation," in *Proceedings of the Third ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*, ser. DIVANet '13. New York, NY, USA: ACM, 2013, pp. 85–92. [Online]. Available: http://doi.acm.org/10.1145/2512921.2512924

[28] W. Sha, D. Kwak, B. Nath, and L. Iftode, "Social vehicle navigation: Integrating shared driving experience into vehicle navigation," in *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications*, ser. HotMobile '13. New York, NY, USA: ACM, 2013, pp. 16:1–16:6. [Online]. Available: http://doi.acm.org/10.1145/2444776.2444798

[29] B. E. Kolko, E. Johnson, and E. Rose, "Mobile social software for the developing world," in *Online Communities and Social Computing*, D. Schuler, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 385–394.

[30] J. Quain, "Social networking for cars," 2012. [Online]. Available: http://wheels.blogs.nytimes.com/2010/07/20/socialnetworking-for-cars/

[31] W. Sha, D. Kwak, B. Nath, and L. Iftode, "Social vehicle navigation: Integrating shared driving experience into vehicle navigation," in *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications*, ser. HotMobile '13. New York, NY, USA: ACM, 2013, pp. 16:1–16:6. [Online]. Available: http://doi.acm.org/10.1145/2444776.2444798

[32] E. Bulut and M. C. Kisacikoglu, "Mitigating range anxiety via vehicle-to-vehicle social charging system," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, June 2017, pp. 1–5.

[33] N. Ullah, X. Kong, L. Wan, H. Chen, Z. Wang, and F. Xia, "A social utility-based dissemination scheme for emergency warning messages in vehicular social networks," *The Computer Journal*, vol. 61, no. 7, pp. 971–986, 2018. [Online]. Available: http://dx.doi.org/10.1093/comjnl/bxy026

[34] A. Rahim, T. Qiu, Z. Ning, J. Wang, N. Ullah, A. Tolba, and F. Xia, "Social acquaintance based routing in vehicular social networks," *Future Generation Computer Systems*, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167739X17304326

[35] G. Sun, Y. Zhang, D. Liao, H. Yu, X. Du, and M. Guizani, "Bus trajectory-based street-centric routing for message delivery in urban vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2018.

[36] A. Rahim, X. Kong, F. Xia, Z. Ning, N. Ullah, J. Wang, and S. K. Das, "Vehicular social networks: A survey," *Pervasive and Mobile Computing*, vol. 43, pp. 96 – 113, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574119217300792

[37] H. Nguyen-Minh, A. Benslimane, and M. Radenkovic, "Social delay tolerant approach for safety services in vehicular networks," in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, Aug 2015, pp. 1199–1204.

[38] R. Umesh, "A technique for cluster formation," *Pattern Recognition*, vol. 21, no. 4, pp. 393 – 400, 1988. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0031320388900520

[39] S. Busanelli, G. Ferrari, and S. Panichpapiboon, "Efficient broadcasting in ieee 802.11 networks through irresponsible forwarding," in *GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference*, Nov 2009, pp. 1–6.

[40] F. J. Martinez, M. Fogue, M. Coll, J.-C. Cano, C. T. Calafate, and P. Manzoni, "Evaluating the impact of a novel warning message dissemination scheme for vanets using real city maps," in *NETWORKING 2010*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 265–276.

[41] S. M. Abuelenin and A. Y. Abul-Magd, "Moment analysis of highway-traffic clearance distribution," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2543–2550, Oct 2015.

[42] S. Durrani, X. Zhou, and A. Chandra, "Effect of vehicle mobility on connectivity of vehicular ad hoc networks," in *2010 IEEE 72nd Vehicular Technology Conference - Fall*, Sep. 2010, pp. 1–5.

[43] A. Soua, W. Ben-Ameur, and H. Afifi, "Analysis of information relay processing in inter-vehicle communication: A novel visit," in *2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct 2012, pp. 157–164.

[44] A. N. Hassan, O. Kaiwartya, A. H. Abdullah, D. K. Sheet, and R. S. Raw, "Inter vehicle distance based connectivity aware routing in vehicular adhoc networks," *Wireless Personal Communications*, vol. 98, no. 1, pp. 33–54, Jan 2018.

[45] R. Amici, M. Bonola, L. Bracciale, A. Rabuffi, P. Loreti, and G. Bianchi, "Performance Assessment of an Epidemic Protocol in VANET Using Real Traces," in *Proc. of International Conference on Selected Topics in Mobile and Wireless Networking (MoWNeT)*, vol. 40, September 2014, pp. 92–99.

[46] N. Lütteken, M. Zimmermann, and K. J. Bengler, "Using gamification to motivate human cooperation in a lane-change scenario," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Nov 2016, pp. 899–906.

[47] R. Drakoulis, F. Bellotti, I. Bakas, R. Berta, P. K. Paranthaman, G. R. Dange, P. Lytrivis, K. Pagle, A. D. Gloria, and A. Amditis, "A gamified flexible transportation service for on-demand public transport," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3, pp. 921–933, March 2018.

[48] A.-L. Barabási, *Network Science*. Cambridge University Press, 2016.

[49] P. Manzoni, E. Hernández-Orallo, C. T. Calafate, and J.-C. Cano, "A proposal for a publish/subscribe, disruption tolerant content island for fog computing," in *Proc. of the 3rd Workshop SMARTOBJECTS*. ACM, 2017, pp. 47–52.

[50] J. A. Khan, Y. Ghamri-Doudane, and D. Botvich, "Inforank: Information-centric autonomous identification of popular smart vehicles," in *2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*, Sept 2015, pp. 1–6.

[51] J. A. Khan and Y. Ghamri-Doudane, "Car rank: An information-centric identification of important smart vehicles for urban sensing," in *2015 IEEE 14th International Symposium on Network Computing and Applications*, Sept 2015, pp. 184–191.

[52] J. A. Khan, Y. Ghamri-Doudane, and D. Botvich, "GRank - An Information-Centric Autonomous and Distributed Ranking of Popular Smart Vehicles," in *2015 IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–7.

[53] J. A. Khan and Y. Ghamri-Doudane, "Saving: Socially aware vehicular information-centric networking," *IEEE Communications Magazine*, 2016.

[54] M. Jalali and N. G. Aghaee, "A fuzzy reputation system in vehicular ad hoc networks," *Procedia Computer Science*, vol. 5, pp. 951 – 956, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050911004595

[55] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE Simulator for DTN Protocol Evaluation," in *SIMUTools '09: Proceedings of the 2nd International Conference on Simulation Tools and Techniques*. New York, NY, USA: ICST, 2009.

[56] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine, "Maxprop: Routing for vehicle-based disruption-tolerant networks," in *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, April 2006, pp. 1–11.

[57] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," Duke University, Tech. Rep., April 2000.

[58] C. Liu and J. Wu, "An optimal probabilistic forwarding protocol in delay tolerant networks," in *Proceedings of the Ten-th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '09. New York, NY, USA: ACM, 2009, pp. 105–114. [Online]. Available: http://doi.acm.org/10.1145/1530748.1530763

[59] M. Báguena, C. T. Calafate, J. Cano, and P. Manzoni, "Towards realistic vehicular network simulation models," in *2012 IFIP Wireless Days*, Nov 2012, pp. 1–3.

[60] J. Dede, A. Förster, E. Hernández-Orallo, J. Herrera-Tapia, K. Kuladinithi, V. Kuppusamy, P. Manzoni, A. bin Muslim, A. Udugama, and Z. Vatandas, "Simulating opportunistic networks: Survey and future directions," *IEEE Communications Surveys Tutorials*, vol. 20, no. 2, pp. 1547–1573, Secondquarter 2018.