

## **Social Evolving Graph-Based Connectivity Model for Vehicular Social Networks**

**Mahmoud Hashem Eiza and Qi Shi**

### **Abstract**

Over the last two decades, social networking over the Internet has attracted the attention of millions of users and becomes a new factor that drives business intelligence and economic growth worldwide. People have strong social instincts and they tend to socialise and communicate with each other in every possible scenario thus, social networks have been extended to different types and forms beyond the Internet. Mobile Social Networks (MSN) represent a new form of social networks that take advantage of the close proximity and the unique features offered by mobile devices to establish social links among mobile users. More recently, Vehicular Social Networks (VSN), a special type of MSNs, have emerged as a new communication paradigm for social networking on the roads. VSNs are decentralised opportunistic communication networks formed among vehicles (Vegni and Loscri 2015), where the communication takes place in three dimensions: human to human, human to machine, and machine to machine communications. The flourish of vehicular networks development over the last decade has made the social interaction on the roads possible promising more enjoyable experience for vehicular travellers.

Prior to discussing VSNs and their connectivity patterns, it is important to develop an understating of vehicular networks on one hand and the human social behaviour in VSNs from a social theory viewpoint on the other hand. Vehicular networks are characterised with high mobility and frequent network topology changes, which make the connectivity between vehicles tends to be vulnerable. On the other side, in addition to being within the communication range of each other, sharing the same social interests and/or characteristics is a prerequisite for any two vehicles to socially interact in VSNs.

This chapter is intended to present a novel social connectivity model for VSNs by utilising the evolving graph theory. First, the properties of VSNs are briefly introduced in the context of vehicular networks. The social metrics of the communicating vehicles are then reviewed using the concepts of social theory along with the conventional connectivity metrics in vehicular networks. Thereafter, a novel social evolving graph-based connectivity model that considers both social and conventional metrics of the communicating vehicles is developed using an extended version of the evolving graph. Moreover, the proposed connectivity model suggests new social links with vehicles that enter the communication area of other vehicles with similar social interests. Finally, the developed connectivity model is investigated in a highway scenario to demonstrate its abilities in capturing the evolving characteristics of social interactions among vehicles and selecting the best paths to forward data. Data forwarding decisions are made based on a combination of social and communication metrics of the communicating vehicles. Simulation results showed that the proposed connectivity model facilitates the social interactions among vehicles and is able to establish reliable social paths among the communicating vehicles.

## **Introduction**

Nowadays, social networking over the Internet has become one of the most popular methods for social interactions among people thanks to the modern and ubiquitous communication technologies/devices. Besides the traditional online social networks, which are offered by service providers such as Facebook, Twitter, LinkedIn, etc., mobile social networks (MSN) have emerged as a new platform over which participants interact within a virtual social network using their mobile devices. These mobile devices take advantage of their close proximity and leverage different communication technologies such as Bluetooth, Wi-Fi Direct, etc. Thus, MSNs offer the possibility of opportunistic social interaction where opportunistic networking is utilised to allow

each node to send, receive and relay information without a server dictating the communications. This feature makes MSNs an attractive option for supporting social interactions and collaborations among people in a number of mobile environments where MSN can take advantage of both infrastructure-based wireless networks, e.g. the mobile Internet, and opportunistic networks, e.g., wireless mobile ad-hoc networks (Hu, et al. 2015).

Vehicular social networks (VSN) are one of the main application domains of MSNs. VSNs are defined as decentralised opportunistic communication networks that facilitate social interactions including content creating and sharing between travellers on roadways. Due to the lack of high rate Internet connections on roadways, especially on highways and rural areas, VSNs encourage vehicles' travellers to create, share and relay information using the available low-cost communication links in vehicular networks including Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications. Direct inquiry of others with similar experience in proximity over social networks tends to be the most convenient and efficient approach to acquire an up-to-date, specialised and domain-specific content and information for travellers (Luan, Lu, et al. 2015). Furthermore, recent TripAdvisor survey of more than 1700 US respondents reveals that 76% of travellers share their travel experience including photos and clips via social networks and 52% do that while travelling/driving back home (TripAdvisor, Inc. 2012). Thus, VSNs represent a unique form of localised mobile social network that exploit the vehicular communication links and offer vehicular travellers the opportunity to engage in social activities along the road.

Given the unique features it provide, VSNs can serve as a platform for various vehicular and traffic related applications. Therefore, VSNs have received more attention and research efforts from academia and industry worldwide. These efforts resulted in developing many applications and frameworks that can operate upon VSNs. RoadSpeak (Smaldone, et al. 2008), Verse (Luan,

Shen, et al. 2015), Clique Trip (Knobel, et al. 2012), NaviTweet (Sha, et al. 2013), and Toyota Friend (Kageyama 2011) are few examples of these applications.

While VSNs promise a new communication platform for social interactions along the roads, it inherits the connectivity problems that already exist in vehicular networks. This includes the high mobility of network nodes and the frequent changes of the network topology. In vehicular networks, the network topology could vary when vehicles change their velocities and/or lanes. These changes depend on the drivers' behaviours, i.e., human factors, and road situations and are normally not scheduled in advance. Here, we assume that vehicles are driven by humans. Self-driving vehicles can be considered as part of a VSN however, this case needs more investigation and is left for future work. Thus, in other words, the VSN can be defined as a vehicular network that takes the social characteristics of human beings such as human mobility, human selfish status and human preferences into account. Therefore, the current connectivity models, which are designed for vehicular networks, cannot guarantee to capture the social evolving connectivity patterns in VSNs. This problem is the subject of this chapter.

### **Basics of Social Theory**

As we have mentioned above, the human factor has a significant impact on the operations and consequently the performance of VSNs. The human factor in VSNs can be considered from two different points of view: the passengers' social behaviours and the drivers' social behaviours since driving itself has been constructed as a set of social practices, embodied disposition, cybernetic associations and physical affordances (Smyth and King 2006). However, in this chapter, we consider the social behaviour of travellers in general, i.e., both drivers and passengers, and the social aspects of vehicles as network nodes in the social network. Considering different behaviours of different travellers in the same vehicle in the context of VSNs is an open research challenge and

is left for future work.

In social theory, there are several indexes that can be used to localise the most significant nodes and quantify their relative importance to other nodes (Batallas and Yassine 2006, Lu 2012). These indexes and measures are similar to those utilised in graph theory since the social network itself represents a communication graph. Thus, in this section, we follow the categorisation of graph theory to these social indexes and metrics (Pallis, et al. 2009) which are concerned with local, network-wide and community-wide metrics in VSN. In the following, we describe these measures along with the model of social morality of vehicular travellers.

### ***Local Metrics***

#### *Propinquity*

Under equal conditions, propinquity means that if two vehicles are geographically near to each other, they are more likely to be connected.

#### *Homophily*

In social theory, homophily is defined as the common social attributes, i.e., the similarity, between two users such as having the same favourites, working for the same organisation, having the same travelling destination, etc. Thus, it is more likely that travellers with the same social attributes, i.e., high homophily, to have a connection. Thus, the higher the homophily the more likely two vehicles will be socially connected (Lu 2012).

Let  $HP_i$  be the social attributes of an entity  $n_i$  where each item in  $HP_i$  is a binary variable that indicates whether  $n_i$  has an interest in the corresponding item or not. For instance, let assume the following social attributes/interests {Football, Rap music, Thai food, Mountain climbing} and  $HP_i = \{1, 1, 1, 0\}$ , i.e.,  $n_i$  likes watching football, listening to rap music and eating Thai food but he/she does not like mountain climbing. In order to match the similarity between two entities in terms of

their social attributes/interests, we adopt the vector space model (VSM) as described in (Li, Guo and Zhao 2008). Let  $HP_i = \{S_{i,x}\}$  where  $x \in 1 \dots n$  and  $S_{i,x} \in \{0,1\}$  be the social attributes/interests profile of  $n_i$ . The homophily between two entities  $n_i$  and  $n_j$ , i.e., the social interests similarity, can be evaluated as follows (Luan, Lu, et al. 2015):

$$SHP_{ij} = \frac{\sum_{k=1}^n S_{i,k} S_{j,k}}{\sqrt{\sum_{q=1}^n S_{i,q}^2} \cdot \sqrt{\sum_{y=1}^n S_{j,y}^2}} \quad (1)$$

If  $SHP_{ij} = 1$ , then users  $n_i$  and  $n_j$  have exactly the same social attributes/interests and consequently they are more likely to socially communicate, i.e., create a social connection in VSNs. Otherwise, if  $SHP_{ij} = 0$ , then  $n_i$  and  $n_j$  have no interests in common and it is unlikely for them to have a social connection in VSNs. The evaluation of the homophily  $SHP_{ij}$  factor between two travellers in two different vehicles is the first step to determine the likelihood of establishing a social connection between them in the context of VSN.

#### *Degree Centrality*

In definition, a central node is the one that relates to a large number of nodes in the network, i.e., have a large number of in-links and out-links with other nodes. The degree of a node  $n_i$  can be measured by counting the number of links incident to it and is represented by  $d(n_i)$  (Snijders and Borgatti 1999). Since the distinction between in-links and out-links in social networks is not needed, the centrality of a node  $n_i$ , denoted as  $C_D(n_i)$ , can be calculated as follows:

$$C_D(n_i) = d(n_i) = \sum_{\forall i \neq j} x_{ij} \quad (2)$$

where  $x_{ij} = 1$  if  $i$  is incident to  $j$  and  $x_{ij} = 0$  otherwise. It can be noted that  $C_D(n_i)$  depends on the size of the network and it becomes complex to use when comparing different networks. Let  $N$  be the total number of nodes in the network, one way to standardise the degree centrality metric  $C_D(n_i)$

is to divide (2) by the maximum number of nodes that  $n_i$  can be connected to as follows:

$$C'_D(n_i) = d(n_i) = \frac{\sum_{\forall i \neq j} x_{ij}}{(n-1)} \quad (3)$$

In the context of VSNs, choosing nodes with high degree centrality index to forward a message means that the chance of delivering this message to its destination will be high.

### *Social Link Duration*

In order to have a social link  $SL_{ij}$  between two nodes  $n_i$  and  $n_j$ , a communication link  $l_{ij}$  should first exist, i.e., both vehicles should be within the transmission range of each other. Since the social attributes of travellers are less likely to change over the road, i.e., their homophily, the social link duration mainly depends on the communication link duration between two vehicles. Let  $H$  denote the wireless transmission range and  $v_i(t)$  and  $v_j(t)$  the velocities of  $n_i$  and  $n_j$  at time  $t$ , respectively, the social link duration  $SL_{ij}(t)$  can be accurately estimated as follows:

$$SL_{ij}(t) = \frac{H - \theta \sqrt{(y_i(t) - y_j(t))^2 + (x_i(t) - x_j(t))^2}}{|v_i(t) - \vartheta v_j(t)|} \quad (4)$$

where  $\theta = -1$  and  $\vartheta = 1$  when  $n_j$  overtakes  $n_i$ ,  $\theta = 1$  and  $\vartheta = 1$  when  $n_i$  moves forward in front of  $n_j$ ,  $\theta = -1$  and  $\vartheta = -1$  when  $n_i$  and  $n_j$  are moving toward each other, and  $\theta = 1$  and  $\vartheta = -1$  when  $n_i$  and  $n_j$  are moving away from each other.

However, the calculation of  $SL_{ij}(t)$  in (4) does not take into consideration the possible changes in vehicles' velocities values. Therefore, we utilise the concept of link reliability, which is introduced in (Eiza and Ni 2013), to accurately estimate the expected social link duration. The link reliability is defined as the probability that the communication link between two nodes  $n_i$  and  $n_j$  will stay continuously available over a specified time period. Given  $SL_{ij}(t)$ , the estimated duration for the continuous availability of a social link  $SL_{ij}$  between two vehicles at time  $t$  as calculated in (4), the

link reliability value  $r_t(SL_{ij})$  is expressed as follows:

$$r_t(SL_{ij}) = P\{\text{To continue to be available until } t + SL_{ij} \mid \text{available at } t\}$$

We assume that the velocity of vehicles has a normal distribution (Niu, et al. 2006, Schnabel and Lohse 1997). Thus,  $r_t(SL_{ij})$  can be calculated as follows:

$$r_t(SL_{ij}) = \begin{cases} \int_t^{t+SL_{ij}} f(T) dT & \text{if } SL_{ij} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where  $f(T)$  is defined as follows:

$$f(T) = \frac{4H}{\sigma_{\Delta v} \sqrt{2\pi}} \frac{1}{T^2} e^{-\frac{\left(\frac{2H}{T} - \mu_{\Delta v}\right)^2}{2\sigma_{\Delta v}^2}} \quad (6)$$

where  $\mu_{\Delta v}$  and  $\sigma_{\Delta v}^2$  denote the average value and the variance of relative velocity  $\Delta v = |v_i - v_j|$ , respectively. Hence, the expected social link duration  $ET(SL_{ij})$  can be estimated as follows:

$$ET(SL_{ij}) = r_t(SL_{ij}) \cdot SL_{ij}(t) \quad (7)$$

## ***Network-wide Metrics***

### *Closeness Centrality*

It can be noticed in (2) that the degree centrality metric does not consider the indirect connections that a node can establish with other nodes using the available paths in the network. Thus, the degree centrality metric is not enough to recognise the most important nodes in VSN. We define a node as a central-close if it can reach other nodes through short distance paths. Hence, the closeness centrality metric is related to the inverse of distance between nodes, e.g., the higher the distance, the less central-close. In social theory, the shortest distance path between two nodes is defined as a geodesic. Thus, the closeness centrality index should consider the geodesics that a given node has to all other nodes in the network. Let  $d(n_i, n_j)$  be the geodesic between two nodes  $n_i$  and  $n_j$ , the

standardised closeness centrality  $C'_C(n_i)$  of a node  $n_i$  can be calculated as follows:

$$C'_C(n_i) = \frac{n-1}{\left( \sum_{j=1, i \neq j}^n d(n_i, n_j) \right)} \quad (8)$$

In (8), it can be noticed that the closeness centrality index for  $n_i$  will be zero if there is at least one node  $n_j$  that is unreachable from  $n_i$ , i.e., its geodesic will be infinite. In the context of VSNs, choosing nodes with high closeness centrality index to forward a message will optimise the resources needed to deliver it and ensures a faster delivery as well.

#### *Betweenness Centrality*

Betweenness is another measure of centrality that focuses on nodes that lie in the path between other nodes. In order to calculate the betweenness centrality index, it is assumed that nodes prefer to communicate via the shortest paths in the network. Thus, the standardised betweenness centrality  $C'_B(n_i)$  of a node  $n_i$ , which expresses the number of shortest paths, i.e., geodesics, that pass through  $n_i$ , is calculated as follows

$$C'_B(n_i) = \frac{\sum_{j < k, i \neq j, i \neq k} \frac{g_{jk}(n_i)}{g_{jk}}}{\frac{(n-2)(n-1)}{2}} \quad (9)$$

where  $g_{jk}(n_i)$  is the number of geodesics linking  $n_j$  and  $n_k$  that contain  $n_i$  in between and  $g_{jk}$  is the total number of geodesics linking  $n_j$  and  $n_k$ . A node with a high betweenness centrality index plays the role of “broker” and has a great influence on the data dissemination in VSN.

#### *Bridging Centrality*

The bridging centrality of a node  $n_i$  is expressed as the product of its betweenness centrality and a bridging coefficient  $\beta(n_i)$ . The bridging centrality metric defines nodes that are central to the network graph, connecting two highly connected sub-graphs and have a low number of direct

connections relative to their neighbour connections. The bridging coefficient  $\beta(n_i)$  is the ratio of the inverse of a node degree to the sum of the inverses of all its neighbours' degrees. The bridging centrality can be calculated as follows:

$$C'_R(n_i) = C'_B(n_i)\beta(n_i) \quad (10)$$

where  $\beta(n_i)$  is calculated as follows

$$\beta(n_i) = \frac{C'_D(n_i)^{-1}}{\left( \sum_{j=1, j \neq i} C'_D(n_j)^{-1} \right)} \quad (11)$$

### *Social Path Duration*

When two nodes  $n_i$  and  $n_j$  are not adjacent to each other, the social path  $SP_{ij}$  is defined as the set of social links that connect  $n_i$  and  $n_j$  via multiple hops in the network. Without loss of generality, for any given path  $SP_{ij}$ , let us denote the number of its formed links by  $k$ , i.e.,  $SP_{ij} = \{SL_1, SL_2 \dots SL_k\}$ . The expected social path duration  $SP_{ij}(t)$  is defined as the minimum of social links durations that comprise this path, i.e.,

$$SP_{ij}(t) = \min_{SL_\omega \in SP_{ij}} ET(SL_\omega) \quad \text{where } \omega = 1 \dots k \quad (12)$$

while the reliability of the social path, denoted as  $R_t(SP_{ij})$ , can be calculated as follows:

$$R_t(SP_{ij}) = \prod_{\omega=1}^k r_t(SL_\omega) \quad (13)$$

### *Community-wide Metrics*

#### *Number of Clusters*

Considering the mobility patterns of vehicles on the roads and the traffic conditions, the formation of non-connected clusters is inevitable in VSNs. We define the cluster as a sub-graph of the whole network that contains a number of connected vehicles where there is a path between any pair of

nodes. The higher the number of clusters in VSNs the lower the chance of creating new social connections among vehicles in the network.

### *Number of Social Groups*

Different from the cluster approach, the social group/community in VSNs is a sub-graph of the whole network that still has a connection with the rest of network. However, the number of intra-group links is larger than the number of inter-group links. Nodes within a social group usually share the same social attributes/interests. The formation of social groups is quite important in VSNs to attract users' attention and encourage them to join. The higher the number of social groups the better the experience users will have when joining the VSN.

### *Model of Social Morality*

In a fully autonomous system, users independently behave based on the rational calculation of expediency (Fukuyama 1996). Generally, users take their decision to act in social interactions from two points of view: 1) economic and motivated by self-interest; and 2) noneconomic and motivated by collective interest and moral obligation. In reality, when users violates a deeply internalised norm, which governs users' behaviour in economic and noneconomic spheres of activity, they will feel guilty to some extent and would likely punish themselves in some manner whether anyone else knew of their actions or not. This is known as social morality (Liang, et al. 2012).

In the context of social networks, both cooperative and non-cooperative behaviours of users have a significant impact on the performance of the social network. It has been shown that users who experienced a feeling of guilt after a non-cooperative behaviour tend to show higher level of cooperation in the later social interactions (Ketelaara and Aub 2003). As a matter of fact, the feeling of guilt encourages users to depart from their typical non-cooperative behaviour. In VSNs, the cooperation is highly desirable among users to promote social interactions and consequently

help with delivering data packets. Since the VSN users are autonomous and intelligent individuals, it is reasonable to assume that they are rational and their behaviours are driven by personal profit and morality. Thus, we are interested in observing two forms of social morality that are resulted based on the decisions taken by users whether to accept and forward data packets, i.e., cooperative behaviour, or to reject and drop data packets, i.e., non-cooperative behaviour. These two forms of social morality are guilt and high-mindedness where users feel high-minded when they choose to cooperate and they feel guilty otherwise.

Let  $g(n_i)$  be a self-regulated morality factor for a node  $n_i$ , which quantitatively depicts the internal moral force of the user. This factor is based on the following two elements (Liang, et al. 2012):

- Morality state  $mt(n_i)$ . This element has a variable value and reflects the behaviour history of the user. It increases by one level for a single cooperation behaviour and decreases by one level due to a single defection conduct. The initial state is set to 0, which means neither guilty nor high-minded. States with a positive index are high-minded states that imply frequent cooperative behaviour in the past. On the other hand, states with a negative index are guilty states that imply an overwhelming defection conduct in the past.
- Sociality strength  $st(n_i)$ . This element is related to the user's personal experience such as education and habitation. It is less independent and stabilised measure with short-term behaviour changes. If the sociality strength of a user is significant, the user experiences a significant increment of guilt toward a single defection behaviour and a significant increment of high-mindedness toward a single cooperation behaviour. The  $st(n_i)$  value is chosen in the range  $[0, 1]$ .

In order to evaluate the morality factor  $g(n_i)$  the current morality state  $mt(n_i)$  and the sociality strength  $st(n_i)$  are utilised as follows:

$$g(n_i) = \begin{cases} f(mt(n_i), st(n_i)) & \text{if } mt(n_i) < 0 \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

where the value of  $g(n_i)$  increases as  $mt(n_i)$  decreases or  $st(n_i)$  increases. The function  $f$  can be selected from the following three morality functions the linear function  $f_1$ , the natural logarithm function  $f_e$ , and the common logarithm function  $f_{10}$

$$\begin{aligned} f_1(mt(n_i), st(n_i)) &= \delta \cdot st(n_i) \cdot (-mt(n_i)) \\ f_e(mt(n_i), st(n_i)) &= \delta \cdot \ln(1 + st(n_i) \cdot (-mt(n_i))) \\ f_{10}(mt(n_i), st(n_i)) &= \delta \cdot \log_{10}(1 + st(n_i) \cdot (-mt(n_i))) \end{aligned} \quad (15)$$

where  $\delta$  is a tuneable coefficient in the range  $[0, +\infty[$ . These three morality functions represent three different levels of morality force that affect user cooperation behaviour, respectively. They always output a nonnegative value. In the context of VSNs, choosing a node with a high morality factor value is very important to ensure the delivery of data packets. Moreover, this will attract more users to join the VSN and promote the cooperative nature of these networks. This issue is further discussed in Section 4.

### The Evolving Graph Theory

The evolving graph theory is proposed as a formal abstraction for dynamic networks (Ferreira 2002). The evolving graph is an indexed sequence of  $\lambda$  sub graphs of a given graph, where the sub graph at a given index corresponds to the network connectivity at the time interval indicated by the index number, as shown below in Fig. 1.

It can be observed from Fig. 1 that edges are labelled with corresponding presence time intervals. For instance, in Fig. 1,  $\{A, D, C\}$  is not a valid journey since edge  $\{D, C\}$  exists only in the past with respect to edge  $\{A, D\}$ . Hence, the journey in the evolving graph is the path in the underlying graph where its edges time labels are in increasing order (Eiza & Ni, 2013). In Fig. 1, it is easy to

find that  $\{A, B, E, G\}$  and  $\{D, C, E, G\}$  are valid journeys while  $\{D, C, E, G, F\}$  is not.

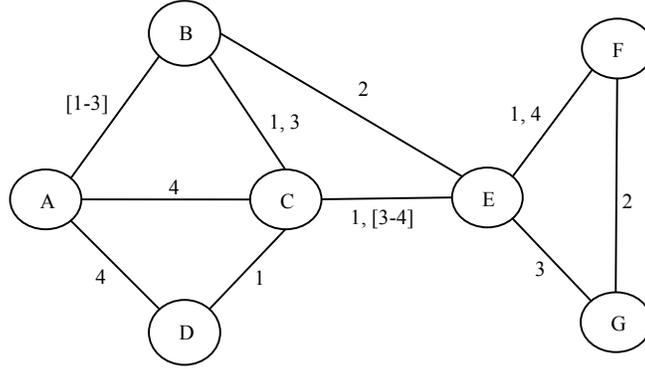


Fig 1. Basic Evolving Graph Model (Monteiro 2008)

Let  $G(V, E)$  be a given graph and an ordered sequence of its sub graphs,  $S_G = G_1(V_1, E_1), G_2(V_2, E_2), G_3(V_3, E_3) \dots G_\lambda(V_\lambda, E_\lambda)$  such that  $\bigcup_{i=1}^{\lambda} G_i = G$ . The evolving graph is defined as  $G' = (S_G, G)$  where the vertices set of  $G'$  is  $V_{G'} = \bigcup V_i$  and the edges set of  $G'$  is  $E_{G'} = \bigcup E_i$ . Suppose that the sub graph  $G_i(V_i, E_i)$  at a given index  $i$  is the underlying graph of the network during time interval  $F = [t_{i-1}, t_i]$  where  $t_0 < t_1 < \dots < t_\tau$ , the time domain  $\check{T}$  is now incorporated in the model.

Let  $\Omega$  be a given path in the evolving graph  $G'$  where  $\Omega = e_1, e_2, e_3 \dots e_k$  with  $e_i \in E_{G'}$  in  $G$ . Let  $\Omega_\sigma = \sigma_1, \sigma_2, \sigma_3 \dots \sigma_k$  with  $\sigma_i \in \check{T}$  be the time schedule indicating when each edge of the path  $\Omega$  is to be traversed. We define a journey  $J = (\Omega, \Omega_\sigma)$  if and only if  $\Omega_\sigma$  is in accordance with  $\Omega, G'$  and  $F$ . This means that  $J$  allows the traverse from node  $n_i$  to node  $n_j$  in  $G'$ .

In the current evolving graph theory, three journey metrics are defined: the foremost, shortest, and fastest journey. They are introduced to find the earliest arrival date, the minimum number of hops, and the minimum delay (time span) path, respectively. Let  $J = (\Omega, \Omega_\sigma)$  be a given journey in  $G'$  where  $\Omega = e_1, e_2, e_3 \dots e_k$  and  $\Omega_\sigma = \sigma_1, \sigma_2, \sigma_3 \dots \sigma_k$  then:

- The hop count  $h(J)$  or the length of  $J$  is defined as  $h(J) = |\Omega|$ .
- The arrival date of the journey  $a(J)$  is defined as the scheduled time for the traversal of the

last edge in  $J$ , plus its traversal time, i.e.,  $a(J) = \sigma_k + t(e_k)$ .

- The journey time  $t(J)$  is defined as the past time between the departure and the arrival, i.e.,

$$t(J) = a(J) - \sigma_1.$$

## **Social Evolving Graph-based Connectivity Model for VSNs**

### ***Motivation***

The current evolving graph theory cannot be directly applied to VSNs because the evolving social properties of the VSN communication graph cannot be scheduled in advance. Moreover, the current evolving graph model does not consider the social metrics of the communicating nodes. In order to facilitate the establishment of social links/paths in VSNs and the data forwarding process, we extend the current evolving graph model to develop a social evolving graph-based connectivity (*SEGC*) model for VSNs. The *SEGC* model has two main goals in the context of VSNs. First, it captures the social characteristics of the existing nodes and, by considering both social and connectivity metrics, it establishes social links/paths among these nodes. Secondly, the *SEGC* model facilitates the data forwarding among the socially connected vehicles using the social theory indexes we have mentioned above along with the conventional connectivity metrics. In the following, we introduce the proposed *SEGC* model and explain the data forwarding mechanism that takes advantage of the developed *SEGC* model.

### ***Social Evolving Graph-based Connectivity (SEGC) Model***

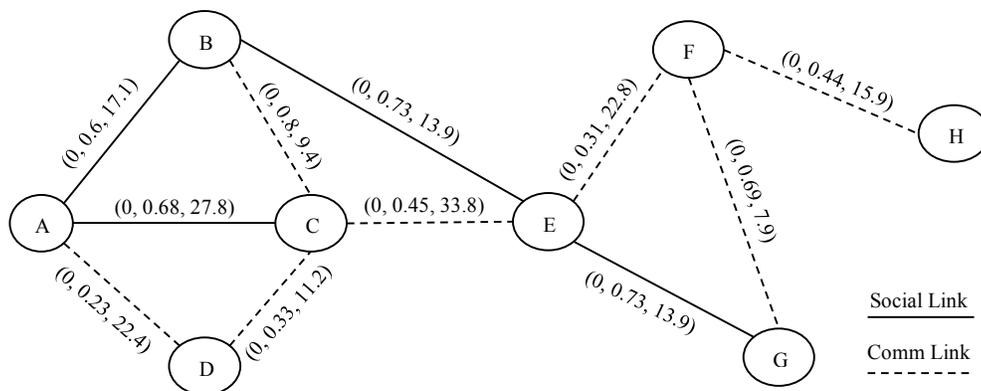
As we have mentioned before, establishing new social connections between two vehicles does not only depend on being within the transmission range of each other but also on their social attributes and interests. Thus, in the proposed *SEGC* model, each link is characterised with a set of attributes that include all the connectivity and social indexes we have mentioned before. The social link is

only established between two vehicles  $n_i$  and  $n_j$  if it satisfies the following two conditions. First, the  $SHP_{ij}$ , i.e., the homophily, should be higher than a predefined threshold  $\Psi_H$  thus the users share the minimum level of interest. Secondly, the expected social link duration  $ET(SL_{ij})$  should be higher than a predefined time threshold  $\Psi_L$ . The value of  $\Psi_H$  can be defined/advertised by the vehicle itself, e.g., high value of  $\Psi_H$  indicates that the user is only interested in communicating with other users that have a lot in common with himself/herself. On the other hand, the time threshold  $\Psi_L$  can be determined by the current application. For instance, in order to share a video clip with other vehicles, the connection time should be long enough to watch/download the video file.

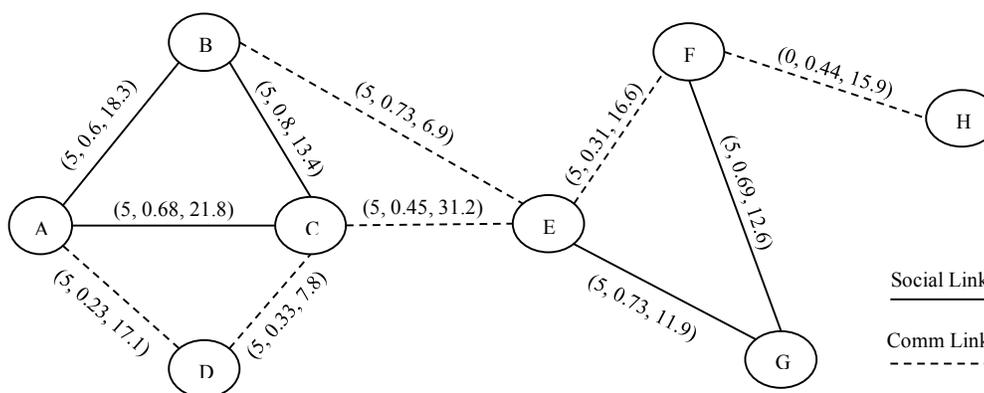
Fig. 2 shows an example of the *SEGC* model on a highway at two time instants  $t = 0$  s and  $t = 5$  s where  $\Psi_H = 0.5$ ,  $\Psi_L = 10$  s. Each node in Fig. 2 shows a vehicle on the highway. It can be seen in Fig. 2 that unlike the corresponding presence time intervals for each link as shown in Fig.1, we associate the following tuple  $(t, SHP_{ij}, ET(SL_{ij}))$  with each link where  $t$  denotes the current time,  $SHP_{ij}$  denotes the homophily between  $n_i$  and  $n_j$  and  $ET(SL_{ij})$  denotes the expected social link duration.

In the *SEGC* model, the social link between two vehicles is not available if  $SHP_{ij} < \Psi_H$  or  $ET(SL_{ij}) < \Psi_L$ . Therefore, even if the communication link exists between two vehicles and satisfies the connection time threshold, e.g., the communication link between vehicles  $A$  and  $D$  in Fig.2 (a), the social link is not established since it does not satisfy the condition of the homophily as  $0.23 < 0.5$ . Fig. 2(a) shows the *SEGC* status and the corresponding  $SHP_{ij}$  and  $ET(SL_{ij})$  values associated to each link at  $t = 0$  s. It can be noticed that the following social links are established  $\{A, B\}$ ,  $\{A, C\}$ ,  $\{B, E\}$  and  $\{E, G\}$ . After 5 seconds, in Fig. 2 (b), the set of the established social links changes and becomes as follows  $\{A, B\}$ ,  $\{A, C\}$ ,  $\{B, C\}$ ,  $\{E, G\}$  and  $\{F, G\}$ . It is worth noting that all links

in Fig. 2 are eligible to be traversed. However, if the link is eligible to be traversed, it does not necessarily mean that a social link will be established.



(a)



(b)

Fig 2. Social Evolving Graph-based Connectivity (SEGC) Model at (a)  $t = 0$  s and (b)  $t = 5$  s

In VSNs, we assume that each vehicle along the road has its own version of the *SEGC* model shown in Fig. 2. This is possible using the information received within the basic safety messages (BSMs) that are periodically exchanged in vehicular networks when the 5.9-GHz dedicated short-range communication (DSRC) standard is deployed (Kenney 2011). In this way, each vehicle  $n_i$  can be only concerned with vehicles of interest, i.e., vehicles that share the same social attributes/interests with  $n_i$ . It can be noticed that unlike the conventional evolving graph, the

presence time of the social link in the *SEGC* model is continuous and depends on the current vehicular traffic status and the social attributes of vehicles. In this case, there is no need to check the order of the presence times of the link when searching for a valid journey.

In order to establish a social path, between two non-adjacent vehicles in the network, the same principle is applied. However, forwarding the data along the established multi-hop social path from the source to the destination should take into account different parameters than those that were considered while establishing the social link. These different parameters are related to the relay vehicles along the established path and are illustrated in the following section.

### ***SEGC-based Data Forwarding Mechanism***

In order to forward data packets in VSNs among non-adjacent connected vehicles, social paths should be established. In this section, we propose a new forwarding data mechanism that can benefit from the *SEGC* model advantages and properties. The proposed mechanism utilises the *SEGC* model and considers both social and connectivity metrics while searching for a path from the source to the destination. The considered metrics are degree centrality, morality factor, closeness centrality, betweenness centrality, and bridging centrality. However, establishing the social path for data forwarding between two vehicles subject to these multiple metrics features a multi constrained path (MCP) selection, which is proven to be an NP-hard problem (Wang and Crowcroft 1996) if the constraints are mutually independent (Reeves and Salama 2000). Therefore, we propose the following evaluation function  $EF(SL_{ij})$  that considers these metrics and its weights:

$$EF(SL_{ij}) = \gamma_g g(n_j) + \gamma_D C_D(n_j) + \gamma_C C_C(n_j) + \gamma_B C_B(n_j) + \gamma_R C_R(n_j) \quad (16)$$

Where  $\gamma_g, \gamma_D, \gamma_C, \gamma_B,$  and  $\gamma_R$  are weighting factors for the morality factor, degree centrality, closeness centrality, betweenness centrality and bridging centrality, respectively. These factors are chosen in the range [0, 1]. We worked out this function by experimentation and its validity is illustrated

by the simulation results presented later. In order to explain the purpose of these weighing factors, let assume that the source and the destination vehicles belong to different social communities/groups. In this case,  $\gamma_B$  is given high value because it is important to forward data packets through a vehicle with a high bridging centrality value. On the other hand,  $\gamma_g$  is always given a high value because data packets should be forwarded through vehicles with high morality factor, i.e., vehicles that showed cooperative behaviour in the past, to ensure a successful data packets forwarding. When the source vehicle has data to send at time  $t$ , it evaluates the communication links in the current *SEGC* model and assigns each link with a single value  $EF(SL_{ij})$  as estimated in (16). Finding the optimal path in the *SEGC* model according to  $EF(SL_{ij})$  value is equivalent to finding the optimal journey in the underlying graph where a modified version of Dijkstra's algorithm can be applied (Eiza and Ni 2013). The modified Dijkstra's algorithm scans all the network nodes in the *SEGC* model and returns the optimal route according to  $EF(SL_{ij})$  value.

### **Performance Evaluation**

The main objective of this performance evaluation is to identify the impact of high dynamics of network topology changes in VSNs on the establishment of social connections among the communicating vehicles. In addition, we want to check the benefits of using the proposed *SEGC* model in the highway scenario. We construct our performance evaluation using the OMNet++ network simulator (Varga 2003). OMNet++ is an extensible modular component based C++ simulation library and framework. The simulations are run on a six-lane traffic simulation scenario of a 10 km highway with two independent driving directions in which vehicles move. For each simulation, we perform 20 runs to obtain its average results. The results are compared to those when *SEGC* model is not involved, i.e., greedy forwarding mechanism is applied.

In our simulation scenario, the average velocity of vehicles in the first two lanes are 40 and 60

$km/h$ , respectively, while we change the average velocity of the vehicles in the third lane only from 60 to 130  $km/h$ . We use the highway mobility model developed in (Eiza, Ni and Owens, et al. 2013), which is built based on traffic theory rules and considers the drivers' behaviours. The number of vehicles on the highway is 120 vehicles and the data packet size is 2  $KB$ . The social attributes profile of each vehicle  $HP_i$  is generated randomly to match the following set {Travel destination, Rap music, Mountain climbing, Thai food, Work place, Football, Jogging, Cooking}. This set is imaginary and designed for the purpose of this simulation. In a real-world scenario, this set could contain over 100 elements. The social attributes profile is assumed to be transmitted periodically for vehicles that want to participate in social interactions along the road over VSNs. The morality factor  $g(n_i)$  is evaluated using the linear function  $f_1$  in (15) where  $\delta = 1$ . The sociality strength  $st(n_i)$  value for each vehicle is randomly selected in the range  $[0, 1]$ . The weighting factors in (16) are set as follows:  $\gamma_g = 1$ ,  $\gamma_D = 0.7$ ,  $\gamma_C = 0.5$ ,  $\gamma_B = 0.5$ , and  $\gamma_R = 1$  if the source and the destination belong to different social groups, otherwise  $\gamma_R = 0.1$ . Finally, the homophily threshold  $\Psi_H$  is randomly selected in the range  $[0, 1]$  for each vehicle at the beginning of the simulation run and stays fixed for the rest of the simulation time. The time threshold  $\Psi_L$  is set to 10 seconds. When  $SHP_{ij} \geq \Psi_H$  and  $ET(SL_{ij}) \geq \Psi_L$  between two vehicles, data packets transmission takes place. The simulation parameters are summarised in Table I.

TABLE I – SUMMARY OF THE SIMULATION PARAMETERS

Simulation Area	1 $km$ x 10 $km$
Mobility Model	Highway
Communication Range	450 $m$
MAC Layer	IEEE 802.11p
Vehicles' velocities	Normally distributed
Vehicles' distances	Exponentially distributed
Number of runs	20
Simulation duration	300 seconds
Morality function	$f_1(mt(n_i), st(n_i)) = \delta \cdot st(n_i) \cdot (-mt(n_i))$
Morality function coefficient	$\delta = 1$
Weighting factors	$\gamma_g = 1, \gamma_D = 0.7, \gamma_C = 0.5, \gamma_B = 0.5, \gamma_R = 1$ or $\gamma_R = 0.1$

### ***Performance Metrics***

The following performance metrics are considered for the simulations.

- Packet delivery ratio (PDR): It represents the average ratio of all successfully received data packets at the destination node over all data packets generated by the application layer at the source node.
- Social connections: It represents the average number of social connections that are established among the communicating vehicles.
- Link failures: It represents the average number of communication link failures during the data forwarding process. This metric shows the efficiency of the data forwarding algorithm in avoiding link failures.
- Social path lifetime: It represents the average lifetime of the established social path between two vehicles. A longer lifetime means a more stable and more reliable path.

### ***Simulation Results***

In Fig. 3, it can be seen that the average PDR reduces noticeably when the average velocity in the third lane starts to exceed  $80 \text{ km/h}$ . This reduction comes from the fact that the network topology becomes more dynamic, and thus links/paths are more vulnerable to disconnection. In this particular case, it is important to establish reliable social paths among the communicating vehicles. The utilisation of *SEGC* model ensures that only reliable paths are established among the socially connected vehicles. These paths are calculated using the *SEGC* model where the evolving characteristics of the network topology are considered via (13). Moreover, the evaluation function in (16) ensures that data packets are relayed through vehicles with high morality factor and high degree centrality. Thus, the probability of a successful data delivery is high.

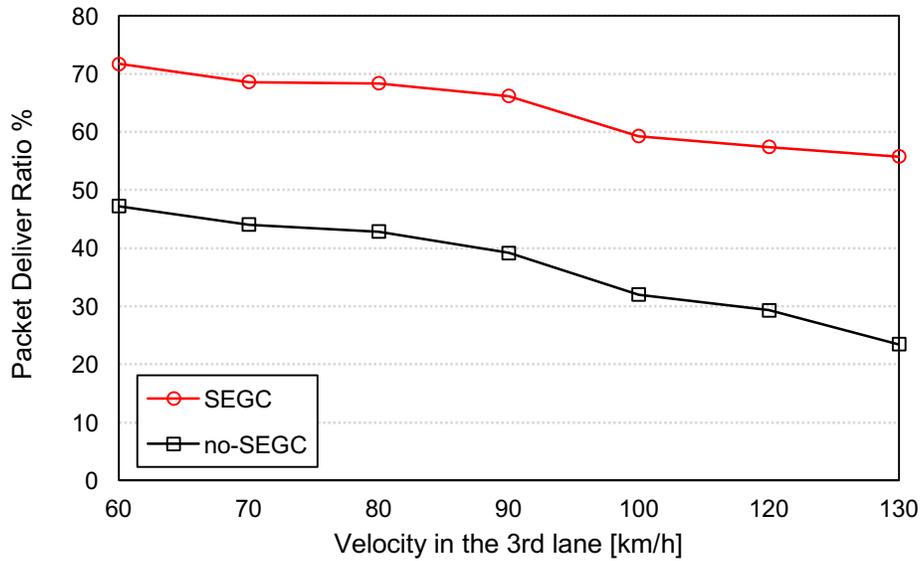


Fig 3. Average Packet Delivery Ratio

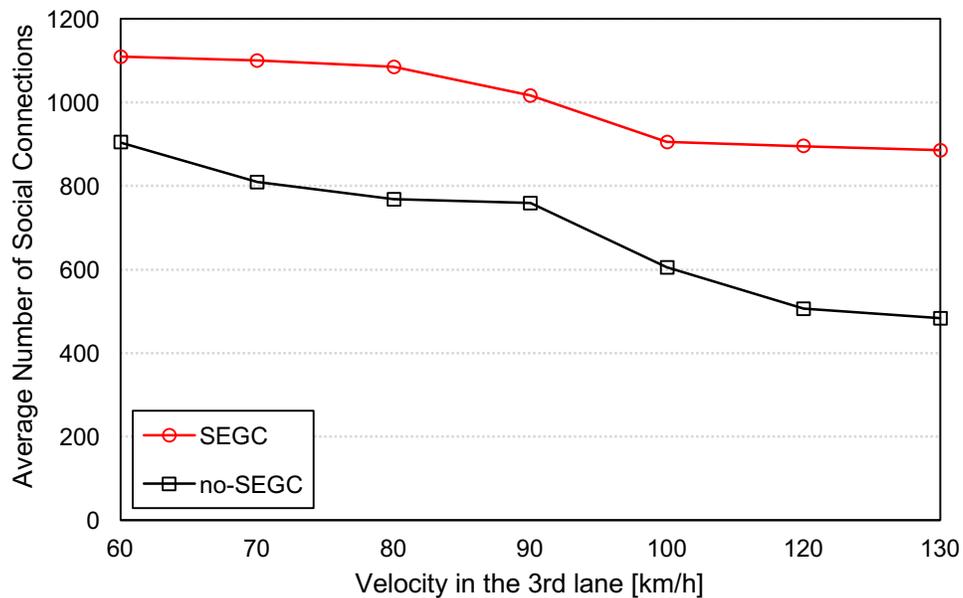


Fig 4. Average Number of Established Social Connections

The average number of established social connections among the communicating vehicles is shown in Fig. 4. It can be noticed that when *SEGC* model is utilised, the number of social connections is high in comparison to the case where *SEGC* is not presented. The reason is that each vehicle has its own *SEGC* model which is updated regularly when a new vehicle enters the

communication range of that vehicle. If the homophily exceeds the defined threshold, the *SEGC* establishes the social link/path between the two vehicles and commences data packets transmission.

In Fig. 5, the utilisation of *SEGC* helps obtaining a very low number of social link failures in comparison to the case when *SEGC* is not utilised. The number of link failures increases when the velocity increases. In this case, it is essential to accurately capture the changes of vehicular velocities and establish reliable paths between the communicating vehicles. Furthermore, choosing relay vehicles with high morality factor is crucial to guarantee that the established path will not break if one of the vehicles along the established path refuse to cooperate in the data packets forwarding process.

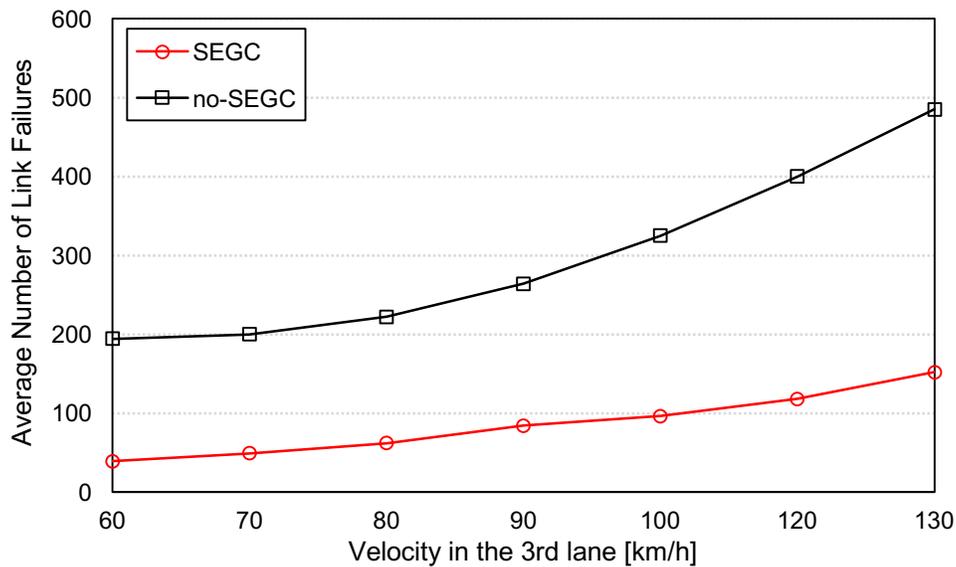


Fig 5. Average Number of Link Failures

In Fig. 6, we show the average social path lifetime obtained in this performance evaluation. When the *SEGC* model is utilised, longer social path lifetimes are achieved thanks to establishing the most reliable paths in the network and utilising the social indexes to forward the data among the communicating vehicles effectively. This observation explains the high PDR shown in Fig. 3.

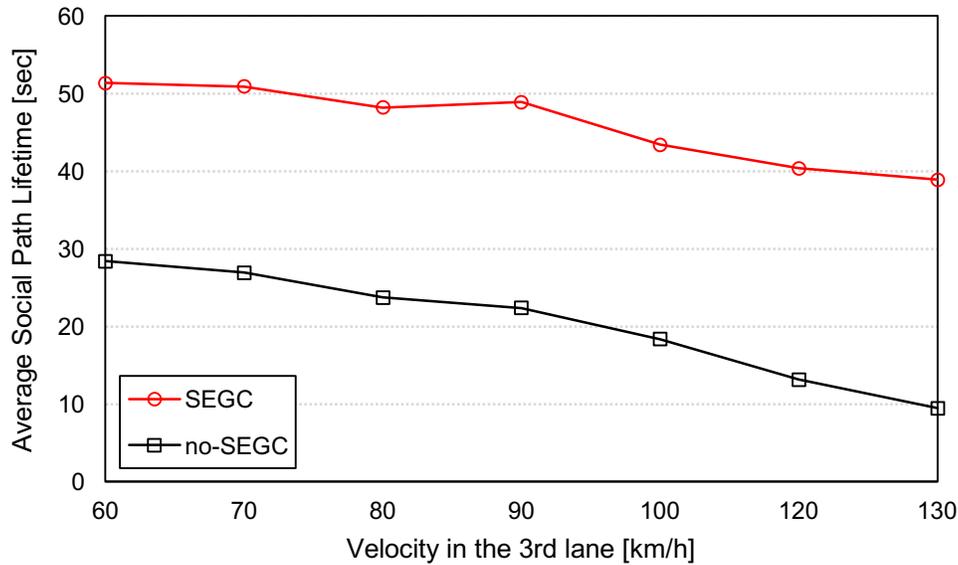


Fig 6. Average Social Path Lifetime

## Conclusion

In this chapter, we have extended the evolving graph theory and utilised the social theory concepts to develop a novel social evolving graph-based connectivity (*SEGC*) model for VSNs. The proposed connectivity model considers both social metrics of the communicating vehicles and the conventional connectivity issues in VSNs. Therefore, the social links/paths are established between the communicating vehicles based on their social characteristics and interests rather than just their kinematic information. The performance of *SEGC* has been compared with the one when greedy data forwarding mechanism is utilised through our simulations. The simulation results showed that the utilisation of *SEGC* model helped achieving higher PDR and establishing stable social paths with longer lifetimes. Since it establishes the most reliable social path between the source and the destination, it also achieves the lowest number of social link failures. The *SEGC* model shows promising results in the context of VSNs. However, more investigation and therefore more simulations are needed to validate the *SEGC* model in different traffic scenarios with

different traffic parameters. In the future work, we will investigate the role of self-driving vehicles and their effects on the connectivity patterns in VSNs. Moreover, we will develop a model to consider different social attributes profiles for individuals inside the vehicles including the travellers and the drivers.

## References

- Batallas, Diego A., and Ali A. Yassine. 2006. "Information Leaders in Product Development Organizational Networks: Social Network Analysis of the Design Structure Matrix." *IEEE Transactions on Engineering Management* 570-582.
- Eiza, M.H., and Q. Ni. 2013. "An evolving graph-based reliable routing scheme for VANETs." *IEEE Transactions on Vehicular Technology* 1493-1504.
- Eiza, M.H., Q. Ni, T. Owens, and G. Min. 2013. "Investigation of routing reliability of vehicular ad hoc networks." *EURASIP Journal on Wireless Communications and Networking* 1-15.
- Ferreira, A. 2002. "On models and algorithms for dynamic communication networks: The case for evolving graphs." *The 4e rencontres francophones sur les ALGOTEL*. Mèze, France.
- Fukuyama, F. 1996. *Trust: Social Virtues and the Creation of Prosperity*. New York: The Free Press.
- Hu, Xiping, Terry H. S. Chu, Victor C. M. Leung, Edith C.-H. Ngai, Philippe Kruchten, and Henry C. B. Chan. 2015. "A Survey on Mobile Social Networks: Applications, Platforms, System Architectures, and Future Research Directions." *IEEE Communication Surveys & Tutorials* 1557-1581.
- Kageyama, Yuri. 2011. 'Toyota Friend' Social Networking Service Is A Twitter For Car Owners. 23 May. Accessed Nov 22, 2015. [http://www.huffingtonpost.com/2011/05/23/toyota-friend-social-network\\_n\\_865437.html](http://www.huffingtonpost.com/2011/05/23/toyota-friend-social-network_n_865437.html).
- Kenney, J. B. 2011. "Dedicated Short-Range Communications (DSRC) Standard in the United States." *Proceedings of the IEEE* 1162-1182.
- Ketelaara, Timothy, and Wing Tung Aub. 2003. "The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction." *Cognition and Emotion* 429-453.
- Knobel, Martin, Marc Hassenzahl, Melanie Lamara, Tobias Sattler, Josef Schumann, Kai Eckoldt, and Andreas Butz. 2012. "Cliques Trip: Feeling Related in Different Cars." *Proceedings of the Designing Interactive Systems Conference*. Newcastle, UK: ACM. 29-37.
- Li, Xin, Lei Guo, and Yihong (Eric) Zhao. 2008. "Tag-based Social Interest Discovery." *Proceedings of the 17th international conference on World Wide Web*. ACM. 675-684.
- Liang, Xiaohui, Xu Li, Tom H. Luan, Rongxing Lu, Xiaodong Lin, and Xuemin Shen. 2012. "Morality-Driven Data Forwarding With Privacy Preservation in Mobile Social Networks." *IEEE Transactions on Vehicular Technology* 3209-3222.

- Lu, Rongxing. 2012. *Security and Privacy Preservation in Vehicular Social Networks*. PhD Thesis, Waterloo, Canada: University of Waterloo.
- Luan, Tom H., Rongxing Lu, Xuemin (Sherman) Shen, and Fan Bai. 2015. "Social on The Road: Enabling Secure and Efficient Social Networking On Highways." *IEEE Wireless Communications* 44-51.
- Luan, Tom H., Xuemin (Sherman) Shen, Fan Bai, and Limin Sun. 2015. "Feel Bored? Join Verse! Engineering Vehicular Proximity Social Networks." *IEEE Transactions on Vehicular Technology* 1120-1131.
- Monteiro, J. 2008. "The use of evolving graph combinatorial model in routing protocols for dynamic networks." *XV Concorso Latinoamericano de Tesis de Maestría*. 1–17.
- Niu, Z., W. Yao, Q. Ni, and Y. Song. 2006. "Link reliability model for vehicle ad hoc networks." *London Communication Symposium*. London. 1-4.
- Pallis, George, Dimitrios Katsaros, Marios D. Dikaiakos, Nicholas Loulloudes, and Leandros Tassioulas. 2009. "On the Structure and Evolution of Vehicular Networks." *IEEE International Symposium on Modeling, Analysis & Simulation of Computer and Telecommunication Systems*. London, UK: IEEE. 1-10.
- Reeves, D.S., and H.F. Salama. 2000. "Distributed algorithm for delay-constrained unicast routing." *IEEE/ACM Transactions on Networking* 239–250.
- Schnabel, W., and D. Lohse. 1997. *Grundlagen der Straßenverkehrstechnik und der Verkehrsplanung*. Berlin: Aufl Verlag für Bauwesen.
- Sha, Wenjie, Daehan Kwak, Badri Nath, and Liviu Iftode. 2013. "Social Vehicle Navigation: Integrating Shared Driving Experience into Vehicle Navigation." *Proceedings of 14th Workshop on Mobile Computing Systems and Applications*. New York, US: ACM. 161-166.
- Smaldone, Stephen, Lu Han, Pravin Shankar, and Liviu Iftode. 2008. "RoadSpeak: enabling voice chat on roadways using vehicular social networks." *Proceedings of the 1st Workshop on Social Network Systems*. Glasgow: ACM. 43-48 .
- Smyth, Tanya L., and Mark J. King. 2006. "Driver-vehicle interactions in 4WDs: A theoretical review." *Australasian Road Safety Research, Policing Education Conference*. Queensland, Australia: Queensland University of Technology.
- Snijders, T., and S. Borgatti. 1999. "Non-Parametric standard errors and tests." *Connections* 161–170.
- TripAdvisor, Inc. 2012. *TripAdvisor Survey Reveals Three Quarters Of U.S. Travelers Sharing Trip Experiences On Social Networks*. 20 Sept. Accessed June 19, 2015. [http://www.tripadvisor.co.uk/PressCenter-i5414-c1-Press\\_Releases.html](http://www.tripadvisor.co.uk/PressCenter-i5414-c1-Press_Releases.html).
- Varga, Andras. 2003. *OMNeT++ - Discrete Event Simulator*. Accessed 2011. <https://omnetpp.org/>.
- Vegni, Anna Maria, and Valeria Loscri. 2015. "A Survey of Vehicular Social Networks." *IEEE Communicaitons and Tutorials* 2397-2419.
- Wang, Z., and J. Crowcroft. 1996. "Quality-of-service routing for supporting multimedia applications." *IEEE Journal on Selected Areas in Communications* 1228–1234.