Throughput analysis of one-dimensional vehicular ad hoc networks

Chi Wan Sung*

Department of Electronic Engineering,
City University of Hong Kong,
Tat Chee Ave, Kowloon, Hong Kong
E-mail: albert.sung@cityu.edu.hk
*Corresponding author

Kenneth W. Shum

Institute of Network Coding,
The Chinese University of Hong Kong,
Shatin, Hong Kong
E-mail: wkshum@inc.cuhk.edu.hk

Wing Ho Yuen

Rutgers, The State University of New Jersey,
Technology Centre of New Jersey,
671 Route 1 South, North Brunswick, NJ 08902-3390, USA
E-mail: esquare2003@yahoo.com

Abstract: We consider content distribution in a one-dimensional vehicular ad hoc network. We assume that a file is encoded using fountain code, and the encoded message is cached at infostations. Vehicles are allowed to download data packets from infostations, which are placed along a highway. In addition, two vehicles can exchange packets with each other when they are in proximity. As long as a vehicle has received enough packets from infostations or from other vehicles, the original file can be recovered. In this work, we derive closed-form expressions for the average per-node throughput, under both discrete and continuous velocity distributions for the vehicles. Our result shows that the average per-node throughput can be expressed as a linear function of the densities of different classes of nodes in the highway. This result implies that the average per-node throughput scales linearly with vehicle arrival rate, and the average system throughput scales quadratically with vehicle arrival rate. Besides, system throughput reduces when overall mobility increases.

Keywords: vehicular ad hoc networks; VANETs; throughput; fountain codes; mobility.

Reference to this paper should be made as follows: Sung, C.W., Shum, K.W. and Yuen, W.H. (2012) 'Throughput analysis of one-dimensional vehicular ad hoc networks', *Int. J. Systems, Control and Communications*, Vol. 4, No. 3, pp.150–163.

Biographical notes: Chi Wan Sung received his BEng, MPhil and PhD in Information Engineering from the Chinese University of Hong Kong in 1993, 1995, and 1998, respectively. He joined the faculty at City University of Hong Kong in 2000, and is now an Associate Professor with the Department of Electronic Engineering. He has been an Editor of the *ETRI Journal* since 2009. His research interests include wireless communications, network coding and algorithm design.

Kenneth W. Shum received his BEng in Information Engineering from the Chinese University of Hong Kong in 1993, and his MS and PhD in Electrical Engineering from the University of Southern California in 1995 and 2000 respectively. He is now a Research Fellow in the Institute of Network Coding, the Chinese University of Hong Kong. His research interests include information theory and coding theory.

Wing Ho Yuen obtained his Bachelor in Electrical Engineering and Master in Information Engineering, respectively from Hong Kong University of Science and Technology and Chinese University of Hong Kong. He obtained his PhD in Electrical Engineering from Rutgers University, in 2004. He worked as a Postdoc in Columbia University from 2005 to 2007, under the supervision of Prof. Henning Schulzrinne. From 2007 to 2010, he worked as a Senior Research Scientist in Telcordia Technologies. He is currently an Assistant Vice-president in Citigroup.

1 Introduction

Vehicular ad hoc network (VANET) is an important class of wireless networks. It consists of cars, trucks, motorcycles, and all sorts of vehicles on the road. A major characteristic of VANET is its highly dynamic topology. Nodes are intermittently connected when they encounter one another on the road. If traffic density is low, the proportion of time that a node is connected to another node may be small, which may result in large delay. On the other hand, the instantaneous transmission rate can be very high, especially if transmission proceeds only when two nodes are close to each other. Due to its nature, VANET is particularly suitable for delay tolerant applications with large bandwidth requirement. An example is that a content provider allows its subscribers to download movies, music, or news from an infostation at the roadside when they pass by, and to exchange contents among themselves when they encounter one another on the road. A user can simply run an application programme in the background, without the aware of download schedule. To facilitate the development of these applications, content distribution protocols are needed.

Although many content distribution protocols were developed for wired network, they may not be suitable for all kinds of wireless systems. For example, in wireless ad hoc networks, a packet typically has to traverse multiple hops from source to destination. It was shown in Gupta and Kumar (2000) that per-node throughput changes at a rate $O(1/\sqrt{n \ln n})$, which drops to zero for large *n*. Consequently, the multi-hop strategy is intrinsically unscalable, no matter what protocols are used in the network layer and above. To maintain scalability, a two-hop relay strategy was considered in

Grossglauser and Tse (2002). It was shown that with node mobility, per-node throughput becomes $O(1)$. As a result, system capacity scales linearly with the number of nodes. This drastic difference motivates the design of many mobility-assisted data transfer protocols. Some examples are presented in Papadopouli and Schulzrinne (2001), and Yuen et al. (2003a). In particular, for VANET, a BitTorrent-like protocol called CarTorrent was proposed in Nandan et al. (2005). Protocols were designed in Ahmed and Kanhere (2006), Cataldi et al. (2009), Lee et al. (2006), Li et al. (2011), Sardari et al. (2009a, 2009b) and Wang et al. (2012) based on the idea of network coding (Ahlswede et al., 2000) and rateless code. Investigations on the throughput of a VANET can be found in Blaszczyszyn et al. (2009), Javanmard and Ashtiani (2009), and Wang et al. (2009).

In this paper, we consider a one-dimensional VANET, which can be used to model a highway scenario. Road traffic is modelled by Poisson processes. A multichannel medium access control (MAC) protocol is assumed. Under such a protocol, transmissions between vehicles are assumed to be performed usually different channels so that interference effect can be ignored. For content distribution, we adopt the fountain code approach (MacKay, 2005). Encoding is performed at infostations but not at vehicles. This method can reduce processing time at vehicles, and reduce decoding complexity if a suitable fountain code is used. Based on this model, a performance analysis is performed. This analysis reveals the relationship between coding, delay, and throughput. Besides, exact formulae for throughput are derived, from which insights on how node density affects throughput can be gained. Our approach is similar to that in Yuen et al. (2003b), but with some major differences in modelling.

2 System model

We consider a one-dimensional vehicular network, which models the scenario where many cars are running on a highway. Suppose that a portion of car users subscribes to a content distribution network. The content provider distributes content such as video and audio files via infostations (Frenkiel et al., 2000), which are installed at the entrances of the highway. The highway is divided into equal-length segments. The length of each segment is d and there is an infostation at the ends of each segment. We consider the group of car users who are interested in downloading a common file from the content provider.

2.1 Road-traffic modelling

We assume that cars arrive at the highway following a Poisson process with rate λ. Each of them travels in the highway at constant velocity. Those coming from the left have positive velocity and are collectively called the *forward traffic*. Those coming from the right have negative velocity and are called the *reverse traffic*.

Let V be the velocity of a car. It is a random variable with a certain probability distribution. In this work, we consider two different models for vehicular velocity. The first one is discrete velocity distribution, which means that V can take on values from a finite set, $\{v_1, v_2, \ldots, v_M\}$, with probability $p_1, p_2, \ldots, p_M$, respectively, where $\sum_{m=1}^{M} p_m = 1$. We denote the set $\{1, 2, \ldots, M\}$ by $\mathcal{M}$. This model is applicable to the scenario where the highway consists of M lanes, including both the forward and reverse directions. Nodes using different lanes are of different speeds. A node, when entering the highway, chooses a suitable lane and then travels in that lane without changing its speed.

The second model is continuous velocity distribution. This model, called the *wide motorway model* in Kingman (1993), is applicable to the scenario where there are multiple lanes and moderate traffic. Since nodes can overtake others at different lanes, there is no interaction among the nodes even if they travel in the same direction. A node can have any speed the driver likes, subject to the speed limit. We let the probability density function of the continuous random variable V be $f_V(v)$, defined for $v \in [a, b]$. We divide the interval $[a, b]$ into many intervals, each of length Δv. Each interval is approximated by a constant function. We assume that f_V is a continuous function (and hence uniformly continuous over $[a, b]$), so that we can approximate f_V as close as we like by increasing the number of intervals.

2.2 Connectivity and medium access

When a car comes close to an infostation, it can download information from it. Besides, a car can exchange information with another car in proximity. We refer a car or an infostation as a *node* and say that a *node encounter* occurs when two nodes are approaching to within a transmit range r from each other. In other words, two nodes are connected if the distance between them is less than or equal to r.

We adopt a multichannel MAC protocol (So and Vaidya, 2004). It means that the available bandwidth is divided into a certain number of channels. A node who has data to transmit selects an idle channel randomly. We assume that the car density is low or moderate so that the number of nodes connected together does not exceed the number of available channels. In that case, an idle channel is always available and each node can transmit its data without interfering other nodes. The concept is similar to frequency division multiple access, but channel selection is done in a distributed way through carrier sensing. Besides, full duplex transmission is assumed. A node can transmit data to other nodes and receive data from them at the same time using different channels. For example, if there are three nodes within the transmit range of a certain node, then that node can broadcast data to all the three nodes and at the same time, receive data from them.

The amount of data received through a particular channel during a node encounter depends on the transmission bit rate R_b and the connection time. We assume that non-adaptive radio is used so that R_b is constant throughout the encounter period. We also assume forward channel coding is used so that the probability of transmission error is negligible. The maximum number of packets that can be received from a particular channel during an encounter is R_pT_c, where T_c is the connection time and R_p is the transmission rate in packets per second and is equal to R_b divided by the packet size in bits. The connection time between two nodes, T_c, depends on their relative speed and is given by

$$T_c = \frac{2r}{|v - v'|}, \tag{1}$$

where v and v' are the velocities of the two nodes. Note that the difference of velocity $v - v'$ may be negative, and the sign depends on their directions. Besides, the velocity of an infostation is zero.

2.3 Coding scheme and forwarding strategy

The file is split into K smaller blocks $W_1, W_2, \ldots, W_K$, each of which consists of L bits. We adopt a fountain code approach (MacKay, 2005) for file distribution at infostations. When a car is within the transmit range of an infostation, the infostation generates and transmits some encoded messages to the car. Each encoded message is obtained by linearly combining the original message blocks:

$$\sum_{k=1}^{K} c_k W_k, \tag{2}$$

where each c_k is either 0 or 1, and the addition is performed over $\mathbb{F}_2$. The vector $c = (c_1, c_2, \ldots, c_K)$ is called the *encoding vector*, which is generated randomly. There are various ways to generate it. One simple way is to pick a vector uniformly at random over $\mathbb{F}_2^K$. Another way is to generate it according to the robust soliton distribution in LT codes (Luby, 2002). Each packet consists of an encoded message as in (2) and the corresponding encoding vector.

The protocol that we propose for packet exchange follows a two-hop strategy. When two cars are within the transmit range of each other, they will exchange those packets that were directly downloaded from infostations. Those packets that are received from other cars will not be forwarded again. In other words, each packet is transmitted in at most two hops: from an infostation to a car, and from that car to another car. Besides, it should be noted that encoding is not performed at cars, but at infostations only. This feature is desirable from a practical standpoint, since the processors at cars may not have a high processing speed.

A car can recover the original file if the encoding vectors in the received packets span the vector space $\mathbb{F}_2^K$, which happens when K of the encoding vectors received are linearly independent. Indeed, if $c_1, c_2, \ldots, c_K$ are encoding vectors that are linearly independent, the file can be decoded by inverting the $K \times K$ matrix whose i^{th} row is c_i for $i = 1, 2, \ldots, K$.

The number of packets that must be received before K linearly independent encoding vectors are obtained depends on the probability distribution of encoding vectors. If random linear fountain code is used, i.e., when the encoding vectors are uniformly generated, the original file can be decoded with probability at least $1-\epsilon$, for some small constant ϵ, after $K + \log_2(1/\epsilon)$ packets are received (MacKay, 2005, Section 3). If we use LT code with robust soliton distribution, the number of packets needed is $K + 2S\log_2(S/\epsilon)$, where $S = c\sqrt{K}\log_e(K/\delta)$ and c is a parameter of order 1

(MacKay, 2005; Section 5). Given the probability of decoding failure ϵ, the downloading time is obtained by dividing the required number of packets by the packet rate.

3 Throughput analysis

As there is an infostation at every entrance of the highway, when a car enters the highway system, it collects some encoded message blocks from the infostation. As cars usually enter the highway at low speed, they should have picked up enough packets to be exchanged during any future encounter. Since the velocity of each node is assumed constant, two nodes meet each other at most once as they travel along the highway. We can therefore guarantee that any newly received packet by a car is statistically independent of the packets already stored in its buffer. Consequently, the packets received by a vehicle are all statistically independent. From the point of view of a particular car, the incoming packets are statistically the same as if there were only one fountain source. This decouples the mobility and traffic model from the analysis of fountain code.

In this section, our objective is to estimate the average downloading time of the file in VANET by analysing the packet rate. In the sequel, we will call it *throughput*. We will first consider the case where the velocity distribution is discrete, and then extend the results to the continuous case.

3.1 Discrete velocity distribution

We consider a specific node, called the *observer node*, or simply the *observer*, travelling between a segment of highway between two consecutive infostations A and B. We will analyse the throughput of the observer in this segment of the highway.

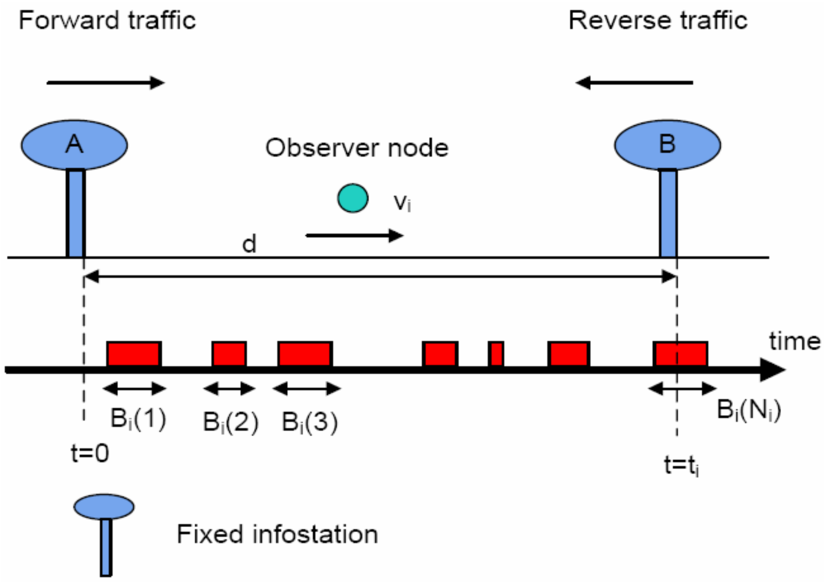
Suppose that the observer belongs to class i for some $i \in \mathcal{M}$, and moves at speed v_i in the forward direction from A to B. The travelling time of the observer in this segment is given by $t_i = d / v_i$. We denote N_i as the number of *node encounters* for the observer when travelling in this segment of highway. Furthermore, for $k = 1, 2, \ldots, N_i$, we denote $B_i(k)$ as the number of packets received from the k^{th} encounter. The total number of packets received by the observer in this highway segment is

$$B_i = \frac{R_p 2r}{v_i} + \sum_{k=1}^{N_i} B_i(k). \tag{3}$$

The first term corresponds to the packets directly downloaded from infostation A and the second the total number of packets from other vehicles.

It is possible that the amount of data received from the last encounter may be less than $B_i(N_i)$ if the connection time of that encounter overshoots the end of the travelling time (see Figure 1). Nevertheless, this boundary effect is negligible when the length of the highway segment is large.

Figure 1 Illustration of the VANET model (see online version for colours)



In order to find the expected value of B_i, we split the Poisson arrival process into M independent Poisson streams with rate $p_m\lambda$, where $m = 1, 2, \ldots, M$. Let $\tilde{N}_{i,m}$ be the number of encounters of the observer with nodes in class m, so that

$$N_i = \tilde{N}_{i,1} + \tilde{N}_{i,2} + \ldots + \tilde{N}_{i,M}, \tag{4}$$

where $\tilde{N}_{i,i}$ is defined as zero. The following lemma gives the expected value of $\tilde{N}_{i,m}$.

Lemma 1: $\tilde{N}_{i,m}$ is Poisson distributed with mean

$$E\left[\tilde{N}_{i,m}\right] = \lambda p_m \left|t_m - t_i\right|, \tag{5}$$

where $t_m = d / v_m$.

Proof: Without loss of generality, suppose the observer enters the highway segment at time 0 and departs at time t_i. We consider its encounter with forward traffic and reverse traffic separately.

For forward traffic, consider a node of velocity $v_m > 0$, which enters the highway segment at time t and departs at time $t + t_m$. Suppose the speed of the node is lower than that of the observer, that is, $t_m > t_i$. It will encounter the observer if and only if it enters the highway earlier than the observer (i.e., $t < 0$) and it departs the highway later than the observer (i.e., $t + t_m > t_i$). In other words, an encounter occurs if and only if $-(t_m - t_i) < t < 0$. Since the arrival process is Poisson with rate λp_m and an encounter occurs when the arrival time of a node falls within an interval of length $| t_m - t_i |$, the number of encounters is Poisson distributed with mean $\lambda p_m(t_m - t_i)$. Next suppose $t_m < t_i$. An encounter occurs if and only if the node enters after time 0 (i.e., $t > 0$) and it departs before t_i (i.e., $t + t_m < t_i$). Again, the number of encounters is Poisson distributed with mean $\lambda p_m| t_m - t_i |$.

For reverse traffic, consider a node of velocity $v_m < 0$. If it enters the highway before time 0 (i.e., $t < 0$), it will encounter the observer if $t + | t_m | > 0$. If it enters the highway after time 0 (i.e., $t > 0$), it will encounter the observer if it enters before t_i (i.e., $t < t_i$). Combining the two cases, we can see that an encounter occurs if $-| t_m | < t < t_i$. Hence, the number of encounters is Poisson distributed with mean also equal to $\lambda p_m| t_m - t_i |$. □

Let $\mathcal{M}_{-i}$ be the set $\mathcal{M} \setminus \{i\}$. We next obtain an expression for the mean of B_i.

Lemma 2:

$$E[B_i] = \frac{R_p 2rt_i}{2}\left[1 + \frac{\lambda}{2}\sum_{m\in\mathcal{M}_{-i}} p_m |t_m|\right]. \tag{6}$$

Proof: The observer will only encounter a node in class m for $m \neq i$. When the observer meets another node of velocity v_m, $v_m \neq v_i$, the number of packets received is equal to $R_p 2r / | v_m - v_i |$. We sum over all $m \in \mathcal{M}_{-i}$ and obtain

$$B_i = \frac{R_p 2r}{v_i} + \sum_{m\in\mathcal{M}_{-i}} \frac{\tilde{N}_{i,m} R_p 2r}{|v_m - v_i|}. \tag{7}$$

Taking expectation and using Lemma 1, we have

$$E[B_i] = \frac{R_p 2r}{v_i} + \sum_{m\in\mathcal{M}_{-i}} \frac{E\left[\tilde{N}_{i,m}\right] R_p 2r}{|v_m - v_i|} \tag{8}$$

$$= R_p 2r\left[\frac{1}{v_i} + \sum_{m\in\mathcal{M}_{-i}} \frac{\lambda p_m |t_m - t_i|}{|v_m - v_i|}\right] \tag{9}$$

$$= R_p 2r\left[\frac{t_i}{d} + \sum_{m\in\mathcal{M}_{-i}} \frac{\lambda p_m t_i |t_m|}{d}\right]. \tag{10}$$

Note that t_m can be negative due to the existence of the reverse traffic. □

Define $C_i = B_i / t_i$ as the average throughput of the observer during its travelling time on the highway segment. Then we have

$$E[C_i] = \frac{R_p 2r}{d}\left[1 + \lambda \sum_{m\in\mathcal{M}_{-i}} p_m |t_m|\right]. \tag{11}$$

Consider a particular time instant t. A car of velocity v_m will be on this highway segment if it enters this segment within the interval $[t - | t_m |, t]$. Therefore, the number of cars of velocity v_m that are on the highway is Poisson distributed with mean equal to $\lambda p_m| t_m |$. The density of the cars belonging to class m, denoted by ρ_m, is equal to $\lambda p_m| t_m | / d$. Since there is one infostation per highway segment, we define the density of infostations on the highway as ρ_0, which is equal to $1 / d$. The above equation can then be rewritten in terms of car density as follows:

Theorem 3:

$$E[C_i] = R_p 2r\left(\rho_0 + \sum_{m\in\mathcal{M}_{-i}} \rho_m\right). \tag{12}$$

The first term within the parenthesis in Theorem 3 is the density of infostation in the highway segment, and the second term is the sum of car densities over all classes except the observer's class. It is interesting to find that the individual throughput depends only on the density of other nodes. Note that the density of nodes belonging to the same class is irrelevant because there will not be any intra-class encounter.

The per-node throughput can also be expressed as

$$E[C_i] = R_p 2r\left(\sum_{m=0}^{M} \rho_m - \rho_i\right). \tag{13}$$

We observe the following:

- *Low-density gain*: The class of cars that has the lowest density get the largest average per-node throughput.

Now let C be the average per-node throughput. By averaging the per-node throughput in Theorem 3 over all velocity classes, we have

$$E[C] = \sum_{i=1}^{M} p_i E[C_i] \tag{14}$$

$$= R_p 2r\left[\rho_0 + \sum_{i=1}^{M} p_i \left(\sum_{m\in\mathcal{M}_{-i}} \rho_m\right)\right], \tag{15}$$

which can be rewritten in the following two equivalent forms:

Theorem 4:

$$E[C] = R_p 2r\left(\sum_{m=0}^{M} \rho_m - \overline{\rho}\right), \tag{16}$$

$$= R_p 2r\left[\frac{1}{d} + \frac{\lambda}{2}\sum_{i\neq j} p_i p_j \left(\frac{1}{|v_i|} + \frac{1}{|v_j|}\right)\right], \tag{17}$$

where $\overline{\rho} = \sum_{i=1}^{M} p_i \rho_i$.

Based on the above result, the following facts can be observed:

- *Incrementally linear scalability*: The average per-node throughput increases with the node arrival rate, λ, in an incrementally linear fashion.

- *Mobility reduces throughput*: If all cars move faster, then the average per-node throughput decreases. For example, suppose all cars double their speeds. Then the car density of each velocity class decreases by one half. According to Theorem 3, the throughput of all users decreases. Hence, the system throughput decreases.

Although the velocity of the cars cannot be controlled by the system, it is interesting to know which probability mass function maximises system throughput, for a given velocity vector $\{v_1, v_2, \ldots, v_M\}$. We answer this question in the Appendix.

3.2 Continuous velocity distribution

The analysis for discrete velocity can be extended to the case where the velocity distribution is continuous. The next theorem is analogous to Theorems 3 and 4.

Theorem 5: Let C_i denote the throughput of a particular observer node with velocity v_i, and C the average per-node throughput. Let ρ be the density of cars on the highway. Then, for all i,

$$E[C] = E[C_i] = R_p 2r\left(\rho_0 + \lambda E\left[\frac{1}{|V|}\right]\right) \tag{18}$$

$$= R_p 2r(\rho_0 + \rho). \tag{19}$$

Proof: We partition $[a, b]$ into M intervals and apply the result for the discrete-velocity case. Following (11), we have

$$E[C_i] = R_p 2r\left(\rho_0 + \lambda \sum_{m \in \mathcal{M}_{-i}} \frac{p_m}{v_m}\right) \tag{20}$$

When M approaches infinity, the above equation becomes

$$E[C_i] = R_p 2r\left(\rho_0 + \lambda \int_a^b f_V(v) \frac{1}{|v|} dv\right). \tag{21}$$

Note that the above expression does not depend on v_i. Therefore, (18) is proven. Let N_m be the number of class-m cars on the highway segment of length d. Note that N_m is Poisson distributed with mean $\lambda p_m| t_m |$. Then we rewrite (20) as

$$E[C_i] = R_p 2r\left(\rho_0 + \sum_{m \in \mathcal{M}_{-i}} \frac{E[N_m]}{d}\right). \tag{22}$$

Let N be the number of cars on the highway segment, i.e., $N = \sum_m N_m$. When M goes to infinity, we have

$$R_p 2r\left(\rho_0 + \frac{E[N]}{d}\right). \tag{23}$$

Note that $E[N] / d$ is the density of cars on the highway. □

Based on this theorem, we have the following observations. The first one is the same as that in the case of discrete speed. The second one is similar but not exactly the same. The last two observations are different.

- *Incrementally linear scalability*: The average per-node throughput increases with the node arrival rate, λ, in an incrementally linear fashion.
- *Mobility reduces throughput*: The average per-node throughput changes at a rate $O(E[1 / | V |])$. It means that the higher the mobility, the lower the car density, and the lower the average per-node throughput.
- *Perfect fairness*: $E[C_i]$ is independent of v_i. It means that given the same background traffic on the highway, the throughput of a node is independent of its own speed. In other words, all nodes yield the same average throughput, which is different from the case of discrete speed.
- *Equivalence of forward and reverse traffics*: The average throughput of a particular node yielded by encountering with forward traffic is the same as that with reverse traffic, provided that the arrival rates and speed distributions of the two directions are the same.

4 Conclusions

We have analysed the throughput performance of a VANET. Based on the Poisson arrival process, we derive simple formulae for throughput under both discrete and continuous velocity distribution. There are two major results: First, the average per-node throughput increases linearly with the arrival rate of vehicles, λ. As the number of users on the highway is directly proportional to λ, we conclude that system throughput increases quadratically with λ. In other words, the system is quadratically scalable. Second, the average per-node throughput decreases when all vehicles increase their speeds, implying that higher overall mobility is not beneficial.

We have also investigated the throughput of individual nodes. For the discrete velocity case, the class of nodes having lower density has higher throughput. In contrast, for the continuous velocity case, all nodes have the same throughput.

References

Ahlswede, R., Cai, N., Li, S-Y.R. and Yeung, R.W. (2000) 'Network information flow', *IEEE Trans. Inform. Theory*, July, Vol. 46, No. 4, pp.1204–1216.

Ahmed, S. and Kanhere, S.S. (2006) 'VANETCODE: network coding to enhance cooperative downloading in vehicular ad-hoc networks', in *Proc. ACM Int. Conf. on Commun. and Mobile Comput.*, pp.527–532.

Blaszczyszyn, B., Muhlethaler, P. and Toor, Y. (2009) 'Maximizing throughput of linear vehicular ad-hoc networks (vanets) – a stochastic approach', in *Proc. European Wireless Conference*, May, pp.32–36.

Cataldi, P., Tomastis, A., Grilli, G. and Gerla, M. (2009) 'A novel data dissemination method for vehicular networks with rateless codes', in *IEEE Wireless Comm. and Networking Conf. (WCNC)*, Budapest, April, pp.1–6.

Frenkiel, R.H., Badrinath, B.R., Borras, J. and Yates, R. (2000) 'The infostations challenge: balancing cost and ubiquity in delivering wireless data', *IEEE Personal Communications*, April, Vol. 7, No. 2, pp.66–71.

Grossglauser, M. and Tse, D. (2002) 'Mobility increases the capacity of ad hoc wireless networks', in *IEEE/ACM Trans. Networking*, August, Vol. 10, No. 4, pp.477–486.

Gupta, P. and Kumar, P.R. (2000) 'The capacity of wireless networks', *IEEE Trans. Inform. Theory*, March, Vol. 46, No. 2, pp.388–404.

Javanmard, A. and Ashtiani, F. (2009) 'Analytical evaluation of average delay and maximum stable throughput along a typical two-way street for vehicular ad hoc networks in sparse situations', *Computer Communications*, October, Vol. 32, No. 16, pp.1768–1780.

Kingman, J.F.C. (1993) *Poisson Processes*, Oxford University Press, New York.

Lee, U., Park, J-S., Yeh, J., Pau, G. and Gerla, M. (2006) 'CodeTorrent: content distribution using network coding in VANET', in *Proc. of the 1st Int. Workshop on Decentralized Resource Sharing in Mobile Computing and Networking*, ACM Press, Los Angeles, California, pp.1–5.

Li, M., Yang, Z. and Lou, W. (2011) 'CodeOn: cooperative popular content distribution for vehicular networks using symbol level network coding', *IEEE J. on Sel. Areas in Comm.*, January, Vol. 29, No. 1, pp.223–235.

Luby, M. (2002) 'LT codes', in *Proc. of the 43rd Annual IEEE Symp. on Foundations of Comp. Sci. (FOCS'02)*, pp.271–282.

MacKay, D.J.C. (2005) 'Fountain codes', *IEE Proc. Commun.*, December, Vol. 152, No. 6, pp.1062–1068.

Nandan, A., Das, S., Pau, G., Sanadidi, M. and Gerla, M. (2005) 'Cooperative downloading in vehicular ad hoc wireless networks', in *Proc. of IEEE/IFIP Int. Conf. Wireless on Demand Network Systems and Services*, pp.32–41.

Papadopouli, M. and Schulzrinne, H. (2001) 'Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices', in *Proc. of the 2nd ACM Int. Symp. on Mobile Ad Hoc Networking and Computing*, Long Beach, California, pp.117–127.

Sardari, M., Hendessi, F. and Fekri, F. (2009a) 'DMRC: dissemination of multimedia in vehicular networks using rateless codes', in *IEEE Infocom Workshops*, Rio de Janeiro, April, pp.1–6.

Sardari, M., Hendessi, F. and Fekri, F. (2009b) 'Infocast: a new paradigm for collaborative content distribution from roadside units to vehicular networks', in the *6th Annual IEEE Comm. Society Conf. on Sensor, Mesh and Ad Hoc Comm. and Networks*, Rome, June, pp.1–9.

So, J. and Vaidya, N.H. (2004) 'Multi-channel MAC for ad hoc networks: handling multi-channel hidden terminals using a single transceiver', in *Prof. of the 5th ACM Int. Symp. on Mobile Ad Hoc Networking and Computing*, pp.222–233.

Wang, Q., Fan, P. and Letaief, K.B. (2012) 'On the joint V2I and V2V scheduling for cooperative VANETs with network coding', *IEEE Trans. on Veh. Tech.*, January, Vol. 61, No. 1, pp.62–73.

Wang, Z., Lu, W. and Jin, W. (2009) 'A study on information throughput of inter-vehicle communications in a unidirectional traffic stream', in *Proc. Communications and Mobile Computing*, January, pp.396–401.

Yuen, W.H., Yates, R.D. and Mau, S-C. (2003a) 'Exploiting data diversity and multiuser diversity in mobile infostation networks', in *Proc. IEEE INFOCOM*, April, Vol. 3, pp.2218–2228.

Yuen, W.H., Yates, R.D. and Sung, C.W. (2003b) 'Effect of node mobility on highway mobile infostation networks', in *Proc. ACM Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, September, pp.82–91.

Appendix

Optimal probability mass function for the case of discrete velocity

Given $\{v_1, v_2, \ldots, v_M\}$, we would like to know what the optimal probability mass function is. The problem can be formally stated as follows:

$$\text{Maximise } F(\boldsymbol{p}) = \sum_{i \neq j} p_i p_j \left(\frac{1}{|v_i|} + \frac{1}{|v_j|} \right)$$

$$\text{subject to } \sum_{m=1}^{M} p_m = 1, \text{ and } p_i \geq 0 \ \forall i.$$

To check whether this is a convex optimisation problem, we first eliminate the equality constraint. We consider the function

$$G(p_1, \ldots, p_{M-1}) \equiv F\left(p_1, \ldots, p_{M-1}, 1 - \sum_{i=1}^{M-1} p_i \right). \tag{24}$$

It can be shown that the Hessian of G is given by

$$-2diag\left(\frac{1}{|v_1|}, \ldots, \frac{1}{|v_{M-1}|} \right) - \frac{2}{|v_M|} \boldsymbol{J}_{M-1}, \tag{25}$$

where $diag(\boldsymbol{x})$ is the diagonal matrix with diagonal elements given by $\boldsymbol{x}$, and $\boldsymbol{J}_n$ is the $n \times n$ all-one matrix. Since the first matrix is negative definite and the second one is negative semidefinite, the Hessian of G is negative definite. Hence, the function G is strictly concave, and there is one unique optimal point, $\boldsymbol{p}^*$.

Proposition 6: If $|v_1| \leq |v_2| \leq \cdots \leq |v_m|$, then at the optimal point $\boldsymbol{p}^*$, we must have

$$p_1 \geq p_2 \cdots \geq p_M. \tag{26}$$

Proof: Suppose $p_i^* < p_j^*$, where $i < j$. Consider another point $\boldsymbol{p}'$, with all components the same as $\boldsymbol{p}^*$ except the i^{th} and the j^{th} components swapped. By definition, it can be seen that $F(\boldsymbol{p}^*) < F(\boldsymbol{p}')$. □

Now we try to solve the optimisation problem. Introduce the Lagrangian multipliers λ_i's, $i = 1, 2, \ldots, M$, for the non-negative constraints and ν for the equality constraint. The KKT conditions after eliminating λ_i's become

$$p_i \left(\nu - \sum_{k \neq i} p_k \alpha_{ik} \right) = 0 \quad i = 1, \ldots, M, \tag{27}$$

$$\sum_{m=1}^{M} p_m = 1, \text{ and } p_i \geq 0 \quad i = 1, \ldots, M, \tag{28}$$

where $\alpha_{ik} = 1 / |v_i| + 1 / |v_k|$.

Let $\boldsymbol{A}$ be the $M \times M$ matrix whose diagonal components are all zero and $(i, j)^{\text{th}}$ component equal to α_{ik} for $i \neq k$. Let $\boldsymbol{A}_k$ be its leading principal submatrix of order k, that is, its last $M - k$ rows and $M - k$ columns are deleted. Let $\mathbf{1}_k$ be the $k \times k$ all-one vector, and $\boldsymbol{p}_k$ be the first k components of $\boldsymbol{p}$. Without loss of generality, we assume that $| v_1 | \leq | v_2 | \leq \cdots \leq | v_m |$. The optimal solution can be found by the following algorithm.

Algorithm 1 (Initialisation): Let $n := M$.

1 Compute

$$\boldsymbol{p}_n = \frac{\boldsymbol{A}_n^{-1}\mathbf{1}_n}{\left\|\boldsymbol{A}_n^{-1}\mathbf{1}_n\right\|_1},$$

where $\|\cdot\|_1$ is the l_1 norm.

2 If all components of $\boldsymbol{p}_n$ are non-negative, then output $\boldsymbol{p}_M$. Otherwise, let $p_n := 0$ and $n := n - 1$. Repeat Step 1.

This algorithm produces the correct solution because it simply tries to solve (27), assuming that $p_i \neq 0$ for all i. Note that the multiplier ν is adjusted such that the equality constraint is satisfied. This corresponds to the normalisation factor in the algorithm. If all the components are non-negative, then the KKT conditions are satisfied. Otherwise, one of the p_i's must be zero because of (27). By Proposition 6, the last component must be zero. Hence, we reduce the dimension of the problem by one and then repeat.

For example, we consider the situation where the nodes can be divided into five classes. Given the velocity values, we can compute the optimal probability mass function. Three examples are shown in Table 1. It can be seen that some of the p_i's can be equal to zero. However, this occurs only for some extreme cases. We have tested many other cases. Typically, all p_i's will be greater than zero.

Table 1 Optimal probability mass functions for different velocity values

Speed	80	90	100	110	120
Distribution	0.26	0.23	0.2	0.17	0.14
Speed	50	60	70	80	130
Distribution	0.3077	0.2692	0.2308	0.1923	0
Speed	20	30	40	110	120
Distribution	0.3889	0.3333	0.2778	0	0

Besides, it can be shown that at least two classes must have probabilities strictly greater than zero, for otherwise no encounter in the system can occur. According to Proposition 6, they must be p_1 and p_2. If there are only two classes or $p_i = 0$ for $i \geq 3$, then it can be easily shown that $p_1^* = p_2^* = 0.5$ is optimal.