

## DDRC: Data Dissemination in Vehicular Networks Using Rateless Codes\*

MOHSEN SARDARI, FARAMARZ HENDESSI<sup>†</sup> AND FARAMARZ FEKRI

*School of Electrical and Computer Engineering  
Georgia Institute of Technology  
Atlanta, GA 30332, U.S.A.*

*E-mail: {mohsen.sardari; fekri}@ece.gatech.edu*

<sup>†</sup>*Department of Electrical and Computer Engineering  
Isfahan University of Technology  
Isfahan, Iran*

*E-mail: hendessi@cc.iut.ac.ir*

In this paper, we address the problem of delivering multimedia content to a sparse vehicular network from roadside info-stations, using efficient vehicle-to-vehicle collaboration. Due to the highly dynamic nature of the underlying vehicular network topology, we depart from architectures requiring centralized coordination, reliable MAC scheduling, or global network state knowledge, and instead adopt a distributed paradigm with simple protocols. In other words, we investigate the problem of reliable dissemination from multiple sources when each node in the network shares a limited amount of its resources for cooperating with others. By using *rateless* coding at the Road Side Unit (RSU) and using vehicles as data carriers, we describe an efficient way to achieve reliable dissemination to all nodes (even disconnected clusters in the network). In the nutshell, we explore vehicles as mobile storage devices. We then develop a method to keep the density of the rateless coded packets as a function of distance from the RSU at the desired level set for the target decoding distance. We investigate various tradeoffs involving buffer size, maximum capacity, and the mobility parameter of the vehicles.

**Keywords:** ad-hoc networks, broadcasting, vehicular networks, data dissemination, rateless codes

### 1. INTRODUCTION

Vehicular ad hoc networks (VANETs) have recently received considerable attention. VANETs are envisioned to provide us with several services in near future. Vehicular communications have many different facets. Applications range from safety support [2], to entertainment for passengers, to local news delivery and advertisement [3]. Advertisements are one of the most important sources of revenue for companies. Conventional multimedia advertisement has moderate amount of data with lenient latency constraints. The advertisement application is a type of data dissemination from an information source to a large number of clients. In this work, by considering a simple but basic case we take some steps to answer questions concerning data dissemination in general vehicular networks such as the impact of mobility on throughput and reliability, robustness, and latency supported by such networks. Specifically, we will address those questions in the context of disseminating information packets from a large array of Roadside Units (RSU)

---

Received March 31, 2009; accepted September 30, 2009.

Communicated by Chih-Yung Chang, Chien-Chung Shen, Xuemin (Sherman) Shen, and Yu-Chee Tseng.

\* Results in this paper have appeared in part in [1].

to a bidirectional linear highway vehicular network. We assume each RSU is an independent source having a block of information packets to disseminate to all vehicles. Ideally, every vehicle must recover information data belonging to each RSU at any distance. However, this is not feasible given the resources. Hence, we are particularly interested in optimizing the distance from the source (RSU) where a typical vehicle can recover its data. This distance depends on the maximum throughput that can be achieved using efficient collaborative store-carry-forward routing. The application motivated this problem is the sale advertisement by the stores in an urban area.

Although disseminating data has been visited in the past in the network community many unique characteristics of the VANET open up new research challenges. First, VANETs can be considered as a category of partitioned ad hoc networks [4-7]. Since density of vehicles is highly variable with space and time, the network changes from a sparsely disconnected network to a densely connected one in a short period of time. As a result of these topology variations, traditional routing and forwarding methods do not perform well in VANETs. Furthermore, many structures for efficient data dissemination such as trees and grids are extremely hard to set up and maintain in VANETs. In this paper we present a new approach that merges the *vehicle-to-vehicle* and *roadside-to-vehicle* communication topologies in order to support reliable data dissemination without the need of complex routing protocols. We suggest the application of a new class of packet-level coding schemes referred as *rateless* codes (section 3) for the reliable and efficient data dissemination in VANETs. Several aspects of rateless codes make them suitable for such applications. First, their rateless property avoids issues regarding the choice of rates even in the presence of varying link loss conditions. Second, rateless codes require very low coding overhead to recover the message while having low encoding and decoding complexity [8]. Using simulations, we show that the gains of rateless coding over classical store-and-forward multihop routing strategies is significant, as measured by the number of packets received at a vehicle as a function of the distance between the vehicle and the nearest RSU. We discuss as to how the limited resources (*e.g.* buffer size) available at each node in the network must be utilized for cooperation efficiently. Physical layer details, specific protocol formats, and all other details that are not relevant to the fundamental properties are coarsely modeled, while mobility will be addressed in more detail.

The rest of the paper is organized as follows. In section 3 we provide an introduction to rateless codes. Section 4 presents network assumptions, models and the details of our proposed scheme. Simulation results and discussion follow in section 5. We present future directions and conclude our paper in section 6.

## 2. RELATED WORK

The problem of distributing data from an info-station to vehicles on a highway is examined in [9]. The paper focuses on a simplified scenario where mobile nodes adopt a slotted ALOHA MAC strategy to cooperatively distribute content provided by a single info-station. The proposed architecture is designed for a dense network and will fail under realistic vehicle mobility. In [10], the performance of network coding as a solution for video streaming in vehicular networks is investigated. The paper adopts the 802.11 DCF MAC with network coding and studies, by means of computer simulations, its packet

delivery ratio compared to the on demand multicast routing protocol (ODMRP) [11].

Based on the structure of the network and the amount of data generated in the network, broadcasting easily leads to severe congestion and significantly reduces the data delivery ratio. To solve such problems, Zhao *et al.* [12] proposed Data Pouring (DP) and buffering on the road scheme. The DP scheme explores the partially predictable vehicle mobility limited by the road layout. In this scheme it was assumed that the vehicle density on the road is large enough to maintain network connectivity.

It should be noted that all these results were obtained in a scenario characterized by single source and infinite storage buffers. The problem of data dissemination involving multiple sources with finite resources has not been investigated. In this work we include the effect of limited resources for cooperation by considering a limited cooperation-buffer at each node in the network. We note that although the vehicles may have large storage buffers available, they may only use a small portion for helping any particular RSU. Our work is motivated by some of the fundamental questions that rise here:

- (1) Is there any scheme that can handle both dense and sparse scenarios?
- (2) What is the optimal strategy for deleting packets from the buffers to satisfy the spatial relevance of the information?
- (3) What is the optimal method to use the buffers to handle a multisource dissemination scenario?
- (4) What are broadcast throughput and reliability and how one can improve them?

We try to take initial steps towards answering such questions.

### 3. OVERVIEW OF RATELESS CODES

Also known as fountain codes, rateless codes were first conceived by Michael Luby [8]. As the name suggests, these codes are unlike conventional codes generated from algebraic and combinatorial means in the sense that they do not possess any fixed rate. From a finite set of data packets, the encoder can potentially generate an infinite stream of encoded packets. In [8], it was also shown that for  $k$  data (information) packets, on the average, the destination requires  $k\Gamma_k$  encoded packets, where  $\Gamma_k = 1 + O(k^{-1/2} \log \frac{k}{\delta})$  is the overhead, to decode all the  $k$  data packets with a probability of  $1 - \delta$ . Moreover, the encoding and decoding processes introduce very low computational complexity and are performed in the following manner. A parameter that is key in the design of rateless codes is the degree distribution polynomial  $\Omega(x) = \sum_{1 \leq x \leq k} \Omega(i)x_i$  where  $\Omega(i) \in [0, 1]$  for  $i = 1, \dots, k$ . This degree distribution induces a probability distribution on the set of data packets  $\{p_1, \dots, p_k\}$  in the following manner. For any subset  $V$  of packets,  $P_\Omega(V) = \frac{\Omega(|V|)}{\binom{k}{|V|}}$ .

To generate a packet, the encoder generates an instance of a random variable  $Z$  that selects each subset  $V$  of packets with the aforementioned probability. Such a selection can be effected by equivalently selecting the weight  $|V|$  of the selection using the distribution  $\Omega$  and then selecting  $|V|$  packets uniformly at random from set of  $k$  data packets. To generate the encoded packet, the encoder does a packet-level XOR of the selected packets and appends each packet with the indices (or IDs) of all the packets XORed to generate

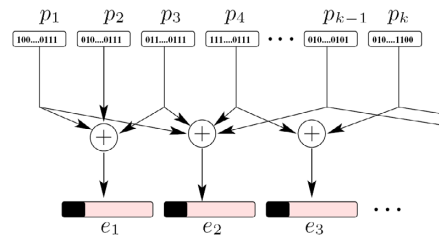


Fig. 1. Encoding at each source.

the encoded packet. Fig. 1 illustrates the encoding process. In this example, the encoded packets  $e_1$ ,  $e_2$  and  $e_3$ , are formed by selecting 3, 4 and 2 packets, respectively, from the set of  $k$  information packets. In each encoded packet  $e_i$ , the dark portion (called the packet overhead) contains information of the indices of the packets used for generating  $e_i$ .

To decode the data packets from the received packets  $e_i$ , the decoder employs iterative message passing algorithm. To decode  $k$  data packets, at least  $k$  encoded packets must be collected at the receiver. However, in practical coding schemes with small  $k$  (of the order of  $10^3$ ), more than  $k$  packets are needed for successful decoding with high probability. The ratio of the required number of distinct encoded packets to the number of data packets  $k$  is referred as the *coding overhead*  $\Gamma_k$ .

#### 4. PROPOSED APPROACH FOR DATA DISSEMINATION IN VANETS

In this section, we introduce various aspects of our proposed scheme. First, we present the coding technique which play a key role in the dissemination problem. Next, we describe the packet transfer protocol, *i.e.*, the set of rules that govern the packet transfer during contacts.

##### 4.1 Efficiency and Packet Transfer Technique

The conventional approach to reduce collisions and the hidden node problem in network is using the request to send/clear to send (RTS/CTS) handshakes. This scheme can improve the data delivery ratio. However, bandwidth will still be wasted by back off timers, control messages, and RTS/CTS handshakes, which effectively reduce the dissemination throughput and increase the latency for time sensitive applications.

Vellambi *et al.* [13] showed that by applying rateless codes at the source where message is generated, even in the presence of packet expiry and intermittent connectivity, one can effect reliable message delivery with improved latency in the unicast scenario. For dissemination scenarios, likewise, we expect that we can gain similar improvements by applying rateless codes.

In [14], authors considered schemes where data is first encoded with a replication factor of  $r_p$  and then packetized into  $sr_p$  chunks for some integer  $s$ . Simulations in [14] reveal the superiority of schemes using erasure-coding over simple replication although there are several drawbacks with their approach. First, employing a fixed-rate erasure

coding scheme raises this question as, “what rate is optimal?”. It is clear that the higher the redundancy, the better the expected delay. However, this comes with the cost of using more network resources such as bandwidth and storage capacity. Moreover, realistic assumptions such as finite packet expiry and finite buffer sizes make this scheme practically inefficient.

#### 4.2 DDRC

In order to resolve the issues we discussed before, we introduce *DDRC*. In this approach, when a message arrives at an RSU, the RSU packetizes the message into smaller data packets. These packets are then encoded into a set of slightly bigger size using the described rateless encoding scheme. Then the RSU broadcasts the set of encoded packets. As shown in Fig. 2, for each RSU, we divide the vehicles on the road into two classes:

**Collector.** These are the vehicles that are approaching toward a specific RSU.

**Carriers.** These are the vehicles that have been successful in decoding a specific RSU’s message (by collecting sufficient packets either directly from the RSU or from other carrier nodes). Note that as shown in Fig. 2, the distance between carrier vehicles and the RSU increases with time. The carrier node broadcasts the RSU’s message after encoding with rateless code as RSU. This would help collector nodes to recover the RSU’s message much faster.

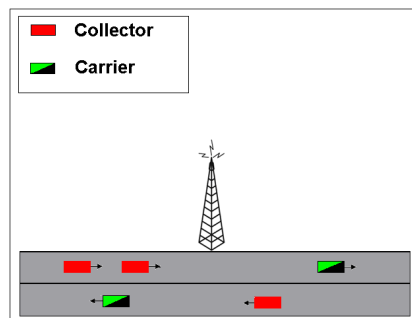


Fig. 2. For each source, we classify the vehicles on the road into two classes: *collectors* and *carriers*.

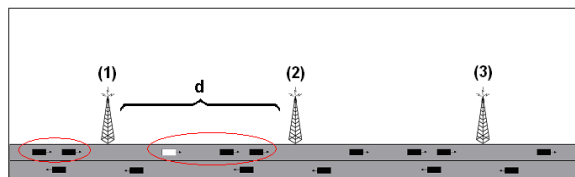


Fig. 3. Each carrier node can potentially carry packets from several sources simultaneously. It can act as a virtual source for an RSU and collector for other RSUs simultaneously. For example, the vehicle marked in white is carrier for (1) and collector for (2) and (3). The vehicles grouped in red form a cluster. Inter-cluster spacing is greater than communication range.

Collector vehicles switch to become *carrier* after they pass an RSU. Note that associated with every RSU, we attribute each vehicle as collector or carrier. Thus, it is possible that a single vehicle acts as a *carrier* for a specific RSU while the same vehicle acts as a *collector* for another RSU. A vehicle knows its location through GPS device and is pre-loaded with digital maps. Fig. 3 shows the basic network model that we focus on; with a sequence of sources  $\Phi_i$  placed uniformly with distance  $d$  from each other. We refer to the space between two adjacent sources as *segments*. We say a node is in segment  $\phi_{i,j}$  if the node is located in  $j$ th segment from source  $\phi_i$ . Also, we denote  $|\Phi_{i,j}| = j$ . Furthermore,

we assume that all sources have equal amount of information packets of size  $\mathcal{I}$ .

In our dissemination problem, every node is interested in data from all the RSUs in the network. Each carrier node is potentially a new virtual source in the network and hence can encode the information packets and broadcast them to the collector nodes. In this scheme, each carrier node can potentially carry packets from several RSUs simultaneously. Thus, it can act as a virtual source and destination for different RSUs at the same time. Every time a collector node listens to a carrier node it receives packets which are innovative (by the rateless encoding property). That is, the encoded packets collected from various carriers are innovative although they are all coded from the same information packets belonging to an RSU. This is the best possible usage of contact opportunity which is an important factor in sparse scenarios.

By using other specifications of VANETs, we can further improve the DDRC. We stress the location specific nature of the advertisement. Hence, the information provided by RSU is only useful in a nearby geographical area and can be discarded outside that area. Our scheme is apt for incorporating this feature because carrier nodes can delete packets from far sources without any need for coordination with other nodes while maintaining the performance intact for nearby sources. We define  $\Delta$  as the maximum distance from an RSU that we want packets pertaining to an RSU present (or *domain* of that RSU).

In the sequel, we investigate the effect of vehicle mobility. By using the empirical data provided by [15], we study the effect of node clustering on the network parameters.

### 4.3 Impact of Mobility

One of the distinguishing features of vehicular networks is mobility. Disconnectivity is possible in such networks. Thus clustering effect of nodes in the network must be investigated. Here, we first consider sparse highway scenarios where network connectivity is low. Let  $s$  be the spacing between vehicles in a cluster. Empirical studies show that the inter-vehicle spacing may be assumed exponential with parameter  $\lambda_s$  as [15].

$$f_s(s) = \lambda_s e^{-\lambda_s s}$$

**Theorem 1** Let  $V_0$  be the average speed of every vehicle on the road. Let  $M_n(L)$  denote the number of clusters a collector vehicle meets during a travel along a road of length  $L$  (note that collector and carrier vehicles move in opposite directions, with respect to an RSU). Further, let  $M_t$  denote the time duration that a collector vehicle spends in contact with a cluster of carrier vehicles. Given that the inter-vehicle spacing follows an exponential distribution, we have

$$E[M_n(L)] \approx \frac{2L}{(e^{\lambda_s R} - 1) \left( \frac{1}{\lambda_s} - \frac{R e^{-\lambda_s R}}{1 - e^{-\lambda_s R}} \right) + R + \frac{1}{\lambda_s}} \quad (1)$$

and

$$E[M_t] = \frac{(e^{\lambda_s R} - 1) \left( \frac{1}{\lambda_s} - \frac{R e^{-\lambda_s R}}{1 - e^{-\lambda_s R}} \right)}{2V_0}. \quad (2)$$

**Proof:** We define  $C_{length}$  and  $S_{inter}$  as the cluster length and inter-cluster spacing. Then, it is easy to see that

$$E[M_n(L)] \approx 2L/(E[C_{length}] + E[S_{inter}])$$

and

$$E[M_l] = E[C_{length}]/2V_0.$$

The proof follows by using the statistical properties of clusters in VANETs [15] and results summarized in the following lemmas. We only give the proof for Lemma 4. The proof for the rest is straightforward.  $\square$

**Lemma 1** Let  $P_{last}$  be the probability that the distance between two successive vehicles is longer than the communication range  $R$  (i.e., the probability of being the last vehicle in cluster). Then it can be shown that

$$P_{last} = Pr\{S > R\} = 1 - F_S(R) = e^{-\lambda_s R}.$$

**Lemma 2** The PDF of inter-cluster spacing  $S_{inter}$  (i.e., the space between the last vehicle of the leading cluster and the first vehicle of the following cluster) can be expressed as

$$f_{S_{inter}}(S_{inter}) = \lambda_s e^{-\lambda_s(S_{inter}-R)}.$$

Therefore, we have

$$E[S_{inter}] = R + \frac{1}{\lambda_s}.$$

**Lemma 3** The expected number of vehicles in a cluster is given by

$$E[C_{size}] = \frac{1}{P_{last}}.$$

**Lemma 4** The average cluster length can be obtained as

$$E[C_{length}] = (e^{\lambda_s R} - 1) \left( \frac{1}{\lambda_s} - \frac{R e^{-\lambda_s R}}{1 - e^{-\lambda_s R}} \right). \quad (3)$$

**Proof:** One can show

$$C_{length} = \sum_{i=1}^{C_{size}-1} S_i.$$

Since  $s$  follows an exponential distribution, we have

$$\begin{aligned}
E[C_{length}] &= E[E[C_{length} | C_{size}]] \\
&= E[E[\sum_{i=1}^{C_{size}-1} S_i]] \\
&= E[C_{size} - 1]E[S | S < R].
\end{aligned} \tag{4}$$

Substituting the required expressions in Eq. (4) completes the proof.  $\square$

Results in Theorem 1 is consistent with the notion that as mobility increases, the number of encounters remains the same but the duration of encounters decreases. From Eqs. (1) and (2), it can be shown that the expected number of packets that a collector vehicle obtains from  $\Phi_i$  during travel in a segment  $\phi_{i,j}$  is directly proportional to the number of packets from the corresponding source that carriers possess in that segment.

#### 4.4 Decoding Distance (DD)

The parameter  $DD$  is basic performance metric we consider which can provide insight to the throughput. As mentioned earlier, each node has a limited buffer space for cooperation, *i.e.*, each carrier node keeps only a limited number of encoded packets from sources it has met. Here, we address a simplified version of the problem of allocating the limited buffer to a number of sources. The objective is to maximize  $DD$  for all sources.

Upon crossing an RSU, every node that has been successful in decoding the RSU's message act as a carrier for that source. Then, every carrier node generates some encoded packets from the RSU's *information* packets and stores them. The number of stored packets to be determined for maximum performance, given that the storage buffer is limited to  $B$ .

By meeting each cluster  $C_i$ , the collector has the opportunity to gather packets. Consider a collector vehicle meets a cluster of vehicles in  $\phi_{i,j}$ , the  $j$ th segment from  $\Phi_i$ . Assume that, on the average, there are  $m_{i,j}$  packets from  $\Phi_i$  in the buffers of the carriers in the cluster. Then, the number of broadcast packets during the meet time  $M_t$ , follows the Poisson distribution with mean  $\rho$ , because carriers send encoded packets randomly and without coordination. Thus, using the throughput relations of the ALOHA network [16], we verify that the maximum throughput occurs when  $\rho$  is equal to  $1/2$  of packet transfer time and is equal to  $1/2e$ . Further, the probability that a received packet is of collector's interest (*i.e.*, the packet is from  $\Phi_i$ ) is equal to  $m_{i,j}/B$ . Hence, by using Eqs. (1) and (2), the maximum expected number of packets collected from a cluster  $N_j^c$ , when carriers possess  $m_{i,j}$  packets from  $\Phi_i$  and total number of collected packets  $N_j^T$ , that a typical collector can obtain from segment  $\phi_{i,j}$  of length  $d$ , are given by

$$E[N_j^c] = \frac{(e^{\lambda_s R} - 1) \left( \frac{1}{\lambda_s} - \frac{R e^{-\lambda_s R}}{1 - e^{-\lambda_s R}} \right) m_{i,j}}{4eV_0 B} \tag{5}$$

and

$$E[N_j^T] = M_n(d) \times E[N_j^c]. \tag{6}$$



Therefore,  $DD$  is directly proportional to the number of packets from the corresponding source that carriers possess per each segment. In order to maximize  $DD$  for an arbitrary source  $\Phi_i$ , we need to find a solution for  $m_{i,j}$ 's subject to the buffer limit constraint. Since sources are all the same, we can omit the first index in  $m_{i,j}$  and find a generic solution  $m_j$  for all sources. We assume that the number of packets a carrier possesses from  $\Phi_i$  cannot be increased. Also, buffer updating for a carrier node occurs when it crosses a new source and enters a new segment (e.g., from  $\phi_{i,j}$  to  $\phi_{i,j+1}$ ). Then, it makes a decision on how many packets it should drop such that the buffer constraint is met. Thus, just after crossing  $\Phi_i$  the carrier node has  $m_0$  encoded packets from the source and reduces them gradually as  $m_0 \geq m_1 \geq m_2 \geq \dots$  and  $m_j = 0$  for  $j \geq \Delta$ . To find the maximizer distribution for  $m_j$ 's, we use the symmetricity in the problem and consider a collector which enters the source  $\Phi_i$ 's domain. Hence,  $DD$  can be formally stated as

$$DD = \min_{m_j} d \tag{7}$$

$$s.t. \sum_{i=0}^{\Delta} E[N_{\Delta-i}] \geq \mathcal{I},$$

where  $e[N_{\Delta-i}]$  is the expected number of packets obtained by a collector in segment  $(\Delta - i)$ . In other words, we are looking for the first segment that the total number of collected packets is greater than  $\mathcal{I}$ . Since Eq. (7) suggests that only the tail behaviour of the distribution of  $m_j$ 's is important and because  $m_j$  is non-increasing with  $j$ , one can see that the maximum value of  $DD$  is achieved when  $m_{\Delta}$  (and hence all the previous segments) has its maximum value. Further, the buffer limit constraint implies that  $\sum_{i=0}^{\Delta} m_i \leq B$ . Therefore, a solution can be formulated as

$$m_0 = m_1 = m_2 = \dots = m_{\Delta} = \frac{B}{\Delta + 1}. \tag{8}$$

The result on a single linear highway can be extended to the two-dimensional case, using the trajectory of vehicles, and we can still classify nodes into two different groups as collector nodes and carrier nodes. Then, as depicted in Fig. 4, we can transform a two-dimensional scenario into a one-dimensional case by mapping sources on nodes' path. Therefore, the analysis is quite similar to the one-dimensional case but with non-uniformly distributed sources.

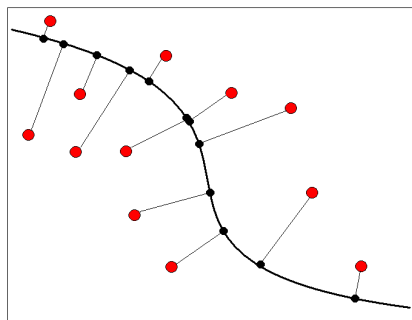


Fig. 4. Two-dimensional scenario.

## 5. RESULTS OF SIMULATIONS AND DISCUSSIONS

### 5.1 Simulation Setup

In this section, we evaluate the performance of a simple DDRC setup. Vehicles enter the road from one end with inter arrival times drawn using instances of exponential distribution with parameter  $\lambda$ . We considered a random placement of 50 points on a 2-D plane which are mapped on the diameter as the positions of the RSUs along the road. Using the results from IV-D we suggest two different buffer management schemes and evaluate their performance.

We developed an NS-2 [17] based simulator to evaluate the proposed schemes. We use the transmission range  $R$  as the unit of distance in our evaluations. We focus on a collector vehicle departs from one end of the road and travels along it until it reaches the other end. The record of all the collected packets versus distance is maintained for analysis. Most experiment parameters are listed in Table 1.

**Table 1. Simulation setup.**

Parameter	Value
Simulation time	1000 seconds
Communication range (R)	200 m
Vehicle velocity	20-40 m/s
Inter-arrival of vehicles ( $\lambda$ )	0.1 veh/second
Simulation road length	20000 m
Broadcast interval	100/second
Drop factor ( <b>D</b> ) for scheme A	5-30
Window size ( <b>N</b> ) for scheme B	1-50
Buffer size	100-1500
Number of RSU's along the road	50

The performance of the protocols is measured by the following metrics:

- *Mean Decoding Distance (MDD)*, which is the expected value  $DD$  over all active sources.
- $P_{Success}$  is the probability that a random message generated at a (random) source  $\Phi$  is available at node  $v$  before it enters the communication range of  $\Phi$ . We measure  $P_{Success}$  as a function of distance to the source. The presented graphs are the average of  $P_{Success}$  over all active sources.
- *Deployment Capacity (DC)*, which is the maximum number of active sources on the road such that the  $P_{Success}$  for a collector vehicle at an average distance  $\eta$  from a source is in  $\epsilon$  neighborhood of 1 ( $\epsilon \ll 1$ ).

In order to meet the optimal distribution of  $m_j$ 's as presented in Eq. (8) and compare the effect of different parameters, two different buffer management schemes are proposed for simulations.

- (1) Scheme A: In this scheme, the buffer space is shared by *all the sources* (RSUs) a vehicle has met thus far. When reaching a new source, a fraction  $D$  of the packets in the buffer are dropped to store the new source packets. Upon crossing the new source and entering a new segment, fraction  $D$  of the packets from every source will be dropped randomly.
- (2) Scheme B: In this scheme, the buffer space is shared by a fixed number of *the most recent sources* a carrier has met. Upon reaching a new source, all the packets from the oldest source in the buffer will be removed and replaced by the incoming packets from the new source.

We perform the simulation for different values of  $N$ , the number of sources from which a carrier keeps packets in its buffer ( $N = 1$  is equivalent to flushing the buffer upon reaching every new source).

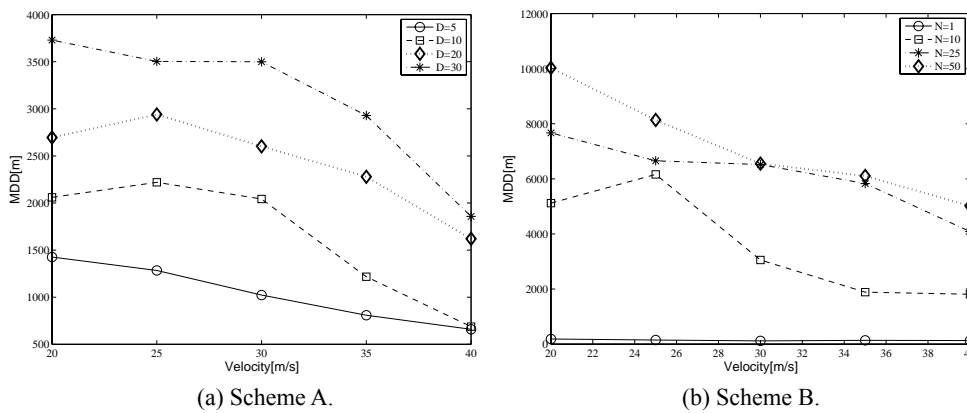


Fig. 5. Mean decoding distance (MDD) vs. velocity for different values of  $N$  and  $D$ .

Fig. 5 compares the MDD for both schemes. As Fig. 5 (b) suggests, for high values of  $N$ , scheme B outperforms scheme A. This result conforms the intuitive solution presented in Eq. (8). Because, for large values of  $N$ , the carrier node keeps packets from previous sources for a longer amount of time and in fact shares the available buffer space on the road between all sources in a window of size  $N$ . When  $N$  is large, packets from far source stay longer at the carrier buffer and hence giving more chance to the collectors to gather packets from the far source. It is worth noting that large  $N$  results in relative fairness between sources.

To plot  $P_{Success}$ , we set the value  $\mathcal{I}$  first and count the number of collected packets  $P_\eta$  (versus distance  $\eta$  from the source) that a collector node can accumulate before entering the communication range of the source. Define the indicator random variable as

$$S_\eta = \begin{cases} 1 & \text{if } P_\eta \geq \mathcal{I} \\ 0 & \text{if } P_\eta < \mathcal{I} \end{cases}$$

The value of  $Pr\{\mathcal{S}_R\}$  is considered as  $P_{Success}$ . Fig. 6 shows  $P_{Success}$  for a number of distances (in multiples of communication range  $R$ ) from an arbitrary source. The collector nodes would like to recover the source information from a far distance. One observation to make from Fig. 6 (b) is that in scheme B, by using a larger value of  $N$  we can ensure the success for collector nodes. Hence, scheme B provides a very good coverage area around each source by enabling collector nodes to decode the source message with high probability even before entering its communication range.

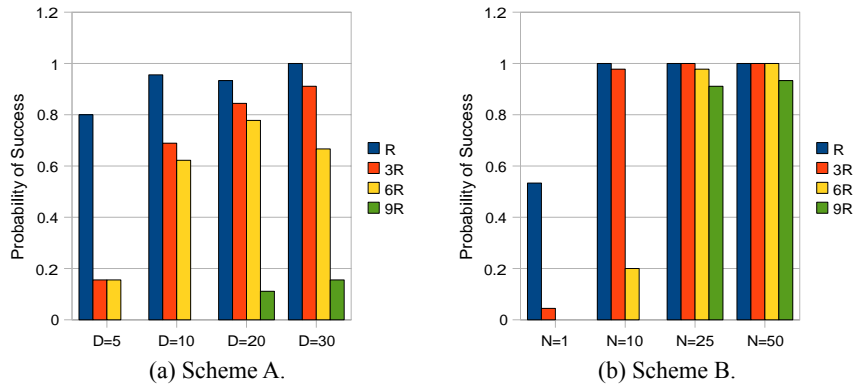


Fig. 6. Probability of success for various distances (in multiples of communication range,  $R$ ) from source.

**Table 2. Deployment capacity evaluation – scheme A.**

	$\eta = R$	$3R$	$6R$	$9R$	$12R$
$M = 10$	1817.2	1699.9	1570.8	1462.7	1395.3
$M = 20$	915.9	832.65	754.6	677.55	628.3
$M = 30$	708.5	625.73	568.63	452.37	400.63
$M = 40$	754.7	677.48	590.1	446.33	388.63
$M = 50$	933.2	662.31	608.39	472.95	423.68

**Table 3. Deployment capacity evaluation – scheme B.**

	$\eta = R$	$3R$	$6R$	$9R$	$12R$
$M = 10$	2068	2067.4	1997.4	1825.1	1673.7
$M = 20$	1459.2	1450.5	1267	1080.2	937.05
$M = 30$	1196.4	1152.3	1128.3	915.9	779.8
$M = 40$	1117.6	1059.3	1031.7	831.8	709.1
$M = 50$	1091.1	1134.6	1007.3	810.3	700.85

Tables 2 and 3 show the average of total number of collected packets per source for a collector node as a function of distance  $\eta$  and different number of active sources on the road. To determine the “Deployment Capacity” (DC), we need to fix a distance and a positive value  $\epsilon$  which indicates the bound on acceptable  $P_{Success}$ . Then, by using the tables we are able to find the range of the maximum number of sources that satisfies the condition on decoding probability at desired distance. Then,  $DC$  can be obtained by per-

forming more simulations on that range and altering the number of active sources incrementally. Suppose that we are interested in  $DC$  for scheme B at  $\eta = 6R$  with  $\epsilon \ll 1$  when  $\mathcal{I} = 1000$ . Since total number of collected packets per source when  $M = 50$  is slightly higher than  $\mathcal{I}$  (from Table 3), we conclude that  $40 < DC(\epsilon) < 50$ . Using the graph in Fig. 7, we conclude that  $DC(\epsilon < 0.05) = 44$ ,  $DC(\epsilon = 0.05) = 46$  and  $DC(\epsilon = 0.25) = 48$ .

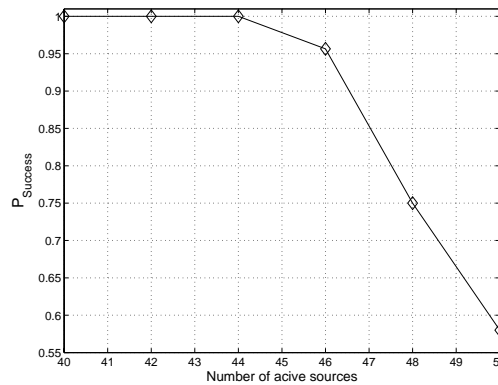


Fig. 7.  $P_{Success}$  vs. number of active sources in scheme B for  $\epsilon = 6R$ .

## 6. CONCLUSIONS AND FUTURE WORK

We presented a novel scheme based on rateless codes for collaborative content distribution from road side units to vehicular networks. In DDRC, we divide the vehicles on the road into two groups of collector and carrier nodes per source. The collector nodes have the opportunity to recover the source message from a far distance using the carrier nodes. We introduced Decoding Distance and Deployment Capacity for performance evaluation and provided analytical models to explore them. The presented models capture the effect of various parameters in the network and provide guidelines on choosing the parameters to maximize the performance metrics.

To the best of our knowledge, this is the first paper to study the effect of finite buffer constraint in the dissemination problem in a VANET. The proposed scheme can seamlessly handle both sparse and dense scenarios. Future directions are to introduce more realistic traffic models for urban areas where the exponential assumptions for inter-arrival time are not valid. Further, adapting the analysis for such scenarios and finding the optimal distribution for buffer allocation are our immediate goals. Finally, deriving better approximations for performance metrics in more general setting and relaxing some of the assumptions remain as open problems.

## REFERENCES

1. M. Sardari, F. Hendessi, and F. Fekri, "DMRC: Dissemination of multimedia in vehicular networks using rateless codes," in *Proceedings of INFOCOM Workshops*, 2009, pp. 1-6.

2. J. Yin, T. Eibatt, G. Yeung, B. Ryu, S. Habermas, H. Krishnan, and T. Talty, "Performance evaluation of safety applications over dsrvehicular ad hoc networks," in *Proceedings of ACM VANET*, 2004, pp. 1-9.
3. S. B. Lee, G. Pan, J. S. Park, M. Gerla, and S. Lu, "Secure incentives for commercial ad dissemination in vehicular networks," in *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2007, pp. 150-159.
4. V. Namboodiri, M. Agarwal, and L. Gao, "A study on the feasibility of mobile gateways for vehicular ad-hoc networks," in *Proceedings of ACM VANET*, 2004, pp. 66-75.
5. V. Naumov, R. Baumann, and T. Gross, "An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces," in *Proceedings of ACM MobiHoc*, 2006, pp. 108-119.
6. H. Wu, R. Fujimoto, R. Guensler, and M. Hunter, "MDDV: A mobility centric data dissemination algorithm for vehicular networks," in *Proceedings of ACM VANET*, 2004, pp. 47-56.
7. M. Fiore and J. Harri, "The networking shape of vehicular mobility," in *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2008, pp. 261-272.
8. M. Luby, "Lt codes," in *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, 2002, pp. 271-282.
9. M. Johnson, L. D. Nardis, and K. Ramchandran, "Collaborative content distribution for vehicular ad hoc networks," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, 2006.
10. J. Park, U. Lee, S. Oh, M. Gerla, and D. Lun, "Emergency related video streaming in vanet using network coding," in *Proceedings of ACM VANET*, 2006, pp. 102-103.
11. S. J. Lee, W. Su, and M. Gerla, "On-demand multicast routing protocol in multihop wireless mobile networks," *Mobile Networks and Applications*, Vol. 7, 2002, pp. 441-453.
12. J. Zhao, Y. Zhang, and G. Cao, "Data pouring and buffering on the road: A new data dissemination paradigm for vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, Vol. 56, 2007, pp. 3266-3277.
13. B. N. Vellambi, R. Subramanian, F. Fekri, and M. Ammar, "Reliable and efficient message delivery in delay tolerant networks using rateless codes," in *Proceedings of 1st International MobiSys Workshop*, 2007.
14. Y. Wang, S. Jain, M. Martonosi, and K. Fall, "Erasure-coding based routing for opportunistic networks," in *Proceedings of ACM SIGCOMM Workshop on Delay-Tolerant Networking*, 2005, pp. 229-236.
15. A. N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar, and O. Tonguz, "Routing in sparse vehicular ad hoc wireless networks," *IEEE Journal on Selected Areas in Communications*, Vol. 25, 2007, pp. 1538-1556.
16. N. Abramson, "The aloha system – another alternative for computer communications," in *Proceedings of AFIPS*, 1970, pp. 281-285.
17. "The network simulator NS-2," <http://www.isi.edu/nsnam/ns/>.



**Mohsen Sardari** received the B.Sc. degree from Sharif University of Technology, Tehran, Iran, in 2007. He is currently pursuing Ph.D. in Electrical and Computer Engineering at the Georgia Institute of Technology, Atlanta, GA. His research interests include network optimization, network coding, coding and information theory.



**Faramarz Hendessi** received the M.Sc. degrees from Isfahan University of Technology, Isfahan, Iran and the Ph.D. degree from the Carleton University, Ottawa, in 1994. Since 1994, Dr. Hendessi has been with the faculty of the Electrical and Computer Engineering at the Isfahan University of Technology (<http://www.iut.ac.ir>) where he currently holds an Associate Professor position. He serves on the editorial board of the International Journal of Information and Communication Technology, Iran, and on the Technical Program Committees of several conferences. Prof. Hendessi's current research interests are in the areas of wireless communications and networking. Dr. Hendessi is Head of one of the famous IT Company in Isfahan, Iran (<http://www.Payamnet.Com>) which is established in 1995 by him.



**Faramarz Fekri** received Ph.D. degree from the Georgia Institute of Technology in 2000. Since 2000, Dr. Fekri has been with the faculty of the School of Electrical and Computer Engineering at the Georgia Institute of Technology where he currently holds an Associate Professor position. He serves on the editorial board of the IEEE Transactions on Communications, and on the Technical Program Committees of several IEEE conferences. Prof. Fekri's current research interests are in the area of communications and signal processing, in particular coding and information theory, information processing for wireless and sensor networks, and communication security. Dr. Fekri received the National Science Foundation CAREER Award (2001), and Southern Center for Electrical Engineering Education (SCEEE) Research Initiation Award (2003), Outstanding Young Faculty Award of the School of ECE (2006). He is a Senior Member of the IEEE.