Efficient Data Dissemination in Vehicular Ad Hoc Networks

Fei Ye, *Student Member, IEEE*, Sumit Roy, *Fellow, IEEE*, Haobing Wang, *Student Member, IEEE*

***Abstract*—Data services for in-vehicle consumption are expected to become a primary driver in the development of future vehicular networks. Due to download rate limitations of present wide-area cellular connectivity such as 3G (the likely 'pipe' to/from vehicles for long range connectivity), direct peer-to-peer data sharing among vehicles can supplement *vertical downloading* with *horizontal dissemination*. This paper studies the inter-vehicle data dissemination problem based on a WAVE[1]/802.11p vehicular ad hoc network, using network coding. We first derive the probability mass functions (PMFs) of dissemination completion time in a prototypical three-node case for both random broadcast and with network coding, to quantify the benefits of the latter. For a one dimensional (1-D) infinite lattice network, we next provide analytical results for the steady state dissemination velocity of a data set, using network coding. The gains from such network coding, relative to the baseline scheme of random broadcast, and with perfect feedback, in presence of Rayleigh fading wireless links for this network are estimated using simulations.**

***Index Terms*—Data dissemination, Vehicle-to-Vehicle, Dissemination Velocity, Coupon Collection, Ad Hoc Network**

I. Introduction

RECENT research, development and standardization advances in vehicular ad hoc networks (VANETs) [1]–[6] have motivated increasing interest in various data services for in-vehicle consumption, such as 'commerce- and entertainment-on-the-wheel'. These include a wide variety of applications: local information (e.g., traffic notification, map updates, location-based advertisements) pushed to vehicles; or specific data pulled from Internet servers (e.g., neighborhood parking, reviews of local restaurants, and video clips of local attractions). Direct download of services to vehicles may be provided over existing wide-area cellular infrastructure (3G/4G) and/or proposed new roadside infrastructure based on short-range Dedicated Short Range Communications (DSRC) links [2]. However, both these approaches have their own challenges: the modest data rate of present 3G links and the cost of large data downloads to the end-user under current cellular data pricing models on one hand, and the intermittent hot-spot type roadside coverage envisaged with DSRC on the other[3]. This leads to the following simple premise for content dissemination: if some content is available at a subset of vehicles (typically via prior cellular vertical download) that is desired by (many) others in the network, peer-to-peer content distribution using vehicle-to-vehicle (V2V) ad hoc communications, is time and cost efficient.

Consider the scenario in Fig. 1 - a group of vehicles in geographic vicinity that have a common interest in some data on the Internet, such as traffic surveillance video, map updates, or other geo-related information. When a reference vehicle successfully downloads the data file via vertical network access, it may then disseminate the newly obtained content to all other peers (horizontally) through V2V mode. For purposes of lightweight data exchange, packets in WAVE can be transmitted as 1-hop broadcast with no feedback, as defined in IEEE 802.11p [6]. In order to improve latency, WAVE does provide priority mechanisms for propagation of emergency warning messages via multi-hop relaying [7], [8]; however, these forwarding/rebroadcasting mechanisms are not suitable for sharing *large amounts* of non-prioritized content.

Due to the absence of feedback in WAVE broadcasts, a common way of disseminating data is via gossiping [9], [10] that mimicks the spread of a rumor in society. In a gossip protocol, a node blindly distributes a randomly chosen packet among those in its buffer, at a transmission opportunity. Its inefficiency is illustrated by a simple example in Fig. 2. Two vehicles N_2 and N_3 within geographical proximity are each waiting to receive two packets X_1 and X_2 from N_1. At this time, N_2 (N_3) has only X_2 (X_1) from prior transmissions. Even with lossless wireless links, N_1 needs two more broadcast transmissions to complete the transfer *if packets are sent individually*. However, it is evident that when N_1 is broadcasting X_1, N_3 does not benefit from overhearing this transmission; similarly for N_2 when N_1 is broadcasting X_2.

Network coding (NC) leverages the above unexploited redundancies due to the broadcast nature of wireless for more efficient data dissemination. We use Fig. 2 to demonstrate how linear NC [11], [12] may apply in this simple case. Node N_1 only needs to broadcast $Y = X_1 \oplus X_2$, then both N_2 and N_3 can both extract the desired packet after one broadcast via $Y \oplus X_1$ ($Y \oplus X_2$), respectively.

Successful application of NC to a realistic VANET scenario requires going beyond the simple scenario of Fig. 2 and confronting several non-trivial challenges [13]–[15] such as: 1) If a source-destination pair are not within direct range of

Manuscript received 15 May 2011; revised 27 October 2011. This work was supported in part by AFOSR award FA-9550-09-1-0298.

F. Ye, S. Roy, and H. Wang are with the Department of Electrical Engineering, University of Washington, Seattle, WA (e-mails: {fye, sroy, hbwang89}@u.washington.edu).

Digital Object Identifier 10.1109/JSAC.2012.120511.

[1]Wireless Access for Vehicular Environments

[2]DSRC in North America is based on 802.11p wireless link access in 5.9 GHz band, subsequently folded into the IEEE 1609 Wireless Access for Vehicular Environments (WAVE) standards.

[3]DSRC's spectrum allocation has four 10MHz service channels, and each supports a per channel data rate from 3 to 27 Mbps, with the optional capability to combine two service channels into one 20Mhz channel.

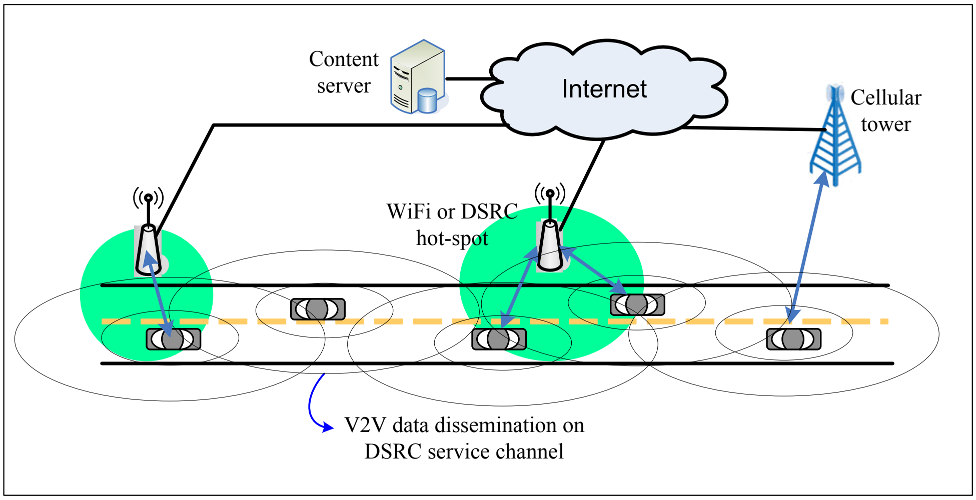

Fig. 1. Inter-vehicle horizontal data dissemination supplements vertical direct downloading.

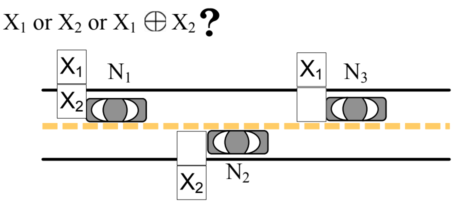

Fig. 2. A 3-node example. Using network coding at N_1 achieves the maximum data dissemination efficiency.

each other, data must be routed in a multi-hop manner; this is not supported in WAVE broadcast;
2) Relaying by intermediate nodes leads to increased multi-access interference, which in turn impacts the reception probability on any link. In other words, the unreliability of packet reception at the link level must be incorporated in any effective wireless networking coding design.

This paper analyzes the V2V horizontal data dissemination part of data services in 1-D VANETs: 'pushing' Internet content to vehicles (see Fig. 3). Our contributions are three-fold:

- We mathematically derive the probability mass functions (pmf's) of the dissemination completion time for both random and NC based broadcast in a prototypical three-node case.
- The steady state data dissemination velocity in a 1-D infinite lattice network using NC is explicitly derived.
- We investigate via simulations the gains from NC relative to random broadcast and a perfect feedback scheme for 1-D lattice and irregular topologies, in presence of Rayleigh fading wireless links.

The remainder of this paper is arranged as follows. Related works are discussed in Sec. II. We develop the system model in Sec. III. The pmf's of dissemination completion time in a three-node case are derived in Sec. IV-A, and the dissemination velocity for 1-D lattice network in Sec. IV-B. Simulation results are provided in Sec. V and we conclude with final remarks in Sec. VI.

II. Related Work

There is a growing literature on content dissemination within vehicular networks, some of which have also considered the application of NC this problem. We briefly discuss the key relevant references next, and highlight their difference from our approach.

Peer-to-peer collaboration via vehicular communication is explored in [16], [17], [18] for content sharing. SPAWN, a swarming protocol for vehicular ad hoc networks was proposed in [17] to extend wired Internet peer-to-peer content distribution (a.k.a. file swarming) protocols (e.g., BitTorrent) to VANET. However, peer discovery and peer & content selection processes have high overhead due to VANET's dynamic nature. The assumption of reasonable transport layer bandwidth (either TCP or UDP over multi-hop wireless links) in [16], [17] does not fit into the WAVE broadcast paradigm. Similar observation was made in [18] for CodeTorrent, a NC aided file swarming protocol with no underlying routing support.

Network coding has been applied to the full data dissemination problem [14] and the reliable multicasting problem [13]. In [14], each node has some packets to share with all other nodes, and a gossip based algorithm is used to diffuse information in the ad hoc network that is optimized for energy efficiency. In [13], reliable multicasting in a single cell wireless network and over a tree topology constructed by several cells are studied. In each cell, coded packets are repeatedly broadcast by an access point till all nodes have correctly received all packets.

The most relevant prior art to this work are [18]–[22], which study linear NC in VANET settings. Our work and [18] are similar in that both apply NC to MAC broadcasting. However, in [18], source nodes send out a file description and other nodes have to keep requesting neighbors for packets (data 'pull'), while in our work, the file is 'pushed' to all nodes. VANET relaying protocols are devised in [21], [22]. In [22], the region is divided into segments, and each segment selects a coordinator which designates a relay node for that segment in a centralized way. In CodeOn [21], nodes share 'content reception status', and then choose a transmission backoff delay

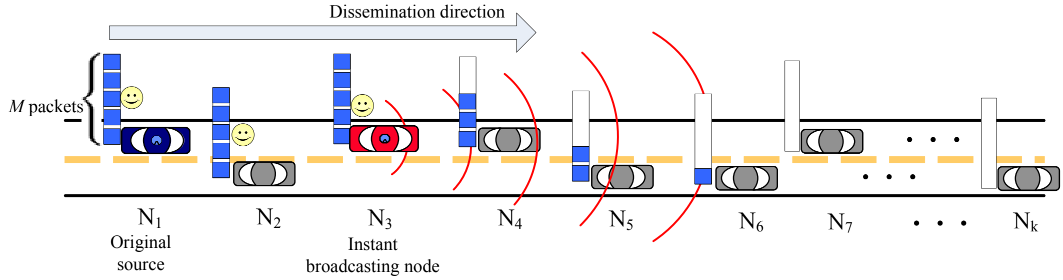

Fig. 3. Data dissemination from one source vehicle to all others in the $+x$ direction in a 1-D VANET.

that is inversely proportional to the amount of useful content they possess. As a result, the node with the most useful content gets to relay. In [21], [22], all broadcasts by the access point and at all relay nodes use symbol level NC. [19] studies an iterative vehicle-to-infrastructure (V2I) download and V2V content distribution process. Vehicles independently download NC coded packets from the server in V2I phase, followed by NC based many-to-many content distribution in next V2V phase. In contrast, the V2V data dissemination is a one-to-many dissemination process in our work. While [19], [21], [22] explores the design via simulation, our work provides *analytical results* for V2V data dissemination under a simple scheduling model. [20] analyzes the multicast throughput of NC aided content distribution in a linear VANET. Their result is based on ideal reception and collision models, in which a packet within a fixed distance from the source is received with probability 1, otherwise it is deemed lost. Such binary characterization of the wireless link is grossly inadequate for purpose; we integrate link reliability to deduce the pmf for dissemination completion duration and velocity in steady state.

III. System Model

A. Network Coding Operations

Linear network coding is a block code conducted over finite field $\mathbb{F}_q$, where q is the finite field size. Every packet x_i consisting of $\log(q)$-bit symbols, can be treated as a vector in $\mathbb{F}_q$. During any transmission opportunity, the source broadcasts a linear combination of all M packets it has stored, $x_1, x_2, \ldots, x_M$, i.e.,

$$y_i = \sum_{k=1}^{M} \gamma_{i,k}\, x_k, \tag{1}$$

where $\gamma_{i,k}$'s are the random NC coefficients selected uniformly in $\mathbb{F}_q$ at the i-th transmission. Any received coded packet at a receiver can be expressed as a linear combination of the M original packets in the network, regardless of the actual sender of that packet. Specifically, if a node receives K coded packets $y_1, \ldots, y_K$, these can be written in the following vector product form:

$$Y_{K\times 1} = \mathbf{A}_{K\times M} X_{M\times 1}, \tag{2}$$

where $\mathbf{A}_{K\times M}$ is the coefficient matrix containing all NC coefficients. In the case $K \geq M$, the matrix $\mathbf{A}$ is full rank almost surely due to random choice of NC coefficients from $\mathbb{F}_q$, and the receiver can recover the original data set $X = \mathbf{A}^{-1} Y$.

For decoding, clearly the receiver needs to know the NC coefficients $\mathbf{A}$ used during all transmissions. We allow the transmitter to embed these coefficients in each packet. For instance, linear NC of M packets over finite field $\mathbb{F}_q$ needs $M \log_2(q)$-bit long coefficients in each packet, which represents 3.3% overhead when $M = 100$, $q = 16$, and packet length $L = 1500$ Bytes.

Defn.: A newly received coded packet is called *innovative* if it increases the rank of $\mathbf{A}$. Only innovative packets are preserved, otherwise a packet is discarded. The probability P that a new packet being innovative satisfies $P \geq 1 - 1/q$, which converges in probability to 1, $\operatorname{plim}_{q\to+\infty} P = 1$. Thus, for practical operations, with reasonably large finite field size, each received coded packet can be considered innovative [23].

B. Scheduling Model

We model a vehicle platoon on a stretch of highway as a static 1-D lattice network of equal inter-vehicle spacing, corresponding to motion with identical velocity. Initially (at time $t = 0$), there is a single file (M packets denoted by $x_1 \ldots x_M$) at a single source node (vehicle) to be disseminated to all other vehicles in a single direction (see Fig. 3). We assume that vehicles have the knowledge of neighbors' relative position[4] to filter out packets coming from the undesired direction. Nodes in the network are assumed synchronized, and time is slotted so that each packet transmission consumes exactly one time slot.

An ideal scheduler is assumed throughout this work, i.e. there is no distributed medium contention prior to transmission (and hence any analysis of MAC efficiency is not relevant to this work). At a given time, there is a single source (the disseminator) that broadcasts the reference data in the $+x$ (dissemination) direction. When the nearest recipient (relative to current disseminator) successfully receives the full set, a 'hand-off' occurs as shown in Fig. 4, i.e. the new node takes on the role of the disseminator (and the prior disseminator stops broadcasting) [5].

[4]Such relative position information can be obtained from: (1) directly comparing the global positioning system (GPS) location information embedded in the packet with receiver's local GPS location, or (2) sender's ID embedded in the packet and a neighborhood map constructed from received beacons (contain location information) from neighbors.

[5]In 802.11 networks, this may be achieved by allowing the new node to seize the channel (from the prior disseminator) by using a smaller contention window value.

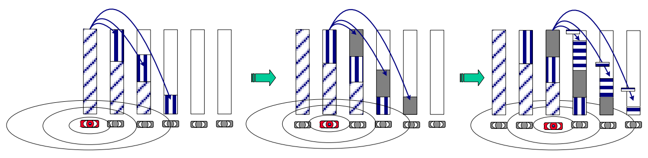

Fig. 4. Disseminator role is handed-off node by node in the ideal scheduling model.

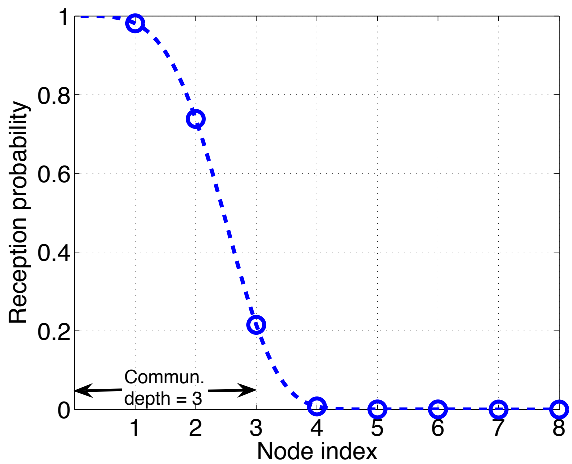

Fig. 5. One example of the reception probabilities $Q_1, Q_2, \cdots$ (at different recipients) and the communication depth N (θ=5%).

C. Communication Depth

When current disseminator broadcasts a packet on the shared wireless medium, it is received by the 1st, 2nd, $\cdots$, N-th, $\cdots$, recipients (relative to disseminator) with probabilities $Q_1 > Q_2 > \cdots, Q_N, \cdots$, respectively. Such reception probability sequence depends on topology and communication parameters, i.e., inter-vehicle spacing, transmit power, wireless channel, etc. Nodes which are sufficiently far away from the disseminator have negligible reception probabilities, i.e., $Q_k = 0$, $\forall k > N$, where we define $N = \arg\max_k \{Q_k \geq \theta\}$, θ is a sufficiently small threshold, say 5%. N will be referred to as the communication depth. An example of the reception probability and communication depth under fading channel is shown in Fig. 5.

IV. Analysis

In the ideal scheduler, the disseminator role is handed-off node by node in the desired propagation direction, each time the complete data set is received at a new node. To characterize the gain of NC, in Sec. IV-A, we first derive the pmf's of the dissemination completion time in a three-node case, with and without NC. In Sec. IV-B, we first prove that the NC dissemination process has a steady state, and then explicitly derive the reference data set's dissemination velocity in steady state as a function of reception probabilities and communication depth.

A. A Three-Node Case

We first demonstrate the (potential) gains from NC by comparing pmf's of the completion time with and without NC via this prototypical three-node example in Fig. 6 with communication depth $N = 2$. At $t = 0$, Node 0 (the disseminator)

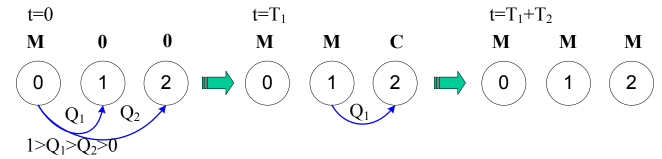

Fig. 6. Data dissemination in a prototypical three-node network in presence of wireless link unreliability.

has the M packets to be disseminated whereas Nodes 1 and 2 start with none. The reception probabilities from Node 0 at the two nodes satisfy $1 > Q_1 > Q_2 >> 0$. Let T_1 denote the completion time for Node 1, i.e., the duration required for Node 1 to receive all M packets successfully. Now, at T_1, Node 2 has accumulated $C \in \{0, 1, \ldots, M\}$ packets due to overhearing the broadcasts. At T_1, Node 1 assumes the disseminator role, and the one-hop link probability between Nodes 1 and 2 is now Q_1. Let T_2 denote the *additional* time for Node 1 to complete the transfer to Node 2. In the following, we explicitly derive pmf's for T_1 and T_2 with and without NC.

Random broadcast (RND): In random broadcast scheme, the disseminator transmits one packet per slot, randomly chosen from those in buffer, till its immediate neighbor has all packets. The lack of any feedback implies that multiple copies of a packet may be transmitted, and then then discarded by receiver. The completion time calculation is a generalization of the classic *coupon collection problem (CCP)*[6] studied in [24], [25].

In the following lemma, we first state an intermediate result for the *conditional* completion time in a single source-receiver pair case with random broadcast.

Lemma 1. *Consider a source-receiver pair where the source has the M packets for dissemination and the receiver has a subset of C packets (as a result of prior overhearing). The link success probability equals Q_1 and there is no feedback. T^{RND}, the* conditional *completion time for random broadcast has the following pmf:*

$$\begin{aligned}
&\Pr(T^{RND} = n \mid C = c) \\
&= \sum_{k=M-c-1}^{n-1} \sum_{j=0}^{M-c-1} (-1)^j \binom{M-c-1}{j} \left(1 - \frac{j+1}{M-c}\right)^k \\
&\quad \binom{n-1}{k} \bar{Q}_1^{k+1} \left(1 - \bar{Q}_1\right)^{n-k-1}, \qquad \forall n \geq M - c,
\end{aligned} \tag{3}$$

where $\bar{Q}_1 = Q_1 \left(1 - \frac{c}{M}\right)$.

Proof: See Appendix A. ∎

In the special case when $C = 0$ and the link is perfect $Q_1 = 1$, (3) reduces to the stopping time T_{CCP} in classic coupon collection problem (see Eq. (3) in [25]).

The following theorem characterizes the pmf of dissemination completion time for random broadcast in a three-node case.

[6]Coupon collection problem studies the number of sample trials $T_{CCP}(M)$ needed to collect all M coupons, in the case that one coupon is randomly drawn with replacement at each trial. $E[T_{CCP}(M)]$ is known to be the M-th Harmonic number.

$$\Pr(T_1^{RND} = t_1) = \sum_{k=M-1}^{t_1-1} \sum_{j=0}^{M-1} (-1)^j \binom{M-1}{j} \left(1 - \frac{j+1}{M}\right)^k \binom{t_1-1}{k} Q_1^{k+1} (1-Q_1)^{t_1-k-1}, \quad t_1 \geq M, \tag{4}$$

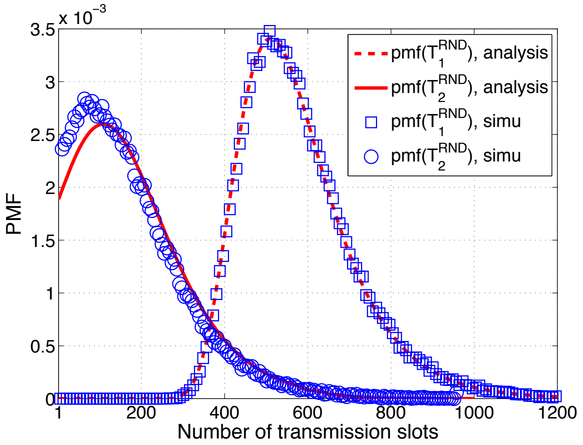

Fig. 7. Pmf's of T_1^{RND} and T_2^{RND} when M=100, Q_1=0.9, Q_2=0.7. Analytical curves of pmf(T_1^{RND}) and pmf(T_2^{RND}) are plotted from (4) and (5). $\Pr(T_2^{RND} = 0)$ is an impulse which is not plotted in this figure. Impulse's height is 0.216 from analysis, and is 0.222 in simulation.

Theorem 1. *In a three-node case (Fig. 6), the two dissemination completion components $T_1^{RND} + T_2^{RND}$ for random broadcast have the following (unconditional) pmf's in (4) and (5):*

Proof: See Appendix B. ■

The pmf's of T_1^{RND} and T_1^{RND} are validated by simulation in Fig. 7. $E[T_1^{RND}] = 576 \gg M = 100$, and $E[T_2^{RND}]$=137.

Network coding based broadcast (NC): In dissemination via NC, a listening node stores all packets and keeps track of the rank of the NC coefficient matrix $\mathbf{A}$ (see Eq. 2). When a new packet is successfully received, the random coefficient is extracted and appended to $\mathbf{A}$ as a new row and the receiver calculates rank($\mathbf{A}$). If $\mathbf{A}$ reaches full rank, the receiver can successfully decode the stored packets to recover the original data set. Otherwise, this receiver waits for more coded packets.

Let D_v denote the dimension of the subspace spanned by the network coded packets successfully received at node v, where $D_v = \text{rank}(A_v)$. Hence, "rank" and "subspace dimension" are used interchangeably throughout this paper. When v receives a coded packet from the disseminator, D_v increases if and only if the newly received packet is not spanned by v's subspace L_v. The following proposition estimates the probability that a node's subspace dimension increases by one after it receives a new coded packet.

Proposition 1. *In NC based broadcast, the conditional probability that the receiver v's dimension increases by one upon receiving a coded packet from a disseminator, given that it's current subspace dimension equals D_v, equals*

$$P_{D_v} = 1 - q^{-(M-D_v)}, \tag{6}$$

where q is the finite field size, M is the number of packets in the reference data set.

Proof: Consider the probability that a newly received coded packet is not spanned by v's subspace L_v. ■

If we adopt a reasonably large finite field in NC operations, $P_{D_v} \approx 1$ for all values of D_v. For example, $P_0 > \cdots > P_{98} = 0.9998 > P_{99} = 0.984$ when $q = 64$, $M = 100$. In other words, every received coded packet can be assumed innovative in practical operations. Therefore, we have the following intermediate result on NC scheme's conditional completion time in a single source-receiver pair case.

Lemma 2. *Consider one source-receiver pair where the source has the full-set of M packets and the receiver has a subset of C packets, initially. Each packet transmission on the link has a success probability of Q_1. The conditional pmf for T^{NC}, the completion time for the transfer by sending one network coded packet per transmission opportunity, is approximated by a negative binomial random variable:*

$$\Pr(T^{NC} = n | C = c) \approx \binom{n-1}{n-M+c} Q_1^{M-c} (1-Q_1)^{n-M+c}, \quad n \geq M-c, \tag{7}$$

The following theorem characterize the pmf's of NC based broadcast's dissemination completion time in a three-node case.

Theorem 2. *In a three-node case (Fig. 6), the two dissemination completion components $T_1^{NC} + T_2^{NC}$ for NC based broadcast have the following (unconditional) pmf's:*

$$\Pr(T_1^{NC} = t_1) = \binom{t_1-1}{t_1-M} Q_1^M (1-Q_1)^{t_1-M}, \quad t_1 \geq M, \tag{8}$$

$$\begin{aligned} &\Pr(T_2^{NC} = t_2) \\ &= \sum_{c=0}^{M} \binom{t_2-1}{t_2-M+c} Q_1^{M-c} (1-Q_1)^{t_2-M+c} \\ &\quad \sum_{t_1=M}^{\infty} \binom{t_1}{c} Q_2^c (1-Q_2)^{t_1-c} \binom{t_1-1}{t_1-M} Q_1^M (1-Q_1)^{t_1-M}, \\ &\qquad t_2 \geq 0. \end{aligned} \tag{9}$$

Proof: See Appendix C. ■

The pmf's of T_1^{NC} and T_2^{NC} are validated by simulation in Fig. 8. $E[T_1^{NC}]$=111, and $E[T_2^{NC}]$=24.7. The gains from NC in data dissemination is clear from the contrast between Fig. 7 and Fig. 8. Pmf's of T_1^{NC} and T_2^{NC} from NC based broadcast in Fig. 8 are much more concentrated, and the mean dissemination completion times are nearly an order of magnitude less than the random broadcast case, i.e., 111 versus 576 for node 1's completion time, and 24.7 versus 137 for node 2's additional time.

$$\begin{aligned}
&\Pr(T_2^{RND} = t_2) \\
&\approx \sum_{c=0}^{M} \sum_{k_1=M-c-1}^{t_2-1} \sum_{j_1=0}^{M-c-1} (-1)^j \binom{M-c-1}{j} \left(1 - \frac{j+1}{M-c}\right)^{k_1} \binom{t_2-1}{k_1} \left(Q_1 \frac{M-c}{M}\right)^{k_1+1} \left(1 - Q_1 \left(\frac{M-c}{M}\right)\right)^{t_2-k_1-1} \\
&\sum_{t_1=M}^{\infty} \sum_{k_2=c}^{t_1} \binom{M}{c} \sum_{i=0}^{c} (-1)^i \binom{c}{i} \left(\frac{c-i}{M}\right)^{k_2} \binom{t_1}{k_2} Q_2^{k_2} (1-Q_2)^{t_1-k_2} \\
&\sum_{k_3=M-1}^{t_1-1} \sum_{l=0}^{M-1} (-1)^l \binom{M-1}{l} \left(1 - \frac{l+1}{M}\right)^{k_3} \binom{t_1-1}{k_3} Q_1^{k_3+1} (1-Q_1)^{t_1-k_3-1}, \quad t_2 \geq 0.
\end{aligned} \tag{5}$$

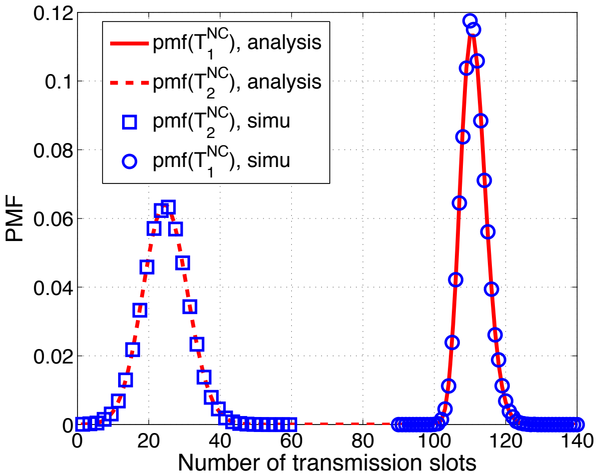

Fig. 8. Pmf's of T_1^{NC} and T_2^{NC} when M=100, Q_1=0.9, Q_2=0.7. Analytical curves of pmf(T_1^{NC}) and pmf(T_2^{NC}) are plotted from (8) and (9).

B. An Infinite 1-D Lattice Network

We have compared the pmf's of dissemination completion time in a three-node network, with and without NC. In a sufficiently long 1-D network, where the source data (at the origin) is to be propagated to all nodes, the average *dissemination velocity* in steady state essentially determines the time needed for completion (beyond any initial transient period). The dissemination velocity is a function of individual reception probabilities $Q_1, Q_2, \ldots, Q_N$ within communication depth. These reception probabilities in turn depend on topology and communication parameters, such as inter-vehicle spacing, transmit power, wireless channel, etc. In this subsection, we first prove that the data dissemination in NC based broadcast reaches a steady-state. The dissemination velocity in steady state is then explicitly derived from a set of detailed balance equations. Since it does not appear feasible to derive a similar closed form result for random broadcast, its dissemination performance in a large network is studied using simulation.

Let t_k denote the time instant when node k first assumes the disseminator role. Then according to our scheduling model, nodes $\{1, \cdots, k\}$ all have the complete reference data set by t_k. Let $D_i(t_k)$ denote the subspace dimension (a.k.a. rank($\mathbf{A}$)) at the i-th node at t_k. At time t_k, according to our communication model (nodes beyond the communication depth N have negigible reception probabilities), node $k+N$ and further have received no packets, i.e., $D_{k+N+i}(t_k) = 0, \ \forall i \geq 0$. Starting from t_k, node k keeps broadcasting coded packets till transfer to the (next) recipient node $k+1$ is complete, at time instance denoted by t_{k+1}. Clearly, $D_i(\cdot)$ over time at any node i constitutes a non-decreasing sequence.

At these embedded time instances $\{t_1, \cdots, t_k, \cdots\}$, the system's status is determined by the instantaneous *dissemination wave front* (the dissemination slope), represented by the subspace dimensions (ranks) at the 2nd, 3rd,$\cdots$, N-th recipients (relative to current disseminator). As the upper figure in Fig. 9 shows, the system status at t_k is solely determined by the vector $S_k = \{D_{k+1}(t_k), \cdots, D_{k+N-1}(t_k)\}$, the dimensions of the $N-1$ recipients at t_k.

All possible system status (states) at these embedded time instances and the transition probabilities among them form a discrete-time Markov chain. We use $\mathbf{S}$ to record the system status over time:

$$\mathbf{S} = \{S_1, S_2, \cdots, S_k, \cdots\}, \tag{10}$$

in which

$$S_k = \{s_1^k, s_2^k, \cdots, s_{N-1}^k\}, \ s_i^k = D_{k+i}(t_k), \quad \forall 1 \leq i \leq N-1\,. \tag{11}$$

Note that S_k is used to emphasize the state at time t_k. Since S_k is not necessarily a unique Markov chain state, it is possible that $S_{k_1} = S_{k_2}$, when $k_1 \neq k_2$. For convenience, we overload $\mathbf{S}$ by calling it Markov chain $\mathbf{S}$ from now on.

The following proposition states an important property of this discrete-time chain $\mathbf{S}$.

Proposition 2. *A transition $S_k \rightarrow S_{k+1}$ is of non-zero probability if and only if*

$$s_i^{k+1} \geq s_{i+1}^k, \quad \forall i = \{1, 2, \cdots, N-2\}. \tag{12}$$

Since s_{i+1}^k and s_i^{k+1} are the subspace dimensions at the same node at time instances t_k and t_{k+1}, Proposition 2 states the fact that a node's subspace dimension is non-decreasing. Fig. 9 graphically shows the $S_k \rightarrow S_{k+1}$ transition, which is composed of *"right shift"* and *"add"* operations.

The next theorem proves that the NC based data dissemination in an infinity 1-D lattice network will reach equilibrium.

Theorem 3. *The discrete-time Markov chain $\mathbf{S}$ is irreducible and ergodic, and hence possesses a steady state distribution (equilibrium).*

Proof: See Appendix D. ∎

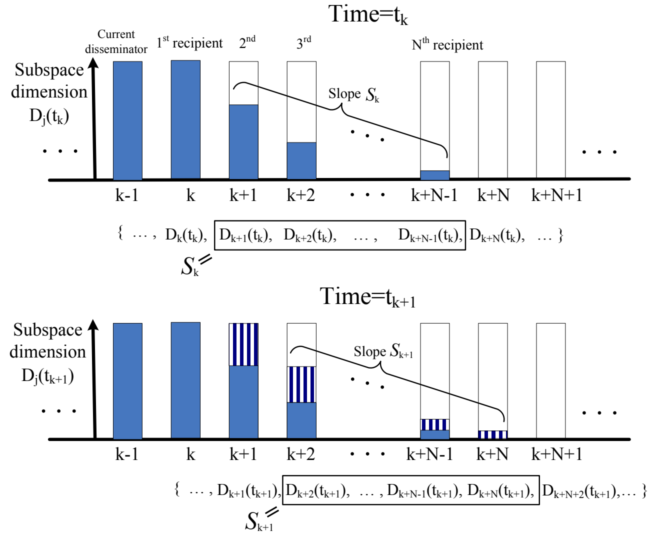

Fig. 9. Definitions of dissemination wave front (dissemination slope) and the Markov chain in data dissemination problem in an infinity 1-D lattice network.

Using mean value approach, we can explicitly derive the mean value $\bar{S}$, as well as the average dissemination velocity in equilibrium.

Corollary 1. *$\bar{S} = \{\bar{s}_1, \cdots, \bar{s}_{N-1}\}$, the average subspace dimensions at the $N-1$ recipients (2nd to N-th recipient relative to current disseminator) in equilibrium, are*

$$\bar{s}_k = M \frac{\sum_{i=k+1}^{N} Q_i}{\sum_{j=1}^{N} Q_j}, \quad \forall 1 \le k \le N-1, \tag{13}$$

which are solutions to the following set of equations:

$$\begin{cases} \bar{s}_2 + (M - \bar{s}_1)\dfrac{Q_2}{Q_1} = \bar{s}_1, \\ \bar{s}_3 + (M - \bar{s}_1)\dfrac{Q_3}{Q_1} = \bar{s}_2, \\ \cdots \\ \bar{s}_{N-1} + (M - \bar{s}_1)\dfrac{Q_{N-1}}{Q_1} = \bar{s}_{N-2}, \\ (M - \bar{s}_1)\dfrac{Q_N}{Q_1} = \bar{s}_{N-1}, \end{cases} \tag{14}$$

where Q_j is the reception probability at the j-th recipient.

Proof: Assuming each new packet is innovative, the time needed for a disseminator to transfer to its immediate next node is a Binomial random variable with mean $\frac{M-\bar{s}_1}{Q_1}$. The number of packets received by the j-th ($2 \le j \le N-1$) recipient by the end of the dissemination step is $(M-\bar{s}_1)\frac{Q_j}{Q_1}$. The first $N-2$ equations in (14) come from the *detailed balance condition*:

$$\bar{s}_j + (M - \bar{s}_1)\frac{Q_j}{Q_1} = \bar{s}_{j-1}, \quad 2 \le j \le N-1. \tag{15}$$

The last equation is the boundary condition, i.e., the N-th receiver (node $k+N$ in Fig. 9) has no prior packets when it first enters the communication range of a disseminator (node k). ■

Given $\bar{S}$, we have the average data dissemination velocity of a reference data set measured by hops per slot in the following theorem.

Theorem 4. *Using NC based broadcast, the dissemination velocity of a reference data set in steady state in an infinite 1-D VANET is:*

$$V_{NC}\,(\text{hop/slot}) = \frac{1}{M}\sum_{i=1}^{N} Q_i, \tag{16}$$

Proof: We omit any inter-frame spacings as negligible overheads. Under the ideal scheduling model, a disseminator on average needs $M - \bar{s}_1$ successful transmissions to complete transfer to the first recipient. Therefore, the dissemination velocity is $\frac{Q_1}{M-\bar{s}_1}$ and the result follows by plugging in $\bar{s}_1 = M\left(1 - \frac{Q_1}{\sum_{j=1}^{N} Q_j}\right)$ from (13). ■

V. Simulation

In this section, simulation is used to validate our analytical results for NC based dissemination, and also round out our investigation by demonstrating gains from NC relative to 1) random broadcast, and 2) a perfect feedback scheme.

Vehicle movement introduces non-zero Doppler spread. However, the 156kHz carrier spacing in current 10MHz 802.11p channel meets the requirements imposed by the observed 1.6kHz maximum Doppler spread in several measurement studies of V2V wireless channel [26], [27]. Based on empirical measurements on received signal strength in vehicular environment [28], our simulation adopts Rayleigh fading [29] links. All nodes use the same transmission power p and modulation scheme to broadcast packets. The wireless channel undergoes Rayleigh fading, and the path loss exponent is α. Let γ_k be the received power for a transmission from a source that is k hops (inter-vehicle spacing is d meters) away from the receiver; then γ_k is an exponentially distributed random variable with mean $p(kd)^{-\alpha}$ with the following probability distribution function (pdf):

$$f_k(\gamma_k) = \frac{1}{p(kd)^{-\alpha}} \exp\left(-\frac{\gamma_k}{p(kd)^{-\alpha}}\right), \quad \forall \gamma_k \ge 0. \tag{17}$$

There is no multi-access interference according to our ideal scheduling model, and the receiver can decode the packet successfully if and only if its received signal-to-noise-ratio (SNR) exceeds a decoding threshold, i.e.,

$$P_{\text{succ}} = \Pr\left(\gamma_k / n_0 \ge z\right), \quad z > 1, \tag{18}$$

where n_0 is the noise power, and z is the capture threshold whose value depends on the channel coding and modulation. In this 1-D lattice network, the reception probability at the k-th recipient relative to current disseminator, is simply:

$$Q_k = \exp\left(-z\, n_0\, p^{-1}\,(kd)^{\alpha}\right), \quad k \ge 1. \tag{19}$$

We compare three schemes - random (RND), network coding based (NC), and with perfect feedback. In RND, a disseminator keeps broadcasting random packets from its buffer till the immediate next node receives the complete data set, leading to significant redundancy due to lack of feedback. Perfect feedback scheme simply provides an upper

TABLE I
PARAMETERS USED IN DATA DISSEMINATION SIMULATION

Platoon size	500 vehicles	M	100 pks
p	1e-5 Watt	d	100 m
α	3	q	128
R_{BPSK}	3 Mbps	z_{QPSK}	5 dB
R_{QPSK}	6 Mbps	z_{QPSK}	8 dB

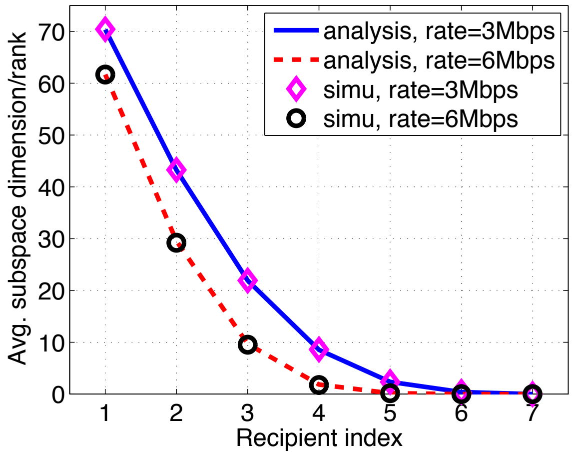

Fig. 10. Dissemination wave front (dissemination slope) in equilibrium. Analytical curves are $\bar{S} = \{\bar{s}_1, \bar{s}_2, \cdots, \bar{s}_{N-1}\}$ from (13) in corollary 1.

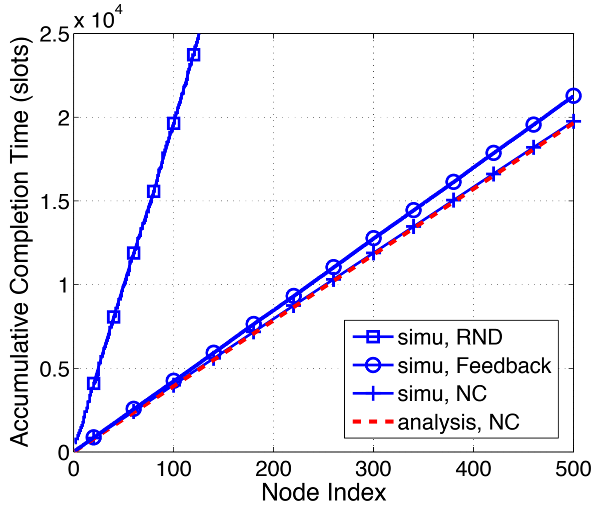

Fig. 11. Cumulative dissemination completion time of all three (RND, Feedback, and NC) schemes. The curves' slopes (also the reciprocals of dissemination velocities) are 205.70, 42.57, and 39.52, respectively. The analytical line has a slope of $\frac{1}{V_{NC}} = 39.31$, where $V_{NC} = 0.0254$ is from (16).

TABLE II
RECEIVED AND TRANSMITTED PACKETS PER NODE IN DATA DISSEMINATION

Schemes	pks rcvd	rcvd non-innovative pks	pks sent
NC	100	<0.01	39.5
RND	524.9	424.9	205.7
Feedback	108.1	8.1	42.6

bound for non-network-coding solutions; with feedback, the disseminator only transmits packets that are not yet received by the first recipient. In NC, a current disseminator broadcasts network coded packets and embeds the NC coefficient in the packet. Before a receiver can decode the data set, it stores all packets and NC coefficients. When a new packet is successfully received, the random coefficient is extracted and appended to the coefficient matrix $\mathbf{A}$ (see Eq. 2) as a new row. Then the receiver calculates the rank of $\mathbf{A}$. If rank($\mathbf{A}$) reaches M, the data set size, the receiver can successfully recover the original data set. Otherwise, it waits for more coded packets. Parameter values used in simulations are summarized in Table I.

We first validate the analytical result of the dissemination wave front (dissemination slope) in equilibrium in Fig. 10. Analytical result $\bar{S} = \{\bar{s}_1, \bar{s}_2, \cdots, \bar{s}_{N-1}\}$ is from the close form (13) in corollary 1. In simulation, $\bar{s}_j$ is computed by averaging node $(k+j)$'s subspace dimension when node k first reaches full buffer, for $k = \{1, 2, \cdots, 500 - j\}$.

We plot the cumulative dissemination completion time versus node index in Fig. 11. The reciprocal of the slope of the cumulative completion time curve is the average dissemination velocity. For the NC based broadcast, the simulation result closely matches our predicted analytical estimate in (16). Note that NC and perfect feedback achieves about the same steady state velocity; however NC provides a better guarantee that a packet is new or innovative as compared to perfect feedback and achieves a slightly higher steady state velocity, as explained below.

Table II shows (1) the average number of packets that a disseminator broadcasts in order to transfer to the immediate next node, (2) the average number of total and useful packets that a node receives. It is clear that over 80% of the transmitted and received packets in random broadcast scheme are duplicates due to the lack of feedback. Perfect feedback cannot eliminate all redundancy (8% redundancy in simulation) because i) feedbacks are not from all the multiple recipients; ii) even with feedbacks from all recipients, there might not exist a packet which is new to all. NC best exploits the broadcast nature of wireless and renders (almost) each received packet at any receipient innovative, and hence achieves the highest efficiency.

Irregular 1-D topology with variable inter-vehicle spacings, a good representation of a practical VANET topology, is also simulated. As Fig. 12 shows, the average per step completion time for both NC based broadcast and broadcast with perfect feedback are slightly larger as the topology dynamics becomes more significant. Standard deviations from both increase a lot. This indicates that our analytical results from 1-D lattice topology are still good predictor for the average dissemination velocity and the cumulative dissemination completion time in 1-D irregular topology, which is validated in Fig. 13.

VI. CONCLUSION

This work has explored horizontal peer-to-peer content distribution using V2V ad hoc communications to supplement vertical vehicular data download from cellular and road side DSRC infrastructures. We studied a single source dissemination scenario with an ideal scheduler to understand the limits (upper bound) of the benefits from network coding. The pmf's of dissemination completion time in the three-node case for

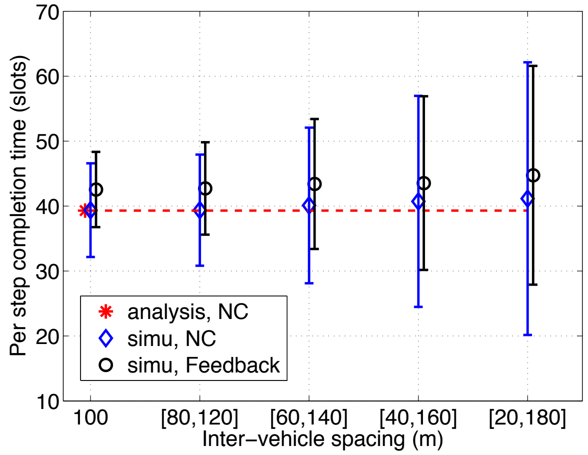

Fig. 12. Average per step completion time with standard deviation error bar in irregular 1-D topology. Inter-vehicle spacing is of uniform distribution.

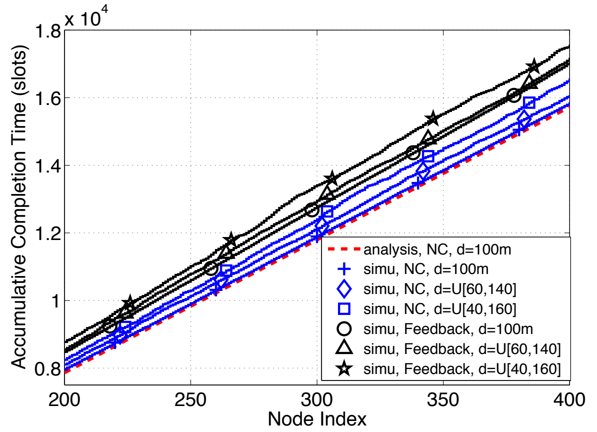

Fig. 13. Cumulative dissemination completion time of Feedback and NC schemes in irregular 1-D topology. Inter-vehicle spacing is of uniform distribution.

both random and NC based broadcast are explicitly derived. For a 1-D infinite network, we proved that NC based broadcast reaches steady-state, and derived a closed form result for the dissemination velocity in equilibrium. Simulation results in large (500-node) 1-D lattice and 1-D irregular network validated our analytical results vis-a-vis gains of NC relative to random broadcast and perfect feedback, in presence of Rayleigh fading links.

ACKNOWLEDGEMENTS

The authors would like to thank Mohammad Hamed Firooz (Electrical engineering, University of Washington) for useful discussions on network coding, and Peizhe Shi (Applied Mathematics, University of Washington) for helpful pointers on Markov models. Finally, the authors wish to thank the reviewers for their thoughtful and insightful comments.

APPENDIX

A. Proof of Lemma 1 - conditional pmf of T^{RND} in a single source-receiver pair case

In order to prove lemma 1, we first prove the following intermediate result from classic coupon collection problem.

Consider a source-receiver pair where the source has the M packets for dissemination and the receiver has none. The link is always successful. Let $U_{M,n}$ be the number of receiver's missing packets after n ($n \geq 1$) trials, its pmf is:

$$\Pr(U_{M,n}=j)=\binom{M}{j}\sum_{k=0}^{M-j}(-1)^k\binom{M-j}{k}\left(\frac{M-j-k}{M}\right)^n, \quad j\in\{\max\{M-n,0\},...,M-1\}. \tag{20}$$

Proof: Let $A_{M,n,i}$ denote the event of missing packet i after n trials when population is M. Using the inclusion-exclusion rule of combinatorics, we can find the probability that at least one packet is missing:

$$\Pr(\cup_{i=1}^{M}A_{M,n,i}) = \sum_{k=1}^{M}(-1)^{(k-1)}\binom{M}{k}\left(\frac{M-k}{M}\right)^n. \tag{21}$$

And also the probability that no one is missing:

$$\begin{aligned}\Pr(\cap_{i=1}^{M}A_{M,n,i}^c) &= 1-\Pr(\cup_{i=1}^{M}A_{M,n,i}),\\ &= \sum_{k=0}^{M}(-1)^k\binom{M}{k}\left(\frac{M-k}{M}\right)^n.\end{aligned} \tag{22}$$

To find the probability that exactly j packets are missing, we first choose j packets to be excluded out of M, then let all $M-j$ remaining packets to be received after n trials. The pmf of $U_{M,n}$ is calculated as follows:

$$\Pr(U_{M,n}=j)=\binom{M}{j}\Pr(\cap_{i=1}^{M-j}A_{M,n,i}^c), \tag{23}$$

where $\Pr(\cap_{i=1}^{K}A_{M,n,i}^c) = \sum_{k=0}^{K}(-1)^k\binom{K}{k}\left(\frac{K-k}{M}\right)^n$ is from (22). The result follows. ∎

Eq. (20) is valid for the case that the receiver has no prior knowledge, and one random packet is reliably received in each trial (transmission). We next generalize (20) to the unreliable link case. Let $\bar{U}_{M,n,Q}$ be the number of missing packets after n unreliable trials, where each trial is of success probability Q. We first choose k trials out of n to be successful, then apply (20) after substituting total trial number k for n. $\bar{U}_{M,n,Q}$'s pmf is:

$$\Pr(\bar{U}_{M,n,Q}=j)=\sum_{k=M-j}^{n}\Pr(U_{M,k}=j)\Pr(B(n,Q)=k), \tag{24}$$

where $B(n,Q)$ is a binomial random variable with success probability Q.

In lemma 1, the receiver possesses c distinct packets before dissemination, hence, only $M-c$ packets are left to be filled. And the wireless link has a success probability of Q_1. In a transmission, if the receiver successfully receive a packet, and this packet is not among the existing c packets, we call this a *valid* transmission. Its probability is simply $\bar{Q}_1 = Q_1\left(1-\frac{c}{M}\right)$. To derive the conditional completion time's pmf, we consider the event that the receiver collects all but one packets after $n-1$ trials, and receive the exact missing packet in the last (n-th) trial. That is,

$$\Pr(T^{RND}=n \mid C=c)=\frac{\bar{Q}_1}{M-c}\cdot\Pr(\bar{U}_{M-c,\,n-1,\,\bar{Q}_1}=1). \tag{25}$$

The result follows after substitution of (24) into the above expression. ■

B. Proof of Theorem 1 - pmf's of T_1^{RND} and T_2^{RND} in a three-node network

Since node 1 has no prior knowledge, (4) immediately follows from (3) after setting $c = 0$.

Let random variable C denote the number of distinct packets received by node 2 when transfer to node 1 is complete. Conditioned on C, we have

$$\Pr(T_2^{RND} = t_2) = \sum_{c=0}^{M} \Pr(T_2^{RND} = t_2 \mid C = c) \Pr(C = c), \tag{26}$$

$\Pr(T_2^{RND} = t_2 \mid C = c)$ is given by (3) in Lemma 1. Conditioned on node 1's completion time T_1^{RND}, $\Pr(C = c)$ can be calculated as following:

$$\begin{aligned}\Pr(C = c) &= \sum_{t_1=M}^{\infty} \Pr(C = c \mid T_1^{RND} = t_1) \Pr(T_1^{RND} = t_1),\\ &\approx \sum_{t_1=M}^{\infty} \Pr(\text{Rx } c \text{ distinct pks out of } t_1 \text{ rnd b'casts})\\ &\quad \Pr(T_1^{RND} = t_1),\\ &= \sum_{t_1=M}^{\infty} \Pr(\bar{U}_{M,t_1,Q_2} = M - c) \Pr(T_1^{RND} = t_1).\end{aligned} \tag{27}$$

In the derivation, we assume that successful packet reception events by nodes 1 and 2 are independent. $\Pr(\bar{U}_{M,t_1,Q_2} = M - c)$ is given by (24). The result follows after substituting (27) and (3) into (26). ■

C. Proof of Theorem 2 - pmfs of T_1^{NC} and T_2^{NC} in a three-node network

Pmf of Node 1's completion time (8) immediately follows from (7) after setting $c = 0$.

Using conditional probability, pmf of T_2^{NC} is calculated as follows:

$$\begin{aligned}\Pr(T_2^{NC} = t_2) &= \sum_{c=0}^{M} \Pr(T_2^{NC} = t_2 \mid C = c) \Pr(C = c),\\ &= \sum_{c=0}^{M} \Pr(T_2^{NC} = t_2 \mid C = c)\\ &\quad \sum_{t_1=M}^{\infty} \Pr(C = c \mid T_1^{NC} = t_1) \Pr(T_1^{NC} = t_1),\end{aligned} \tag{28}$$

where $\Pr(T_2^{NC} = t_2 | C = c)$ and $\Pr(T_1^{NC} = t_1)$ are given by (7) and (8), respectively. Since we assume every coded packet is innovative, conditioned on T_1^{NC} - Node 1's completion time - the pmf of the number of innovative packets received by Node 2 follows a Binomial distribution:

$$\Pr(C = c \mid T_1^{NC} = t_1) = \binom{t_1}{c} Q_2^c (1 - Q_2)^{t_1 - c}. \tag{29}$$

By substituting (7), (29), (8) into (28), the result follows. ■

D. Proof of Theorem 3

It is easy to verify that the chain $\mathbf{S}$ is memoryless, that is, conditioned on the present state S_k, its future S_{k+1} and past S_{k-1} are independent. Because a node's rank $s_i \in \{0, 1, \cdots, M\}$, $\forall 1 \le i \le N - 1$, this Markov chain has $(M + 1)^{N-1}$ (finite) states.

A finite state Markov chain has steady state distribution if it is irreducible and ergodic. We first prove the irreducibility of $\mathbf{S}$ by showing that any state has a non-zero probability path to any other states. Let's consider two states $\{s_1, s_2, \cdots, s_{N-1}\}$ and $\{s_1^*, s_2^*, \cdots, s_{N-1}^*\}$. According to proposition 2, the transition $\{s_1, s_2, \cdots, s_{N-1}\} \rightarrow \{s_2, s_3, \cdots, s_{N-1}, s_1^*\}$ is of non-zero probability, regardless of the value of s_1^*. In an iterative way, the following path of non-zero probability is constructed:

$$\begin{aligned}&\{s_1, s_2, \cdots, s_{N-1}\} \rightarrow \{s_2, \cdots, s_{N-1}, s_1^*\} \rightarrow \cdots\\ &\rightarrow \{s_{N-1}, s_1^*, \cdots, s_{N-1}^*\} \rightarrow \{s_1^*, s_2^*, \cdots, s_{N-1}^*\}.\end{aligned} \tag{30}$$

We have proved that all states of $\mathbf{S}$ communicate with each other, hence $\mathbf{S}$ is irreducible.

Next we prove that $\mathbf{S}$ is ergodic (aperiodic and positive recurrent). It can be verified that the state $\{s_1, s_2, \cdots, s_{N-1}\}$, $s_1 \ge s_2 \ge \cdots \ge s_{N-1}$ has a non-zero probability self-loop. Therefore, this state is aperiodic. Since all states of $\mathbf{S}$ belong to the same class (irreducibility), they are all aperiodic. Further, all states of a finite state irreducible Markov chain are positive recurrent, hence, $\mathbf{S}$ is ergodic. Due to its irreducibility and ergodicity, $\mathbf{S}$ has a steady state distribution that is independent of its initial state. ■

References

[1] H. Hartenstein, K. Laberteaux, and E. Corporation, *VANET: Vehicular Applications and Inter-Networking Technologies*. Wiley Online Library, 2010.

[2] D. Jiang and L. Delgrossi, "IEEE 802.11p: Towards an international standard for wireless access in vehicular environments," in *Proc. IEEE Vehicular Tech. Conf.*, 2008, pp. 2036–2040.

[3] J. Zhu and S. Roy, "MAC for dedicated short range communications in intelligent transport system," *IEEE Commun. Mag.*, vol. 41, no. 12, pp. 60–67, Dec. 2003.

[4] "Dedicated Short Range Communications (DSRC)," 2003. [Online]. Available: http://www.standards.its.dot.gov/Documents/advisories/dsrc_advisory.htm

[5] "IEEE Std 1609 family, IEEE Trial-Use Standard for Wireless Access in Vehicular Environments(WAVE)," Nov. 2006.

[6] "IEEE P802.11p/D5.0, IEEE 802.11 Amendment 7: Wireless Access in Vehicular Environments," Nov. 2008.

[7] S. Biswas, R. Tatchikou, and F. Dion, "Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety," *IEEE Commun. Mag.*, vol. 44, no. 1, pp. 74–82, 2006.

[8] F. Ye, R. Yim, J. Guo, J. Zhang, and S. Roy, "Prioritized broadcast contention control in VANET," in *Proc. IEEE Int'l Conf. on Communications (ICC)*, 2010, pp. 1–5.

[9] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry, "Epidemic algorithms for replicated database maintenance," in *Proc. 6th ACM Symposium on Principles of Distributed Computing*, 1987, pp. 1–12.

[10] Z. Haas, J. Halpern, and L. Li, "Gossip-based ad hoc routing," *IEEE/ACM Trans. Netw.*, vol. 14, no. 3, pp. 479–491, 2006.

[11] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, 2003.

[12] S. Li, R. Yeung, and N. Cai, "Linear network coding," *IEEE Transaction on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003.

[13] M. Ghaderi, D. Towsley, and J. Kurose, "Reliability gain of network coding in lossy wireless networks," in *Proc.IEEE INFOCOM*, 2008, pp. 2171–2179.

[14] C. Fragouli, J. Widmer, and J. Le Boudec, "Efficient broadcasting using network coding," *IEEE/ACM Trans. Netw.*, vol. 16, no. 2, pp. 450–463, 2008.
[15] Y. Evren Sagduyu and A. Ephremides, "On joint mac and network coding in wireless ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3697–3713, 2007.
[16] K. Lee, S. Lee, R. Cheung, U. Lee, and M. Gerla, "First experience with cartorrent in a real vehicular ad hoc network testbed," in *Proc. Mobile Networking for Vehicular Environments*, 2007, pp. 109–114.
[17] A. Nandan, S. Das, G. Pau, M. Gerla, and M. Sanadidi, "Co-operative downloading in vehicular ad-hoc wireless networks," in *Proc. 2nd annual conf. on Wireless On-demand Network Systems and Services*, 2005, pp. 32–41.
[18] U. Lee, J. Park, J. Yeh, G. Pau, and M. Gerla, "Code torrent: content distribution using network coding in vanet," in *Proc. 1st int'l workshop on Decentralized Resource Sharing in Mobile Computing and Networking*, 2006, pp. 1–5.
[19] S. Ahmed and S. Kanhere, "VANETCODE: network coding to enhance cooperative downloading in vehicular ad-hoc networks," in *Proc. int'l conf. on Wireless Communications and Mobile Computing (IWCMC)*, 2006, pp. 527–532.
[20] M. Johnson, L. De Nardis, and K. Ramchandran, "Collaborative content distribution for vehicular ad hoc networks," in *Proc. Allerton Conf. on Communication, Control, and Computing*, 2006.
[21] M. Li, Z. Yang, and W. Lou, "Codeon: Cooperative popular content distribution for vehicular networks using symbol level network coding," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 223–235, 2011.
[22] Z. Yang, M. Li, and W. Lou, "Codeplay: Live multimedia streaming in vanets using symbol-level network coding," in *Proc. 18th IEEE int'l conf. on Network Protocols (ICNP),*, 2010, pp. 223–232.
[23] D. Lucani, M. Médard, and M. Stojanovic, "Random linear network coding for time-division duplexing: field size considerations," in *Proc. IEEE Global Telecommunications Conf.*, 2009, pp. 1–6.
[24] I. Adler, S. Oren, and S. Ross, "The coupon-collector's problem revisited," *Journal of Applied Probability*, vol. 40, no. 2, pp. 513–518, 2003.
[25] H. Rubin and J. Zidek, "A waiting time distribution arising from the coupon collector's problem," 1965.
[26] L. Cheng, B. Henty, R. Cooper, D. Stancil, and F. Bai, "A measurement study of time-scaled 802.11 a waveforms over the mobile-to-mobile vehicular channel at 5.9 ghz," *IEEE Commun. Mag.*, vol. 46, no. 5, pp. 84–91, 2008.
[27] A. Paier, J. Karedal, N. Czink, C. Dumard, T. Zemen, F. Tufvesson, A. Molisch, and C. Mecklenbräuker, "Characterization of vehicle-to-vehicle radio channels from measurements at 5.2 ghz," *Wireless personal communications*, vol. 50, no. 1, pp. 19–32, 2009.
[28] J. Yin, G. Holland, T. Elbatt, F. Bai, and H. Krishnan, "Dsrc channel fading analysis from empirical measurement," in *1st IEEE int'l conf. on Communications and Networking in China*, 2006, pp. 1–5.
[29] H. Zhu and J. Wang, "Chunk-based resource allocation in ofdma systems-part i: chunk allocation," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2734–2744, 2009.

Fei Ye received the B.Eng. and M.Sc. degrees, both in electronic engineering, from Tsinghua University, Beijing, China, in 2005 and 2007, respectively, and the Ph.D. degree in electrical engineering from the University of Washington, Seattle, in 2011.

His main research interest is protocol design in wired and wireless networks including data center network, wireless LAN, vehicular network, focusing on performance optimization at different layers, i.e., network layer (e.g., multi-hop routing), and MAC layer (e.g., IEEE802.11). His current work is on data dissemination in vehicular networks, and data center networks.

Sumit Roy received the B. Tech. degree from the Indian Institute of Technology (Kanpur) in 1983, and the M. S. and Ph. D. degrees from the University of California (Santa Barbara), all in Electrical Engineering in 1985 and 1988 respectively, as well as an M. A. in Statistics and Applied Probability in 1988. His previous academic appointments were at the Moore School of Electrical Engineering, University of Pennsylvania, and at the University of Texas, San Antonio; presently he is Prof. of Electrical Engineering, Univ. of Washington where his research interests include analysis/design of communication systems/networks, with an emphasis on next generation mobile and wireless networks. He spent 2001-03 on academic leave at Intel Wireless Technology Lab as a Senior Researcher engaged in research and standards development for ultra-wideband systems (Wireless PANs) and next generation high-speed wireless LANs. He served as Science Foundation of Ireland Isaac Walton Fellow during a sabbatical at University College, Dublin (Jan-Jun 2008) and was the recipient of a Royal Acad. Engineering (UK) Distinguished Visiting Fellowship during summer 2011. He has over 70 archival journal and 120 conference publications and his research has been consistently funded by various US national agencies. He was elevated to IEEE Fellow by Communications Society in 2007 for his "contributions to multi-user communications theory and cross-layer design of wireless networking standards".

His activities for the IEEE Communications Society includes membership of several technical and conference program committees (notably Technical Committee on Cognitive Networks). He has served as Associate Editor for all the major ComSoc publications in his area at various times, including the IEEE Trans. Communications and IEEE Trans. on Wireless Communications. He currently serves on the Editorial Board for IEEE Trans. Communications, IEEE J. Intelligent Transportation Systems and the IEEE Trans. Smart Grids. His other notable professional activities include reviewing proposals for various international research panels (Singapore, HongKong, Qatar etc.) and serving as external examiner for Ph.D. thesis (Singapore, Australia, Canada etc.).

Haobing Wang received the B.S. degree in electronic engineering, from Shanghai Jiao Tong University, Shanghai, China, in 2010. Since September 2010, he has been working towards the Ph.D. degree in electrical engineering at the University of Washington, Seattle. His current research interest is network coding, cognitive radio networks and vehicular networks.