On Infostation Density of Vehicular Networks

Vinod Kone[†], Haitao Zheng[†], Antony Rowstron[§], and Ben Y. Zhao[†]
[†]Department of Computer Science, U. C. Santa Barbara
[§]Microsoft Research, Cambridge, United Kingdom
{*vinod, htzheng, ravenben*}*@cs.ucsb.edu*, {*antr*}*@microsoft.com*

***Abstract*—Vehicle-to-Vehicle and Vehicle-to-Roadside communications are going to become an indispensable part of the modern day automotive experience. For people on the move, vehicular networks can provide critical network connectivity and access to real-time information. Infostations play a vital role in these networks by acting as gateways to the Internet and by extending network connectivity. In this context, an important question is "What is the minimum number of infostations that need to be deployed in an area in order to support vehicular applications?" Optimizing infostation density is vital to understanding and reducing the cost of deployment and management. In this paper, we examine the required infostation density in a highway scenario using different data dissemination models. We start from a simple analysis that captures the required density under idealized assumptions. We then run detailed QualNet simulations on both controlled and realistic vehicular traces to observe the information density trends in practical environments, and consequently propose techniques to improve dissemination performance and reduce the required infostation density.**

I. Introduction

Vehicular networking is becoming a reality. The current generation of GPS-based navigation and information devices already incorporate wireless connectivity. As they grow to support features such as streaming traffic data, streaming audio and location-aware content, it is important to understand how to provide infrastructure to efficiently support these devices. In addition, as application data grows to include more time sensitive data (*e.g.*, free parking spaces, traffic flow), users must update devices more frequently. A major challenge has been how to get timely data to these devices.

The current generation of devices uses a number of different wireless technologies to achieve this, including leveraging cellular infrastructures or broadcast-based infrastructures such as satellite or FM sub-carrier bands. A cellular approach implies devices "pull" data from the network. Devices can use the cellular infrastructure by either building cellular capability into the device, *e.g.* GM's OnStar system, or by connecting to a mobile phone using BlueTooth, *e.g.* the TomTom One. An alternative is to use a broadcast medium that pushes data to devices, including satellite networks, *e.g.* Sirius, or FM sub-carrier bands, *e.g.* NAVTEQ.

Pragmatically, both cellular and broadcast approaches face disadvantages. The broadcast medium requires significant investments in large-scale infrastructure such as a global satellite network. In contrast, a cellular approach relies on infrastructure owned by a telecom company that controls network access and therefore requires per-country negotiation. Furthermore, incompatible cellular networks partition the user base depending on region. While connections made through users' mobile phones are region-independent, it has the major disadvantage that different data plans control the volume of data that can be pulled to the device. This information needs to be exposed to the information device, or else it can be very expensive for the user!

Since cellular and broadcast approaches are less than ideal, we turn our attention to an alternative based on infostations (or access points) assisted with vehicle-to-vehicle propagation. Infostations are attractive because they can be incrementally deployed by a variety of vendors with comparatively low upfront costs. Zero cost vehicle-to-vehicle communications can decrease the infostation density, thereby significantly reducing cost of deploying infostations. This leads us to the question: *Will vehicle-to-vehicle distribution combined with intermittent connectivity to infostations provide a suitable data dissemination system for vehicular networks?* In particular, could a network of WiFi access points, such as those owned by T-Mobile, McDonalds, or Starbucks be used to deliver data to vehicles? What density of infostations would be required? In this paper we attempt to understand the trade-offs and limitations of infostation-based data dissemination, as the first stage of answering these questions.

We achieve this by modeling two different approaches to infostation-assisted dissemination in a vehicular network: *push* and *pull*. In the push model, infostations act as a data carousel continuously transmitting data. Vehicles receive the data when in transmission range of infostations, and then propagate the data to other vehicles, using epidemic data dissemination to extend data transmission to vehicles beyond the range of infostations. In the pull model, each vehicle has a data set that it requires. It attempts to pull the information from infostations, using multi-hop routing if necessary. We assume that each infostation has access to all the requested content.

The main goal of our paper is to understand the *minimum infostation density* required to support different data dissemination applications. We choose to consider a highway scenario because this is one of the most practical deployment scenarios. We begin by building simple analytical models of the two infostation-based dissemination approaches without considering wireless propagation and contention. These abstract models allow us to study in more detail the impact of transmission range, data size, and data lifetime on each approach in an ideal scenario. To further understand trade-offs such as the impact of wireless channel contention, we perform detailed simulation experiments using QualNet [1] with vehicular traces derived from real-world vehicular traces of a segment of a Californian highway.

For each dissemination scheme, we show that the lifetime of data is a dominating factor on required infostation density. Our QualNet simulations show that a number of additional factors play crucial roles in dissemination performance, including the selection of forwarding neighbors, knowledge of nearby vehicle locations, and pull request density in the pull model. We propose some initial techniques to address these challenges and focus on examining their significant impact on the overall required density of infostation. Finally, experiments on a trace derived from US-101 highway provide interesting insights into the impact of the wireless equipment penetration ratio (or wireless vehicle density) on the infostation density.

II. Related Work

There is increasing interest in vehicular networking, with a number of small scale vehicular networking deployments [6], [15], [21].

However, due to the small scale of these deployed systems, most prior work has focused on evaluating particular strategies and dissemination models. Heavy emphasis has been put on push-based dissemination [14], [16], [20], [22], [13], and dissemination through vehicle-to-vehicle propagation [7], [18]. Some recent work on practical systems [6], [15], [3], [10] hints at the feasibility of pull-based distribution.

Nadeem et al. [16] examine vehicle-to-vehicle dissemination using the push model. They compare two push model variants; flooding and dissemination. In the flooding model information is generated by cars and flooded through the vehicular network, which clearly has scalability issues. Therefore, they also consider a push variant where vehicles perform batching of local information received from neighbors before pushing it to their neighbors. In [16], the authors study the impact of varying the set of vehicles that propagate information, such as to those vehicles traveling in the same direction. Wu et al. [20] use analytical models to analyze push information dissemination, but in particular take into account the clustering of vehicles into partitions and then consider dissemination within and across partitions.

Lochert et al. [14] also examine the feasibility of information dissemination in city environments using a push model. They conclude that the potential for low density of equipped vehicles means some infostations are required, and evaluate where they should be placed in a city environment. They also conclude that networked infostations perform better than infostations that simply act as a repeater for information heard from other vehicles. Zhao et al. [22] also model and simulate push based dissemination using infostations. They model the impact of delay at intersections.

In [19], the authors propose a dissemination protocol for comfort applications based on segmentation of the road. Scalability through aggregation of information and intelligent broadcasts are achieved with the help of a digital road-map. Uichin et al. [11] propose a buy-sell based virtual market place on vehicular networks. Dissemination is through neighbors that come into contact with each other. This scheme suits better to urban environments than highway scenarios.

While prior works summarized in the above have examined the performance of vehicle data dissemination (mostly under the push model), little attention was given on the impact of infostation density. Yet it is the major factor that will affect the deployment of infostations and the performance of dissemination. Different from prior work, in this paper we focus on examining the impact of infostation density in highway scenarios, using two dissemination models.

III. Abstract Models

In this section, we develop simple analytical models for the infostation density required to support data dissemination in vehicular networks. A complete and accurate model of vehicular data dissemination is highly complex, because it must capture the impact of a large number of factors ranging from physical transmission characteristics to data lifetimes. Therefore, we derive simple analytical models for push and pull dissemination schemes that capture their defining characteristics. The goal of our models is to understand the high level trends for infostation density and the impact of *content settings*, *i.e.* data size and lifetime.

Specifically, we make the following assumptions:

- We assume the width of the road is much smaller than the transmission range of infostations and vehicles. This assumption is true in general if infostations and vehicles use WiFi to communicate.
- We assume that vehicular density is high, and all vehicles travel in the same direction at a uniform velocity V, which is much lower than the signal propagation rate.
- We assume that the impact of transmission errors and MAC contention/schedule is abstracted into the bandwidth and range of transmissions between infostation and vehicles and among vehicles. This abstraction allows us to perform the analysis without being confined to any specific MAC and physical device configuration.

Table III summarizes the parameters used in our analysis.

	Infostation Parameters
B_{info}	Bandwidth of an infostation
R_{info}	Transmission range of an infostation
D_{info}	Dissemination range - distance between two neighboring infostations, or $1/D_{info}$ is the density of infostations
	Vehicle Parameters
B_{mobile}	Bandwidth of a vehicle when communicating with a peer (depending on the vehicle density).
R_{mobile}	Transmission range of a vehicle
V	Velocity of a vehicle
$I_{arrival}$	Inter-arrival interval of vehicles
	Data Parameters
S	Size of the data i
T	Lifetime of the data i
S_{pull}	Size of the pull request

A. The Push Model

In the push model, the data dissemination range D_{info} represents the maximum distance that the data i can travel to at the end of i's lifetime. We can separate the range into two parts: r, the distance covered by a broadcast from the infostation; and R, the distance covered by the subsequent forwarding by vehicles (in one direction). As shown in Figure 1, the total coverage is $r + 2R$.

We can determine the distance r as:

$$r = 2(d/2 - (d - R_{info})) = 2R_{info} - d = 2R_{info} - \frac{S}{B_{info}}V$$

where $d = \frac{S}{B_{info}}V$.

Next, we derive the *relative* distance covered by vehicles forwarding data i within its lifetime T, which we refer to as R.

$$T = \underbrace{\left\lceil \frac{R}{R_{mobile}} \right\rceil}_{\text{\# of hops}} \cdot \underbrace{\frac{S}{B_{mobile}}}_{\text{delay per hop}} + \underbrace{\frac{S}{B_{info}}}_{\text{AP broadcast delay}}$$

$$\Longrightarrow \quad R = R_{mobile} \left\lfloor T \cdot \frac{B_{mobile}}{S} - \frac{B_{mobile}}{B_{info}} \right\rfloor$$

Therefore, the total distance covered by the data dissemination within its life time T is

$$\begin{aligned} D_{info} &= r + 2R \\ &= 2R_{info} - \frac{S}{B_{info}}V + 2R_{mobile} \left\lfloor T \cdot \frac{B_{mobile}}{S} - \frac{B_{mobile}}{B_{info}} \right\rfloor \end{aligned}$$

When infostations and vehicles use the same radio configuration, B_{info} = B_{mobile} and R_{info} = R_{mobile}, we have

$$D_{info}^{push} \approx \frac{2R_{mobile}B_{mobile}T}{S} - \frac{SV}{B_{mobile}}. \tag{1}$$

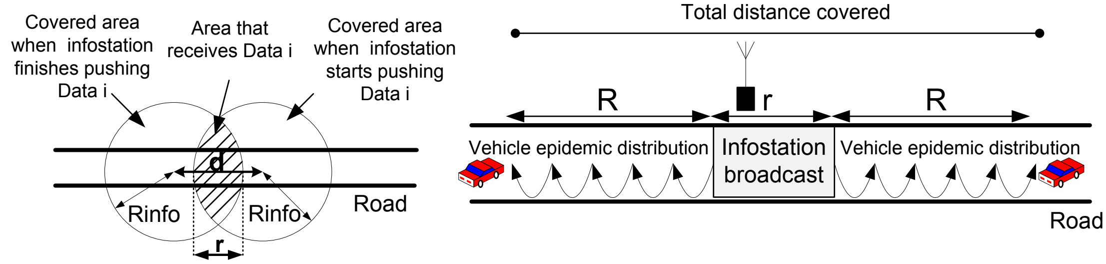


Fig. 1. Coverage of push. (Left) Coverage of infostation's broadcast. Nodes which obtain the complete content of data i are only in the shaded area. $d = \frac{S}{B_{info}} V$. If $d \geq 2R_{info}$ then the shaded area is NULL, *i.e.* no node can download the complete set of data i. (Right) Coverage of subsequent vehicle epidemic content distribution. Vehicles propagate data i to both directions in multi-hops.

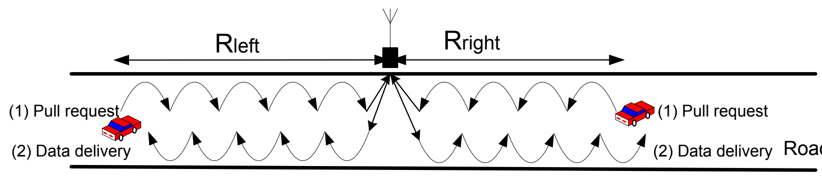


Fig. 2. Coverage of pull. The requesting car (either at left or right of the infostation) first sends out a pull request towards its nearest infostation, forwarded by vehicles in the middle. The infostation then sends the requested data to the vehicle in multi-hops. Note that for both sides, the distance covered by the pull request and the distance covered by the data delivery differ due to mobility.

B. The Pull Model

In this model, infostations never push data. Instead, any vehicle n in need of data i sends a request to nearby infostation to request it on-demand*. Therefore, the data dissemination process consists of two stages: *(i)* vehicles forwarding the pull request to the nearest infostation; and *(ii)* the infostation delivering data i to the source vehicle, if necessary using multi-hop forwarding by other vehicles. We will refer to the time spent in stage one as *the delay of pull* and the time in stage two as *the delay of delivery*. Figure 2 provides a graphic view of the stages.

The distance covered represents the farthest vehicle (from an infostation) which can pull and retrieve data i within the data lifetime. That is, the sum of the delay of pull and the delay of delivery must not exceed the data lifetime. Clearly, whether a vehicle is to the left or right of the infostation matters in this case because the distance traveled by the pull request and its response is different.

First, for a vehicle at R_{right} distance to the right of an infostation, the time required to retrieve data i is:

$$T = \underbrace{\left\lceil \frac{R_{right} - R_{info}}{R_{mobile}} \right\rceil \cdot \frac{S_{pull}}{B_{mobile}} + \frac{S_{pull}}{B_{info}}}_{\text{delay of pull } (t^{pull}_{right})} + \underbrace{\left\lceil \frac{R_{right} - R_{info} + V(t^{pull}_{right} + \frac{S}{B_{info}})}{R_{mobile}} \right\rceil \frac{S}{B_{mobile}} + \frac{S}{B_{info}}}_{\text{delay of delivery}}$$

Similarly, for a node at distance R_{left} to the left of an infostation, the time required to retrieve the data is

$$T = \underbrace{\left\lceil \frac{R_{left} - R_{info}}{R_{mobile}} \right\rceil \cdot \frac{S_{pull}}{B_{mobile}} + \frac{S_{pull}}{B_{info}}}_{\text{delay of pull } (t^{pull}_{left})} + \underbrace{\left\lceil \frac{R_{left} - r_{left} - V(t^{pull}_{left} + \frac{S}{B_{info}})}{R_{mobile}} \right\rceil \frac{S}{B_{mobile}} + \frac{S}{B_{info}}}_{\text{delay of delivery}}$$

where r_{left} is the distance between the infostation and the leftmost vehicles who obtain the complete broadcast of i.

When infostations and vehicles use the same radio configuration, $B_{info} = B_{mobile}$ and $R_{info} = R_{mobile}$ and the data size is limited, *i.e.* $\frac{SV}{B_{mobile} \cdot R_{mobile}} << 1$, we can reduce R_{right} as follows:

$$R_{right} \approx \frac{B_{mobile} \cdot R_{mobile} \cdot T}{S + S_{pull}} - \frac{SV}{B_{mobile}} \frac{S}{S + S_{pull}}. \tag{2}$$

Similarly, we have

$$R_{left} \approx \frac{B_{mobile} \cdot R_{mobile} \cdot T}{S + S_{pull}} + \frac{VS}{B_{mobile}} \frac{S}{S + S_{pull}}.$$

From the above, we can derive the range of dissemination:

$$\begin{aligned} D^{pull}_{info} &= R_{right} + R_{left} \\ &\approx \frac{2B_{mobile} T R_{mobile}}{S + S_{pull}} \end{aligned} \tag{3}$$

We note that the above density measure only applies to scenarios where in a data lifetime T, only one vehicle requests the data. In this case, compared to the push model, the infostation density increases by approximately a factor of $\frac{S+S_{pull}}{S}$. On the other hand, in order to support multiple N requests in parallel, the bandwidth of B_{mobile} (and B_{info}) could drop by a factor of $f(N)$ where $f(N)$ represents the bandwidth degradation from multi-user contention. This maps to an increase of $f(N)$ in the infostation density.

*Note that when vehicles cache data, a node in between the infostation and n can send the data to n. However, our analysis assumes that data i is not available at any vehicles.

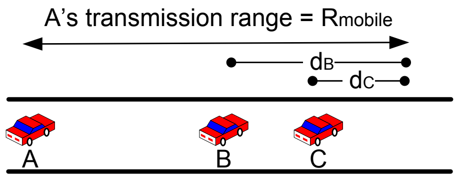


Fig. 3. Position-guided group broadcast. When vehicles B and C receive A's broadcast of data i, both initiate a local timer that is proportional to their distance to the edge of A's transmission range, d_B and d_C.

C. Summary of Observations

The above analysis leads to the following observations:

- Both data size and lifetime are major factors affecting the required infostation density.
- Because the velocity of vehicles is in general much lower than wireless signal propagation speed, it has minimal impact on dissemination performance.
- Under the same bandwidth/range assumption, the infostation density of pull is always higher than that in the corresponding push scheme.

Though the above simplistic analytical models help us understand the general trends, we expect the actual performance of the dissemination schemes to be quite different in reality. Therefore, in the next two sections we design and evaluate realistic dissemination schemes that can work with commodity WiFi.

IV. Making It Work in Practice

Both dissemination models rely on multi-hop data dissemination to cover a large number of vehicles. Effective multi-hop data dissemination in real world systems, however, faces several challenges, including wireless interference, unreliable transmissions, and queuing delays at nodes. In this section, we describe several challenges faced in a realistic wireless setting, and the techniques we proposed to address them and improve transmission bandwidth/range. Because our focus is not on detailed performance optimizations, we focus only on high level techniques that have significant impact on the overall required density of infostations. Some of the proposed techniques are inspired by existing works, and our contribution is to integrate necessary techniques together to enable a reasonable evaluation of the infostation density.

A. Position-guided Group Broadcast

Our first challenge is how to reliably transmit data between two moving vehicles or a vehicle and a nearby infostation. We choose a group-based broadcast model for data transmission rather than a point-to-point routing approach. There are two reasons for this design decision. First, the high rates of mobility in a vehicular network prevent efficient discovery and maintenance of multi-hop routes. Traditional techniques such as AODV [17] or DSR [8] would require frequent network-wide broadcasts as routes quickly degrade and are restored. Second, infostations will often disseminate highly popular data desired by the majority of nearby vehicles, *e.g.* real-time traffic conditions. Such communication naturally lends itself to a broadcast communication model.

Traditional MAC-level broadcasts are clearly insufficient for our application scenario. Network-wide broadcasts would incur extremely high per-message overhead, and uncoordinated broadcast messages would result in significant contention and packet loss. Since there are no reliability mechanisms at the MAC layer, data dissemination would be highly unreliable and data would have extremely limited transmission ranges, thus requiring artificially high density of infostations for effective network coverage.

Position-guided group broadcast. We propose a position-based broadcast that exploits the uni-dimensional nature of vehicular traffic. We assume that wirelessly-connected vehicles can obtain their geographic locations either from widely available GPS devices or through wireless triangulation on cellular networks. A node determines the data dissemination direction based on its relative location to the message source, and broadcasts the message. When a node receives a packet from a neighbor (previous hop), it starts a *propagation timer* whose expiration time is linearly proportional to the number of cars between itself and the maximum transmission range of the previous hop. If it does not hear any retransmissions of this message before its timer expires, it retransmits the data forward. The node cancels its timer if it hears someone else's retransmission † before the timer expires. We illustrate the process in Figure 3, where vehicle C's propagation timer goes off first, since it is the vehicle closest to the edge of the transmission range. Intuitively, this allows the nodes farthest from the previous hop to retransmit first while suppressing redundant messages from closer nodes. This can be seen as a form of *directional gossip*, and is similar in principle to techniques used in the Trickle code dissemination system [12], ExOR [4] and [9], [5].

Limitations. While generally effective, this *directional gossip* method has limitations in practice. First, retransmission at each hop relies on the nodes' ability to receive transmissions from nearby neighbors, and suppressing their own transmissions to reduce interference. However, a node cannot always predict the actions of a neighbor farther down the dissemination path using the propagation timer. For example, wireless interference at a farther node might force it to back off its transmission. But an intermediate node might incorrectly assume that no farther node has received the packet and retransmit the packet itself when its timer expires. This results in redundant transmissions and interference. This is a form of the hidden terminal problem, and cannot be addressed using explicit acknowledgments because they themselves can be lost to interference.

Second, our technique requires each node to know the approximate positions of its neighbors, in order to estimate the value for its propagation timer. While this is trivial for ideal scenarios with uniform distribution of vehicles, it is a challenge for real world scenarios. In practice, we expect nodes to periodically broadcast their location and velocity data to their neighbors. Such data is also required to support widely-proposed traffic safety applications, and can either be broadcasted on a separate control channel or embedded into regular data packets. For our experiments, we evaluate two schemes, an *Oracle* scheme where nodes know the relative positions of other nodes in their neighborhood, and a *Naive* scheme where each node predicts the positions of its neighbors based on an estimate of vehicle density in the entire network.

B. Eliminating Redundant Traffic

Minimizing wireless interference is crucial to improving performance of our data dissemination framework. This is especially relevant for traffic in the pull and hybrid dissemination schemes,

†Only transmissions from nodes closer to the destination than the current node cancel its propagation timer

where multiple pull requests must contend with each other and reply data for available bandwidth. Therefore, we must eliminate excess traffic when possible. First, each message (pull request, pull reply or push message) has an inherent direction. We embed the source and destination addresses and locations inside packet headers to suppress forwarding in the opposite direction. Second, since pull requests and responses have the infostation or requester as a destination, once a pull message reaches its destination (estimated by the destination location in the header), further transmissions are dropped. This *destination-awareness* helps to minimize unnecessary network traffic, thus also reducing congestion and interference.

C. Decoupling Control with Explicit Probes

Our position-guided broadcasts implicitly use data packets for media contention, i.e. nodes listen for neighbor transmissions to determine if they need to forward the data. However, this leads to performance issues in the presence of large data packets. Large packets take longer to transmit and receive, especially if fragmentation and reassembly occur at the wireless MAC layer. Because nodes are only notified of a message when the entire message has been received, timers can expire at nodes listening for neighbor transmissions because a neighbor has not finished transmitting a large message. This results in multiple nodes transmitting the same data and causing interference. The presence of variable sized messages means that propagation timers are often inaccurate, either set too low, expiring before larger messages could be received or set too high, unnecessarily reducing propagation performance.

We address this issue by decoupling data messages from channel contention messages. We introduce an *explicit probe* message that is broadcast by a node to indicate its desire to forward the data packet. The probe packet is small, but includes the hash of its associated data packet. Neighbors receiving this packet understand a neighbor is about to transmit, and cancel their own timers. Because these probes are small and fixed in size, we can accurately set propagation timers to account for their transmission time.

D. Reducing Collisions Between Pull Requests and Replies

In the pull model, there is high probability of collision between the small pull requests and the large data packets traveling in opposite directions. To minimize these collisions, nodes receiving the probe packets set a *busy* time stamp and avoid transmitting any packets as long as this time stamp is valid. The time stamp is valid for the duration of the ongoing data packet transmission. Additionally, every node stores the pull requests it has forwarded in its local buffer. A pull packet is dropped from the buffer of a node if it gets an indirect ack from another node closer to the infostation or the buffer timer expires, whichever comes first. When a node receives a data packet, it forwards any buffered pull packets after giving enough time for the data packet to pass by. The idea is to retransmit those pull packets that might have collided with an oncoming data packet transmission.

While this *pull buffering* technique improves the performance of pull dissemination, it does not achieve optimal results in a real multi-hop network. This is because the interference and backoff timers at the MAC layer may lead a node to make incorrect assumptions about the transmissions in the neighborhood. For example, a node receiving a long data transmission might still experience interference from a nearby pull packet transmission though its busy time stamp is set. Our simulations in Section V-B quantify this effect.

V. QualNet Simulations and Results

We perform realistic evaluation of the dissemination schemes through detailed QualNet simulations that take into account real MAC protocols, wireless interference, queuing delays and propagation effects. Using techniques described in the previous section, we implemented the push and pull schemes as application-level modules in QualNet. We use Constant Bit Rate (CBR) traffic on top of an 802.11 MAC layer. In addition to quantifying performance in a realistic wireless network, we hope to use these experiments to answer three key questions. 1) Do our techniques perform well on realistic vehicular traffic patterns? 2) Can we improve dissemination reliability by increasing infostation deployment density? 3) How is dissemination performance affected by incomplete adoption of our system (less than 100% penetration ratio)?

For all QualNet simulations, we perform our experiments on three separate vehicle traces:

Uniform. Vehicles are uniformly distributed with a spacing of 60m between consecutive vehicles and a uniform constant velocity of 30 m/s ($\approx$67 miles/hr).

Exponential. We introduce vehicles into the network with distance between consecutive vehicles defined as an exponential distribution bounded by the vehicle transmission range. Average distance between consecutive vehicles is 60m and vehicles travel at a uniform constant velocity of 30 m/s. Bounding the inter-vehicle distance minimizes the probability of disconnection in the network.

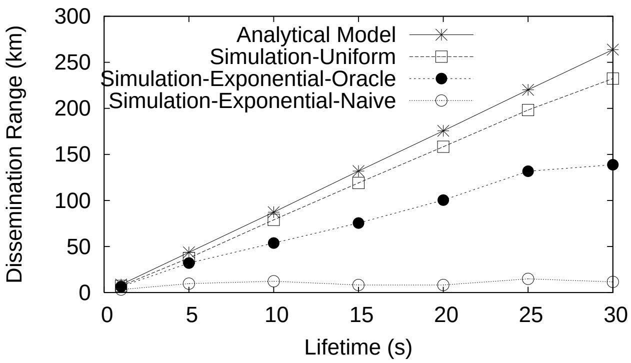
Real Measurements. We downloaded a publicly available highway vehicle trace from NGSIM [2]. This trace consists of data collected on a 2100 ft segment of the 5-lane Southbound US-101 highway over a period of 45 minutes on June 25^{th}, 2005. Vehicle positions were obtained by cross referencing data points from eight synchronized digital video cameras mounted on a 36-story building adjacent to the freeway. Vehicle density in this trace is high, with an average distance of 20m between consecutive vehicles in the same lane. Since the trace itself does not have sufficient data points, we build a complex 21-state Markov model using statistics from the trace, then use the model to generate arbitrarily long synthetic traces while preserving the statistical properties of the original trace. Details on the trace modeling and synthesis are beyond the scope of this paper.

We set both transmission ranges of infostations and vehicles (R_{info},R_{mobile}) to 485m, transmission bandwidth for both infostations and vehicles (B_{info}, B_{mobile}) to 1 Mbps, size of all data objects (S) to 12500 Bytes, and size of pull requests (S_{pull}) to 200 Bytes. Applying these parameters into the analytical results of Section III, we derive the theoretical infostation density. Assuming there is no contention/interference among vehicles, this analytical result (marked as "analytical model") represents the upper bound on the performance.

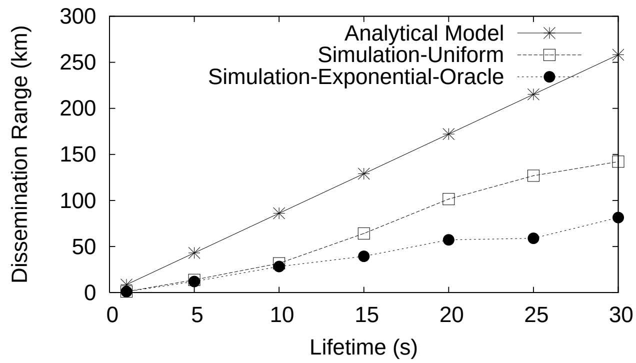
A. Performance of the Push and Pull Schemes

We begin our evaluation by first examining the performance of Push and Pull schemes. Our key metrics are the effective dissemination range, which determines the necessary infostation density, and the reception percentage, ratio of vehicles in the dissemination range that receive their desired data.

Dissemination Range. In these experiments we measure the dissemination ranges of push and pull as data lifetime increases. From Fig. 4(a), we note that both Uniform and Exponential with Oracle perform well, while the Exponential-Naive suffers heavy degradation.

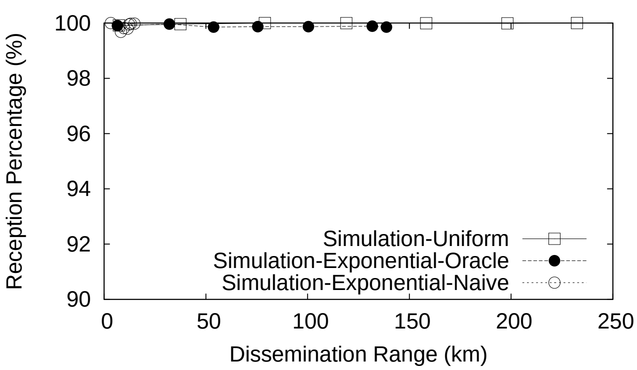


(a) The Push Scheme

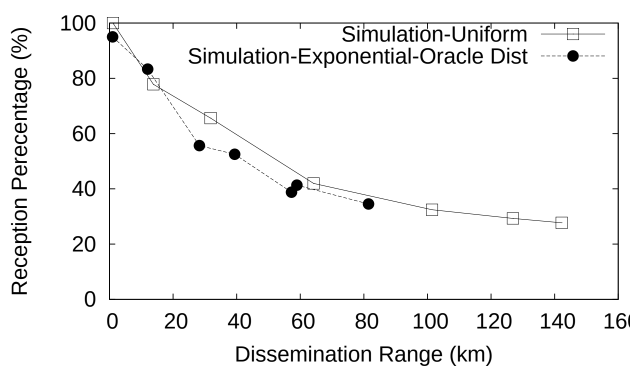


(b) The Pull Scheme

Fig. 4. Dissemination Range (Controlled Traces): Figures 4(a) & 4(b) show that uniform distribution performs better than exponential because the effective transmission range at each hop is likely larger for uniform. Naive performs badly compared to Oracle due to the contention from redundant transmissions resulting from inaccurate propagation timers. The pull scheme has lower dissemination range than the push scheme due to contention among pull requests.



(a) The Push Scheme



(b) The Pull Scheme

Fig. 5. Reception Percentage (Controlled Traces): In the push scheme nearly 100% of the expected cars get their data. In the pull scheme, the reception percentage decreases with increase in the length of road because the number of requests contending for the available bandwidth increases leading to interference and collisions.

The poor performance of Exponential-Naive can be attributed to the fact that, without location information of nearby vehicles, vehicles' propagation timers can be highly inaccurate, resulting in redundant transmissions of the same data. The ensuing contention and interference cause data collisions and propagation delays, resulting in low dissemination ranges. Uniform performs better than Exponential with Oracle because the effective transmission range at each intermediate hop is likely to be higher for uniform than exponential distribution.

In the pull scheme, increasing data lifetimes also increases the segment of the road that an infostation is responsible for, thus increasing the number of incoming pull requests[‡] Figure 4(b) shows that pull follows the same trend as the push scheme albeit with lower values for dissemination ranges. This is because our analytical model for pull assumes no wireless interference between different pull requests, but in reality wireless interference heavily affects the performance. This interference lowers effective bandwidth in the system, and contention leads to increased transmission delays. In contrast, the push scheme has minimum interference since there is no cross traffic at any point in time.

Reception Percentage. Since dissemination range is calculated based on the farthest car reached by data from the infostation, it does not quantify the number of vehicles in the range that actually receive their data. To quantify this, we calculate the *reception percentage*, the percentage of the vehicles expecting data that actually receive their data. In the push scheme, all vehicles in the dissemination range expect data. In the pull model, only those that sent pull requests expect data.

Figure 5(a) plots the reception percentage for the push experiment we did in Figure 4(a). We can see that the reception percentage is nearly 100%, thus validating the effectiveness of the dissemination techniques. For the pull scheme, we generate requests from the observed dissemination ranges of Figure 4(b). We see from Figure 5(b) that the reception percentage decreases with longer segments of the road. Since pull request density per road length is constant, longer roads mean more incoming requests and increased contention.

To investigate the impact of pull requests further, we ran a simulation varying the number of pull requests in the network. Specifically, we fix the lifetime of the data to 30 seconds and vary the number of pull requests from 1 to 100 on a fixed length of road. Figure 6 shows that the reception percentage decreases exponentially as the number of requests increases in the network. This is from increased requests contending for the same bandwidth leading to collisions and large delays.

[‡]Unless specified otherwise, the number of pull requests is directly proportional to the length of the road (1 request/2km) in all our pull experiments.

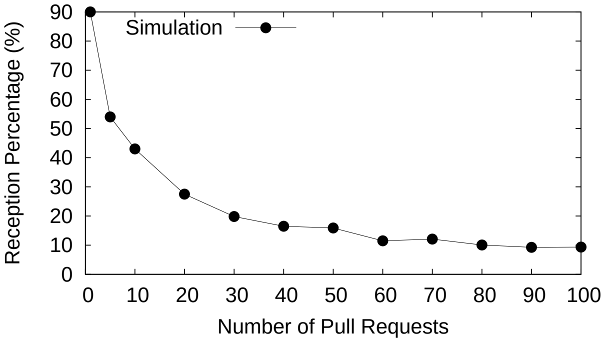


Fig. 6. Effect of varying Pull Requests (Uniform Trace): The reception percentage drops exponentially as number of requests increases. This results from increased contention between the pull requests and responses, leading to collisions and large delays.

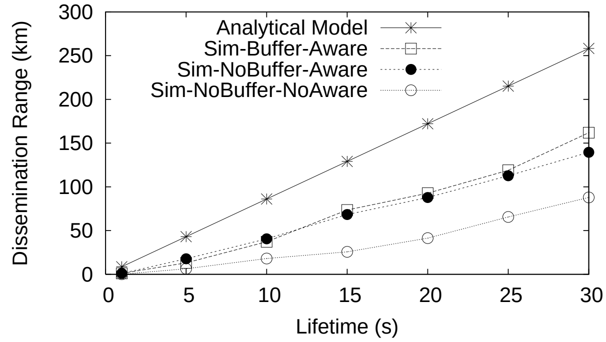
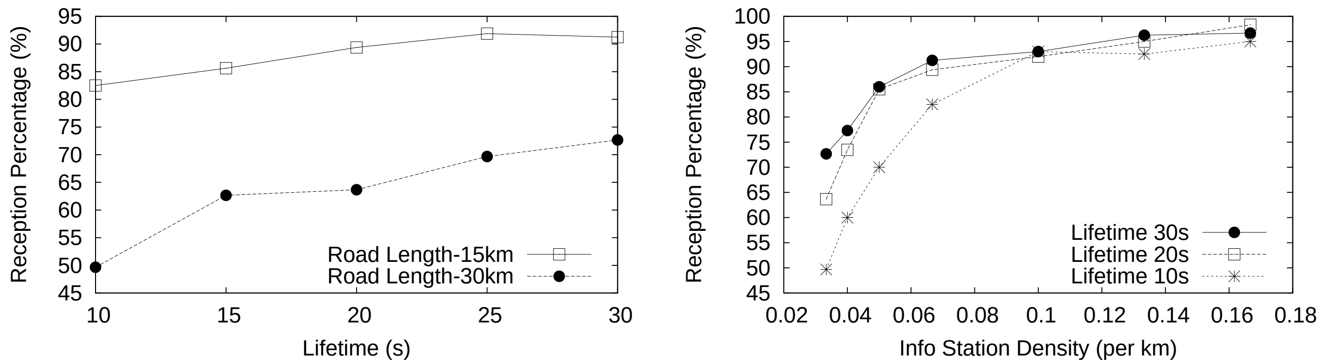


Fig. 7. Performance of various dissemination techniques for the Pull scheme (Uniform Trace). Buffer and destination awareness improve the performance of dissemination by reducing congestion and retransmitting potentially collided pull requests.



(a) Increasing Lifetime

(b) Increasing Info Station Density

Fig. 8. The Pull Scheme (Exponential Trace): Reception Percentage with increasing lifetime and increasing infostation density. Figure 8(a) demonstrates that we can improve the reception percentage by increasing the lifetime because the data has more time to reach the destination. Orthogonally, Figure 8(b) shows that as info station density increases reception percentage improves because there are fewer requests contending for an infostation and hence less delay and collisions.

B. Impact of Dissemination Optimizations

To quantify the effectiveness of our optimizations discussed in Section IV, we measure the impact of *destination-awareness* and *pull buffering* techniques on dissemination ranges in the pull scheme. Vehicles are uniformly distributed. Figure 7 plots the dissemination ranges of the pull scheme using different techniques. Our results show that as expected, utilizing both techniques achieves the best dissemination performance. While *destination-awareness* reduces unnecessary traffic, *pull buffer* retransmits potentially collided pull requests allowing them to reach the infostation. The highest gain was achieved by reducing traffic using destination-awareness. In comparison, pull buffering produced less benefits in the presence of wireless interference. Specifically, a node that did not receive an (indirect) acknowledgment cannot be sure if a request is lost or if the sender backed off due to interference, and might incorrectly retransmit causing congestion.

C. Improving Pull Performance

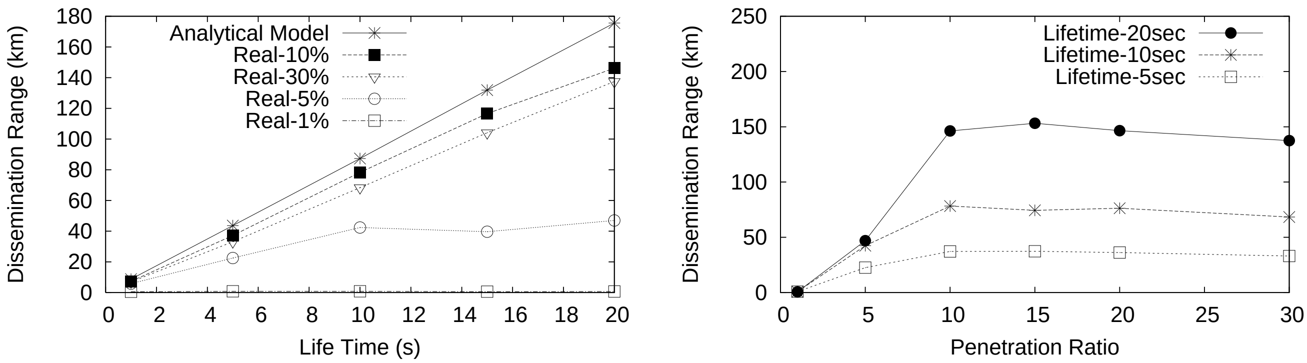
Our earlier results show that reception percentage is poor for pull scenarios with multiple pull requests. We propose two solutions to improve the reception percentage. First, we can increase the lifetime of the data for a given dissemination range, so that data packets have more time to reach their destinations. We experiment with 2 traces: 1) a 30km exponential trace which corresponds to 10s lifetime (From Figure 4(b)), and 2) a 15km exponential trace with 10secs lifetime, which essentially doubles the infostation density. We see from Figure 8(a) that by increasing data lifetime, we have improved reception percentage from 50% to about 73% for the 30km trace and 82% to 92% for the 15km trace.

Alternatively, by increasing the infostation density, the number of pull requests contending for service from any single info station can be reduced. The reduction in interference will lead to an improved reception percentage. To validate this, we experiment with data of lifetimes 10s, 20s and 30s and vary the distance between consecutive infostations from 30km to 5km. The vehicles are exponentially distributed. Figure 8(b) shows that reception percentages for all 3 lifetimes increase with increase in infostation density.

D. Impact of Penetration Ratio

As vehicular networks are initially deployed, the number of vehicles equipped with wireless devices that participate in a particular network will be low. This penetration ratio (active participants as a ratio of all vehicles on the road) plays a crucial part in the performance of dissemination schemes. To understand its impact, we randomly equip x% of vehicles in our real trace with wireless radios and observe the impact on dissemination ranges of the push and pull schemes.

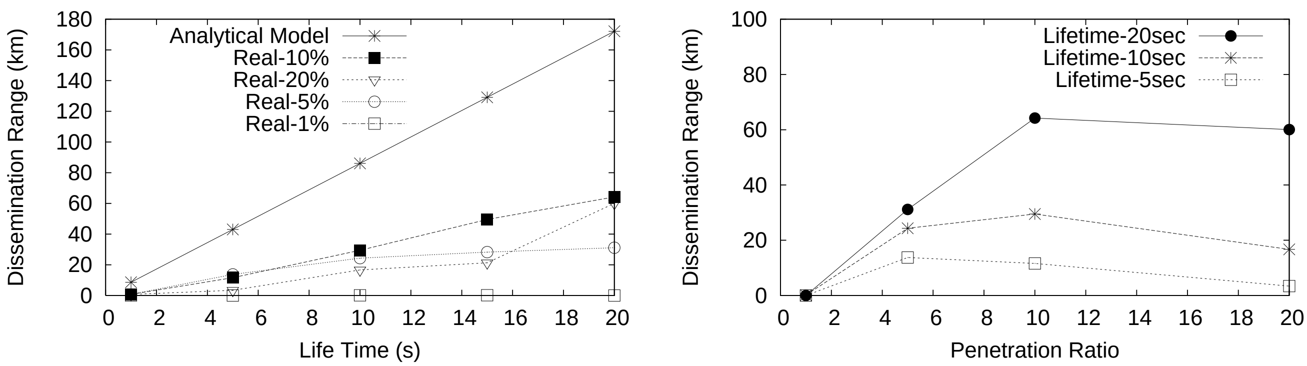
Figure 9(a) plots the push dissemination range of various penetration ratios with increasing lifetime. When the penetration ratio is



(a) Impact of Penetration Ratio with varying Lifetime

(b) Impact of Lifetime with varying Penetration Ratio

Fig. 9. The Push Scheme (Real Trace): Impact of Penetration Ratio. As the penetration ratio increases from 1% to 10% dissemination performance improves because of increased connectivity. Further increase in penetration ratio leads to a slight decrease in the performance because more vehicles are contending to forward the data, thus causing transmission delays.



(a) Impact of Penetration Ratio with varying Lifetime

(b) Impact of Lifetime with varying Penetration Ratio

Fig. 10. The Pull Scheme (Real Trace): Impact of Penetration Ratio. Dissemination range increases as penetration ratio increases from 1-10%. Further increase in penetration ratio leads to increased contention among the equipped vehicles which results in a decrease in the dissemination performance.

5% or less, disconnections in the network lead to low dissemination ranges. With 10% penetration ratio, there are enough participant vehicles to maintain network connectivity, thereby producing dissemination ranges similar to those predicted by our model. Interestingly, with 30% penetration ratio the dissemination range slightly decreases compared to 10%. This is because an increasing number of participants starts to produce contention among neighbors at each hop, leading to transmission delays. Figure 9(b) plots the dissemination range with increasing penetration ratio. We can see that past 10% penetration ratio, dissemination performance slowly starts decreasing due to increased contention among equipped vehicles.

Similarly for the pull scenario, Figure 10(a) shows that the optimal dissemination performance can be observed at a penetration ratio of 10%. Further increase in penetration ratio leads to a decrease in the dissemination range due to increased contention. Figure 10(b) shows that for all the 3 lifetimes, 10s, 20s and 30s the dissemination performance increases as penetration ratio increases from 1-10%, but gradually starts decreasing with a further increase in the penetration ratio.

VI. Conclusion

To evaluate the feasibility of infostation-based data dissemination, we seek to answer two key questions. Given an application's requirements for data lifetime, how densely must we deploy infostations, and what fraction of the vehicles need to be participants in our network? Our study shows that dissemination using the push model provides the best results. With a penetration ratio of only 10%, the system can support data with lifetime of 15 seconds or longer using infostations spaced 100km apart, For the same scenario, the pull model requires a much higher density as expected, with infostations spaced approximately every 10 to 20km. Finally, our results support the hypothesis that vehicle-to-vehicle distribution combined with intermittent connectivity to infostations does indeed provide a suitable data dissemination system for a vehicular network, at least in highway scenarios.

References

[1] QualNet. http://scalable-networks.com.

[2] NGSIM HomePage. http://ngsim.fhwa.dot.gov.

[3] Balasubramanian, A., et al. Web search from a bus. In *CHANTS* (Sept. 2007).

[4] Biswas, S., and Morris, R. Exor: opportunistic multi-hop routing for wireless networks. In *Proc. of SIGCOMM* (2005).

[5] Briesemeister, L., Schafers, L., and Hommel, G. Disseminating messages among highly mobile hosts based on inter-vehicle communication. *Proc. of IEEE IV* (2000).

[6] Bychkovsky, V., et al. A measurement study of vehicular internet access using unplanned 802.11 networks. In *Proc. of MobiCom* (2006).

[7] Chen, Z. D., Kung, H. T., and Vlah, D. Ad hoc relay wireless networks over moving vehicles on highways. In *Proc. of MobiHoc* (2001).

[8] Johnson, D., and Maltz, D. Dynamic source routing in ad hoc wireless networks. In *Ad Hoc Networking* (2001).

[9] KORKMAZ, G., EKICI, E., ÖZGÜNER, F., AND ÜMIT ÖZGÜNER. Urban multi-hop broadcast protocol for inter-vehicle communication systems. In *Proc. of VANET* (2004).

[10] LEE, K., ET AL. First experience with cartorrent in a real vehicular ad hoc network testbed. In *Proc. of MOVE* (2007).

[11] LEE, U., PARK, J.-S., AMIR, E., AND GERLA, M. Fleanet: A virtual market place on vehicular networks. *Proc. of V2VCOM* (July 2006).

[12] LEVIS, P., PATEL, N., CULLER, D., AND SHENKER, S. Trickle: A self-regulating algorithm for code propagation and maintenance in wireless sensor networks. In *Proc. of NSDI* (March 2004).

[13] LITTLE, T., AND AGARWAL, A. An information propagation scheme for vanets. In *IEEE ITSC* (2005).

[14] LOCHERT, C., SCHEUERMANN, B., CALISKAN, M., AND MAUVE, M. The feasibility of information dissemination in vehicular ad-hoc networks. In *Proc. of WONS* (Jan. 2007).

[15] MAHAJAN, R., ZAHORJAN, J., AND ZILL, B. Understanding wifi-based connectivity from moving vehicle. In *Proc. of IMC* (2007).

[16] NADEEM, T., SHANKAR, P., AND IFTODE, L. Comparative study of data dissemination models for vanets. In *Proc. of Mobiquitous* (July 2006).

[17] PERKINS, C., AND ROYER, E. Ad hoc on-demand distance vector routing. In *Proc. of WMCSA* (Feb. 1999).

[18] TORRENT-MORENO, M. Inter-vehicle communications: assessing information dissemination under safety constraints. In *Proc. of WONS* (2007).

[19] WISCHHOF, L., EBNER, A., AND ROHLING, H. Information dissemination in self-organizing intervehicle networks. *IEEE Trans. on ITS 6*, 1 (March 2005), 90–101.

[20] WU, H., FUJIMOTO, R., AND RILEY, G. Analytical models for information propagation in vehicle-to-vehicle networks. In *Proc. of VTC Fall* (2004).

[21] ZHANG, X., ET AL. Study of a bus-based disruption tolerant network: Mobility modeling and impact on routing. In *Proc. of MobiCom* (Sept. 2007).

[22] ZHAO, J., ZHANG, Y., AND CAO, G. Data pouring and buffering on the road: A new data dissemination paradigm for vehicular ad hoc networks. *IEEE Trans. on VT* (Nov. 2007).