Congestion Control Based on Channel Occupancy in Vehicular Broadcast Networks

Yaser P. Fallah, ChingLing Huang, Raja Sengupta
Institute of Transportation Studies, University of California-Berkeley, USA
Email: yaserp@eecs.berkeley.edu

Hariharan Krishnan
Electrical and Controls Integration Laboratory
General Motors R&D Center
Warren, MI, USA

Abstract— Cooperative vehicle safety (CVS) systems rely on vehicular ad-hoc networks operating in broadcast mode to deliver vehicle tracking and safety information to neighboring cars. This information is used to enable collision avoidance and warning systems. One of the main challenges of the eventual large scale deployment of such systems is network congestion, which could critically degrade the quality of a CVS system. In this paper, we present a method for congestion monitoring and control based on limited feedback from the network. We study the relationship between channel occupancy, as a readily available feedback measure, controllable network parameters, and network performance. We describe a performance measure relevant to CVS systems and present a congestion control method, based on channel occupancy measurements, that robustly maintains the system performance near optimal operation points. We examine the convergence properties of the congestion control algorithm and provide guidelines for the design of such systems based on network and traffic density conditions. Through simulation experiments we show significant gains in performance when closed loop congestion control methods are applied.

Keywords-VANET; vehicle safety; congestion control; broadcast

I. INTRODUCTION

Wireless vehicular networks are expected to revolutionize intelligent transportation systems. One of the most important applications of vehicular networks is Cooperative Vehicle Safety (CVS) [1]. In CVS systems, vehicles broadcast their tracking (e.g., position) and safety event (e.g., hard braking) information to other vehicles in their neighborhood over a wireless channel (Figure 1). Vehicles receiving this information process the received data to form a real-time map of their neighborhood and continuously watch for possible threats. The standards for CVS communication are based on Dedicated Short Range Communications (DSRC) and WAVE [2][3][4]. Current recommendations of [5] suggest that CVS messages be sent at a rate of 10 msg/sec and to distances of at least 150m, and typically to 250m. The DSRC channel set aside for this purpose is a single 10Mhz channel, which is available at pre-assigned times [4] and only used for safety messages.

While at the time of this writing CVS systems are still in prototype stage, their large scale deployment seems inevitable and is on the agenda of industry and government (e.g., IntelliDrive program). One of the main challenges of large scale deployment of these systems is network congestion which may lead to excessive tracking inaccuracy and eventually failure of CVS. Congestion is mainly due to the large amount of periodic vehicle tracking messages, which make up the bulk of safety messages. These messages include vehicle position and speed information, amongst other measures.

In this paper, we present a method for monitoring congestion in a vehicular ad-hoc network (VANET) dedicated to CVS systems, and propose a scheme for controlling congestion with the aim of maximizing system performance. For this purpose, we analyze he VANETs of CVS systems and derive the relationship between channel occupancy ratio which is a limited feedback measure, system performance (in terms of information dissemination rate), and controllable parameters of the system such as rate or range of message broadcast. These controllable parameters are currently pre-set in most prototype designs, using the recommendations of the vehicle safety communication consortium (VSCC) [5].

II. SYSTEM MODEL

The methods presented in this paper complement the existing work on improving the performance of CVS systems in particular, and Vehicular Ad-Hoc Networks in general.

In recent years, and following the debut of DSRC technology, several research efforts have focused on dealing with the issues of network congestion in DSRC based systems. For example, the work in [12] proposes to fairly allocate transmission power across all cars in a max-min fashion. This helps reduce the load at every point of a formulated highway, assuming a predefined maximum target load. Thus bandwidth is reserved for higher priority emergency messages. In [6], model based estimators are used to reduce the frequency of message broadcast while maintaining the same level of performance. This method provides compression at the source and independent of the network performance. In addition to these works, in our previous work in [7] we presented a rate control method based on perceived tracking error and channel occupancy. In [8], a joint range and rate control mechanisms has been presented based on the same features. The work in [8] presented a tunable feedback control scheme which could be adjusted to operate robustly in different traffic scenarios.

In this paper, we examine how network performance can be inferred from observing the local feedback measure of “channel occupancy”, and investigate how controllable...
parameters of a network can be adapted accordingly. Based on this analysis an enhanced congestion control mechanism is presented. The algorithm presented here does not rely on power-distance mappings as our previous work in [8] does, and can directly control the transmission power.

II. NETWORK PERFORMANCE AND CHANNEL OCCUPANCY

Congestion due to the high amount of tracking messages is the main cause of network performance degradation for CVS. To understand how congestion can be most efficiently controlled, we need to examine the effect of controllable parameters of CVS on its performance. In a CVS system, we can control some parameters such as rate (frequency) of message broadcast, $r$, and range of message broadcast, $d$. Other parameters such as contention window size can also be adapted, but have more limited effect. $r$ and $d$ are two parameters that are directly related to CVS performance. As a result, almost all enhancements for vehicular ad-hoc networks of CVS systems have focused on adjusting these parameters.

To quantify the effect of these parameters on the overall performance of the system, we need to relate the performance of the broadcast VANET to the performance of the tracking application. For this purpose, a model describing the accuracy of the tracking process at different neighboring nodes as a function of system parameters is needed. Designing such a model requires modeling the interaction of the estimation processes in the application layer with the complex multiple access scheme of VANETs. Considering the significant effect of hidden nodes and complexities it creates for network modeling, such an integrated model has proven to be intractable so far. As a result, we have to use an intermediate performance measure from the network that corresponds to CVS performance. Given the well known fact that a higher rate of message arrival at an estimation process will result in more accurate predictions [9][10], it is plausible to rely on a corresponding network performance as a substitute of the more complex overall performance measure. In this case, the network performance measure should consider the fact that in a CVS broadcast network both rate of packet reception at receiver vehicles, and number of receivers are important (i.e., it serves the objective of CVS to include more neighbors in the transmission range, up to some maximum range, e.g. 300m).

A straightforward representation of this performance measure is the broadcast packet delivery rate, defined here as the information dissemination rate (IDR). IDR represents the number of copies of a packet delivered per unit time from a single vehicle to its neighbors up to a given distance $d_{\text{max}}$:

$$
IDR = \sum_{i \in \text{Set of Neighbors}} r_{\text{succ}}(i)
$$

(1)

Where $P_{\text{succ}}(i)$ is the probability of successful packet delivery to vehicle $i$. This definition can be modified by including a weighting function that gives a higher weight to closer vehicles (if for safety purpose a CVS design assumes nearer cars are more important). Here we focus on normal IDR for clarity of discussion. Weighted IDR is similarly treated.

IDR is expected to increase with the increase in transmission rate or range, if such increases do not lead to congestion and increase in packet drop. However, the increase in rate or range significantly affects the performance of the VANET MAC layer, which follows the CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance) protocol. In addition to packet loss due to CSMA/CA collisions, the effect of hidden node interference is also a significant factor in a VANET. In particular, when the near 1-D topology of a highway is considered, hidden node collision becomes the dominant source of packet loss [8].

To quantify how different parameters such as $r$ and $d$ affect IDR, we have conducted an extensive set of simulation experiments. The simulation results presented here are also used for verification of a mathematical model that we have developed for CVS VANETs and reported in [13]; due to space limitation the mathematical model could not be included here. Nevertheless, for the sake of completeness we present the simulation results instead of the mathematical model. We note that the existing CSMA/CA models are not accurate in the case of CVS (e.g., highway topology VANET) due to the high number of interfering hidden nodes.

The simulations have been performed using an event driven 802.11 MAC simulator developed at UCB (for static topology and initial study) and using OPNET (for mobile nodes and final evaluation) considering the 802.11p default parameters [3], with the bitrate of 3Mbps (similar to 6Mbps with 50% duty cycle of 1609). We have considered a 3Km long stretch of a highway (results collected from the middle 1Km of the topology to avoid edge effects). Nodes were placed randomly at intervals uniformly distributed in a $(0,2\rho)$ range, where $\rho$ is the traffic density in vehicle/meter. While in the first set of experiments node locations were fixed (to more precisely observe the MAC performance in relation to $r$ and $d$ parameters), in later experiments we used mobile nodes based on trajectories generated by SHIFT traffic simulator (developed at California PATH, UC Berkeley).

In the first set of experiments we observe IDR, and plot it versus the choice of transmission range $d$ for different values of $r$. We show simulation results for two typical values for road density $\rho$ in Figure 2. As it is observed in these figures, the maximum level of IDR is reached at a different value of range $d$, for any given choice of $r$ and $\rho$. This fact further emphasizes the need for adaptation of rate or range of transmission. Another interesting observation from these simulations is that the maximum value of IDR is the same for all curves. This observation leads us to believe that maximization of IDR can be achieved by adjusting one of the parameters of $r$ or $d$, if the constraints allow. Such an observation relaxes the problem to a one dimensional optimization problem if the other parameter is in an acceptable range.

![Figure 2 IDR vs. range of transmission d for different r and p](image-url)
A. Channel Occupancy as Feedback Measure

In the next step, we observe the channel occupancy ratio, denoted as \( u \), and plot the achieved IDR versus \( u \) for different choices of \( r \), \( d \), and different road densities \( \rho \). Results are depicted in Figure 3. Channel occupancy ratio is measured by computing the ratio of time that the channel is sensed busy in a given window of time \( T \) (in the order of hundreds of milliseconds to seconds). If channel status at each minislot of duration \( T_{slice} \) is measured, \( u \) is found as follows:

\[
 u = \frac{\sum_{i=1}^{[T/T_{slice}]} \Lambda_i}{[T/T_{slice}]} \tag{2}
\]

where \( \Lambda \) is 1 for busy minislots and 0 otherwise, and \([T/T_{slice}]\) is the number of minislots in the measurement window.

Interestingly, it is observed that the resulting IDR for all different choices of parameters fall on the same dome shaped curve (thus not separable in the figure). The shape of the curve is thus not a function of the parameters \( r \), \( d \), or \( \rho \). We have observed in our modeling work in [13] that the shape is a result of the type of protocol (CSMA/CA, Aloha, etc.) and the type of interference (pure CSMA/CA or hidden node affected CSMA/CA), as well as MAC and PHY parameters. For the case of highway topology VANET, the MAC layer is based on CSMA/CA and is heavily affected by hidden node interference; the resulting IDR vs. \( u \) curve is depicted in Figure 3.

![Figure 3 IDR vs. channel occupancy for different values of \( r \) (5-115 msg/sec), \( d \) (20-400m), and \( \rho \) (0.1-0.2 vehicle/m). Points belonging to the same experiment with different values of \( d \) are connected by dotted line and different colors; although due to overlap they are indistinguishable.](image)

The implication of this observation is that channel occupancy \( u \) can be used as a feedback measure to evaluate network performance in terms of IDR. For example, we infer that maximum IDR is achieved when the observed channel occupancy is near \( u^* \approx 0.7 \), regardless of the combination of parameters. This suggests that channel occupancy can be used as an indicator of the system performance. Since channel occupancy cannot be controlled directly, we need to investigate the relationship between channel occupancy ratio and controllable parameters of the system (such as \( d \) and \( r \)), in order to devise methods for controlling channel occupancy and consequently IDR. For this purpose and as an example, we plot the relationship between \( u \) and \( d \), for different road density (\( \rho \)) scenarios in Figure 4 (\( r \) = 10 in this figure, increasing \( r \) will make the curves rise faster, similar to the effect of increasing \( \rho \)). It is seen that the optimal value of channel occupancy is achieved at different values of \( d \), motivating adaptive control.

While modeling the CVS network behavior is a complex task described in our other works [13], it is possible here to use empirical data and derive approximate models that describe the relationship between the parameters of concern in this experiment. For example, it is observed that the shape of the simulation result curves in Figure 4 has a component of exponential form. This can also be intuitively associated to the fact that packets are arriving according to a Poisson process (exponential interarrival), and that heavy hidden node interference adds strong Aloha like behavior to the CSMA/CA MAC. These two factors will contribute to the exponential forms seen from simulations. The exponent is expected to be a function of the load of the network, which, in turn is a product of the number of nodes and \( r \), number of nodes in a domain is also a product of \( \rho \) and \( d \). Thus, we propose a form like the following:

\[
 u = g(r, d, \rho) = 1 - \exp(-C(d - d_o)(\rho - \rho_o)(r - r_o)) \tag{3}
\]

where \( C \), \( \rho_o \), \( r_o \), and \( d_0 \) are constants, dependent on protocol parameters. Using curve fitting, we derive the parameters for the above form and find channel occupancy \( u \) as:

\[
 u = 1 - \exp\left(\frac{-(d - d_o)(\rho - \rho_o)(r - r_o)}{500}\right) \tag{4}
\]

Note that the above expression is derived by curve fitting which has to be redone if protocol settings (e.g., contention window size, packet size, etc.) change. The results from this model are compared with simulation results in Figure 4, which show that the model is relatively accurate. Nevertheless, we do not require high accuracy from the model since we only need it for stability test of range adaptation algorithms in the next section, where large design margins will be considered.

III. Congestion Control Based on Channel Occupancy

Depending on how much information about the network status and behavior of other nodes is available, different control mechanisms can be employed. Obviously, if the network topology and density \( \rho \) are known, a precise parametric network model (e.g. [13]) could be used to find optimal pairs of \( r \) and \( d \), or if one of \( r \) or \( d \) is set through other methods and is known, the other parameter can be calculated using the model (open loop control).

The issue with open loop control, as described above, is that traffic density, \( \rho \), is usually known and changes over time. The rate or range of transmission of other nodes, though could be synchronized and predicted through protocols, may be determined through other means and be unknown in most cases (e.g., [6][7][8]). Fortunately, the findings of previous section suggests that even if these parameters are unknown, maximizing IDR performance can still be done by trying to achieve the optimal value of channel occupancy by controlling at least one of the controllable parameters \( d \) or \( r \). Thus, we could use closed loop control schemes and observe channel...
occupancy, then adjust only one of the parameters $r$ or $d$ (or both if possible) to achieve optimal IDR. For example, if optimal channel occupancy is $u^*$, and rate of transmission is preset (as in [5]) or changes according to methods such as [6][7][8], and the road density $\rho$ is also changing, we can adapt the range of transmission, $d$, using gradient descent methods as follows to achieve optimal IDR:

$$d_{k+1} = d_k + \eta (u^* - u_k)$$

(5)

In the above equation, $d_k$ denotes the value for range at $k^{th}$ update, and $u_k$ is the resulting measured channel occupancy. The time interval between updates is $T$, and $u$ is calculated as in (2). Clearly, when the algorithm converges the value of $d$ will be such that $u = u^*$ and maximum IDR is achieved. However, for the above algorithm to converge, the value of $\eta$ has to be selected in a way that: 1) $d$ converges quickly to near optimal value, before the value of $\rho$ or average value of $r$ changes considerably 2) the system does not overshoot too much or oscillate and stay in a region that yields significantly low value of IDR (e.g. $u>0.95$ or $u<0.3$ in Figure 3).

To study how $\eta$ could be set, we note that the system that is being controlled is described by one of the curves in Figure 4, depending on the value of $\rho$ or $r$. Finding the precise mathematical model that describes this relationship (plant model) is discussed in our upcoming publications [13], but for the current paper the approximate model of equation (4) is sufficient (given that the precise relationship is not required by the closed loop controller). Substituting this model in (5) leads to a non-linear difference equation which cannot be easily analyzed. However, we can redesign (5), using the concept of feedback linearization [14], and create a controller that turns the feedback compensated system into a linear one. This can be achieved by the following update equation:

$$d_{k+1} = d_k + \eta \ln\left(\frac{1-u_k}{1-u^*}\right)$$

(6)

Analyzing the stability conditions to find $\eta$ can be done in a straightforward way by substituting (4) in (6) and taking a $z$-transform. To simplify expressions, we rewrite (4) as follows, if $r$ and $\rho$ are considered fixed for the duration of adaptation of $d$ (thus, only one of the curves describes the network):

$$u_k = 1-e^{-\alpha(d_k-\sigma)}$$

(7)

Here, all the constants in the exponent are represented by $\alpha$ and $\sigma$. Using (7) in (6) and taking a $z$-transform we get:

$$zD(z) = D(z) - \alpha \eta D(z) + \eta \ln(\alpha \sigma/(1-u^*))$$

(8)

Given that the last term on right is a constant, the transfer function of the feedback controlled system will have this form: $H(z)=C/(z-1+\eta\alpha)$, with $C$ representing the constants, and has only one pole at $z=1-\eta\alpha$. To ensure the stability and convergence of the feedback controlled system the pole must be located within the unit circle and we have:

$$|1-\eta\alpha| < 1 \Rightarrow 0<\eta<2/\alpha$$

(9)

The above describes bounds on the value of $\eta$. The value of $\alpha$ depends on $r$ and $\rho$, and changes with the change of these variables, thus we need to consider the maximum possible value $\alpha$ to set $\eta$, to ensure the algorithm remains robust when values of $r$ and $\rho$ change. To find a good value within these bounds that ensures fast enough convergence, while avoiding excessive overshoot, we investigate the performance of the adaptive system in the next subsection.

A. Convergence Requirements

To find requirements on convergence speed, we consider the fact that $\rho$ is not expected to change considerably in a $T$ interval and usually takes tens of seconds before a change in its value becomes significant (for example, it takes tens of seconds before traffic goes from free flow to congested). Therefore, it is safe to assume that $\rho$ remains almost constant in a window of less than 5 seconds. If $r$ is not fixed (e.g. to 10Hz as in [5]), the average rate of $r$ is also determined by vehicle dynamics which on average changes at the same rate as $\rho$ [6]. Given that a good estimate of $u$ requires $T$ to be in the order of hundreds of milliseconds (e.g. 300-500 msec), with a typical $T=0.5\text{sec}$, $\eta$ should be selected so that the algorithm converges to the optimal value in less than 10 steps. Typically, higher number of steps is also fine, but we consider the worst case here.

To examine the system convergence performance, we use MATLAB and simulate a case where the value of $\rho$ changes at 10 iteration intervals from 0.1 to 1.0 (nodes/meter), which approximately covers all possible cases of the highway road density (0.899 1-D density is derived for an 8-lane congested highway moving at 14Mph). At each $\rho$ change instance we reset $d$ to 50m, in order to see how fast the algorithm converges from an initial state. For this test, we assume fixed $r=10$, but a variable $r$ that changes not faster than $\rho$ will yield similar results. Here we calculate $\alpha_{\text{max}}$ based on $\rho=1.0$ and $r=10$. As it is shown in Figure 5, for value of $\eta=0.99(2/\alpha_{\text{max}})$, system gets close to instability when $\rho$ reaches its highest levels, but the system is very fast in rising $u$ to near $u^*=0.7$. On the contrary, $\eta=0.10(2/\alpha_{\text{max}})$ is stable, but very slow to react to changes in $\rho$. A value of $\eta=0.75(2/\alpha_{\text{max}})$ seems to provide a good balance. We use this value in the next section for OPNET simulations.

The algorithm (with all choices of $\eta$ as above) is slow for small values of $\rho$, because we set $\eta$ to a small value based on $\alpha_{\text{max}}$, which was calculated for maximum $\rho$.

IV. PERFORMANCE EVALUATION

For OPNET simulations, we consider two more constraints for the adaptive algorithm, based on the considerations of the CVS application. Given that transmission to over 300-400 meters is not required, and a minimum range of 50-100 meters is plausible from safety viewpoint, the algorithm is enhanced with these limits, that is: if resulting $d$ from (6) is greater than $d_{\text{max}}$, it is set to $d_{\text{max}}$, and if less than $d_{\text{min}}$ it is set to $d_{\text{min}}$. $d_{k+1} = \min(d_{\text{max}}, \max(d_{k+1}, d_{\text{min}})$).

In OPNET simulations we use a ricean fading model for the physical layer, based on the empirical results in [11]. One major advantage of the method presented here is that it does...
not require knowledge of the propagation model, i.e., power-distance mappings, as our previous work in [8] did, and can adjust power through any approximate mapping of d to power and still achieve optimal channel occupancy.

The network topology considered here is a bi-directional highway, 4-lane in each direction. Vehicle trajectories are generated by SHIFT and fed to OPNET that simulates DSRC (we modified 802.11a for this purpose). Other parameters were set as in previous experiments. The scenarios tested were 2 traffic scenarios of congested highway (14mph) and light traffic (54mph), which in an 8 lane highway lead to road density of 0.89 and 0.34 nodes/meter respectively, assuming car length of 4 meters and 0.8sec of distance between vehicles.

To evaluate tracking performance we measure the tracking error (of each estimation instance at sampling rate 20Hz) of all neighboring cars in different neighborhood distance bins, and derive the error histogram (of all instances). The tracking measure is then defined as the 95% cut-off error that is the value below which 95% of error population lies. Since the safety application deals with worst case scenarios, and theoretical worst case in a wireless network is unbounded. Instead, we use the 95% cut off error as a statistical substitute.

Figure 6 Channel Occupancy versus time and location X (along the road) of all vehicles (14mph case), to show convergence, system is reset at 10 sec

Figure 7 OPNET and SHIFT simulation results for congested and light traffic scenarios (14mph and 54mph): adaptive vs. fixed range method

OPNET and SHIFT simulation results are shown in Figure 6 and Figure 7. We expect the adaptive algorithm to adjust the range and operate at optimal channel occupancy; the evolution of channel occupancy for the case of 14mph is shown Figure 6. From Figure 7 we observe that in all network conditions the adaptive method manages to keep the tracking error at much lower levels than the fixed power design. It is also seen that as network becomes more crowded (14mph), the adaptive algorithm does not provide service to farther nodes (see distances above 120m in Figure 7. This is in fact a desired behavior from CVS perspective, since in crowded networks it is desired that the tracking accuracy of nearer neighbors be maintained at the cost of farther nodes which are several vehicles away and less of a safety concern.

I. CONCLUSIONS

Channel occupancy can be used as a readily available feedback measure to infer the performance of a broadcast VANET. In this paper, we presented the relationship between channel occupancy and controllable parameters of a CVS system, and demonstrated how closed loop congestion control can be achieved by using channel occupancy as a feedback measure. Simulation results demonstrate the effectiveness of using adaptive methods to control congestion and improve the performance of applications like CVS.

The method presented in this paper is inherently robust and does not rely on the knowledge of the propagation model, road density or rate of transmission of other nodes. Nevertheless, the adaptation algorithm can be made faster, if such knowledge can be extracted or estimated from locally available information.

REFERENCES


