Characterizing Urban Vehicle-to-Vehicle Communications for Reliable Safety Applications

Feng Lyu, *Member, IEEE*, Hongzi Zhu, *Member, IEEE*, Nan Cheng, *Member, IEEE*, Haibo Zhou, *Senior Member, IEEE*, Wenchao Xu, Minglu Li, *Senior Member, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*— The IEEE 802.11p-based dedicated short range communication (DSRC) is essential to enhance driving safety and improve road efficiency by enabling rapid cooperative message exchanging. However, there is a lack of good understanding on the DSRC performance in urban environments for vehicle-to-vehicle (V2V) communications, which impedes its reliable and efficient application. In this paper, we first conduct intensive data analytics on V2V performance, based on a large amount of real-world DSRC communications trace collected in Shanghai city, and obtain several key insights as follows. First, among many context factors, the non-line-of-sight (NLoS) link condition is the major factor degrading V2V performance. Second, the durations of line-of-sight (LoS) and NLoS transmission conditions follow power law distributions, which indicate that the probability of experiencing long LoS/NLoS conditions both could be high. Third, the packet inter-reception (PIR) time distribution follows an exponential distribution in the LoS conditions but a power law in the NLoS conditions, which means that the consecutive packet reception failures rarely appear in the LoS conditions but can constantly appear in the NLoS conditions. Based on these findings, we propose a context-aware reliable beaconing scheme, called *CoBe*, to enhance the broadcast reliability for safety applications. The CoBe is a fully distributed scheme, in which a vehicle first detects the link condition with each of its neighbors by machine learning algorithms, then exchanges such link condition information with its neighbors, and finally selects the minimal number of helper vehicles to rebroadcast its beacons to those neighbors in bad link condition. To analyze and evaluate the *CoBe* performance, a two-state Markov chain is devised to model beaconing behaviors. The extensive trace-driven simulations are conducted to demonstrate the efficacy of *CoBe*.**

***Index Terms*— IEEE 802.11p, DSRC, non-line-of-sight, V2V communication, reliable safety beacons, machine learning.**

I. Introduction

DRIVING safety has been the first priority for people' daily commute as a large number of traffic accidents happen every year. As indicated in the most recent report of U.S. Department of Transportation, in U.S. of 2016, there were an estimated 37,461 people were killed and 3,144,000 people were injured in police-reported 7,277,000 traffic crashes [2]. In addition, autonomous driving has been paid tremendous efforts in recent years, while two recent deaths involved by autonomous vehicles from Tesla and Uber have also raised the debate on safety, which may threaten to significantly delay of the technology adoption [3]. For drivers and autonomous driving systems, the inability of reacting to emergency situations timely is the major reason leading to the traffic crashes. Building cooperative Vehicle Safety Communication (VSC) systems [4]–[9] is a prospective solution to enhance driving safety by providing danger warnings to drivers or autonomous driving systems in advance. Meanwhile, IEEE 802.11p-based Dedicated Short Range Communication (DSRC) [10], [11] has been a standard, customized for severe-fading and highly mobile vehicular environments. Based on DSRC, V2X (broadly including vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-pedestrian (V2P), etc.) communications become the essential component to enable cooperative VSC systems. Understanding the characteristics of 802.11p-based DSRC, especially in urban environments, is of capital importance to vehicular network protocols and VSC applications.

However, to characterize the behavior of V2V[1] communications in urban environments, is very challenging for three main reasons. First, as urban environments are complex and highly dynamic, too many uncontrollable factors such as time-varying traffic conditions, various types of roads, and all different surrounding trees and buildings [12]–[14], could affect V2V link performance; it is normally hard to separate the

Manuscript received November 23, 2018; revised March 4, 2019; accepted May 17, 2019. Date of publication June 17, 2019; date of current version May 29, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61420106010, Grant 91638204, Grant 61472255, and Grant 61772340, and in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada. This paper was presented in part at the IEEE International Conference on Sensing, Communication and Networking (SECON) 2016 [1]. The Associate Editor for this paper was F. Qu. *(Corresponding author: Nan Cheng.)*

F. Lyu, W. Xu, and X. Shen are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: f2lyu@uwaterloo.ca; w74xu@uwaterloo.ca; sshen@uwaterloo.ca).

H. Zhu and M. Li are with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: hongzi@cs.sjtu.edu.cn; li-ml@cs.sjtu.edu.cn).

N. Cheng is with the School of Telecom Engineering, Xidian University, Xi'an 710071, China (e-mail: nancheng@xidian.edu.cn).

H. Zhou is with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China (e-mail: haibozhou@nju.edu.cn).

Digital Object Identifier 10.1109/TITS.2019.2920813

[1] In this paper, we mainly focus on V2V performance since it is highly related to the driving safety due to the fast speed of moving vehicles.

impact of each factor. Second, to conduct realistic studies on V2V communications in urban scenarios, experiments should involve different traffic conditions, road types, and cover a sufficiently long time, which are labor-intensive and time-consuming. The lack of real-world trace is the hurdle to achieve efficient protocols and precise modeling. Third, to well capture the link variation in the moving, various metrics should be comprehensively investigated. Limited-metric analytics not only gives one-sided communication knowledge, but also may confuse researchers and application designers without providing multi-perspective clues.

In the literature, some measurement-based DSRC studies have been carried out. Meireles *et al.* [15] and Boban *et al.* [16] focused on investigating the impact of obstacles between the communication link. They confirmed that line-of-sight (LoS) and non-line-of-sight (NLoS) conditions could deeply affect the DSRC performance, based on which, they then designed V2V propagation models with taking LoS and NLoS conditions into consideration. However, they conducted experiments by fixing two communicating vehicles, and did not investigate further when vehicles move. Similarly, physical layer measurements on DSRC channels are conducted [17]–[20], in which the characteristics of the path loss, coherence time, Doppler spectrum, etc., were investigated. All these findings could be very different when vehicles move. In our paper, we do not model a LoS/NLoS channel but focus on investing their impacts on V2V communications, and we pay little attention to physical layer features as they vary dramatically in the moving and are impossibly characterized in patterns. By collecting communication trace form moving vehicles, Bai *et al.* [21] investigated the metric of packet delivery ratio (PDR), and Martelli *et al.* [22] studied the metric of packet inter-reception (PIR) time, which refer to the probability of successfully receiving a packet and the interval of time elapsed between two successfully received packets, respectively. However, in both pieces of work, very limited metrics are evaluated; in addition, they did not discriminate between different channel conditions in terms of LoS and NLoS and drew their conclusions based on all aggregated measurements, which could bias from the ground truth. Nevertheless, there is no statistical study on the impact of channel conditions in terms of LoS and NLoS and how these two conditions interact in the moving under urban scenarios.

In this paper, we further improve our previous empirical study on urban 802.11p-based V2V communications [1]. Specifically, based on 802.11p-compatible onboard units (OBUs), we implement a V2V communication testbed and collect large volumes of beaconing traces together with their simultaneous environmental context information in Shanghai city. The data collection covers three typical road environments, i.e., urban, suburban and highway, and lasts more than two months, with a total traveling distance of over 1,500 km. Moreover, with the whole data collection recorded by cameras, we visually label out all LoS and NLoS situations for all traces. By analyzing the collected traces, we have observed that the V2V communication performs generally well, but the NLoS conditions, if encountered, may severely degrade the V2V communication performance in terms of PDR. Given the importance of NLoS conditions, we then examine the durations of LoS and NLoS conditions and find that *both LoS and NLoS durations follow a power law distribution*, which implies that not only the probability of meeting long LoS conditions is high but also the probability of seeing long NLoS conditions is also high. We further investigate the interactions between LoS and NLoS conditions by examining the distribution of PIR and packet inter-loss (PIL) times (referring to the interval of time elapsed between two dropped packets). We have two key insights as follows. First, *PIR time follows an exponential distribution in LoS conditions but a power law in NLoS conditions*. It means that consecutive packet reception failures can rarely appear in LoS conditions but can constantly appear in NLoS conditions. This is cross verified by the observation that PIL time follows a power law distribution in LoS conditions but an exponential in NLoS conditions. Second, unlike the observation that PIR time follows a power law distribution reported in work [22], *the overall PIR time distribution is actually a mix of an exponential distribution of small PIR times in LoS conditions and a power law distribution of PIR times in NLoS conditions*.

As safety applications rely on reliable beacon exchanges, we then propose a context-aware reliable beaconing strategy, called *CoBe*, to enhance the broadcast reliability when meets harsh NLoS conditions. In CoBe, the link states (LoS or NLoS) among neighbors are first detected in real time by supervised machine learning algorithms; in the payload of each beacon, in addition to application data, vehicles also include the information of their link states to its one-hop neighbors; upon identifying a NLoS condition, the sender selects *a helper vehicle* with the best link quality with both the sender and the receiver among all the optional helpers, to rebroadcast its beacons. As *CoBe* runs at the application layer and no additional environment input or cross-layer information is required, it is easy and feasible to implement in practice. To analyze the performance of *CoBe*, we devise a two-state Markov chain to model beaconing behaviors with taking LoS/NLoS channel conditions into account, based on which we derive the beaconing reliability and the corresponding cost, evaluated by the metrics of *Beacon Reception Ratio (BRR)* and *Broadcast Utility (BU)*, respectively. Beyond that, we also conduct extensive trace-driven simulations to evaluate the performance of *CoBe*. Both numerical and simulation results demonstrate the efficacy of *CoBe*, where BRR can be greatly enhanced with very slight degradation of BU. Additionally, highly analogous performance between the modeling and simulation results verify the accuracy of the proposed Markov chain model, which can be of great value for other vehicular ad hoc network (VANET) researches such as performance analysis, model establishment, and parameter tuning.

In the following, we summarize our major contributions:

- We implement a 802.11p-based V2V communication testbed and collect large-volume beaconing traces under three different urban scenarios. In addition, we have labeled LoS and NLoS conditions based watching the recorded videos and opened the data for public access[2].

[2]The labeled trace can be downloaded from our website "http://lion.sjtu.edu.cn/project/projectDetail?id=14".

Fig. 1. Illustration of an experiment car.

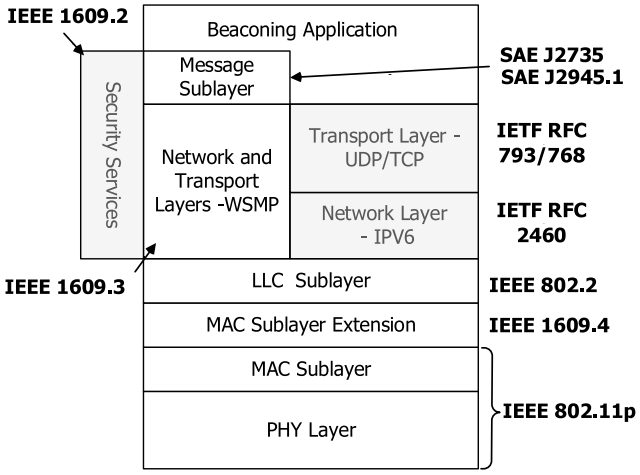


Fig. 2. Beaconing application implemented on WAVE protocol stack, where grey blocks are not involved.

- We find that 802.11p works very reliably in urban settings with a wide range of "perfect zone" (i.e., the portion of PDR larger than 80%) found, and the impact of signal power attenuation on the link performance is not obvious at least within a sufficiently long range of 500 meters, which indicate that DSRC radio is adequate to deliver safety beacons.
- In particular, LoS and NLoS channel conditions have very opposite characteristics with respect to the PIR and PIL time distributions, e.g., PIR times following an exponential distribution in LoS conditions but turning out to be a power law in NLoS conditions. Therefore, they should be distinctively treated in data analysis, modeling, protocol design, etc.
- We propose a context-aware reliable beaconing strategy, called *CoBe*, which is a fully distributed scheme and integrates three major techniques: 1) online NLoS detection; 2) link status exchange; 3) beaconing with helpers. In addition, a two-state Markov chain model is devised for performance analysis and extensive trace-driven simulations are conducted; both results demonstrate its efficacy.

The remainder of this paper is organized as follows. Section II describes the experiment platform and data-collection campaigns. We check the overall performance of 802.11p and delve into the key factor of link performance degradation in Section III. In Section IV, we further investigate the interaction of LoS and NLoS channel conditions and their impacts on 802.11p. Section V elaborates *CoBe* design. We analyze the performance of *CoBe* in Section VI. Performance evaluation is carried out in Section VII. We review the related work in Section VIII. Finally, we conclude and direct future work in Section IX.

II. Collecting V2V Trace

A. Experiment Platform Description

In this section, we introduce our V2V communication testbed and the data collection campaign. As shown in Fig. 1, the testbed includes two experimental vehicles, each of which has the following components:

1) DSRC Module: The off-the-shelf Arada LocoMateTM OBU [23] is adopted as the DSRC module and it is mounted on the roof of the experimental vehicle. In the DSRC module, IEEE 802.11p and IEEE 1609 standards are implemented for wireless access in vehicular environments (WAVE). Fig. 2 shows the WAVE protocol stack, where IEEE 802.11p serves as the physical and MAC layer to cope with fast fading and Doppler frequency shift. The DSRC radio operates in the frequency ranging from 5.700 GHz to 5.925 GHz, and supports one Control Channel (CCH) and multiple Service Channels (SCHs) with two optional bandwidths of 10 MHz and 20 MHz. For a 10 MHz and 20 MHz channel, the supported data rates range from 3 Mbps to 27 Mbps, and 6 Mbps to 54 Mbps, respectively. The transmission power can be dynamically specified with the maximum value up to 14 dBm. To achieve the most reliable V2V communication, in our experiments, we adopt the 10 MHz channel, with the lowest data rate of 3 Mbps and the maximum transmission power of 14 dBm. In addition, the DSRC OBU has a 64 MB memory, a 16 MB Flash, one 680 MHz MIPS processor (running in Linux), and one Gigabit Ethernet interface.

2) GPS Module: A high-performance GPS receiver is integrated in each OBU with an external RF antenna. The GPS receiver can help obtain the location information (including the latitude, longitude, altitude, velocity, etc.) of the experimental vehicle. Besides, the GPS modules are utilized to synchronize both OBUs every 200 ms.

3) Mobile Computer: We use a ThinkPad X240 laptop, to connect and control the OBU via its Gigabit Ethernet interface by running the `telnet` protocol. In addition, as the storage and memory of the OBU are very limited, we buffer all transmitted and received packets at the OBU temporarily, and periodically download those packets to the laptop, in order to keep the data collection program running all the time.

4) Camera Recorders: As the urban environments are highly dynamic and complex, we deploy two cameras on each vehicle with one mounted on the front window and the other fixed on the rear window, in order to record the whole data collection process for offline analysis. The time of cameras

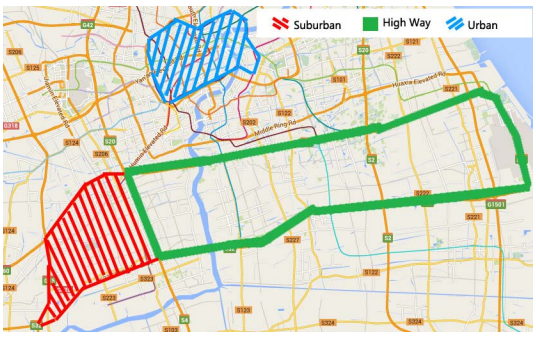


Fig. 3. Various urban environments are selected to conduct data collection.

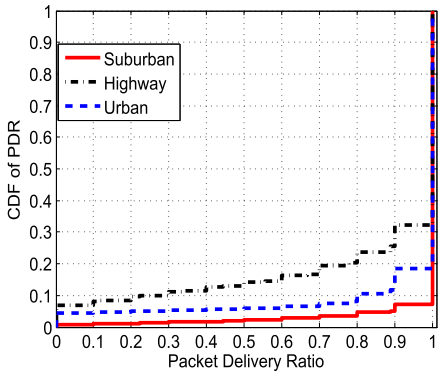


Fig. 4. CDFs of PDR over all traces.

are synchronized to the OBU within a precision of one second level.

5) V2V Beaconing Application: By adopting the Wave Short Message Protocol (WSMP), we implement a beaconing application on the WAVE protocol stack which is shown in Fig. 2. WSMP is a transport layer protocol, in which there is no retransmission or ACK mechanisms (similar to UDP). There are two programs, i.e., one *transmitter* and one *receiver*, in our beaconing application. Specifically, the receiver keeps listening the channel and the transmitter transmits a 300-byte beacon (the maximum payload of a WSMP packet reaches 1,300 bytes) every 100 ms in accordance with the driving safety requirement. Each beacon contains a sequence number as well as the latitude, longitude, altitude and speed information of the transmitter. Both the transmitter and receiver logs the beacon transmission/reception record, and by offline comparing the difference between the transmitted and received beacons, we are able to evaluate the V2V communication performance.

B. Data Collection Campaign

To cover all typical urban road conditions, we consider three major road types: 1) *urban*: roads can be unidirectional 1- or 2-lane wide and bidirectional 4- to 8-lane wide, with a large number of tunnels, overhead bridges, tall buildings, and elevated roads, as well as heavy traffic; 2) *suburban*: roads are normally bidirectional and 4- to 6-lane wide with open lands, remote houses and light traffic; 3) *highway*: bidirectional 8-lane urban freeway with a large number of walls and time-varying traffic.

We conduct our data collection campaign within areas of the above three road types in Shanghai, and the collection areas are shown in Fig. 3. For each road type, the data collection lasts for at least ten days, and in each data collection, we conduct data collection during two different time periods, i.e., rush hour (from 5:00 pm) and off-peak time (from 8:00 pm). To guarantee valid communication, during experiments, we control the distance between two communicating vehicles to be no more than 500 meters.[3] To mimic realistic driving conditions, there is no additional requirement on how the drivers drive.

[3]We implement an application running in mobile phone to calculate the distance of two vehicles by exchanging the GPS information, and report the distance to the driver every 3 seconds.

The overall campaign lasts for more than two months with an accumulated distance of over 1,500 kilometers. As a result, for each road condition, we obtain a trace, denoted as **trace** $\mathcal{U}$ (urban condition), **trace** $\mathcal{S}$ (suburban condition) and **trace** $\mathcal{H}$ (highway condition). The total amount of all traces adds up to 110 GB. In addition, we concatenate all three traces of different environments together to form a universal trace, denoted as **trace** $\mathcal{A}$.

III. Overall Urban V2V Performance Analysis

To gain an overall picture of V2V communication, we first examine the PDR performance in different urban environments. In practice, the PDR is often calculated as the ratio of the number of data packets received at the receiver to the total number of packets transmitted at the transmitter within a pre-defined time window.[4]

A. Observing Prevalent Perfect Zone

Fig. 4 shows the cumulative distribution functions (CDFs) of PDR for all traces and it can be seen that the ideal case of V2V communication could frequently happen. Ideal case means all packets are successfully received (i.e., PDR = 100%), and in urban, suburban and highway environments, it happens with the probability of 81.4%, 92.9% and 67.8%, respectively. On the contrary, in the respective environments, the probability of worst case (i.e., PDR = 0%) drops to 4.3%, 0.7% and 6.9%. It is interesting to compare our results with previous work [21] that studies communication characteristics in rural and suburban vehicular networks. From Fig. 2 of [21], the authors observed the "gray-zone phenomenon" where intermediate reception (20%≤PDR≤80%) prevails throughout the whole communication range. The probability of this happening reaches over 50.6% while the perfect reception (PDR≥80%) zone is not always guaranteed with the probability 35.2%. *Unlike their observation, we find that 802.11p performs rather reliably in urban environments and "perfect zone" prevails with a wide communication range up to 350 meters.* For instance, in urban, suburban and highway environments, the probability of perfect reception can reach above 89.6%, 95.4% and 76.2%, respectively.

[4]In this paper, we calculate PDR using a time window of one second.

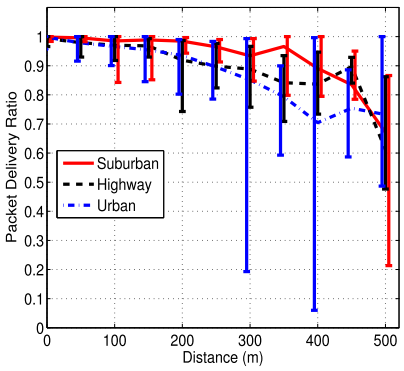


Fig. 5. PDR vs. distance between a pair of vehicles.

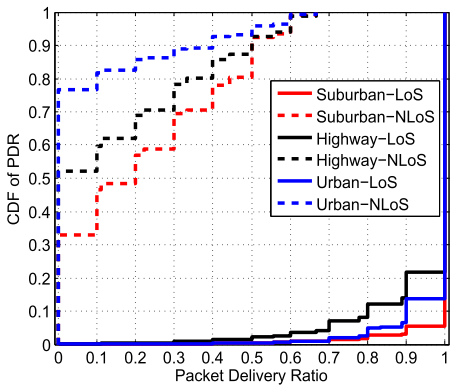


Fig. 6. CDFs of PDR in LoS/NLoS conditions.

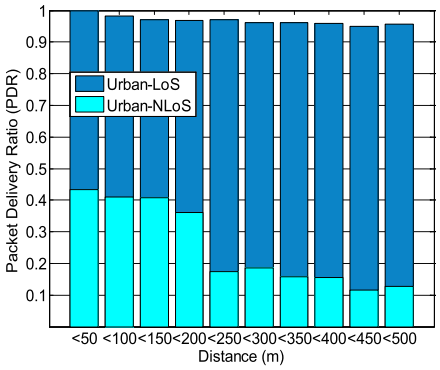


Fig. 7. PDR vs. distance under LoS/NLoS conditions.

Furthermore, from Fig. 4, we can observe that compared with the suburban environment, in urban and highway environments, multi-path fading effects are much more severe. Particularly, in the urban and highway environment, the probability of poor reception (PDR$\leq$20%) is about 5% and 9.6%, respectively, while in the suburban environment, the probability falls to 1.2%. It is reasonable as in the suburban environment, there are few vehicles or obstacles that could cause multi-path effects. On the contrary, there are a large number of mobile scatters (high-speed vehicles) and numerous stationary scatters (buildings) in the urban and highway environments, which could inject multiple paths into the channel, resulting in poor PDRs in both environments.

B. Analyzing Key Factors of Performance Degradation

To derive the key factor of performance degradation, the impact of the communication distance is first intuitively investigated and we plot the average PDR within different distance ranges, which is shown in Fig. 5. It can be seen that in all studied environments, with the distance increasing, the average PDR drops gradually. However, it is surprising to find that the PDR variation increases dramatically as the communication distance increases, especially for the urban environment. Particularly, supposed at a communication distance of 400 meters in the urban environment, the average PDR can often reach up to 100% but can also fall to below 10%. To figure out the reason for such large PDR variations, we check with the recorded videos and observe that packets are frequently lost when two vehicles are blocked by obstacles, i.e., encountering NLoS situations. To this end, based on watching videos, we then mark all NLoS situations when two vehicles cannot visually see each other,[5] and divide the original trace into LoS and NLoS two categories. In real driving scenarios, between two communicating vehicles, there may be slopes, big obstacles such as trucks and buses, and turns, which could result in NLoS situations.

1) NLoS Conditions Instead of Separation Distance Affect Link Performance Most: Fig. 6 shows the CDFs of PDR in LoS and NLoS conditions, respectively, and we can see that most packet reception failures happen under NLoS conditions. For example, in the urban, suburban and highway environments, the probability of poor reception (PDR$\leq$20%) under NLoS conditions reaches over 82.6%, 48.3% and 62.1%, respectively, and the probability of perfect reception is zero in all environments. On the contrary, in the urban, suburban and highway environments, the probability of perfect reception under LoS conditions is 93.5%, 96.9% and 86%, respectively, and the probability of poor reception is less than 1% in all environments. We can conclude that *it is NLoS conditions instead of separation distance that lead to most packet reception failures*. Although the separation distance is not the direct reason of poor PDR, it is true that the probability of encountering a NLoS condition increases as the separation distance increases, which explains the large PDR variations at long separation distances. The insight can be further verified by Fig. 7, which shows results of Fig. 5 in LoS and NLoS conditions, respectively.[6] We can see that the average PDRs in LoS conditions seem to be rather stable (all above 95%) while the average PDRs in NLoS conditions have poor performance (all below 40%) regardless of the distance variation.

IV. Interactions Between LoS and NLoS

Given the importance of NLoS conditions, we further investigate the interactions between LoS and NLoS conditions by

[5]Note that, although NLoS conditions found by cameras are not necessarily to be NLoS for RF radios, those visually NLoS conditions are still good approximations of real radio NLoS conditions and valuable for analysis.

[6]Note that, results under suburban and highway scenarios are omitted due to the similar observations and space limitation.

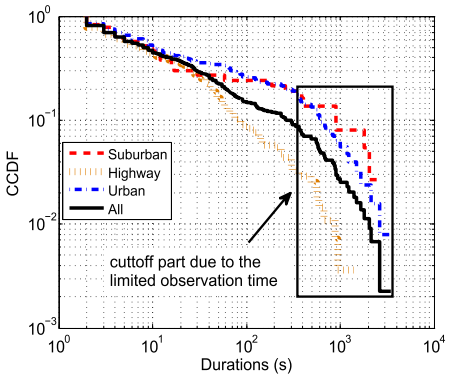


Fig. 8. CCDFs of LoS durations.

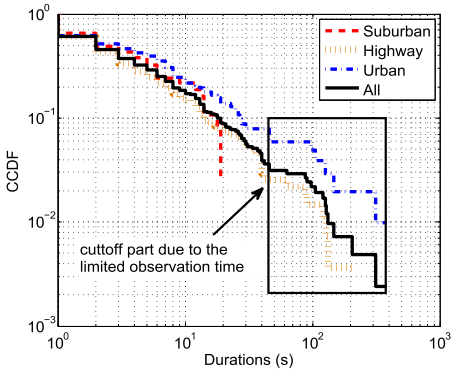


Fig. 9. CCDFs of NLoS durations.

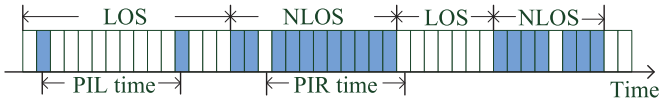


Fig. 10. An example sequence of packets, where white blocks denote successfully received packets and dark blocks denote packet reception failures.

examining the metrics of PIR and PIL times as illustrated in Fig. 10. In the figure, the PIR time refers to the interval of time elapsed between two successfully received packets, i.e., the duration between two adjacent white blocks, while the PIL time refers to the time interval elapsed between two dropped packets, i.e., the duration between two adjacent dark blocks.

A. Power Law Distributions of NLoS and LoS Durations

We first examine LoS and NLoS durations to check how they appear in real driving conditions and plot their tail distributions, which are shown in Fig. 8 and Fig. 9, respectively. Two main observations can be achieved. First, *both LoS and NLoS durations follow a power law distribution*, as linear plots in log-log scale are found in both figures. It indicates that not only the probability of meeting long LoS conditions is high but also the probability of meeting long NLoS conditions is also high. It should be noted that, the cutoff part of the tail distribution should not be considered due to the effect of limited observation duration, which has also been pointed out in previous studies on characterizing inter-contact

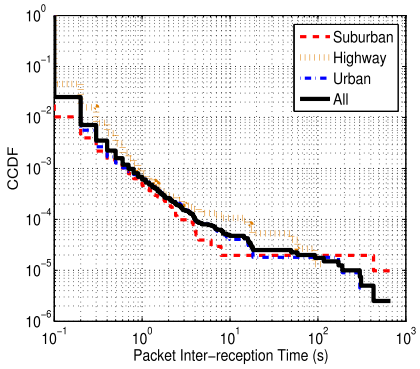


Fig. 11. CCDFs of overall PIR times in log-log scale.

time distribution of human [24] and vehicular [25] mobility. In addition, linearity tests and distribution fittings are conducted, and results are shown in Section VII. Second, LoS durations are in general longer than NLoS durations. For instance, the proportion of durations longer than ten seconds reaches 50% in LoS conditions while drops to only 18% in NLoS conditions. Nevertheless, the finding of heavy-tailed of NLoS durations is critical for the beacon-based application design, which has to cope with the relatively long and constant communication blackouts when vehicles move.

B. Mixed Distributions of PIR Times

We plot the complementary cumulative distribution functions (CCDFs) of PIR times of all traces in Fig. 11 (log-log scale). It is also interesting to compare our results with previous work [22] that studied 802.11p-based beaconing performance based on data collected during trips traveled among several Italian cities. As shown in Fig. 4 of [22], the authors observed that the CCDF of PIR times satisfies a power law (identified by linear plots in log-log scale) and had the conclusion that the PIR time distribution is heavy tailed, which means that the probability of having relatively long PIR time is relatively high. *Unlike their observation, we find that the CCDF of PIR time appears linear in log-log scale only for large PIR times and has a much faster decay for small PIR times, which implies that the PIR time only partially follows a power law.* For example, as shown in Fig. 11, it can be seen that the CCDF of PIR times is not linear when PIR time is smaller than one second.

To explain it, we can see from Fig. 6, where large proportion of PDRs are greater than 80% in LoS conditions while in NLoS conditions, very poor PDR is witnessed, indicating that small PIR times are common in LoS conditions and large PIR times normally appear in NLoS conditions. To this end, we check the distribution of PIR times in LoS and NLoS conditions, and plot their CCDF results in Fig. 12 (linear-log scale) and Fig. 13 (log-log scale), respectively. Clear linear plots are seen in both figures, which means that *PIR time in LoS conditions follows an exponential distribution and that in NLoS conditions follows a power law distribution*. It implies that short PIR times (consecutive successfully-received packets) are more likely to happen in LoS conditions whereas

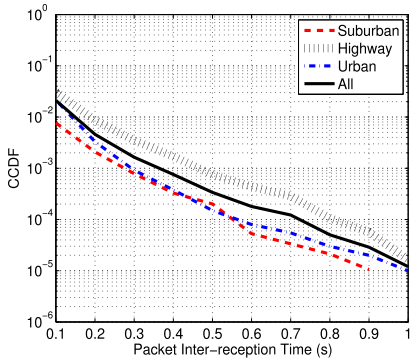


Fig. 12. CCDFs of PIR times in LoS conditions in linear-log scale.

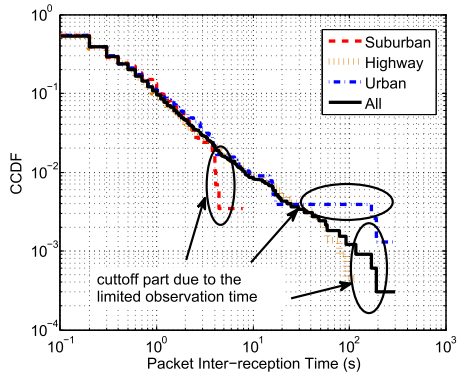


Fig. 13. CCDFs of PIR times in NLoS conditions in log-log scale.

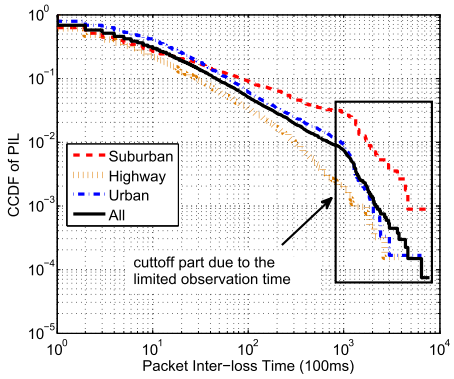


Fig. 14. CCDFs of PIL times in LoS conditions in log-log scale.

the probability of having long PIR time is relatively high under NLoS conditions. Then, we can well explain why the CCDF of overall PIR times in Fig. 11 has a much faster decay than a power law distribution when the PIR times are small. It is because the CCDF result is a combination of an exponential distribution of small PIR times in LoS conditions, and a power law distribution of PIR times in NLoS conditions. On the other hand, we have a similar but opposite observation on the distribution of PIL times. In particular, as shown in Fig. 14 and Fig. 15, we find that *PIL time in LoS conditions follows a power law distribution whereas that in NLoS conditions follows an exponential distribution*. It means that short PIL times (consecutive packet losses) are common

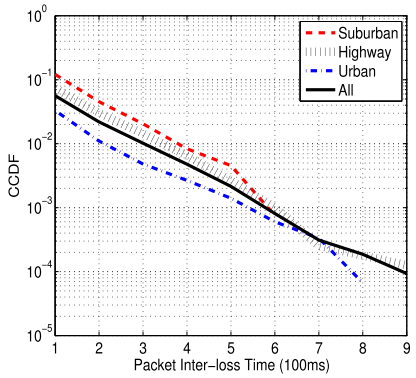


Fig. 15. CCDFs of PIL times in NLoS conditions in linear-log scale.

in NLoS conditions while in LoS conditions, relatively long PIL times are more likely to happen.

V. Context-Aware Reliable Beaconing Design

A. Overview

Reliable beaconing is an essential building block in VSC applications, where periodical "status" messages (containing information of the vehicles position, speed, acceleration, braking status, etc.) broadcasted by each vehicle should be well received by all neighbors in its vicinity. From above analysis, beacon-based VSC applications could benefit from link context information. For instance, when encounters NLoS conditions, a better beaconing strategy is to find proper neighbor vehicles to help rebroadcast the beacons, to further enhance the beaconing reliability. To this end, we propose a context-aware cooperative beaconing strategy, called *CoBe*, which can enhance the beaconing reliability when encounters harsh NLoS conditions. In essence, *CoBe* integrates three major techniques as follows: 1) online NLoS detection; 2) link status exchange; 3) beaconing with helpers.

B. Online NLoS Detection

1) Using Physical Layer Hints: With more advanced antenna (e.g., MIMO systems) and physical layer techniques (e.g., OFDM modulation), NLoS conditions can be accurately detected in real time. For example, the power delay profile [26], which profiles arriving signals from multi-path channels and gives the power strength of a received signal, can be utilized to detect NLoS conditions.

2) Perceiving NLoS at Application Layer by Machine Learning Algorithms: In cases where physical layer hints are not available, it is still possible for upper-layer applications to perceive NLoS conditions. From Fig. 6, it can be seen that all visually identified NLoS conditions have very low PDRs which are less than 70%, while LoS conditions have relatively high PDRs and 95% of them are greater than 70%. This observation indicates that PDR values have some latent relations to the underlying LoS or NLoS channel conditions. Therefore, we can adopt the application layer PDR values to infer the underlying LoS/NLoS condition. To learn those unknown relations, supervised machine learning algorithms

can be well applied since we have marked all LoS and NLoS conditions. More specifically, given the consecutive time slots (last for one second), we can extract their historical PDR values, i.e., the PDR value within previous 1, 5 and 10 seconds, as input training features. Together with their current LoS/NLoS labels (i.e., the training targets), classical machine learning algorithms such as support vector machine (SVM), k-nearest neighbor (KNN), decision tree and random forest, are leveraged to train the process and then output a binary classifier. By calculating the previous 1-second, 5-second and 10-second PDR values and inputting them into the classifier, NLoS conditions can be detected in real time. Particularly, if the classifier outputs a result 1 (i.e., meaning a LoS condition), the channel is considered to be in good condition; otherwise, a NLoS condition is found.

C. Link Status Exchange

To be compliant with the broadcast requirement for most of safety applications, each vehicle is required to broadcast every 100 ms [8], [27]. By logging the beacon reception records for each one-hop neighbor, vehicles are able to calculate the up-to-date historical PDR values of each one-hop link and then perceive their real-time link status (LoS or NLoS). Once a vehicle detects a NLoS condition for a link, the vehicle should select a helper vehicle to help rebroadcast the beacon, in order to recover the harsh link condition. To help vehicles choose an optimal neighbor in a distributed way, in each beacon, in addition to application data, vehicles should also include their one-hop link status information (0 means NLoS and 1 means LoS), in order to exchange such link status among neighbors. Specifically, for each vehicle (say vehicle x), it maintains a circular recording queue for each one-hop link and calculates the up-to-date historical PDRs to get the link status. It then includes the information of (neighbor vehicle ID, LoS/NLoS flag) in each beacon and broadcasts out. In this way, by receiving beacons, vehicle x is able to understand the link status between its one-hop neighbor (say vehicle y) and the neighbors of y, i.e., seeing two-hop links, which is beneficial to helper selection.

D. Beaconing With Helpers

Upon identifying one or more NLoS links between a vehicle and its neighbors, the vehicle seeks for helper vehicles from its neighbors to rebroadcast the beacon. Such a helper would be selected if it has a LoS condition with both the sender and the receiver. In general, we define the *helper selection problem* as, in case that there are multiple NLoS links, how to select helpers so that all NLoS links are covered and the number of helpers are minimized at the same time? The following theorem can be concluded.

Theorem 1: The helper selection problem is NP-hard.

Proof: To prove its NP-hardness, we devise a polynomial reduction from a classic NP-hard problem, i.e., *Max k-cover* [28], to the helper selection problem. In the *Max k-cover* problem, there is a collection of subsets, $F = \{S_1, S_2, \ldots, S_i\}$, each of which is a set of n points; the objective is to select k subsets from F to maximize the total number of points contained in their union.

The instance of the *Max k-cover* problem can be taken as an input for reduction as follows. In specific, assuming that the link conditions between each pair of vehicles within the communication range of a sender vehicle are known, we can construct a graph, $G(N, E)$, where N is the set of nodes and E is the set of edges. In the graph, each vehicle within the communication range of the sender is a node, and there is an edge between a pair of nodes if they have a LoS links. Denote each node as n_i, and all its LoS neighbors, i.e., the nodes having a LoS link with n_i, as $S_i \subseteq N$, for $i = 1, \ldots, |N|$. With this graph, given a k, to find k different S_j for $j = 1, \ldots, k$ such that their union contains as many nodes as possible, is equivalent to the *Max k-cover* problem, which is NP-hard. The problem is then to find the smallest number of k to cover all nodes. Therefore, the helper selection problem is a NP-hard problem, which concludes the proof. ■

Given the NP-hardness of the helper selection problem, in *CoBe*, we adopt a greedy heuristic to select preferable helpers. In specific, neighbors are ranked according to the size of S_i, for $i = 1, \ldots, |N|$. The node with the largest $|S_i|$, say node n_l, is first selected as a helper. Then, for each node n_i, for $i = 1, \ldots, |N|$ and $i \neq l$, S_i is updated to remove nodes appearing in S_l, i.e., $\{n_j | n_j \in S_i \wedge n_j \notin S_l\}$ and re-ranked to select the second helper. This procedure repeats until that all nodes are covered or there is no node left.

E. Protocol Overhead Analysis

The main overhead of *CoBe* is the required link status information in each beacon, including the vehicle ID and the according LoS/NLoS flag of one-hop neighbors. Denote by N_{max} and N^2_{max} the maximum number of vehicles in one-hop set and two-hop set of one particular vehicle, respectively. To label all vehicles in the two-hop st (i.e., interference range), the short ID is devised to replace the MAC address of each vehicle, in order to decrease overhead. Vehicles decide their IDs randomly and will update it if the ID is detected already in use by another vehicle. Therefore, at least $\lceil \log_2 N^2_{max} \rceil$ bits are required, in order to label each individual vehicle with a unique short ID, and $\lceil . \rceil$ is the ceil function symbol. Hence, the maximum overhead of *CoBe* (in bits) is

$$overhead = N_{max} \times (\lceil \log_2 N^2_{max} \rceil + 1). \tag{1}$$

For a vehicle, as its one-hop set area is a circle with the radius R, the N_{max} on a road can be calculated by

$$N_{max} = (\frac{2R}{length_{vehicle} + distance_{safety}}) \times L, \tag{2}$$

where R, $length_{vehicle}$, $distance_{safety}$ and L is the valid communication range (in meters), length of a vehicle (normally reaching 5 meters for sedans), safety following distance in driving and the number of lanes on a road, respectively. Considering a normal urban case, R and L could be empirically set to be 300 and 6, respectively [1]. In addition, according to the 2 second driving rule that drivers should drive at least 2 seconds behind the preceding vehicle even under ideal

driving situations, the $distance_{safety}$ is approximately about $2 \times (60/3.6) \approx 33$ m when gives a normal urban vehicle speed of 60 km/h. Then, $N_{max} = (\frac{600}{5+33}) \times 6 \approx 95$. Given the ID length of 10 bits, which is able to label more than 1024 vehicles (rich enough for the size of N^2_{max}), the overhead then is $overhead = 95 \times 11 = 1045$ bits ≈ 131 bytes.

As beacon application data is normally no more than 500 bytes [29], it is *acceptable* to include such extra 131 bytes of coordination data in each beacon to enable the upper-layer *CoBe* functions since the overall packet size is far smaller than the maximum payload (normally above $1,400$ bytes) that the MAC layer can support. In addition, adding extra coordination information in beacons to achieve distributed protocol design is common in VANETs, e.g., designing TDMA MACs [8], [27], [30], in which each vehicle needs to broadcast the ID and the according time slot index of all one-hop neighbors, in order to negotiate the time slot usage. If these MACs are adopted, the additional overhead of *CoBe* would be only LoS/NLoS flags of one-hop neighbors, i.e., 95 bits $\approx$ 12 bytes in above analysis, which is very small and easy to implement.

VI. Performance Analysis

In this section, we devise a two-state Markov model to analyze the performance of *CoBe*, and compare it with the following two benchmark schemes:

- **Conventional 802.11p:** In broadcast mode of 802.11p protocol, vehicles broadcast beacons every 100 ms and there is no any rebroadcast mechanisms;
- **Random Forwarding:** In this scheme, a vehicle not only broadcasts beacons as described above but also randomly chooses one or more of its neighbors to rebroadcast its beacons every time.

We analyze the performance of each scheme considering the following two metrics:

- **Beacon Reception Ratio (BRR):** refers to the ratio of beacon reception calculating by the number of neighbors having received the beacon to the total number of one-hop neighbors, which is defined to evaluate the beaconing reliability;
- **Broadcast Utility (BU):** refers to the ratio of the BRR to the total number of broadcast beacons, which is defined to evaluate the rebroadcast cost. For example, if the BRR of 0.9 is achieved when no helper rebroadcasts (i.e., the beacon is broadcasted only once by the transmitter), the BU becomes 0.45 if one more helper is sought (i.e., the beacon is broadcasted twice).

A. Two-State Markov Chain Model

As the channel switches between LoS and NLoS states, a two-state Markov model can be devised as shown in Fig. 16, where the transition probabilities are P_L and P_N, respectively. The likelihood of successfully receiving a packet heavily relies on the current link state, which is P_{good} ($0 < P_{good} < 1$) when the link is in LoS state, and is P_{bad} ($0 < P_{bad} < P_{good}$) when the link is in NLoS state. To further derive variables in this model, we start by stating a known property of the two-state Markov chain:

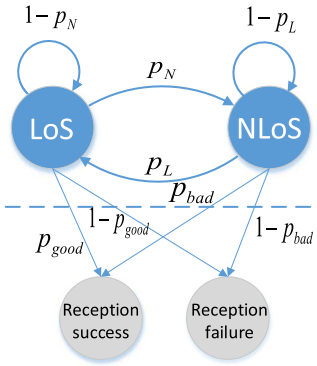


Fig. 16. Two-state Markov chain model.

Proposition 1 (see, e.g., [31]): If $0 < P_L, P_N < 1$, *the unique stationary distribution (or initial state distribution) of the two-state Markov chain in Fig. 16 is*

$$\pi = (P_L = \frac{P_L}{P_L + P_N}, P_N = \frac{P_N}{P_L + P_N}),$$

where P_L and P_N (equals to $1-P_L$) represent the stationary probabilities of the link being in state LoS and NLoS, respectively. Then we derive the probability $P(Rx)$, that the packet can be successfully received at the kth time slot, $k \geq 1$. According to the law of total probability, $P(Rx)$ can be represented as

$$\begin{aligned} P(Rx) &= P(L) \cdot P(Rx|L) + P(N) \cdot P(Rx|N) \\ &= P(L) \cdot P_{good} + P(N) \cdot P_{bad}, \end{aligned} \tag{3}$$

where $P(L)$ and $P(N)$ is the probability of being in LoS state and in NLoS state, respectively. To get the value of $P(L)$ and $P(N)$ at the kth time slot, we can consider all possible unfolding of the Markov chain during k steps. Particularly, at the ith step, $i \leq k$, we denote P_{iL} and P_{iN} as the probability of being LoS state and NLoS state, respectively. According to the unfolding rules of our Markov chain, the recursive relations can be achieved as follows

$$\begin{cases} P_{iL} = P_{(i-1)L} \cdot (1 - P_N) + P_{(i-1)N} \cdot P_L \\ P_{iN} = P_{(i-1)L} \cdot P_N + P_{(i-1)N} \cdot (1 - P_L), \end{cases} \tag{4}$$

with $P_{1L} = P_L$ and $P_{1N} = P_N$. As $P_L + P_N = 1$, from Eq. (4), we can obtain

$$\begin{cases} P_{iL} = P_L \\ P_{iN} = P_N. \end{cases} \tag{5}$$

As a result, $P(Rx)$ can be derived as follows

$$P(Rx) = P_L \cdot P_{good} + P_N \cdot P_{bad}. \tag{6}$$

B. Theoretical Results

We derive the expected BRR and BU of all schemes with respect of one sender vehicle, one single receiver vehicle, and one or multiple helper vehicles. To make the analysis tractable, we further assume that all links are *independent and identically distributed*. In specific, we denote BRR as the probability that

one packet can be well received by the receiver and BU as the ratio of BRR to the number of broadcasts.

Performance of Conventional 802.11p: As in 802.11p, there is no helper during broadcasting, BRR and BU can be easily obtained as

$$\begin{cases} BRR = P(Rx) \\ BU = P(Rx). \end{cases} \tag{7}$$

In the remainder of this subsection, we derive the results of *CoBe* and random forwarding scheme with one or multiple helpers.

a) Using One Helper: Performance of CoBe*:* To derive the BRR of *CoBe*, we divide the analysis into two parts with respective to LoS and NLoS conditions. We can get

$$\begin{aligned} BRR &= P_L \cdot P(BRR|L) + (1 - P_L) \cdot P(BRR|N) \\ &= P_L \cdot P_{good} + (1 - P_L) \cdot (1 - (1 - P_{bad}) \cdot P_{fail}), \end{aligned} \tag{8}$$

where P_{fail} means the probability that the retransmission from the helper is unsuccessful. The failure of retransmission can happen due to the two reasons: one is that the helper does not hear the notification of retransmission from the sender; the other one is that the retransmitted packet is dropped. Considering n neighbors exist around the sender and receiver, *CoBe* may choose the helper with the probabilities p_1, p_2 and p_3 in terms of link conditions of sender-to-helper and helper-to-receiver; p_1 represents the probability that both two links are in LoS conditions, p_2 is the probability that one link is in LoS while the other one is in NLoS, and p_3 is the probability of both being in NLoS conditions. They can be calculated as follows

$$\begin{cases} p_1 = 1 - (1 - P_L^2)^n \\ p_2 = 1 - p_1 - p_3 \\ p_3 = (1 - P_L)^{2n}. \end{cases} \tag{9}$$

Then, P_{fail} can be written as

$$\begin{aligned} P_{fail} = {} & p_1 \cdot (1 - P_{good}^2) + p_2 \cdot (1 - P_{good} \cdot P_{bad}) \\ & + p_3 \cdot (1 - P_{bad}^2). \end{aligned} \tag{10}$$

To investigate the cost of *CoBe*, we define the average transmitted packets by the sender and the helper as num. Apparently, BU can be written as

$$BU = \frac{BRR \cdot 1}{num} = \frac{BRR}{num}. \tag{11}$$

Similarly, we compute num in LoS and NLoS two situations, respectively. We can obtain

$$num = P_L \cdot 1 + (1 - P_L) \cdot (1 + PKT_{helper}), \tag{12}$$

where PKT_{helper} is the average transmitted packets from the helper. It can be calculated as follows

$$PKT_{helper} = p_1 \cdot P_{good} + \frac{p_2}{2} \cdot (P_{good} + P_{bad}) + p_3 \cdot P_{bad}. \tag{13}$$

Then, BU can be derived.

Performance of Random Forwarding: For fair comparison in terms of using the same number of helpers with *CoBe*, in this scheme, the sender will always randomly select a helper from its neighbors to retransmit the beacon.

To derive the BRR in this scheme, we consider two events, i.e., the helper does not receive the retransmission notification and the helper receives the notification then retransmits the packet. Then BRR can be calculated as

$$BRR = (1 - P(Rx)) \cdot P(Rx) + P(Rx) \cdot (1 - (1 - P(Rx)^2), \tag{14}$$

Similarly, we calculate the average transmitted packets to derive BU. Then BU can be easily obtained as

$$BU = \frac{BRR \cdot 1}{1 + P(Rx) \cdot 1} = \frac{BRR}{1 + P(Rx)}. \tag{15}$$

b) Using H Helpers: In fact, even with a helper, reception failures may also encounter due to the dynamic channel fading. In this case, we analyze seeking multiple helpers to expand the overlap of covering sets, which can conservatively enhance the reliability.

Performance of CoBe*:* When using H helpers, $1 \leq H \leq N$, Eq. (8) can also be satisfied. P_{fail} means the probability that the retransmissions from H helpers are unsuccessful simultaneously. It can be written as

$$P_{fail} = \Pi_{i=1}^{H} p_i, \tag{16}$$

where p_i for $1 \leq i \leq H$, is the probability of unsuccessful retransmission from the ith helper. The value of p_i depends on the link conditions from the sender to the ith helper and from the ith helper to the receiver. The two links can be in LoS-LoS, LoS-NLoS/NLoS-LoS or NLoS-NLoS three categories. We define P_{LL}, P_{NL} and P_{NN} as the probability of unsuccessful retransmission from the helper under link conditions of this three category, respectively, and thus they can be achieved as follows

$$\begin{cases} P_{LL} = 1 - P_{good}^2 \\ P_{NL} = 1 - P_{good} \cdot P_{bad} \\ P_{NN} = 1 - P_{bad}^2. \end{cases} \tag{17}$$

In addition, given two links, they can be LoS-LoS, LoS-NLoS/NLoS-LoS and NLoS-NLoS conditions with the probability P^{LL}, P^{NL} and P^{NN}, respectively. They can be obtained as

$$\begin{cases} P^{LL} = P_L^2 \\ P^{NL} = 1 - P_L^2 - (1 - P_L)^2 \\ P^{NN} = (1 - P_L)^2. \end{cases} \tag{18}$$

For H helpers, we define the event that i helpers, j helpers and $H - i - j$ helpers are with link conditions LoS-LoS, LoS-NLoS/NLoS-LoS and NLoS-NLoS, respectively; the event happens with the probability $P(i, j)$ for $i, j, H - i - j \geq 0$. Then Eq. (16) can be replaced as

$$P_{fail} = \Sigma P(i, j) \cdot P_{LL}^i \cdot P_{NL}^j \cdot P_{NN}^{(H-i-j)}, \quad i, j \geq 0, i + j \leq H \tag{19}$$

TABLE I
NLoS DETECTION RESULTS

Algorithms	Accuracy (%)			Precision (%)			Recall (%)			F-Score (%)			FPR (%)		
	$\mathcal{H}$	$\mathcal{S}$	$\mathcal{U}$	$\mathcal{H}$	$\mathcal{S}$	$\mathcal{U}$	$\mathcal{H}$	$\mathcal{S}$	$\mathcal{U}$	$\mathcal{H}$	$\mathcal{S}$	$\mathcal{U}$	$\mathcal{H}$	$\mathcal{S}$	$\mathcal{U}$
Decision Tree	90.86	97.52	97.57	94.39	98.71	98.77	93.68	98.74	98.58	94.03	98.73	98.67	6.45	1.26	1.43
Random Forest	92.08	97.91	98.08	95.22	99.04	99.16	94.37	98.83	98.77	94.79	98.93	98.96	5.73	1.18	1.24
SVM	93.67	98.31	98.28	98.11	99.65	99.63	93.93	98.67	98.51	95.94	99.16	99.07	6.56	1.35	1.51
KNN	93.0	98.15	98.21	96.6	99.31	99.4	94.34	98.82	98.67	95.45	99.07	99.03	5.9	1.19	1.35

The value $P(i, j)$ depends on the rule of helper choosing. According to *CoBe* design, $P(i, j)$ can be achieved as follows

$$P(i,j) = \begin{cases} \Sigma_{m=H}^{n} C_n^m (P^{LL})^m (1-P^{LL})^{n-m}, \\ \quad i = H; \\ \Sigma_{m=H-i}^{n-i} C_n^i C_{n-i}^m (P^{LL})^i (P^{NL})^m (P^{NN})^{n-i-m}, \\ \quad 0 \le i < H, j = H-i; \\ C_n^i C_{n-i}^j (P^{LL})^i (P^{NL})^j (P^{NN})^{n-i-j}, \\ \quad 0 \le i < H, 0 \le j < H-i. \end{cases} \tag{20}$$

Then BRR can be derived. To calculate BU in this condition, Eq. (11) and Eq. (12) can also be utilized. However, computing PKT_{helper} is different from Eq. (13), it follows

$$PKT_{helper} = \Sigma P(i,j) \cdot (i \cdot P_{good} + \frac{j}{2} \cdot (P_{good} + P_{bad}) + (N-i-j) \cdot P_{bad}). \tag{21}$$

Then, BU can be derived.

Performance of Random Forwarding: For H helpers, we define the event that i helpers for $i \in [0, H]$ receive the retransmission notification and retransmits the packet while the remaining $H-i$ helpers do not receive the retransmission notification; the event happens with a probability $P(i)$. Then, BRR can be calculated as

$$BRR = \Sigma_{i=0}^{H} P(i) \cdot (1-(1-P(Rx))^{i+1}), \quad 0 \le i \le H, \tag{22}$$

and BU can be obtained by

$$BU = \frac{BRR}{1+\Sigma_{i=0}^{H} P(i) \cdot i}, \quad 0 \le i \le H. \tag{23}$$

In addition, the probability $P(i)$ follows

$$P(i) = C_H^i \cdot P(Rx)^i \cdot (1-P(Rx))^{H-i}, \quad 0 \le i \le H. \tag{24}$$

VII. PERFORMANCE EVALUATION

In this section, we first carry out the NLoS detection results and then evaluate the performance of *CoBe* in terms of BRR and BU.

A. NLoS Detection Accuracy

We adopt the cross-validation scheme to investigate NLoS detection accuracy in different machine learning algorithms. Specifically, as there are about 16, 425, 16, 033 and 27, 439 labeled LoS/NLoS condition samples in **trace** $\mathcal{H}$, $\mathcal{S}$ and $\mathcal{U}$, respectively, for each data set, we first split them into 10 subsets. For each round, one different subset is chosen as the testing set and the other 9 subsets are aggregated as the training set. Based on the basic testing results of True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN), we evaluate the following five metrics:

- *Accuracy:* the probability that a condition is correctly identified, i.e., $\frac{\text{TP+TN}}{\text{TP+FP+TN+FN}}$;
- *Precision:* the probability that an identified NLoS conditions is correctly identified, i.e., $\frac{\text{TP}}{\text{TP+FP}}$;
- *Recall:* the probability that all NLoS conditions in ground truth are correctly identified, i.e., $\frac{\text{TP}}{\text{TP+FN}}$;
- *F-Score:* combining the precision and recall metric together, i.e., $2 \times \frac{Precision \times Recall}{Precision + Recall}$;
- *False Positive Rate (FPR):* the probability that a LoS condition is identified as a NLoS condition, i.e., $\frac{\text{FP}}{\text{FP+TN}}$.

c) Precise NLoS Detection: Fig. 17 shows the average F-Score and FPR results in respective environments, and we can easily observe that machine learning algorithms are very suitable for online NLoS detection, especially for SVM and KNN algorithms. Specifically, for the average F-Scores, the SVM algorithm achieves the result about 95.9%, 99.2% and 99.1% in highway, suburban and urban environment, respectively, and the KNN algorithm achieves the respective result about 95.5%, 99.1% and 99%, which are precise enough to intelligently react to the channel variation. In contrast, the average FPRs are very small in both algorithms; for instance, in the SVM algorithm, they are about 6.6%, 1.4% and 1.5% in highway, suburban, and urban environment, respectively. The reason of relatively higher FPRs triggered in highway environments, is that packets may accidentally dropped due to the fast speed of vehicles in highways, which are mistakenly identified as NLoS conditions. Table I shows all metric results, demonstrating that NLoS conditions can be precisely detected in real time. In addition, SVM achieves the best performance among all machine learning algorithms and in the following simulations, we adopt it as the NLoS detection algorithm for *CoBe* performance evaluation.

B. Efficiency of CoBe

In this subsection, we conduct trace-driven simulations of all candidate schemes and carry out both numerical and simulation results.

1) Synthesizing V2V Communication Trace: Even though we have collected a large volume of V2V traces for a pair of vehicles, they are still not enough for extensive trace-driven simulations when involves more vehicles in the network. As we have disclosed PIR and PIL time distributions in both

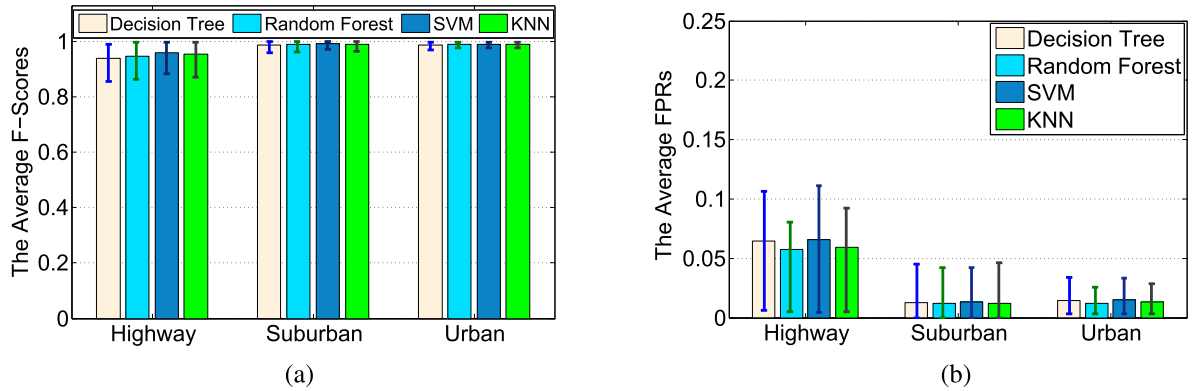


Fig. 17. Precise NLoS detection achieved by machine learning algorithms. (a) The average F-scores. (b) The average FPRs.

TABLE II
FITTING RESULTS OF LOS DURATION DISTRIBUTION IN DIFFERENT ENVIRONMENTS

Fitting Results	Suburban	Highway	Urban
Coefficient C (with 95% confidence bounds)	$1.098(1.092, 1.103)$	$1.014(1.01, 1.019)$	$1.081(1.073, 1.089)$
Coefficient a (with 95% confidence bounds)	$0.3768(0.3751, 0.3784)$	$0.3771(0.3758, 0.3785)$	$0.3293(0.327, 0.3316)$
SSE	0.0008859	0.0004907	0.001941
R-square	0.9993	0.9996	0.9985
Adjusted R-square	0.9993	0.9996	0.9985
RMSE	0.003086	0.002273	0.00467

LoS and NLoS conditions, it is possible to synthesize abundant V2V traces based on their distribution fitting results. As the distribution of LoS durations is a power law, the CCDF of LoS durations can be fitted by adopting the following general model

$$P(LoS \quad duration > x) = C \cdot x^{-a}. \tag{25}$$

Fitting metrics of SSE, RMSE, R-square and Adjusted R-square, are evaluated, and Table II shows the fitting results, in which precise fitting results demonstrate our findings in the data analytics. Similarly, the CCDF of NLoS durations can be fitted by adopting the following model

$$P(NLoS \quad duration > x) = C \cdot x^{-a}, \tag{26}$$

as NLoS duration also follows a power law distribution. Given the fitting results of a and C, LoS and NLoS durations can be generated in demand.

After that, we fit PIR and PIL time distribution in LoS and NLoS conditions, respectively. As PIR times follow an exponential distribution in LoS conditions, to fit PIR time distribution, we adopt the following model

$$P(PIR > x) = e^{-\lambda x}, \tag{27}$$

to output fitting parameters and in order to generate PIR times under a given LoS duration. Similarly, the exponential distribution model

$$P(PIL > x) = e^{-\lambda x}, \tag{28}$$

is adopted to fit PIL time distribution under a give NLoS duration. The fitting results of NLoS durations, PIR times and PIL times are omitted due to the space limitation.

Specifically, we first generate a LoS and NLoS duration list, denoted as $\mathcal{LOS} = [los_1, los_2, \ldots, los_n]$ and $\mathcal{NLOS} = [nlos_1, nlos_2, \ldots, nlos_n]$, respectively. For each los_i, $i \in \{1, 2, \ldots, n\}$, a PIR time list $\mathcal{PIR}$ is generated. Within a certain PIR time duration, the packets at the beginning and ending slots are well received while packets at other slots are lost. Similarly, for each $nlos_i$, $i \in \{1, 2, \ldots, n\}$, a PIL time list $\mathcal{PIL}$ is also generated. In contrast, within a certain PIL time duration, the packets at the beginning and ending slots are lost while packets at other slots are well received. With this scheme, we can synthesize abundant V2V trace by setting the value of n, i.e., the length of LoS and NLoS duration list.[7]

2) Using One Helper: For choosing one helper, we investigate the impact of the number of neighbors on the choosing decision.

a) Simulation Setup: With the fitting results of **trace** $\mathcal{U}$, we first synthesize 100 links of trace and each of them lasts for 1000 min. For each round of simulation, two links trace are first randomly chosen to mimic the sending/receiving process between the sender and receiver. To add a neighbor in, another two pairs of link trace are randomly chosen, one pair of which represent the sending/receiving process between the sender and neighbor and the another pair for the link between the neighbor and receiver. We range the number of neighbors n from 1 to 10 and conduct the simulation over each value of n. For other parameters, we set them according to the fitting results of **trace** $\mathcal{U}$. Particularly, P_L and P_N are set to be 0.8 and 0.2, respectively. According to the average PDR

[7]Note that, synthesizing V2V communication trace can be also valuable for other VANET researches, e.g., tuning models or validating protocol and scheme designs.

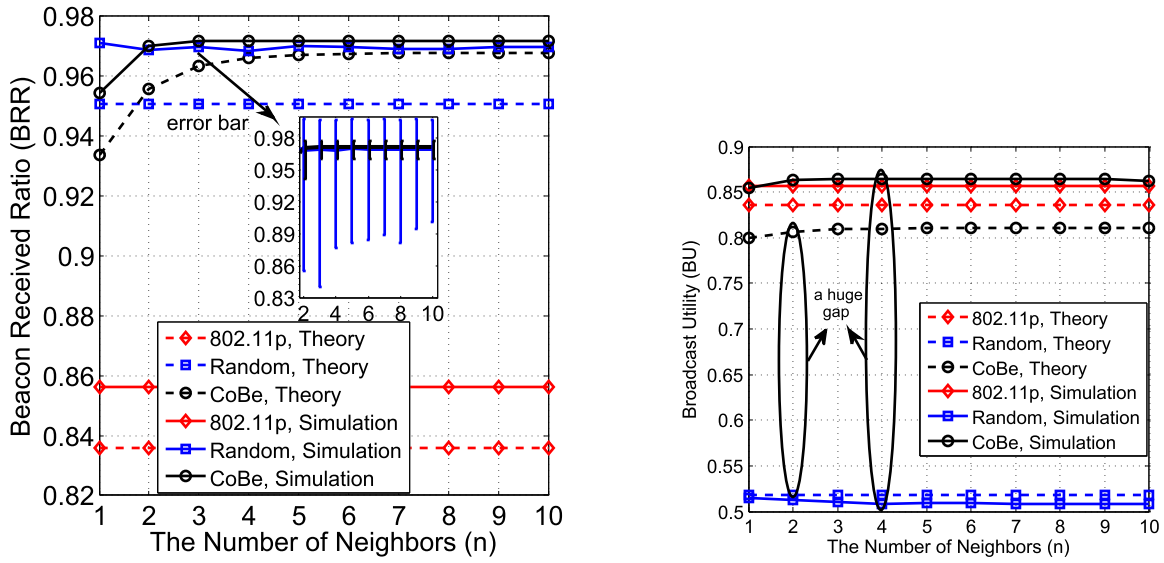


Fig. 18. The performance of using one helper. (a) BRR vs. the number of neighbors. (b) BU vs. the number of neighbors.

TABLE III
SIMULATION PARAMETERS

Parameters	Value	Parameters	Value
Channel number	178	Neighbors n	$[1, 10]$ or 10
Channel bandwidth	10 MHz	Number of links	100
Transmission power	14 dBm	Chosen links	$[6, 42]$ or 42
Data rate	3 Mbps	P_L	0.8
Environment	Urban	P_N	0.2
Number of lanes	4-8	P_{good}	0.97
Mean distance	147.796 m	P_{bad}	0.3
Mean Tx speed	5.107 m/s	Simulation time	1000 min
Mean Rx speed	5.041 m/s	Simulation rounds	30
Helpers H	1 or $[1, 10]$		

under LoS and NLoS conditions, we set the P_{good} and P_{bad} to be 0.97 and 0.3, respectively. Table III shows the detailed simulation parameters. For each simulation setup, we run the simulation for 30 rounds to achieve statistically significant results.

b) Performance Comparison: Fig. 18 (a) shows the average BRRs achieved by three beaconing strategies when chooses one helper from different number of neighbors. The dashed lines represent the numerical results and the solid lines denote the simulation results. We can observe that numerical and simulation results present highly similar trends and other two major insights can be achieved. First, with seeking a helper to rebroadcast beacons, the beaconing reliability can be greatly enhanced. Specifically, the *CoBe* and random approach always outperforms the 802.11p and can enhance its BRR from 83.6% to about 96% and 95%, and from 85% to about 97.1% and 96.8% in numerical and simulation results, respectively. Second, *CoBe* is able to achieve a better performance when the number of neighbors in the environment increases (i.e., more helper selections appear), while the 802.11p and random approach fail to react to the environment well. Particularly, in simulation results, when the number of neighbors increases, the BRRs keep constant in 802.11p and fluctuate within a small range in random approach while they increase gradually in *CoBe*. In addition, when there are more than two neighbors in the environment, the BRR in *CoBe* could reach up to 97% which outperforms the random approach of about 96.8%, and the error bar of two schemes demonstrate the stability of *CoBe*.

Fig. 18 (b) shows the average BUs and we can observe that the random approach achieves the lowest broadcast utility (below 52%) comparing with the other two strategies (more than 80%) in both numerical and simulation results. It is a huge gap and can incur extra burden to the system. Besides, it can be seen that with more neighbors in the environment, *CoBe* can achieve a better BU. Specifically, in simulation results, the BUs gradually increase and can reach more than 86.2% in *CoBe* when there are more than two neighbors, which surpass the value of 85.6% in 802.11p.

In general, compared with two benchmark strategies, *CoBe* can enhance the BRR dramatically with the slight BU degradation, especially when there are more than two neighbors in the environment.

3) Using H Helpers: For choosing H helpers, we investigate the upper-bound of reliability that proposed schemes can support and their corresponding cost.

a) Simulation Setup: The detailed simulation parameters are shown in Table II. For each round of simulation, we first choose two links trace representing the sending/receiving process between the sender and receiver; then, 10 neighbors with 20 pairs of link trace are added in. We range the number of helpers H from 1 to 10 and conduct the simulation over each value of H. For each simulation setup, we run the simulation for 30 rounds to achieve the statistically significant results.

b) Performance Comparison: Fig. 19 (a) and (b) show the average BRRs and BUs of three beaconing strategies when using H helpers. Numerical and simulation results present

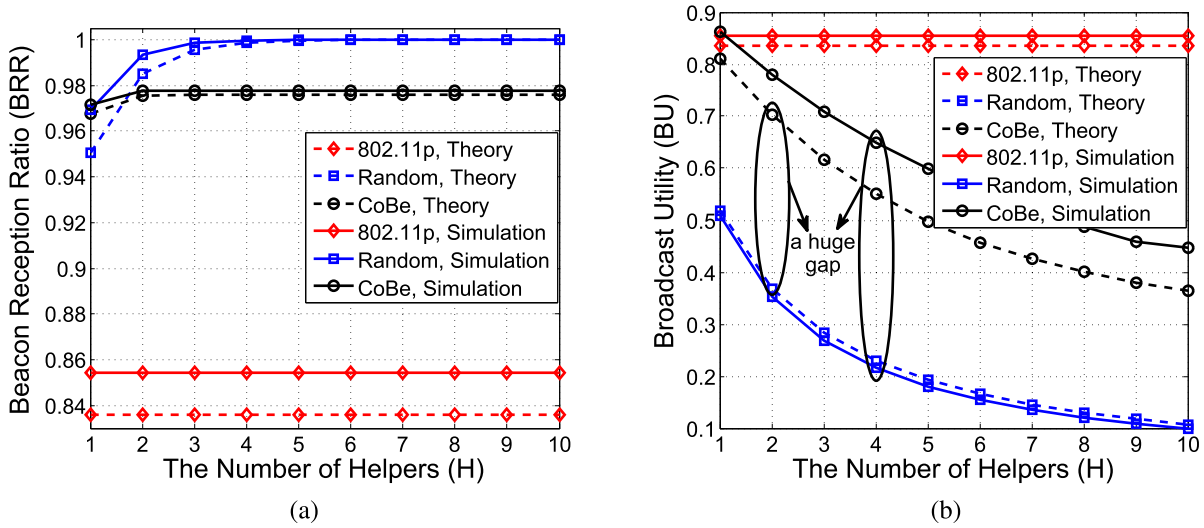


Fig. 19. The performance of using H helpers. (a) BRR vs. the number of helpers. (b) BU vs. the number of helpers.

highly similar trends and we can observe that, with more helpers, in both *CoBe* and random approach, average BRRs increase gradually while BUs decrease dramatically. Specifically, in numerical results, the average BRR increases from the value 96.8% to 97.6%, and 95.1% to 100% in *CoBe* and random approach respectively, while the BU decreases from the value 81.1% to 36.5%, and 51.8% to 10.7%, respectively. However, without any retransmission, the average BRRs and BUs keep a constant value 83.6% in 802.11p. This observation demonstrates that enhancing communications to an extreme reliability, e.g., above 97%, means a huge cost should be paid, especially for the random approach. Even though the upper-bound reliability of random approach can reach 100% when seeks more than 4 helpers, there is a huge gap of BUs compared with *CoBe*, reaching over 30%, which will aggravate the resource shortage problem in VANETs [32]–[34]. This observation indicates that when meets harsh NLoS conditions, at most 2 helpers should be sought in *CoBe* as they can cover all packet failures with an acceptable cost.

In summary, *CoBe* can be a smart strategy for beaconing enhancement with coping with harsh NLoS conditions. In most case, seeking one helper is enough as it can provide 96.8% and 97.2% of the average BRRs, 81% and 86.2% of the average BUs in numerical and simulation results, respectively.

VIII. Related Work

A. Characterizing V2V Communications

In [21], Bai *et al.* presented an extensive analysis on PDR by investigating the impacts of transmission distance and power, mobility and propagation environment on it. However, only PDR is evaluated, which is insufficient to capture the features of intermittent and time-varied V2V communications especially in complex urban scenarios. In addition to PDR, Martelli *et al.* studied the metric of PIR time in [22], and revealed the PIR time distribution being a power-law. However, they drew the conclusion based on all aggregated measurements without discriminating LoS and NLoS conditions, which could bias from the ground truth. In [35], the authors investigated the reliability of DSRC communication at application level and disclosed that the reliability is adequate since most of the time, packets do not drop in bursts, while the data analysis is not fine-grained enough to guarantee ultra-reliable V2V communications for driving safety applications. There are also some measurements on physical layer of DSRC channels [17], [18], in which the path loss, coherence time and Doppler spectrum are investigated. However, all their findings could be very different when in different LoS and NLoS link conditions. In this paper, we pay little attention to physical layer features as they vary dramatically in the moving, which can be hardly characterized in patterns. In regard to LoS and NLoS affects, Meireles *et al.* [15] conducted the experimental study and confirmed that the channel quality is heavily influenced by LoS and NLoS conditions. Specifically, they collected the PDR and together with the received signal power information in several important scenarios, based on which, the impact of obstructions are quantified, e.g., NLoS conditions effectively halving the usable communication range within which 90% of communications can be successful. This insight is valuable, however they collected data when vehicles and obstructions are statistic, and did not investigate the interaction of LoS and NLoS conditions in the moving, and thus cannot provide comprehensive knowledge about V2V communications. For LoS and NLoS modeling, Boban *et al.* [16] designed a Geometry-based Efficient propagation Model for V2V communication (GEMV2) with taking LoS and NLoS conditions into account. In our work, we do not model a LoS or NLoS channel but research on the LoS and NLoS interactions during the driving and characterize V2V performance under two distinct channel conditions.

B. Relay Schemes for Reliable Communications

1) Receiver-Oriented Schemes: In the receiver-oriented schemes, all vehicles that have received the beacon, will contend for being relays employing ether probability-based or waiting-time-based mechanisms. Wisitpongphan *et al.* [36] proposed three probability-based

schemes, i.e., slotted p-persistence, slotted 1-persistence and weighted p-persistence. They works with a similar rule, in which, if the receiver j receives a packet from the sender i for the first time, it will rebroadcast the packet with the probability p; otherwise, it will drop the packet. The difference between each scheme is the setting of value p, in 1-persistence and p-persistence scheme, the value of which is set to be 1 and a pre-determined probability, respectively. Differently, the probability p_{ij} in weighted p-persistence is set to be $\frac{D_{ij}}{R}$ where D_{ij} is the transmission distance between vehicles i and j, and R is the average communication range. Regarding waiting time-based schemes, [37] is the first waiting time-based forwarding scheme, in which each candidate relay determines their waiting time solely based on their distance d to the source vehicle, and the relay with larger d will be associated with a smaller waiting time. In addition, in the protocol of ABSM (Acknowledged Broadcast from Static to highly Mobile) [38], upon receiving a beacon, instead of retransmitting it immediately, the vehicle will wait to check if retransmissions from other neighbor vehicles already have covered its whole neighborhood.

Aforementioned relay schemes are easy to implement since algorithms are processed locally and without complex negotiation. However, as communication contexts are not taken into account, they cannot react to dynamic environments well, which would lead to the broadcast storm problem easily and cause channel resource wasting with broadcasting duplicated beacons excessively.

2) Sender-Oriented Schemes: Conversely, in the sender-oriented relay schemes, the source broadcaster explicitly selects potential vehicles to be relays, and only those vehicles that have received the beacon, are listed as potential relays. As the sender-oriented selections initially limits the number of contending relays, efficient channel bandwidth utilization can be guaranteed. However, they requires frequent updating information of neighboring vehicles and the performance heavily relies on information input, i.e., what kind of information is leveraged and how these information can be achieved in real driving scenarios. For example, Rehman *et al.* [39] proposed BDSC (Bi-directional Stable Communication), in which they investigated the relations between the estimated link qualities and the transmission distance; after that, the relay were selected based on quantitative representation of link qualities. However, this study is only evaluated by theoretical analysis while in practice, the distance is not the only factor that affect link performance. In TVR (Tall Vehicle Relaying) [40], Boban *et al.* indicated that on tall vehicles, the elevated position of antennas can improve communication performance since tall vehicles are more likely to encounter LoS conditions. Therefore, they distinguished between short and tall vehicles, and chose tall vehicles as next hop relays. The scheme normally can achieve a good performance; however it fails to work robustly when no tall vehicles exist in the scenario, or there are slopes (i.e., the road is not flat) between vehicles, which are common NLoS conditions during our data collections. Differently, in our scheme, we leverage high-frequency beacon exchange among neighbors to perceive real-time link conditions, based on which, the helper-selection strategy is designed. Compared with the current literature, it advances in two aspects: 1) it can work with an efficient channel resource utilization since rebroadcast action can be only triggered by those vehicles who encounter harsh NLoS conditions where rebroadcast is really in need; 2) the scheme solely relies on the beacon receiving records and is independent from underlying driving environments, which can be well adopted to support universal safety applications.

IX. Conclusion and Future Work

In this paper, based on real-world urban DSRC communication traces, we have presented intensive data analytics on V2V performance, characterized the V2V channel, and proposed *CoBe* to enhance the broadcast reliability by coping with harsh NLoS conditions. To evaluate its performance, a two-state Markov chain model has been devised for performance analysis and extensive trace-driven simulations have been conducted; both results demonstrate the efficacy of *CoBe*. In the future, we will exploit these analytic results to benefit other advanced vehicular techniques, such as routing, MAC design, and architecture design. In addition, we will collect more V2V communication traces with more involved vehicles, over multiple locations.

References

[1] F. Lv *et al.*, "An empirical study on urban IEEE 802.11p vehicle-to-vehicle communication," in *Proc. IEEE SECON*, Jun. 2016, pp. 1–9.

[2] NHTSA. (Sep. 2018). *2016 Summary of Motor Vehicle Crashes*. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812580

[3] *Wharton School of the University of Pennsylvania. Autonomous Car Crashes: Who—Or What—Is to Blame?*. Accessed: Nov. 2018. [Online]. Available: http://knowledge.wharton.upenn.edu/article/automated-car-accidents

[4] W. Xu *et al.*, "Internet of vehicles in big data era," *IEEE/CAA J. Automatica Sinica*, vol. 5, no. 1, pp. 19–35, Jan. 2018.

[5] X. Cheng, R. Zhang, and L. Yang, "Wireless toward the era of intelligent vehicles," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 188–202, Feb. 2019.

[6] S. Zhang, J. Chen, F. Lyu, N. Cheng, W. Shi, and X. Shen, "Vehicular communication networks in the automated driving era," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 26–32, Sep. 2018.

[7] F. Abbas, P. Fan, and Z. Khan, "A novel low-latency V2V resource allocation scheme based on cellular V2X communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2185–2197, Jun. 2019. doi: 10.1109/TITS.2018.2865173.

[8] F. Lyu, H. Zhu, H. Zhou, L. Qian, W. Xu, M. Li, and X. Shen, "MoMAC: Mobility-aware and collision-avoidance MAC for safety applications in VANETs," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10590–10602, Nov. 2018.

[9] X. Shen, X. Cheng, L. Yang, R. Zhang, and B. Jiao, "Data dissemination in VANETs: A scheduling approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2213–2223, Oct. 2014.

[10] *Standard Specification for Telecommunications and Information Exchange Between Roadside and Vehicle Systems—5-GHz Band Dedicated Short-Range Communications (DSRC), Medium Access Control (MAC), and Physical Layer (PHY) Specifications*. Standard ASTM E2213-03, 2018. [Online]. Available: http://www.astm.org/Standards/E2213.htm

[11] M. A. Togou, L. Khoukhi, and A. Hafid, "Performance analysis and enhancement of WAVE for V2V non-safety applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 8, pp. 2603–2614, Aug. 2018.

[12] X. Zheng *et al.*, "Big data for social transportation," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 3, pp. 620–630, Mar. 2016.

[13] N. Cheng *et al.*, "Big data driven vehicular networks," *IEEE Netw.*, vol. 32, no. 6, pp. 160–167, Nov./Dec. 2018.

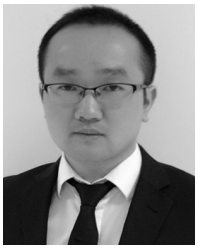
[14] X. Cheng, C. Chen, W. Zhang, and Y. Yang, "5G-enabled cooperative intelligent vehicular (5GenCIV) framework: When benz meets marconi," *IEEE Intell. Syst.*, vol. 32, no. 3, pp. 53–59, May 2017.

[15] R. Meireles, M. Boban, P. Steenkiste, O. Tonguz, and J. Barros, "Experimental study on the impact of vehicular obstructions in VANETs," in *Proc. IEEE VNC*, Dec. 2010, pp. 338–345.
[16] M. Boban, J. Barros, and O. K. Tonguz, "Geometry-based vehicle-to-vehicle channel modeling for large-scale simulation," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4146–4164, Nov. 2014.
[17] I. Tan, W. Tang, K. Laberteaux, and A. Bahai, "Measurement and analysis of wireless channel impairments in DSRC vehicular communications," in *Proc. IEEE ICC*, May 2008, pp. 4882–4888.
[18] L. Cheng, B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige, "Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 GHz dedicated short range communication (DSRC) frequency band," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 8, pp. 1501–1516, Oct. 2007.
[19] X. Yin, X. Ma, K. S. Trivedi, and A. Vinel, "Performance and reliability evaluation of BSM broadcasting in DSRC with multi-channel schemes," *IEEE Trans. Comput.*, vol. 63, no. 12, pp. 3101–3113, Dec. 2014.
[20] M. D. Soltani, M. Alimadadi, Y. Seyedi, and H. Amindavar, "Modeling of Doppler spectrum in V2V urban canyon oncoming environment," in *Proc. 7th IST*, Sep. 2014, pp. 1155–1160.
[21] F. Bai, D. D. Stancil, and H. Krishnan, "Toward understanding characteristics of dedicated short range communications (DSRC) from a perspective of vehicular network engineers," in *Proc. ACM MobiCom*, Sep. 2010, pp. 329–340.
[22] F. Martelli, M. E. Renda, G. Resta, and P. Santi, "A measurement-based study of beaconing performance in IEEE 802.11p vehicular networks," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 1503–1511.
[23] Arada Systems. *LocoMate Classic On Board Unit OBU-200*. Accessed: Mar. 2015. [Online]. Available: http://www.aradasystems.com/locomate-obu/
[24] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, pp. 606–620, Jun. 2007.
[25] H. Zhu, L. Fu, G. Xue, Y. Zhu, M. Li, and L. M. Ni, "Recognizing exponential inter-contact time in VANETs," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–5.
[26] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity WiFi," in *Proc. ACM MobiCom*, Sep. 2015, pp. 53–64.
[27] H. A. Omar, W. Zhuang, and L. Li, "VeMAC: A TDMA-based MAC protocol for reliable broadcast in VANETs," *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1724–1736, Sep. 2013.
[28] Z. Zhang, J. Willson, Z. Lu, W. Wu, X. Zhu, and D. Du, "Approximating maximum lifetime k-coverage through minimizing weighted k-cover in homogeneous wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 24, no. 6, pp. 3620–3633, Dec. 2016.
[29] DSRC Committee, "Dedicated short range communications (DSRC) message set dictionary," Soc. Automot. Eng., Warrendale, PA, USA, Tech. Rep. J2735 200911, Nov. 2009.
[30] F. Lyu *et al.*, "SS-MAC: A novel time slot-sharing MAC for safety messages broadcasting in VANETs," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3586–3597, Apr. 2018.
[31] A. E. F. Clementi, C. Macci, A. Monti, F. Pasquale, and R. Silvestri, "Flooding time of edge-Markovian evolving graphs," *SIAM J. Discrete Math.*, vol. 24, no. 4, pp. 1694–1712, Dec. 2010.
[32] N. Cheng, N. Lu, N. Zhang, X. S. Zhang, X. Shen, and J. W. Mark, "Opportunistic WiFi offloading in vehicular environment: A game-theory approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 7, pp. 1944–1955, Jul. 2016.
[33] X. Cheng, L. Yang, and X. Shen, "D2D for intelligent transportation systems: A feasibility study," *IEEE Trans. Intell. Trans. Syst.*, vol. 16, no. 4, pp. 1784–1793, Jan. 2015.
[34] H. Zhou *et al.*, "TV white space enabled connected vehicle networks: Challenges and solutions," *IEEE Netw.*, vol. 31, no. 3, pp. 6–13, May/Jun. 2017.
[35] F. Bai and H. Krishnan, "Reliability analysis of DSRC wireless communication for vehicle safety applications," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Sep. 2006, pp. 355–362.
[36] N. Wisitpongphan, O. K. Tonguz, J. S. Parikh, P. Mudalige, F. Bai, and V. Sadekar, "Broadcast storm mitigation techniques in vehicular ad hoc networks," *IEEE Wireless Commun.*, vol. 14, no. 6, pp. 84–94, Dec. 2007.
[37] L. Briesemeister and G. Hommel, "Role-based multicast in highly mobile but sparsely connected ad hoc networks," in *Proc. ACM MobiHoc*, 2000, pp. 45–50.
[38] F. J. Ros, P. M. Ruiz, and I. Stojmenovic, "Acknowledgment-based broadcast protocol for reliable and efficient data dissemination in vehicular ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 1, pp. 33–46, Jan. 2012.
[39] O. Rehman, M. Ould-Khaoua, and H. Bourdoucen, "An adaptive relay nodes selection scheme for multi-hop broadcast in VANETs," *Comput. Commun.*, vol. 87, pp. 76–90, Aug. 2016.
[40] M. Boban, R. Meireles, J. Barros, P. Steenkiste, and O. K. Tonguz, "TVR—Tall vehicle relaying in vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 5, pp. 1118–1131, May 2014.

Feng Lyu (M'18) received the Ph.D. degree from Shanghai Jiao Tong University, in 2018. He is currently a Post-Doctoral Fellow with the Broadband Communications Research (BBCR) Group, University of Waterloo. His research interests include vehicular ad hoc networks, space-air-ground integrated networks, big data driving application design, and cloud/edge computing.

Hongzi Zhu received the B.S. and M.Eng. degrees from Jilin University, in 2001 and 2004, respectively, and the Ph.D. degree from Shanghai Jiao Tong University (SJTU), in 2009. He was a Post-Doctoral Fellow at the University of Waterloo and The Hong Kong University of Science and Technology, in 2009 and 2010, respectively. He is currently an Associate Professor with the Department of Computer Science and Engineering, SJTU. His research interests include wireless networks, mobile sensing and mobile computing, and vehicular ad hoc networks.

Nan Cheng (M'16) received the B.E. and M.S. degrees from the Department of Electronics and Information Engineering, Tongji University, Shanghai, China, in 2009 and 2012, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, in 2016. He was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, and also with the Department of Electrical and Computer Engineering, University of Waterloo, from 2017 to 2018. He is currently a Professor with the School of Telecommunication Engineering, Xidian University, Shaanxi, China. His current research interests focus on space-air-ground integrated system, big data in vehicular networks, and self-driving systems. His research interests also include performance analysis, MAC, opportunistic communication, and application of AI for vehicular networks.



Haibo Zhou (M'14–SM'18) received the Ph.D. degree in information and communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2014. From 2014 to 2017, he was a Post-Doctoral Fellow with the Broadband Communications Research Group, Electrical and Computer Engineering Department, University of Waterloo. He is currently an Associate Professor with the School of Electronic Science and Engineering, Nanjing University, Nanjing, China. His research interests include resource management and protocol design in cognitive radio networks and vehicular networks.

Wenchao Xu received the B.E. and M.E. degrees from Zhejiang University, Hangzhou, China, in 2008 and 2011, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His interests include wireless communications, with the emphasis on resource allocation, network modeling, and mobile data offloading.

Minglu Li (M'08–SM'19) received the Ph.D. degree in computer software from Shanghai Jiao Tong University in 1996. He is currently a Full Professor and the Director of the Network Computing Center, Shanghai Jiao Tong University. He has published over 350 papers in academic journals and international conferences. His research interests include vehicular networks, big data, cloud computing, grid computing, and wireless sensor networks. He has also served as a PC member for over 50 international conferences, including the IEEE INFOCOM 2009–2016 and the IEEE CCGrid 2008. He was the Chairman of the Technical Committee on Services Computing (TCSVC) (2004–2016) and the Technical Committee on Distributed Processing (TCDP) (2005–2017), and also the Chairman of the IEEE Computer Society in Great China region. He has served as the General Co-Chair for the IEEE SCC, the IEEE CCGrid, the IEEE ICPADS, and the IEEE IPDPS, and the Vice Chair of the IEEE INFOCOM.

Xuemin (Sherman) Shen (M'97–SM'02–F'09) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990. He is currently a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research interests focus on resource management, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He is an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society. He received the R.A. Fessenden Award from the IEEE, Canada, in 2019; the James Evans Avant Garde Award from the IEEE Vehicular Technology Society in 2018; the Joseph LoCicero Award in 2015; and the Education Award from the IEEE Communications Society in 2017. He has also received the Excellent Graduate Supervision Award in 2006, the five-time Outstanding Performance Award from the University of Waterloo, and the Premier's Research Excellence Award (PREA) from the Province of Ontario, Canada, in 2003. He has served as the Technical Program Committee Chair/Co-Chair for the IEEE Globecom'16, the IEEE Infocom'14, the IEEE VTC'10 Fall, the IEEE Globecom'07, the Symposia Chair for the IEEE ICC'10, the Tutorial Chair for the IEEE VTC'11 Spring, the Chair for the IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He is the Editor-in-Chief of the IEEE INTERNET OF THINGS JOURNAL and the Vice President on Publications of the IEEE Communications Society. He is a registered Professional Engineer of the Province of Ontario, Canada.