Abstract—Life-critical warning message, abbreviated as warning message, is a special event-driven message that carries emergency warning information in Vehicle-to-Everything (V2X). Three important characteristics that distinguish warning messages from ordinary vehicular messages are sporadicity, crowding, and ultra-time-criticality. In other words, warning messages come only once in a while in a sporadic manner; however, when they come, they tend to come as a crowd and they need to be delivered in short order. This paper puts forth a medium-access control (MAC) protocol for warning messages. To circumvent potential inefficiency arising from sporadicity, we propose an override network architecture whereby warning messages are delivered on the spectrum of the ordinary vehicular messages. Specifically, a vehicle with a warning message first sends an interrupt signal to pre-empt the transmission of ordinary messages, so that the warning message can use the wireless spectrum originally allocated to ordinary messages. In this way, no exclusive spectrum resources need to be pre-allocated to the sporadic warning messages. To meet the crowding and ultra-time-criticality aspects, we use advanced channel access techniques to ensure highly reliable delivery of warning messages within an ultra-short time in the order of 10 ms. In short, the overall MAC protocol operates by means of interrupt-and-access. We investigate the use of spread spectrum sequences as interrupt signals over the 5.8 GHz ISM band. Simulation results show that the missed detection rate (MDR) of the interrupt signals can be very small given sufficient sequence length, e.g., when the signal-to-interference ratio is $-32$ dB, a 0.43 ms sequence (64512 symbols, 150 MHz) can guarantee an MDR of $10^{-4}$. For channel access, we investigate two uncoordinated channel access schemes for reliable multiple access. Targeting for a $10^{-4}$ message loss rate in our set-up, a simple multi-replica ALOHA scheme can support up to 11 active nodes with a warning message to transmit. If the number of transmitters exceeds 11, a more complex coded ALOHA scheme using successive interference cancellation to extract messages can support up to 120 nodes.

Index Terms—V2X, wireless interrupt, spread spectrum, ISM band, coded ALOHA.

I. INTRODUCTION

With the explosive growth of vehicles on road, safety has become a major concern for future intelligent transportation systems (ITS) [1]. Statistical data show that the number of crashes in the United States is nearly 6 million each year [2]. Vehicle-to-Everything, abbreviated as V2X, is a promising means to cut the road toll [3]. Through V2X, all the entities on the road (e.g., vehicles, road side units and pedestrians) are connected, hence, they can exchange safety messages and cooperate to prevent road accidents or cut down fatality and injury rates when they do occur.

Safety messages in V2X can be classified into two categories [3]: 1) heartbeat messages. Each on-road node periodically broadcasts heartbeat messages to declare its existence, current state and environment information. Receiving nodes can then evaluate whether there are hazards from information disseminated by transmitters and data gathered from the environment. 2) event-driven messages. Safety in V2X is not limited to passive evaluation of the received heartbeats. An on-road node encountering unexpected events could actively broadcast event-driven messages so that the surrounding nodes can respond quickly. Typical events that may induce event-driven messaging include lane change, roadwork, ambulance approach, to name a few.

Among event-driven messages, life-critical warning messages (hereinafter, referred to as warning messages) deserve particular attention [4]. Warning messages are triggered by extreme traffic emergencies that are likely to cause casualties, e.g., hard braking on the highway, imminent crash, and swerving vehicles at the opposite lane. Typically, a warning message contains the following data [5]: node ID (4 bytes), message generation time (4 bytes, modulo one minute, with resolution 1 $\mu$s), message type (2 bytes, e.g., braking, acceleration, steering) and message attributes (14 bytes, e.g., for braking message type, the attributes could contain brake force, current vehicle speed and wheel state) for an aggregate of 24 bytes.

Three important characteristics that distinguish warning messages from ordinary vehicular messages are as follows:

1) They are rare and sporadic. Statistics indicate that there are on average 1.04 fatal crashes every 100 million miles a vehicle travels [2].

2) Warning messages are short but multiple warning messages may arrive as a crowd in a batch. This is because a single emergency event can trigger multiple emergency responses from multiple nearby nodes. As a result, these emergency nodes (typically less than 30) can broadcast multiple warning messages simultaneously.

3) They must be delivered with high certainty in short order. According to the automotive white paper from 5G-PPP [1], the maximum tolerable end-to-end delay of these safety-of-life messages is 10 ms, and the maximum tolerable message loss rate within 10 ms is $10^{-4}$.

We refer to these three message characteristics as sporadicity, crowding, and ultra-time-criticality.

In V2X, safety messages are disseminated by simple means of one-hop broadcast. Multiple access control (MAC) designs are especially crucial if the stringent delay and reliability

Y. Shao, S. C. Liew and J. Liang are with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong (e-mail: {sy016, soung, lj015}@ie.cuhk.edu.hk). This paper was presented in part in ICC 2018.
requirements are to be met. As shown in Fig. 1, the Federal Communications Commission (FCC) in the United States allocates 75 MHz “5.9 GHz band” for V2X communication [5], [6], based on which many MAC protocols have been proposed and developed to support safety-message broadcasting [5]–[9]. However, existing schemes are designed primarily for heartbeat messages and conventional event-driven messages. When it comes to sporadic ultra-time-critical crowd messaging, none of them can meet the stringent delay and reliability requirements.

To fill this gap, this paper puts forth a medium-access control protocol tailored for the delivery of life-critical warning messages without exclusive allocation of wireless spectrum to them. Two underpinnings of our MAC protocol are as follows:

1) To address sporadicity efficiently, we build our MAC protocol upon an override network architecture whereby wireless spectrum originally allocated to regular vehicular network is used to deliver warning messages only when they appear. Specifically, no wireless spectrum is dedicated exclusively to warning messages since they rarely occur. A vehicle with warning messages will first send an interrupt message to pre-empt the transmission of regular vehicular messages so that the warning message can follow after that.

2) To address crowding and ultra-time-criticality, we use advanced channel access techniques to ensure delivery of warning messages within the stringent delay and reliability targets. In particular, in a life-threatening situation, multiple vehicles may have life-critical messages to send. A channel access protocol that does not incur excessive hand-shaking overhead to coordinate the transmissions of these vehicles on the shared spectrum is critical if the stringent delay target is to be met.

In short, the overall MAC protocol operates by means of interrupt-and-access. For wireless interrupt, we devise an interrupt mechanism for V2X in which the interrupt signals are spread spectrum sequences. Simulation results show that the missed detection rate (MDR) of the interrupt signals can be very small provided that the interrupt sequences are long enough, e.g., when the signal-to-interference ratio (SIR) is $-32$ dB, a 0.43 ms sequence (64512 symbols, 150 MHz) can guarantee an MDR of $10^{-4}$. For channel access, we investigate two uncoordinated channel access schemes for reliable multiple access. Targeting for a $10^{-4}$ message loss rate in our set-up, a simple multi-replica ALOHA scheme can support up to 11 nodes. If the number of transmitters exceeds 11, a more advanced (and more complex) coded ALOHA scheme can potentially support up to 120 nodes while keeping the message loss rate lower than $10^{-4}$.

The remainder of this paper is organized as follows: Section II reviews the state-of-the-art MAC protocols for V2X. Section III outlines our interrupt-and-access MAC protocol tailored for life-critical warning messages. Section IV presents our wireless interrupt protocol, and the design of interrupt signals on the ISM band. Section V presents two random channel access protocols for warning messages. Section VI concludes this paper.

II. STATE-OF-THE-ART V2X MAC PROTOCOLS

Existing MAC protocols for V2X communications operate in either a distributed or a centralized manner. Centralized MAC designs, e.g., LTE-based MAC [7], have certain limitations: 1) Infrastructure could be a single point of failure. These MAC protocols may not function when infrastructure failure occurs or when vehicles are out of the coverage of the infrastructure (e.g., in blind zone, tunnels, and underground parking lots). 2) The coordination-based framework, e.g., schedule-before-transmit, does not fit delay-sensitive applications, owing to the extra delay and overhead consumed. Distributed self-organizing MAC designs are in general more suitable for ultra-delay-sensitive warning messages [1].

A. IEEE 802.11p

Dedicated short-range communication (DSRC) [5] refers to the sets of standards on the 5.9 GHz band. The MAC protocol in DSRC, i.e., IEEE 802.11p [10], is an amendment from IEEE 802.11a with enhanced distributed channel access (EDCA) Quality-of-Service (QoS) extension. In 802.11p, both heartbeat and event-driven messages share the 10 MHz control channel (CCH) by means of carrier sensing multiple access (CSMA). In particular, different types of messages are assigned with different priorities: high-priority messages have smaller interframe spacing and backoff waiting time so that they have priority over low-priority messages in channel access.

There are three main reasons why 802.11p is not suitable for sporadic ultra-time-critical crowd messaging, even if we assign the highest priority to warning messages.

1) Delay concern – In 802.11p, messages with different priorities share the same control channel. When high-priority warning messages are generated, a low-priority message may be in the midst of occupying the channel. As a result, the warning messages must have to wait until the channel is idle. Furthermore, even if the channel is idle, multiple warning messages with the same high priority may compete for the channel simultaneously, leading to a high collision rate that may significantly increases the delay.

2) Lack of acknowledgment (ACK) – Warning messages are broadcasted for all vehicles in the vicinity of the

---

1 Some hybrid MAC designs, e.g., LTE-ProSe [6], support both distributed and centralized modes.
2 The sets of standards in Europe are referred to as C-ITS [6]. DSRC and C-ITS share similar PHY and MAC layers.
warning-message generating vehicle. However, requiring an ACK from each one-hop neighbor of the broadcaster (potentially hundreds of nodes) can be highly costly. As a result, 802.11p does away with ACK for broadcast messages. This means that nodes cannot detect collisions and there is no retransmission. In other words, collisions mean packet loss, and this makes 802.11p highly unreliable. A simple calculation shows that, if we set the contention window to $7$ ($7$ is already the maximum contention window for the access category with the highest priority $[10]$, the collision rate is as high as $50.95\%$ when there are $10$ warning-message transmitters, and this number increases to $92.11\%$ when there are $30$ transmitters.

3) Hidden node problem – To tackle the hidden node problem, RTS/CTS handshaking is implemented in conjunction with CSMA in IEEE 802.11a $[11]$. However, in 802.11p, the hidden node problem is left unsolved, because for broadcast messages, the frequent RTS/CTS handshakes consume too much resources. As a result, a warning broadcast message may collide with another warning broadcast message two hops away, leading to packet loss.

B. TDMA-based MAC

Vehicular MAC protocols may also be based on time division multiple access (TDMA). Two representative examples are ADHOC MAC $[12]$ and its multi-channel evolution VeMAC $[9]$.

As with IEEE 802.11p, VeMAC use the control channel (CH 178) in the 5.9 GHz band for both heartbeat and event-driven messages. This 10 MHz channel is assumed to be time-slotted, and every $M$ slots are grouped together as a frame. In VeMAC, each node occupies at least one slot in every frame for the broadcast of its heartbeat message. If a node has an event-driven message to broadcast, it will need to acquire one more slot. In particular, the slots a node occupies must be different from the slots occupied by any of its neighbors within two hops (this guarantees that there is no hidden node problem). How nodes within two hops coordinate with each other and occupy different slots is an essence of VeMAC.

To enable sporadic ultra-time-critical crowd messaging in the context of VeMAC, all nodes can reserve one slot every 10 ms to cater for the rare occasion when they have warning messages to broadcast. However, simple calculation indicates that this is not viable. Assuming the slot duration is $50 \mu s$, and therefore, there are $200$ slots available in $10$ ms, if there are $200$ nodes within two hops (in practice could be up to 1000), then all the slots are reserved by these nodes for warning messaging alone. Even if we assumed sparse nodes, reserving resources for warning messaging is quite inefficient, because warning messages are rare and sporadic.

Instead of exclusive reservation of slots, a node could attempt to acquire a slot only upon the generation of a warning message. However, slot acquisition under VeMAC takes one or more frames (a frame usually lasts for $100$ ms $[9]$), because the transmitter must wait for all its one-hop neighbors’ ACKs to make sure the new slot is free for it to use. Worse still, when there are $K$ nodes with warning messages, the interaction process for them to acquire $K$ different new slots can take an inordinate amount of time.

III. AN OVERRIDE ARCHITECTURE

Let us consider a typical V2X scenario where $K_{\text{all}}$ on-road nodes are communicating with each other on the 5.9 GHz band in an ad-hoc manner. Each node is equipped with two sets of half-duplex transceivers TRX1 and TRX2. TRX1 is aligned to the 10 MHz control channel (CH 178), on which nodes exchange heartbeat or conventional event-driven messages to get an overall perception of the environment. TRX2 is aligned to the 40 MHz service channels (CH 175 and 181), on which nodes exchange non-safety messages, e.g., infotainment messages.

As illustrated in Fig. 2, an accident suddenly happens: a runaway vehicle $O$ violates the traffic light and runs to an opposite lane. This accident triggers the urgent reactions of $K$ nearby nodes, and each of them generates a life-critical warning message to warn its nearby nodes. For example, node $A$ brakes hard, triggering a warning message informing its neighbors (e.g., nodes $B$ and $C$) of its emergency braking caused by the runaway node $O$, so that they could react in time to avoid further crashes.

These life-critical warning messages have stringent delay and reliability requirements. In this sense, we may need to assign them sufficient time-frequency resources, so that the stringent QoS requirements can be met. On the other hand, warning messages are sporadic and arrive once in a long while, hence, assigning them exclusive resources is highly inefficient, because these resources are wasted most of the time in the absence of life-critical events. This motivates us to build the warning-message MAC protocol upon an override network architecture, where life-critical warning messages share the 40 MHz service channels with non-safety messages. In non-emergency situations, non-safety messages are the primary users on the service channels. When emergency arises, high-priority warning messages will override non-safety messages and seize the service channels $[3]$.

As shown in Fig. 3, our override architecture operates by means of interrupt-and-access: nodes with incoming warning messages send interrupt signals to nodes transmitting non-safety messages to pre-empt them so that the nodes with

\footnote{In practice, warning messages can override the whole 40 MHz bandwidth or part of the bandwidth of the service channels, e.g., override only the 20 MHz channel 181, so that non-safety messages would not be totally deprived of services.}
warning messages can broadcast on the service channels. Sections IV and V provide the details of wireless interrupt and channel access, respectively.

IV. WIRELESS INTERRUPT

Interrupt is a technique widely used in computer systems for multitasking with different priorities. Specifically, an incoming high-priority task triggers an interrupt signal to the central processor, so that the processor can suspend the currently ongoing low-priority task and process the high-priority task immediately. Interrupt is rarely used in conventional wireless communication systems. It is, however, useful for our application scenario.

A. Interrupt in V2X

In the vehicular network, if a node wants to broadcast a warning message successfully on the service channels, a prerequisite is that all its one-hop and two-hop neighbors are silent on these channels. The objective of wireless interrupt in V2X is to silent these neighbors (i.e., neighbors within two hops) in case they are halfway transmitting on the service channels.

To this end, two interrupt signals, a primary interrupt signal (PIS) and a secondary interrupt signal (SIS), are defined. For our MAC protocol, interrupts proceed as follows:

i) A node with a warning message to transmit first broadcasts a PIS.

ii) Any node detecting the PIS then broadcasts a SIS.

iii) Any node detecting a PIS or a SIS (except for nodes in iv) below) keeps silent on the service channels for 10 ms, including nodes that are halfway transmitting on a service channel when the PIS or SIS is detected.

iv) Any node that issues PIS, regardless of whether it receives a PIS/SIS from another node, then transmits its warning message during the channel access period following the interrupt period.

An example is given in Fig. 4, where the emergency node A sends a PIS to pre-empt other nodes from using the service channels. The one-hop neighbors of A are $B = \{B_1, B_2, B_3, B_4\}$, and the two-hop neighbors of A are $C = \{C_1, C_2, C_3, C_4\}$. First, node A broadcasts the PIS. The instant A’s one-hop neighbors B detect the PIS, they keep silent on the service channels. Further, the detection of PIS triggers each of them to broadcast an SIS, so that A’s two-hop neighbors C can detect the SIS and keep silent on the service channels as well. Once they have received a PIS or a SIS, node A’s one-hop and two-hop neighbors will be silent on the service channels in the next 10 ms, and node A can broadcast the warning message in the channel access period safely.

Note that multiple overlapped PISs and SISs may be broadcasted simultaneously (e.g., when there are multiple warning-message transmitters). A node may detect the multiple PIS and SIS separately, and respond to each of the interrupt signal following rule iii).

B. Interrupt signal design

We now present our designs for the PIS and SIS. Potentially, there are two alternatives: in-band interrupt and out-of-band interrupt. We could interrupt in-band\( ^3 \) by exploiting special features of non-safety signal on the service channels. For instance, assuming the non-safety messages are carried by OFDM signals, we could transmit the interrupt signal on the guard-band subcarriers.

This paper considers out-of-band interrupt. Specifically, we transmit interrupt signals on the 5.8 GHz ISM band. PIS and SIS are designed as spread-spectrum sequences \( ^4 \) on this 150 MHz band, so that they can be detected in the presence of interference.

1) Interference on the 5.8 GHz band: The 5.8 GHz ISM band (5.725-5.875 GHz) is a free radio band centered on 5.8 GHz, the channel characteristics of which is similar to that of the 5.9 GHz vehicular band. The primary traffic on the 5.8 GHz band is Wi-Fi signal, and Wi-Fi are commonly deployed indoors.

To evaluate the interference of indoor Wi-Fi signal to our outdoor interrupt signal, we conducted an experiment over our campus to capture 5.8 GHz Wi-Fi signal using USRP X310 (with BasicTX daughter board). The experimental data indicates that in the outdoor environment, 1) most of the time, the 5.8 GHz band was calm and quiet, and nothing can be detected; 2) when Wi-Fi signal was detected on the 5.8 GHz band, the signal power was much lower than the indoor power.

In one experiment, an access point (AP, Linksys EA6900) was deployed indoors. The AP used Wi-Fi channel 153 (5.765 GHz) with 20 MHz channel bandwidth. We measured the

\( ^3 \) Another alternative to realize in-band interrupt is full duplex communication. However, full duplex communication requires dedicated full duplex transceivers, i.e., tailored-made RF chips with self-interference cancellation. Interrupt via full duplex techniques is overkill, because unlike the receiver of a full-duplex link, the receiver of an interrupt does not need to receive a data stream in the reverse direction; it only needs to be able to detect the presence of an interrupt signal.

\( ^4 \) For any neighbor within two hops of an emergency node, interrupt is successful as long as at least one interrupt signal is detected, whether it is PIS or SIS.
received Wi-Fi signal intensities from the AP at two locations. The first location was indoor (5 meters from the AP, LOS), and the second location was outdoor (straight distance 20 meters from the AP, NLOS).

The Power Spectral Densities (PSDs) of the received signals captured indoors and outdoors are plotted in Fig. 5. As can be seen, there is a 32-dB gap between them. In particular, for the outdoor signal, the signal-to-noise ratio (SNR) is about 11.3 dB over the 150 MHz ISM band.

Remark: The PIS and SIS designed in this paper are spread spectrum signals over the 150 MHz ISM band. As will be shown later, the interference from the Wi-Fi captured in the experiment is negligible for the designed PIS/SIS signal as far as missed detection rate and false alarm rate are concerned.

2) Interrupt signal design: This subsection presents the design of PIS and SIS on the 5.8 GHz ISM band. We only explain the generation and detection of PIS in the following, SIS is generated and detected similarly.

The PIS consists of $Q N$-point Zadoff-Chu (ZC) sequences \[14\] embedded in a $Q$-point maximum-length-sequence (m-sequence \[15\]), for a total of $QN$ samples. Denote the m-sequence by $c_p$. Let $z$ be a ZC sequence given by

$$z[n] = \begin{cases} 
\exp\left(-j\pi n(n+1) / N\right) & \text{for } N \text{ odd}, \\
\exp\left(-j\pi n^2 / N\right) & \text{for } N \text{ even}, 
\end{cases}$$

(1)

where $n = 0, 1, 2, ..., N - 1$, and $M$ is a positive integer coprime to $N$. For our application, we set $M = 1$ (the reason for choosing $M = 1$ will be explained later).

Then, the $QN$-point PIS $I_p$ is generated by

$$I_p = c_p \otimes z,$$

(2)

where $\otimes$ is the Kronecker product, and each element in $I_p$ is given by

$$I_p[i] = c_p[[i/N]z[i \mod N]$$

for $i = 0, 1, 2, ..., QN - 1$. In (2), the ZC sequence acts like a spread spectrum sequence with rate 150 MHz, thereby spreading the power of the m-sequence over the 150 MHz band.

The receiver computes two cross-correlations to detect the PIS. Given the received sequence $r$ (i.e., the 150 MHz samples after ADC), the receiver first cross-correlates $r$ and $z$ as follows:

$$y[i] = \sum_{j=0}^{N-1} z[j] r[i + j].$$

(3)

Note that the target interrupt signal is embedded in $r$. Thus, in the presence of an interrupt signal, the operation in (3) produces $Q$ peaks if we look at the absolute values of the resulting sequence $y$, thanks to the correlation property of ZC sequences. Then, we make use of the m-sequence $c_p$ modulated on the ZC sequence, and accumulate the power of all $Q$ peaks, yielding

$$u[i] = \sum_{j=0}^{Q-1} c_p[j] y[i + Nj].$$

(4)

Finally, a sharp peak emerges from the absolute values of sequence $u$. The capture of this peak results in successful detection of PIS.

For the SIS, the same ZC sequence is used, but in place of $c_p$, another $Q$-point m-sequence $c_s$ is used.

Remark: ZC sequences have a nice correlation property: the periodic autocorrelation function of a ZC sequence is zero everywhere except at a single maximum per period \[14\]. However, when we modulate m-sequence onto ZC sequence, this nice correlation property no longer holds. To be specific, let us consider the first two ZC sequences in PIS.

1) If these two ZC sequences are modulated by same values 1 or −1, then $\|y[l]\| = 0$ for $l = 1, 2, 3, ..., N - 1$, because a ZC sequence is orthogonal to its cyclic shift.

2) If these two ZC sequences are modulated by opposite values 1 and −1, we show in Appendix A that

$$\|y[l]\| = 2 \times \left| \frac{\sin(\pi M^2/N)}{\sin(\pi M/2)} \right|,$$

(5)

where $l = 1, 2, 3, ..., N - 1$.

As can be seen, when the two adjacent ZC sequences in PIS are modulated by opposite values, the resulting cross-correlated signal is in general nonzero at $l = 1, 2, 3, ..., N - 1$. The cross-correlated signals for $M = 1$ and $M = 3$ are shown in Fig. 6. Among all possible $M$, we found that setting $M = 1$ minimizes the maximal interference $\max_l \|y[l]\|$ as well as the
C. Performance evaluation

The three components in the received sequence $r$: the target interrupt signal, the Wi-Fi signal on the 5.8 GHz band as interference, and noise. If we fix the noise power, then the successful detection of interrupt signal depends on the amount of interference, or more precisely, the signal-to-interference ratio (SIR).

To evaluate the detection performance of our scheme under various SIRs, we simulated the following single interrupter case: an interrupt node $A$ broadcasts a PIS, and this PIS triggers three SISs by one-hop neighbors of $A$. For node $A$’s one-hop neighbors, detection of the PIS peak means a successful interrupt; for node $A$’s two-hop neighbors, detection of at least one SIS peak means a successful interrupt.

Performance metrics in our simulation are missed detection rate (MDR) and false alarm rate (FAR) [16]. We set a threshold $\gamma_{th}(Q, N)$ commensurate with the PIS (SIS) length (longer sequence corresponds to higher thresholds). Any entry in sequence $|u|$ above $\gamma_{th}$ is considered as a peak. Moreover, the real data we collected outdoors only contains one Wi-Fi signal (SNR 11.3 dB). For the simulation, we deliberately added additional Wi-Fi signal so that we can vary the SIR, and show the robustness of our system under even stronger interferences from Wi-Fi signals.

The MDR versus SIR (dB) are shown in Fig. 7 where we fix the received power of PIS and SIS to be equal to the noise power (that is, $-90$ dBm/Hz $\times 150$ MHz $= -8.24$ dBm), and set different interference power $P_I$ to obtain the target SIR (e.g., for SIR $= -28$ dB, we set $P_I = 19.76$ dBm).

Two observations from Fig. 7 are as follows: 1) The MDR of the interrupter $A$’s two-hop neighbors outperforms that of $A$’s one-hop neighbors, given the same PIS (SIS) length. This is intuitive because each of $A$’s two-hop neighbors has three chances to capture the SIS, while $A$’s one-hop neighbors only have one chance to capture the PIS. 2) If we set $0.0001$ as the required MDR, then setting $Q = 31$ can meet the requirement when SIR $\geq -28$ dB; setting $Q = 63$ can meet the requirement when SIR $\geq -32$ dB. Moreover, when $Q = 63$, the number of symbols in PIS (SIS) is $NQ = 64512$, and the overall time consumed by interrupt is about 0.43 ms (the signal processing time is ignored).

We now evaluate the FAR versus interference power with the same simulation setup as for Fig. 7. In this simulation, there is no interrupt signal, and the received sequence contains only interference and noise. The noise power is fixed to $-8.24$ dBm as in Fig. 7 and the interference powers are set as in Fig. 7 (e.g., for SIR $= -28$ dB, we set $P_I = 19.76$ dBm; for SIR $= -40$ dB, we set $P_I = 31.76$ dBm). FAR is defined as the probability that we detect a false alarm within a sample sequence of PIS (SIS) length (i.e., $QN$ samples). Thus, given a FAR, we can calculate the number of false alarms per hour by FAR $\times 3600 \times 150$ MHz/$QN$.

The FAR versus interference power under different thresholds are shown in Fig. 8. As can be seen, higher threshold yields better FAR performance. If we set $Q = 63$, $N = 1024$, the number of false alarms per hour is less than 1 when $P_I = 21.76$ dBm (corresponding to SIR $= -30$ dB in Fig. 7). Note that the consequence of a false alarm is only to keep silent on the service channel for 10 ms.

Remark: If we use a long $NQ$-point ZC sequence instead of our design in this paper (i.e., $Q$ cascaded $N$-point ZC sequences), the detection performance will be the same. However, long ZC sequences greatly increases the computational complexity. Specifically, 1) for our design, the two-step cross-correlation takes $N+Q$ multiplication and $N+Q-2$ addition; 2) for a long $NQ$-point ZC sequence, one $NQ$-point cross-correlation is needed, and it takes $NQ$ multiplication and $NQ-1$ addition.

A caveat here is that if a one-hop neighbor $B_1$ did not detect the PIS, it is possible for it to detect SIS sent by another one-hop neighbor $B_2$ who detected the PIS, if $B_1$ and $B_2$ are within transmission range of each other. Thus, the MDR of a one-hop neighbor presented here is conservative.
V. CHANNEL ACCESS

After interruption, the service channels are set aside for ultra-time-critical crowd messaging. The next problem is the channel access of multiple emergency nodes. Overall, we can summarize the problem as follows:

- There are \( K \) (out of \( K_{all} \)) active nodes. Typically, \( K \in [0, 30] \) and \( K_{all} \in [0, 1000] \).
- All the active nodes intend to transmit a message (24 Bytes) within \( T = 10 - T_I \) ms, where \( T_I \) is the time consumed by interrupts.
- The available bandwidth is 40 MHz. In practice, we may override only 20 MHz so that the primary traffic of the service channels would not be clipped suddenly, and can still transmit on the other 20 MHz channels.

Schedule-based channel access protocols, e.g., TDMA, FDMA, CDMA, OFDMA, requires pre-allocating orthogonal resources for the overall \( K_{all} \) nodes [17]. Let us take CDMA for instance. When operated with CDMA, all the nodes within two hops are pre-assigned different spread spectrum codes, e.g., PN codes, so that the spread spectrum signals from distinct nodes will not interfere with each other. For one thing, a background coordinator must run in all time to guarantee all the nodes within two hops use different PN codes; for another, since there is no prior information on the potential transmitters, a receiving node must despread the received signal using all the potential PN codes (up to a few thousands). This poses great challenges to the processing capacity of the receiver. In this context, random channel access protocols are preferable in our framework.

A. Multi-replica ALOHA

A simple random-access protocol is ALOHA [18]. However, ALOHA requires ACK to inform the transmitter whether the previous transmission is successful or not. As stated in the introduction, ACK is not viable for the broadcast scenario, because each broadcast requires feedback from all one-hop neighbors, incurring excessive overhead when the network is dense, hence compromising the ability to meet the critical time constraint.

One alternative is multi-replica ALOHA. The basic idea is that, since transmitters cannot determine whether their transmissions are successful or not given the lack of ACK, they can replicate their warning packet \( d \) times and randomly broadcasts these \( d \) replicas within \( T \) ms. If one or more replicas from a node are broadcasted without any collision, then the delivery of the warning message is considered successful. An example is given in Fig. 9 in which \( d = 4 \), and three transmitters \( A, B \) and \( C \) broadcast four replicas, respectively. In this example, only replica \( B_3 \) is clean. Thus, only node \( B \) successfully broadcasts its warning message while nodes \( A \) and \( C \) fail.

1) Message loss rate: To analyze the performance, we first derive a probability \( P_0 \); for any two nodes \( A \) and \( B \), \( P_0 \) is defined as the probability that a particular \( A \)’s replica, say \( A_1 \), does not collide with any of \( B \)’s \( d \) replicas. For multi-replica ALOHA, \( P_0 \) as follows is derived in Appendix E

\[
P_0 = \left( \frac{T - (d + 1)T_p}{T} \right)^d \left( \frac{d+1}{d} \right) \left( \frac{T - T_p}{T} \right).
\]

(6)

\( P_0 \) is the probability that \( A_1 \) is clean with respect to \( B \)’s messages. By our independence assumption, \( P_0 \) is also the probability that \( A_1 \) is clean with respect to any other node’s messages and that \( P_0^{K-1} \) is the probability that \( A_1 \) is clean with respect to all other node’s messages.

As a result, the probability that node \( A \) can broadcast its warning message successfully, denoted by \( P_{1, succ} \), is equal to the probability that one or more replicas from node \( A \) are clean. That is,

\[
P_{1, succ} = 1 - (1 - P_0^{K-1})^d,
\]

(7)

wherein an approximating assumption is made that all \( d \) replicas from node \( A \) have independent collision probabilities. This approximation is valid when \( dT_p \ll T \).

We note that \( P_{1, succ} \) is the “message success rate” \( R_{succ} \), defined as the number of successful nodes over the number of all emergency nodes. Thus, the “message loss rate” is

\[
R_{loss} = 1 - R_{succ} = 1 - P_{1, succ}.
\]

(8)

2) Numerical results: To evaluate the multiple-access performance of Multi-replica ALOHA, we consider a specific OFDM-based PHY layer, the parameters of which are given in Table I [5]. In particular, 1) we assume warning messages override half of the service channels, i.e., 20 MHz, so that the non-safety messages would not be totally deprived of services. 2) A typical 24 Byte warning message occupies two OFDM symbols, leading to a 24 \( \mu \)s warning packet at the PHY layer (each OFDM symbol is 8 \( \mu \)s and the preamble is 8 \( \mu \)s). 3) The time for interruption is \( T_I = 0.5 \) ms, hence, the available time for channel access is \( T = 9.5 \) ms.

Following [5], the numerical results of \( R_{loss} \) is plotted in Fig. 10. We also simulate a system to verify the accuracy of \( R_{loss} \) given by (8). It turns out that the simulated \( R_{loss} \) matches with Fig. 10 very well, and the approximation in (7) is very accurate in our set-up.

Two observations from Fig. 10 are as follows:

1) For different numbers of emergency nodes \( K \), Fig. 10 shows that the optimal performance (i.e., the minimum \( R_{loss} \)) is obtained by different duplication factors \( d^* \). For example, when \( K = 10 \), the optimal \( d^* = 14 \), and when \( K = 20 \), the optimal \( d^* = 7 \). We show in Appendix C that the optimal \( d^* \) can be approximated by

\[
d^* = \frac{\ln 2}{2(K - 1)} T_p.
\]

(9)
classical ALOHA, the transmission attempt rate 

\[ G \]

Remark: 

offered load (i.e., new arrivals) and the retransmissions. The 

\[ KT \]

than maximizing the throughput. In particular, the offered load 

\[ K \]

Fig. 10. The message loss rate 

\[ R \]

of the 

\[ KT = 0.24 \]

Each replica is 

\[ R > 2 \]

Given 

\[ R_{loss} = 10^{-1} \]

as the target performance, the 

\[ K^* = 11 \]

More generally, given a target message loss rate 

\[ R_{loss}, \]

the maximal number of sustainable nodes in multi-replica 

\[ K^* \]

\[ K \]

\[ T \]

+ 1. \hspace{1cm} (10)

The derivations and insights are presented in Appendix C.

Remark: Appendix C also draws an analogy between classical ALOHA and our problem using Multi-replica ALOHA. In classical ALOHA, the transmission attempt rate 

\[ G \]

node to jack up the transmission attempt rate to lower the loss probability. The effective 

\[ G \]

is therefore 

\[ KT_{loss} \]

(i.e., number of attempts per packet duration). Assuming large 

\[ K \]

so that 

\[ K - 1 \approx K \]

, equation (9) and 

\[ G = \frac{dKT}{T} \]

imply that we have to achieve an effective 

\[ G = \frac{\ln 2}{2} \]

to obtain the minimum message loss rate. Note that the expression 

\[ G = \frac{\ln 2}{2} \]

is independent of 

\[ d \]

and 

\[ K \]

under the adoption of the optimal 

\[ d \]

for the given 

\[ K \]

This expression in turn implies that we need to modify the optimal transmission attempt rate of classical ALOHA, 

\[ G = \frac{1}{2} \]

, by a factor of 

\[ \ln 2 = 0.693 \]

in order to arrive at 

\[ G = \frac{\ln 2}{2} \]

as the optimal transmission attempt rate for our problem.

Overall, multi-replica ALOHA is a simple channel-access technique. Compared with other advanced techniques (e.g., the coded ALOHA introduced below), simplicity is its most attractive property. In particular, the signal processing of multi-replica ALOHA will not consume much additional time. Thus, for a target 

\[ R_{loss} \]

, if the number of transmitters 

\[ K \]

is no more than 

\[ K^* \]

given in (10), we would recommend Multi-replica ALOHA for reliable channel access with message loss rate less than 

\[ R_{loss} \]. On the other hand, for the target 

\[ R_{loss} \], if 

\[ K \]

exceeds 

\[ K^* \] in (10), we need to resort to more advanced techniques using more complex signal processing. In subsection V-B below, we explore the use of coded ALOHA to increase the sustainable 

\[ K \].

B. Coded ALOHA

At the transmitter, as with the multi-replica approach, each emergency node repeats its broadcast for 

\[ d \]

times to increase the success rate. At the receiver, successive interference cancellation (SIC) [19] can be used to boost performance. In this paper, by Coded ALOHA, we means that the SIC technique is used at the receiver to extract messages. This includes the same set-up that we studied in subsection V-A where a transmitter just repeats its message 

\[ d \]

time (i.e., repetitive code is used), with the difference that SIC is used at the receiver to reduce message loss rate.

Consider the example in Fig. 9 again. Only 

\[ B_3 \]

can be decoded with the previous multi-replica reception mechanism. With coded ALOHA, the receiver stores all the signal received during the 

\[ T \]

ms, and make use of SIC to recursively cancel the interference caused by the decoded nodes. First, the clean replica 

\[ B_3 \]

can be used to cancel other replicas of node 

\[ B \]

, i.e., 

\[ B_1, B_2, \] and 

\[ B_4. \] As a result, the interference from node 

\[ B \]

to other nodes is removed. Moreover, this interference cancellation process creates a new clean replica 

\[ C_2 \]

, and all node 

\[ C \]’s replicas can be removed accordingly. Finally, only replicas from node 

\[ A \] is left, and they are all clean and decodable.

This scheme, multi-replica ALOHA with SIC (or for simplicity, coded ALOHA), is similar to coded slotted ALOHA [20], [21], except for the absence of the concept of slotted time in the former. Practically, a time-slotted system causes two problems in our application: 1) The slot must be short, e.g., as short as 

\[ 24 \mu s \]. However, small slot duration means larger overhead on slot alignment/synchronization among nodes. The required guard time between slots will eat up a large portion of the slot time. 2) Alignment and synchronization of slots

| Types          | Description                  | Value      |
|----------------|------------------------------|------------|
| OFDM           | modulation                   | QPSK       |
| PHY            | channel code rate            | 1/2        |
|                | CP duration                  | 1.6 µs     |
|                | OFDM symbol duration         | 8 µs       |
|                | preamble duration            | 8 µs       |
| potential transmitters | \( K \in [0, 30] \)       |            |
| warning message size | 24 Bytes               |            |
| warning packet duration | 24 µs               |            |
| warning message | Overall TTL of warning messages | 10 ms  |
| message        | Time consumed by interruption | \( T_I = 0.5 \) ms |
|                | Time left for channel access  | \( T = 9.5 \) ms |

Fig. 10. The message loss rate 

\[ R_{loss} \]

of multi-replica ALOHA, where each of the \( K \) emergency nodes broadcasts \( d \) replicas of their warning packets. Each replica is 

\[ 24 \mu s \], and the overall time available for channel access is 

\[ T = 9.5 \] ms. The results presented here are based on [5]. Simulation results, not showing here, are almost exactly the same as the analytical results.

The approximate analytical expression in (9) gives the right ballpark of the empirical optimal 

\[ d^* \]

shown in Fig. 10.

2) Given 

\[ R_{loss} = 10^{-1} \]

as the target performance, the maximal number of sustainable nodes in multi-replica ALOHA is 

\[ K^* = 11 \]. More generally, given a target message loss rate 

\[ R_{loss} \], the maximal number of sustainable nodes in multi-replica ALOHA can be approximated by

\[ K^* = -\frac{T}{\ln 2} \cdot \frac{\ln 2}{2\ln R_{loss}} + 1. \hspace{1cm} (10) \]

The parameter settings are presented in Table I.

| Types          | Description                  | Value      |
|----------------|------------------------------|------------|
| available bandwidth | 20 MHz              |            |
| subcarrier spacing    | 156.25 KHz           |            |
| available data subcarriers | 96               |            |
in the slotted system must be maintained all the time since we cannot predict the arrival of emergencies. That is, nodes will need to participate in the slot synchronization process whether they currently have urgent messages to transmit, in preparation for possible arrivals of urgent messages—performing synchronization only after the arrivals of messages will likely to cause unacceptable latency.

The simulation results for coded ALOHA are plotted in Fig. 11, where the PHY-layer parameter settings are given in Table I, and we assume perfect interference cancellation at the PHY layer. For different number of active nodes $K = 10, 20$, and 30, the message loss rate $R_{\text{loss}}$ is simulated. As can be seen, when $d \geq 3$, the $10^{-4}$ requirement is satisfied for all $K \leq 30$.

To measure the maximal number of sustainable nodes $K^*$ in coded ALOHA systems, we keep increasing the number of active nodes $K$ and simulate $R_{\text{loss}}$ given different degree $d$. Fig. 12 shows the $R_{\text{loss}}$ when $K = 120$. Approximately, the maximal number of sustainable nodes in coded ALOHA systems is $K^* = 120$ given a target performance $R_{\text{loss}} = 10^{-4}$.

Remark (The optimal degree distribution): The performance of coded (unslotted) ALOHA is analytically intractable due to the lack of mathematical tools to characterize the embedded SIC process. On the other hand, in coded slotted ALOHA, the SIC decoding process can be analytically described by iterative message passing (i.e., the evolution of the erasure probabilities) on a bipartite graph [20], [22]. The bipartite graph consists of Burst Nodes (BNs), Sum Nodes (SNs) and edges. For example, a BN is a warning message transmitter, a SN is a slot, and an edge connects a BN and a SN if and only if a replica of the BN is transmitted in the SN/slot. The number of edges connected to a BN is referred to as the degree of a BN. The optimal degree distribution is given by

$$
\Lambda_D(x) = \sum_{d=1}^{D} \lambda_d x^d.
$$

The problem is then to discover the optimal degree distribution $\Lambda_D(x)$ to minimize the message loss rate $R_{\text{loss}}$.

The degree distributions simulated in Fig. 11 and 12 are regular distributions $\Lambda_d(x) = x^d$ for fixed $d = 1, 2, 3, 4, ...$. It is shown that the regular distribution $\Lambda_4(x)$ has already met the reliability requirements of warning messages. However, for the problem itself, the optimal degree distribution $\Lambda_D(x)$ is yet unknown due to the lack of mathematical tools for coded slotted ALOHA.

We note that our problem is different from the problem studied previously in the context of coded slotted ALOHA [22]. The most obvious difference is that ours is an unslotted system while [22] studied a slotted system. A more subtle difference is that the degree distribution obtained in [22] is one that optimizes the throughput in the asymptotic limit when the number of active nodes $K$ goes to infinity. For our problem set-up, $K$ is finite, and the offered load $KT_p/T$ (therefore the target throughput) is low. For a given $K$ and offered load $KT_p/T$, our problem is to find the optimal degree distribution that minimizes $R_{\text{loss}}$. For example, with $T_p/T = 24/9500$ and $K = 30$, the offered load is only 0.076. In essence, we are trying to achieve low latency (low $T$) and high reliability (low $R_{\text{loss}}$) with a finite node population (finite $K$); whereas in [22], the aim is to study the asymptotic throughput in the limit that $K$ (and therefore $T$) goes to infinity. Because of these fundamental differences, it is not clear that the degree distribution optimal for the problem set-up in [22] is also optimal in the context of high reliability with low latency such as in our problem set-up. As we will see, the answer is no.

Additional simulations are performed by us to verify if the optimal degree distributions designed for coded slotted
ALOHA can be applied to our problem of coded unslotted
ALOHA. The simulation results are presented in Table II
in which \( K = 30 \). The first three rows of Table II are
irregular distributions (with maximal degrees 4, 8, and 16)
designed for coded slotted ALOHA [22]. With greater
maximal degree, higher asymptotic threshold of throughput,
denoted by \( G_5^* \), can be achieved (if the offered load is
smaller than \( G^* \), the messages can be recovered from
the SIC process with a probability close to 1 in the
asymptotic limit when \( K \rightarrow \infty \)). The last two rows in
Table II are regular distributions used in Fig. 11. As shown,
the regular distributions outperforms irregular distributions
by much.

VI. CONCLUSION

This paper studies the problem of life-critical warning
messaging in V2X. Our main contributions are as follows.

1) We put forth an interrupt-and-access MAC protocol for
warning messaging that takes into account the three
characteristics of warnings: sporadicity, crowding, and
ultra-time-criticality requirement. The idea is to
interrupt the regular wireless services only when warning
messages arrive so as to acquire usage of the spectrum
ordinarily allocated to the regular services. In this way,
precious wireless spectrum does not have to be pre-
allocated to warning messaging, which occurs only once
in a long while in a sporadic manner.

2) For wireless interrupt, we devised an interrupt
mechanism for V2X and presented an out-of-band
interrupt signal design where the interrupt signals are
spread spectrum sequences on the ISM band. Simulation
results validate the nice detection performance of our
design, e.g., for a 0.43 ms (64512 symbols, 150 MHz) sequence,
the missed detection rate can be kept lower than 0.0001 when SIR
\( \geq -32 \) dB.

3) For wireless access, we investigated different uncoor-
dinated channel access schemes to meet the stringent
delay and reliability requirements of warning message.
A simple multi-replica ALOHA scheme can support up
to 11 nodes in our set-up with message loss rate lower
than \( 10^{-4} \). If the number of transmitters in the system
exceeds 11, a more advanced coded ALOHA scheme with
successive interference cancellation can support up to 120
nodes in our set-up while keeping the message loss rate
lower than \( 10^{-4} \).

**Appendix A**

**Deriving the Cross-Correlation Results**

In this Appendix, we derive the cross-correlation results
when the two adjacent ZC sequences in PIS are modulated
by distinct values.

First, from (2), the PIS is given by \( I_p = c_p \otimes z \), where \( c_p \)
is a Q-point m-sequence and \( z \) is an N-point ZC sequence
given by (1). In the following derivations, we consider even
N (odd N yields the same results).

As with (3), at the receiver, we cross-correlate PIS \( I_p \) with
the conjugate of ZC sequence \( z \), yielding

\[
\bar{y}[l] = \sum_{j=0}^{N-1} z^*[j] I_p[i + j].
\]

Given sequence \( y \), we find peak in its modulus \(|y|\). Without
loss of generality, we now focus on the first two ZC sequences
in PIS.

If these two ZC sequences are modulated by same values, then

\[
y[l] = \sum_{n=0}^{N-l-1} z^*[n] z[n + l] + \sum_{n=N-l}^{N-1} z^*[n] z[n-N+l] = 0,
\]

where \( l = 1, 2, 3, ..., N - 1 \). Eq. (12) follows since a ZC
sequence is orthogonal to its cyclic shift.

If these two ZC sequences are modulated by distinct values,
say 1 and \(-1\), respectively. We have

\[
y[l] = \sum_{n=0}^{N-l-1} z^*[n] z[n + l] - \sum_{n=N-l}^{N-1} z^*[n] z[n-N+l]
\]

\[
= 2 \sum_{n=0}^{N-l-1} z^*[n] z[n + l].
\]

Substituting (1) into (13), yields,

\[
|y[l]| = 2 \times \left| \frac{\alpha^{2} - \alpha^{-2}}{1 - \alpha^{2l}} \right|,
\]

where \( \alpha = \exp(-j\pi M/N) \).

Notice that \( |\alpha^{2} - \alpha^{-2}| \) and \(|1 - \alpha^{2l}| \) are two strings of
a unit circle on the complex plane. According to the Law of
cosines, we have

\[
|\alpha^{2} - \alpha^{-2}| = \sqrt{1 + 1 - 2 \cos(\frac{\pi M}{N} * 2l)}
\]

\[
= 2 \sin\left(\frac{\pi M}{N}l\right),
\]

\[
|1 - \alpha^{2l}| = \sqrt{1 + 1 - 2 \cos(\frac{\pi M}{N} * 2l)}
\]

\[
= 2 \sin\left(\frac{\pi M}{N}l\right),
\]

Thus, (14) can be written as

\[
|y[l]| = 2 \times \left| \frac{\sin(\pi M^{2}/N)}{\sin(\pi M/N)} \right|.
\]
APPENDIX B

DERIVING \( P_0 \) FOR MULTI-REPLICA ALOHA

This appendix derives the probability \( P_0 \) of multi-replica ALOHA in (6). For any two nodes \( A \) and \( B \), \( P_0 \) is the probability that one of \( A \)'s replicas, say \( A_1 \), does not collide with \( B \)'s \( d \) replicas \( \{B_1, B_2, \ldots, B_d\} \).

Let us consider a simple case where \( d = 1 \), and \( P_0 \) is the probability that \( A_1 \) does not collide with \( B_1 \). Denote by \( t(A_1) \) the transmission start time of packet \( A_1 \), and \( t(B_1) \) the transmission start time of packet \( B_1 \). To avoid collisions, \( t(A_1) \) and \( t(B_1) \) must satisfy the following constraints:

\[
\begin{align*}
|t(A_1) - t(B_1)| &\geq T_p, \\
0 &\leq t(A_1) \leq T - T_p, \\
0 &\leq t(B_1) \leq T - T_p.
\end{align*}
\]

Fig. 13 illustrates these constraints, wherein the shaded regions are the regions that satisfy the constraints. \( P_0 \) is then the proportion of the shaded area to the total area, giving,

\[
P_0 = \frac{(T - 2T_p)^2}{(T - T_p)^2}.
\]

Next, we consider the case \( d = 2 \), and \( P_0 \) is the probability that \( A_1 \) does not collide with \( B_1 \) and \( B_2 \). To avoid collisions, the transmission start times \( t(A_1), t(B_1) \) and \( t(B_2) \) must satisfy

\[
\begin{align*}
|t(A_1) - t(B_1)| &\geq T_p, \\
|t(A_1) - t(B_2)| &\geq T_p, \\
|t(B_1) - t(B_2)| &\geq T_p, \\
0 &\leq t(A_1) \leq T - T_p, \\
0 &\leq t(B_1) \leq T - T_p, \\
0 &\leq t(B_2) \leq T - T_p.
\end{align*}
\]

In particular, the condition \( |t(B_1) - t(B_2)| \geq T_p \) is met by default because node \( B \) will not transmit two overlapping packets. \( P_0 \) can be derived as

\[
P_0 = \frac{\Pr(A_1, B_1, B_2 \text{ do not collide})}{\Pr(B_1, B_2 \text{ do not collide})}.
\]

Fig. 14 illustrates the regions associated with the numerator and denominator of (15). As can be seen, the region associated with the numerator of (15) is essentially a cube with side length \( T - 3T_p \). On the other hand, the region associated with the denominator of (15) is a cuboid with length \( T - 2T_p \) (x-axis), width \( T - 2T_p \) (y-axis) and height \( T - T_p \) (z-axis). As a result,

\[
P_0 = \frac{(T - 3T_p)^3}{(T - 2T_p)^2(T - T_p)}.
\]

In general, for the general case where node \( B \) transmit \( d \) replicas, we have

\[
P_0 = \frac{\Pr(A_1, B_1, \ldots, B_d \text{ do not collide})}{\Pr(B_1, \ldots, B_d \text{ do not collide})}.
\]

where the numerator represents a \((d + 1)\)-dimensional regular polyhedron with side length \( T - (d + 1)T_p \), and the denominator represents a \((d + 1)\)-dimensional polyhedron with side length \( T - dT_p, T - dT_p, \ldots, T - dT_p, T - T_p \) (i.e., only one side is of length \( T - T_p \)). Thus, we have

\[
P_0 = \frac{(T - (d + 1)T_p)^{d+1}}{(T - dT_p)^d(T - T_p)}.
\]

APPENDIX C

APPROXIMATING THE OPTIMAL DEGREE \( d^* \) AND THE MAXIMAL NUMBER OF SUSTAINABLE NODES \( K^* \) IN MULTI-REPLICA ALOHA

Consider the situation faced by one particular node, say, node \( A \). On a line of length \( T \), there are \((K - 1)d\) points corresponding to the beginnings of the \((K - 1)d\) replicas of the other \( K - 1 \) nodes. We denote this set of points by \( S \). To the extent that \( K \) is large, approximately the points in \( S \) form a Poisson process on the line. In other words, the inter-point distance is exponentially distributed with mean

\[
\mu = \frac{T}{(K - 1)d}.
\]

A. The optimal degree \( d \) for a given \( K \)

Consider a replica of node \( A \), say \( A_1 \), that is randomly placed on the line of length \( T \). Refer to the beginning of this packet as point \( t(A_1) \). Ignoring the edge effects at the two ends of the line, the probability that the distance of point \( t(A_1) \) to the next point of \( S \) to the right is more than \( T_p \) is \( e^{-T_p/\mu} \).
Similarly, the probability the distance of point $t(A_1)$ to the next point of $S$ to the left is more than $T_p$ is $e^{-T_p/\mu}$. The probability of no collision is therefore $T_p$ is $e^{-2T_p/\mu}$. Thus, $R_{\text{loss}}$ is approximately given by

$$R_{\text{loss}} = (1 - e^{-\rho d})^d,$$

(16)

where $\rho = \frac{2(K-1)T_p}{T}$. For the regime of our interest (i.e., $T \gg dT_p$), and $K \gg 1$), the $R_{\text{loss}}$ in (16) is approximately equal to that in (8). See Appendix D for more details.

Differentiating $\ln R_{\text{loss}}$ with respect to $d$ and setting the derivative to zero gives us the following equation:

$$e^{-\rho d^*} \ln e^{-\rho d^*} = (1 - e^{-\rho d^*}) \ln(1 - e^{-\rho d^*})$$

This is satisfied by $e^{-\rho d^*} = 1 - e^{-\rho d^*}$, which gives

$$e^{-\rho d^*} = \frac{1}{2}. $$

From (17), the optimal $d^*$ is given by

$$d^* = \frac{1}{\rho} \ln 2 = \frac{\ln 2}{2(K-1) T_p}. $$

(17)

For our experiments where $\frac{T_p}{T} = \frac{9500}{24}$, we have

$$d^* \approx \frac{137}{K}, $$

which gives

$$d^* \approx \begin{cases} 15.2, & K = 10, \\ 7.2, & K = 20, \\ 4.7, & K = 30. \end{cases} $$

(18)

Compared with Fig. [10] (18) approximates the optimal degree very well.

More importantly, the above analysis reveals a fundamental relation between $K$ and $d$ in (17). That is, the optimal $d^*$ is inversely proportional to $K - 1$.

B. On the optimal transmission rate

In classical ALOHA, $G$ is the the number of transmission attempts by all nodes per packet duration (including both the new arrivals and the retransmissions). The throughput of ALOHA is $G e^{-2G}$ packets per packet duration. Thus, the optimal $G$ to maximize throughput is $G = \frac{1}{2}$, and the corresponding optimal throughput is $G e^{-2G} = \frac{1}{2e}$. The study of classical ALOHA is to achieve this optimal throughput. If packets can be retransmitted indefinitely after back-offs until success, then there is no loss in the system as long as the offered load is less than $\frac{1}{2e}$ (subject to a proper backoff method). This may incur excessive delay, however.

In our problem, the offered load is fixed to $\frac{K T_p}{T}$, and we allow $d$ attempts per node. In other words, the effective $G$ in our problem is $\frac{K T_p d}{T}$. Assuming large $K$ so that $K - 1$ is approximately $K$, equation (17) implies that we have to achieve $G = \frac{\ln 2}{2}$ to minimize the message loss rate. That is, we need to modify the optimal transmission attempt rate of classical ALOHA, $G = \frac{1}{2}$, by a factor of $\ln 2 = 0.693$.

Note that, in our problem set-up, we do not adjust the offered load to try to meet the maximum throughput – our offered load is already fixed (in fact smaller than the best sustainable offered load). We try to reduce the loss probability for a fixed offered load lower than the sustainable offered load of ALOHA, while bounding the delay to within $T$. The optimal $G$ will therefore be different.

C. The maximal number of sustainable nodes $K^*$

As far as one of the replicas is concerned, its success rate is given by $e^{-2G} = e^{-\ln 2} = \frac{1}{2}$ in the optimal setting – i.e., half chance of success for each trial. Note that this success rate for a replica is independent of $K$ and $d$ because $K$ and $d$ have been optimized to give $G = \frac{\ln 2}{2}$. Thus, regardless of $K$, under the optimal setting, the failure rate after $d$ attempts of the $d$ replicas is $(\frac{1}{2})^d$. Of course, for a fixed $G$, the larger the $K$, the smaller the $d$. For a given $K$, the minimum message loss rate is given by

$$R^*_{\text{loss}} = (\frac{1}{2})^d = (\frac{1}{2})^{\frac{\ln 2}{2(K-1) T_p}}. $$

(19)

Eq. (19) gives us an insight on how the minimum message loss rate $R^*_{\text{loss}}$ depends on $K$ with the optimized $d$. In the log scale, we have

$$\ln R^*_{\text{loss}} = -\frac{\ln 2}{2(K-1) T_p} T_p. $$

Given a target message loss rate $R_{\text{loss}}$, the maximum $K^*$ is

$$K^* = \frac{T_p}{2 \ln 2} \ln R_{\text{loss}} + 1. $$

(20)

For our settings where $R_{\text{loss}} \leq 10^{-4}$, the maximum $K$ is $K^* = 11$. This is consistent with the numerical results in Fig. [10].

APPENDIX D

RECONCILING (8) WITH (16)

To reconcile the $R_{\text{loss}}$ derived in (8) and (16), we want to show that, for the regime of our interest (i.e., $T \gg dT_p$, and $K \gg 1$), $P_0^{K-1}$ in (16) is approximately equal to $e^{-\rho d}$ given in (8), where $\rho = \frac{2(K-1)T_p}{T}$.

From (6), we have

$$P_0^{K-1} = \left(1 - \frac{(d+1)T_p}{T}T_p \right)^{d(K-1)} \left(1 - \frac{T_p}{T}T_p \right)^{K-1} $$

\begin{align*}
&= \left(1 - \frac{d+1}{d+1} \right)^{d(K-1)} \left(1 - \frac{T_p}{T} \right)^{K-1} \\
&\approx \left(1 - \frac{T_p}{T} \right)^{d(K-1)} \left(1 - \frac{T_p}{T} \right)^{K-1}.
\end{align*}

where the approximation follows because $\frac{T_p}{T} \ll 1$.

As $K$, $Kd \rightarrow \infty$, we have

$$P_0^{K-1} \rightarrow e^{-\frac{d(K-1)T_p}{T}} e^{-\frac{d(K-1)T_p}{T}} \rightarrow e^{-\frac{2d(K-1)T_p}{T}}. $$
REFERENCES

[1] 5G-PPP, “5G automotive vision,” The 5G infrastructure public private partnership, available online: https://5g-ppp.eu/white-papers/, 2015.

[2] NHTSA, “The national highway traffic safety administration,” available online: https://www.nhtsa.gov/research-data, 2017.

[3] R. Popescu-Zeletin, I. Radusch, and M. A. Rigani, “Vehicular-2-X communication: state-of-the-art and research in mobile vehicular ad hoc networks,” Springer Science & Business Media, 2010.

[4] Y. Shao, S. C. Liew, and J. Liang, “Sporadic ultra-time-critical messaging in V2X,” in IEEE ICC, 2018.

[5] J. B. Kenney, “Dedicated short-range communications (DSRC) standards in the United States,” Proc. IEEE, vol. 99, no. 7, pp. 1162–1182, 2011.

[6] E. G. Strom, “On medium access and physical layer standards for cooperative intelligent transport systems in Europe,” Proc. IEEE, vol. 99, no. 7, pp. 1183–1188, 2011.

[7] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, “LTE for vehicular networking: a survey,” IEEE Commun. Mag., vol. 51, no. 5, pp. 148–157, 2013.

[8] X. Lin, J. Andrews, A. Ghosh, and R. Ratasuk, “An overview of 3GPP device-to-device proximity services,” IEEE Commun. Mag., vol. 52, no. 4, pp. 40–48, 2014.

[9] H. A. Omar, W. Zhuang, and L. Li, “VeMAC: A TDMA-based MAC protocol for reliable broadcast in VANETs,” IEEE Trans. Mobi. Comput., vol. 12, no. 9, pp. 1724–1736, 2013.

[10] D. Jiang and L. Delgrossi, “IEEE 802.11p: Towards an international standard for wireless access in vehicular environments,” in IEEE VTC Spring, 2008, pp. 2036–2040.

[11] A. Doufexi, S. Armour, M. Butler, A. Nix, D. Bull, J. McGeehan, and P. Karlsson, “A comparison of the HIPERLAN/2 and IEEE 802.11a wireless LAN standards,” IEEE Commun. Mag., vol. 40, no. 5, pp. 172–180, 2002.

[12] F. Borgonovo, A. Capone, M. Cesana, and L. Fratta, “ADHOC MAC: New MAC architecture for ad hoc networks providing efficient and reliable point-to-point and broadcast services,” Wireless Networks, vol. 10, no. 4, pp. 359–366, 2004.

[13] A. J. Viterbi, CDMA: principles of spread spectrum communication. Addison Wesley Longman Publishing Co., Inc., 1995.

[14] D. Chu, “Polyphase codes with good periodic correlation properties (corresp.),” IEEE Trans. Inf. Theory, vol. 18, no. 4, pp. 531–532, 1972.

[15] J. J. Komo and S.-C. Liu, “Maximal length sequences for frequency hopping,” IEEE J. Sel. Areas Commun., vol. 8, no. 5, pp. 819–822, 1990.

[16] T. D. Wickens, Elementary signal detection theory. Oxford University Press, USA, 2002.

[17] A. Goldsmith, Wireless communications. Cambridge university press, 2005.

[18] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, “An ALOHA protocol for multihop mobile wireless networks,” IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 421–436, 2006.

[19] S. P. Weber, J. G. Andrews, X. Yang, and G. De Veciana, “Transmission capacity of wireless ad hoc networks with successive interference cancellation,” IEEE Trans. Inf. Theory, vol. 53, no. 8, pp. 2799–2814, 2007.

[20] E. Paolini, G. Liva, and M. Chiani, “Coded slotted ALOHA: A graph-based method for uncoordinated multiple access,” IEEE Trans. Inf. Theory, vol. 61, no. 12, pp. 6815–6832, 2015.

[21] S. Yang, Y. Chen, S. C. Liew, and L. You, “Coding for network-coded slotted ALOHA,” in IEEE Inf. Theory Workshop (ITW), 2015.

[22] C. Xie, “Graph-based analysis and optimization of contention resolution diversity slotted aloha,” IEEE Trans. Commun., vol. 59, no. 2, pp. 477–487, 2010.

[23] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, “Design of capacity-approaching irregular low-density parity-check codes,” IEEE Trans. Inf. Theory, vol. 47, no. 2, pp. 619–637, 2001.