Ultra-Reliable Low-Latency Vehicular Networks: Taming the Age of Information Tail

Mohamed K. Abdel-Aziz*, Chen-Feng Liu*, Sumudu Samarakoon*, Mehdi Bennis*, and Walid Saad†
*Centre for Wireless Communications, University of Oulu, Finland
†Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA
E-mails: {mohamed.abdelaziz, chen-feng.liu, sumudu.samarakoon, mehdi.bennis}@oulu.fi, walids@vt.edu

***Abstract*—While the notion of *age of information* (AoI) has recently emerged as an important concept for analyzing ultra-reliable low-latency communications (URLLC), the majority of the existing works have focused on the *average* AoI measure. However, an average AoI based design falls short in properly characterizing the performance of URLLC systems as it cannot account for extreme events that occur with very low probabilities. In contrast, in this paper, the main objective is to go beyond the traditional notion of average AoI by characterizing and optimizing a URLLC system while capturing the AoI tail distribution. In particular, the problem of vehicles' power minimization while ensuring stringent latency and reliability constraints in terms of probabilistic AoI is studied. To this end, a novel and efficient mapping between both AoI and queue length distributions is proposed. Subsequently, extreme value theory (EVT) and Lyapunov optimization techniques are adopted to formulate and solve the problem. Simulation results shows a nearly two-fold improvement in terms of shortening the tail of the AoI distribution compared to a baseline whose design is based on the maximum queue length among vehicles, when the number of vehicular user equipment (VUE) pairs is 80. The results also show that this performance gain increases significantly as the number of VUE pairs increases.**

***Index Terms*—5G, age of information (AoI), ultra-reliable low-latency communications (URLLC), extreme value theory (EVT), vehicle-to-vehicle (V2V) communications.**

I. Introduction

Vehicle-to-vehicle (V2V) communication will play an important role in the next generation (5G) mobile networks, and is envisioned as one of the most promising enabler for intelligent transportation systems [1]–[3]. Typically, V2V safety applications (forward collision warning, blind spot/lane change warning, and adaptive cruise control) are known to be *time-critical*, as they rely on acquiring real-time status updates from individual vehicles. In this regard, the European telecommunications standards institute (ETSI) has standardized two safety messages: cooperative awareness message (CAM) and decentralized environmental notification message (DENM) [4]. One key challenge in V2V networks is to deliver ultra-reliable and low-latency communications for such status update messages.

Indeed, achieving ultra-reliable low-latency communication (URLLC) represents one of the major challenges facing 5G and vehicular networks [5]. In particular, a system design based on conventional average values is not adequate to capture the URLLC requirements, since averaging often ignores the occurrence of extreme events (e.g., high latency events) that negatively impact the overall performance. To overcome this challenge, one can resort to the robust framework of *extreme value theory* (EVT) that can allow a full characterization of the probability distributions of extreme events, defined as the tail of the latency distribution or queue length [6]. Remarkably, the majority of the existing V2V literature which address latency and reliability, focus only on average performance metrics [7]–[11], which is not sufficient to enable URLLC. Only a handful of recent works have considered extreme values for vehicular networks [12], [13]. In particular, the work in [12] focuses on studying large delays in vehicular networks using EVT, via simulations using realistic mobility traces, without considering any analytical formulations. In [13], the authors study the problem of transmit power minimization subject to a new reliability measure in terms of maximal queue length among all vehicle pairs. Therein, EVT was utilized to characterize the maximal queue length.

Since V2V safety applications are time-critical, the freshness of a vehicle's status updates is of high importance [14], [15]. A relevant metric in quantifying this freshness is the notion of *age of information* (AoI) proposed in [15]. AoI is defined as the time elapsed since the generation of the latest status update received at a destination. Thus, providing quality-of-service (QoS) guarantees in terms of AoI is essential for any time-critical application. It should be noted that, minimizing AoI is fundamentally different from delay minimization or throughput maximization. In [15], the authors derive the minimum AoI at an optimal operating point that lies between the extremes of maximum throughput and minimum delay. Recently, minimizing the *average* AoI in vehicular networks was studied (e.g., see [15]–[18], and references therein). However, while interesting, a system design based on average AoI cannot enable the unique requirements of URLLC. Instead, the AoI distribution needs to be considered especially when dealing with time-critical V2V safety applications.

The main contribution of this paper is to go beyond the conventional average AoI notion by developing a novel framework to characterize and optimize the tail of AoI in vehicular networks. In particular, our key goal is to enable vehicular user equipment (VUE) pairs to minimize their transmit power while ensuring stringent latency and reliability constraints based on a probabilistic AoI measure. To this end, we first derive a novel relationship between the probabilistic AoI and

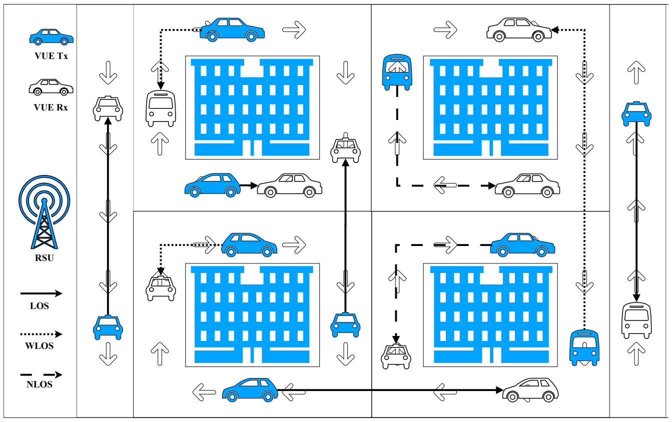

Figure 1. System and path loss models of the considered V2V communication network.

the queue length of each VUE. Then, we use the fundamental concepts of EVT to characterize the tail and excess value of the vehicles' queues, which is then incorporated as a constraint within our optimization problem. To solve the formulated problem, a roadside unit (RSU) is used to cluster VUEs into disjoint groups, thus mitigating interference and reducing the signaling overhead between VUEs and RSU. Subsequently, by leveraging Lyapunov stochastic optimization techniques, each VUE can locally optimize its power subject to probabilistic AoI constraints. Simulation results corroborate the usefulness of EVT in characterizing the distribution of AoI. The results also show over two-fold performance gains in AoI and queue length compared to a baseline whose design is based on the maximum queue length [13]. Our results also expose an interesting tradeoff between the arrival rate of the status updates and the average and worst AoI achieved by the network.

The rest of this paper is organized as follows. In Section II, the system model is presented. The reliability constraints and the studied problem are formulated in Section III, followed by the proposed AoI-aware resource allocation policy in Section IV. In Section V, numerical results are presented followed by conclusions in Section VI.

II. System Model

As shown in Fig. 1, we consider a V2V communication that uses a Manhattan mobility model [19], which is composed of a set $\mathcal{K}$ of K VUE transmitter-receiver pairs under the coverage of a single RSU. During the entire communication lifetime, the association of each transmitter-receiver is fixed. We consider a slotted communication timeline which is indexed by t, and the duration of each slot is denoted by τ. Additionally, all VUE pairs share a set $\mathcal{N}$ of N orthogonal resource blocks (RBs) with bandwidth ω per RB. We further denote the RB usage as $\eta_k^n(t) \in \{0,1\}, \forall k \in \mathcal{K}, n \in \mathcal{N}$, in which $\eta_k^n(t) = 1$ indicates that RB n is used by VUE pair k in time slot t. Otherwise, $\eta_k^n(t) = 0$. The transmitter of pair k allocates a transmit power $P_k^n(t) \geq 0$ over RB n to serve its receiver subject to $\sum_{n\in\mathcal{N}} \eta_k^n(t) P_k^n(t) \leq P_{\max}$, where $P_{\max}$ is the total power budget. Moreover, $h_{kk'}^n(t)$ is the instantaneous channel gain, including path loss and channel fading, from the transmitter of pair k to the receiver of pair k' over RB n in slot t. We consider the 5.9 GHz carrier frequency and adopt the path loss model in [20].

For our model, we express an arbitrary transmitter's and an arbitrary receiver's Euclidean coordinates as $\boldsymbol{x} = (x_i, x_j) \in \mathbb{R}^2$ and $\boldsymbol{y} = (y_i, y_j) \in \mathbb{R}^2$, respectively. When the transmitter and receiver are on the same lane, we consider a line-of-sight (LOS) path loss value $l_0 \left\|\boldsymbol{x} - \boldsymbol{y}\right\|^{-\alpha}$, where $\|.\|$ is the l_2-norm, l_0 is the path loss coefficient, and α is the path loss exponent. If the transmitter and receiver are located separately on perpendicular lanes with one near the intersection within a distance $\mathscr{D}$, then we consider a weak-line-of-sight (WLOS) path loss model $l_0 \left\|\boldsymbol{x} - \boldsymbol{y}\right\|_1^{-\alpha}$ with the l_1-norm $\|.\|_1$. Finally, if both transmitter and receiver are located on the perpendicular lanes, but the distances to the intersection are larger than $\mathscr{D}$, then we adopt a non-line-of-sight (NLOS) path loss value $l_0^{'} \left(|x_i - y_i|.|x_j - y_j|\right)^{-\alpha}$, with the path loss coefficient $l_0^{'} < l_0 \left(\frac{\mathscr{D}}{2}\right)^{\alpha}$. Fig. 1 briefly illustrates these three path loss cases. For this network, the data rate of VUE pair k in time slot t (in the unit of packets per slot) is expressed as

$$R_k(t) = \frac{\tau}{Z} \sum_{n\in\mathcal{N}} \omega \log_2 \left(1 + \frac{P_k^n(t) h_{kk}^n(t)}{N_0\omega + I_k^n(t)}\right), \tag{1}$$

with Z the packet length in bits and N_0 is the power spectral density of the additive white Gaussian noise. Here, $I_k^n(t) = \sum_{k'\in\mathcal{K}/k} \eta_{k'}^n(t) P_{k'}^n(t) h_{k'k}^n(t)$ indicates the received aggregate interference at the receiver of VUE pair k over RB n from other VUE pairs operating over the same RB n.

Furthermore, each VUE transmitter has a queue buffer to store the data to be delivered to the desired receiver. Denoting VUE pair k's queue length at the beginning of slot t as $Q_k(t)$, the queue dynamics are given by,

$$Q_k(t+1) = \max\left(Q_k(t) - R_k(t), 0\right) + A, \tag{2}$$

where A is the constant packet arrival rate per slot, under the assumption of deterministic periodic arrivals. Hence, packet i's arrival time instance can be denoted as $\frac{i}{A}\tau$. It can be noted that the indices of the packets that arrive during slot t satisfy $i \in [tA, (t+1)A-1]$, while the packets that are served during the same slot fulfill

$$tA - Q_k(t) \leq i \leq tA - 1 - \max\left(Q_k(t) - R_k(t), 0\right). \tag{3}$$

III. Reliability Constraints based on Age of Information

Providing real-time status updates in mission critical applications (e.g., CAMs) is a key use case for V2V networks. Further, these applications rely on the "freshness" of the data, which can be quantified by the concept of AoI [15]:

$$\Delta_k(T) \triangleq T - \max_i \left(T_k^{\mathrm{A}}(i) \mid T_k^{\mathrm{D}}(i) \leq T\right). \tag{4}$$

Here, $\Delta_k(T)$ is the AoI of VUE pair k at a time instant T. $T_k^{\mathrm{A}}(i)$ and $T_k^{\mathrm{D}}(i)$ denotes the arrival and departure instant of packet i of VUE pair k, respectively. As a reliability requirement, we impose a probabilistic constraint on the AoI for each VUE pair $k \in \mathcal{K}$, i.e.,

$$\lim_{T\to\infty} \Pr\left\{\Delta_k(T) > d\right\} \leq \epsilon_k, \forall k \in \mathcal{K}, \tag{5}$$

where d is the age limit, and $\epsilon_k \ll 1$ is the tolerable AoI violation probability. It was shown in [14] that for a given age limit d with $\frac{A}{\tau} \geq \frac{1}{d}$, the steady state distribution of AoI for a D/G/1 queue can be characterized as,

$$\lim_{T\to\infty} \Pr\Big\{\Delta_k(T) > d\Big\} = \lim_{T\to\infty} \Pr\Big\{T_k^{\mathrm{D}}(\hat{i}) > T\Big\}, \tag{6}$$

where $\hat{i} \triangleq \lceil \frac{A}{\tau}(T-d) \rceil$ is the index of the packet that first arrives at or just after time $T-d$. Next, in Lemma 1 we propose a mapping between the steady state distribution of the departure instant of a given packet and the queue length.

Lemma 1. *Assuming that T is observed at the beginning of each slot $t+1$, i.e., $T = \tau(t+1)$, then*

$$Pr\Big\{T_k^{D}(\hat{i}) > T\Big\} \leq Pr\left\{Q_k(t) > R_k(t) - \psi\right\},$$

will be satisfied with $\psi = 2 - (\frac{d}{\tau} - 1)A$.

Proof: Since $T = \tau(t+1)$, then $\Pr\Big\{T_k^{\mathrm{D}}(\hat{i}) > T\Big\} = \Pr\Big\{\hat{i}$ is NOT served before time $\tau(t+1)\Big\}$, which means that $\hat{i}$ is not served at or before time slot t. Subsequently, we apply (3) and derive

$$\begin{aligned}
&\Pr\Big\{T_k^{\mathrm{D}}(\hat{i}) > T\Big\} \\
&\overset{(a)}{=} \Pr\Big\{\hat{i} > tA - 1 - \max\left(Q_k(t) - R_k(t), 0\right)\Big\} \\
&\leq \Pr\left\{\frac{A}{\tau}(\tau(t+1) - d) + 1 > tA - 1 - (Q_k(t) - R_k(t))\right\} \\
&= \Pr\left\{Q_k(t) > R_k(t) - \psi\right\},
\end{aligned}$$

where $\hat{i} = \lceil \frac{A}{\tau}(T-d) \rceil \leq \frac{A}{\tau}(\tau(t+1)-d)+1$ is used in step *(a)*. It also should be noted that if $\hat{i}$ departs after $\tau(t+1)$, then $Q_k(t) - R_k(t) > 0$, which is used in the same step. ■

Combining the results of Lemma 1 and (6), the probabilistic constraint (5) can be rewritten as,

$$\lim_{C\to\infty} \frac{1}{C} \sum_{t=0}^{C-1} \Pr\left\{Q_k(t) > R_k(t) - \psi\right\} \leq \epsilon_k, \forall k \in \mathcal{K}. \tag{7}$$

As previously discussed, enabling URLLC requires the characterization of the tail of the AoI distribution. Therefore, we further investigate the event $Q_k(t) > R_k(t) - \psi$ and study the tail behavior of AoI in the following part. First, we need to introduce the *Pickands–Balkema–de Haan theorem* for exceedances over threshold [6]. Consider a random variable Q whose cumulative distribution function (CDF) is denoted by $F_Q(q)$. As a threshold δ closely approaches $F_Q^{-1}(1)$, the conditional CDF of the excess value $X = Q - \delta > 0$ is $F_{X|Q>\delta}(x) \approx G(x;\sigma,\xi)$ where

$$G(x;\sigma,\xi) = \begin{cases} 1 - (\max\{1 + \frac{\xi x}{\sigma}, 0\})^{-\frac{1}{\xi}}, & \xi \neq 0, \\ 1 - e^{-\frac{x}{\sigma}}, & \xi = 0. \end{cases}$$

Here, $G(x;\sigma,\xi)$ is the generalized Pareto distribution (GPD) whose mean and variance are $\frac{\sigma}{1-\xi}$ and $\frac{\sigma^2}{(1-\xi)^2(1-2\xi)}$, respectively. Moreover, the characteristics of the GPD depend on the scale parameter $\sigma > 0$ and the shape parameter $\xi < \frac{1}{2}$.

The *Pickands–Balkema–de Haan theorem* states that for a sufficiently high threshold δ, the distribution function of the excess value can be approximated by the GPD. In this regard, considering constraint (7), we define the conditional excess queue value of each VUE pair $k \in \mathcal{K}$ at time slot t as $X_k(t)|_{Q_k(t)>R_k(t)-\psi} = Q_k(t) - R_k(t) + \psi$. Thus, we can approximate the mean and variance of $X_k(t)$ as

$$\mathbb{E}\left[X_k(t)|Q_k(t) > R_k(t) - \psi\right] \approx \frac{\sigma_k}{1-\xi_k}, \tag{8}$$

$$\mathrm{Var}\left[X_k(t)|Q_k(t) > R_k(t) - \psi\right] \approx \frac{\sigma_k^2}{(1-\xi_k)^2(1-2\xi_k)}, \tag{9}$$

with a scale parameter σ_k and a shape parameter ξ_k. Note that the smaller the σ_k and ξ_k, the smaller the mean value and variance of the GPD. Hence, we further impose the thresholds on the scale and the shape parameters, i.e., $\sigma_k \leq \sigma_k^{th}$ and $\xi_k \leq \xi_k^{th}$ [21]. Subsequently, applying both parameter thresholds and $\mathrm{Var}(X_k) = \mathbb{E}\left[X_k^2\right] - \mathbb{E}\left[X_k\right]^2$ to (8) and (9), we consider the constraints for the time-averaged mean and second moment of the conditional excess queue value, i.e.,

$$\bar{X}_k = \lim_{C\to\infty} \frac{1}{C} \sum_{t=0}^{C-1} \mathbb{E}\left[X_k(t)|Q_k(t) > R_k(t) - \psi\right] \leq H, \tag{10}$$

$$\bar{Y}_k = \lim_{C\to\infty} \frac{1}{C} \sum_{t=0}^{C-1} \mathbb{E}\left[Y_k(t)|Q_k(t) > R_k(t) - \psi\right] \leq B, \tag{11}$$

where $H = \frac{\sigma_k^{th}}{1-\xi_k^{th}}$, $B = \frac{2(\sigma_k^{th})^2}{(1-\xi_k^{th})(1-2\xi_k^{th})}$ and $Y_k(t) := \left[X_k(t)\right]^2$. In light of this, our problem can be formulated as minimizing the total power consumption of VUEs while ensuring the QoS in terms of reliability constraints based on AoI. By denoting the RB usage and power allocation vectors as $\boldsymbol{\eta}(t) = [\eta_k^n(t)]_{k\in\mathcal{K}}^{n\in\mathcal{N}}$ and $\boldsymbol{P}(t) = [P_k^n(t)]_{k\in\mathcal{K}}^{n\in\mathcal{N}}$, $\forall t$, respectively, we formulate a network-wide optimization problem which is written as follows:

$$\begin{aligned}
\mathbb{P}_1 : \min_{\boldsymbol{\eta}(t),\boldsymbol{P}(t)} \quad & \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{N}} \bar{P}_k^n \\
\text{subject to} \quad & (7), (10), \text{and } (11), \\
& \lim_{C\to\infty} \frac{1}{C} \sum_{t=0}^{C-1} R_k(t) > A, \forall k \in \mathcal{K}, && \text{(12a)} \\
& \sum_{n\in\mathcal{N}} \eta_k^n(t) P_k^n(t) \leq P_{\max}, \ \forall k \in \mathcal{K}, && \text{(12b)} \\
& 0 \leq P_k^n(t) \leq P_{\max} \mathbb{1}\{\eta_k^n(t) = 1\}, \\
& \qquad \forall t, \, k \in \mathcal{K}, \, n \in \mathcal{N}, && \text{(12c)} \\
& \eta_k^n(t) \in \{0, 1\}, \ \forall t, \, k \in \mathcal{K}, \, n \in \mathcal{N}, && \text{(12d)}
\end{aligned}$$

where $\bar{P}_k^n = \lim_{C\to\infty} \frac{1}{C} \sum_{t=0}^{C-1} P_k^n(t)$ is the time-averaged power consumption of VUE pair k over RB n, and $\mathbb{1}\{.\}$ is the indicator function. Additionally, constraint (12a) ensures queue stability. In order to find the optimal resource $\boldsymbol{\eta}(t)$ and power $\boldsymbol{P}(t)$ allocation vectors that solve problem $\mathbb{P}_1$,

$$J_k^{(X)}(t+1)=\max\left(J_k^{(X)}(t)+(X_k(t)-H)\mathbb{1}\left\{Q_k(t)>R_k(t)-\left(2-(\frac{d}{\tau}-1)A\right)\right\},0\right), \tag{14}$$

$$J_k^{(Y)}(t+1)=\max\left(J_k^{(Y)}(t)+(Y_k(t)-B)\mathbb{1}\left\{Q_k(t)>R_k(t)-\left(2-(\frac{d}{\tau}-1)A\right)\right\},0\right), \tag{15}$$

$$J_k^{(R)}(t+1)=\max\left(J_k^{(R)}(t)-R_k(t)+A,0\right), \tag{16}$$

$$J_k^{(Q)}(t+1)=\max\left(J_k^{(Q)}(t)+R_k(t)\mathbb{1}\left\{Q_k(t)>R_k(t)-\left(2-(\frac{d}{\tau}-1)A\right)\right\}-R_k(t)\epsilon_k,0\right). \tag{17}$$

we invoke techniques from Lyapunov stochastic optimization [22].

IV. AoI-Aware Resource Allocation

A. Lyapunov Optimization Framework

We first rewrite (7) as

$$\lim_{C\to\infty}\frac{1}{C}\sum_{t=0}^{C-1}R_k(t)\mathbb{1}\left\{Q_k(t)>R_k(t)-\psi\right\}\leq\bar{\epsilon}_k, \tag{13}$$

where $\bar{\epsilon}_k=\lim_{C\to\infty}\frac{1}{C}\sum_{t=0}^{C-1}R_k(t)\epsilon_k$ is the product of the time-averaged rate and the tolerance value. Using Lyapunov optimization, the time-averaged constraints (10), (11), (12a), and (13) can be satisfied by converting them into virtual queues and maintaining their stability [22]. In this regard, we introduce the corresponding virtual queues with the dynamics shown in (14)–(17). Denoting $\boldsymbol{J}(t)=\left\{J_k^{(X)}(t),J_k^{(Y)}(t),J_k^{(Q)}(t),J_k^{(R)}(t),Q_k(t):k\in\mathcal{K}\right\}$ as the combined physical and virtual queue vector, the conditional Lyapunov drift-plus-penalty for slot t is given by

$$\mathbb{E}\left[\mathscr{L}\left(\boldsymbol{J}(t+1)\right)-\mathscr{L}\left(\boldsymbol{J}(t)\right)+\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}}VP_k^n(t)|\boldsymbol{J}(t)\right], \tag{18}$$

where $\mathscr{L}\left(\boldsymbol{J}(t)\right)=\frac{\boldsymbol{J'}(t)\boldsymbol{J}(t)}{2}$ is the Lyapunov function. Here, $V\geq 0$ is a parameter that controls the trade-off between power consumption and queue stability. By leveraging the fact that $\left(\max\left(a-b,0\right)+c\right)^2\leq a^2+b^2+c^2-2a(b-c)$, $\forall a,b,c\geq 0$, and $\left(\max\left(x,0\right)\right)^2\leq x^2$ on (2) and (14)–(17), an upper bound on (18) can be obtained as follows:

$$\begin{aligned}(18)\leq\mathfrak{C}+\mathbb{E}\Bigg[\sum_{k\in\mathcal{K}}\Bigg(&\Big(J_k^{(Q)}(t)-J_k^{(X)}(t)-2\left(Q_k(t)+\psi\right)^3\\&-\left(2J_k^{(Y)}(t)+1\right)\left(Q_k(t)+\psi\right)\Big)\cdot\mathbb{1}\left\{Q_k(t)>R_k(t)-\psi\right\}\\&-\left(J_k^{(R)}(t)+A+Q_k(t)+J_k^{(Q)}(t).\epsilon_k\right)\Bigg)R_k(t)\\&+\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}}VP_k^n(t)|\boldsymbol{J}(t)\Bigg].\end{aligned} \tag{19}$$

Here, $\mathfrak{C}$ is a bounded term that does not affect the system performance.[1] Note that the solution to problem $\mathbb{P}_1$ can be obtained by minimizing the upper bound in (19) in each slot t [22], i.e.,

$$\begin{aligned}\mathbb{P}_2:\min_{\boldsymbol{\eta}(t),\mathbf{P}(t)}\sum_{k\in\mathcal{K}}\Bigg[\sum_{n\in\mathcal{N}}VP_k^n(t)-\Bigg[&J_k^{(R)}(t)+A+Q_k(t)\\&+J_k^{(Q)}(t).\epsilon_k+\Big(-J_k^{(Q)}(t)+J_k^{(X)}(t)\\&+(2J_k^{(Y)}(t)+1)(Q_k(t)+\psi)+2(Q_k(t)+\psi)^3\Big)\\&\times\mathbb{1}\left\{Q_k(t)>R_k(t)-\psi\right\}\Bigg]R_k(t)\Bigg]\end{aligned}$$

subject to (12b) - (12d).

To solve $\mathbb{P}_2$ in each time slot t, the RSU needs full global channel state information (CSI) and queue state information (QSI). This is clearly impractical for vehicular networks since frequently exchanging fast-varying local information between the RSU and VUEs can yield a significant overhead which is not acceptable. To alleviate the information exchange burden, we propose a two-timescale resource allocation mechanism which is performed in two stages. Briefly speaking, RBs are allocated over a long timescale at the RSU whereas each VUE pair decides its transmit power over a short timescale.

B. Two-Stage Resource Allocation

1) Spectral Clustering and RB Allocation at the RSU: Before allocating RBs to VUE pairs, we note that if nearby vehicles transmit over the same RBs, co-channel transmission can lead to severe interference. In order to avoid the interference from nearby VUEs, the RSU first clusters VUE pairs into $g>1$ disjoint groups, in which the nearby VUE pairs are allocated to the same group, and then orthogonally allocates all RBs to the VUE pairs in each group. Vehicle clustering is done by means of spectral clustering [23]. We adopt the VUE clustering and RB allocation technique as in [13], denoting $\boldsymbol{v}_k\in\mathbb{R}^2$ as the Euclidean coordinate of the midpoint of VUE transmitter-receiver pairs k, we use a distance-based Gaussian similarity matrix $\boldsymbol{S}$ to represent the geographic proximity information, in which the (k,k')-th element is defined as

$$s_{kk'}:=\begin{cases}e^{-\|\boldsymbol{v}_k-\boldsymbol{v}_{k'}\|^2/\gamma^2}, & \|\boldsymbol{v}_k-\boldsymbol{v}_{k'}\|\leq\phi,\\ 0, & \text{otherwise,}\end{cases}$$

where ϕ captures the neighborhood size, while γ controls the impact of the neighborhood size. Subsequently, $\boldsymbol{S}$ is used

[1]Due to the space limitations, the derivations are omitted. The interested reader may refer to [22] for more details.

to group VUE pairs using spectral clustering as shown in Algorithm 1. After forming the groups, the RSU orthogonally allocates RBs to the VUEs inside the group. Hereafter, we denote the set $\mathcal{N}_k$ as the allocated RB of each VUE pair $k \in \mathcal{K}$.

Algorithm 1 Spectral Clustering for VUE Grouping

1: Calculate matrix $\boldsymbol{S}$ and the diagonal matrix $\boldsymbol{D}$ with the diagonal $d_j = \sum_{q=1}^{K} s_{jq}$.
2: Let $\boldsymbol{U} = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_g]$ in which $\boldsymbol{u}_g$ is the eigenvector of the g-th smallest eigenvalue of $\boldsymbol{I} - \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{S}\boldsymbol{D}^{-\frac{1}{2}}$.
3: Numerically, use the k-means clustering approach to cluster K normalized row vectors (which represent K VUE pairs) of matrix $\boldsymbol{U}$ into g groups.

Moreover, to overcome the signaling overhead issue caused by frequent information exchange between the RSU and VUE pairs, the RSU clusters VUE pairs and allocate RBs in a longer time scale, i.e., every $T_0 \gg 1$ time slots, since the geographic locations of the vehicles do not change significantly during the slot duration τ (i.e., coherence time of fading channels). In such, VUE pairs send their geographic locations to the RSU only once every T_0 slots instead of every slot.

2) Power Allocation at the VUE: Since VUE pair k can only use the set $\mathcal{N}_k$ of allocated RBs to send information, we modify the power allocation and RB usage constraints, i.e., (12b)–(12d), $\forall k \in \mathcal{K}$, as

$$\begin{cases} \sum\limits_{n\in\mathcal{N}_k} P_k^n(t) \leq P_{\max},\ \forall t, \\ P_k^n(t) \geq 0,\ \forall t,\, n \in \mathcal{N}_k, \\ P_k^n(t) = 0,\ \forall t,\, n \notin \mathcal{N}_k. \end{cases} \tag{20}$$

Now the RBs in $\mathcal{N}_k$ are reused by distant VUE transmitters in different groups. We approximately treat the aggregate interference as a constant term I and rewrite the transmission rate as

$$(1) \approx \frac{\tau}{Z} \sum_{n\in\mathcal{N}_k} \omega \log_2\left(1 + \frac{P_k^n(t) h_{kk}^n(t)}{N_0\omega + I}\right). \tag{21}$$

Subsequently, applying (20) and (21) to $\mathbb{P}_2$, the VUE transmitter of each VUE pair k locally allocates its transmit power by solving the following convex optimization problem $\mathbb{P}_3$ in each slot t:

$$\mathbb{P}_3: \min_{P_k^n(t)} \sum_{n\in\mathcal{N}_k} V P_k^n(t) - \Im_k(t) \log_2\left(1 + \frac{P_k^n(t) h_{kk}^n(t)}{N_0\omega + I}\right)$$
subject to (20),

where $\Im_k(t) = \frac{\tau\omega}{Z}\Big[J_k^{(R)}(t) + A + Q_k(t) + J_k^{(Q)}(t)\epsilon_k + \Big(-J_k^{(Q)}(t) + J_k^{(X)}(t) + (2J_k^{(Y)}(t)+1)(Q_k(t)+\psi) + 2(Q_k(t)+\psi)^3\Big) \times \mathbb{1}\{Q_k(t) > R_k(t) - \psi\}\Big]$. Based on the Karush-Kuhn-Tucker (KKT) conditions, the optimal VUE transmit power $P_k^{n*}(t), \forall n \in \mathcal{N}_k$, of $\mathbb{P}_3$ satisfies,

$$\frac{\Im_k(t) h_{kk}^n(t)}{(N_0\omega + I + P_k^{n*}(t) h_{kk}^n(t)) \ln 2} = V + \zeta,$$

if $\frac{\Im_k(t) h_{kk}^n(t)}{(N_0\omega+I)\ln 2} > V + \zeta$. Otherwise, $P_k^{n*}(t) = 0$. Moreover, the Lagrange multiplier ζ is 0 if $\sum_{n\in\mathcal{N}_k} P_k^{n*}(t) < P_{\max}$, and we have $\sum_{n\in\mathcal{N}_k} P_k^n(t) = P_{\max}$ when $\zeta > 0$. Note that, given a small value of V, the derived power $P_k^{n*}(t)$ provides a sub-optimal solution to problem $\mathbb{P}_1$, whose optimal solution is asymptotically obtained by increasing V. After sending its status update, VUE pair k updates the physical and virtual queues as per (2) and (14)–(17).

Table I
SIMULATION PARAMETERS [13].

Parameter	Value	Parameter	Value
N	20	H	0.8334
ω	180 KHz	B	0.7576
τ	3 ms	g	10
$P_{\max}$	23 dBm	γ	30 m
Z	500 Byte	ϕ	150 m
N_0	−174 dBm/Hz	T_0	100
Arrival rate	0.5 Mbps	α	1.61
d	60 ms	$\mathscr{D}$	15 m
ϵ_k	0.001	l_0	−68.5 dB
ψ	−3.25	l_0'	−54.5 dB
V	0		

V. Simulation Results and Analysis

For our simulations, we use a 250×250 m^2 area Manhattan mobility model as in [8], [13]. The average vehicle speed is 60 km/h, and the distance between the transmitter and receiver of each VUE pair is 15 m. Unless stated otherwise, the remaining parameters are listed in Table I. The performance of our proposed solution is compared to that of [13], where a power minimization is considered subject to a reliability measure in terms of maximal queue length among all VUEs.

In Fig. 2, we verify the accuracy of using EVT to characterize the distribution of the excess value $X_k(t)|_{Q_k(t)>R_k(t)-\psi} = Q_k(t) - R_k(t) + \psi$. In particular, Fig. 2 shows the complementary cumulative distribution function (CCDF) of the excess value for various densities of VUEs, along with the GPD distributions with the specified parameters. A near-perfect fitting can be noted, which verifies the accuracy of using EVT to characterize the distribution of the excess value.

In Fig. 3, we show the CCDF of the queue length of all VUEs, in which our approach achieves a better distribution in terms of the tail, compared to the baseline with a reduction in the worst queue length by more than two-fold for all cases. The AoI distribution, which is the main reliability measure in our model is shown in Fig. 4 in terms of its CCDF for various densities of VUEs. Again, the AoI distribution in our approach outperforms that of the baseline with a much shorter tail (higher reliability). By inspecting both Fig. 3 and Fig. 4, we can observe that both queue length and AoI distributions follow the same trend, which verify the proposed mapping between the probabilistic AoI and the queue length (Lemma 1). It is worth mentioning that, these sudden changes in the distributions are due to the adopted path loss model, ranging from LOS, WLOS and NLOS. To further validate this, we have varied the distance between the transmitter and receiver

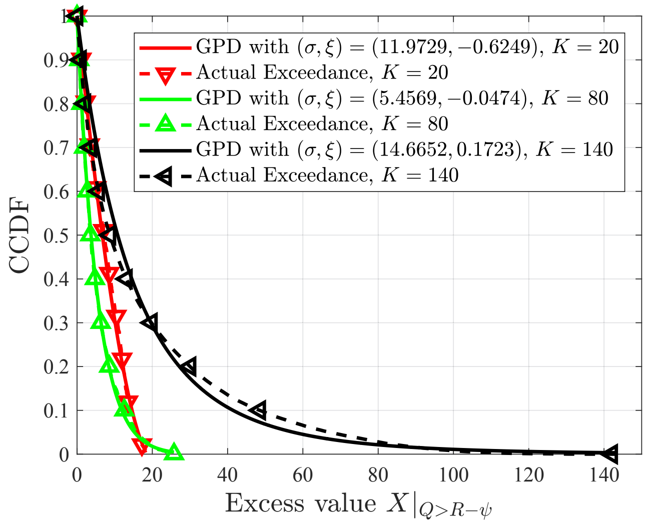

Figure 2. CCDF of the exceedance value fitted to GPD for various densities of VUEs (K).

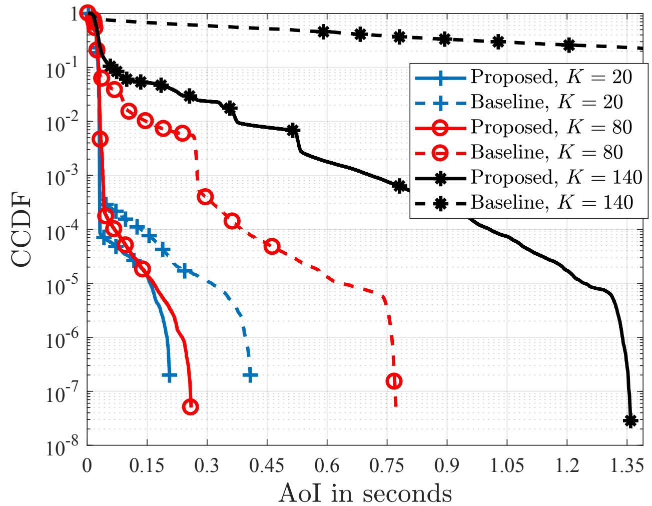

Figure 4. CCDF of the AoI for various densities of VUEs (K).

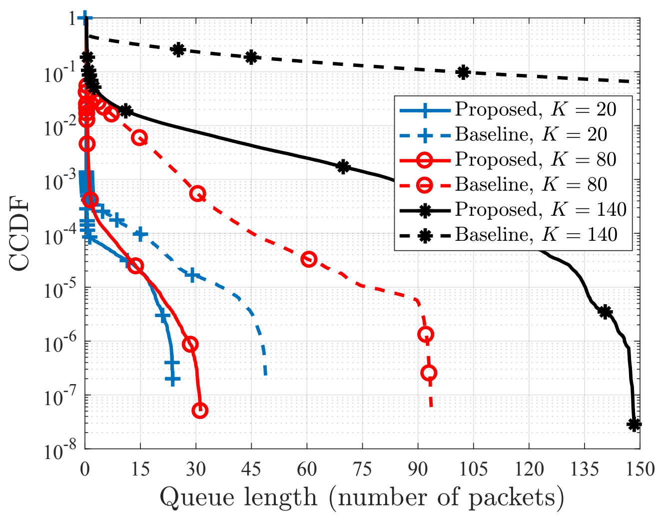

Figure 3. CCDF of the queue length for various densities of VUEs (K).

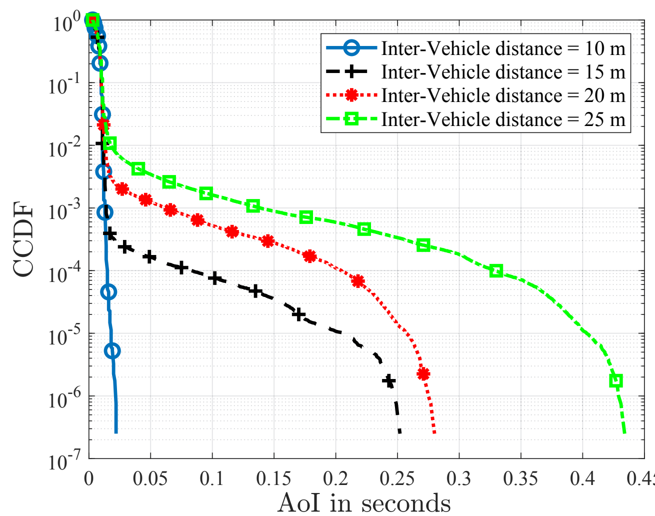

Figure 5. CCDF of the AoI for various inter-vehicle distance with $K = 80$.

of each VUE pair from 10 m to 25 m. Fig. 5 shows that the tail of AoI distribution decreases as the inter-vehicle distance decreases until it vanishes at a 10 m distance.

Finally, in Fig. 6 we study the impact of the arrival rate of the status updates for two different VUE densities, $K = 80$ and $K = 20$, for both average and worst AoI. When the arrival rate is small, the network incurs higher average AoI due to the lack of status updates sent by VUEs. Hence, increasing the arrival rate will reduce that lack of status updates, leading to a better average AoI, up to a certain point, after which the average AoI starts increasing again due to interference. Note that, at low VUE density ($K = 20$) the network can withstand higher arrival rates. Fig. 6 also highlights that the increase of the worst AoI with the arrival rate shows a slight increase up to the aforementioned optimal arrival rate for the average AoI, and exhibits a rapid increase afterwards. Moreover, we note that the worst AoI is about 10 fold higher than the average AoI. Although the AoI achieves a small average, the worst AoI is heavy-tailed. In this situation, relying on the average AoI is inadequate for ensuring URLLC. This discrepancy between these two metrics demonstrates that the tail characterization is instrumental in designing and optimizing URLLC-enabled V2V networks.

VI. Conclusion

In this paper, we have studied the problem of ultra-reliable and low-latency vehicular communication. For this purpose, we have defined a new reliability measure, in terms of probabilistic AoI. Second, we have established a novel relationship between the AoI and queue length probability distributions. Then, we have shown that characterizing the queue length tail

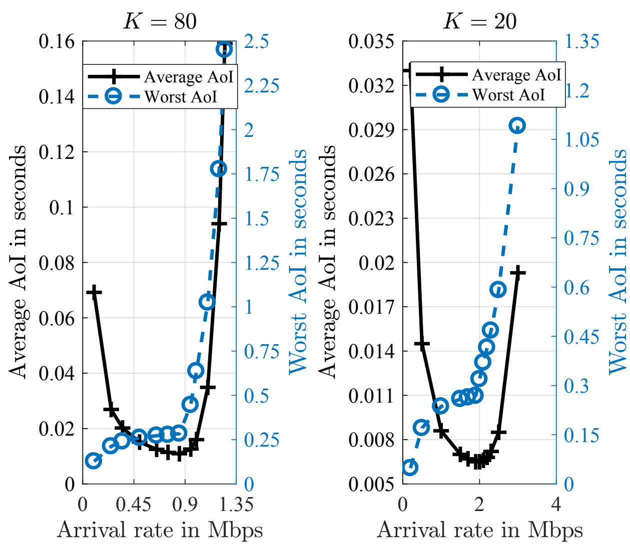

Figure 6. Arrival rate versus AoI trade-off, with $K = 80$ and $K = 20$ VUEs.

distribution can be effectively done using EVT. Subsequently, we formulated the problem as a power minimization problem subject to the probabilistic AoI and solved it using Lyapunov optimization. Simulation results show that the proposed approach yields significant improvements in terms of AoI and queue length, when compared to a baseline that only considers the maximum queue among vehicles. In our future work, we will extend the current framework to other use cases such as autonomous aerial vehicles (UAVs) [24], [25] and industrial control [26]. Another interesting extension pertains to studying the impact of short packets and finite blocklength on our framework and its performance [27].

Acknowledgment

This work was supported in part by the Academy of Finland project CARMA, and 6Genesis Flagship (grant no. 318927), in part by the INFOTECH project NOOR, in part by the U.S. National Science Foundation under Grants CNS-1513697 and CNS-1739642, and in part by the Kvantum Institute strategic project SAFARI.

References

[1] 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, May 2013.

[2] T. Zeng, O. Semiari, W. Saad, and M. Bennis, "Joint communication and control for wireless autonomous vehicular platoon systems," *CoRR*, vol. arXiv/1804.05290, 2018.

[3] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, "5g for vehicular communications," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 111–117, Jan. 2018.

[4] ETSI EN Std 302 637-2, "Intelligent transport systems; vehicular communications; basic set of applications; part 2: Specification of cooperative awareness basic service," Aug. 2013.

[5] M. Bennis, M. Debbah, and H. V. Poor, "Ultra-reliable and low-latency wireless communication: Tail, risk and scale," *Proc. IEEE*, 2018, to be published.

[6] S. Coles, *An introduction to statistical modeling of extreme values*. Springer, 2001.

[7] M. I. Ashraf, M. Bennis, C. Perfecto, and W. Saad, "Dynamic proximity-aware resource allocation in vehicle-to-vehicle (V2V) communications," in *Proc. IEEE Global Commun. Conf. Workshops*, Dec. 2016, pp. 1–6.

[8] M. I. Ashraf, C.-F. Liu, M. Bennis, and W. Saad, "Towards low-latency and ultra-reliable vehicle-to-vehicle communication," in *Proc. Eur. Conf. Netw. Commun.*, Jun. 2017, pp. 1–5.

[9] T. Liu, Y. Zhu, R. Jiang, and Q. Zhao, "Distributed social welfare maximization in urban vehicular participatory sensing systems," *IEEE Trans. Mobile Comput.*, vol. 17, no. 6, pp. 1314–1325, Jun. 2018.

[10] J. Mei, K. Zheng, L. Zhao, Y. Teng, and X. Wang, "A latency and reliability guaranteed resource allocation scheme for LTE V2V communication systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3850–3860, Jun. 2018.

[11] C. Perfecto, J. D. Ser, and M. Bennis, "Millimeter-wave v2v communications: Distributed association and beam alignment," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2148–2162, Sep. 2017.

[12] A. Mouradian, "Extreme value theory for the study of probabilistic worst case delays in wireless networks," *Ad Hoc Networks*, vol. 48, pp. 1–15, Sep. 2016.

[13] C.-F. Liu and M. Bennis, "Ultra-reliable and low-latency vehicular transmission: An extreme value theory approach," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1292–1295, Jun. 2018.

[14] J. Champati, H. Al-Zubaidy, and J. Gross, "Statistical guarantee optimization for age of information for the D/G/1 queue," in *Proc. IEEE Conf. Comput. Commun. Workshops*, Apr. 2018, pp. 130–135.

[15] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, Jun. 2011, pp. 350–358.

[16] A. Baiocchi and I. Turcanu, "A model for the optimization of beacon message age-of-information in a VANET," in *Proc. IEEE 29th Int. Teletraffic Congress*, Sep. 2017, pp. 108 – 116.

[17] B. Zhou and W. Saad, "Joint status sampling and updating for minimizing age of information in the internet of things," *CoRR*, vol. arXiv/1807.04356, 2018.

[18] M. A. Abd-Elmagid and H. S. Dhillon, "Average age-of-information minimization in UAV-assisted IoT networks," *CoRR*, 2018.

[19] F. Bai, N. Sadagopan, and A. Helmy, "IMPORTANT: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks," in *Proc. IEEE Conf. Comput. Commun.*, vol. 2, Mar. 2003, pp. 825–835.

[20] M. Abdulla and H. Wymeersch, "Fine-grained vs. average reliability for V2V communications around intersections," in *Proc. IEEE Global Commun. Conf. Workshops*, Dec. 2017, pp. 1–7.

[21] C.-F. Liu, M. Bennis, and H. V. Poor, "Latency and reliability-aware task offloading and resource allocation for mobile edge computing," in *Proc. IEEE Global Commun. Conf. Workshops*, Dec. 2017, pp. 1–7.

[22] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan and Claypool Publishers, Jun. 2010.

[23] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, Dec. 2007.

[24] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3949–3963, Jun. 2016.

[25] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience," *IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on Human-In-The-Loop Mobile Networks*, vol. 35, no. 5, pp. 1046–1061, May 2017.

[26] O. N. C. Yilmaz, Y. . E. Wang, N. A. Johansson, N. Brahmi, S. A. Ashraf, and J. Sachs, "Analysis of ultra-reliable and low-latency 5G communication for a factory automation use case," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, Jun. 2015, pp. 1190–1195.

[27] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.