Article:

Bute, MS, Fan, P, Luo, Q et al. (2 more authors) (2023) QoS-Aware Content Dissemination Based on Integrated Social and Physical Attributes Among Cellular and V2V Users. IEEE Transactions on Vehicular Technology. ISSN 0018-9545

https://doi.org/10.1109/TVT.2023.3267793

QoS-Aware Content Dissemination Based on Integrated Social and Physical Attributes among Cellular and V2V Users

Muhammad Saleh Bute, *Student Member, IEEE,* Pingzhi Fan, *Fellow, IEEE,*
Quyuan Luo, Li Zhang, *Senior Member, IEEE,* and Fakhar Abbas, *Senior Member, IEEE*

***Abstract*—With the recent advancement in intelligent transportation systems (ITS), the vehicular network has become the major force in content dissemination, and it is essential to improve efficiency in content delivery and quality and service (QoS). The rapid growth in vehicular data traffic and limited resources in the vehicular network have created a bottleneck in content dissemination. In addition, the demand of high data rate, low latency and user's satisfaction on content make content dissemination more challenging. To address the above problem which involves vehicle to vehicle (V2V) peer discovery and resource allocation, we propose an efficient content dissemination scheme which integrates both social and physical attributes of the users. V2V users are paired to share content based on the similarity in their preferences and link reliability, while considering the QoS of vehicle to network (V2N) users in both uplink and downlink phases. Moreover, we present an improved matching algorithm for channel allocation, which maximizes the sum rate of V2V pairs weighted by the intensity of their social relationship. Simulation results demonstrate the performance of the proposed scheme in terms of weighted sum rate and user's satisfaction on content.**

***Index Terms*—Vehicle to everything (V2X), content, resource block, agumented reality (AR).**

I. Introduction

WITH the emergence of the intelligent transportation system (ITS), many vehicular applications are introduced [1], increasing the demand for high data rates in the vehicular ad-hoc networks (VANETs). The huge demand on the limited wireless network bandwidth signifies the need for effective resource allocation. The vehicle to everything (V2X) technology, is one of the major features of the fifth generation (5G) network that supports the vehicular communications framework. The third generation partnership project (3GPP) presents a long-term evolution (LTE) network architecture that supports the V2X services, which is mainly categorized into vehicle to vehicle (V2V) communication and vehicle to network (V2N) communication. The existing cellular networks augmented with the V2V communications enable direct communication between nodes without traversing through the eNodeB. The V2V communication paradigm significantly improves the performance of V2V and V2N links in terms of resource management, quality of service requirement (QoS), channel capacity and minimization of the effect of vehicular mobility [2], [3]. To utilize the limited bandwidth in the wireless network, researchers in academia and industries attempt to exploit new technologies to improve efficiency and capacity of the cellular network. The cellular V2X network enables spectrum reuse between the V2V and V2N users in the licensed cellular band under control of the eNodeB [4], [5]. Thanks to the proximity of the V2V communication, it provides high data rates, and minimizes end to end latency and power consumption [6]. With the rapid increase in the number of connected vehicles [7] the global vehicular data traffic is growing exponentially due to data intensive applications such as digital map navigation systems, augmented reality (AR) aided driving, autonomous driving, games and other infotainment applications. Conventionally, content dissemination is handled by the V2N communication link [8]. However, the cellular network is getting overloaded and congested with the increasing demand for content dissemination. The V2V communication can support the V2N communication for efficient content dissemination among the vehicular nodes. Vehicles are operated by humans who naturally have social attributes such as preferences. The social relationship between users involves discovering the interaction pattern among users with common interest through the information they share i.e., social networks, computation offloading, traffic information, and the road situations etc. With the adoption of social networks and microblogging sites as an integral part of our lives, the social attributes of users can be obtained from the historical data of social platforms. The information obtained can be explored to improve network performance and QoS in order to satisfy the users' demands. Several vehicles might request for the same content such as map guide or local news, then the eNodeB has to transmit the content to the requesting vehicles through multiple re-transmissions [9]. Using the V2V communication, vehicles with similar request can be paired together to share content. This approach not only relieves the huge data traffic in the cellular infrastructure but also increases spectrum

This work was supported by the NSFC Project No.61731017 and No.62020106001. The work of Quyuan Luo was supported by NSFC Project No.62101463

Muhammad Saleh Bute, Pingzhi Fan, and Quyuan Luo are with the Key Laboratory of Information Coding and Transmission, School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China (e-mail: msbute@my.swjtu.edu.cn; pzfan@swjtu.edu.cn; qyluo@swjtu.edu.cn.).

Li Zhang is with the Institute of Communication and Power Networks, School of Electrical and Electronic Engineering, University of Leeds, Leeds LS2 9JT, U.K (email: l.x.zhang@leeds.ac.uk).

Fakhar Abbas is with the Centre for Trusted Internet & Community (CITIC), National University of Singapore (NUS), Singapore (email:fakhar.14@hotmail.com).

efficiency. Content dissemination schemes were presented in [10] and [17], where a central vehicle is selected to share content to other vehicles through V2V communication. Due to the centrality of the content transmitter, information can be disseminated within a short period of time. However, in reality each user's preference might differ on content. For example, two vehicles request for local news from the eNodeB. Vehicle A may be interested in only the news headline, while Vehicle B prefers the entire news presentation. Consideration of only physical attributes such as distance between V2V pair and centrality in peer discovery might not pair the optimal V2V users to achieve efficient content dissemination. Thus, to improve QoS and user satisfaction we have to consider some social information of users such as preferences, social ties, similarity and link reliability based on their historical data. This approach will facilitate effective peer discovery for V2V content sharing. The work in [6] and [12] consider both physical and social attributes in their proposed schemes. They only consider social ties between users for content distribution. However, without considering link reliability between users with similar preferences, content might not be delivered successfully.

In this paper, we propose a QoS aware content dissemination based on integrated social and physical attributes among cellular and V2V users. In the physical domain unlike the works in [7], [11], [18] we consider spectrum reuse in both uplink and downlink phase. To minimize mutual interference due to channel sharing, V2V users with minimum interference level are grouped in to a cluster. In the social domain, different from previous studies [4], [5], [8] in order to improve users' satisfaction, we jointly consider user preference on content and link reliability in the peer discovery process. The Bayesian nonparametric statistics is applied in estimating the probability of selecting similar content among users. An optimization problem is formulated to maximize the weighted sum rate of V2V users while considering QoS of V2N users. The problem is solved in three phases, which include peer discovery, power control, and channel assignment. The main contributions of this paper are summarized as follows

- *Model*: A QoS aware content dissemination scheme based on integrated social and physical attributes among cellular and V2V users is proposed. Leveraging the direct D2D paradigm, the V2V link augments V2N link in content sharing. V2V peers are established based on physical and social attributes. A non-parametric model is applied on the historic data of users to obtain the intensity of social relationship between users. To improve reliability, spectrum reuse is considered in both uplink and downlink channels.
- *Optimization Problem*: To maximize user satisfaction on content, we consider similarity in preference and link reliability between users while establishing V2V pair. We then formulate the objective function of the optimization problem as a weighted sum rate, which is weighted by the degree of social similarity and link reliability between V2V users.
- *Algorithm Design*: An improved matching algorithm is proposed for channel assignment. To enhance spectral efficiency by resource sharing among cellular and V2V users, interference is minimized by grouping vehicles into clusters. Vehicles with minimum interference and relative velocity are grouped in to a cluster.
- *Validation*: Based on real world scenarios, simulation results demonstrate the performance of the proposed scheme in comparison with the existing schemes.

The remainder of this paper is organized as follows. Section II presents the related work. Section III describes the proposed system model. Section IV presents problem formulation. Section V presents the proposed efficient content dissemination scheme. Section VI presents simulation results and discussions. Finally, Section VII concludes the paper.

II. Related Work

Most content dissemination schemes consider only the physical layer data [7], [10], [11], [15] such as vehicle centrality, the distance between V2V pair, and the mobility path of the vehicles. This approach enhances content distribution to many users as possible. However, social attributes such as content delivery success and the preference of users on contents are vital metrics that need to be considered. To improve content delivery performance, the proposed scheme jointly considers physical and social attributes.

A. Physical Attribute-based Schemes

The physical attribute-based content dissemination schemes aim at minimizing content delivery latency. In [10], the authors presented a content dissemination scheme based on the vehicular cloud to minimize cost and delay. The authors exploit the name-centric mechanism in content delivery. Q. Luo et al. [17] proposed an intelligent algorithm-based content dissemination in vehicular networks. In the proposed scheme, to improve system utility, dual importance evaluation was applied in determining users' priority on content. At the same time, a fuzzy logic algorithm was used in selecting the optimal number of seed vehicles for onward re-distribution. The authors in [11] studied a proactive content seeding mechanism where contents are cached on some selected vehicles for efficient dissemination. To maximize the system throughput of content dissemination, [15] proposed a joint V2N and V2V content distribution and scheduling in vehicular networks. Content transmitting vehicles are selected based on a utility function using a content delivery path selection algorithm. R. Hou et al. [7] presented a vehicle tracking content dissemination scheme in vehicular networks. The authors exploit the Tabu search mechanism in tracking target vehicle's mobility, hence, maximizing content delivery. To minimize content delivery latency and network traffic, A. Pal et al. [18] proposed a neighborhood caching and interest content dissemination scheme. The authors exploited a low complexity Bloom Filter algorithm to optimally cache and remove contents in the network. However, all the research works in this section consider only the physical layer data and do not utilize the users' social information. In contrast to the above articles, in the proposed scheme to enhance efficiency, we consider interference management and power control by grouping V2V pairs into clusters based on minimum interference. Moreover, to improve reliability, content can be distributed in both uplink and downlink channels.

B. Physical and Social Attribute-based Schemes

Some researchers explore physical and social attributes to improve content dissemination in vehicular networks. Xiabo et al. [4] presented an entity and sociality-aware content dissemination solution in vehicular networks. The authors consider entity and sociality trust among users to enhance content distribution, and a fuzzy logic algorithm was applied to obtain a tradeoff between social and entity trust. The authors in [5] investigated a social aware content delivery scheme in mobile edge computing, where content dissemination and caching were jointly considered using the social relationship among users. Similarly, the work in [13] studied social trust-based emergency message dissemination in vehicular networks. To avoid the spread of distorted content by malicious users, the authors exploit social utility in identifying trustworthy users to participate in content distribution. In [6], the authors presented a double-layer game theory-based socially aware content caching and distribution scheme. Their approach selects some strategically located users and allocates content to them for subsequent transmission to the requesting nodes. The social attribute of users was exploited to generate the utility function for content allocation and distribution. N. Vo et al. [12] proposed a social-based spectrum sharing and caching helper selection strategy. The authors applied a genetic algorithm in selecting optimal nodes based on social layer information for content caching and sharing. L. Yao et al. [8] studied social aware cooperative caching and content dissemination in vehicular networks. Contents are cached and distributed based on users' social similarity and mobility patterns. The authors explore a hidden Markov algorithm to predict the mobility of vehicles. The authors in [14] proposed sociality-based reliable content dissemination for vehicular networks. For efficient content distribution, the authors applied the Weiner process to estimate connection probability between users and convolutional neural networks was used to obtain the social ties between users.

However, most works focus only on social similarity in V2V pairing, which might not be the optimal combination. Link reliability is an important metric that needs to be considered because even if users with similar preferences are paired, without a reliable link, content delivery may fail. The authors in [14] considered connection probability among users using neural networks, the centralized algorithm they proposed incurs excessive computational overhead, which is unsuitable for vehicular networks. Unlike the research works discussed above, we propose a distributed algorithm that jointly considers user preference on content and link reliability to minimize cost and improve efficiency.

III. System Model

We consider a single cell along a multilane road, and the vehicles moving on the road form a vehicular network which consists of N V2V pairs comprising the transmitting vehicles denoted as $VT = \{VT_1, VT_2, VT_3, ..., VT_N\}$, and the receiving vehicles denoted as $VR = \{VR_1, VR_2, VR_3, ..., VR_N\}$. We also have M V2N users denoted as $CU = \{CU_1, CU_2, CU_3, ..., CU_M\}$, and the total available bandwidth in the system is divided in to K resource blocks denoted as $RB = \{RB_1, RB_2, RB_3, ..., RB_K\}$. We consider a cellular V2X network as illustrated in Fig. 1, where V2V and V2N operates in the licensed cellular band [16]. To improve the spectral efficiency, both the uplink and downlink spectrum of cellular users are re-used for V2V communication. Typically, the number of V2V users are higher than the V2N users, hence resource block is orthogonally allocated to V2N users. Vehicles are grouped to form ϕ clusters denoted as $ZH = \{ZH_1, ZH_2, ZH_3, ..., ZH_\phi\}$ where $\phi \leq K$. Each V2N user shares a resource block with multiple V2V users. To minimize interference, only V2V users of the same cluster can share a resource block.

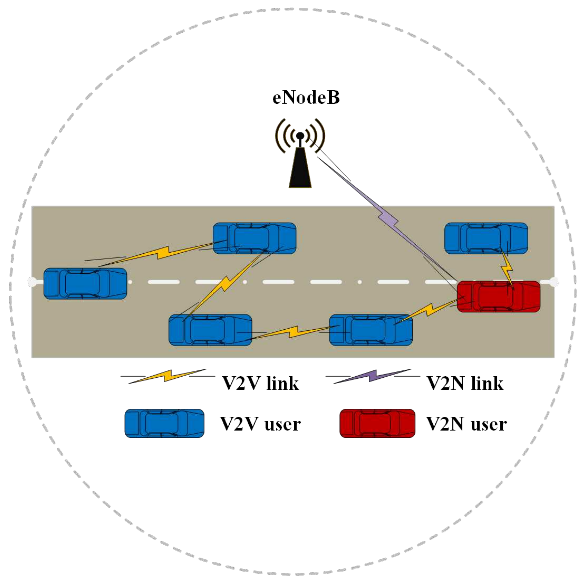


Fig. 1: System Model

For security and privacy, a central controller at the eNodeB coordinates the V2V peer discovery and content dissemination. The credentials of the new and existing vehicles are registered, only authenticated vehicles are eligible to request for information or participate in content sharing.

A. Physical Domain

For the proposed scheme, we utilize both the uplink and downlink channels for spectrum re-use. Since interference is caused by the resource sharing, for the channel modeling we consider both small scale and large-scale fadings.

1) Uplink Channel Transmission Rate

The recieved signal to interference and noise ratio (SINR) of a transmission link between mth V2N user and the eNodeB [17] can be expressed as

$$UL^c_{m,e} = \frac{P^c_m g^c_{m,e}}{\sum\limits_{i \in VT} P^d_i g^d_{i,e} + \sigma^2} \tag{1}$$

TABLE I: Summary of Related Works

	Parameters	References					
		[5]	**[8]**	**[12]**	**[13]**	**[14]**	**Proposed Scheme**
S/N	**Physical Attribute**						
1	Distance between V2V pair	Yes	Yes	Yes	Yes	Yes	Yes
2	Interference management	No	No	Yes	No	No	Yes
3	Power control	No	No	No	No	No	Yes
4	Uplink channel distribution	Yes	Yes	No	Yes	Yes	Yes
5	Downlink channel distribution	No	No	Yes	No	No	Yes
S/N	**Social Attribute**						
1	Content similarity	Yes	Yes	No	Yes	Yes	Yes
2	User satisfaction	No	No	No	No	No	Yes
3	Link reliability	No	Yes	No	No	Yes	Yes
4	Social relationships	Yes	Yes	Yes	Yes	Yes	Yes

TABLE II: Main Notation Summary

Symbol	Description
CU	Set of V2N users
VT	Set of V2V transmitters
VR	Set of V2V receivers
RB	Set of resource blocks
$UL^c_{m,e}$	The SINR of mth V2N user in uplink phase
$UL^d_{i,j}$	The SINR of ith V2V user in uplink phase
RC^{UL}_m	The data transmission rate of mth V2N user in uplink phase
RD^{UL}_n	The data transmission rate of ith V2V user in uplink phase
$DL^c_{m,e}$	The SINR of mth V2N user in downlink phase
$DL^d_{i,j}$	The SINR of ith V2V user in downlink phase
RC^{DL}_m	The data transmission rate of mth V2N user in downlink phase
RD^{DL}_n	The data transmission rate of ith V2V user in downlink phase
P^d_i	The transmit power of ith V2V transmitter
P^c_m	The transmit power of mth V2N transmitter
P^e	The transmit power of eNodeB
$g^c_{m,e}$	The desired channel gain between mth V2N transmitter and eNodeB
$g^d_{i,e}$	The interfering channel gain between ith V2V transmitter and eNodeB
$g^d_{i,j}$	The desired channel gain between ith V2V transmitter and jth V2V receiver
$g^c_{m,j}$	The interfering channel gain between mth V2N transmitter and jth V2V receiver
$g^c_{e,j}$	The interfering channel gain between eNodeB and jth V2V receiver
$R^{UL}_{\min}$	The required data rate for V2N users in uplink phase
$R^{DL}_{\min}$	The required data rate for V2N users in downlink phase
$UL_{\min}$	The required SINR to successfully decode a packet in uplink phase
$DL_{\min}$	The required SINR to successfully decode a packet in downlink phase
$\rho(i,j)$	The degree of social relationship between V2V pair
$RD^{sum}_{i,j,k}$	The weighted sum rate of V2V users
$x_{i,j,k}$	The binary variable indicating resource allocation

where P^c_m is the transmission power of the mth V2N transmitter, P^d_i denotes the transmission power of the ith V2V transmitter, $g^c_{m,e}$ is the channel gain between the mth V2N transmitter and the eNodeB, $g^d_{i,e}$ denotes the interfering channel gain between the ith V2V transmitter and the eNodeB, and σ^2 is the additive white Gaussian noise. The channel gain can be defined as $g^c_{m,e} = h_{m,e}\chi_{m,e}\theta d^{-\gamma}_{m,e}$. Where $h_{m,e}$ denotes the small-scale fading component, which is exponentially distributed with unit mean. $\chi_{m,e}$ represents the lognormal shadow fading random variable, θ denotes the pathloss component, $d_{m,e}$ denotes the distance between the mth V2N transmitter and the eNodeB, and γ represents the decay component. Note that this definition is applicable to channel gain variables throughout this work. The data transmission rate of the mth V2N user in uplink phase is given as

$$RC^{UL}_m = w_0 \log_2(1 + UL^c_{m,e}) \tag{2}$$

where w_0 represents the allocated bandwidth. The received SINR of a transmission link between ith V2V transmitter and jth V2V receiver can be expressed as

$$UL^d_{i,j} = \frac{P^d_i g^d_{i,j}}{\sum\limits_{j\in VR} P^c_m g^c_{m,j} + \sigma^2} \tag{3}$$

where P^d_i is the transmission power of the ith V2V transmitter, P^c_m denotes the transmission power of the mth V2N transmitter, $g^d_{i,j}$ is the channel gain between the ith V2V transmitter and the jth V2V receiver, $g^c_{m,j}$ denotes the interfering channel gain between the mth V2N transmitter and jth V2V receiver, and σ^2 is the additive white Gaussian noise. The data transmission rate of ith V2V transmitter in the uplink phase is given as

$$RD^{UL}_n = w_0 \log_2(1 + UL^d_{i,j}) \tag{4}$$

where w_0 denotes the allocated bandwidth.

2) Downlink Channel Transmission Rate

Similar to the uplink channel modelling, the downlink phase transmission rate can also be defined in the same approach. The SINR of the transmission link between the eNodeB and mth V2N transmitter can be expressed as

$$DL_{m,e}^{c} = \frac{P^e g_{m,e}^{c}}{\sum\limits_{i \in VT} P_i^d g_{i,e}^{d} + \sigma^2} \tag{5}$$

where P^e is the transmission power of the eNodeB, P_i^d denotes the transmission power of the ith V2V transmitter, $g_{m,e}^{c}$ is the channel gain between the mth V2N transmitter and the eNodeB, $g_{i,e}^{d}$ denotes the interfering channel gain between the ith V2V transmitter and the eNodeB, and σ^2 is the additive white Gaussian noise. The data transmission rate of the mth V2N user in the downlink phase is given as

$$RC_m^{DL} = w_0 \log_2(1 + DL_{m,e}^{c}) \tag{6}$$

where w_0 represents the allocated bandwidth. In the downlink phase, the received SINR of a transmission link between ith V2V transmitter and jth V2V receiver can be expressed as

$$DL_{i,j}^{d} = \frac{P_i^d g_{i,j}^{d}}{\sum\limits_{j \in VR} P^e g_{e,j}^{c} + \sigma^2} \tag{7}$$

where P_i^d is the transmission power of the ith V2V transmitter, P^e denotes the transmission power of the eNodeB, $g_{i,j}^{d}$ is the channel gain between the ith V2V transmitter and the jth V2V receiver, $g_{e,j}^{c}$ denotes the interfering channel gain between the eNodeB and jth V2V reciever, and σ^2 is the additive white Gaussian noise. The data transmission rate of the ith V2V transmitter in the downlink phase is given as

$$RD_n^{DL} = w_0 \log_2(1 + DL_{i,j}^{d}) \tag{8}$$

where w_0 denotes the allocated bandwidth.

B. Social Domain

The social characteristics of users determine the degree of their relationship [19]. Thus, it is essential to analyze the user preference in the process of social modeling in order to derive maximum satisfaction for the users, but it is difficult to model these attributes appropriately. Bayesian methods can be applied to obtain the probability of users having the same preference. The Bayesian non parametric models (BNM) can estimate the probability density function (PDF) directly from the samples obtained from user's historical data.

1) Bayesian Non-Parametric Model

A Bayesian non parametric approach is proposed to predict the user's preference on content. The prior and posterior distribution are non-parametric distributions, which is a stochastic process [20], [21]. Applying statistical distributions which approximates a practical distribution from the estimated sample represents the probability density function (PDF) parametrically, therefore we have to adopt a suitable method that can approximate actual distribution to estimate the parameters. The Bayesian non-parametric model estimates PDF without any assumption for the distribution. In this process, accuracy of the estimation depends on the number of data observed. The Dirichlet process is mostly adopted in Bayesian non-parametric statistics because in a Dirichlet distribution, prior and posterior distributions are stochastic processes [22].

2) Dirichlet Process

The Dirichlet process (DP) is termed as a distribution over distributions. Dirichlet distribution is a multivariate generation of beta distribution [23]. Let assume that an event L_i has been observed $\xi_i - 1$ times, where $i = \{1, 2, 3, ..., \vartheta\}$, and $(l_1, l_2, l_3, ..., l_\vartheta)$ denotes the probability of $(L_1, L_2, L_3, ..., L_\vartheta)$. The Dirichlet distribution of order $\vartheta \geq 2$ with parameters $\xi_1, \xi_2, \xi_3, ..., \xi_\vartheta > 0$ has a PDF with respect to Lebesgue measure on the Euclidean space $\mathbb{R}^{\vartheta-1}$:

$$\begin{aligned} Dir(\xi_1, \xi_2, \xi_3, ..., \xi_\vartheta) &= f(l_1, l_2, l_3, ..., l_{\vartheta-1}; \xi_1, \xi_2, \xi_3, ..., \xi_\vartheta) \\ &= \frac{1}{B(\xi)} \prod_{i=1}^{\vartheta} l_i^{\xi_1 - 1} \end{aligned} \tag{9}$$

where $l_1, l_2, l_3, ..., l_{\vartheta-1} > 0$ satisfies $l_1 + l_2 + l_3 + ... + l_{\vartheta-1} < 1$, l_ϑ represents the abbreviation for $1 - l_1 - l_2 - l_3 - ... - l_{\vartheta-1}$. The density is zero outside $(\vartheta - 1)$ dimensional simplex. $B(\xi)$ represents the normalizing constant, and is expressed in form of gamma function since it is a multinomial beta function

$$B(\xi) = \frac{\prod_{i=1}^{\vartheta} \Gamma(\xi_i)}{\Gamma(\sum_{i=1}^{\vartheta} \xi_i)} \tag{10}$$

where $\xi_0 = \sum_{i=1}^{\vartheta} \xi_i$. The marginal distribution of the Dirichlet distribution which is a beta distribution can be expressed as $l_i \sim Beta(\xi_i, \xi_0 - \xi_i)$. Subsequently, the DP over a set Φ, is a stochastic process and its realization is a probability distribution on the set Φ. Let H denote a random distribution on a set Φ. The distribution is a DP distributed with F and δ, which is defined as $H \sim DP(\delta, F)$. For the finite measurable partition $A_1, A_2, A_3, ..., A_R$ of Φ, we have:

$$\begin{aligned} &(H(A_1), H(A_2), H(A_3), ..., H(A_R)) \sim \\ &Dir(\delta F(A_1), \delta F(A_2), \delta F(A_3), ..., \delta F(A_R)) \end{aligned} \tag{11}$$

where $F(.)$ is the mean of the Dirichlet process, while δ denotes the inverse of variance. Therefore, for any measurable partition $A \subset \Phi$ we have

$$\begin{aligned} E[H(A)] &= F(A) \\ V[H(A)] &= \frac{F(A)[1 - F(A)]}{\delta + 1} \end{aligned} \tag{12}$$

where δ is also seen as a strength parameter, which indicates the strength of the prior when the DP is used as prior information in Bayesian non-parametric methods. Thus, the variance decreases as δ grows, hence the DP will pay more emphasis on the mean. Since H is a random distribution over Φ, we can use the independent samples from H which is expressed as a sequence $c_1, c_2, c_3, ..., c_n$. The values c_i are in Φ, $n_r = \#\{i : c_i \in A_r\}$ represents the observed values in

A_r, and $r = 1, 2, 3, ..., R$. Considering the conjugacy between the multinominal and Dirichlet distributions, then we have

$$\begin{aligned}&(H(A_1), H(A_2), H(A_3), ..., H(A_R)) \,|c_1, c_2, c_3, ..., c_n \sim \\ &Dir(\delta F(A_1) + n_1, ..., \delta F(A_R) + n_R)\end{aligned} \tag{13}$$

Similarly, for all the finite measurable partitions, the posterior distribution over H is a Dirichlet process which can be given as

$$H\,|c_1, c_2, c_3,\ ..., c_n \sim DP\left(\delta + n, \frac{\delta}{\delta + n}F + \frac{n}{\delta + n}\frac{n_r}{n}\right) \tag{14}$$

where $n_r = \sum_{r=1}^{n} \delta_r(A_r)$. The weighted average between the prior base distribution F is proportional to δ. The weight of the empirical distribution n_r is proportional to the number of observations n. If δ is equal to zero, then F is no more essential because the posterior distribution can be deduced from the empirical distribution. Let the predictive distribution of c_{n+1} be expressed in a sequence $c_1, c_2, c_3, ..., c_n$. We know that $c_{n+1}\,|c_1, c_2, c_3, ..., c_n \sim H$ for a measure $A \in \Phi$, then we have

$$\begin{aligned}P(c_{n+1} \in A\,|c_1, c_2, c_3, ..., c_n) &= E[H(A)\,|c_1, c_2, c_3, ..., c_n] \\ &= \frac{1}{\delta + n}(\delta F(A) + n(A))\end{aligned} \tag{15}$$

Subsequently, H is marginalized out. The concept of kernel is applied in order to smooth the distribution and derive its density distribution. Let $f(x\,|\tau)$ denote the family of kernels (densities), and τ is the density index. The smoothed non parametric density is given as $P(x) = \int f(x\,|\tau)H(\tau)d\tau$.

3) *Probability Distribution Estimation*

Assuming Z represent the users in the system, a user can be represented as $z \in Z$ and U denotes observation sets associated with the probabilities of requesting similar content, which can be extracted from user's history records [24]. Let Y be the observation set at a particular instance $U \in Y$. A user z requests for a similar content with the probability p_{Yz}. The probability density function of p_{Yz} can be expressed as $P_{Yz}(p_{Yz})$ over the space $\Phi = [0, 1]$. For each observation set, Q_{Yz} observations are obtained and can be expressed as

$$Q_{Yz} = \{p_{Yz}^1, p_{Yz}^2, p_{Yz}^3, ..., p_{Yz}^{Q_{Yz}}\}, \forall z \in Z, U \in Y \tag{16}$$

To obtain the predictive PDF of the subsequent observation $p_{Yz}^{Q_{Yz}+1}$, we apply the concept of Dirichlet process based on equation (16), then we have

$$\begin{aligned}&P_{Yz}(p_{Yz}^{Q_{Yz}+1} \in \mu\,|p_{Yz}^1, p_{Yz}^2, p_{Yz}^3,\ ..., p_{Yz}^{Q_{Yz}}) \\ &= \frac{1}{\alpha + Q_{Yz}}(\alpha G(\mu) + \sum\nolimits_{Q=1}^{Q_{Yz}} \beta_{p_{Yz}^Q}(\mu))\end{aligned} \tag{17}$$

where G denotes the prior distribution, μ represents the measurable partition of Φ, and α denotes the strength of the prior distribution in respect of the posterior estimation. By marginalizing out the Dirichlet process, if the concentration parameter α and the base distribution are unknown, the predictive PDF of the subsequent observation is given as

$$P_{Yz}(p_{Yz}^{Q_{Yz}+1} \in \mu\,|p_{Yz}^1, p_{Yz}^2, p_{Yz}^3,\ ..., p_{Yz}^{Q_{Yz}}) = \frac{\sum_{Q=1}^{Q_{Yz}} \beta_{p_{Yz}^Q}(\mu)}{Q_{Yz}} \tag{18}$$

Similarly,

$$\beta_{p_{Yz}^Q}(\mu) = \begin{cases} 1, & p_{Yz}^Q \in \mu \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

where $\beta_{p_{Yz}^Q}$ is the point mass located at p_{Yz}^Q. To smooth out the distribution, we apply kernel from the Dirichlet process in order to obtain the continuous estimate $\tilde{P}_{Yz}$ of P_{Yz}. When the number of observations is small, a different method is considered to improve the estimates. For a user $z \in Z$, let $X \subseteq Y$ denote a subset in the observation set $U \in Y$, and $X_{Yz} = X \backslash \{U\}$ represents the rest of the observation sets in X except U. By integrating the observation set U with the verified priors X_{Yz}, the PDF of the new observation is given as

$$P_{Yz}^X = \ell_U \tilde{P}_{Yz}(\mu) + \sum_{T \in X_{Yz}} \ell_T \tilde{P}_{Tz}(\mu) \tag{20}$$

where ℓU quantifies the PDF of p_{Yz}^X, ℓT quantifies the PDF of $T \in X_{Yz}$. The weights of the quantifiers are set to be proportional to the number of observations, considering the equal availability of the observation sets, then $P_z = P_{Yz}^X$. The probability for two users (i, j) to select similar content is applied to obtain the normalized correlation, which can be defined as

$$\omega_{i,j} = \frac{(cor(p_i, p_j) + 1)}{2} \tag{21}$$

where $\omega_{i,j} \in [0, 1]$, $p_i \sim P_i(p)$, $p_j \sim P_j(p)$, p_i and p_j denotes the estimated correlative PDFs.

4) *Success of Content Delivery*

In this scheme, we also consider link reliability in V2V peer discovery. Due to the high mobility of vehicles, communication link can be unstable and easily be affected by the driver's behavior. A V2V pair established by users in communication range cannot guarantee the success of data transmission [5], [25]. Therefore, evaluating the success probability of the V2V link is necessary in order to ascertain link reliability. Vehicles encounter history can be analyzed to extract their communication characteristics. The minimum required duration to transmit a packet over a V2V link can be expressed as $cd_{i,j} = \frac{o}{RD_{i,j}}$. Where o denotes the size of a packet to be transferred over a V2V link, $RD_{i,j}$ represents the data rate. Let $\tilde{cd}_{i,j}$ denote the average contact duration between V2V links, the success probability of data transmission can be given as

$$P_s = P\left\{\tilde{cd}_{i,j} \geq \frac{o}{RD_{i,j}}\right\} \tag{22}$$

Encounter time is referred to as the duration of each encounter between vehicles. The distribution of encounter duration is obtained based on encounter history. The encounter time of two vehicles is modelled as a power law distribution with a heavy tail index [26], [27] denoted as ε. Then we have

$$P\left(cd_{i,j} > a\right) = \left(\frac{a}{a_1}\right)^{-\varepsilon} \tag{23}$$

where a_1 denotes minimum encounter duration between two vehicles. The average encounter duration $\tilde{cd}_{i,j}$ can be computed using

$$\tilde{cd}_{i,j} = \begin{cases} a_{\min}\left[1 + \frac{1}{\varepsilon-1}\left(\frac{a_1}{a_{\min}}\right)^{\varepsilon}\right], \text{ if } a_{\min} > a_1 \\ \\ \frac{\varepsilon}{\varepsilon-1}a_1, \text{ otherwise} \end{cases} \tag{24}$$

where $a_{\min}$ denotes the minimum required communication duration for a V2V link. We can see that P_s is related to the data rate $RD_{i,j}$. With available statistical characteristics of channel gain at the eNodeB we can obtain the distribution of data rate $RD_{i,j}$, which is given as

$$\begin{aligned} P_s &= P\left\{RD_{i,j} > \frac{o}{\tilde{cd}_{i,j}}\right\} \\ &= P\left\{\frac{P_j^d g_j^d}{P_i^c g_{i,j}^c + \sigma^2} 2^{\left(\frac{o}{\tilde{cd}_{i,j}}\right)} - 1\right\} \end{aligned} \tag{25}$$

Thus, when $a_{\min} > a_1$ we have

$$P_s = P\left[RD_{i,j} \geq \frac{o}{a_{\min}\left[1 + \frac{1}{\varepsilon-1}\left(\frac{a_1}{a_{\min}}\right)^{\varepsilon}\right]}\right] \tag{26}$$

Otherwise, when $a_{\min} \leq a_1$ then we have

$$P_s = P\left[RD_{i,j} \geq \frac{o(\varepsilon-1)}{\varepsilon a_1}\right] \tag{27}$$

where $P_s \in [0,1]$. When a V2V pair is established for content dissemination, the behavior of the V2V users affects the success of content transmission. The behavior of the user could be objective or subjective. The probability of content similarity, which is obtained based on the preference of users. However, a V2V pair cannot have a successful content delivery when the established V2V link is not stable. In the proposed scheme, the social attributes including similarity in preference of users denoted as $\omega_{i,j}$ and link reliability denoted as $P_s(i,j)$ are considered for efficient content sharing. Thus, the degree of social relationship between V2V pair (i,j) can be expressed as $\rho(i,j) = \omega_{i,j}.P_s(i,j)$.

IV. Problem Formulation

The objective is the maximization of the sum rate of V2V pairs, taking into account both the social and physical attributes of users in the network. To achieve content dissemination with user satisfaction [30], we formulate the objective function as a weighted channel rate, which is weighted by the degree of social relationship between V2V users. The weighted sum rate of V2V pairs can be given as

$$RD_{i,j,k}^{sum} = WRD_{i,j,k}^{UL} + WRD_{i,j,k}^{DL} \tag{28}$$

where $WRD_{i,j,k}^{UL} = SI(\rho_{i,j})\rho_{i,j}RD_n^{UL}$ and $WRD_{i,j,k}^{DL} = SI(\rho_{i,j})\rho_{i,j}RD_n^{DL}$. V2V pair is established and allowed to share content only when the degree of relationship between the V2V pairs is above or equal to the predefined threshold η. Therefore, a potential V2V pair VT_i and VR_j can connect through V2V link only when $\rho_{i,j} \geq \eta$, thus $SI(\rho_{i,j})$ is a function that indicates degree of social relationship between V2V users. Then we have

$$SI(\rho_{i,j}) = \begin{cases} 1, \text{ when } \rho_{i,j} \geq \eta \\ 0, \text{ otherwise} \end{cases} \tag{29}$$

A binary variable $x_{i,j,k} = 1$ indicates V2V pair reusing a V2N resource block is established, and when $x_{i,j,k} = 0$ otherwise. The binary variables are denoted by a set $X = \{x_{i,j,k}\}$. The resource allocation problem can be formulated as

Problem 1:

$$\begin{aligned} &\underset{\{X,P\}}{\text{maximize}} \quad \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{K}(RD_{i,j,k}^{sum})x_{i,j,k} \\ &\text{subject to} \quad C1: RC_m^{UL} \geq R_{\min}^{UL}, \forall m \in CU, \\ &\qquad C2: RC_m^{DL} \geq R_{\min}^{DL}, \forall m \in CU, \\ &\qquad C3: UL_{i,j}^d \geq UL_{\min}, \forall i \in VT, j \in VR, \\ &\qquad C4: DL_{i,j}^d \geq DL_{\min}, \forall i \in VT, j \in VR, \\ &\qquad C5: 0 \leq P_i^d \leq P_{\max}^d, \forall i \in VT, \\ &\qquad C6: 0 \leq P_m^c \leq P_{\max}^c, \forall m \in CU, \\ &\qquad C7: x_{i,j,k} \in \{0,1\}, \forall i \in VT, j \in VR, k \in RB, \\ &\qquad C8: \sum_{zh=1}^{\phi} x_{i,j,k} \leq 1, \forall i \in VT, j \in VR, k \in RB \end{aligned} \tag{30}$$

where constraint C1-C2 guarantee the data rate of V2N users in both uplink and downlink channels, C3-C4 ensure that V2V pairs can successfully transmit data in uplink and downlink phases. C5-C6 guarantee that the transmission power of V2V and V2N users did not exceed the maximum transmit power, C7 ensures that a V2V transmitter is paired to at most one V2V receiver and vice versa, C8 ensures that only V2V pair of thesame cluster can share a resource block. The problem is MINLP thus the complexity increases with higher number of vehicles in the network. For simplicity, we decompose the optimization problem into two subproblems; power allocation and channel assignment.

Lemma 1: Only when constraints C1-C4 are satisfied with equality, then we can obtain the optimal solution of problem 1.

Proof: let the optimal solution of problem 1 be denoted as X^*, P^* that satisfies constraints C1-C4: $RC_m^{UL} > R_{\min}^{UL}$, $RC_m^{DL} > R_{\min}^{DL}$, $UL_{i,j}^d > UL_{\min}$ and $DL_{i,j}^d > DL_{\min}$. Consequently, the transmit powers of both V2V and V2N users are set to make sure that $RC_m^{UL} = R_{\min}^{UL}$, $RC_m^{DL} = R_{\min}^{DL}$, $UL_{i,j}^d = UL_{\min}$, and $DL_{i,j}^d = DL_{\min}$. This implies that there exists a solution X, P. Hence, the initial assumption that X^*P^* is optimal solution is not true.

V. The Proposed Efficient Content Dissemination Scheme

A. Cluster Formation

The major aim of forming vehicle clusters is for resource sharing with minimum interference. The proposed clustering algorithm assigns V2V pairs into several clusters by considering their mutual interference and relative speed [28]. Vehicles are grouped into ϕ clusters. The V2V pair in each cluster shares resource block with a certain V2N user while a resource block cannot be shared with the V2V users of different clusters. Therefore, the number of clusters is less than or equal to the number of V2N links in the nework [29]. This clustering approach does not only minimize intra-cluster interference by supporting V2V pairs to operate on less transmission power but also improves system performance by maximizing the capacities of V2N users. A summary of the clustering algorithm is given as follows

- At the initial stage, V2V pairs are randomly assigned to each cluster based on location and vehicular density.
- For efficient V2V communication, the variation in velocity between cluster members should be below a threshold. $Rv_{i,j} \leq \partial$, where $Rv_{i,j}$ is the relative velocity between vehicles and ∂ denotes the variation threshold.
- In the final stage, the interference level of each cluster is computed assuming the V2V pair is added to the cluster. Then the V2V pair will be assigned to the cluster with least interference.

Algorithm 1: Clustering Algorithm

```
inputs: TW, φ, SIF
output: ZH
randomly assign V2V pair into φ clusters based on
  location and vehicular density
compute the average velocity of each cluster
while |TW| ≠ 0 do
    foreach V2V pair not assigned into a cluster do
        if Rv_{i,j} ≤ ∂ then
            compute the sum of intra-cluster
              interference of each cluster after adding
              the V2V pair
            SIF = Σ_{zh∈ZH}^{φ} In + In'_i
            Add V2V pair to the cluster with lowest
              SIF
        else
            request not accepted
        end
    end
end
```

B. Power Allocation

For the power control of V2V pair and V2N user, suppose the ith V2V transmitter reuses the RB of mth V2N user. Let $h^{UL} = e^{\frac{R_{\min}^{UL}}{w_0}} - 1$ and $h^{DL} = e^{\frac{R_{\min}^{DL}}{w_0}} - 1$, then we can rewrite constraints C1-C4 in problem 1 as

$$\begin{aligned} &P_m^c g_{m,e}^c \geq h^{UL}\left(P_i^d g_{i,e}^d + \sigma^2\right) \\ &P^e g_{m,e}^c \geq h^{DL}\left(P_i^d g_{i,e}^d + \sigma^2\right) \\ &P_i^d g_{i,j}^d \geq UL_{\min}\left(P_m^c g_{m,j}^c + \sigma^2\right) \\ &P_i^d g_{i,j}^d \geq DL_{\min}\left(P^e g_{e,j}^c + \sigma^2\right) \end{aligned} \tag{31}$$

The power allocation problem for a V2N user and V2V pair can be considered as

Problem 2:

$$\begin{aligned} \underset{\{P\}}{\text{maximize}} \quad & RD_{i,j,k}^{sum} x_{i,j,k} \\ \text{subject to} \quad & 0 \leq P_i^d \leq P_{\max}^d, \forall i \in VT\ , \\ & 0 \leq P_m^c \leq P_{\max}^c, \forall m \in CU\ , \\ & h^{UL} \leq \frac{P_m^c g_{m,e}^c}{P_i^d g_{i,e}^d + \sigma^2}\ , \\ & h^{DL} \leq \frac{P^e g_{m,e}^c}{P_i^d g_{i,e}^d + \sigma^2} \end{aligned} \tag{32}$$

To maximize the weighted sum rate of V2V pairs, the transmit power P_i^d need to be increased, and the transmit powers P_m^c and P^e have to be minimized as much as possible. The optimal point can be obtained when h^{UL} and h^{DL} take equalities, thus we have $P_m^{*c} = \frac{h^{UL}}{g_{m,e}^c}\left(P_i^d g_{i,e}^d + \sigma^2\right)$ and $P^{*e} = \frac{h^{DL}}{g_{m,e}^c}\left(P_i^d g_{i,e}^d + \sigma^2\right)$. By substituting P_m^{*c} and P^{*e} in the objective function, then we have

$$\begin{aligned} & w_0 \log_2\left(1 + \frac{g_{i,j}^d}{\frac{\sigma^2}{P_i^d}\left(1 + \frac{h^{UL} g_{m,j}^c}{g_{m,e}^c}\right) + \left(1 + \frac{h^{UL} g_{i,e}^d g_{m,j}^c}{g_{m.e}^c}\right)}\right) \\ & + w_0 \log_2\left(1 + \frac{g_{i,j}^d}{\frac{\sigma^2}{P_i^d}\left(1 + \frac{h^{DL} g_{e,j}^c}{g_{m,e}^c}\right) + \left(1 + \frac{h^{DL} g_{i,e}^d g_{e,j}^c}{g_{m,e}^c}\right)}\right) \end{aligned} \tag{33}$$

We can see that equation (33) is monotonically increasing with the transmit power P_i^d. To maximize the objective function we need to set the transmit power of V2V transmitter as high as possible, $P_i^{*d} = \{P_{\max}^d\}$. From equations (32-33) we can compute the optimal transmit powers of users.

C. Channel Assignment

To obtain the optimal channel assignment for V2V pairs reusing V2N resource blocks that can maximize the weighted sum rate of all V2V links, we have the following optimization problem

Problem 3:

$$\begin{aligned} \underset{\{x_{i,j,k}\}}{\text{maximize}} \quad & \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{K} RD_{i,j,k}^{sum} x_{i,j,k} \\ \text{subject to} \quad & \\ & x_{i,j,k} \in \{0,1\}, \forall i \in VT, j \in VR, k \in RB \end{aligned} \tag{34}$$

To solve problem 3 with a lower computational complexity, we propose a distributed matching algorithm.

1) Matching Algorithm for Channel Assignment

Problem 3 is a typical three-dimension (3D) matching problem [31], which consists of V2V transmitters $VT's$, V2V receivers $VR's$, and resource blocks $RB's$. Due to the computational complexity of 3D matching we transform the problem into a two-dimensional (2D) matching problem. Since the resource blocks are orthogonally allocated to V2N users $CU's$, we merge the V2V pair with the resource block they reuse i.e. a VR mapped to a RB is denoted as RR. Initially we already have N number of $VR's$ and M number of $CU's$. Therefore, $RR = \{RR_1, RR_2, RR_3, ..., RR_{NM}\}$. Matching model is designed to allocate resources to users. A typical solution to a matching problem is referred to as two-sided stable matching, which involves two finite and disjointed sets, i.e. users and resources respectively. Each element in the set has its preference over elements in the opposite set.

- **Definition 1**: consider a set of users $US = \{us_1, us_2, us_3, ..., us_n\}$ and a set of resources $RS = \{rs_1, rs_2, rs_3, ..., rs_m\}$. Matching ψ is the one to one mapping for the set $US \cup RS$ to itself, which can be presented as $\psi : US \cup RS \rightarrow US \cup RS$, given that $\psi(us) = rs$ denotes us and rs are matched, and $\psi(us) = us$ denotes that us is yet to be matched. Matching is strictly based on defined preference of each element in the set.
- **Definition 2**: a matching is stable if there exists no blocking pair (BP). Considering that us and rs are already matched to different partners, but they prefer to be matched together. Therefore us and rs form a BP for matching ψ, hence due to the BP the matching ψ is not stable.

In our proposed scheme, we already have two finite and disjointed sets of V2V transmitter VT and resource unit RR respectively. Matching is obtained based on maximum weighted sum rate of V2V pairs. Matching ψ can be given as $VT \cup RR \rightarrow VT \cup RR$ such that $\psi(VT_i) = RR_{j,k}$.

2) Matching Preference

Preference sets the priority for selection among several alternatives [32]. In matching models an individual in one set proposes to individual in the other set based on their preference ranking [33]. To obtain the preference list of each V2V transmitter VT on resource units RR, we have to compute the weighted sum rate of each V2V pair. At first, we have to temporarily pair VT_i with every $RR_{j,k}$. Since each RB may generate different data rate and user satisfaction, for each combination the data rates are sorted in the descending order. VT_i will prefer the RR with higher data rate over those with lower data rate, therefore the preference list of VT_i can be presented as $pl = \{pl_1, pl_2, pl_3, ...pl_n\}$, and the preference list of all $VT's$ can be denoted in a set $PL = \{PL_1, PL_2, PL_3, ..., PL_n\}$.

3) Pay Off Function

To obtain the payoff value of each VT_i, we introduce matching cost, which is the price for services offered by the RR unit. The initial price of each $RR_{j,k}$ is set to zero. The prices of RR units can be denoted as $PR = \{PR_1, PR_2, PR_3, ..., PR_{NM}\}$. Let $AB_{i,j,k}$ denote the valuation of $RD^{sum}_{i,j,k}$, and $PR_{j,k}$ represent the matching cost for V2V transmitters $VT's$. The payoff of VT_i when matched with $RR_{j,k}$ is given as $AB_{i,j,k} - PR_{j,k}$. In a situation where more than one VT proposed to any RR, the RR increases their prices by a step sz, then the $VT's$ will modify their preferences due to the change in matching cost, which is determined based on the highest payoff [34] given as

$$PO_{j,k} = \arg \max_{\{j \in VR, k \in RB\}} (AB_{i,j,k} - PR_{j,k}) \quad (35)$$

Algorithm 2: Preference Formulation Algorithm

```
inputs: VT, VR, CU, RB, P
output: PL
foreach i ∈ VT do
    foreach j ∈ VR do
        foreach k ∈ RB do
            obtain the maximum weighted rate of each
              pair
        end
    end
end
foreach i ∈ VT do
    sort the weighted rate in descending order
    generate preference list pl_i for VT_i based on
      sorted the weighted rate
end
return PL
```

4) Proposed Stable Matching Algorithm

The proposed matching algorithm is based on Gale Shapley technique [31], to obtain stable matching in a distributed manner. We obtained the preferences of each set using algorithm 2. At the first stage each VT_i would propose to its most preferred $RR_{j,k}$. In a circumstance where $RR_{j,k}$ receives a proposal from more than one VT_i, the $RR_{j,k}$ adjusts their prices to determine a winner. The stable matching process operates as follows:

- For each VT_i that has not been paired with any $RR_{j,k}$, it proposes to its most preferred candidate based on the payoff function in equation (35).
- If $RR_{j,k}$ receives a proposal from more than one VT_i, then there exists a conflicting set denoted as ζ. Thus, each element in ζ would increase their price by step sz.
- Each VT_i that proposed to an element in ζ will modify its preference due to increase in matching cost and renew its proposal. The matching cost keeps increasing until when each $RR_{j,k}$ receives only one request.
- The algorithm terminates when there is no more request from VT to RR.

D. Stability and Complexity Analysis

In this section, we analyse some vital properties of the proposed algorithms. We derive the stability of the proposed algorithms and their computational complexity.

Algorithm 3: Channel Allocation Algorithm

inputs: $VT, RR, PL, PR, sz, \zeta$
output: ψ
initialization: set all $i \in VT$ as free, all $j, k \in RR$ as free, $\psi = \emptyset$, $\zeta = \emptyset$, $sz = 0.1$
while $VT \neq \emptyset$ **do**
  foreach *$i \in VT$ that is free and has not proposed to all elements in pl_i* **do**
    propose to the most preferred $RR_{j,k}$
  end
  if $\zeta \neq \emptyset$ **then**
    foreach $RR_{j,k} \in \zeta$ **do**
      $Increase$ matching cost
      $PR'_{j,k} = PR_{j,k} + sz$
      $PR'_{j,k} > PR_{j,k}$
      $update$ the pl_i for each $i \in VT$ base on equation (35)
      VT_i renews proposal using the updated pl_i
    end
  else
    $\zeta = \emptyset$
    there exists no conflicting set
    $assign$ VT_i to $RR_{j,k}$
  end
  $update$ PL and PR
  return ψ
end

1) Stability

Lemma 2: the proposed algorithm converges to a stable matching in a finite number of iterations.

Proof: a matching is stable only if a blocking pair doesn't exist. Let's assume that we have $VT_i \in VT$, $RR_{j,k} \in RR$ such that VT_i and $RR_{j,k}$ are not paired together under matching ψ, but they prefer to be matched to each other. Since the goal of each VT_i is to maximize the payoff $\max_{j \in VR, k \in RB} PO_{j,k}, VT_i \in VT$. We earlier assume that VT_i prefers $RR_{j,k}$. VT_i must have proposed to $RR_{j,k}$. Consequently, due to the effect of pricing in algorithm 3; if $\psi(VT_i) \neq RR_{j,k}$ then the payoff $PO_{j,k} = 0$ and VT_i couldn't win $RR_{j,k}$ during contention. Hence considering the analysis above, the initial assumption that VT_i and $RR_{j,k}$ are blocking pair cannot hold. Therefore, the matching is stable.

2) Complexity

The computational complexity of the proposed scheme includes the complexity of algorithms 1-3. In algorithm 1, to group vehicles in clusters, we need to calculate the interference level of each cluster when a V2V pair is added. Therefore, the complexity is $O(NM)$ to check all possibilities. In algorithm 2, each $VT_i \in VT$ obtains its preference on $RR_{j,k}$. The computation complexity of the process is $O(NM)$. VT_i forms its preference list by sorting the preference value of each $RR_{j,k}$. To perform such task, the computation complexity is $O(NM \log(NM))$. For the matching process presented in algorithm 3, each VT_i that have not been paired proposes to its preferred $RR_{j,k}$, and the computational complexity for this operation is $O(N^*)$, where N^* denotes the number of iterations during the contention period. The matching of contending elements is completed when $\varsigma \neq \emptyset$. However, if $\varsigma = \emptyset$ then $N^* = 1$. Thus, the computational complexity of the proposed stable matching model is $O(NN^*)(N \geq M)$.

VI. Simulation Results and Discussions

To validate the proposed scheme, simulations were conducted. We consider a freeway road with multilane of 5km as described in 3GPP TR [35]. The road passes through a cell where an eNodeB is located. Vehicles arrival is based on spatial Poisson process and the vehicular density is determined by the vehicles speed. For the experiments, mobility traces are generated in SUMO and the performance evaluation is conducted using NS3 and MATLAB 2017b simulators. The detailed simulation parameters [36], [37], [40] are given in Table III.

TABLE III: Summary of Simulation Parameters

Simulation Parameter	Value
Cell Radius	500m
Bandwidth	10Mhz
Carrier Frequency	2Ghz
Vehicle Antenna	1.5m (3dBi)
eNodeB Antenna	25m (8dBi)
Noise Power	-114 dBm
Number of V2V Users	30
Number of V2N Users	10
Maximum Transmission Power	23dBm
Number of Lanes	4
Small Scale Fading	Rayleigh Fading
V2V Pathloss Model	LOS in WINNER II B1
V2N Pathloss Model	128.1 + 37.6 log10d
V2V Shadowing Model	Log-normal 3dB
V2N Shadowing Model	Log-normal 8dB

The performance of the proposed scheme is compared in terms of weighted sum rate of V2V pairs with that of the following schemes:

- Uplink based scheme [18]: this is a variant of the proposed scheme, herein only the uplink channels is considered for resource block reuse. The V2V pairs optimally selects the V2N resource block to reuse.
- Random based scheme [38]: in this resource management scheme, the V2V users randomly select a V2N resource block to reuse. In this scheme only the uplink channels are considered for resource sharing.
- Fixed dimension clustering scheme [39]: in the scheme number of vehicles in each cluster are equal and vehicles with the worst interference are added to the same cluster. However, vehicles in the same cluster are not allowed to share the same resource block.

To evaluate the significance of the proposed scheme, we consider six different simulation experiments as discussed below.

A. Convergence

The optimality of the proposed algorithm is evaluated by the weighted sum rate of V2V pairs vs. the number of iterations. For the purpose of comparison, we use a graph-theory based algorithm to obtain an optimal solution for the problem. After almost 40 iterations, the proposed scheme converges to a stable value. As illustrated in Fig. 2 the weighted sum rate improves with increase in number of users and the proposed scheme achieves a close optimal performance when $N = 40$.

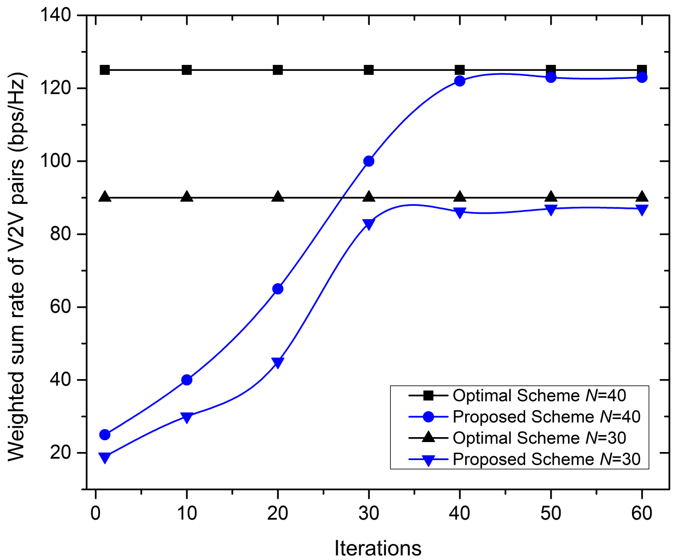


Fig. 2: Convergence of algorithms over number of iterations

B. The impact of different number of V2V pairs on performance

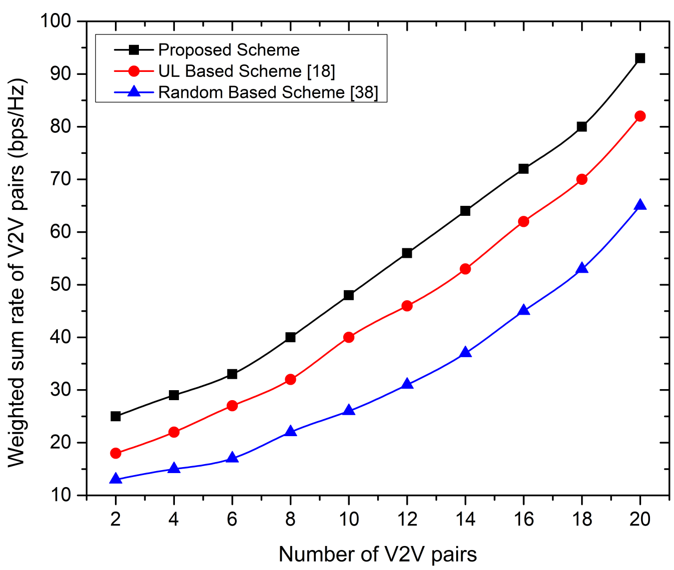


Fig. 3: Weighted sum rate of V2V pairs over number of V2V pairs

The weighted sum rate depends on the number of V2V pairs. The proposed scheme is evaluated under different number of V2V pairs. Initially we set the number of V2V pairs to two, subsequently we keep increasing the number of the V2V pairs as illustrated in Fig. 3. We observe that the weighted sum rate of V2V pairs is proportional to the number of V2V pairs in all schemes. The proposed scheme achieves the highest rate because data can be transmitted in both uplink and downlink phases, which provides more opportunity for the V2V pairs to choose an optimal channel. While scheme [18] transmits data only in the uplink phase, scheme [38] randomly selects any available channel for transmission.

C. The impact of V2N user's data rate requirement on performance

To evaluate the impact of V2N data rate requirement, we obtain the weighted sum rate of V2V pairs by varying the V2N data rate requirement from 100 kbps to 800 kbps as illustrated in Fig. 4. It is observed that with the increase in V2N data rate requirement, the weighted sum rate of V2V pairs decreases in all schemes. This is because the V2N increases transmit power in order to guarantee the V2N data rate. The increase in the V2N transmit powers generates more interferences to the V2V pairs, hence degrades the system performance. However, the proposed scheme achieves the highest rate in comparison with the two other schemes.

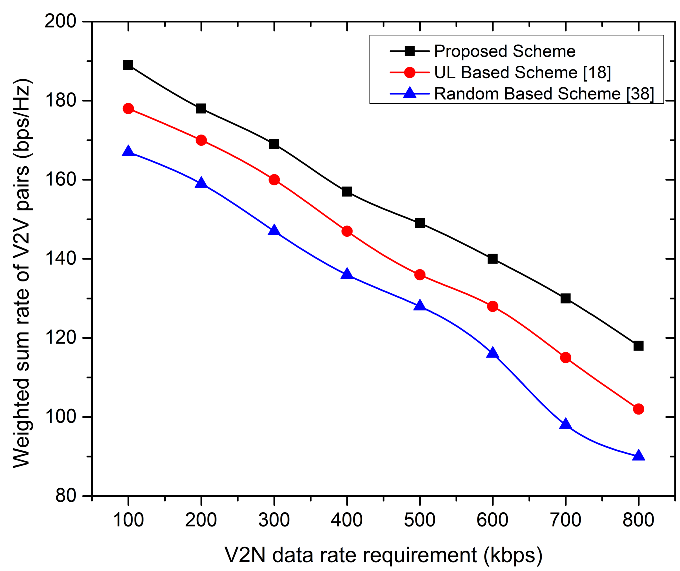


Fig. 4: Weighted sum rate of V2V pairs over data rate requirement of V2N users

D. The impact of vehicle clustering on performance

In Fig. 5, we evaluate the impact of vehicle clustering on the system performance in terms of the weighted sum rate of V2V pairs. The proposed scheme achieves the highest rate than the other two schemes, because the clusters are formed by minimizing intra cluster interference. The fixed dimension scheme [39] also considers intra cluster interference, where each cluster is formed with equal number of vehicles. However, this approach could not minimize the interference to the barest level. The random clustering recorded the worst performance, because it doesn't consider intra cluster interference in the cluster formation process, rather vehicles are clustered based on uniform distribution.

E. The impact of vehicle speed on performance

Fig. 6 illustrates the weighted sum rate of V2V pairs with an increasing vehicle speed. We observe that the weighted sum

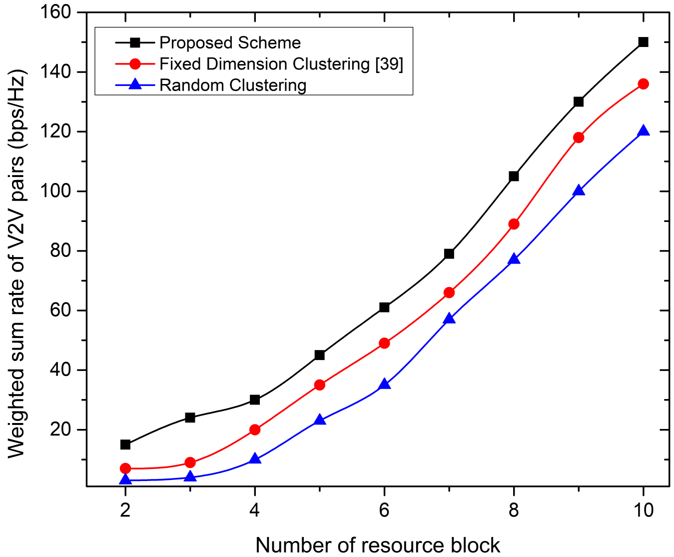


Fig. 5: Impact of vehicle clustering on the system performance

rate of V2V pairs in all schemes is inversely proportional to the vehicle speed. This is because the increase of the vehicle speed induces a sparse vehicle traffic, which reduces transmission data rates and the overall system performance. The proposed scheme performs better than other schemes.

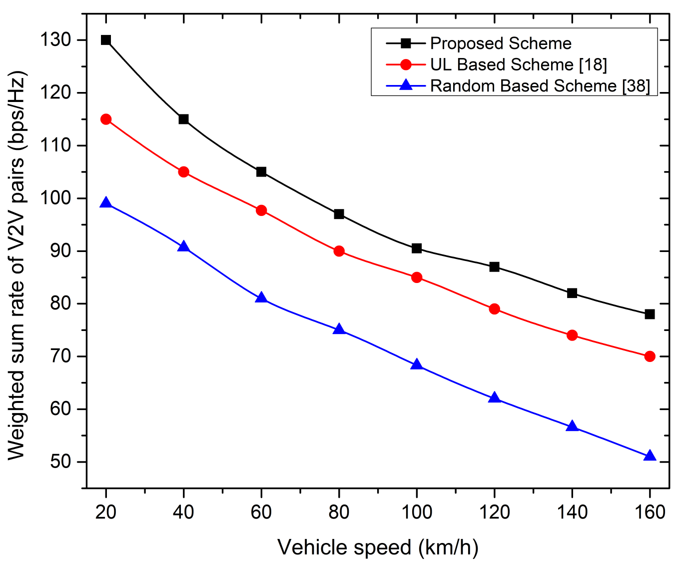


Fig. 6: Weighted sum rate of V2V pairs with varying vehicle speed

F. *V2V user's satisfaction on received content*

The cumulative density function (CDF) of the satisfaction of V2V content receivers $RT's$ is illustrated in Fig. 7, which is based on the similarity of user's preference on content. This is determined by the level of social relationship between V2V content transmitter and content receiver. We evaluate the satisfaction of socially aware and socially unaware V2V pairs by varying the threshold of relationship intensity. The percentage of V2V receivers with satisfaction of 90% and above are as follows; for the socially unaware scheme we have 14%, and for the socially aware scheme when the threshold is set at 0.5 and 0.7, we have 24% and 33%, respectively.

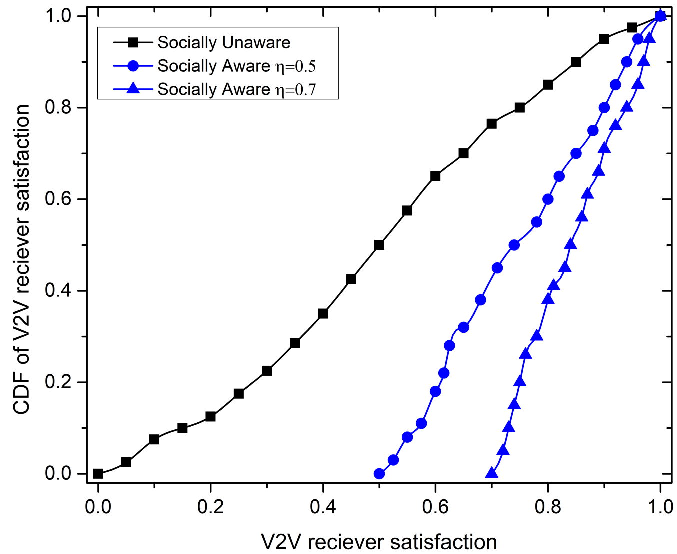


Fig. 7: Satisfaction of V2V recievers on content

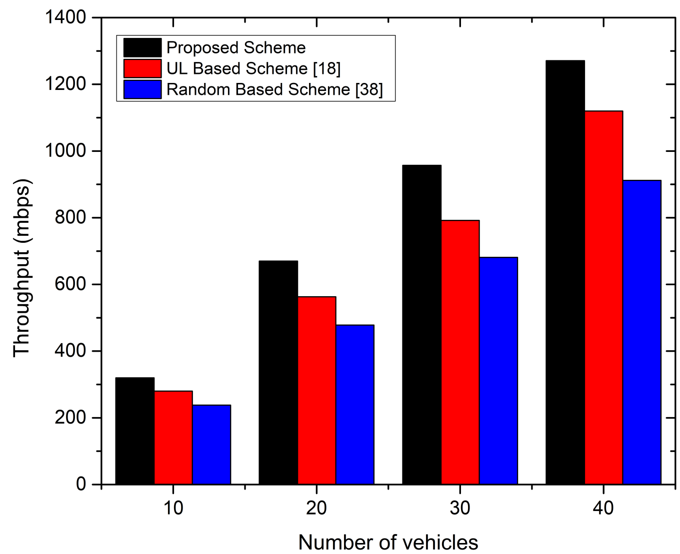


Fig. 8: Impact of number of vehicles on throughput

G. *Impact of number of vehicles on throughput*

Fig. 8 shows the impact of the number of vehicles on system throughput. The throughput is proportional to the number of vehicles in all schemes. When the number of vehicles increases, the system throughput increases. Because with a small number of vehicles, there are few vehicles to share content through V2V communications. Most vehicles have to receive content through the eNodeB using the V2N communication, which results in low throughput. When the number of vehicles increases, more vehicles share content through V2V communication, eventually improving throughput. The proposed scheme achieves the highest throughput than the existing schemes.

VII. Conclusions

In this paper, we propose an efficient content dissemination scheme which is designed based on the exploitation of social attributes and physical attributes of the users. We apply the Bayesian non-parametric methods to obtain the similarity in preference of the users, which is used to pair the V2V users for content dissemination. To improve the system performance, we consider users' QoS in both uplink and downlink channels. Furthermore, for spectrum reuse in the underlaying cellular network, we propose an improved matching algorithm for channel allocation. Simulation results show that the proposed scheme achieves a better performance than other existing schemes.

References

[1] Z. Li, L. Xiang and X. Ge, "Age of Information Modeling and Optimization for Fast Information Dissemination in Vehicular Social Networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5445-5459, May 2022.

[2] F. Abbas, X. Yuan, M. S. Bute and P. Fan, "Performance Analysis Using Full Duplex Discovery Mechanism in 5G-V2X Communication Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 11453-11464, Aug. 2022.

[3] K. Sehla, T. M. T. Nguyen, G. Pujolle and P. B. Velloso, "Resource Allocation Modes in C-V2X: From LTE-V2X to 5G-V2X," *IEEE Internet of Things Journal*, vol. 9, no. 11, pp. 8291-8314, June 2022.

[4] K. Zhang, J. Cao, H. Liu, S. Maharjan and Y. Zhang, "Deep Reinforcement Learning for Social-Aware Edge Computing and Caching in Urban Informatics," *IEEE Transactions on on Industrial Informatics*, vol. 16, no. 8, pp. 5467-5477, Aug. 2020.

[5] Y. Zhao, W. Liu, B. Li, X. Zhou, Z. Ning, T. Qiu, and M. Atiquzzaman, "Entity and Sociality Trust-Aware Model for Content Distribution in Social Internet of Vehicles," *IEEE Transactions on Vehicular Technology*, pp. 1-11, Early Access, 2022.

[6] J. Li, M. Liu, J. Lu, F. Shu, Y. Zhang, S. Bayat, and D. N. K. Jayakody "On Social-Aware Content Caching for D2D-Enabled Cellular Networks With Matching Theory," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 297-310, Feb. 2019.

[7] R. Hou, S. Zhou, M. Cui, L. Zhou, D. Zeng, J. Luo, and M. Ma, "Data Forwarding Scheme for Vehicle Tracking in Named Data Networking," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 7, pp. 6684-6695, July 2021.

[8] L. Yao, Y. Wang, X. Wang and G. Wu, "Cooperative Caching in Vehicular Content Centric Network Based on Social Attributes and Mobility," *IEEE Transactions on Mobile Computing*, vol. 20, no. 2, pp. 391-402, Feb. 2021.

[9] C. Xu, C. Gao, Z. Zhou, Z. Chang and Y. Jia, "Social Network-Based Content Delivery in Device-to-Device Underlay Cellular Networks Using Matching Theory," *IEEE Access*, vol. 5, pp. 924-937, 2017.

[10] X. Wang, and Y. Li, "Content Delivery Based on Vehicular Cloud," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2105-2113, Feb. 2020.

[11] I. Turcanu, T. Engel and C. Sommer,"Adaptive content seeding for information-centric networking under high topology dynamics," *IEEE Vehicular Technology Magazine*, vol. 16, no. 2, pp. 68-75, June 2021.

[12] N. Vo, T. Phan, M. Bui, X. Dang, N. T. Viet and C. Yin, "Social-Aware Spectrum Sharing and Caching Helper Selection Strategy Optimized Multicast Video Streaming in Dense D2D 5G Networks," *IEEE Systems Journal*, pp. 1-12, 2020.

[13] N. Ullah, K. Xiangjie, N. Zhaolong, T. Amr, A. Mubarak, and X. Feng,"Emergency warning messages dissemination in vehicular social networks: A trust based scheme," *Vehicular Communications*, vol. 22, pp. 100199-100215, April 2020.

[14] A. Gulati, G. S. Aujla, R. Chaudhary, N. Kumar, M. S. Obaidat and A. Benslimane,"DiLSe: Lattice-Based Secure and Dependable Data Dissemination Scheme for Social Internet of Vehicles," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 6, pp. 2520-2534, Dec. 2021.

[15] L. Su, Y. Niu, Z. Hou, B. Ai, R. He, Y. Wang, N. Wang, and X. Su, "Content Distribution Based on Joint V2I and V2V Scheduling in mmWave Vehicular Networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 3201-3213, March 2022.

[16] P. Pawar, and A. Trivedi, "Joint Uplink and Downlink Resource Allocation for D2D Communication Underlying Cellular Network," *IEEE Transactions on Communications*, Early Access, Oct. 2021.

[17] Q. Luo, C. Li, T. H. Luan, and W. Shi, "EdgeVCD: Intelligent Algorithm-Inspired Content Distribution in Vehicular Edge Computing Network," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5562-5579, June 2020.

[18] A. Pal and K. Kant, "A Neighborhood Aware Caching and Interest Dissemination Scheme for Content Centric Networks," *IEEE Transactions on Network and Service Management*, vol. 18, no. 3, pp. 3900-3917, Sept. 2021.

[19] D. Wu, L. Zhou, Y. Cai and Y. Qian,"Optimal Content Sharing Mode Selection for Social-Aware D2D Communications," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 910-913, Dec. 2018.

[20] Y. Li, E. Schofield, and M. Gönen, "A Tutorial on Dirichlet Process Mixture Modeling," *Journal of mathematical psychology*, vol. 91, pp. 128–144, 2019.

[21] S. J. Gershman and D. M. Blei, "A Tutorial on Bayesian Nonparametric Models," *Journal of mathematical psychology*, vol. 56, no. 1, pp. 1–12, 2012.

[22] Y. Lai, Y. Ping, K. Xiao, B. Hao, and X. Zhang, "Neurocomputing Variational Bayesian Inference for a Dirichlet Process Mixture of beta Distributions and Application," *Neurocomputing*, vol. 278, pp. 23–33, 2018.

[23] Z. Ma, Y. Lai, W. B. Kleijn, Y. Song, L. Wang and J. Guo, "Variational Bayesian Learning for Dirichlet Process Mixture of Inverted Dirichlet Distributions in Non-Gaussian Image Feature Modeling," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 2, pp. 449-463, Feb. 2019.

[24] H. Dai, W. Wang, Q. Xu, Y. Xiong, and D. Qing, "Estimation of Probability Distribution and Its Application in Bayesian Classification and Maximum Likelihood Regression," *Interdiscip. Sci. Comput. Life Sci.*, vol. 11, no. 3, pp. 559–574, 2019.

[25] G. Khanna, S. K. Chaturvedi, and S. Soh, "Reliability Evaluation of Mobile Ad Hoc Networks by Considering Link Expiration Time and Border Time," *Int. J. Syst. Assur. Eng. Manag.*, vol. 10, no. 3, pp. 399–415, 2019.

[26] W. Li, L. Liu, X. Cao, X. Tian, and Y. Cheng, "Sociality-Aware Resource Allocation for Device-to-Device Communications in Cellular Networks." *IET Communications*, vol. 9, no. 3 pp. 342-349, 2015.

[27] T. Karagiannis, J. Le Boudec and M. Vojnović, "Power Law and Exponential Decay of Intercontact Times between Mobile Devices," *IEEE Transactions on Mobile Computing*, vol. 9, no. 10, pp. 1377-1390, Oct. 2010.

[28] F. Yang, J. Han, X. Ding, Z. Wei and X. Bi, "Spectral Efficiency Optimization and Interference Management for Multi-Hop D2D Communications in VANETs," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6422-6436, June 2020.

[29] A. A. Siddig, A. S. Ibrahim and M. H. Ismail, "Optimum Resource Allocation for Full-Duplex Vehicular Communication Networks," *IEEE Access*, vol. 8, pp. 146683-146696, 2020.

[30] R. Geng, H. Ren and J. Yan,"User Satisfaction-Aware Resource Allocation for 5G Green Vehicle Platooning," *in Proc. 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Atlanta, GA, USA, 2019, pp. 868-873.

[31] J. Ren, F. Xia, X. Chen, J. Liu, M. Hou, A. Shehzad, N. Sultanova, and X. Kong, "Matching Algorithms: Fundamentals, Applications and Challenges," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 3, pp. 332-350, June 2021.

[32] H. Wu, J. Chen, W. Xu, N. Cheng, W. Shi, L. Wang, and X. Shen, "Delay-Minimized Edge Caching in Heterogeneous Vehicular Networks: A Matching-Based Approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6409-6424, Oct. 2020.

[33] M. S. Bute, P. Fan, G. Liu, F. Abbas and Z. Ding,"A Collaborative Task Offloading Scheme in Vehicular Edge Computing," *in Proc. 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, Helsinki, Finland, 2021, pp. 1-5.

[34] Y. Wang, P. Lang, D. Tian, J. Zhou, X. Duan, Y. Cao, and D. Zhao, "A Game-Based Computation Offloading Method in Vehicular Multiaccess Edge Computing Networks," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 4987-4996, June 2020.

[35] *3rd Generation Partnership Project: Technical Specification Group Radio Access Network: Study LTE-Based V2X Services: (Release 14)*, Standard 3GPP TR 36.885 V2.0.0, Jun. 2016.

[36] L. Liang, G. Y. Li and W. Xu,"Resource Allocation for D2D-Enabled Vehicular Communications," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 3186-3197, July 2017.

[37] C. Guo, L. Liang and G. Y. Li,"Resource Allocation for Vehicular Communications With Low Latency and High Reliability,"*IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 3887-3902, Aug. 2019.

[38] Y. Sun, F. Wang, and Z. Liu, "Coalition Formation Game for Resource Allocation in D2D Uplink Underlaying Cellular Networks," *IEEE Communications Letters*, vol. 23, no. 5, pp. 888-891, May 2019.

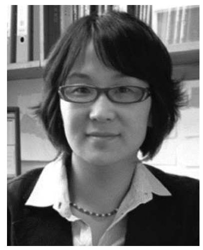
[39] W. Sun, D. Yuan, E. G. Ström and F. Brännström, "Cluster-Based Radio Resource Management for D2D-Supported Safety-Critical V2X Communications,"*IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2756-2769, April 2016.

[40] M. S. Bute, P. Fan, G. Liu, F. Abbas and Z. Ding "A cluster-based cooperative computation offloading scheme for C-V2X networks," *Ad Hoc Networks*, vol. 132, pp. 102862-102874, April 2022.

Muhammad Saleh Bute (Student Member, IEEE) received an M.Sc. degree in Computer Science from the University of Ilorin, Kwara State, Nigeria, in 2014. He is currently pursuing a Ph.D. degree with the Key Laboratory of Information Coding and Transmission, School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China. He is also an academic staff with the Computer Science Department, Gombe State University, Gombe State, Nigeria. His research interests include mobile edge computing (MEC), vehicular ad-hoc networks (VANETs), 5G/6G technologies, cellular vehicle-to-everything (C-V2X) technology, intelligent transportation system (ITS), game theory, fuzzy systems, auction theory, radio resource management and federated learning for wireless networks. Muhammad is a member of the IEEE Young Professionals and The Nigeria Computer Society (NCS). He is also an Active Reviewer of the IEEE TRANSACTIONS ON MOBILE COMPUTING and the IEEE Vehicular Technology Conference (VTC).

Pingzhi Fan (Fellow, IEEE) received the M.Sc. degree in computer science from Southwest Jiaotong University, Chengdu, China, in 1987, and the Ph.D. degree in electronic engineering from the University of Hull, Kingston upon Hull, U.K., in 1994. He is currently a Distinguished Professor and the Director of the Institute of Mobile Communications, Southwest Jiaotong University, China. He has been a Visiting Professor with the University of Leeds, Leeds, U.K., since 1997, and a Guest Professor with Shanghai Jiaotong University since 1999. He has authored or coauthored more than 300 research articles in various international journals and eight books. He is the inventor of 25 granted patents. His research interests include high mobility wireless communications, massive random access techniques, and signal design and coding. He is an IEEE VTS Distinguished Speaker from 2019 to 2022 and a fellow of IET, CIE, and CIC. He was a Board Member of IEEE Region 10, IET(IEE) Council, and IET Asia-Pacific Region. He was a recipient of the U.K. ORS Award in 1992, the NSFC Outstanding Young Scientist Award in 1998, the IEEE VTS Jack Neubauer Memorial Award in 2018, the IEEE SP Soc SPL Best Paper Award in 2018, the IEEE WCSP 10-Year Anniversary Excellent Paper Award from 2009 to 2019, and the IEEE or CIC ICCC Best Paper Award in 2020. He was the general chair or the TPC chair of a number of international conferences. He is the Founding Chair of IEEE VTS Beijing Chapter and IEEE ComSoc Chengdu Chapter, and IEEE Chengdu Section.

Quyuan Luo received the PhD degree in communication and information system from Xidian University, Xi'an, China, in 2020. He had been a visiting scholar with computer science, Wayne State University, USA from 2019 to 2020. He is currently an assistant professor with the School of Information Science and Technology, Southwest Jiaotong University. His current research interests include intelligent transportation systems, content distribution, edge computing and resource allocation in vehicular networks.



Li Zhang (Senior Member, IEEE) received her PhD in Communications at the University of York in 2003. Currently she is an Associate Professor and leads the Wireless Communication Group at the school of Electronic and Electrical Engineering, University of Leeds, UK. Her research interest is focused on wireless communications and signal processing techniques, such as massive MIMO, mmWave communications, Heterogeneous Network, Device to Device communications and 5G systems etc. She has served on the Technical Programme Committees of most major IEEE conferences in communications since 2006 and also an associate editor of IEEE journal. She has been selected as a member of the prestigious UK EPSRC Peer Review College since 2006, and regularly helps reviewing grant applications for Research councils and book proposals. She has been PhD examiner for numerous Universities. In 2005, she received a Nuffield award for a newly appointed lecturer. In 2006, she became a fellow of Higher Education Academy. In 2011, she was awarded as IEEE exemplary reviewer and in 2012 she was promoted as senior IEEE member.

Fakhar Abbas (Senior Member, IEEE) received the Ph.D. degree in Information and Communication Engineering from the School of Information Science and Technology, Southwest Jiaotong University (SWJTU), China, in 2020 and the M.S. degree in Information and Communication Engineering from Harbin Engineering University, Harbin, China, in 2015. He is currently a Research Fellow with the Centre for Trusted Internet & Community (CITIC), National University of Singapore (NUS), Singapore. He has been a reviewer for several IEEE journals and major conferences. He was a Senior Lecturer at the Computer Science Department, COMSATS University Islamabad (CUI), Islamabad Campus, Pakistan. His current research interests include Artificial Intelligence (AI), Machine learning (ML), deepfake detection and generation algorithm design, Data Science, governance of autonomous vehicles (AVs), wireless communications, Resource management for 5G/6G cellular networks, full duplex, and algorithm designs for V2X networks. He was a recipient of the Best Presentation Award of the International Academy of Computer Technology (IACT) International Conference 2015.