Distance-Based Resource Allocation for Vehicle-to-Pedestrian Safety Communication

Usman Ali Khan and Sang Sun Lee *

Department of Electronics & Computers, Hanyang University, Seoul 04763, Korea; usmanali@hanyang.ac.kr
* Correspondence: ssnlee@hanyang.ac.kr

check for updates

Abstract: Cellular Vehicle to Everything (V2X) has redefined the vehicular communication architecture as something that needs an ultra-reliable link, high capacity, and fast message delivery in vehicular networks. The V2X scenarios are broadly categorized as Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I), Vehicle to Pedestrians (V2P), and Vehicle to Network (V2N). Vulnerable pedestrians belong to the V2P category and hence require an ultra-reliable link and a fast message delivery in case the moving vehicle is in the close proximity of the pedestrian. However, congestion in the network calls for an optimized resource allocation that would allow a fast and secure connection between a vehicle and the pedestrian. In this paper, we have proposed a distance-based resource allocation that classifies the pedestrians in different categories, performs a one-to-many weighted bipartite matching, and finally a reinforcement learning based power allocation.

1. Introduction

Vehicular communication has seen a recent shift in terms of technology with the introduction of the cellular V2X technology proposed by the 3rd Generation Partnership Project (3GPP). The 5th Generation cellular technology has paved a way for the inclusion of the vehicles in the cellular domain [1]. Release 14 introduced the concept of V2X communication for the first time by including the air interfaces and core network technologies to support the V2X communication. Cellular V2X technology will not only resolve the data rate issues for onboard services but will also improve the safety communication by allowing two users in the vicinity of each other to have direct communication with each other. The device-to-device (D2D) communication (introduced earlier in Release 12 of the 3GPP) allows any two cellular users (within a certain range) to share safety information with each other. The D2D communication is basically the driving force for the V2X communication with out-of-coverage and partial coverage scenarios, the D2D communication ensures the safety of the cellular users in remote and out-of-bound areas where the cellular coverage is low to none.

Since, the inception of the concept of cellular-based V2X, there is a new room for research in the intelligent transportation system (ITS) area. A shift from the Wi-Fi-based dedicated short-range communication (DSRC) to the seamlessly connected vehicles in C-VTX, has been seen in the research community. A comprehensive survey on the vehicular technologies [2] has highlighted the various technologies along with their standardization bodies, use cases, security aspects, software defined vehicular networks (SDVN), and internet of vehicles (IoV) among other things. An area of particular interest is the radio resource management (RRM) for vehicular communication using cellular technology. The author in [3] has reviewed the challenges involved in the paradigm of radio resource management by an extensive survey of the objectives, techniques, and their achieved performances. According to

the survey V2X communication is still in its preliminary stage and requires an extensive research in areas of RRM, effective handoff mechanisms, and reliable channel signal information.

The V2X technology is broadly categorized in V2V, V2I, V2P, and V2N scenarios. V2V scenario considers two vehicles in vicinity of each other to share safety information in order to avoid accidents or to alert the other vehicle about traffic jams etc. The V2I scenario is simply a vehicle communicating with the infrastructure. Like V2V, V2P scenario is also for safety communication between a vehicle and pedestrian. Whereas, V2N is a vehicle which is accessing the network for cloud-based services. V2V and V2I require ultra-reliability and high capacity, respectively. Other aspects of the V2X communication like vehicle-to-sensors (V2S), vehicle-to-grid (V2G), and vehicle-to-human (V2H) are studied in [4] as an emerging technology in current R&D.

V2P being one of the most crucial scenarios for vulnerable road users (VRUs); i.e., users which are in close proximity of a moving vehicle and are prone to an accident. Pedestrians on the verge of crossing a road, a cyclist nearby or on the road, and pedestrians on a busy street all qualify as VRUs. A recent survey on VRUs in [5] shows the increasing number of fatalities in different countries. The author classifies the VRUs into categories such as cyclist, motorized two wheelers, and pedestrians. Furthermore, the author performs a case study for pre-crash scenarios under different technologies.

While most of the research consider scenarios that include V2I and V2V users, we have aimed to present a roadside scenario, that specifically focus on the V2P links. In our scenario, vehicular user equipment (VUE) share resources with cellular user equipment (CUE). A VUE and CUE form a D2D pair when they are in certain range of each other. This pair is referred to as V2P. V2P which is similar to a V2V link is aimed for highly reliable safety communication. However, our aim here is to categories V2P users based on their vulnerability by using a distance-based classification, and then allocate them resources, accordingly. Unlike traditional network scenarios, our scenario focuses only on the safety communication of VRUs. A user in the vicinity of the D2D range should receive the information regarding the approaching vehicle. The feasible range for D2D Communication is studied in [6]. In this paper, we have taken the range to be 50 meters, which means that any user lying in 50m radius of another user can involve in D2D communication.

In this paper, we have proposed a distance-based resource allocation scheme in a D2D-based V2P scenario. Therefore, our major contribution is:

- Optimum power allocation using Q-learning based reinforcement learning with different discount factors.
- A one-to-many weighted bipartite matching scheme (with maximum flow) to create connections between users for frequency assignment.

The rest of the paper is organized as follows: Section 2 is related to current research trends in this area. Section 3 describes the network and system model. Section 4 is the classification of the VRUs. Section 5 is regarding optimal power assignment based on the classification of users in Section 4. Section 6 introduces a one-to-many bipartite matching scheme. Section 7 will present the simulation setup and results. Finally, Sections 8 and 9 discusses future directions and conclusions, respectively.

2. Related Work

Keeping in mind the differentiated quality of service (QoS) of the different communication scenarios, resource allocation among users becomes a major milestone. Resource allocation in LTE is the distribution of resources like power and frequency among the users, such that the method is optimum and solves the network congestion problem. A survey on radio resource allocation for V2X communication [7] thoroughly investigated the communication modes, radio resource management techniques, and discussions on different methodologies. Resource allocation for D2D-based vehicular communication has been deeply researched in the recent past with techniques like traditional optimization techniques [8], matching [9], graph theory [10], and machine learning [11] being used to investigate the performance of the network under varying scenarios. A resource allocation based on

an immune algorithm for D2D-based vehicular networks has been in investigated in [12], where the authors have incorporated an adaptive cloning and population update strategy to provide a highly efficient method for solving resource sharing among vehicles. Power control using a distributed deep deterministic policy gradient method [13], introduced two models to solve a multi-agent energy-efficient power allocation problem. The author proposed a model that employs neural networks to overcome problems with existing approaches.

Resource allocation with low-latency vehicular communication has been studied by several authors. Low latency with packet retransmission is investigated in [14], where the authors have presented the queuing analysis and then derived an expression for average packet sojourn time. A twin-timescale scheme for low-latency vehicular networks [15], aimed for reduction of maximum transmission latency by using a two-stage process, by first minimizing the worst-case transmission latency and then the base station allocating a total power at short-term timescale. The idea for using a twin-timescale is to provide a more realistic approach for avoiding frequent exchange of near instantaneous channel state information. hybrid strategy, whereby, cellular-based V2V communications is introduced to IEEE 802.11p-based vehicular networks [16] to improve network latency, is also being studied.

Matching and graph theory are among the most researched areas in D2D-based vehicular communication. Both techniques have primarily been used for frequency assignment problems. Frequency resource blocks (RB) are shared among users in the cellular network. In order to avoid communication impairment due to interference, the RB are assigned in an optimum manner. Graph theory-based resource assignment has been studied in both two and three dimensions. Interference is usually taken as a metric to define the edges between the vertices in a graph. Interference hypergraph-based 3D matching resource allocation for V2X [17] using a weighted 3-partite interference hypergraph based on greedy and iterative matching, has been shown to improve the network throughput.

Dynamic proximity-aware resource allocation in V2V communications [18] employs zone formation based on traffic patterns for vehicles. Then, a matching game is proposed to allocate resources to V2V pairs within each zone. The zone formation reduces interference and signaling overhead and hence satisfies the quality of service in terms of SINR. Distance-based power control schemes have been employed using stochastic geometry [19] in a D2D setting by using distance-dependent path loss parameters. Location information has been used to monitor the QoS requirements in [20] and dynamically modifies resource assignment to interfering links.

Most of the research has focused on creating a network scenario with V2V and V2I users and then the problem is formulated using an objective function and constraints. The constraints usually include power, latency, SINR, and outage probability. The formulated problem is then solved using the techniques mentioned above. While matching and graph theory have specifically been used for frequency assignment problems, other optimization techniques along with machine learning has been used to solve a wide variety of resource sharing problems.

3. Network and System Model

We consider a single cell network scenario with one evolved node B(eNb) in the center of the cell, a road that runs in the middle of the cell, vehicular users on that road, and cellular users distributed randomly throughout the cell. Let c = {1 . . . . . . , C} by the CUE involved in direct communication with the infrastructure and let v = {1 . . . . . . , V} be the V2P users (a pair consisting of a cellular and a vehicular user).

3.1. Network Architecture

The network architecture consists of a vehicle surrounded by several road-side users. The vehicle is involved in direct D2D communication with the road users in the vicinity (V2P). In this network, we also have some cellular users involved in direct communication with the infrastructure. Figure 1 shows the vehicle with road-side users. If the distance between a cellular user and the vehicle is more

than 50 meters, the cellular user is involved in direct communication with the infrastructure (CUE). When the distance becomes 50 meters or less, V2P communication starts.
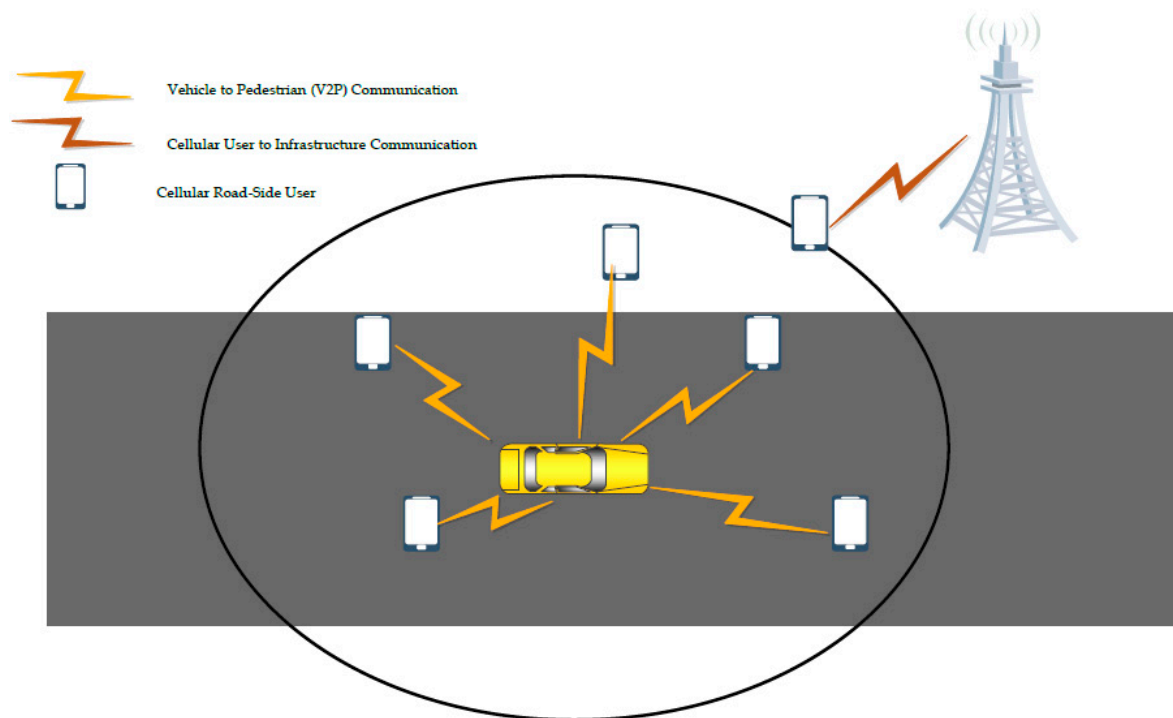

Figure 1. A Road-Side Network Scenario.

3.2. System & Channel Model

Since the vehicle is assumed to be moving, it will encounter both small- and large-scale fading and therefore will have different channel model compared to a cellular user which is static or moving very slowly. The SINR of the two types of users, is given by Equations (1) and (2):

$$\gamma_c = \frac{P_{c,x}\,h_{c,B}}{\sigma^2 + P_{v,x}\,h_{v,c}} \tag{1}$$

$$\gamma_v = \frac{P_{v,x}\,h_v}{\sigma^2 + P_{c,x}\,h_{c,v}} \tag{2}$$

The channel power gains are denoted by *h* in Equations (1) and (2). The subscript '*x*' denotes the frequency resource block (RB) shared among users. We assume that the RB of V2P user can be shared by a CUE, but resource blocks cannot be shared between VUEs. The gain $h_{c,B}$ is the desired channel gain from a CUE user to the base station, $h_{v,c}$ is the interference gain from the D2D pair to the CUE user, h_v is the desired gain between the D2D based V2P pair, and $h_{c,v}$ is interference gain from CUE to the D2D pair. The channel gain, in turn is dependent on pathloss, shadowing, distance, and small-scale fading component; given by Equation (3):

$$h^{a,b} = PL \times Y \times \ell^{-\gamma} \times g^{a,b} \tag{3}$$

where *PL* is the path loss, *Y* is the log-normal shadowing component, ℓ is the distance between transmitter and receiver, and γ is the decay exponent. The path loss models for V2P and CUE are given in Table 1.

Table 1. Path Loss Model.

Parameter	V2P Link	CUE Link
Path loss model	$127 + 30 \log_{10}(\ell)$	$128.1 + 37.6 \log_{10}(\ell)$ [21]

3.3. Problem Formulation

In this section we will define the optimization problem; objective function and the constraints. Traditionally, most problems [9] assign high capacity to the V2I link and high reliability to V2V link. However, our scenario is particularly focused on the V2P links, so our aim is to provide a high capacity and reliable link to the V2P, while providing a minimum threshold SINR to the CUEs:

$$
\begin{aligned}
Max \sum_{v=1}^{V} \sum_{x=1}^{X} & E[log_2(1 + \gamma_v)], \\
0 \leq & P_c \leq Pmax^c \\
0 \leq & P_v \leq Pmax^v \\
Pr\{\gamma_v & \leq \gamma_0\} \leq p \\
& \gamma_c \geq \gamma_0
\end{aligned}
\tag{4}
$$

Equation (4) shows the objective function while equations (4) shows the constraints. In the next section, we have further decomposed Equation (4) into three different parts for each category of user. While all the CUEs get a uniform power depending on the threshold SINR. The maximum allowed power is 23 dBm. Table 2 summarizes the list of symbols with their respective definitions.

Table 2. List of Symbols.

Symbol	Definition
$h_{c,B}$	Desired channel Gain between CUE and BS
$h_{v,c}$	Interference channel gain from V2P to CUE
h_v	Desired channel gain between V2P
$h_{c,v}$	Interference channel gain from CUE to V2P
$P_{c,x}$	CUE transmit power
$P_{v,x}$	V2P transmit power
$h^{a,b}$	Channel gain between two users
PL	Path loss
Y	Log normal shadowing
$\ell^{-\gamma}$	ℓ is the distance between two users and γ is the decay exponent
$g^{a,b}$	Small scale fading component
$Pmax^c$	Maximum allowable transmit power for CUE
$Pmax^v$	Maximum allowable transmit power for V2P
p	Probability of outage
γ_0	Threshold SINR for CUE

4. Classification of Vulnerable Road Users

In this section we will classify the V2P users in three different clusters as shown in Table 3. Type 1 (very critical) V2P users are in most crucial state which requires an ultra-reliable link, Type 2 (critical) users are slightly less vulnerable, and Type 3 users are also classified according to their requirements. The classification is distance dependent (distance between the vehicle and cellular user). A 50-meter distance initiates the V2P communication and thereafter the users are classified accordingly. In the next two sections, we will use this classification for resource allocation.

Table 3. Categories of V2P Users.

	Very Critical	Critical	Normal
V2P Distance	<20 m	20–35 m	35–50 m

In Figure 2, we can see the V2P users classified into three types, depending on their vulnerability. Users involved in direct communication with infrastructure are referred as CUEs. The shortest distance between two users is found using the coordinates during simulation.

In a real-life scenario, a number of methods can be used to calculate the distance between devices. The first and foremost is the global positioning system (GPS) that can be used to find the coordinates. However, accuracy of this method is still low especially when we talk about safety communication. Secondly, received signal strength indicator (RSSI) can be used to estimate the distance between two devices. The author in [22] has used the RSSI-based technique, assuming the absence of the GPS, to find the distances between the vehicles for short-range communication. In our scenario distance is calculated between two users once they enter a D2D range. User equipment (UE) location reporting is defined in the 3GPP TS 23.303 standard, where UE reports its location on periodic basis to their corresponding servers [23]. This information is available at the network level.
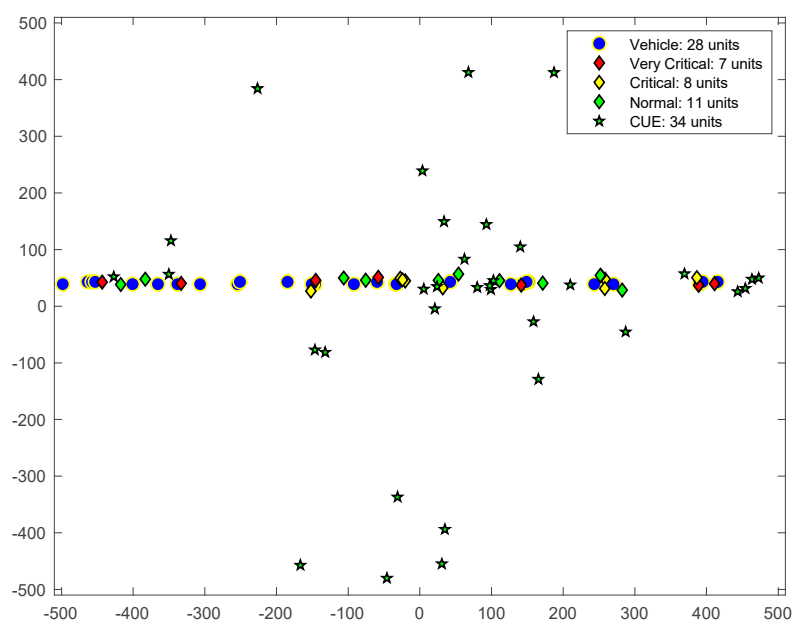

Figure 2. V2P user classification.

5. Reinforcement Learning Based Power Allocation

Reinforcement larning (a type of machine learning) has recently been used in the domain of resource allocation in wireless networks. Q-learning, which is basically a subtype of reinforcement learning has found its way to wireless communication problems. If we specifically talk about D2D communication, some authors have researched on that including the author in [24] has used deep reinforcement learning for allocating resources to vehicles in a network scenario composed of V2V links with stringent latency and V2I links with high capacity requirements. The author in [25] has used a distributed reinforcement learning method, where each agent in the network keeps its own table of Q-values.

Q-Learning involves agents, actions, states, and rewards. The learning in this method is based on interaction with the environment. It is represented by a tuple < *S, A, T, R(s, a)* >, where *S* is the set of states in the environment, *A* is the agent, *T* is a state transition probability, and *R* is the reward function.

5.1. Learning Process

An agent interacts with the environment by sensing the current state and taking an action (according to a policy). This results in environment transitioning in a new state and an agent is rewarded accordingly. The process is repeated till we have an optimum policy. The learning process under Q-learning [26] is given by Equations (5)–(10).

Under a policy π, a state is given by Equation (5):

$$V^{\pi}(s) = E_{\pi}\{r_t|s_t = s\} = E_{\pi}\left\{\sum_{k=1}^{\infty} \eta^k r_{t+k+1}|s_t = s\right\} \tag{5}$$

where η is the discount factor with value ranging between 0 and 1. The discount factor helps to determine the future reward. A discount factor of 0 will make the agent strive for only current rewards, whereas, a discount factor near to 1, will make the agent strive for long term rewards. $V\pi(s)$ is the expected discounted reward. There exists at least one optimal policy π^*, such that:

$$\boldsymbol{V}^*(\boldsymbol{s}) = \boldsymbol{V}^{\pi}(\boldsymbol{s}) = \boldsymbol{max}_a\left\{\boldsymbol{R}(\boldsymbol{s},\boldsymbol{a}) + \boldsymbol{\eta} \sum_{s' \in S} \boldsymbol{P}_{s,s'}(\boldsymbol{a})\boldsymbol{V}^*(\boldsymbol{s})\right\} \tag{6}$$

$P_{s,s'}(a)$ is a transition probability from state s to s'. The action values or the Q-values is the expected return of taking action 'a' in state 's' under the policy 'π':

$$Q^{\pi}(s,a) = R(s,a) + \eta \sum_{s' \in S} P_{s,s'}(a)Q(s',a) \tag{7}$$

and the optimal policy $Q^*(s,a)$ is given by:

$$\boldsymbol{Q}^*(\boldsymbol{s},\boldsymbol{a}) \equiv \boldsymbol{Q}^{\pi^*}(\boldsymbol{s},\boldsymbol{a}), \forall \boldsymbol{s}, \boldsymbol{a} \tag{8}$$

and so, we get:

$$V^*(s) = max_a Q^*(s,a) \tag{9}$$

The update rule, that helps the Q-learning algorithm (by adjusting the Q-values) is given by:

$$Q_{t+1}(s_t,a_t) = \begin{cases} Q_t(s_t,a_t) + \alpha[r_{t+1} + \gamma max_{a'} Q_t(s_{t+1},a_{t+1}) - Q_t(s_t,a_t)] & if \ s = s_t \ and \ a = a_t \\ Q_t(s_t,a_t) & otherwise \end{cases} \tag{10}$$

Here, α denotes the learning rate. The values of learning rate are between 0 and 1. A learning rate of 0 means that the agent has learnt nothing, while a learning rate of 1 means the agent only considers the most recent information. The update rule in Equation (10) is based on the steps taken by an agent to reach a terminal point. The Q-values are adjusted based on the difference between discounted new values and the old values. The parameter gamma is used to discount the new values and the step size is adjusted using learning rate.

5.2. Proposed Methodology

In this section, we will solve the optimization problem presented in Section 3 using the distance-based classification in Section 4 to find the optimum power. A decentralized approach, where each user maintains its own Q-values table has been used since a centralized approach would cause an excessive overhead [23]. Firstly, we need to define the states, agents, actions, and reward function in our scenario.

Agent: The V2P users will act as agents. We have v numbers of users engaged in V2P communication. Therefore, $1 \leq v \leq V$ number of agents in the network.

State: User v in a state S, on resource block x, at time t with an interference level of I_t^x is given by:

$$S_t^{v,x} = \{I_t^x\} \tag{11}$$

The interference level is given by:

$$I_t^x = \begin{cases} 1 & \gamma_c \geq \gamma_0 \\ 0 & otherwise \end{cases} \tag{12}$$

The equation above guarantees the minimum required SINR for the cellular users.

Action: The actions consists of the uplink power of the V2P users.

$$P = \left(P_1^v, P_2^v, \ldots\ldots\ldots P_n^v \right) \tag{13}$$

where the subscript n denotes that we have 'n' number of powers.

Reward Function: The reward function consists of our objective function, that was mentioned in Section 3.

$$R^{v2p} = \begin{cases} \frac{1}{C} \log_2 (1 + \gamma_v), & \gamma_c \geq \gamma_0 \\ 0, & otherwise \end{cases} \tag{14}$$

Learning Rate and Discount Factors: We have used a constant learning rate $\alpha = 0.5$. While we used three different discount factors ($\eta = 0.8, 0.5, 0.3$) based on the vulnerability of the users.
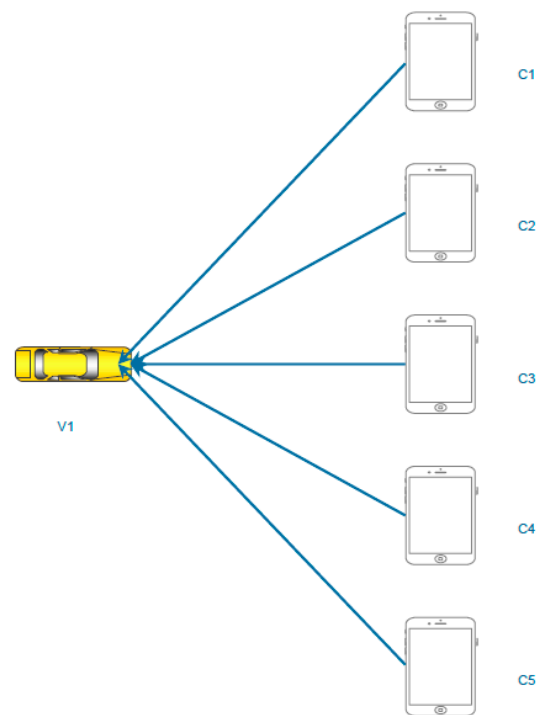
6. One-to-Many Bipartite Matching

In the last section, distance-based optimum power was assigned to the users. However, since the user share the uplink frequency resource blocks, interference can degrade the performance of the system. Therefore, we aim to perform a pair matching technique that can cater the interference by using matching. Matching has been used extensively in various wireless network problems [27]. It has mostly found its way in problems related to resource sharing. Most matching problems are two dimensional that employ one-to-one matching like the Hungarian algorithm used in [9]. While some authors have used three-dimensional matching for resource allocation in more complex problems [28,29]. In this paper, we have used a variant of the bipartite matching known as one-to-many bipartite matching with maximum flow. The motivation behind using the one-to-many bipartite matching is inspired by the coalition formation for multi robot task in [30], where a single task is assigned to multiple robots. We have taken the key idea from here to create connections and assign weights.

6.1. Graph Structure

A single vehicle is communicating with multiple cellular users. For the sake of simplicity, we have assumed that one vehicle can communicate with up to five cellular users, at a given time. Let $V = \{v_1, v_2, \ldots, v_n\}$ be the set of vehicles and $C = \{c_1, c_2, \ldots, c_n\}$ be the set of cellular users. A cellular user is connected to a vehicle when it enters into a 50m radius from the vehicle, forming an edge. Let $G = (\{X,Y\},E,W)$ be a bipartite graph with X vertices corresponding to cellular users and Y vertices corresponding to the vehicles. E is an edge set that consists of pairs with one vertex from each X and Y. The weight W of the edge is defined by the distance between the vehicle and cellular user. A cellular user is allocated to at most one vehicle at a time. Our main objective is to assign multiple edges to a single vertex (vehicle). The idea is illustrated in Figure 3. The vertex v1 is shared by c1, . . . , c5 and the edges between the vehicle and cellular user are denoted by the blue line.

Figure 3. One-to-many bipartite matching.

6.2. Algorithm

The algorithm takes a bipartite graph (W(C,V)) as input. The output is an array that maps a vehicle to the cellular users. For each V that is available, a C is connected to that V (if the vehicle has less than 5 connections). A group is complete once a V connects to five Cs at a given time. We decomposed the bipartite matching problem by converting it into maximum flow problem using the greedy Ford Fulkerson maximum flow method [31]. The pseudo code is shown in Algorithm 1.

A vehicular user acts as a central controller with a W(C,V) matrix. This matrix has rows and columns showing a distance relationship between the cellular user and a vehicular user. The matching algorithm runs on each vehicle. For every cellular user (C), the assignVeh(C) function is called. A vehicle (V) is available ($V_{available}$) if the W(C,V) matrix has an entry greater than zero and the vehicle has not been visited before. Each available vehicle is then put into three different groups based on the number of connections ($Conn^v$) it has with the cellular users. If number of connections is less than 1 (min_connection), available vehicle is directly assigned to Group 1. If the number of connections is more than 1 and less than 5 (max_connection), the available vehicle is assigned to Group 2. If number of connections is greater than equal to 5, than the available vehicle is assigned to Group 3. If a group is full but there is a cellular user (C′) that is more vulnerable than any of the current members of the group, then it is assigned to that group with a replacement weight, replacing the least vulnerable user of that group. Once the vehicles are all assigned to the groups, the C is connected to the optimum vehicle ($V_{optimum}$). Once each group is full (maximum connections), distance wise sorting is done. Maximum weight is assigned to the C closest in distance to the vehicle.

Algorithm 1. One-to-Many Bipartite Matching

Procedure: createConnections

        Inputs: $W(C,V)$: distance relation between all V and C

        for each C unit **do**

        CALL "assignVeh(C)" to connect C with optimum V unit

Procedure: assignVeh(C)

    for each V **do**

          if $W(C,V) > 0$ AND (NOT previously visited)

          $V_{available} \leftarrow V$

        for each $V_{available}$ **do**

            $Conn^V \leftarrow$ Number of C connected to $V_{available}$

            if $Conn^V <$ min_connection

            $V_{group(1)} \leftarrow V_{available}$

            elseif $Conn^V \geq$ min_connection AND $Conn^V <$ max_connection

            $V_{group(2)} \leftarrow V_{available}$

            else $Conn^V \geq$ max_connection

            $V_{group(3)} \leftarrow V_{available}$

        if $V_{group(1)}$ is NOT empty

           $V_{optimum} \leftarrow V_{group(1)}$ with max weight (lowest distance from current C)

           Connect C and $V_{optimum}$

           RETURN TRUE

        if $V_{group(2)}$ is NOT empty

           $V_{optimum} \leftarrow V_{group(2)}$ with max weight (lowest distance from current C)

           Connect C and $V_{optimum}$

           RETURN TRUE

        if $V_{group(3)}$ is NOT empty

           Sort $V_{group(3)}$ according to their distance cluster in ascending order

           for each $V_{group(3)}$ **do**

              if $Conn^{V(group\_3)} \leq$ replacement weight AND assignVeh(C') == TRUE

                $V_{optimum} \leftarrow V_{group(3)}$

                Connect C and $V_{optimum}$

              RETURN TRUE

      RETURN FALSE

6.3. Complexity Analysis

The motivation behind using this method was the time complexity. The Ford Fulkerson method decomposes the bipartite matching problem into a maximum flow problem. So, the Ford Fulkerson problem can find a maximum flow matching in bipartite graph in *O(EF)* time, where *E* is the number of edges and *F* is maximum flow. The complexity of finding an augmenting path is *O(E)* and computing the bottleneck capacity is *O(F)*. So, we can find maximum partite matching in *O(EF)* time by reducing the problem to network flow.

7. Simulations and Results

In this section, we will check the performance of the methodologies used in this paper. In order to ensure the reliability of users according to their distances, we have assigned three different discount factors in the Q-learning based power control: very critical (v.c = 0.8), critical (c = 0.5), and normal (n = 0.3). We will check the performance in terms of vehicular speed, number of vehicular users, and the threshold SINR mentioned above. Table 4 shows the simulation parameters used in our setup. The simulation was carried out in MATLAB and Python.

Table 4. Simulation parameters.

Parameters	Value
Operating Frequency	2 GHz
Bandwidth	10 MHz
Cell Radius	500 m
Vehicular Speed	60–140 km/h
Discount Factor (η)	v.c = 0.8, c = 0.5, n =0.3
Maximum Transmit Power	23 dBm
AWGN Noise	−114 dBm
SINR Thresholds (γ_0)	5, 7, 10, 15 dBs
Number of Vehicular User Equipment (VUEs)	Variable
Number of Cellular User Equipment (CUEs)	Variable
Learning Rate (α)	0.5
D2D Communication Radius	50 m
Highway Length	1 km
Number of Lanes	1
Number of Resource Blocks	50

7.1. Vehicular Speed vs. V2P Capacity

For total V2P throughput we have taken the SINR of each V2P link and calculated the sum from all V2P UE. Using the equation below:

$$\sum_{v=1}^{V}\left[\log_2(1+\gamma_v)\right] \tag{15}$$

An increasing speed of vehicles lowers the throughput of the V2P links. Figure 4 below shows the performance of increasing speed. The very critical and critical pedestrians still get a higher capacity compared to a normal pedestrian. The benchmark technique used for comparison is the method proposed in [9], where the pair matching is done using Hungarian Algorithm and optimum power is found using bisection search. We have used the same method in our scenario and compared the result with our proposed method.
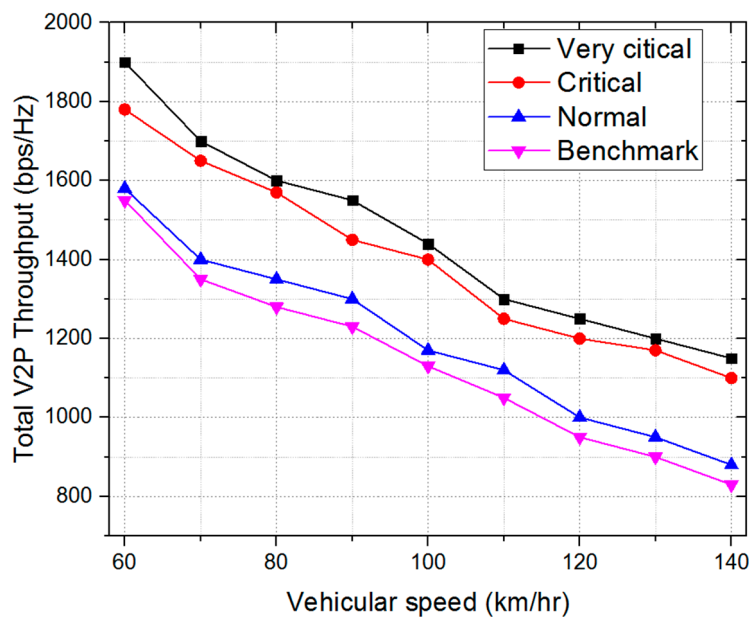
Figure 4. Vehicular Speed vs. V2P throughput.

7.2. Number of Vehicles vs. System Throughput under Variable SINR

The system throughput (Mbps) is calculated by the summation of the total throughput of V2P users and total throughput of CUE given by Equation (16):

$$\left(\sum_{v=1}^{V} \left[\log_2(1 + \gamma_v) \right] + \sum_{c=1}^{C} \left[\log_2(1 + \gamma_c) \right] \right) \tag{16}$$

Cellular users with greater SINR threshold contribute better to the overall system throughput as evident from the formula and shown in Figure 5.
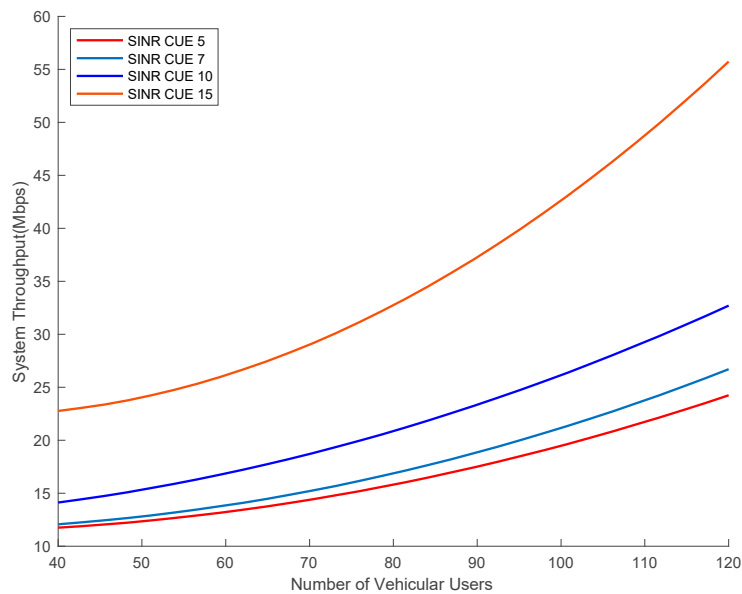

Figure 5. Number of Vehicular Users vs. System Throughput.

7.3. Number of Vehicles vs. V2P Throughput

Increasing the number of vehicles in the network improves the V2P throughput especially for the critical users the Q-learning algorithm performs well as shown in Figure 6 below.
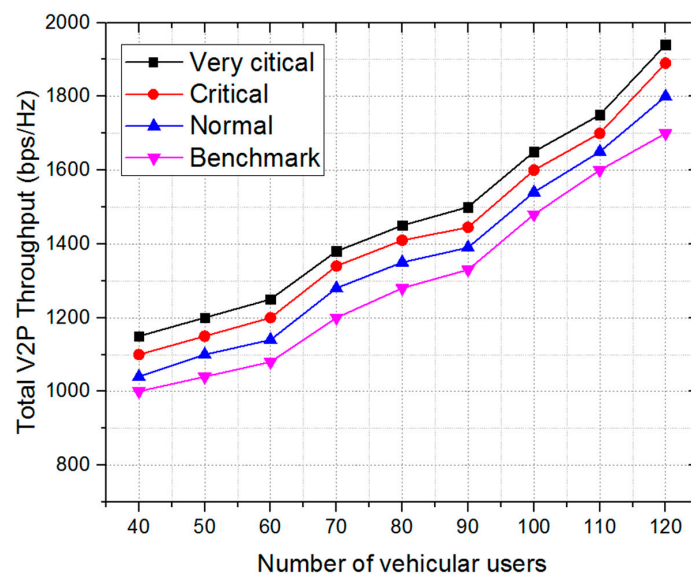

Figure 6. Number of Vehicular Users vs. V2P Throughput.

7.4. Variable Threshold SINRs of CUEs

By varying the SINR threshold, we look at the V2P link capacity under a variable vehicular speed. At high speeds the V2P Capacity decreases. An SINR of 5dBs is therefore chosen as the base threshold, since the capacity is higher compared to other SINR thresholds as shown in Figure 7.
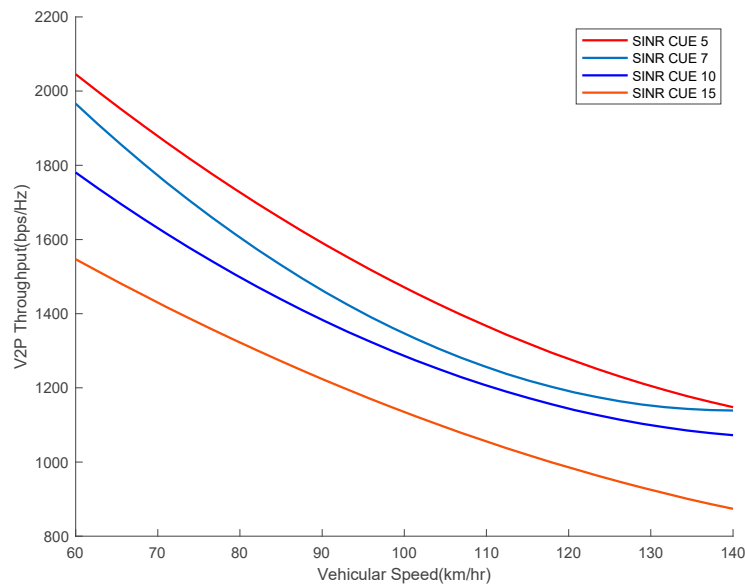

Figure 7. Vehicular Speed vs. V2P throughput.

7.5. Number of VUEs and Reliability

In order to ensure the reliability of the VUEs, we define the term SINR probability as the ratio of VUEs with SINR more than 5 dBs divided by the total number of VUEs.

SINR Probability = Amount of UE with SINR more than 5 db/total amount of UE

Figure 8 shows that higher number of VUEs results in better SINR probability.
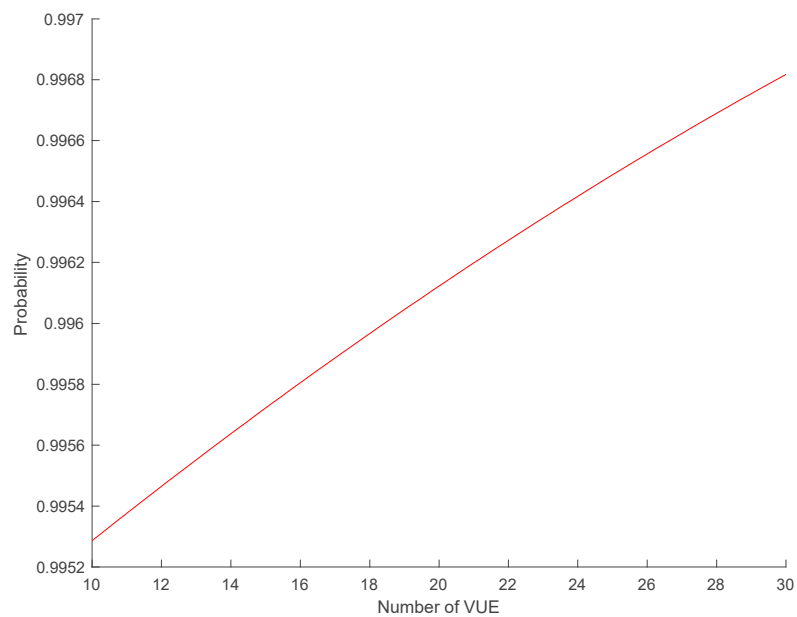

Figure 8. Number of VUE vs. Reliability.

8. Discussion and Future Work

The objective of this paper was to establish a roadside scenario where vulnerable road users are classified according to their distances and allocated resources accordingly. Simulation results show that the performance of our method is efficient compared to a one-to-one matching scheme. We have aimed to show results under varying vehicular speed and number of users. Under varying vehicular speed, our method has performed better than the benchmark [9] scheme. A higher threshold SINR for cellular users can result in better overall system throughput but can result in increased interference to the V2P links and hence decreases the capacity of these links. Therefore, an SINR threshold of 5dB was chosen. Finally, a higher number of VUEs contribute better to the SINR probability of the V2P, which ensures a better capacity of the links. In future, we want to extend this approach for other crucial scenarios like an out-of-coverage or relay assisted scenario. The integration of this scenario with mobile edge computing (MEC) technology will bring newer challenges while overcoming the latency issues. These issues include interoperability between different standardization organizations, deployment, handover, and offloading decision [32].

The resource allocation in D2D-based vehicular networks has seen a tremendous research interest in the recent past. Methods ranging from heuristic optimization [12] to machine learning [33], have been used under varying sets of users and their requirements. Machine learning, being the current hot topic in this domain; centralized vs. distributed reinforcement learning, multi-agent deep reinforcement learning [34], and other methods that could provide a faster convergence with better performance than traditional methods. Relay-assisted [35], out-of-coverage scenarios [36] and energy efficient resource allocation [37] is currently being researched upon. Resource allocation in UAV- enabled vehicular communication is another new area of research being investigated for emergency coverage [38]. Resource allocation for virtual reality (VR) content sharing in D2D multicast communication [39] is another interesting area of research where the authors have created D2D-based multicast clusters for video content sharing. This approach could be integrated in a vehicular network scenario in the future. Hybrid network architectures [40] including the already in-service DSRC along with emerging LTE V2X could call for a more complex resource management. Channel modeling for V2X need more research due to the delay in the availability of channel state information (CSI) at the base station (due to high vehicular speed). Moreover, interference mitigation techniques to overcome interference from adjacent frequencies and co-occurring technologies, need further research and consideration. Nevertheless, there is a vast room for research in the area of cellular V2X, which will facilitate in the implementation of advanced applications and become a major enabler for intelligent transportation system (ITS) and autonomous driving.

9. Conclusions

In this paper we have proposed a D2D-based V2P resource allocation based on the distance between vehicles and road-side pedestrians. We have performed two separate tasks after classifying the users on the basis of their vulnerability; firstly, we have performed a Q-learning-based reinforcement learning for power allocation. A one-to-many bipartite matching has been performed for pair assignment. The matching allows the users to form the connections based on their weights (depending on distance) and the reinforcement learning allocates the optimum power by giving a higher discount factor to the most vulnerable pedestrians. Users outside the D2D range are treated as normal cellular users with a minimum guaranteed quality of service. The results have shown that our method provides a higher throughput to the most vulnerable links while taking other factors such as threshold SINR, vehicular speed, and number of users into account.

References

1. Höyhtyä, M.; Apilo, O.; Lasanen, M. Review of Latest Advances in 3GPP Standardization: D2D Communication in 5G Systems and Its Energy Consumption Models. *Future Internet* **2018**, *10*, 3. [CrossRef]
2. Singh, P.K.; Nandi, S.K.; Nandi, S. A tutorial survey on vehicular communication state of the art, and future research directions. *Veh. Commun.* **2019**, *18*, 100164. [CrossRef]
3. Loussaief, F.; Marouane, H.; Koubaa, H.; Zarai, F. Radio resource management for vehicular communication via cellular device to device links: Review and challenges. *Telecommun. Syst.* **2020**, *73*, 607–635. [CrossRef]
4. Lozano Domínguez, J.M.; Mateo Sanguino, T.J. Review on V2X, I2X, and PSX Communications and Their Applications: A Comprehensive Analysis over Time. *Sensors* **2019**, *19*, 2756.
5. Sewalkar, P.; Seitz, J. Vehicle-to-Pedestrian Communication for Vulnerable Road Users: Survey, Design Considerations, and Challenges. *Sensors* **2019**, *19*, 358. [CrossRef]
6. Ding, H.; Ma, S.; Xing, C. Feasible D2D communication distance in D2D-enabled cellular networks. In Proceedings of the 2014 IEEE International Conference on Communication Systems, Macau, China, 19–21 November 2014; pp. 1–5.
7. Masmoudi, A.; Mnif, K.; Zarai, F. A Survey on Radio Resource Allocation for V2X Communication. *Wirel.Commun. Mob. Comput.* **2019**, *2019*, 2430656. [CrossRef]
8. Yu, N.; Mei, J.; Zhao, L.; Zheng, K.; Zhao, H. Radio resource allocation for D2D-based V2V communications with Lyapunov optimization. In Proceedings of the 2017 IEEE/CIC International Conference on Communications in China (ICCC), Qingdao, China, 22–24 October 2017; pp. 1–6.
9. Liang, L.; Li, G.Y.; Xu, W. Resource Allocation for D2D-Enabled Vehicular Communications. *IEEE Trans. Commun.* **2017**, *65*, 3186–3197. [CrossRef]
10. Liang, L.; Xie, S.; Li, G.Y.; Ding, Z.; Yu, X. Graph-Based Resource Sharing in Vehicular Communication. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 4579–4592. [CrossRef]
11. Souhir, F.E.; Belghith, A.; Zarai, F. A Reinforcement Learning-based Radio Resource Management Algorithm for D2D-based V2V Communication. In Proceedings of the 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 24–28 June 2019; pp. 1367–1372.
12. Li, X.; Sun, Y.; Zhou, L.; Qi, A.; Zhou, S. A Resource Allocation Scheme Based on Immune Algorithm for D2D-Based Vehicular Communication Networks. *IEEE Access* **2019**, *7*, 122536–122543. [CrossRef]
13. Nguyen, K.K.; Duong, T.Q.; Vien, N.A.; Le-Khac, N.; Nguyen, L.D. Distributed Deep Deterministic Policy Gradient for Power Allocation Control in D2D-Based V2V Communications. *IEEE Access* **2019**, *7*, 164533–164543. [CrossRef]
14. Guo, C.; Liang, L.; Li, G.Y. Resource Allocation for High-Reliability Low-Latency Vehicular Communications With Packet Retransmission. *IEEE Trans. Veh. Technol.* **2019**, *68*, 6219–6230. [CrossRef]
15. Yang, H.; Zheng, K.; Zhao, L.; Hanzo, L. Twin-Timescale Radio Resource Management for Ultra-Reliable and Low-Latency Vehicular Networks. *IEEE Trans. Veh. Technol.* **2020**, *69*, 1023–1036. [CrossRef]
16. Abbas, F.; Fan, P.; Khan, Z. A Novel Low-Latency V2V Resource Allocation Scheme Based on Cellular V2X Communications. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 2185–2197. [CrossRef]
17. Wang, B.; Zhang, R.; Chen, C.; Cheng, X.; Yang, L.; Jin, Y. Interference Hypergraph-Based 3D Matching Resource Allocation Protocol for NOMA-V2X Networks. *IEEE Access* **2019**, *7*, 90789–90800. [CrossRef]
18. Ashraf, M.I.; Bennis, M.; Perfecto, C.; Saad, W. Dynamic Proximity-Aware Resource Allocation in Vehicle-to-Vehicle (V2V) Communications. In Proceedings of the 2016 IEEE Globecom Workshops (GC Wkshps), Washington, DC, USA, 4–8 December 2016; pp. 1–6.

19. Abdallah, A.; Mansour, M.M.; Chehab, A. A Distance-Based Power Control Scheme for D2D Communications Using Stochastic Geometry. In Proceedings of the 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall), Toronto, ON, Canada, 24–27 September 2017; pp. 1–6.

20. Lucas-Estan, M.C.; Gozalvez, J. Distance-Based Radio Resource Allocation for Device to Device Communications. In Proceedings of the 2017 IEEE 85th Vehicular Technology Conference (VTC Spring), Sydney, Australia, 4–7 June 2017; pp. 1–5.

21. 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects. V9.0.0, TR 36.814. 2010. Available online: http://www.3gpp.org/ftp/Specs/html-info/36814.htm (accessed on 28 September 2020).

22. Yokoyama, R.S.; Kimura, B.Y.L.; Villas, L.A.; Moreira, E.D.S. Measuring Distances with RSSI from Vehicular Short-Range Communications. In Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, Liverpool, UK, 26–28 October 2015; pp. 100–107.

23. 3GPP. Technical Specification Group Services and System Aspects; Proximity-Based Services (ProSe); Stage 2. TS 23.303, Version 13.3.0. 2016. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=840 (accessed on 28 September 2020).

24. Ye, H.; Li, G.Y.; Juang, B.F. Deep Reinforcement Learning Based Resource Allocation for V2V Communications. *IEEE Trans. Veh. Technol.* **2019**, *68*, 3163–3173. [CrossRef]

25. Nie, S.; Fan, Z.; Zhao, M.; Gu, X.; Zhang, L. Q-learning based power control algorithm for D2D communication. In Proceedings of the 2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Valencia, Spain, 4–8 September 2016; pp. 1–6.

26. Otterlo, M.V.; Wiering, M. *Reinforcement Learning: State-of-the-Art*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 12.

27. Gu, Y.; Saad, W.; Bennis, M.; Debbah, M.; Han, Z. Matching theory for future wireless networks: Fundamentals and applications. *IEEE Commun. Mag.* **2015**, *53*, 52–59. [CrossRef]

28. Wei, Q.; Sun, W.; Bai, B.; Wang, L.; Ström, E.G.; Song, M. Resource allocation for V2X communications: A local search based 3D matching approach. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017; pp. 1–6.

29. Khan, U.A.; Lee, S.S. Three-Dimensional Resource Allocation in D2D-based V2V Communication. *Electronics* **2019**, *8*, 962. [CrossRef]

30. Dutta, A.; Asaithambi, A. One-to-many bipartite matching based coalition formation for multi-robot task allocation. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 2181–2187.

31. Heineman, G.T.; Pollice, G.; Selkow, S. *Algorithms in a Nutshell*; Oreilly Media: Sebastopol, CA, USA, 2008; pp. 226–250.

32. Chen, S.; Hu, J.; Shi, Y.; Zhao, L.; Li, W. A Vision of C-V2X: Technologies, Field Testing, and Challenges With Chinese Development. *IEEE Internet Things J.* **2020**, *7*, 3872–3881. [CrossRef]

33. Ye, H.; Liang, L.; Li, G.Y.; Kim, J.; Lu, L.; Wu, M. Machine Learning for Vehicular Networks: Recent Advances and Application Examples. *IEEE Veh. Technol. Mag.* **2018**, *13*, 94–101. [CrossRef]

34. Li, Z.; Guo, C. Multi-Agent Deep Reinforcement Learning Based Spectrum Allocation for D2D Underlay Communications. *IEEE Trans. Veh. Technol.* **2020**, *69*, 1828–1840. [CrossRef]

35. Huang, W.; Chen, M.; Yang, Z.; Huang, N.; Pei, L. Resource Allocation for Relay-Assisted D2D Communications with Network Coding. In Proceedings of the 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 27–30 August 2018; pp. 1–5.

36. Wang, J.; Rouil, R.A.; Cintron, F.J. Distributed Resource Allocation Schemes for Out-of-Coverage D2D Communications. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019; pp. 1–7.

37. Xiao, H.; Zhu, D.; Chronopoulos, A.T. Power Allocation With Energy Efficiency Optimization in Cellular D2D-Based V2X Communication Network. *IEEE Trans. Intell. Transp. Syst.* **2019**, 1–11. [CrossRef]

38. Deng, L.; Wu, G.; Fu, J.; Zhang, Y.; Yang, Y. Joint Resource Allocation and Trajectory Control for UAV-Enabled Vehicular Communications. *IEEE Access* **2019**, *7*, 132806–132815. [CrossRef]

39. Yang, Y.; Feng, L.; Zhang, C.; Ou, Q.; Li, W. Resource allocation for virtual reality content sharing based on 5G D2D multicast communication. *EURASIP J. Wirel. Commun. Netw.* **2020**, *2020*, 112. [CrossRef]

40. Abbas, F.; Fan, P. A Hybrid Low-Latency D2D Resource Allocation Scheme Based on Cellular V2X Networks. In Proceedings of the 2018 IEEE International Conference on Communications Workshops (ICC Workshops), Kansas City, MO, USA, 20–24 May 2018; pp. 1–6.