Congestion Control in V2V Safety Communication: Problem, Analysis, Approaches

Xiaofeng Liu † and Arunita Jaekel *†

School of Computer Science, University of Windsor, Windsor, ON N9B 3P4, Canada; liu1ew@uwindsor.ca
* Correspondence: arunita@uwindsor.ca; Tel.: +1-519-253-3000 (ext. 2996)
† These authors contributed equally to this work.

Received: 26 January 2019; Accepted: 27 March 2019; Published: 13 May 2019

Abstract: The emergence of Vehicular Ad Hoc Networks (VANETs) is expected to be an important step toward achieving safety and efficiency in intelligent transportation systems (ITS). One important requirement of safety applications is that vehicles are able to communicate with neighboring vehicles, with very low latency and packet loss. The high mobility, unreliable channel quality and high message rates make this a challenging problem for VANETs. There have been significant research activities in recent years in the development of congestion control algorithms that ensure reliable delivery of safety messages in vehicle-to-vehicle (V2V) communication. In this paper, we present a comprehensive survey of congestion control approaches for VANET. We identify the relevant parameters and performance metrics that can be used to evaluate these approaches and analyze each approach based on a number of factors such as the type of traffic, whether it is proactive or reactive, and the mechanism for controlling congestion. We conclude this paper with some additional considerations for designing V2V communication protocols and interesting and open research problems and directions for future work.

Keywords: congestion control; V2V; VANET

1. Introduction

Without doubt, vehicular safety is an important problem in modern society. Each year, nearly 1.25 million people die in road crashes, on average 3287 deaths a day [1]. Over the years, various technologies have emerged in the vehicular traffic sector to help reduce accidents. Since 2001, when Vehicle Ad Hoc Network (VANET) was first introduced in [2] as a term, it has been widely perceived by government, car manufacturing industries, and academia as a promising concept for future realization of Intelligent Transportation System (ITS) thereby achieving safety and efficiency in our nearly overcrowded motorways [3]. To date, a number of standards have been implemented to accommodate the Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications for safety-related applications. Short to medium range wireless communication known as Dedicated Short Range Communication (DSRC) [4] has been proposed for such vehicular communication. In the United States, the Federal Communication Commission (FCC) allocated 75 MHz spectrum for DSRC. Collectively the IEEE 1609 family, IEEE 802.11p and the Society of Automotive Engineers (SAE) J2735 form the key parts of the currently proposed Wireless Access in Vehicular Environment (WAVE) protocol stack [5], which provides the protocols from physical level to application level for vehicular communication. Although several safety-related standards (e.g., channel spectrum, safety message format) have been defined, there remain many aspects of DSRC performance requirements, e.g., defining sending rate, transmit power control, adaptive message rate control that need to be investigated thoroughly. There are significant DSRC-related social and technical challenges that have to be dealt with before large-scale deployment.
A key requirement of V2V communication is the reliable delivery of safety messages. These messages are typically broadcast to neighboring vehicles using DSRC/WAVE technology, based on CSMA/CA in the media access layer. Due to the multi-access wireless channel, many factors can cause the delay or failure of the dissemination of safety messages. The limited channel bandwidth available to transmit the safety messages means that the shared radio channels can become easily congested as vehicle density increases. In [6], experiments have demonstrated that channel congestion can occur even in relatively simple traffic scenarios. Channel congestion is a critical factor that leads to delayed or failed messages delivery. With higher vehicle density, it is not clear if the channel capacity will be sufficient to support the data load generated by both beacons and event-driven safety messages. Therefore, the development of effective congestion control strategies for V2V communication is of utmost importance and has been an area of intense research interest in recent years.

Congestion control techniques for traditional wired and wireless communication has been well investigated in the literature [7–9]. For example, TCP [10] handles congestion with end-to-end control by adjusting the data rate on the source node, when detecting a change in the acknowledgments (ACKs) received from the destination node. In Mobile Ad hoc Networks (MANETs), researchers focus on routing and backward compatibility in multi-hop networks [11]. However, the unique challenges and requirements of VANET mean that these existing techniques are typically not suitable to be applied directly for VANET communication. For example, since safety messages are broadcast to all neighboring vehicles, collecting the ACKs from all the receiving vehicles becomes infeasible, as these ACK messages will consume bandwidth and further exacerbate the channel congestion. The single-hop based safety messages broadcasting in VANET also has completely different requirements compared to the multi-hop communication in MANETs. Furthermore, congestion control in VANET becomes even more challenging due to the high node mobility and channel fading constraints.

This paper will give a comprehensive overview of the requirements, performance metrics and current approaches for designing a congestion/awareness control protocol for V2V communication. The remainder of the paper is organized as follows. In Section 2, we introduce the relevant background information on V2V safety messages communication using DSRC/WAVE and the effects of higher channel load on such communication. In Section 3 we identify some of the relevant parameters and performance metrics for evaluating congestion control algorithms in VANET. In Section 4, we analyze and discuss the different congestion control protocols for VANET available in the literature and present our conclusions along with some open research problems and directions for future work in Section 5.

2. Background

Vehicular safety means not only strengthening the body of the vehicle to withstand crashes, but also improving a vehicle’s awareness of surrounding vehicles and vice versa. Currently available camera-based safety techniques work within a limited distance, limited scope and require good visibility. Inter-vehicle wireless communication can help to overcome many of these limitations. The prompt delivery of safety messages, e.g., velocity, position, can be very helpful collision avoidance and other emergency situations. Since the 1920s, radio communications between vehicles have been used to improve safety [12]. Today, a wide variety of safety applications can be developed based on ITS technologies. One popular method to classify these applications is Time-To-Collision (TTC) [13], which measures the traffic conflict probability as the time required for two vehicles to collide if they continue at their present speed and on the same path [14]. In [15], authors argue that one of the congestion control principles should be deference to safety applications’ requirements in different driving contexts. This means the applications will disseminate safety messages in different ways based on the vehicle’s driving context.
In V2V communication safety-related messages are disseminated by broadcasting and consist of two main messages by two main types—(i) periodic messages and (ii) event-driven messages. The Basic Safety Messages (BSMs) defined by SAE J2735 [16] in the U.S. and Cooperative Awareness Messages (CAMs) specified by ETSI TC ITS [17] in Europe, also called as beacons, are transmitted periodically to announce the vehicle status. The event-driven messages are transmitted at the detection of a traffic event or road hazard, e.g., Decentralized Environmental Notification Message (DENM) developed by ETSI TC ITS [18]. In this paper, we primarily focus on BSMs for periodic broadcasts. BSMs consist of the following two parts:

- Part I includes basic vehicle state which is mandatory in each BSM, e.g., time, position, motion, vehicle size.
- Part II includes vehicle safety extension, which is optional for V2V safety applications, e.g., event flags, path history.

The average size of a BSM is from 320 to 350 Bytes, but if security-related overheads are considered, it could increase to around 800 Bytes. The default transmit rate is 10 times per second.

2.1. DSRC Spectrum for V2V Communication

FCC has allocated a 75 MHz spectrum in the 5.850–5.925 GHz (5.9 GHz band) divided into seven 10 MHz channels, with a 5 MHz guard band at the low end, for DSRC communication. In this spectrum, Channel 172 is designated for safety [19] and vehicle safety messages are expected to be exchanged on this channel. In [4], the authors propose that pairs of 10 MHz channels (channel 172 and 174) can also be combined into a 20 MHZ channel, to achieve increased capacity with shorter transmitting time, thus reducing the collision probability for the frame transmission. But a 20 MHz channel has been reported to lead to higher inter-symbol interference (ISI) and inter-carrier interference (ICI) that can significantly affect the successful reception of packets [20]. So far, the testing of DSRC and other related research mainly focus on 10 MHz channels.

Based on orthogonal frequency division multiplexing (OFDM) and different modulation techniques, DSRC provides 8 different data transfer rates (3, 4.5, 6, 9, 12, 18, 24, 27 Mbps). The higher data rate, the higher is the signal-to-interference-plus-noise ratio (SINR) threshold (dB) required for packet reception. There has been only limited research work reported on how to optimally select the data rate for safety communications. In [21], the authors proposed a 6 Mbps data rate after analyzing the work in [22], which concluded the lowest probability of collision is with 24 Mbps data rate in the channel under low overall channel load and without considering channel fading. A 6 Mbps data rate has been widely adopted in many congestion control simulations in VANET [4,11,15,23], as well as in some standardization activities [24]. Other data rates have also been considered, e.g., 3 Mbps is used in [25–27] for its lowest SINR requirement.

Figure 1 shows the current DSRC channel band and how BSM is transmitted on channel 172. Each vehicle will run some safety applications, e.g., emergency electronic brake lights (EEBL), left turn assist (LTA). These safety applications will need to disseminate the vehicle’s BSMs periodically and acquire the other vehicles’ BSMs to complete their assigned tasks. All these BSMs are transmitted on channel 172 plus some event-driven messages as well.
2.2. Channel Congestion in DSRC

The 802.11p wireless protocol is currently the only protocol for low-latency V2V safety communication in VANET. All the safety messages are transmitted with 802.11p, which uses carrier sensing multiple access with collision avoidance (CSMA/CA) to reduce collisions and provide fair access to the channel. Due to the low latency requirements, the safety messages (both beacons and event-driven messages) are always transmitted using one-hop broadcasting communications. CSMA/CA is a contention-based random access protocol that can cause simultaneous transmissions, i.e., packet collisions, which significantly lower communication reliability [28,29]. For broadcast messages, re-transmission is usually not an option because receiving nodes do not send any acknowledgement (ACK) to avoid ACK flooding. Furthermore, the important packet-collision avoidance mechanisms like the request-to-send/clear-to-send (RTS/CTS) handshake are disabled in 802.11p MAC layer, resulting in hidden terminal problems.

The requirement to frequently broadcast BSMs can quickly lead to very high channel loads for the single 10 MHz channel that has been allocated for safety messages. For example, let us consider the following scenario: a 4-lane bidirectional highway with a vehicle density of 25 vehicle/km/lane, BSM transmissions every 100ms from each vehicle, where each BSM is 500 Bytes. The amount of data generated per second for mutual awareness could easily exceed the available data rate of 3 Mbps. Similar calculation are also given in [20,27]. The work in [14] lists three important consequences of high channel load for V2V communication:

- successful packet receptions decrease and transmission delay increases with increasing message generation rates;
- effective transmission ranges decrease under high channel load conditions;
- communication range of a transmitter decreases under interference (lower SINR) due to hidden stations or simultaneous sending.

The authors in [14] also identify the following reasons for packet loss under high channel load, which need to be considered in the design of congestion control protocols:
transmission collision due to simultaneous sending;
lower SINR caused by the interference from single/multiple hidden stations;
packet dropped locally due to the failure of medium access;
lower SINR caused by the nearby transmitting stations.

Both the MAC transmission delay and the number of packet collisions have been shown to grow rapidly when the channel load increases over a threshold of the channel access capacity, e.g., above a channel busy ratio (CBR) of 40% [14]. The MAC transmission delay will cause safety messages to arrive late, and a high number of packet collisions lead to a lower reception probability and hence an effective radio range reduction. When this occurs, it is referred to as channel congestion. In order to ensure reliable and timely delivery of safety messages, effective countermeasures need to be taken. In the following sections, we will discuss and compare some of the important congestion control strategies and their objectives for V2V communication.

3. Performance Metrics and Simulation Parameters

A number of different strategies for controlling congestion, in the allocated wireless channel for V2V communication, have been proposed in recent years. However, there is a wide divergence in terms of the objectives, performance criteria, simulation scenarios and other characteristics associated with these approaches, as reported in the literature. In order to have a meaningful comparison of the various approaches, we need to first identify the relevant metrics used to evaluate them, as well as the specific scenarios and channel characteristics, for which each approach was evaluated. In this section, we first consider the metrics that have been used to evaluate the performance of V2V congestion control algorithms. We then discuss some of the different scenarios, e.g., traffic patterns, channel characteristics that affect the evaluation of these algorithms through simulations.

3.1. Performance Metrics

There are many different performance metrics for evaluating congestion control algorithms. Some of the most commonly used ones for V2V communication are discussed below.

3.1.1. Channel Busy Ratio (CBR)

CBR (Channel Busy Ratio) or CBT (Channel Busy Time) is defined as the ratio between the time the channel is sensed as busy and the total observation time (e.g., 100 ms). It is a measure for the channel load perceived by a vehicle, and depends on the number of vehicles in its transmission range and their individual message generation rates. Many congestion control algorithms, such as periodically updated load sensitive adaptive rate control (PULSAR) [15], random transmit power control (RTPC) [30], distributed fair transmit power adjustment (D-FPAV) [25], distributed network utility maximization (D-NUM) [31], packet-count based decentralized data-rate congestion control algorithm (PDR-DCC) [32] use CBR/CBT as one of the metrics to analyze the communication benefit of the algorithms. CBR has been shown to be a suitable metric to increase packet delivery performance [33]. Simulations show that high message rates cannot be supported for all vehicles in case of high vehicle densities or large packets [14]. In [34], the authors investigate the relationships between the CBR locally perceived by vehicles, i.e., the one-hop CBR, and the two-hops CBR. CBR can also be an input to congestion control mechanism. CBR has been referred to in other documents as the Channel Busy Function (CBF) [11] and the Channel Busy Percentage (CBP) in (V2V Safety Communication Scalability based on the SAE J2945/1 Standard). For analytical research about CBR, paper [35] gives a mathematical model in terms of CBR for channel load at a single position, load contribution from a single transmitter and channel load distribution on road respectively. Paper [36] introduces an analytical model based on a straightforward Markov reward chain to obtain transient measurements of the idle time of the link between two VANET nodes.
3.1.2. Packet Loss Rate (PLR)

Packet loss may be caused by different reasons, e.g., collisions, fading or control channel interval (CCHI) expiry time. PLR measures the percentage of packets lost as a result of collisions and/or fading. The percentage of dropped packets due to the CCHI expiry time is the ratio of dropped packets to the total packets generated by the application layer [37]. There are other similar metrics like PLR used in literature. Packet error rate is used in [38]. Packet collision rate, which is defined as the number of packet collisions normalized in time and space, is used in [30]. Reception probability is used in [21,25]. The inverse of PLR is packet delivery ratio (PDR), the analytical expression for PDR is given in [39].

3.1.3. Inter-Packet Delay (IPD)

The inter-packet delay (IPD) is the average delay between successive packets received from surrounding vehicles. It is also referred to as packet inter-reception time (IRT) in [40], which is the time between two sequential packet receptions for a particular sender-receiver pair, and this metric is used to assess the temporal aspect of awareness in each transmission power control algorithm. A higher IRT value means older information is being received from the neighboring vehicles.

Both PLR and IPD are important metrics for VANET, as they both give a measure of the level of awareness of surrounding vehicles. As such, both metrics have been widely used in performance evaluation of V2V communication protocols [21,25,30–40]. Although both PLR and IPD provide information on vehicular awareness, they give different perspectives on awareness. PLR is used to estimate what percentage of packets are reaching a vehicle, which gives a coarse-grained measure of vehicular awareness. On the other hand, IPD can provide a more in-depth view, e.g., difference in awareness for cases where the PLR may be the same. This is illustrated using the two scenarios in [41], where 50% of the beacons transmitted by vehicle A with a 10Hz frequency in an interval of 10 s are received at vehicle B, i.e., PLR = 50% in both cases.

- Scenario 1: Alternate beacons are received correctly. In this case IPD = 200 ms, meaning information for vehicle A is outdated for at most 200 ms;
- Scenario 2: Beacons are received in batches—25 beacons are received in the first two seconds of the interval, then no beacon is received for 5 s, and the remaining 25 beacons are received in the last three seconds of the interval. In this case, there is a situation awareness blackout of at least 5 s. Considering that a vehicle position can change by over 100 m in 5 s at highway speeds, it is clear that situation awareness is severely impaired in Scenario 2, resulting in possibly undetected dangerous situations.

3.1.4. Additional Metrics

Besides the aforementioned metrics, some other metrics have also been used to evaluate the performance of congestion control techniques. In [27] the probability of successful reception of beacon message is used with respect to the distance to assess the algorithm’s effectiveness and the average channel access time (CAT) for claiming fairness. Update delay is defined in [30] as the elapsed time between two consecutive CAMs/BSMs successfully received from the same transmitter, which aims to assess the awareness of vehicle from a communication perspective. In [26] a 95% Euclidean cut-off error (in meters) is defined and a 95% Tracking Error in [23] is used to represent the tracking accuracy of the communication policy, which possesses a statistical sense similar to a confidence interval (CI). The information dissemination rate (IDR), i.e., the number of packets received successfully by a node’s neighbours, is used in [33] as a metric for application performance. The authors conclude that CBR is a suitable feedback metric for maximizing the number of IDR, with a recommended CBR target value of 0.65.
3.2. Simulation Parameters

Due to the high cost and safety factors, simulation is an effective and popular way to evaluate congestion control approaches. In this section, we discuss some the important simulation tools and parameters that can affect the simulation outcomes.

3.2.1. Traffic Scenarios

Channel congestion can happen both on the highway or in urban areas, when vehicle density reaches a certain threshold. According to [42], a typical density for a 3-lane per direction on the highway could be 25 vehicles/Km/lane, which will result in a 300 vehicles within a 1000 m communication range. Authors in [14] claim that more challenging scenarios may be found on highways, because a simple velocity-based rate adaptation would solve the problem in urban area. Under normal low-traffic conditions, vehicles float in free flow. Due to perturbations at junctions or on-ramps, the state changes to synchronized flow, where vehicles on all lanes in the same direction have nearly the same speed. Due to further slight distortions, severe road traffic congestion occurs, also referred to as wide moving jams. These transitions of traffic flow are most relevant for decentralized congestion control, as the traffic density is high while at the same time the velocity is also considerably high. Traffic perturbations may cause emergency braking and accidents, which should be mitigated with the aid of inter-vehicle communication. Most of the congestion control approaches focus on the highway scenarios. But still some researches consider the urban scenarios [34,43]. The traffic scenarios are either deterministic, i.e., the vehicles are equally spaced with distance $d$ m, or randomly positioned according to a Poisson distribution of average density $\rho$ vehicles/m.

3.2.2. Channel Fading Model

The carrier signals are always degraded during transmission, which can limit both the carrier sense range and transmission range, and also cause hidden terminal problem. Appropriate channel fading models should be used to simulate the realistic environment, e.g., the deterministic Two-Ray Ground model, the probabilistic Log-Normal Shadowing model, Rayleigh model and Nakagami model. Among them, Nakagami model [44] is the most used one. Rayleigh model is a special case of Nakagami model, where the Nakagami fading parameter $m = 1$ and models a rough Non-Line-Of-Sight scenario. For fading parameter $m > 1$, Nakagami models an increased Line-Of-Sight scenario. In [45], the authors give an empirical conclusion: the lower the $m$ is, the longer is the transmission distance used in the simulation. For example, when $m = 1$, the model is suitable for the transmission distances higher than 150 m. Since longer distance causes more fading, in [27] the authors refer to Nakagami $m = 1$ as severe fading conditions, to Nakagami $m = 3$ as medium fading conditions, and to Nakagami $m = 5$ as low fading conditions. We note that the suitable configuration values for channel fading can be highly dependent on different simulation parameters such as vehicle density, distance, presence of obstacles etc. and different papers have used different values.

3.2.3. Simulation Tools

There are many network simulators that can be utilized to help researchers understand the behavior and performance of the networks and their protocols. Here we list three of the most popular open source network simulators currently being used: ns-2, ns-3, OMNet++.

1. ns-2: Network Simulator-2 (ns-2) [46] is an open source, discrete event network simulator for both wired and wireless networks. In ns-2, arbitrary network topologies can be defined that are composed of routers, links and shared media [47]. The physical activities of the network are processed and queued in the form of events, in a scheduled order. These events are then processed as per the scheduled time, which increases along with the processing of events. However, the simulation is not real time, it is considered virtual [48]. ns-2 was extended by [49] with: (a) node mobility, (b) a realistic physical layer with a radio propagation model, (c) radio network
Electronics 2019, 8, 540

interfaces, and (d) the IEEE 802.11 Medium Access Control (MAC) protocol using the distributed coordination function (DCF). After revised by [50], the resulting PHY is a full featured generic module capable of supporting any single channel frame based communications. The key features include cumulative signal to interference plus noise ratio (SINR) computation, preamble and physical layer convergence procedure (PLCP) header processing and capture, and frame body capture. The MAC now accurately models the basic IEEE 802.11 carrier sense multiple access with collision avoidance (CSMA/CA) mechanism, as required for credible simulation studies.

2. ns-3: The ns-3 project [51] started from mid 2006 and is still under development. The latest release is ns-3.29. It is treated as a replacement instead of upgrading of ns-2. It supports parallel simulation and emulation using sockets. ns-3 provides a realistic environment and its source code is well organized compared to ns-2 [52]. In ns-3, vehicle mobility and network communication are integrated through events. User-created event handlers can send network messages or alter vehicle mobility each time a network message is received and each time vehicle mobility is updated by the model. The authors in [53] implemented a straight highway model in ns-3 that manages vehicle mobility, while allowing for various user customizations. The revised ns-3 has two main classes: Highway and Vehicle. Vehicles are fully-functional ns-3 nodes that contained additional information regarding their current acceleration, velocity, and position. The Highway class uses Model and LaneChange objects attached to Vehicles to move vehicles based on IDM (Intelligent Driver Model [54]) and MOBIL lane change model (Minimizing Overall Braking Induced by Lane Changes [55]). In addition, Highway used ns-3 callbacks to enable simulation developers to take control of Vehicles based on network messages, overriding, if need be, the standard controls used in Highways. More improvements for ns-3 can be found in [56].

3. OMNet++: Unlike ns-2 and ns-3, OMNet++ supports more than network simulation. It can also be used for multiprocessors modelling, distributed hardware systems, etc. As a general discrete event, component-based open architecture simulation framework, OMNet++ can combine with SUMO (Simulation of Urban MObility) [57], a traffic simulator and Veins (Vehicles In Network Simulation) [58], which couples the network and traffic simulator to simulate the vehicle communication. With Veins each simulation is performed by executing two simulators in parallel: OMNeT++ (for network simulation) and SUMO (for road traffic simulation). Both simulators are connected via a TCP socket. The protocol for this communication has been standardized as the Traffic Control Interface (TraCI). This allows bidirectionally-coupled simulation of road traffic and network traffic. Movement of vehicles in the road traffic simulator SUMO is reflected as movement of nodes in an OMNeT++ simulation. Nodes can then interact with the running road traffic simulation.

Paper [59] compared the performance of these simulators with the ad hoc on demand distance vector (AODV) routing algorithm and concluded that ns-3 has the fastest performance in terms of computation time. Both ns-2 and ns-3 fully utilize the CPU, while OMNet++ can provide more flexible functions. All three simulators can satisfy most of the vehicle communication simulation needs. They all can set up the common simulation parameters, e.g., noise floor, carrier sense threshold, packet reception SINR, payload size, data rate, Tx rate and power. More details about the survey and comparison of the simulators can be found in [60–62].

Although simulation-based VANET research has already produced important results, it has to be considered that wireless simulators are not entirely accurate representations of the real life situation, e.g., simplified propagation and interference models, inaccurate application data patterns [38]. A validation of simulation results using implementations on actual hardware (and vice versa) is necessary in VANET research. There are several existing testbeds for studying various VANET based protocols and applications. DRIVE [63] (Demonstrator for Intelligent Vehicular Environments) is an experimental implementation featuring both advanced vehicular services and vehicular networking (i.e., intra-vehicle, V2V, V2I communications). It provides a modular, re-configurable architecture based on carPC, wireless sensors and rooftop antennae and uses wireless and mobile technologies,
e.g., WiFi, UMTS and Bluetooth. In [38], the authors discussed how a wireless testbed could be suitable for VANET research by using only software adjustments, e.g., an approximation of the IEEE 802.11p standard using IEEE 802.11a hardware, the emulation of mobility based on link impairment and the use of low transmit power together with manual topology configuration. With these techniques, IBBT w-i-lab.t can support high vehicle density, multi-hop and emulation of mobility in a three-floor building with some extensive facilities. In [64], SD-VANET (Software Defined Architecture for VANET) is proposed to solve the high complexity in terms of the control and management of the network infrastructure due to the coexistence of multiple different access technologies (e.g., already deployed Wi-Fi networks, DSRC and cellular network). The authors use software defined networking to enhance VANET with Wi-Fi access capability as both roadside units and vehicles. Additionally, a Wireless Access Management (WAM) protocol is proposed to provide wireless host management and basic flow admission. SD-VANET is proposed to solve the high complexity in terms of the control and management of the network infrastructure due to the coexistence of multiple different access technologies (e.g., already deployed Wi-Fi networks, DSRC and cellular network). The authors use software defined networking to enhance VANET with Wi-Fi access capability as both roadside units and vehicles. Additionally, a Wireless Access Management (WAM) protocol is proposed to provide wireless host management and basic flow admission with respect to the available bandwidth to validate the capability of the offered architecture. MDX VANET [65] is a prototype VANET network built on the Middlesex University Hendon Campus, London. It is used to investigate better propagation models, road-critical safety applications as well as algorithms for traffic management. Some researchers use testbeds to measure the beaconing performance for congestion control. In [41], the authors use two IEEE 802.11p compliant devices, namely the LinkBird-MX v3 units produced by NEC, to measure two important metrics for congestion control: beacon (packet) delivery rate (PDR), and the packet inter-reception (PIR) time. Authors in [66] implemented the congestion control (CC) protocol defined in (SAE) J2945/1 standard [67] using a congestion generation testbed of Remote Vehicles (RVs). The test uses a DSRC-enabled Host Vehicle (HV) to wirelessly transmit BSMs, a Ground Truth Equipment (GTE) mounted on the HV to accurately capture its position and a DSRC sniffer to remotely capture the BSMs received over-the-air (OTA). Furthermore, The authors used a Congestion Generation Tool (CGT) that can emulate up to 160 Remote Vehicles (RVs) in the setup (which could be expanded) and up to 80% Channel Busy Percentage (CBP) (i.e., a measure of channel occupancy) by transmitting a mix of BSM and WSM (WAVE Short Message) packets. For congestion control research with simulators, different comprehensive scenarios can be researched in a low-cost way, experiments can be repeatable and parameters (e.g., vehicle movement patterns, vehicle density) can be easily controlled. But simulation can’t represent the real-life performance. Several studies [68–70] have discussed this problem in detail. Regarding signal propagation and interference, several aspects are often simplified in wireless network simulators. Besides networking aspects, other elements such as traffic patterns and end-user models can also be incomplete in simulations. The advantages of experimental implementations (i.e., testbeds) over the above-mentioned approaches are the gain of hands-on experience in close-to-real or real performance and behavioral issues, as well as favoring intuition shape practice. The drawback of Field Operational Tests (FOTs) is the practical limitations, e.g., difficulty of implementing repeatable experiments, high cost, specific traffic model requirements. More valid evaluation methods still need to be developed.

4. Analysis and Approaches

Many approaches have been proposed to tackle the congestion problem in VANET, using a range of different control mechanisms, parameters, metrics and problem definitions. In spite of this diversity of approaches, some common topics are generally considered when implementing congestion control techniques, e.g., the safety message type, transmission parameters, fairness. In this section, we will discuss some topics that congestion control measures need to consider and give an overview of the related approaches. We will also give a comparison of the various approaches in terms of the criteria discussed in this paper.

4.1. Beacons (BSMs/CAMs) vs. Event-driven Safety Messages

As discussed in Section 2, beacons are used to periodically broadcast vehicle status information, while event-driven messages typically disseminate more urgent information. Therefore, ideally, event-driven safety messages should be delivered with shorter delay and lower collision probability,
over a larger area. So, a higher transmission power is more suitable for event-driven safety messages. Beacons are normally more relevant to nearby vehicles, typically within a 150 m radius [26]. In [25], the authors note that there must be a strategy to balance event-driven messages and beacons and conclude that the amount of load resulting from beacons should be limited, to leave some available bandwidth for event-driven messages. Since beacons are transmitted by every vehicle on the road periodically, most congestion control approaches only consider the beacons as the messages transmitting on the channel [71]. In [72] the authors propose a V2V communication protocol specially for emergency warning dissemination, which is a common example of event-driven safety messages. Obviously both periodic beacons and event-driven messages should be considered in congestion control approaches. But due to the difference in priority and dissemination mode (one-hop or multi-hop), it is difficult to integrate the dissemination for both messages in one algorithm.

4.2. Reactive vs. Proactive

Congestion control mechanisms can be categorized as proactive (or adaptive in [14]) (i.e., feed forward) or reactive (i.e., feedback). Proactive approaches try to prevent congestion before it occurs; whereas reactive approaches try to handle congestion after it is detected. In [20], the authors use traditional control theory to depict the congestion control mechanism, as shown in Figure 2. The objective is usually determined based on the requirements of the safety application, e.g., a threshold channel load value. The controller is open- or closed-loop depending on whether it accepts the feedback from the result, which can be from the network or from another individual node. The controller will adjust the transmission parameters based on the traffic situation or scenario and the objectives.

![Figure 2. Congestion control mechanism from a control theory point of view [20].](image)

Proactive approaches usually set up a target value for a specific of channel load metric, e.g., CBR and adapt the transmission parameters to drive the metric toward the target value. Then the algorithm can ensure the channel load will not exceed the target value. Some important proactive approaches include D-FPAV [27], linear message rate control algorithm (LIMERIC) [11], error model based adaptive rate control (EMBARC) [23], proactive ad hoc on-demand distance vector routing (Pro-AODV) [73], additive increase and multiplicative decrease based decentralized congestion control (AIMD-Based DCC) [74], D-NUM [31] and RTPC [30].

An alternative approach is to react to congestion after it has actually occurred. Such a reactive approach is used in PULSAR [15], Adaptive transmission control [75], AC3 [76], etc.

Reactive strategy reacts directly to the measured channel load. The transmit behavior of can be pre-determined for any given channel load value. The function between channel load (e.g., CBR) and transmission parameters (e.g., Tx rate) determines the behavior and convergence of the control loop, when the number of nodes in range or their message rate changes. This freedom makes them very flexible, but analytically difficult to handle and understand. Additionally, since these algorithms react after congestion occurred, some BSMs may already get lost, which potentially can cause danger.
Proactive strategy can ensure that the channel load does not exceed a given threshold (e.g., CBR target value), but it needs to calculate the real-time information of the channel load and share with its neighbours. If the shared information is too much and too frequent, the algorithm itself can contribute to channel congestion.

4.3. Selection of Control Parameters

Based on the channel capacity analysis theory in [77], when vehicle density increases, each node (i.e., vehicle) needs to either decrease the transmission rate (Tx rate) or reduce the transmission power (Tx power), so that the limited channel load can be shared by all the nodes. So, transmission power and rate are the two basic parameters that can be controlled to reduce channel congestion. Existing VANET congestion control techniques either adjust the rate of BSM packet transmission, or the transmit power or a combination of both to achieve acceptable levels of congestion.

4.3.1. Rate Control

The default Tx rate for BSMs/CAMs is 10 Hz, i.e., ten messages per second. At this rate, the channel load will exceed channel capacity when the vehicle density is high, leading to packet loss and delay. Rate control based approaches adapt the Tx rate to reduce the number of safety messages transmitted by each node per unit time, if necessary. The safety applications based BSMs/CAMs typically require a high probability information reception in a so called target range, i.e., 100 to 300 m [78]. In this context the inter-packet delay (IPD) is a common metric for application performance. Safety applications usually will ask for a minimum and a maximum value of IPD within the target range. Using rate control typically results in increased IPD, and may lead to unacceptable IPD values, if congestion is high. A number of recent congestion control algorithms use rate control, including PULSAR [15], which is based on the observation that the optimal choice of Tx rate depends on node density and the optimal Tx range does not. A similar binary message rate control approach is given in [79]. Other popular rate control approaches include LIMERIC [11], EMBARC [23] and a tracking error based approach presented in [26].

4.3.2. Power Control

Tx power determines how far a message can be sensed and be received successfully, i.e., the carrier sense (CS) range and the Tx range (normally 500 m with Tx power 10.21 dBm) respectively. Normally, the carrier sense range is longer (e.g., double) than the Tx range, which may decrease further due to channel fading and the interference from neighbours or hidden terminals. The goal of Tx power control is to adapt the Tx range, based on some performance metrics. Reducing the transmit power instead of transmission rate means that nearby vehicles will still see frequent BSMs. However, distant vehicles will not see any packets at all from a vehicle that is far away. So, power control sacrifices awareness of distant vehicles, but can maintain full awareness of nearby vehicles. Rate control, on the other hand, reduces awareness for both near and far vehicles; but, all vehicles within the transmission range will receive the BSMs. One of the most cited algorithms is D-FPAV [25], which uses a fully distributed, asynchronous and localized approach. Other power control based algorithms include SPAV [80], OSC [81], SBAPC [82], RTPC [83], NOPC [84] and AC3 [76], etc.

4.3.3. Hybrid Control

The ultimate objective for congestion control is to improve the awareness of the vehicles. Decreasing the Tx power or rate may mitigate congestion, but lowers the quality of awareness. Conversely, increasing the awareness by transmitting with maximum Tx power and rate increases congestion. As indicated in [85], Tx power and rate are inversely correlated at constant target load on the wireless channel, i.e., decreasing the Tx power allows an increase in the Tx rate and vice versa. The approaches discussed so far can only set one single power/rate pair, which results in the transmission power/rate trade-off dilemma [30], as shown in Figure 3. To fulfill the awareness range requirement (OPrange) by
increasing the power, the rate has to be reduced, failing to achieve the required quality. On the other hand, to fulfill the quality requirement \((OP_{\text{quality}})\) by increasing the rate, the power has to be reduced, failing to achieve the required awareness range. To overcome this limitation, some joint power/rate congestion control approaches have been proposed. These approaches [30,75,86,87], attempt to jointly adjust rate and power to achieve the required level of awareness for all neighboring vehicles.

![Figure 3. The trade-off dilemma by using constant transmit powers [30].](image)

4.4. Overview of Different Congestion Control Approaches

In this section, we provide an overview of some of the leading congestion control algorithms in VANET, in terms of the parameters discussed in the previous sections. Table 1 lists the main characteristics of the leading congestion control approaches available in the literature.

**PULSAR:** PULSAR [15] (Periodically Updated Load Sensitive Adaptive Rate control) is a reactive approach, where each vehicle measures CBR at the end of a fixed time interval and compares the measured value against the target value. When the measured value is higher than the target value, i.e., congestion is observed, the vehicle will decrease its Tx rate accordingly. The PULSAR algorithm first fixes the Tx range based on the required target range and then adapts the Tx rate, with respect to CBR measurement. In PULSAR, the arrival of a new CBR measurement triggers a Tx rate adaption. The interval between CBR measurements is called channel monitoring and decision interval (CMDI) and is a fixed time for all nodes. When a new measurement arrives at the end of CMDI, PULSAR compares the measured value against the target value \((\text{CBR} = 0.6)\) and adapts the Tx rate using Additive Increase Multiplicative Decrease (AIMD).

**LIMERIC:** LIMERIC [11] (Linear Message Rate Control algorithm) uses linear feedback to adapt the message rate instead of the limit cycle behavior inherent in PULSAR. In LIMERIC vehicles in a given region, i.e., a single collision domain, where all the nodes measure the same channel load, sense and adapt their Tx rates to meet a specified CBR. LIMERIC is a distributed algorithm, executed independently by each station in a neighborhood. If \(r_j(t)\) is the message rate of the \(j\)th station at time \(t\), and \(\text{CBR}(t)\) is the CBR value measured at time \(t\), then at each update interval LIMERIC adapts the message rate according to:

\[
r_j(t + 1) = (1 - \alpha) r_j(t) + \beta (\text{CBR}_{\text{target}} - \text{CBR}(t))
\] (1)
In this equation, $\alpha$ and $\beta$ are algorithm parameters that influence stability and adaptation speed. The final Tx rate is highly dependent on the number of vehicles in the region.

**AIMD**: In [74], the authors use the TCP’s AIMD (Additive Increase and Multiplicative decrease) congestion control strategy, described in [88]. The AIMD rule is applied on the transmit rate $r$. Exceeding the channel load threshold $C_{\text{thresh}}$ serves as decrease criterion. As long as the perceived channel load is below the threshold, the transmit rate is increased by an increment $\alpha$, otherwise decreased with a factor $\beta$, where $0 < \beta < 1$. After the update function is evaluated, the resulting new transmit rate might be bounded by a predefined minimum and a maximum. If the CBR measurement interval is large or subject to averaging or smoothing, the CBR update might be slower than the transmit rate updates. Then a second decrease step is likely after a first decrease. This can be prevented by resetting the measured CBR.

**MD-DCC**: In [89], the authors propose a decentralized combined message-rate and data-rate congestion control (MD-DCC) scheme, which provides a fair and effective way of message-rate and data-rate allocation among vehicles to avoid congestion and satisfy application requirements. MD-DCC keeps the beacon frequency above the required minimum value by lowering the message-rate. However, the traffic density may reach a point where minimum frequency value is reached and traffic density still keeps increasing. At that point, MD-DCC increases the data-rate, resulting in more channel capacity. Adjusting data-rate for more channel bandwidth is a novel idea, but if the vehicles transmit message with different data-rates, the synchronization between the sender and receiver could be a big problem. Also higher data-rate needs higher SNR which is more difficult when vehicle density is getting higher.

**EMBARC**: EMBARC [23] (Error Model Based Adaptive Rate Control) extends the work of LIMERIC with the capability to preemptively schedule messages based on the vehicle’s movement, i.e., a kind of suspected tracking error technique. This approach can provide extra transmit opportunities for highly dynamic vehicles, hence reduce incidences of large tracking error.

**TRC for CASS**: Another rate control algorithm TRC for CASS (Transmission Rate Control for Cooperative Active Safety System) considering tracking error is presented in [26], which focuses on the dissemination of real-time tracking information of neighbouring vehicles while avoiding network congestion and failure based on multiple scalar linear time-invariant (LTI) dynamical systems. The authors first evaluate the decentralized information dissemination policies for tracking multiple dynamical systems and use a Err-Coll-Dep [90] policy to get a better tracking performance, then calculate the transmission probability for each vehicle every 50 ms based on the suspected tracking error. Based on the transmission probability, the vehicle stochastically generates a packet with its latest state information and places it on the MAC queue for transmission. This approach does not transmit the safety message as a periodic beacon, but in a stochastic way based on tracking information instead of CBR.

**D-FPAV**: Paper [27] proposes a method called D-FPAV (Distributed Fair transmit Power Adjustment for Vehicular Ad Hoc Networks) to control the load of beacons and another method EMDV (Emergency Message Dissemination for Vehicular environments) for dissemination of event-driven messages. D-FPAV achieves congestion control by setting the node transmission power based on the prediction of application-layer traffic and the observed number of vehicles in the surrounding area. As a proactive approach, D-FPAV uses a predefined Max Beaconing Load (MBL) and calculates the network-wide optimal transmission power to keep the channel load under this value. By periodically gathering the information about AL (Application-layer Load) for nodes within maximum carrier sense range, a node first uses FPAV [71] to calculate the maximum Tx power for all the nodes within its maximum carrier sense range, without violating the MBL condition. Then the nodes exchange this power by broadcasting to the neighbours within the maximum carrier sense range. After getting the same message from the nodes in its vicinity, each node computes the final power level, which is the minimum value among the local maximum Tx power levels in its vicinity. In FPAV, all the nodes increase their Tx power simultaneously, by the same amount, as long as the MBL condition is satisfied.
The authors of D-FPAV also consider strategies to deal with event-driven messages, i.e., emergency messages to effectively balance both low-priority traffic (beacons), and high-priority messages. D-FPAV shows a clear prioritization of emergency over beacon messages and reception probability for both message types are significantly higher at close distances, compared to no congestion control being applied. In D-FPAV, a key problem is how to acquire the knowledge about the presence of other nodes in the maximum carriers sense range. The carriers sense range is normally larger than the transmission range. This means it is not possible to get this information of the vehicles located outside the transmission range with one hop broadcasting. D-FPAV provides this information by piggybacking the aggregated status information in only 1 out of 10 beacon messages, in order to reduce the overhead; but, the overhead can still grow very high.

**SPAV:** Paper [80] proposes a distributed algorithm SPAV (Segment-based Power Adjustment for Vehicular environments) that adjusts the Tx power, based on averaged position and communication values for the neighbour information instead of using detailed neighbour information about each single node in [27]. To get the minimum ‘required knowledge’ of positional information on neighboring nodes that is necessary to guarantee a preconfigured maximum beaconing load in the network, the authors propose a distributed vehicle density estimation (DVDE) strategy to provide an approximation of the surrounding traffic situation to each vehicle. This approximation is then locally used by a separate power control SPAV. SPAV is executed locally at each vehicle in order to compute a common maximum transmission power for all vehicles in its environment. Therefore, SPAV derives a so called catchment area, the maximum area in which vehicles are allowed to transmit beacons with a maximum common power only affecting the center of the segment without violating segment’s MBL threshold, for each segment in its environment partition using the merged vehicle density histogram provided by DVDE.

**RTPC:** In [83], the authors propose to increase the Awareness Quality by randomly selecting the TX power for each CAM transmission and vehicle. Each vehicle controls its current TX power by using a certain probability distribution over a valid TX power interval. Using random TX power selection the radio propagation conditions are shifted with each transmission as well as the collision and interference areas. This mitigates the recurrence of collisions and interference, caused by the periodicity of CAM transmissions in combination with slow relative speeds. The random power selection mechanism provides statistical fairness as all vehicles use the same probability distribution and thus the same average TX power. Hence, local fairness is maintained as well.

**OSC:** In [81], the authors present a new algorithm to handle congestion by altering the value of transmission power. They propose a proactive approach, where transmit power of BSMs oscillate between two levels. A specified number of low-power packets are transmitted followed by a single high-power packet. This pattern is repeated until a change in network conditions (e.g., channel congestion) is detected. If necessary, a rate control algorithm such as LIMERIC can be combined with this method for further control.

**SBAPC:** SBAPC (Speed Based Adaptive Power Control) [82] presents a novel speed-based approach for controlling transmission power of BSMs. In SBAPC, each vehicle dynamically adjusts the transmit power of BSM packets, depending on its current speed. The main idea for SBAPC is that time to collision (TTC) with neighboring vehicles decreases as the speed of the vehicle increases. Therefore, BSMs from vehicle at higher speeds should be received by more distant vehicles, and such vehicles should use a higher transmit power. The results indicate a slight decrease in CBR compared to OSC. The IPD is similar to OSC for vehicles within 150 m of the ego vehicle. Between 150 m–260 m SBAPC has lower IPD compared to OSC; however, the IPD increases quickly beyond 260 m.

**NOPC:** NOPC (Non-cooperative beacon Power Control) [84] proposes a distributed congestion control algorithm based on non-cooperative game theory to find the equilibrium point, i.e., Nash Equilibrium (NE). The authors consider the power control is a submodular game, hence has a greatest and a least NE. In NOPC, each player, i.e., each vehicle, should pay higher price at higher congestion for increasing its power in terms of twice differentiable pay-off function which is a function
Electronics 2019, 8, 540 15 of 24

of CBR. Each vehicle measures its local CBR to update its BSM power based on gradient method and
does not require to exchange the algorithm related information with its neighbors. NOPC algorithm
also has a per vehicle parameter that every vehicle can adjust, without communicating it with other
vehicles to meet its application requirement. The simulation results show that NOPC is more efficient
in terms of bandwidth, it also has good stability and convergence to the unique NE with different
initial values of power for vehicles.

AC3: In [76], the authors propose Adaptive transmit power Cooperative Congestion Control
(AC3) to allow vehicles to select their transmit power autonomously with respect to their local
channel congestion. AC3 uses the notion of marginal contributions of vehicles towards congestion and
determines a fair power decrease for vehicles using a shapely value system model. This model requires
the vehicles with the highest marginal contributions, i.e., the highest level of transmit power used
by vehicles, to reduce the most transmit power and vice versa. Unlike [84], AC3 models cooperative
congestion control approach on the principles of cooperative game theory in [91], which allows us
to construct models of conflict and cooperation among players (i.e., vehicles) in order to maximize
individual payoffs.

Adaptive Transmission Control: In [75], the authors adjust the BSM Tx rate using AIMD (additive
increase multiplicative decrease), based on MAC blocking when congestion happens, e.g., the CBR
exceeds a defined threshold. The authors observe that any combination of Tx rate and Tx power results
in the same IDR vs CBR curve, which means CBR can be a suitable feedback for maximizing IDR.
They propose a methodology which first fixes Tx rate, based on vehicle tracking performance, and then
adapts Tx power based on CBR. One drawback is that the optimal choice of Tx power depends not
only on the current Tx rate, but also on vehicle density, which is hard to estimate correctly.

TPRC: In [87], the authors propose a TPRC (Transmit Power/Rate Control) strategy by first
analyzing the spatio-temporal characteristics of awareness and identifying the average or kth percentile
of IRT at a target distance \( d_t \), as a suitable choice to assess the safety benefit associated with a particular
Tx parameter (power or rate) setup. To find the optimal Tx parameter combinations, the authors first
set the Tx power to the optimal value for the current \( d_t \) of the individual vehicle, then adapt the Tx
rate with TRC (Transmit Rate Control). When \( d_t \) changes, TPC (Transmit Power Control) is applied.
The main advantage of the suggested joint optimization strategy is that it efficiently approximates the
identified optimal Tx parameter configurations for a given channel load target and \( d_t \), while reducing
complexity due to its independence of vehicle density.

RTPC + TRC: Considering awareness as in [83], the authors in [30] propose a new awareness/congestion
control strategy. The authors first use RTPC (Random Tx Power Control) featuring spatial awareness
control to reduce the channel load and then combine RTPC with transmit rate control (TRC) by
subsequently increasing the current Tx rate until the target load is reached. This approach is similar
with TPRC [87], but relies on randomized transmit power to provide a heterogeneous awareness
quality in space and to relax the strict transmission range to power mapping. RTPC is to find an
optimal Tx power distribution that in turn represents an optimal behavior of the awareness quality in
space instead of the optimal constant Tx power in TPRC.

CPRC: CPRC (Combined Power and Rate Control) [86] proposes a mechanism for the combination
of power and rate adjustment in a single algorithm loop, unlike the the two-phase joint power/rate
control approaches in [30,87,92]. CPRC’s policy is a cooperative behavior, where nodes that are not
directly involved in a potentially dangerous situation decrease their Tx power to allow nodes affected
by dangerous situations to increase their sending rate. It allows applications to send periodic messages
at higher rates, when required by contextual factors like speed and higher probability of collision,
without having the network load exceed the predefined reliability-based threshold.

4.5. Comparison of Different Approaches

Table 1 provides a comparison of leading congestion control approaches discussed in this section,
in terms of the parameters discussed in the previous sections.
Table 1. Classification of congestion control algorithms.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Rate/Power Control</th>
<th>Proactive/Reactive</th>
<th>Message Type</th>
<th>Metrics Used</th>
<th>Channel Fading Model</th>
<th>Simulator</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIMD [74]</td>
<td>Rate</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Reception rate, IPD, CBR</td>
<td>Not mentioned</td>
<td>own simulator</td>
</tr>
<tr>
<td>EMBARC [23]</td>
<td>Rate</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Tracking error</td>
<td>Not mentioned</td>
<td>ns-2</td>
</tr>
<tr>
<td>TRC for CASS [25]</td>
<td>Rate</td>
<td>Proactive</td>
<td>Beacon</td>
<td>95% cutoff error</td>
<td>Rayleigh</td>
<td>OPNET</td>
</tr>
<tr>
<td>SPAV [80]</td>
<td>Power</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Beaconing load</td>
<td>Nakagami, Two Ray Gound</td>
<td>ns-2</td>
</tr>
<tr>
<td>RTPC [83]</td>
<td>Power</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Packet collision rate, CBR, IPD, Update delay</td>
<td>Not mentioned</td>
<td>ns-3</td>
</tr>
<tr>
<td>OSC [81]</td>
<td>Power</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Beacon error rate</td>
<td>Not mentioned</td>
<td>OMNet+++, Veins</td>
</tr>
<tr>
<td>SBAPC [82]</td>
<td>Power</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Beacon error rate, CBR, IPD</td>
<td>Not mentioned</td>
<td>OMNet+++, Veins</td>
</tr>
<tr>
<td>NOPC [84]</td>
<td>Power</td>
<td>Proactive</td>
<td>Beacon</td>
<td>CBR</td>
<td>Nakagami</td>
<td>OMNet++, SUMO</td>
</tr>
<tr>
<td>AC3 [76]</td>
<td>Power</td>
<td>Reactive</td>
<td>Beacon</td>
<td>CBR</td>
<td>Not mentioned</td>
<td>OMNet+++, Veins</td>
</tr>
<tr>
<td>Hybrid transmission control [75]</td>
<td>Hybrid</td>
<td>Reactive</td>
<td>Beacon</td>
<td>95% cutoff error</td>
<td>Rayleigh</td>
<td>OPNET</td>
</tr>
<tr>
<td>TPRC [87]</td>
<td>Hybrid</td>
<td>Proactive</td>
<td>Beacon</td>
<td>IRT, CBR</td>
<td>Power-law</td>
<td>ns-2</td>
</tr>
<tr>
<td>RTPC+TRC [30]</td>
<td>Hybrid</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Packet collision rate, CBR, IPD, Update delay</td>
<td>Not mentioned</td>
<td>ns-3</td>
</tr>
<tr>
<td>CPRC [86]</td>
<td>Hybrid</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Channel Busy Time</td>
<td>Not mentioned</td>
<td>ns-2.31</td>
</tr>
<tr>
<td>MD-DCC [89]</td>
<td>Hybrid</td>
<td>Proactive</td>
<td>Beacon</td>
<td>Channel Busy Time</td>
<td>Not mentioned</td>
<td>ns-3, SUMO</td>
</tr>
</tbody>
</table>

Table 1 shows that there is a fair balance between rate control and power control algorithms, although there appears to be a growing consensus that a joint approach combining both will likely lead to better results in the future. The vast majority of the current approaches focus mainly on periodic beacons. There is a clear need to implement strategies that allow high priority event-driven messages to be accommodated, even with high channel loads. Finally, in terms of performance evaluations, a wide range of metrics have been used. However, some metrics such as CBR, IPD and packet loss seem to be emerging as widely-used performance indicators. There is a need to identify a set of standard metrics that can be used to evaluate and compare different approaches consistently.

5. Additional Design Considerations for V2V Congestion Control

5.1. Fairness

Fairness is important in vehicular networks to ensure that all vehicles in the network have similar opportunities to communicate with nearby nodes [20]. It implies that no vehicle should be allocated arbitrarily less resources than its neighbours. However, there is a trade-off between fairness and efficiency [93]: more fairness usually results in a less efficient use of the shared resource. Unfortunately, efficiency is critical for safety communication, so it is important to find a suitable balance between fairness and efficiency. Although many congestion control approaches consider fairness as an important objective, there is currently no universally used measure of fairness. This makes it difficult to evaluate and compare different approaches for fairness. However, in this section we highlight different ways in which some of the existing algorithms have considered fairness.
Fairness includes both local fairness and global fairness. Local fairness means nodes near each other share the same channel, i.e., they have similar congestion controls levels. In LIMERIC [11], local fairness is implemented among immediate neighboring vehicles by assuming all the vehicles are in a single collision domain. When the algorithm converges, all the vehicle will use the same Tx rate to transmit. The concept of “fair power control”, where all vehicles in a certain area must restrict their beacons’ transmission power by the same ratio is introduced in [71]. A max-min principle is used to achieve fairness, which aims to find a per-node maximal power assignment and the minimum of the nodes’ transmit power levels is maximized as long as network load conditions are satisfied. However, the max-min principle is very difficult to apply, due to the unbounded and probabilistic nature of vehicular wireless communication. Authors in [15] deem that it is not necessary to require nearby nodes to have the same Tx parameters.

For global fairness, all the vehicles contributing to congestion will be involved, which means all the vehicles in the carries sense (CS) range need to exchange the channel state information. The CS range is always larger than the Tx range, so one-hop is not sufficient to notify all the nodes within CS range. In [15], the authors propose a two-hop piggybacking to share the channel state with the CS range, hence to fulfill the global fairness principle. This maximizes the Tx rate of the nodes that are throttled the most by congestion control, while not unnecessarily constraining other vehicles which do not contribute to congestion. One drawback is that nodes measuring congestion would decrease their Tx rates first, while others contributing to congestion might not yet be aware of doing so, and increase their Tx rates even further. Effective ways to react to the same state of the system at the same time is still a critical problem for global fairness.

5.2. Awareness Control

The vast majority of the approaches discussed so far aim to control congestion by appropriately reducing the channel load. However, the ultimate objective of congestion control is to improve the overall vehicle awareness, hence to achieve the safety of the vehicles. With this view, the authors in [30] define awareness in the context of vehicular safety communication as follows:

Definition 1 (Awareness). It is the ability of an application to know the status, e.g., position, speed, heading, of neighboring vehicles. Awareness is qualified by its range, i.e., distance at which the application at most becomes aware of vehicles, and its quality, i.e., accuracy/up-to-dateness of the status information.

Most of the papers we have discussed so far focus on how to control the channel load by adapting the transmission parameters, so called congestion control protocols. The concept of awareness control is explicitly proposed in [20], which aims at ensuring each vehicle’s capacity to detect as well as to communicate with the relevant vehicles in their local neighborhood. Awareness control focuses more on the satisfaction of its individual safety applications and the traffic information of relevant local neighborhood. The issues have also been considered in some of the congestion control protocols. Based on observation that congestion and awareness control protocols have been typically designed and evaluated separately, the authors propose INTERN [94], which dynamically adapts the Tx power and rate of beacons of each vehicle to satisfy its application requirements, while controlling the channel load as the same time. The authors also claim that awareness control can achieve both local and global fairness. But when different safety applications have different transmission requirements, it is very difficult to guarantee that all vehicles in the neighborhood will have similar transmission parameters. Awareness control and congestion control can even be in conflict in some cases, e.g., when some emergency event happens. The author of [94] proposed COMPASS (cross-layer coordination of multiple vehicular protocols) in [95] which try to coordinate awareness control and congestion control protocols. This approach has to get the transmission parameters firstly from the protocols and find a intersection configuration set between awareness control and congestion control protocols which might fail.
To show how awareness considerations can influence the delivery of different safety messages in VANET, we will examine several typical scenarios.

In Figure 4a we assume a low vehicle density on the road, where each vehicle is sending the safety messages periodically in a radio range with some transmission power and rate. At this time, the channel load is under a “safe” level, usually 40% of the theoretical maximum channel capacity [14], which ensures that all the safety messages can be delivered to the other vehicles in the transmission range. As the vehicle density increases, as shown in Figure 4b, the channel load will exceed the “safe” level, which means MAC transmission delay and the number of packet collisions will grow rapidly. The MAC transmission delay will cause safety messages to arrive late, and a high number of packet collisions will lead to a lower reception probability and hence an effective radio range reduction. To decrease the channel load, all the vehicles have to take some countermeasure, i.e., congestion control, to reduce the number of safety message packets. Since there are no “special” events, fairness in this scenario implies that every vehicle should have the same chance to share its current state, i.e., same Tx rate and power.

In Figure 4c we list two typical maneuvers: acceleration and lane changing. Vehicle A and B are driving in the higher density context, so the maneuvers should be noticed instantly by their neighbours to prevent an accident by enhancing the transmission power or rate. However, vehicle C and D have different driving context, they can finish the maneuvers without any special adjustment of the transmission parameters. In [20] the authors defined awareness control to differentiate it from congestion control. Awareness control aims at ensuring each vehicle’s capacity to detect, and possibly communicate with the relevant vehicles and infrastructure nodes present in their local neighborhood as needed, through the dynamic adaptation of their transmission parameters. Since Vehicle A and B’s safety applications require more channel resources than the nearby vehicles to maintain adequate awareness, they may use a higher Tx power, compared to the surrounding vehicles, as shown in Figure 4c. Finally, in Figure 4d a car crash is assumed. The safety message here is event-driven and should be delivered to the other vehicles at the largest radio range at one time.

There are still many challenging issues in implementing appropriate awareness control strategies, such as how to define the “relevant” vehicles for a particular safety application. For example, in Figure 4c, even if vehicles C and D are on the opposite lanes as A and B, their signals may still contribute to the congestion of the area where vehicle A and B are located. The concept of awareness control is an important research problem for V2V safety communication. The goal is to find

![Figure 4. Driving Scenarios.](image-url)
a suitable balance between a single vehicle’s safety application’s requirement, while maintaining low channel load and fairness among vehicles.

6. Conclusions

V2V safety communication enhances transportation safety, by improving the awareness of the vehicles on the road. The available 10 MHz channel cannot guarantee the delivery of safety messages without latency and error when vehicle density increases. The dynamic topology of vehicular networks makes the dissemination of safety messages even more difficult and error-prone. In this paper, we reviewed the congestion problem in V2V communication and discussed the important approaches for congestion control proposed in the literature. We analyzed how congestion occurs and discussed the most commonly used metrics when evaluating the channel performance. We also identified different simulation parameters that affect the evaluation of the algorithms. Finally, we classified and compared the leading congestion control approaches in terms of different relevant criteria, such as control mechanism, parameter being adjusted, and message type.

Congestion control in VANET is an important and challenging problem. Although, some interesting and effective strategies have been reported in the literature, there are still many challenges and open research problems for congestion control in V2V safety communication in the future. A few directions for future research are given below.

- Joint power/rate control: Existing hybrid approaches that combine Tx power and rate adaptation typically implement congestion control in two different phases, e.g., fix a Tx power first and then adapt the Tx rate. A real time combined Tx power and rate control based on detailed safety benefit calculations can lead to improved performance and is a promising direction for research.
- Improved awareness control: As discussed in Section 5.2, awareness control focuses more on the relevant vehicles and local vehicles’ information. We need to be able to accurately identify “relevant” vehicles and acquire detailed local traffic information, e.g., awareness needs more specific information about the vehicle’s position, speed. So far only a few papers have considered the tracking error when implementing congestion control.
- Relative fairness: Each vehicle may have a very different driving context. So absolute fairness for both local or global fairness is not realistic. More specific relative fairness criteria are needed, when adapting the transmission parameters. It is an important problem to develop suitable metrics for evaluating fairness in different contexts and design approaches that maximize fairness.
- Standardization: In this paper, we reviewed many different metrics and approaches for V2V safety communication. It is not realistic to follow one unified process to deal with congestion problem, however at least the metrics used in the approaches should be normalized. For example, the widely used CBR metric has been referred to by a number of different names in different papers. There is a need for adoption of a common terminology and method of calculation for the performance metrics, to ensure consistent and fair evaluation of the various approaches.

Author Contributions: X.L., A.J. designed the paper, X.L. collected the data, performed the analysis, and wrote the paper under the supervision of A.J. A.J. supervised the research and provided guidance and key suggestions in writing the paper.

Funding: This research was funded by NSERC DG, Grant# RGPIN-2015-05641.

Acknowledgments: The work of A. Jaekel has been supported by a research grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Conflicts of Interest: The researcher claims no conflict of interests.

References


5. Li, Y. An Overview of the DSRC/WAVE Technology; Springer: Berlin/Heidelberg, Germany, 2012; Volume 74. [CrossRef]


56. Paranthaman, V.V.; Canberk, B.; Duong, T.Q.; Shu, L. Software defined architecture for VANET: A testbed implementation with wireless access management. *IEEE Commun. Mag.* 2017, 55, 135–141. [CrossRef]


75. Huang, C.L.; Fallah, Y.P.; Sengupta, R.; Krishnan, H. Adaptive intervehicle communication control for cooperative safety systems. IEEE Netw. 2010, 24, 6–13. [CrossRef]


95. Sepulcre, M.; Gozalvez, J. Coordination of Congestion and Awareness Control in Vehicular Networks. Electronics 2018, 7, 335. [CrossRef]