Flexible Packet Generation Control for Multi-Application V2V Communication

I. Khan†, J. Härri†
†EURECOM, 450 route des Chappes, 06904 Sophia-Antipolis, France
E-mails:{khanm, haerri}@eurecom.fr

Abstract—Decentralized Congestion Control (DCC) for 802.11p based V2X communication has been widely investigated for a single Cooperative Awareness service. As future connected intelligent vehicles will be based on multiple V2X services with a variety of application traffic pattern, DCC mechanisms should be capable of limiting channel congestion while satisfying the channel resource requirement of a range of heterogeneous applications in each node.

In this paper, we analyze the application layer rate control approach of Facilities DCC, currently being standardized in Europe, and propose a flexible packet generation control, which gives each node more flexibility and agility for resource management among its applications. Our simulation based results show that a flexible allocation can more rapidly serve application packet generation requests while keeping the channel load within the desired level even with high node density. More importantly, it solves the starvation issue, commonly associated with the strict priority based traffic shaping approach of DCC.

I. INTRODUCTION

Cooperative vehicular networks will be deployed in the near future to increase a vehicle’s awareness and improve road safety and traffic efficiency by periodically exchanging awareness information to complement a vehicle’s on board sensors. The awareness messages propagate a vehicle’s status information such as position, speed, heading etc. and share a vehicle’s intelligence to neighboring vehicles and road infrastructure. These messages are expected to be broadcast using a potential wireless communication technology called DSRC in USA and ITS-G5 in Europe, based on IEEE 802.11p a variant of the Wi-Fi standard.

In vehicular networks based on IEEE 802.11p there is no centralized scheduler to regulate the wireless channel access. The absence of any regulation can easily lead to channel congestion due to periodic message broadcast. Thereby each node must regulate its spatial and temporal channel access, adopting a decentralized channel resource allocation strategy with a global objective to maintain the channel load below a threshold for the common benefit of all nodes. Thus, many Decentralized Congestion Control (DCC) mechanisms have been proposed over the years and have been standardized by ETSI in Europe [1]–[3] and SAE in USA [4], to manage periodic message broadcast and ensure adequate awareness range, information freshness, channel access fairness and network stability.

In USA the approach to DCC has been cross-layer, considering multiple sensing parameters, such as vehicular traffic density, packet error rate, neighbor tracking error etc. While, in Europe until lately, DCC has been mainly at the Access Layer. Recently, DCC in Europe is being extended to the upper layers of the ETSI ITS stack and new mechanisms have been proposed to distribute the channel resource among the applications of a vehicle and control the packet generation rate. In this paper, we analyze this aspect of packet generation control and demonstrate issues with the approach proposed in ETSI Facilities DCC. We propose a flexible resource management approach to better serve multiple heterogeneous safety applications on the same channel, balancing their demands and avoiding application starvation during resource scarcity.

The rest of the paper is organized as follows: Section II presents a brief overview of DCC mechanism, followed by Section III which discusses several issues with DCC Facilities. Section IV presents a flexible resource allocation approach, followed by Section V providing performance evaluation results. Lastly, Section VI concludes the paper.

II. CONGESTION CONTROL AND RELATED WORK

Over the years, a variety of avenues have been explored for V2X channel congestion control. The most common approach has been to monitor the channel load and limit individual temporal channel resource usage i.e. transmit rate of each node [5], [6] or individual spatial channel usage i.e. transmit power [7], [8]. Similarly, the work in [9] aims to control transmit rate by additionally considering application requirements. Other approaches try to enhance channel usage efficiency by optimizing the data rate [10] or influence the channel load monitoring by tweaking the carrier sense threshold [11].

Most of these aforementioned works have considered a single type of packet, mainly single hop periodic broadcast of Cooperative Awareness Message (CAM) or Basic Short Message (BSM). Only recently, some works have started analyzing rate control considering multiple applications and highlighting the shortcomings of existing rate control approaches in this regard. The works in [12], [13] illustrate the starvation problem of Access DCC when dealing with multiple traffic classes. However, there is a shortage of work regarding packet generation control aspect of rate control, in particular balancing the needs of multiple applications with diverse traffic pattern. In this paper we focus on this aspect.

Moreover, the motivation for analyzing multiple heterogeneous applications on the same channel is that although 3 channels of 10MHz have been reserved in the 5.9GHz for vehicular usage, other types of technologies, mainly Wi-Fi and cellular V2X could be allowed to operate in these spectrum [14]. In such a scenario, only 1 channel will be left exclusively for DSRC or ITS-G5, which will have to be optimally exploited to serve multiple safety applications.

A. Rate Control via obligatory non-transmission

Transmit Rate Control (TRC), is the most common strategy of V2X channel congestion control. As shown in Fig 1, it is controlled by either limiting the number of packets released
into the medium via flow control at the access layer or by limiting the number of packets generated by the applications of a node in the upper layers. The standards on congestion control do not impose the exact flow control algorithm, but specify the channel usage limit for a node, typically by enforcing an obligatory gap between two transmit opportunities, which can be applied for flow control at the access layer using a leaky-bucket or an obligatory period of no packet generation after granting packet generation opportunity to an application.

1) Flow Control: According to ETSI standards [2], [15], the channel resource limit per node follows the relation:

\[ \text{ChannelResourceLimit}_{\text{perNode}} = \frac{\text{ChannelUsageLimit}}{\#\text{Neighbors}} \] (1)

The Channel Resource Limit (CRL) for each node is defined as a duty cycle, as the ratio of the transmit duration or packet airtime denoted as \( T_{on} \), to the sum off \( T_{on} \) and an obligatory non-transmission period \( T_{off} \) according to:

\[ \frac{T_{on}}{T_{on} + T_{off}} = CRL \quad T_{off} = T_{on} \times \frac{1 - CRL}{CRL} \] (2)

Therefore, to stay within the CRL, after each transmission of duration \( T_{on} \), the flow control mechanism at the access layer, prevents any transmission for at least a duration of \( T_{off} \) via a leaky-bucket Gate Keeper or any other mechanism.

2) Packet Generation Control: A problem with pure flow control without controlling the generation of packets is that applications might generate excess packets, which would queue up and age in the flow control queues, causing old packets being transmitted or dropped. Therefore, influencing the generation of packets at the application layer allows to control applications to generate packets according to the channel capacity and prevent excess packet generation.

In this regard, DCC Facilities is being standardized as ETSI TS 103 141 [3], which operates at the Facilities Layer [16] below the Application Layer, to control the packet generation. It sub-divides the total resource available for the node to individual resource for each application, as shown in Eq. 3, prioritizing resource allocation by application traffic class (TC).

\[ \frac{T_{on}}{T_{on} + T_{off}} = \frac{T_{on,app1}}{T_{on,app1} + T_{off,app1} + ... + T_{on,appN} + T_{off,appN}} \] (3)

The allocated resource to an application is translated into packet generation limits for that application. After generating a packet of airtime such as \( T_{on,app} \), DCC Facilities allows it to generate the next packet only after the corresponding non-generation period \( T_{off,app} \) of that application, to restrain it within its allocated resource limit. DCC Facilities does not involve queuing or dropping packets. It only delays allocating packet generation approval to applications using the \( T_{off,app} \).

The various components of rate control have been described in [17]. In this paper we analyze in particular the application packet generation control, as we focus mainly on the upper two blocks of Figure 1, and not the access layer flow control.

Moreover, the dynamicity of external conditions such as neighbor density variation and dynamic variation of resource available per node is not considered in this recent results paper, and are left for future work.

III. Packet Generation Control - Issues

As discussed in the previous section, DCC Facilities controls application packet generation by imposing an obligatory non-generation of duration \( T_{off,app} \) after each packet generation of airtime \( T_{on,app} \). Similarly generation opportunity is prioritized solely via application TC. These two features can be problematic, as described in this section.

Resource Monopolization & Starvation: DCC Facilities prioritizes transmit opportunity to applications solely based on TC. A higher TC enjoys absolute priority over a lower TC, and a lower TC is not served unless the demand of a higher TC is fully met. This is problematic during resource shortage, that is, sum of application demands > channel resource limit per node. In such a case, only the applications having higher TC will be served and applications of lower TC will be starved. This is an inherent problem of simple priority based systems such as priority based queuing.

Delaying Packets: The obligatory \( T_{off,app} \) after every generation could suffice for periodic packets such as CAMs but for Day 2 scenario, there will be a variety of applications with heterogeneous packet size, periodicity and traffic generation pattern. One application is exchanging a vehicle’s perception, which will generate bursts in a scenario such as an intersection, to communicate a node’s sensor information to potentially interested neighbors. The obligatory \( T_{off,app} \) period, will not allow bursts, causing information to be outdated and no longer useful for potential neighbors. DCC allows bursts only for emergency messages such as Decentralized Environment Notification Message (DENM), which is not bound by the rate control mechanism of DCC Facilities.

Memoryless: The channel load and resource available per node is updated every 100ms, but there is no notion of any resource allocation period among the applications and the past behavior or an application is not considered when calculating the \( T_{off,app} \), which depends solely on the application’s last packet’s \( T_{on,app} \). Therefore, if a node or an application pauses for a while to accumulate opportunities and transmit a burst, it is not permitted by the obligatory \( T_{off,app} \).
IV. FLEXIBLE PACKET GENERATION CONTROL

In order to increase a node’s control over allocating packet generation opportunity to its applications, we propose a flexible packet generation control, using an approach similar to the token-bucket algorithm. The temporal channel resource for the node per duty cycle, is considered as a resource allocation cycle, such as one second, and each application with respect to its TC is allocated a resource quota for the cycle.

The notion of cycle and quota allows each node more flexible application resource distribution while remaining within the resource limit, and dynamically optimize packet generation control rules based on the transmission history, the node’s context and resource availability. For example, during resource shortage, a starving lower priority application can eventually be served by diverting some resource from higher priority applications. Similarly, unused quota from one cycle can be carried over to the next cycle or quota for one cycle can be increased by borrowing from the next cycle to allow an emergency burst. This approach enables each node to better manage its temporal channel resource and optimally balance applications’ demands, instead of the hard and fast rule of obligatory \( T_{off} \) after every \( T_{off} \).

A flexible packet generation control algorithm is presented in this section to illustrate this approach.

Begin Cycle;

```plaintext
while new request from app do
  if quota available then
    grant transmission request;
    update quota remaining for cycle;
  else
    reduce future cycle quota from lowest priority app;
    if app deferral quota available then
      defer request to next cycle;
      update app deferral quota;
    end
  end
update quota for next cycle;
End Cycle
```

**Algorithm 1:** Flexible Packet Generation Control Algorithm

The algorithm has 4 key aspects:

i) As packet generation requests come, the available quota in the cycle is spent and packets generated without any delay. 
ii) If the quota is insufficient, requests in the current cycle are deferred during the beginning of the next cycle, withing a postponement window having a jitter. The order and the number of allowed deferrals depend on the TC. 
iii) During resource shortage, the quota of an application is reduced for one/several future cycles, based on the TC.
iv) If the quota for an application is zero for a maximum successive cycles, quota is shifted from the next higher priority application via gracious degradation to prevent indefinite starvation.

The main control parameters are: i) number of deferrals allowed ii) normal application resource quota iii) reduced application resource quota iv) reduction persistence duration. By default the parameters prioritize applications based on TC, but can be dynamically modified to prefer an application over another. This aspect of dynamic optimization of the parameters need detailed investigation and is out of the scope of this paper. The goal of this paper is to present the feasibility and benefit of flexible resource allocation compared to strict priority and control via obligatory \( T_{off} \).

To summarize, this flexible approach aims the following objectives: i) Respect the application packet priority ii) Reduce the delay between packet generation request and opportunity iii) Prevent low priority application starvation iii) Respect the resource limit set for the node.

Figure 2 shows an example of the packet generation opportunity for a bursty traffic pattern. Let’s consider a node has a single application, requesting 4 packets of 500 Bytes to be generated at 100ms interval. Each node is limited to use 0.25% channel capacity. Using Eq. 2, the obligatory \( T_{off} \) is 267ms for 500 Bytes at 6 Mbps data rate. The first packet is generated without a delay, while the second packet is delayed by 167ms, the third packet by 334ms and the fourth packet by 501ms, giving a total delay of 167+334+501 = 1002ms.

Using the flexible allocation, a 0.25% of channel resource per node gives a quota of 1875 Bytes per duty cycle of one second. Following Algorithm 1, the first 3 packets are generated without any delay, as shown in Fig 2. When the fourth request arrives at 600ms, the quota left for that second is 1875-1500 = 375 Bytes. Accordingly, the fourth packet is deferred towards the start of the next second, and generated at 1001ms (can vary slightly due to jitter). Therefore, the total delay with respect to granting generation opportunity for the 4 packets using \( T_{off} \) is 1002ms, whereas it is only 401ms using the flexible allocation of Algorithm 1.

V. PERFORMANCE EVALUATION

In the rest of the paper, we refer to the packet generation control via obligatory \( T_{off} \) as the strict method of resource allocation, and our alternative approach using Algorithm 1 as the flexible method. We compare the performance of the two

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmit Rate</td>
<td>CAM: 3-5 [Hz], CPMA: 4 [Hz]</td>
</tr>
<tr>
<td></td>
<td>LDM: 1 [Hz]</td>
</tr>
<tr>
<td>Transmit Power</td>
<td>2.5 dBm</td>
</tr>
<tr>
<td>Packet Size</td>
<td>CAM: 300 Bytes, CPML: 300 Bytes</td>
</tr>
<tr>
<td></td>
<td>LDM: 750 Bytes</td>
</tr>
<tr>
<td>Packet Priority</td>
<td>CAM: DCO Profile (DP2), CPMA: DP3</td>
</tr>
<tr>
<td></td>
<td>LDM: DP4</td>
</tr>
<tr>
<td>Data Rate</td>
<td>6 Mbps</td>
</tr>
<tr>
<td>Number of Nodes</td>
<td>100 to 240</td>
</tr>
<tr>
<td>Mobility</td>
<td>Static</td>
</tr>
<tr>
<td>Simulation Time</td>
<td>30 seconds</td>
</tr>
<tr>
<td>PHY and MAC</td>
<td>ITS-G5 802.11p in 5.9 GHz (10 MHz Control Channel)</td>
</tr>
<tr>
<td>Fading</td>
<td>Winner B1 Urban Microcell (Correlated Gaussian &amp; Ricean)</td>
</tr>
<tr>
<td>Preamble DetectionThreshold</td>
<td>92 [Hz]</td>
</tr>
<tr>
<td>Performance Indicators</td>
<td>pkt Generation Request Delay, Rx Rate, 50%</td>
</tr>
<tr>
<td></td>
<td>95% Confidence Interval</td>
</tr>
</tbody>
</table>
approaches, in terms of how rapidly the demand for packet generation by each application is served. Secondly we analyze the packet reception rate of multiple applications belonging to different TC.

A simple scenario consisting between 100 to 240 nodes equipped with ITS-G5 transmitters and the ETSI ITS stack is simulated on the iTETRIS simulator [18], which has a full ITS-G5 protocol stack implemented on top of NS-3. The channel has fading according to WINNER B1 model, and all nodes are in Line of Sight (LOS) without hidden node. In this recent results paper, we don’t focus on the mobility aspect and the nodes are static in a grid formation with 5m gap between the nodes. Although in reality, vehicular application traffic pattern depends on node mobility, we intend to do a more complete analysis will realistic mobility traces for future work.

Each node runs 3 applications, broadcasting packets on the same channel i.e. 300 Bytes CAM, 500 Bytes Cooperative Perception Message (CPM) and 750 Bytes Local Dynamic Map (LDM), having different traffic pattern. CPM (ETSI TS 103 324), currently being standardized at ETSI to communicate a vehicle’s various sensor information to its neighbors. LDM (ETSI TS 102 863 [19]) is a message which ITS stations will use to exchange their Local Dynamic Map.

CAM transmission requests are generated at a rate between 3 to 5 Hz. This is a rough approximation of CAM triggering conditions [20], in urban mobility scenario. CPMs are requested to be generated at 4Hz, at regular burst of 4 packets within a span of 200 milliseconds. Lastly, requests for LDM are generated at a rate of 1Hz. Considering a data rate of 6Mbps, this traffic pattern demands channel resource between 0.486% and 0.56% per node. The performance of the two application resource allocation methods are compared by allocating each node a channel resource limit between different TC.

The performance is evaluated in terms of access delay, packet reception rate and channel load. As the traffic pattern is heterogeneous and not just periodic, the inter reception time is not considered. The results are average of 50 simulation runs with 95% Confidence Interval. For each run, the exact same packet generation request is used to compare both the strict and flexible allocation mechanisms. Table I summarizes the main simulation parameters.

Compared to the strict Toff, the service delay is much shorter using the flexible packet generation control, as shown in Fig 3b. In the case of 100 neighbors, there is zero delay, whereas the strict method causes a delay between 60 to 200 milliseconds depending on the application. Similarly, the delay for flexible approach is lower for other node densities compared to the strict method. More importantly, even as the neighbor density increases, the LDM application is not starved, which is explained further in the next sub-section.

B. Packet Reception Frequency

Figure 4 shows the packet reception frequency for the 3 applications CAM, CPM and LDM, for both the strict and flexible resource allocation. The transmission and reception rates follow the same trend, so only the reception rate is analyzed due to space limitation.

The average reception frequency is 3.4 Hz for CAM, 3.8 Hz for CPM and 1 Hz for LDM, when demand < resource available per node, and 100 nodes share the channel. Similarly, as the neighbor number increases, following Eq. 1, the resource limit per node decreases, decreasing the transmission and reception rates. From a neighbor number of 160 nodes and onwards, LDM is starved by the strict allocation, causing zero reception. Initially CPM has higher rate than CAM, as the average CPM Tx rate is higher. During resource scarcity, CAM preferentially achieves higher Tx and Rx rates than CPM.

The reception rate of the flexible allocation follows the strict allocation till 120 nodes. Afterwards, the rates of CAM and the strict packet generation control. CAMs have the highest priority, DCC Profile (DP) 2, then CPM with DP 3 and LDM with DP 4, so the delay patterns follow this ratio. As mentioned earlier, the traffic pattern demands between 0.486% and 0.56% channel resource per node. With a target channel load of 60%, when 100 nodes share the channel, using Eq. 1 the allocation is 0.6% channel resource per node. Thereby all 3 applications are satisfied as demand < availability. With 120 nodes, each node has 0.5% resource and the demand begins to exceed the supply, so the LDM delay increases sharply. Similarly, when 160 nodes share the channel, the resource per node is 0.375% which is consumed by the other two applications, while LDM is not given any transmit opportunity at all. Therefore there are no delay values in Fig. 3a for 160 nodes and onwards.

A. Delay between Packet Generation Request & Opportunity

Figure 3a shows that delay between packet generation request and generation opportunity for the 3 applications using...
Although the flexible approach for packet generation control performs well in a static regime, its performance has to be analyzed in dynamic regime with realistic mobility pattern and dynamically changing external conditions, which we leave for future work.

ACKNOWLEDGMENT

The authors acknowledge the support of the SCHEIF project within the French-German Academy for the Industry of the Future (www.future-industry.org). EURECOM acknowledges the support of its industrial members, namely, BMW Group, IABG, Monaco Telecom, Orange, SAP and Symantec.

REFERENCES

[15] “ETSI TR 101 612 intelligent transport systems (its); cross layer ddc management entity for operation in the its g5a and its g5b medium; report on cross layer ddc algorithms and performance evaluation,” 2014.

CPL are slightly lesser for the flexible allocation, as some resource is diverted to transmit 750 Bytes LDM, at least at 0.3 Hz even when 240 nodes are sharing the channel.

Compared to strict allocation, flexible allocation results a 0.5 Hz and 0.2 Hz reduction for CAM and CPM reception frequencies respectively. In terms of balancing the resource, 0.3 Hz 750 Bytes LDM corresponds to a capacity of 225 Bytes, which is earned by sacrificing 0.2 Hz 500 Bytes CPM i.e. 100 Bytes and 0.5 Hz 300 Bytes CAM, i.e. 150 Bytes. Therefore, some resource is shifted from one application to another to prevent starvation. The amount of resource to divert can be controlled by using the quotas for each application, as described in Section IV. Thus, even as the node density increases, the flexible allocation tries to serve all the applications instead of drastically sacrificing the lower priority application as done by the strict allocation.

Lastly, Figure 5 shows the channel load for both the methods. Both the mechanisms can limit the channel load to the set limit of 60%, with slight variation between the two. This proves that while remaining within the channel load limit, the resource allocation via flexible method provides better application performance than the strict allocation.

VI. CONCLUSION

In this paper we present and analyze the packet generation control aspect of DCC transmit rate control. We illustrate that the systematic non-transmission interval after a transmission degrades application performance. This is particularly problematic in case of multiple applications having heterogeneous and non-periodic traffic pattern. Similarly, allocating resource to applications solely based on traffic class can indefinitely starve lower priority applications.

We propose a flexible packet generation control and resource allocation approach, which divides the temporal axis into periodic resource allocation cycle and allocates resources among the applications flexibly without involving any strict non-transmission period. This flexible approach allows each node to change control parameters to dynamically regulate the packet generation delay and amount of resource for any application. Simulation results show that the flexible allocation performs better than using an obligatory non-transmission interval for varying levels of node density, in terms of allocating packet generation opportunity, and preventing lower priority applications from starvation during channel resource scarcity.