Decentralized Cooperative Resource Allocation with Reliability at Four Nines

Morejon, Santiago; Bruun, Rasmus; Sørensen, Troels Bundgaard; Kiilerich Pratas, Nuno; Madsen, Tatiana Kozlova; Lianghai, Ji; E. Mogensen, Preben

Published in:
2021 IEEE Global Communications Conference (GLOBECOM)

DOI (link to publication from Publisher):
10.1109/GLOBECOM46510.2021.9686004

Publication date:
2021

Document Version
Accepted author manuscript, peer reviewed version

Decentralized Cooperative Resource Allocation with Reliability at Four Nines

C. Santiago Morejón García*, Rasmus Liborius Bruun*, Troels B. Sørensen*, Nuno K. Pratas†, Tatiana Kozlova Madsen*, Ji Lianghai† and Preben Mogensen*†
*Wireless Communication Networks Section, Department of Electronic Systems, Aalborg University, Denmark
†Nokia Standardization, Aalborg, Denmark

***Abstract*—Decentralized cooperative resource allocation schemes for robotic swarms represents an alternative to infrastructure-based communications across different commercial, industrial and environmental protection use cases. The cooperative communication schemes, device sequential and group scheduling in [1], have shown superior performance in comparison to 5G NR sidelink mode 2, but have also shown performance issues due to signaling overhead and signaling induced failures. In this paper we introduce different techniques that reduce the failure probability of data packet transmissions and the packet inter-reception (PIR) time. We evaluate two techniques, respectively, of incremental redundancy using hybrid automatic repeat request and link adaptation by aggregation, as well as their combination for our decentralized cooperative resource allocation schemes and sidelink mode 2. Our results show that the introduced enhancements, allow to double the amount of supported swarm members while achieving four nines reliability when compared to the case where the same enhancements are applied to the sidelink mode 2.**

***Index Terms*—decentralized resource allocation, swarm communication, cooperative communication, reliability, link adaptation, HARQ, packet inter-reception**

I. Introduction

Robots will replace humans in even more complex operations of future industrial production. For that purpose, proximity communication will play a vital role in enabling cooperation between collaborating robots. Proximity communication involves collective perception of the environment by sharing video streams. These require a 10 Mbps data rate, with a maximum latency of 10 ms, and a reliability of 99.99 % (equivalent to a 10^{-4} transmission failure probability) as mentioned in [2]. Even though the perception is collective, each individual robot will be governed by a control loop. This control loop is vulnerable to variations in the arrival timing of the input [3], and these timing variations should be kept at a minimum.

To address these requirements, in [4] two decentralized cooperative resource allocation schemes were proposed. In [4] it was shown that these scheme were able to outperform the baseline resource allocation scheme defined for New Radio (NR) sidelink, termed as NR Sidelink Mode 2, by an order of magnitude. The proposed techniques in [1] addressed the issues related to the control plane signaling associated with the resource allocation, in particular control plane signaling blocking reception (half duplex) of user plane data and unsuccessful reception of control plane signaling.

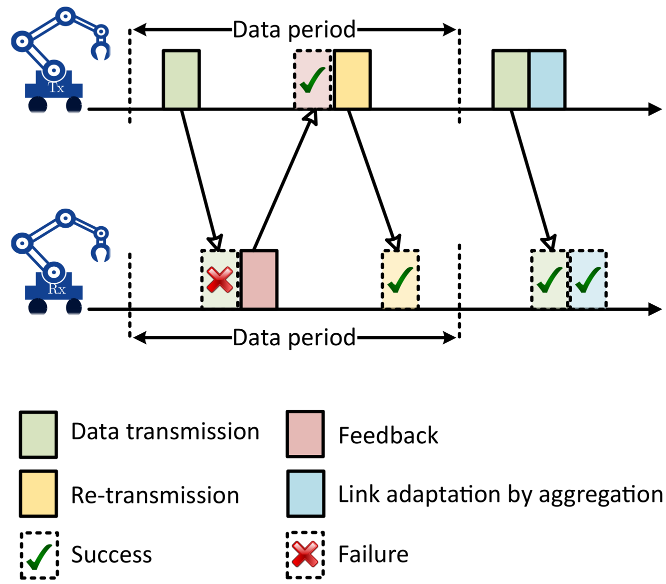

Fig. 1. Data failure reception recovered by HARQ or incremental robustness to subsequent transmissions given by link adaptation by aggregation for UEs sharing high-throughput data

However, even with the proposed enhancements in [1], these two resource allocation schemes were not able to achieve the targeted performance requirements, as in [1] it was shown that after the control plane issues have been addressed, the next performance bottleneck occurred in the user plane.

The aim of this paper is to introduce user plane focused enhancements that enable the proposed resource allocation schemes to reach 99.99 % reliability. In addition, we also evaluate the time variations of the control loop input by using the packet inter-reception (PIR) metric, which was defined in 3GPP [5] as the time in between successive packet receptions.

One well known and widely utilized technique to improve reliability on a per-transmission basis is hybrid automatic repeat request (HARQ). In HARQ, forward error correction is combined with re-transmissions (illustrated by yellow boxes in Figure 1) to flexibly adapt the redundancy in the transmission to cope with the current channel conditions.

To maintain the reliability of subsequent transmissions, link adaptation is utilized. It controls the modulation and coding scheme (MCS) (illustrated by blue boxes in Figure 1) to meet the configured average block error rate in dynamic channel conditions [6].

In this paper we show how HARQ and link adaptation can be adapted to the scenario and evaluate the impact to reliability and PIR in proximity communications. Specifically, the contributions of this paper are:

- Augmentation of the decentralized resource allocation schemes with HARQ and link adaptation by aggregation to reach the four nines reliability
- Characterization of the PIR to validate the suitability for control loop operation

In Section II we explain the HARQ and link adaptation by aggregation techniques in detail. The simulation setup is briefly described in Section III and the results and evaluation follow in Section IV. Concluding remarks are made in Section V.

II. Failure causes and enhancement techniques

Resource allocation is a complex task with impact on communication. The performance of the resource allocation is determined by the available information (obtained through passive or active means) and how the information is being utilized (the algorithm and computation power available). Resource allocation is an NP-hard problem [7] and often appear in a context where time is a limiting factor. Thus, an appropriate solution is dependent on the specific context. We briefly summarize the considered resource allocation schemes from [1], [4] in the following.

1) Sidelink mode 2: The baseline scheme is the current procedure for autonomous decentralized resource allocation in 5G NR called sidelink mode 2. In this procedure, a UE (term used for the communication module attached to a robot) senses the assigned communication resources (resource pool) during the sensing window prior to resource allocation. It then excludes resources from the candidate set based on reoccurring semi-persistently scheduled (SPS) transmissions for the upcoming allocation based on the reference signal received power level (RSRP). However, 20 % of the potential resources must remain in the candidate set. The occupied resources with lowest RSRP may be re-included into the candidate set to meet this criteria. The transmission resource is chosen randomly from the set of candidate resources.

In proximity communication, beside exchange of application data (be it video or any other high data rate stream), discovery between robots (UEs) is paramount, and performed by periodic transmission of discovery messages (DM) in a resource pool dedicated for control-type transmissions. The discovery messages include at least position and heading information but can be optionally extended.

2) Device sequential [4]: Device sequential resource allocation takes advantage of cooperation between UEs in the resource allocation phase. In addition to sensing ongoing SPS transmission from other UEs, each UE includes in their DM the time at which they will initiate resource allocation, denoted by the *trigger time*. Thereby, UEs in proximity will be aware of others' intentions to allocate resources, and the UEs can follow a sequential procedure of allocating a resource,

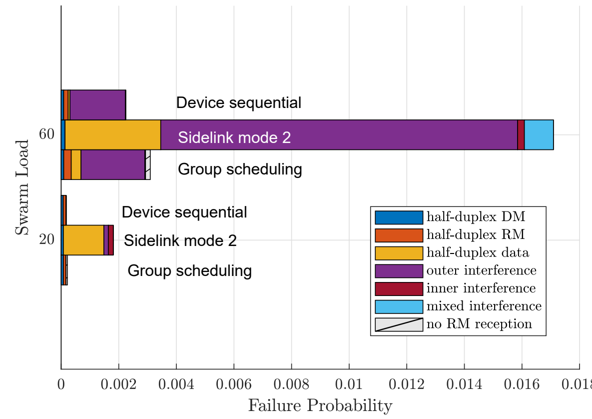

Fig. 2. Failure probability and the causes of data transmission failures (halfduplex of DM, RM and data, inner, outer and mixed interference, and no RM reception) for three resource allocation schemes with enhanced (error-prone) signaling

and publishing the allocation immediately by transmitting a resource selection message (RM) in the control resource pool. Upon reception, the next UE in the sequence proceeds. This cooperative procedure allows UEs to select the resource they seem best fit without relying on a random procedure.

3) Group scheduling [4]: Group scheduling resource allocation builds on the idea to save signaling and build a wider information base by letting local group leaders collect sensing results and perform resource allocation for multiple group members simultaneously. This scheme implies the addition of a leader selection phase. We found in [1] that the required signaling could be contained in the discovery messages with negligible impact on DM reliability. When the leaders need to cooperate, they do so by following the sequential procedure of transmitting the RM, which contains the resource allocation assigned to group members.

A. *Failure causes*

Figure 2 shows the causes of data reception failure and their prevalence in each resource allocation scheme without utilization of enhancement techniques for 2 swarm loads (20 and 60 UEs) following the methodology presented in [1]. Half-duplex issues caused by communicating UEs simultaneously transmitting data is consistently a cause of failures for the sidelink mode 2 resource allocation scheme. Half-duplex issues caused by transmission of control messages (DMs and RMs) account for only a small part of the failures for all schemes. In dense swarms, interference from other data transmissions become the main cause of reception failures for all schemes. Specific to the group scheduling resource allocation scheme is the non-reception of RMs from the leader. This failure cause was greatly reduced by the RM retransmission technique introduced in [1].

The PIR metric will be impacted by data reception failures. During the SPS transmissions the PIR - in absence of failures - will be equal to the 10 ms period of data transmission. PIR can

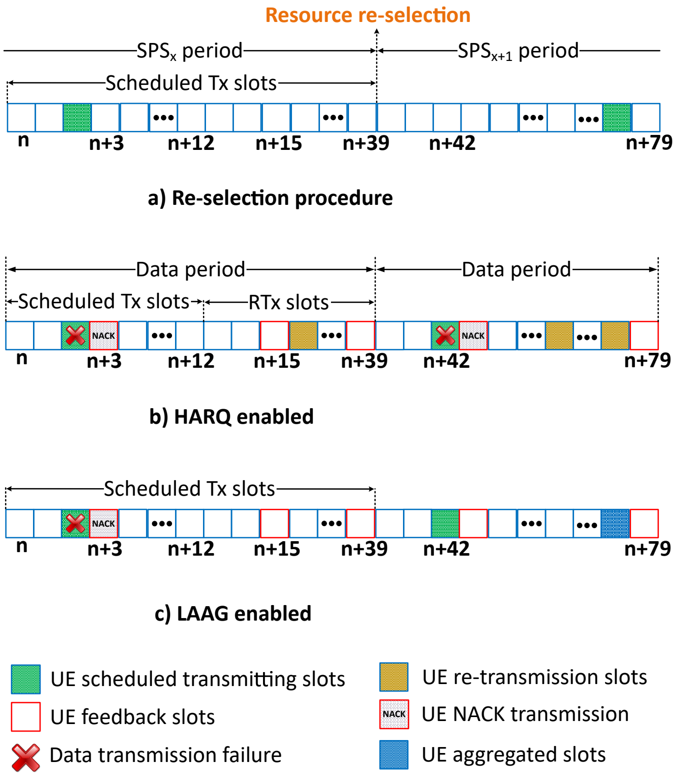

Fig. 3. Resource utilization a) during resource re-selection, b) with HARQ enabled and c) with LAAG enabled

exceed 10 ms due to 3 reasons. When SPS transmissions are reconfigured, the PIR will deviate from the 10 ms in case the SPS resource is re-selected. This behavior can be observed in Figure 3 (a) where the UE re-selects the transmission resource from the third slot (n+2) in SPS period x to the second last slot (n+78) in SPS period $x+1$. This will cause a PIR greater than 10 ms. However, the latency requirement is still fulfilled because it is measured relative to the data period.

The second reason for PIRs above 10 ms is data reception failures which triggers re-transmissions, and the third reason is the combination of the former.

In the following, we will introduce techniques with the potential to tackle the reception failures and elevate the performance of the data transmission to meet the four nines reliability requirement in proximity swarm communication.

B. Enhancement techniques

The mitigation techniques rely on feedback from the receiving to the transmitting UE. 5G NR provides flexibility to customize the slot configuration such that the feedback channel (PSFCH) is included [8]. Specifically, the higher layer parameter sl-PSFCH-Period-r16 (defined in [8]) can be set to 0 (feedback disabled), 1 (feedback in all NR slots), 2 (feedback in every second NR slot), or 4 (every fourth NR slot). Within each feedback slot the receiver UE can send a negative-acknowledgment (NACK) to its transmitter. We have chosen a periodicity of 4 NR slots for our implementation as shown in Figure 3 (b) and (c).

1) NR HARQ re-transmissions: The purpose of hybrid automatic repeat request (HARQ) is to improve reliability at the expense of redundant information added to a transmission or as additional transmissions. In [1] we introduced the *trigger time* as the time at which a UE needs to perform a resource allocation and the data period of 10 ms starts. To avoid the fact that resources are allocated in slot(s) at the end of the data period, and hence do not give enough room to perform re-transmissions, we have divided the data period when the HARQ re-transmissions are enabled into two: scheduled Tx slots and RTx slots. Scheduled Tx slots compose one-third (13 slots as depicted in Figure 3 (b)) of the total data period where resources are allocated to perform the first data transmission attempt. If receiver UEs were not able to decode the transmitted data or have not received data within the scheduled transmission window (deprived transmission explained in [4]), they proceed to send a NACK in the following feedback slot (marked by red outline in Figure 3 (b)). Once the transmitter UE receives the NACK, it randomly selects a slot(s) within the RTx slots to perform a re-transmission. If this re-transmission is not successfully received, the procedure repeats while there are available slots in the RTx slots period.

In our implementation we introduce HARQ with soft combining. It uses chase combining with a combining efficiency factor $\eta = 1$. The resulting SINR, γ_{CC}, is calculated as

$$\gamma_{CC} = \sum_{i=0}^{R} \gamma_i \cdot \eta^R \tag{1}$$

where R is the number of re-transmissions and γ_i is the SINR of the ith (re-)transmission (the original transmission has index 0).

2) Link adaptation by aggregation: Link adaptation by aggregation (LAAG) works by allocating additional resource(s) when a UE fails one of its SPS data transmissions. When a data transmission failure happens, the receiving UE proceeds to send a NACK in the following available feedback slot (see Figure 3 (c)). After the transmitting UE successfully received this NACK, it proceeds to autonomously (i.e. without cooperation) allocate additional resource(s), allowing it to utilize a lower MCS index for the subsequent transmission (slot n+78 in Figure 3 (c)) and until SPS resources are reconfigured. This increases the robustness for subsequent data transmissions and the receiver will be able to decode the data at lower SINR. Successful reception of the transmission with additional resources is dependent on the effective SINR of the aggregated transmission. In our implementation, we determined the effective SINR, γ_{MIC}, by using the SINRs of each resource combined by the mean instantaneous capacity (MIC) [9] calculated as,

$$\gamma_{MIC} = 2^{\frac{1}{K}\sum_{k=1}^{K} log_2(1+\gamma_k)} - 1 \tag{2}$$

where K is the number of resources and γ_k is the SINR of the k^{th} resource.

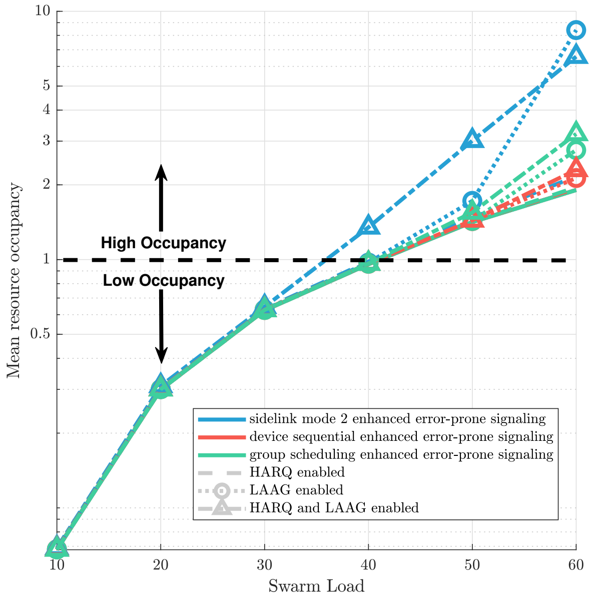

Fig. 4. Mean resource occupancy for swarm load from 10 to 70 UEs. Mean occupancy below one is considered low

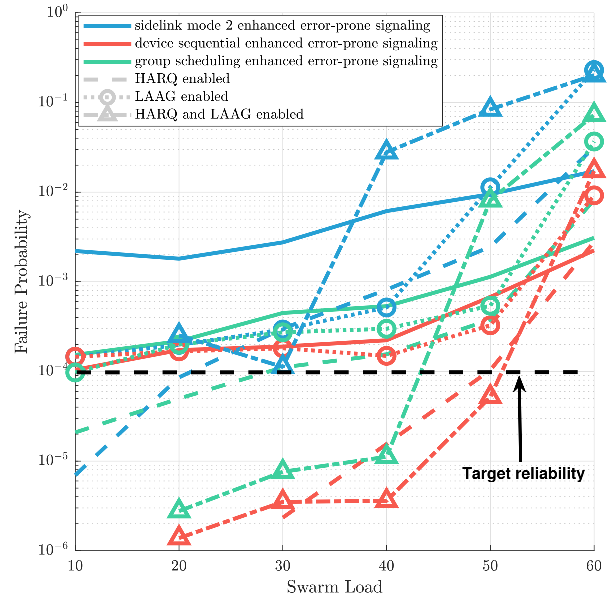

Fig. 5. Sidelink mode 2, device sequential and group scheduling failure probability for: error-prone signaling, signaling with HARQ, signaling with LAAG and signaling with HARQ and LAAG enabled

III. System level evaluation

We went beyond our system level simulator development presented in [1] by implementing the two previously introduced techniques and the evaluation of PIR and resource occupancy. Resource occupancy is defined as the average number of UEs occupying a single time-frequency resource.

The simulation models an indoor factory in which UEs move around following the random way-point model. The pathloss follows the 3GPP indoor factory model found in [10]. In addition we applied correlation to the shadow fading component by following the technique proposed by [11]. When UEs get within a 5 meter distance of another UE, they initiate the proximity communication in which multi-cast is utilized for message exchange. Each UE selects the modulation and coding scheme (MCS) such that a 100 kbit data message can be transmitted with an expected block error rate (BLER) of 0.01 % at the estimated SINR-conditions.

Table I presents the simulator settings. For the evaluations, simulations with the following four configurations were performed:

1) **Enhanced (error-prone) signaling** in which the successful reception of data messages was enhanced by the techniques of RM re-transmissions, non-overlapping and piggybacking as presented in [1]
2) **Signaling with HARQ enabled** in which in addition to 1) the HARQ technique is utilized
3) **Signaling with LAAG enabled** in which in addition to 1) the link adaptation by aggregation technique is utilized
4) **Signaling with HARQ and LAAG enabled** in which both HARQ and LAAG are enabled in addition to 1)

TABLE I
SIMULATION PARAMETERS

Parameter	Value/range
Carrier frequency, f_c	3.5 GHz
Swarm size (number of UEs)	$[10, 20, 30, 40, 50, 60, 70]$
Critical cooperation range, r_c	5 m
Extended Cooperation range, r_e	25 m
Facility dimensions	120×50 m^2 [10]
Transmission power, P_{tx}	0 dBm
Data channel bandwidth	100 MHz
Control channel bandwidth	7.2 MHz
NR slot duration	250 μs
Thermal noise power spectral density	-174 dBm/Hz
Receiver noise figure	9 dB
UE speed	1 m/s
Mobility model	Random waypoint (RWP)
Pathloss model	InF-SL [10]
De-correlation distance δ	20 m [11]
Discovery message periodicity	100 ms
Data message periodicity	10 ms
Data message size	100 kb
sl-PSFCH-Period-r16	1 ms
Scheduled Tx slots window	3.33 ms
RTx slots window	6.67 ms
Simulation time	1000 s

IV. Simulation results

The resource occupancy is illustrated in Figure 4. We see that swarm loads below 40 UEs correspond to low resource occupancy, i.e. in average less than one UE per resource, while swarm loads of 40 and above result in high resource occupancy. The mean resource occupancy is negligibly affected by

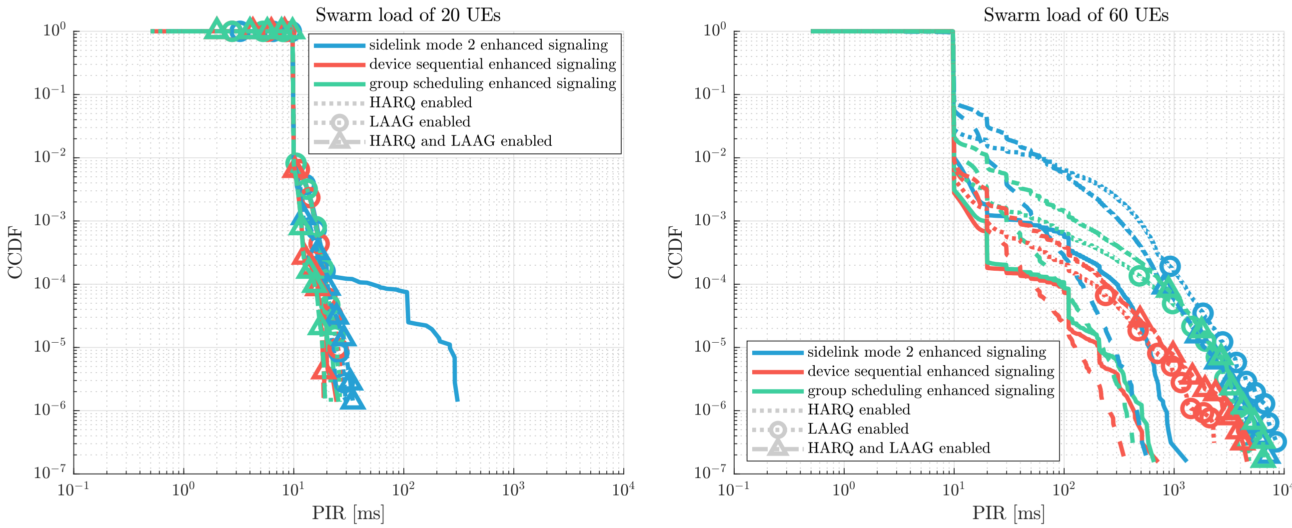

Fig. 6. Packet inter-reception for swarm loads of 20 and 60 UEs for each of the resource allocation schemes for: error-prone signaling, signaling with HARQ, signaling with LAAG and signaling with HARQ and LAAG enabled

the mitigation techniques at low occupancy. At high swarm load, LAAG has the strongest impact on resource occupancy. We observe that at high occupancy, LAAG causes sidelink mode 2 to allocate more resources than the cooperative schemes.

The transmission failure probability as a function of swarm load obtained in simulations is shown in Figure 5. The transmission failure probability is the probability that a single transmission of a 100 kbit data message cannot be delivered within the 10 ms latency requirement. With the enhanced error-prone signaling alone we observe how the transmission failure probability, and hence the reliability of both cooperative schemes, are consistently better than the baseline sidelink mode 2. With increasing swarm load the transmission failure probability increases for all schemes, and the gap between the cooperative schemes and sidelink mode 2 remains. Between the cooperative schemes, device sequential has slightly lower transmission failure probability at lower swarm loads, this advantage remains as swarm load increases. The plot of the device sequential scheme with HARQ and all schemes with HARQ and LAAG enabled starts at swarm load of 30 and 20 devices respectively due to the logarithmic y-axis and absence of errors at lower device loads.

Enabling HARQ is advantageous at low swarm load for all schemes. Interestingly, the relative performance gain is lowest for the group scheduling scheme. The likely explanation for this behavior is failure of reception of RMs from the leader due to the reduced number of scheduled Tx slots, which deprives the transmission from a group member and cause HARQ to be redundant to the RM re-transmission. For the sidelink mode 2 and group scheduling schemes, the four nines reliability requirement can be met until swarm load of 20 and 30 UEs respectively. Device sequential performs better, exceeding the reliability target for an additional 20 UEs. HARQ is the best configuration until swarm loads of 30 UEs for the device sequential scheme. At high occupancy the HARQ technique is detrimental to the transmission failure probability.

At the lowest swarm loads LAAG has negligible impact on the cooperative schemes, but reduces the transmission failure probability of sidelink mode 2 to the level of the cooperative schemes. Interestingly, the LAAG seems to dominate the impact on transmission failure probability, leaving all resource allocation schemes at the same performance until swarm load of 30 UEs. This can be explained by the fact that aggregated resources are allocated autonomously regardless of resource allocation scheme and when MCS is lowered, the autonomously selected resources will quickly dominate the resources selected by the resource allocation schemes. The difference is seen at higher swarm loads, where sidelink mode 2 transmission failure probability drastically increases at 50 UEs swarm load and the failure probability of the cooperative schemes increases at 60 UEs swarm load.

With HARQ and LAAG enabled simultaneously, the group scheduling sees an additional reduction in transmission failure probability at low swarm loads and gives the best performance for the scheme until swarm load of 40 UEs. For device sequential the combination of HARQ and LAAG is the best configuration at swarm loads of 40 and 50 UEs. For sidelink mode 2 the performance at low loads is ambiguous as the failure probability increases from 10 to 20 UEs but then drops from 20 to 30 UEs. Still, the four nines reliability target is only met at 10 UEs swarm load. The complementary effect at low swarm loads turns into a destructive effect a higher swarm loads where the combination of the two techniques perform worse than either of the techniques alone. The destructive effect can be explained by the reduction in scheduling Tx slots imposed by the HARQ, which limits LAAG and quickly saturates resources with transmissions leaving little time for reception.

PIR for two representative swarm loads is depicted in Figure 6. As expected, the majority of PIR is exactly at one data period of 10 ms. Consistently, it is the configuration at the given swarm load with the highest transmission failure probability (in Figure 5) which also experience the longest tail (in Figure 6). This implies that when failures are introduced, they are likely to happen persistently in some UEs rather than sporadically. This is contrary to the conclusion of [12] where they show that transmission failure probability and PIR are only weakly correlated for random transmissions. Our observation is likely coupled with the SPS transmissions, and hence higher determinism in our scenario. At all swarm loads it is a variant of the sidelink mode 2 scheme which exhibits the highest transmission failure probability. Between the cooperative schemes, device sequential experiences both the lowest transmission failure probability and lowest PIR.

V. Conclusion

With the techniques of HARQ and link adaptation by aggregation (LAAG) we are able to reach the four nines reliability at more than twice the load of what is achievable with the sidelink mode 2 when using our proposed cooperative resource allocation schemes.

HARQ has the greatest impact at low swarm loads where resource occupancy is low. LAAG is more helpful at higher swarm loads. When resource occupancy becomes excessive, the techniques do not improve reliability and thus should be enabled in dependence of load.

The best PIR is coupled with the combination of techniques at a given load. At 20 devices, enabling both HARQ and LAAG has lowest transmission failure probability and also lowest PIR. At 60 devices the HARQ technique results in lowest transmission failure probability and shortest PIR. The growth of the PIR tails is correlated with the transmission failure probability, implying that an increased transmission failure probability is caused by additional successive failures in a subset of UEs rather than evenly across all UEs.

References

[1] R. Bruun *et al.*, “Signaling design for cooperative resource allocation and its impact to reliability,” arXiv:2109.07206 [cs.NI].

[2] 3rd Generation Partnership Project (3GPP), “Study on enhancement of 3gpp support for 5g v2x services,” in *3GPP TR 22.886 V16.2.0*, Dec. 2018.

[3] G. Jornod, A. E. Assaad, and T. Kurner, “Packet Inter-Reception Time Conditional Density Estimation Based on Surrounding Traffic Distribution,” *IEEE Open Journal of Intelligent Transportation Systems*, vol. 1, pp. 51–62, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9094651/

[4] S. Morejon *et al.*, “Cooperative Resource Allocation for proximity Communication in Robotic Swarms in an indoor Factory,” in *2021 IEEE Wireless Communications and Networking Conference (WCNC2021)*. Nanjing, China: IEEE, Mar. 2021.

[5] 3GPP, “TR 37.885 V15.3.0 study on evaluation methodology of new Vehicle-to-Everything (v2x) use cases for lte and nr,” Jun. 2019.

[6] M. G. Sarret *et al.*, “Dynamic Outer Loop Link Adaptation for the 5G Centimeter-Wave Concept,” p. 6, 2015.

[7] K. Yao *et al.*, “Self-Organizing Slot Access for Neighboring Cooperation in UAV Swarms,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2800–2812, Apr. 2020, conference Name: IEEE Transactions on Wireless Communications.

[8] 3rd Generation Partnership Project (3GPP), “Radio resource control (rrc) protocol specification (release 16),” in *3GPP TR 38.331 V16.4.1*, Dec. 2021.

[9] J. Andrews, *Fundamentals of WiMAX : understanding broadband wireless networking.* Harlow: Prentice Hall, 2007.

[10] 3rd Generation Partnership Project (3GPP), “Study on channel model for frequencies from 0.5 to 100 ghz,” in *3GPP TR 38.901 V16.1.0*, Dec. 2019.

[11] S. Lu, J. May, and R. J. Haines, “Efficient modeling of correlated shadow fading in dense wireless multi-hop networks,” in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*. Istanbul, Turkey: IEEE, Apr. 2014, pp. 311–316.

[12] M. E. Renda, “IEEE 802.11p VANets: Experimental evaluation of packet inter-reception time,” *Computer Communications*, p. 13, 2016.