

Prioritizing Critical Access in LTE Random Access Procedure Supporting ITS Applications

Ruki Harwahyu¹, Jonathan Loo², Ray-Guang Cheng^{*1}, Yu-Yi Chu¹

1) National Taiwan University of Science and Technology, Taiwan

2) Middlesex University, London, UK

* Corresponding author

LTE-based vehicular to everything (V2X) service is one of promising technologies to support intelligent transportation systems (ITS). LTE adopts a request-grant approach to allocate wireless resource. Each user should follow a contention-based random access (RA) procedure to transmit its request. However, the contention-based RA procedure may result in unbounded delay if too many users compete for limited RA resource. Prioritized critical access in LTE RA procedure is required to differentiate critical V2X services from non-critical services. This work presents a generalized preamble allocation strategy that can be used to offer prioritized channel access for V2X service through flexible preamble allocation. Simulation result shows that the proposed scheme can prioritize the critical ITS service and fulfill its stringent-timing and reliability requirement.

Keyword: ITS, V2X, LTE, random access

1. Introduction

The 3rd Generation Partnership Project (3GPP) is investigating possible enhancements of Long Term Evolution (LTE) to support vehicular to everything (V2X) service. The LTE-based V2X service includes vehicle-to-vehicle (V2V), vehicle-to-pedestrian (V2P), and vehicle-to-infrastructure/network (V2I/N). V2V covers LTE-based communication between vehicles; V2P refers to LTE-based communication between a vehicle and a device carried by a pedestrian, cyclist, driver or passenger; while V2I/N denotes LTE-based communication between a vehicle and a roadside unit/network [1].

The V2X service lays the technicalities to realize Intelligent Transportation Systems (ITS). Many improved services are meant to be built atop it such as road management, efficiency and safety, positioning, parking and roadside assistance, car engine diagnostics and telematics, self-driving car, as well as on-board infotainment (human-centric information and entertainment).

From wireless network perspective, V2X service consists of nodes with high mobility which operates under unreliable channel conditions. It bears various network traffic patterns and quality-of-service (QoS) requirements. For ITS, some V2X

services may need a dedicated connection for delivering periodically generated data (e.g. traffic reporting messages). The other V2X services may only stay in idle mode and transmit infrequent small data whenever needed. For these idle-mode-based V2X services, some of them may require an ultra-reliable and low latency channel to deliver the unexpected emergency alarm message such as collision warning. However, these critical V2X services may have to compete for limited RA resource with the other services.

At the end of 2015, 11% of all data in the network are machine-generated. By 2020, the number is expected to rise up to 40% with about 40 billion connected devices [2]. Although there is no specific report of how much V2X service would contribute to this number, it probably would be significantly as the road is much of our third home after our house and office. Driven by the huge estimated market demand, several notable standards have been proposed.

In 2003, ATSM E2213-03 was approved for ITS standard which specifies application, data link, and physical layers. In 2010, Institute of Electrical and Electronics Engineers (IEEE) 802.11p was incorporated to Wireless Access in Vehicular Environment (WAVE) protocol stack, which is the continuation of ATSM E2213-03. IEEE 802.11p is based on IEEE 802.11 family. It employs Enhanced Distributed Channel Access (EDCA) as the channel access mechanism and uses dedicated short-range communication (DSRC) to form ad hoc connection. Another framework called Communication Access for Land Mobiles (CALM) was introduced which suggests using all access technologies such as 2G/3G/4G, WiMAX, WiFi, Infrared, and IEEE 802.11p to provide seamless connection. Those access technologies are knitted together by IPv6 or CALM-non-IP protocol [3].

As the prominent technology for ITS, IEEE 802.11p has easy deployment, low cost, mature technology, and native V2V and V2I supports with Road Side Units (RSUs). Nonetheless, it suffers from scalability issues, unbounded delays, and lack of deterministic QoS guarantee [4]. Furthermore, due to limited radio range and without pervasive roadside communication infrastructure, IEEE 802.11p only offers intermittent V2I connectivity. These limitations motivate the investigation of LTE network for alternative [5].

In LTE, ITS service can be regarded as machine-type communication (MTC) and the nodes are referred as MTC user equipment (UE). LTE relies on cellular infrastructure offering wide area coverage. This would eliminate the poor and intermittent connectivity as well as network fragmentation for V2V communication with low car density or poor signal propagation conditions. LTE would also enable V2I communications at high car speeds compared to IEEE 802.11p [6]. In addition, LTE is expected to penetrate the market faster than IEEE 802.11p as it is widely adopted in consumer devices and steadily invested by telecom operators. Moreover, LTE has the capacity to support several vehicular nodes per cell. The Rel. 11 LTE-A offers up to 1 Gbps data rate, which is much higher than the 27-Mbps of IEEE 802.11p [5].

There are several challenges for LTE to support vehicular applications. LTE utilizes centralized architecture which may increase latency for V2V communication [5]. LTE uses ALOHA-based access mechanism, which suffers from low access success probability and high mean access delay in presence of higher network load. Furthermore, LTE also serve H2H service. The H2H service may be severely disrupted during massive access from ITS and MTC services in general, which is likely to happen in emergency situation such as road accident.

The development of efficient RA protocols is a challenging task for beyond 4G networks. Vinel, *et al.* [7] proposed an analytical model to derive upper and lower bounds of the capacity for a reservation-based RA system. A distributed queueing random access algorithm has been presented to enhance the performance of the RA system by dividing the collided stations into smaller collision sets [8]. In LTE, all UEs must conduct RA procedure to obtain medium access for data transmission. Figure 1 shows the network traffic burst due to emergency situation in LTE serving ITS. Massive and sporadic access request may cause heavy congestion during RA. It brings low system throughput, higher mean access delay, and poor resource utilization for all services in the same cell. Therefore, only small number of UEs can proceed to the next signalling process before eventually transmit its data. In this case, QoS mechanisms in later signalling stages are less effective.

To support critical ITS service, it is necessary to implement service differentiation and prioritization in RA procedure, which is addressed as QoS-aware RA procedure herein. QoS-aware RA procedure aims to ensure higher access success probability and lower mean access delay for prioritized services, such as safety applications, while giving reasonable access success probability and mean access delay for lower priority services, such as telemetry or infotainment. QoS-aware RA procedure is built on top of the same basic RA procedure defined in LTE. However, it further adopts a preamble allocation strategy to allocate limited preambles to various V2X services by considering their different traffic patterns and QoS requirements.

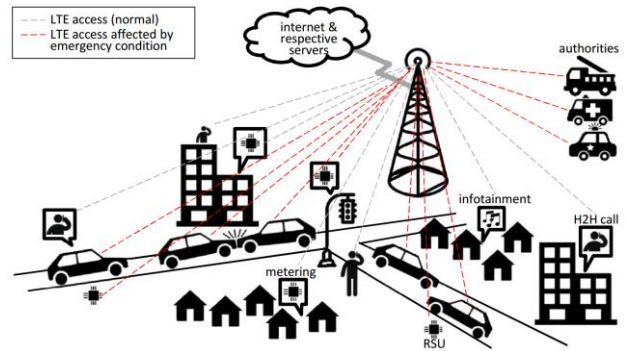


Figure 1 Network traffic burst due to emergency situation in LTE serving ITS.

This paper proposes a generalized preamble allocation strategy to realize the QoS-aware RA procedure. This proposal aims to avoid the said drawbacks for critical ITS services, or other higher-priority services, during high load. This proposal uses an allocation matrix to allocate each preamble independently toward one or more services. Hence, it is flexible to apply various access prioritizations toward multiple services coexisting in the cell.

2. LTE Random Access (RA) Procedure

The RA Procedure is one of the initial signaling steps in LTE network. This section elaborates the contention-based RA procedure, which is performed for initial access from *idle* mode, for RRC Connection Re-establishment, and upon UL data arrival in *connected* mode.

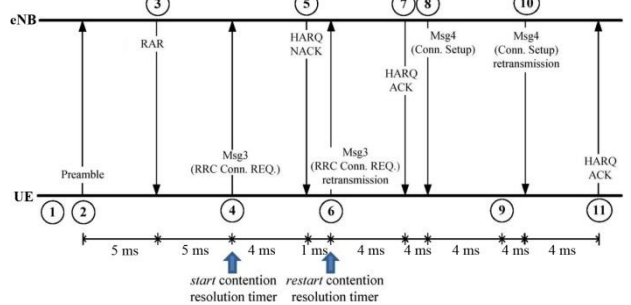


Figure 2 Basic Random Access Procedure in LTE [14]

Figure 2 illustrates the steps and message exchanges in contention-based RA procedure. An UE should transmit a randomly-chosen preamble (Step (1) in Figure 2) from a set of pre-allocated preambles by an eNB through a random access opportunity (RAO). Collision may happen when more than 1 UE transmit the same preamble at the same RAO. This collision is detected by the involved UEs after the transmission of Msg3. eNB broadcast random access response (RAR) in response of the received preambles (3). UE whose preamble is indicated in RAR will transmit Msg3 conveying UE's ID (4). UE whose preamble is not indicated in RAR should perform backoff and retransmit a new preamble. eNB replies the successfully received Msg3 with hybrid automatic repeat

request (HARQ ACK (7) and Msg4 (8). The UE it replies with HARQ ACK (11) if it finds its ID in Msg4. eNB replies a HARQ NACK (5) if Msg3 is collided and unrecovered. It also triggers retransmission of Msg3 (6). If ACK from the intended UE is not received by eNB (9), it retransmits Msg4 (10). Transmission of Msg3 and Msg4 can be conducted until each of them meets its HARQ attempt limit.

Related works

There are several works proposing QoS-aware RA procedure. Several schemes are studied in [9] which consider 2 ACs: H2H and MTC. The schemes are: (i) access class barring, which delays access by MTC UEs based on a barring factor, (ii) separated preambles for H2H and MTC UEs, (iii) dynamic allocation of preambles when network load is predictable, (iv) differentiated backoff scheme for MTC UEs, (v) separated access period for MTC and H2H UEs, and (vi) pull-based schemes, which allows eNB to fully control access attempt of each UE.

Several works have proposed preamble-allocation-based prioritization, which try to manage the allocation of preamble toward multiple ACs to improve the system. Proposal in [10] pre-allocates preambles for different ACs and uses service-dependent backoff procedure along with dynamic access barring. Schemes in [11] prioritize H2H over MTC service by separated preamble groups and assess two scenarios: (i) one group is for H2H and the other for MTC, (ii) one group is for H2H only whereas the other is for both H2H and M2M. The work in [12] extends [11] with simulation study for partially-shared preamble allocation with more than two ACs. In [13], preambles are divided into three groups: contention-free, contention-based, and priority group. With fixed number of contention-free preambles, number of contention-based and priority preambles are adjusted according to some predefined load thresholds to achieve higher access success probability for both contention-free and contention-based UEs.

Different from the existing studies, this work proposes a generalized preamble allocation strategy to deal with prioritization and allocation optimization for various network patterns and QoS requirements.

3. Generalized preamble allocation

The preamble allocation is a tradeoff among the performance of different ACs. The performance of high-priority AC can be enhanced by allocating more preambles, which comes at the cost of degraded performance in lower priorities ACs. The preamble allocation patterns can be simply divided into four categories: *fully-overlap*, *partial fully-overlap*, *partial-overlap*, and *non-overlap* allocations. In *fully-overlap* allocation, each preamble is shared by all ACs as in standard LTE. In contrast, each preamble is dedicated allocated to only one AC in *non-overlap* allocation [11]. In *partial fully-overlap* allocation, the preambles allocated to one AC are also used by higher priority ACs [11]. The *partial-overlap* allocation is the

most general case in which each preamble can either be dedicated allocated to one AC or be shared by two or more ACs. An example illustrating the four patterns is shown in Figure 3. In this figure, ten preambles are allocated to two ACs and '1' denotes the preamble is allocated to an AC.

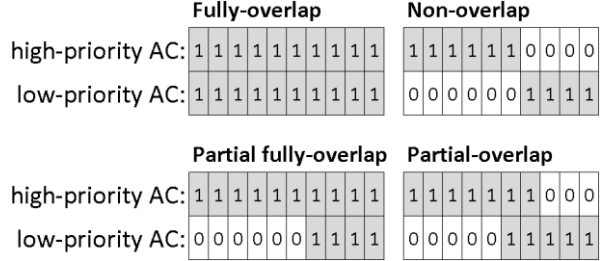


Figure 3 Various preamble allocation patterns

In LTE, the access success probability (i.e., reliability) and the mean access delay (i.e., latency) that a user experience during the RA procedure mainly depend on the number of contending users and number of pre-allocated preambles. The prioritization among different services may be achieved by allocating preambles to each service. The eNB can easily ensure the ultra-reliability and low latency requirement of critical V2X services by reserving all preambles to them. However, reserving all preambles to these critical V2X services also block the opportunity for remaining services to use the LTE network. The principle of preamble-allocation-based is to divide and lower the contention to achieve better performance.

The proposed generalized preamble allocation scheme allows each preamble to be assigned toward one or more ACs independently. In the implementation, the preamble allocation specified in this scheme is then broadcasted by eNB to the UEs in each AC via System Information Block 2 (SIB2) to be used by the UEs in preamble selection (Step (1) in Figure 2). The priority of the AC is defined by the network operator. Individual UE needs to specify a proper AC based on its QoS requirements during service subscription. Consider a cell where eNB reserves R preambles in each RAO for K number of ACs, with a critical ITS service being one of the ACs. A K -by- R preamble allocation matrix X is defined to represent the allocation of R preambles to K ACs as

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,r} & \cdots & x_{1,R} \\ x_{2,1} & x_{2,2} & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{k,1} & x_{k,2} & \cdots & x_{k,r} & \cdots & x_{k,R} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{K,1} & x_{K,2} & \cdots & x_{K,r} & \cdots & x_{K,R} \end{bmatrix}. \quad (1)$$

Element in k th row and r th column, $x_{k,r}$ is set to 1 if the r th preamble is allocated to the k th AC, or 0 otherwise.

The allocation matrix is set based on the estimated peak loading and target QoS requirements of high-priority ACs. The designer has to determine an allocation pattern according to the

arrival process and QoS requirement of the ACs. The fully-overlapped allocation is adopted for network with ACs with unpredictable offered load or similar QoS requirements. In this case, it is efficient for all ACs to share the same pool of preambles. However, it relies on extra overload control scheme to maintain the QoS requirement of high-priority AC(s). Partial-overlap and partial fully-overlap allocations are selected if the high-priority AC(s) with massive and sporadic access nature and high QoS requirement and low-priority AC(s) with relatively stable offered load. In this case, the network may reserve preambles based on the estimated peak load of the ACs to ensure the instantaneous QoS requirement for high-priority AC(s). The unused preambles can also be shared to low-priority AC(s). The amount of preambles to be shared to low-priority AC(s), however, depends on the importance of the high-priority AC(s). The non-overlap allocation is suitable for high-priority AC(s) with stable offered load. In this case, dedicated preambles are allocated to ensure the QoS requirement. With a selected allocation pattern, preambles are first assigned to the highest priority AC under consideration based on its estimated peak loading and QoS requirement.

With a given allocation matrix, matrix permutation is performed to reorder the preambles and divide the preambles into one or more contention domains. A contention domain is a non-empty set of preamble(s) which are allocated to the same set of AC(s). The contention domain can be identified based on matrix X. Take a matrix X in Eq. (2) as an example,

$$X = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}. \quad (2)$$

In this example, preamble 1, shared by AC1 and AC3, forms the first contention domain; preambles 2 and 3, which are shared by AC1 and AC2, form the second contention domain; and preambles 4 and 5, which are shared by AC2 and AC3, form the third contention domain.

In RA procedure, each UE shall randomly select one preamble from the preamble set allocated to its AC. Hence, for a specific AC, the percentage of UEs which join a contention domain is equal to the ratio between number of preambles in the contention domain and total number of preambles allocated for the AC. For each contention domain, we can estimate the loading contributed by each AC (i.e., based on percentage of preambles allocated to the contention domain) and estimate the performance of the AC based on the simple formula presented in [12]. The overall performance of each AC can then be derived as a weighted sum of the performance of the AC experienced in each contention domain.

4. Result and Discussion

The proposed scheme is evaluated via a C-based event-driven LTE RA simulator developed in [14]. Three scenarios are considered in the simulation to demonstrate effect of preamble allocation on the requirements of cooperative ITS

application. Scenarios 1 and 2 illustrate the trade-off among four allocation patterns and their effect in prioritizing services with various network patterns during the RA procedure. Scenario 3 shows a way to optimize preamble allocation for achieving the QoS of critical ITS service.

The RA parameters defined in Table 6.2.2.1.1 in [9] and processing latency specified in Table B.1.1.1-1 in [15] are used in all scenarios. The processing latency used in this paper include: eNB processing delay between the end of preamble reception and the end of RAR transmission is 5 ms; UE processing delay between the end of RAR reception and Msg3 transmission is 5 ms; MAC contention resolution timer is 48 ms; eNB processing delay between the end of successful Msg3 reception and Msg4 transmission is 4 ms; HARQ feedback timeout is 4 ms; processing delay before HARQ retransmission is 1 ms.

Scenarios 1 and 2 compare results of the four allocation patterns (i.e., *fully-overlap/basic* LTE, *non-overlap*, *partial fully-overlap*, and *partial-overlap*) for two ACs with various network patterns. The RA parameters used in these two scenarios are: PRACH Configuration Index 6; backoff indicator is 20 ms; and maximum number of preamble transmission is 10 [9]. In both scenarios, M_1 ranges from 50 to 500, where 500 is the estimated peak load of AC1, and $T_1 = 50$ ms. In both scenarios, that the minimum number of preambles are allocated to AC1 to ensure target QoS requirement of $P_{S,1} \geq 90\%$, except for *fully-overlap* allocation.

Scenarios 1 is designed to let two ACs have identical arrival interval, which gives $M_2 = 250$ and $T_2 = 50$ ms. In this scenario, the preamble allocations are given below: 54 preambles are all shared by AC1 and AC2 in *fully-overlap* allocation; 38 and 16 preambles are exclusively allocated to AC1 and AC2, respectively, in *non-overlap* allocation; 54 preambles are allocated to AC1, and 29 of them are also allocated to AC2 in *partial fully-overlap* allocation; and preambles 1 to 50 are allocated to AC1 and preambles 29 to 54 are allocated to AC2 (i.e., preambles 29 to 50 are shared by AC1 and AC2) in *partial-overlap* allocation. The results of access success probability, $P_{S,k}$, and mean access delay, D_k , are shown in Figures 4(a) and 4(b), respectively. The horizontal dash line shown in Figure 4(a) represents the QoS requirement for AC1. It can be found in Figures 4(a) and 4(b) that AC1 and AC2 have identical access success probability and mean access delay in *fully-overlap* allocation since they share the same pool of preambles. As a result, it cannot satisfy the target QoS requirement of AC1. The results in Figure 4(a) show that the high access success probability of AC1 ($P_{S,1} \geq 90\%$) is ensured by sacrificing the performance of AC2 in *non-overlap* allocation ($P_{S,2} = 58.3\%$). The access success probability of AC2 will not be increased even if the loading of AC1 decreases. It is because that the unused preambles for AC1 cannot be shared with AC2. It can be found that $P_{S,2} = 23\%$ and $P_{S,2} = 67.7\%$ are achieved in *partial fully-overlap* allocation and *partial-overlap* allocation, respectively, under the peak load of

AC1. $P_{S,2}$ increases as the offered load of AC1 decreases in both patterns. From Figure 4(b), it can be found that the mean access delay for AC1 in *non-overlap* allocation is the lowest among the four patterns. The price paid is the highest mean access delay for AC2. The mean access delay of the two ACs in *partial fully-overlap* and *partial-overlap* allocation are quite similar, except that *partial-overlap* allocation obtains a slightly lower mean access delay for AC1 (D_1) but a higher mean access delay for AC2 (D_2) around the peak load of AC1.

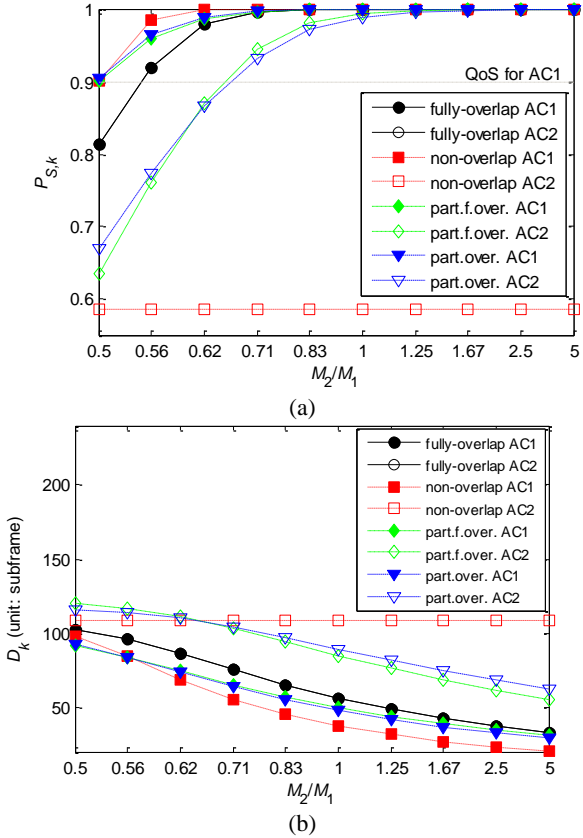


Figure 4 Results for Scenario 1, (a) access success probability, (b) mean access delay

Scenario 2 is designed to have two ACs with different arrival interval, which gives $M_2 = 500$ and $T_2 = 100$ ms. Note that only AC2's arrival duration is different between Scenarios 1 and 2. Hence, only *partial fully-overlap* and *partial-overlap* allocation need to be adjusted to meet the QoS requirement of AC1. The adjusted preamble allocation is given by: 54 preambles are allocated to AC1, and 23 of them are also allocated to AC2 in *partial fully-overlap* allocation; and preambles 1 to 48 are allocated to AC1 and preambles 33 to 54 are allocated to AC2 (i.e., preambles 33 to 48 are shared by AC1 and AC2) in *partial-overlap* allocation. The access success probability and mean access delay of the two ACs are shown in Figures 5(a) and 5(b), respectively. Different to Scenario 1, it can be found in Figure 5(a) that the success probabilities of AC1 and AC2 are not identical even in *fully-*

overlap allocation. The access success probability of low priority AC is higher than that of low priority ($P_{S,2} > P_{S,1}$) when the loading of AC1 is far below the estimated peak load (smaller values of M_2/M_1). It is because that the arrival intervals of AC1 and AC2 are totally overlapped in the first 50 ms and thus, experience a lower access success probability. In contrast, arrivals of AC2 during 50 ms to 100 ms always have a high access success probability. Hence, on average, the overall access success probability of AC2 is higher than that of AC1. The performance of AC2 is not as good as that in Scenario 1 in *non-overlap*, *partial fully-overlap* and *partial-overlap* allocations since we assign a higher offered load to AC 2.

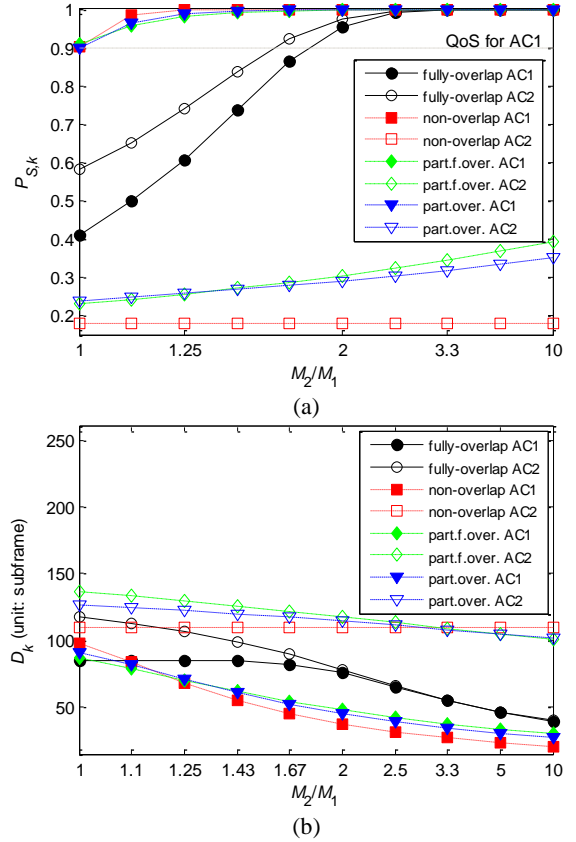


Figure 5 Results for Scenario 2 (a) access success probability, (b) mean access delay

Scenario 3 shows a way that we used to optimize the preamble allocation for achieving the QoS of critical ITS service. A special case of *partial fully-overlap* in [11] is optimized using the proposed scheme. To represent an LTE network which is specially designed for low-latency requirement of ITS, minimum PRACH interval of 1 ms (PRACH Configuration Index 14) and the minimum non-zero backoff indicator of 10 ms are used as the RA parameters. The QoS requirement considered in Scenario 3 is $P_{S,1} \geq 90\%$ and $D_1 \leq 40$ ms. As in Scenario 1, the arrival interval for both ACs, are restricted to the first 50 ms and $M_2 = 250$. M_1 ranging from 100 to 1600 is considered, where 1600 is the estimated peak load of AC1. Figure 6 shows the performance of the proposed

method (which is referred as Optimized in the figure) and compares it with that obtained from [11] (which is referred as Original on the figure). According to [11], 54 preambles are assigned to AC1 while 7 of them are shared with AC2. In the optimized allocation, 54 preambles are for AC1 while 8 of them are also for AC2. Figures 6(a) and 6(b) show the access success probability and mean access delay for each AC. It can be found that $P_{S,1}$ and D_1 in both schemes can meet their QoS requirements. However, the optimized allocation has a higher $P_{S,2}$ and a lower D_2 , which shows the effectiveness of the proposed method.

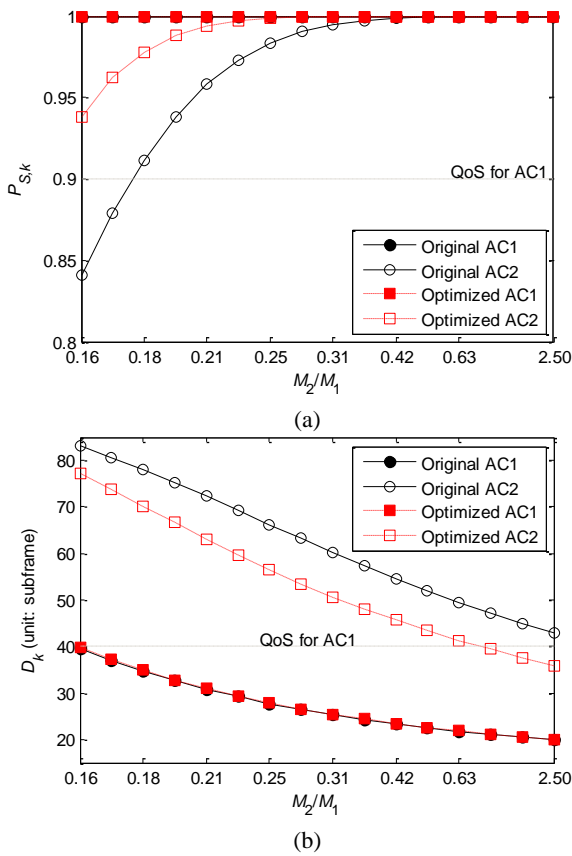


Figure 6 Results for Scenario 3 (a) access success probability, (b) mean access delay

5. Conclusion

Generalized preamble allocation is proposed to support delay-sensitive critical ITS service coexisting with other services in LTE cell, which is likely to be compromised in overload condition when no prioritization is applied. It provides a way to flexibly allocate preambles toward multiple services according to their priority. By properly tuning the preamble allocation matrix, the contention experienced by critical ITS service can be made lower to achieve higher access success probability and lower mean access delay.

This scheme can be applied to any services in general, and prioritizing several services at once when sufficient number of preambles is given. With the introduced flexibility, telecom operators and researchers can easily implement their access prioritization policy for multiple services coexisting in an LTE cell.

References

- [1] 3GPP TR 36.885, "Technical Specification Group Radio Access Network; Study on LTE-based V2X Services," v. 2.0.0, Jun., 2016.
- [2] Oracle, "The Internet of Things: Manage the Complexity, Seize the Opportunity," Redwood, 2015.
- [3] E. Uhlemann, "Introducing Connected Vehicles," *IEEE Veh. Technol. Mag.*, vol. 10, no. 1, pp. 23-32, March 2015.
- [4] M. Amadeo *et al.*, "Enhancing IEEE 802.11p/WAVE to Provide Infotainment Applications in VANETs," *Elsevier Ad Hoc Networks*, vol. 10, no. 2, pp. 253-69, Mar., 2012.
- [5] G. Araniti *et al.*, "LTE for vehicular networking: a survey," *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 148-157, May 2013.
- [6] Z. H. Mir and F. Filali, "LTE and IEEE 802.11p for vehicular networking: a performance evaluation," *EURASIP J. on Wireless Commun. and Networking*, vol. 2014, no. 89, pp. 1-15, 2014.
- [7] A. Vinel *et al.*, "Capacity analysis of reservation-based random access for broadband wireless access networks," *IEEE J. on Selected Areas in Commun.*, vol. 27, no. 2, pp. 172-181, 2009.
- [8] V. Kobliakov *et al.*, "Distributed queue random multiple access algorithm for centralized data networks," *Proc. IEEE International Symposium on Consumer Electronics*, pp. 1-6, 2006.
- [9] 3GPP TR 37.868, "RAN improvements for machine-type communication," v. 11.0.0, Sep., 2011.
- [10] C. H. Lee, T. M. Lin and J. P. Cheng, "Prioritized random access with dynamic access barring for RAN overload in 3GPP LTE-A networks," *Proc. IEEE GLOBECOM*, pp. 368-372, Dec., 2011.
- [11] S. Kim, B. Yi and K. D. Lee, "Throughput comparison of random access methods for M2M service over LTE networks," *Proc. IEEE GLOBECOM*, pp. 373-377, Dec., 2011.
- [12] Y. Y. Chu *et al.*, "Study of generalized resource allocation scheme for multichannel slotted ALOHA systems," *Proc. IEEE PIMRC*, pp. 1702-1706, 2015.
- [13] X. L. Li, Q. N. Ren and N. Hu, "Random access preamble assignment algorithm of TD-LTE," *Advan. Comput., Commun., Control Autom., Lec. Notes in Electr. Eng.*, vol. 121, pp. 701-708, 2012.
- [14] C. H. Wei, R. G. Cheng, S. L. Tsao, "Performance analysis of group paging for machine-type communications in LTE networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3371-3382, 2013.
- [15] 3GPP TR 36.912, "Feasibility study for further advancements for E-UTRA (LTE-Advanced)," v. 13.0.0, Jan., 2016.

Acknowledgements

This work was supported in part by the National Science Council, Taiwan under contracts MOST 102-2219-E-011-001 and 102-2221-E-011-003-MY3.