

Random Access for Network Slicing

Linna Chen, Chunjing Hu, Yong Li

Wireless Signal Processing and Network Laboratory
Key Laboratory of Universal Wireless Communications, Ministry of Education,
Beijing University of Posts and Telecommunications
Beijing, China 100876
chenlinna@bupt.edu.cn

ABSTRACT

Network slicing is an emerging paradigm for the architecture of the upcoming 5G mobile networks. One of the features enabled by network slicing is the isolation between different network slices which can guarantee that one slice would not be interrupted by congestion in other slices. Although data channel isolation has been investigated in the literature, the technique for realizing the isolation of control channel has not been discussed adequately. In this paper, the isolation of uplink random access (RA) channel will be studied, and a novel scheme will be proposed to support efficient network slicing for RA procedures, in which the different requirements of multiple slices and the characteristics of their traffic are considered. The success probability of the proposed scheme is analytically analyzed and simulated, which shows that, compared to the conventional random access scheme in which all users share the same set of RA preambles and the physical isolation scheme which assigns orthogonal RA preamble subsets to different slices, the proposed scheme can achieve efficient isolation between slices as well as efficient resource utility.

KEYWORDS

Network slicing, preamble, random access, URLLC, eMBB

ACM Reference format:

Linna Chen, Chunjing Hu, Yong Li. 2017. Random Access for Network Slicing. In *Proceedings of ACM Mobimedia conference, Chongqing, China, July 2017 (MOBIMEDIA'17)*, 6 pages. DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Instead of traditional one-size-fits-all network architecture, 5G systems will be built to serve the emerging wide range of use cases with different characteristics and requirements [1]. Some of the most mentioned use cases are enhanced mobile broadband (eMBB), massive machine type communication (mMTC), and ultra-reliable and low-latency communication (URLLC). Network slicing is a new paradigm for the next

generation network architecture which can deploy multiple independent logical networks over a single physical substrate. Operators and service providers are allowed to create a dedicated network slice instance customized to meet diverse requirements for their subscribers [2].

There are three important features that should be enabled by network slicing, which are customization, isolation and utilization. Different slices may provide different services with its specific goals, therefore flexible and appropriate policies should be offered for each network slice to achieve their customized performance requirements. Network slicing enables multiple slices to share the same physical resource in a flexible, dynamic manner, which can utilize resource efficiently and reduce the capital expenditure and operational expenditure [3, 4]. However, if the shared common resource is not managed properly, the unexpected events within one slice, e.g., overload or congestion, may impose a negative effect on the performance of other slices. Therefore, isolation among multiple slices is one of important requirements that needs to be ensured in order to handle slices in a robust way [5]. Dedicated resource partition between multiple slices may be a simple solution and it can achieve perfect isolation, but the waste of scarce resource is inevitable due to unpredictable traffic demands of different slices. Therefore, there is a need to make a reasonable tradeoff between two conflicting goals of isolation and efficient resource utilization [6].

The random access (RA) procedure will be used to establish the uplink synchronization and request scheduling resources when no dedicated resources are available for the user, which has a great influence on the performance of the system as well as the quality of experience of the user [7]. Therefore, random access has become one important research of the communication system, and the isolation of random access channel (RACH) resource between different slices also needs to be considered. When all available random access resource is shared between multiple slices, the event caused by the increase of users in one slice should not impose negative effect on the random access performance of users in other slices. For example, if one slice is being overloaded at the time of RA procedure, other slices should still be accessible via RACH.

Isolation of RACH resource can be fully achieved by hard splitting the physical RACH resources among different slices, which can ensure that there is no impact toward each other during the RA procedure. However, it will lead to unefficient utilization of the limited resource. Thus, we propose a novel scheme to support efficient network slicing for RA procedure

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
MOBIMEDIA'17, Chongqing, China
© 2017 Copyright held by the owner/author(s). 123-4567-24-567/08/06...\$15.00
DOI: 10.1145/nnnnnnn.nnnnnnn

in order to achieve satisfactory isolation between slices as well as efficient resource utilization, and the customized performance requirements can also be considered by the proposed scheme. We evaluate our proposed scheme and other two traditional schemes through mathematical analysis, which is then verified to be consistent with the simulation. Finally the simulation results show that the proposed scheme can make a better tradeoff between isolation and utilization compared to other schemes.

The rest of the paper is organized as follows. In Section II, we describe the system model for the RA procedure. In Section III, we present the detailed description of the novel scheme and two baseline schemes, as well as their corresponding mathematical analysis. In Section IV, simulation results are shown and discussed. Finally, conclusions are summarized in Section V.

2 SYSTEM MODEL

Although there are two different random access modes, i.e., contention and non-contention based random access procedures in the radio access network. In this paper, we specifically consider the four-step contention-based random access procedure [8]. For brevity and without loss of generality, we consider in this paper two network slices with distinct requirements on reliability and latency. One slice serves URLLC users and thus requires high reliability and low latency, therefore it is essential to reduce the latency of the random access procedure. While the other slice is created to provide eMBB service in a best-effort manner which is not demanding in reliability and latency.

The total number of available preambles used for contention-based RA procedure in the network is denoted by N . In each slot, the numbers of users making RA attempts in the URLLC slice and the eMBB slices follow Poisson distribution and are denoted by X and Y , respectively, and the average number of RA attempts in the two slices are denoted by λ_1 and λ_2 , respectively. That is, we have $X \sim \Pi(\lambda_1)$ and $Y \sim \Pi(\lambda_2)$. It is assumed that RA procedure of a given user is successful if there are no other users selecting the same preamble with the given user on the same PRACH. The overall RA success probability of each network slice is defined as the ratio of the total number of successful RA events over the total number of RA attempts.

3 RANDOM ACCESS SCHEMES

3.1 Physical Sharing

The current 4G LTE network architecture has not adopted the concept of network slicing yet. Every user of different services can therefore randomly select one RA preamble from the entire set of preambles, which is referred to as physical sharing (PS) scheme in this paper. Specifically, in our exemplary network, a user in either the URLLC or the eMBB slice can randomly select one out of the N preambles with equal probability for a RA attempt.

Conditioned on the assumption that $X = x$ users in the URLLC slice and $Y = y$ users in the eMBB slice are making

their RA attempts simultaneously in a random access slot, the probability that a given user in the URLLC slice can select a preamble without collision with the other $x + y - 1$ users can be written as

$$P_{URLLC}^{\text{PS}}|_{X=x, Y=y} = \left(1 - \frac{1}{N}\right)^{x+y-1}, \quad (1)$$

which means that all the other $x + y - 1$ users can not choose the preamble which is chosen by the given user in the URLLC slice. As X and Y are both Poisson distributed, the average probability of RA success can be given by

$$\begin{aligned} P_{URLLC}^{\text{PS}} &= \frac{\sum_{x=1}^{\infty} \sum_{y=0}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \frac{\lambda_2^y}{y!} e^{-\lambda_2} \cdot x \cdot \left(1 - \frac{1}{N}\right)^{x+y-1}}{\sum_{x=1}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \cdot x} \\ &= \sum_{x=1}^{\infty} \sum_{y=0}^{\infty} \frac{\lambda_1^{x-1}}{(x-1)!} e^{-\lambda_1} \frac{\lambda_2^y}{y!} e^{-\lambda_2} \cdot \left(1 - \frac{1}{N}\right)^{x+y-1} \\ &= \sum_{x=0}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \left(1 - \frac{1}{N}\right)^x \cdot \sum_{y=0}^{\infty} \frac{\lambda_2^y}{y!} e^{-\lambda_2} \left(1 - \frac{1}{N}\right)^y \\ &= \left. \frac{\left(\frac{\lambda_1}{N}\right)^x}{x!} e^{-\frac{\lambda_1}{N}} \right|_{x=0} \cdot \left. \frac{\left(\frac{\lambda_2}{N}\right)^y}{y!} e^{-\frac{\lambda_2}{N}} \right|_{y=0} \\ &= e^{-\frac{\lambda_1 + \lambda_2}{N}}. \end{aligned} \quad (2)$$

Similarly, for the eMBB slice, we also have

$$P_{eMBB}^{\text{PS}} = e^{-\frac{\lambda_1 + \lambda_2}{N}}. \quad (3)$$

It is worth noting from (2) and (3) that, although users in the two slices have different RA rates, their success probabilities are identical, which implies that there is no any isolation effect between the two slices. The RA congestion from one slice will inevitably lower RA success probability of the other slice. Moreover, both slices cannot be offered customized RA success probability to achieve their goals.

3.2 Physical Isolation

A straightforward means to provide isolation of RA procedure is to assign orthogonal subsets of preambles to different slices, which should take their specific performance requirements and characters of traffics into consideration. For this purpose, we assume that N_1 and N_2 preambles are allocated for the URLLC and the eMBB slices in our exemplary network, respectively, with $N_1 + N_2 = N$.

For the URLLC slice, conditioned on the assumption that $X = x$ users in the URLLC slice are making RA attempts simultaneously each of which randomly selects one out of N_1 preambles with equal probability, the probability that any given user within the x users can select a preamble without collision with other $x - 1$ users can be written as

$$P_{URLLC}^{\text{PI}}|_{X=x} = \left(1 - \frac{1}{N_1}\right)^{x-1}. \quad (4)$$

As X is Poisson distributed, the average probability of RA success can be given by

$$\begin{aligned}
P_{URLLC}^{\text{PI}} &= \frac{\sum_{x=1}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \cdot x \cdot \left(1 - \frac{1}{N_1}\right)^{x-1}}{\sum_{x=1}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \cdot x} \\
&= \frac{1}{\lambda_1} \sum_{x=1}^{\infty} \frac{\lambda_1^x}{(x-1)!} e^{-\lambda_1} \left(1 - \frac{1}{N_1}\right)^{x-1} \\
&= \sum_{x=0}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \left(1 - \frac{1}{N_1}\right)^x \\
&= e^{-\frac{\lambda_1}{N_1}}.
\end{aligned} \tag{5}$$

Similarly, for the eMBB slice, we have

$$P_{eMBB}^{\text{PI}} = e^{-\frac{\lambda_2}{N_2}}. \tag{6}$$

In this scheme, isolation and customization can be achieved absolutely by hard splitting all the available random access preambles among multiple slices. Before the RA procedure, the users should obtain the preamble partition information through the broadcast channel, which can increase the overhead of the common signalling. In addition, the preamble resources cannot be utilized efficiently in some cases.

3.3 Multi-preamble

In our proposed scheme, a user in the eMBB slice with lower requirement for reliability and latency is still able to use the entire set of N preambles as candidate. In contrast, a user in the URLLC slice with higher requirement for reliability and latency is allowed to issue multiple RA preambles in one slot, and a user can be successfully detected by the network as long as at least one preamble can be received without collision, which can be considered as a successful RA procedure.

Assuming that each of $X = x$ users in the URLLC slice randomly selects L out of N preambles with equal probability¹ and each of $Y = y$ users in the eMBB slice randomly selects one out of N preambles with equal probability, the probability that a given user in the URLLC slice can select at least one preamble without collision can be written as

$$P_{URLLC}^{\text{MP}}|_{X=x, Y=y} = 1 - \left(1 - \left(1 - \frac{1}{N}\right)^{L(x-1)+y}\right)^L. \tag{7}$$

¹More specifically, it is proposed that a URLLC user should randomly selects L preambles *independently*, which implies that one preamble may be likely to be selected multiple times by a URLLC user. If this happens, the actually number of preambles transmitted from this user will be less than L . We do not propose to let the URLLC user make reselection in this case because transmitting exactly L preambles from each URLLC user will undesirably increase block probability between them.

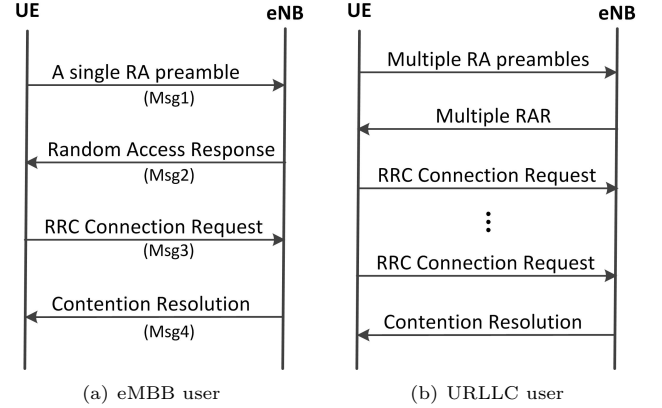


Figure 1: Random access procedure

As X and Y are both Poisson distributed, the average probability of RA success can be given by

$$\begin{aligned}
P_{URLLC}^{\text{MP}} &= \frac{\sum_{x=1}^{\infty} \sum_{y=0}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \frac{\lambda_2^y}{y!} e^{-\lambda_2} \cdot x \cdot \left(1 - \left(1 - \left(1 - \frac{1}{N}\right)^{L(x-1)+y}\right)^L\right)}{\sum_{x=1}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \cdot x} \\
&= 1 - \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \frac{\lambda_2^y}{y!} e^{-\lambda_2} \left(1 - \left(1 - \frac{1}{N}\right)^{Lx+y}\right)^L \\
&= 1 - \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} \sum_{l=0}^L \frac{\lambda_1^x}{x!} e^{-\lambda_1} \frac{\lambda_2^y}{y!} e^{-\lambda_2} C_L^l (-1)^l \left(\left(1 - \frac{1}{N}\right)^{Lx+y}\right)^l \\
&= 1 - \sum_{l=0}^L C_L^l (-1)^l e^{-\lambda_1 \left(1 - \left(1 - \frac{1}{N}\right)^{L}\right)^L - \lambda_2 \left(1 - \left(1 - \frac{1}{N}\right)^L\right)}.
\end{aligned} \tag{8}$$

The probability that a given user in the eMBB slice can select one preamble without collision can be written as

$$P_{eMBB}^{\text{MP}}|_{X=x, Y=y} = \left(1 - \frac{1}{N}\right)^{Lx+y-1}, \tag{9}$$

and hence the average probability of RA success for an eMBB user can be given by

$$\begin{aligned}
P_{eMBB}^{\text{MP}} &= \frac{\sum_{x=0}^{\infty} \sum_{y=1}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \frac{\lambda_2^y}{y!} e^{-\lambda_2} \cdot y \cdot \left(1 - \frac{1}{N}\right)^{Lx+y-1}}{\sum_{y=1}^{\infty} \frac{\lambda_2^y}{y!} e^{-\lambda_2} \cdot y} \\
&= \sum_{y=1}^{\infty} \sum_{x=0}^{\infty} \frac{\lambda_2^{y-1}}{(y-1)!} e^{-\lambda_2} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \cdot \left(1 - \frac{1}{N}\right)^{Lx+y-1} \\
&= \sum_{y=0}^{\infty} \frac{\lambda_2^y}{y!} e^{-\lambda_2} \left(1 - \frac{1}{N}\right)^y \cdot \sum_{x=0}^{\infty} \frac{\lambda_1^x}{x!} e^{-\lambda_1} \left(1 - \frac{1}{N}\right)^{Lx} \\
&= e^{-\frac{\lambda_2}{N}} \cdot e^{-\lambda_1 \left(1 - \left(1 - \frac{1}{N}\right)^L\right)}.
\end{aligned} \tag{10}$$

4 IMPLEMENTATION

In this section, we will introduce the RA implementation of the proposed scheme in detail. In our proposed scheme,

the number of RA preambles simultaneously selected by the URLLC user should be configured by the evolved NodeB (eNB) according to the actual needs. For the user in the eMBB slice, the RA procedure is initiated by randomly selecting a single RA preamble to the eNB, which is the same as the traditional way. The procedure is shown in Fig. 1(a) [8]. However, the URLLC user will perform a different RA procedure as depicted in Fig. 1(b).

First, the URLLC user randomly selects multiple RA preambles which are transmitted to the eNB. Once the random access preamble message (Msg1) is received successfully, the eNB will transmit a response message (Msg2). The Msg2 may contain different random access response (RAR) to each detected preamble. In the third step, the URLLC user shall, according to the uplink grant information in the response, transmit multiple RRC connection request messages (Msg3) on the corresponding resources. It is noticeable that, if multiple users select the same preamble on the same physical random access channel, they will transmit their RRC connection request on the same assignment resource and the collision will occur. Therefore, the UE identity should be provided in Msg3, and the eNB will respond with a contention resolution message (Msg4) to the user whose message is successfully received in the previous step [9]. The UE who detects its own identity echoed in the Msg4 can access the network successfully.

5 RESULTS AND DISCUSSIONS

Extensive simulation experiments have been performed to verify the analytical results in Section III, and it turns out that analytical and simulation results agree very well. In the following, we will consider a network configured with $N = 64$ RA preambles and present RA performance of different schemes.

In Fig. 2 and Fig. 3, it is assumed that the average RA rate in one slot is 1 user/slot for the URLLC slice, while the eMBB slice has much higher average RA rate at 10 users/slot. As a baseline, Fig. 2 illustrates the RA success probabilities of both slices under the physical isolation scheme. It shows that although assigning more preambles to the URLLC slice can improve its RA success probability but at a very limited rate. Even when 60 out of 64 preambles are assigned to the URLLC slice, its RA success probability is still below 99%, and in this case, however, the RA success probability of the eMBB slice falls below 10%. The performance of the eMBB slice is seriously affected by the number of preambles in its available group, which is not expected to happen. Therefore, it is unadvisable to allocate a large amount of preambles to improve the RA success probability of the URLLC slice at the cost of a sharp decline of the RA success probability of the other slice. And in practical situations, when the RA loads of network slices vary dynamically, the assignment information of random access preambles needs to be broadcast frequently in order to achieve better system performance, which can increase signaling overhead and energy consumption.

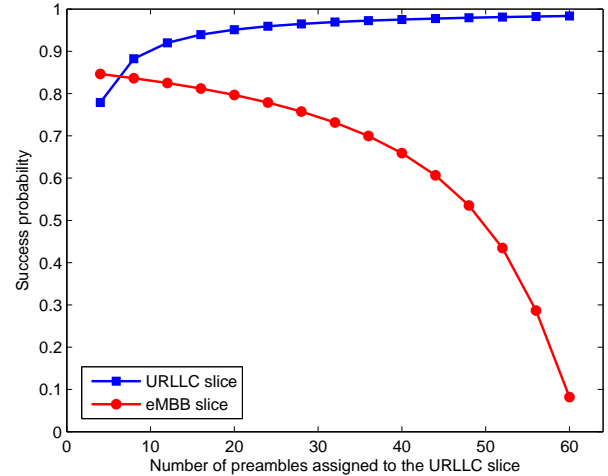


Figure 2: RA success probability under the isolation scheme.

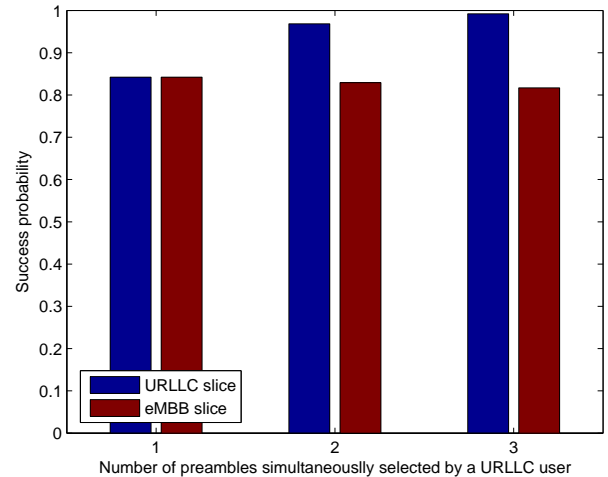


Figure 3: RA success probability under the proposed scheme.

Fig. 3 shows the success probability for the two slices under the proposed scheme. To keep the additional complexity of RA procedures in a reasonably low level for the proposed scheme, the number of RA preambles simultaneously selected by a URLLC user takes values from $\{1, 2, 3\}$. If a URLLC user selects only a single RA preamble, i.e., $L=1$, the proposed scheme is the same as the traditional physical sharing scheme in the legacy network architecture, and it can be observed that the URLLC slice has the same RA performance with the eMBB slice. When multiple preambles can be selected in one slot by a URLLC user, the proposed scheme can significantly

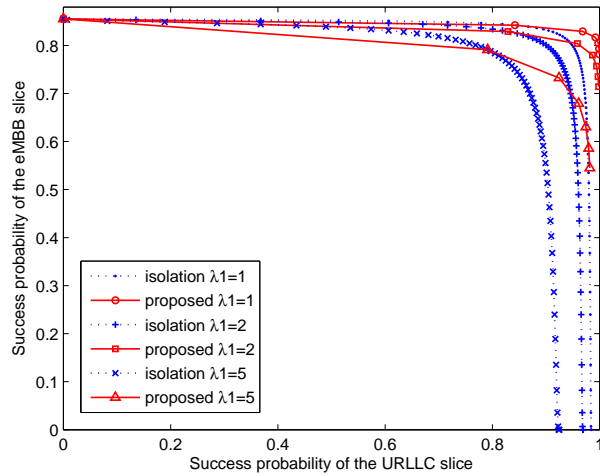


Figure 4: RA success probability of the URLLC slice vs. the eMBB slice.

improve RA success probability of URLLC users; specifically, when $L=2$, the success probability of the URLLC slice can increase sharply to 97% compared to 84% when $L=1$ and, when $L=3$, the RA performance of the URLLC slice can safely exceed 99%. Meanwhile, although more RA load is injected into the network by the URLLC slice, the RA success probability of the eMBB slice only suffers graceful degradation given the fact that the number of eMBB users dominates in practical scenarios. In Fig. 3, the RA performance of the eMBB slice stays well above 80% with a slight degradation from 84% when $L=1$ to 82% when $L=3$, which means that our proposed scheme has a minimal impact on the RA performance of eMBB slice.

Further performance tradeoff between the two slices is shown in Fig. 4 from another perspective. The RA rate of the eMBB slice is still fixed as 10 users/slot while three different RA rates of the URLLC slice, namely, 1, 2, 5 users/slot, are evaluated. For the isolation scheme, the number of preambles assigned to the two slices is varied between 0 and N . It is obvious that more preambles assigned to the URLLC slice will lead to better RA performance for the URLLC users, however, at the cost of significant performance degradation of the eMBB slice, which can be revealed by the large slopes of the three corresponding curves on the right-hand side. For the proposed scheme, the number of RA preambles simultaneously selected by a URLLC user is varied between 0 and 6. It is observed that even when the URLLC slice is enabled to have a higher RA success probability, the eMBB slice can also achieve a stable RA success probability. Moreover, Fig. 4 also shows that RA load of the URLLC slice can have an impact on the RA performance of both schemes.

Fig. 5 depicts the RA success probability under the proposed scheme versus the variation of the arrival rate of RA attempts of the eMBB slice λ_2 , when the arrival rate

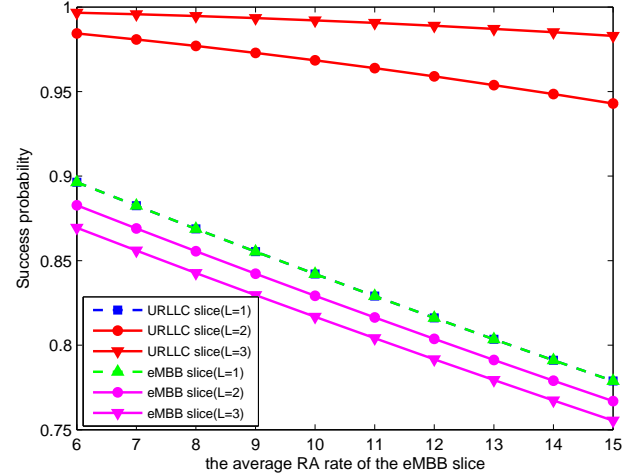


Figure 5: RA success probability vs. λ_2 under the proposed scheme.

of RA attempts of the URLLC slice is fixed, i.e., $\lambda_1 = 1$. If a URLLC user selects only a single RA preamble, i.e., $L=1$, the proposed scheme is the same as the physical sharing scheme, in which URLLC users and eMBB users have identical RA success probability. It can be observed that the RA success probability of each slice decreases as the arrival rate of RA attempts of the eMBB slice increases in the sharing scheme. In addition, for our proposed scheme, the success probability of eMBB slice decreases proportionally as its RA load and the value of L increase, where the effect imposed by RA attempts of the URLLC slice is fixed. While for the URLLC slice, the success probability decreases more slowly, especially in the larger L value case, which means that the load of eMBB slice has a very minor effect on the URLLC slice. Because multiple preambles increase the chance of successful access. In the heavier load of the eMBB slice region, the improvement in the RA performance of the URLLC slice is even greater compared to the physical sharing scheme.

6 CONCLUSIONS

In this paper, we proposed a novel scheme to support efficient network slicing for the RA procedure. In our proposed scheme, two network slices with distinct requirements on reliability and latency were considered, whose traffic also has very different characteristics. Both slices are able to use all available random access preambles as candidates, a user in the slice with lower requirement is allowed to randomly select a single preamble, while a user in the slice with higher requirement is allowed to select multiple preambles in one slot. The mathematical analysis and related simulation results demonstrated that the proposed scheme outperforms the other two schemes in the performance of RA procedure, which can not only satisfy the customized performance need of each slice, but also achieve the isolation between network

slices to a certain extent as well as the efficient utility of the preamble resources.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61671074) and in part by the State Major Science and Technology Special Projects (Grant No. 2016ZX03001017-004).

REFERENCES

- [1] K. Samdanis, X. C.-Pérez, and V. Sciancalepore “From network sharing to multi-tenancy: The 5G network slice broker,” *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 32-39, July 2016.
- [2] X. C.-Pérez, J. Swetina, T. Guo, R. Mahindra, and S. Rangarajan, “Radio access network virtualization for future mobile carrier networks,” *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 27-35, July 2013.
- [3] M. Jiang, M. Condoluci, and T. Mahmoodi, “Network slicing management and prioritization in 5G mobile systems,” in *Proc. European Wireless Conference'16*, Oulu, Finland, May 2016, pp. 1-6.
- [4] R. Kokku, R. Mahindra, H. Zhang, and S. Rangarajan, “CellSlice: Cellular wireless resource slicing for active RAN sharing,” in *Proc. COMSNETS'13*, Bangalore, India, Jan. 2013, pp. 1-10.
- [5] M. Kalil, A. Shami, and Y. Ye, “Wireless resources virtualization in LTE systems,” in *Proc. IEEE INFOCOM WKSHPS'14*, Toronto, Ontario, Canada, Apr. 2014, pp. 363-368.
- [6] R. Kokku, R. Mahindra, H. Zhang, and S. Rangarajan, “NVS: A substrate for virtualizing wireless resources in cellular networks,” *IEEE/ACM Trans. Netw.*, vol. 20, no. 5, pp. 1333-1346, Oct. 2012.
- [7] E. Osipov, L. Riliskis, A. E.-Damlin, M. Burakov, M. Nordberg, and M. Wang “An improved model of LTE random access channel,” in *Proc. IEEE VTC Spring'13*, Dresden, Germany, June 2013, pp. 1-5.
- [8] J. -P. Cheng, C. -H. Lee, and T. -M. Lin, “Prioritized random access with dynamic access barring for RAN overload in 3GPP LTE-A networks,” in *Proc. IEEE GC Wkshps'11*, Houston, Texas, USA, Dec. 2011, pp. 368-372.
- [9] H. -Y. Hwang, S. -M. Oh, C. Lee, J. H. Kim, and J. Shin, “Dynamic RACH preamble allocation scheme,” in *Proc. ICTC'15*, Hangzhou, China, Oct. 2015, pp. 770-772.