ACK Feedback based UE-to-CTU Mapping Rule for SCMA Uplink Grant-Free Transmission

Jiali Shen, Wen Chen, Fan Wei and Yongpeng Wu

Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China

Email: {lisalisa, wenchen, weifan89, yongpeng.wu }@sjtu.edu.cn

Abstract—**Massive machine type communication (mMTC) is one of the three significant scenarios in 5G, which is characterized by massive connection and low energy consumption. To meet these requirements, the contention based grant-free sparse code multiple access (SCMA) transmission which can support a large number of devices and reduce the cost of signaling overhead caused by massive connection is proposed. It allows user equipment (UE) to transmit in the preconfigured radio resources shared by multiple UEs, which is called contention transmission unit (CTU). Since several UEs can be mapped to the same CTU in the transmission, collision may occur. In the conventional set-up, the UE is allocated to a CTU by the fixed mapping rule, which brings the problem of unfairness among the UEs and the probability that UE may collide again in retransmission. In this paper, we propose a new mapping rule, which utilizes the Acknowledgement (ACK) feedback to indicate the radio resource allocation. The theoretical deduction confirms the superiority of the proposed method. The simulation results show that collision probability is reduced by around 20%.**

I. INTRODUCTION

5G wireless networks are expected to support diverse applications and devices [1]. One of the three significant scenarios is massive machine type communication (mMTC), which is characterized by large number of users, small infrequent packets, uplink-dominated transmissions and low user data rates [2]. It is often connected with a plenty of low-complexity low-cost sensors and deployed for Internet of Things applications, agriculture automation and environment monitoring. Hence, missive connection and low energy consumption are two main challenges faced by mMTC.

Several new non-orthogonal multiple access (NOMA) [3] techniques with high overloading factor have been proposed to meet the massive connection requirement. Unlike conventional orthogonal multiple access techniques, NOMA introduces some controllable interference to implement overloading at the cost of the increased receiver complexity. Thus, massive connectivity and high spectral efficiency can be achieved [2]. Specifically, thanks to the sparse codebook design [4], sparse code multiple access (SCMA) [5] can achieve the multiuser detection with the acceptable complexity based on the idea of message passing algorithm (MPA).The recent work in [6] is proposed to further reduce the decoding complexity of MPA where sphere decoding is used to reduce the searching space of conventional MPA. Moreover, dynamic

resource allocation improves the throughput of SCMA system [7].

The problem caused by massive connectivity is the cost of signaling overhead for the very small packets since the ratio of signaling overhead to useful payload is high, which can cause high energy consumption. In the conventional transmission system, a large number of scheduling signals are introduced by the request-grant procedure as in Long Term Evolution (LTE) [8]. To simplify the dynamic scheduling and cut down the signaling overhead, several grant-free data transmission schemes [9]–[12] have been proposed.

Because of the infrequency traffic, grant-free transmission usually adopts a contention based mechanism to put radio resource into more efficient use. It allows several user equipment (UE) to share the same radio resources, so the system can support even more devices. In [9]-[10], SCMA-based grant-free multiple access allows UEs to transmit data in the preconfigured resources which is called contention transmission unit (CTU), i.e., a combination of time, frequency, codebook and pilot. The mapping rule of UE-to-CTU is predefined in advance. Furthermore, in [9], a blind detection solution is introduced to confirm the feasibility of grant-free transmission, which can blindly recognize active users and decode users data without active codebook knowledge.

However, contention based mechanism will incur the collision when multiple UEs who share the same CTU transmit data simultaneously. The excessive retransmission will result in more energy consumption no doubt. Hence, it is necessary to develop a solution to avoid collisions within CTU and reduce the number of retransmissions. There are two aspects to solve the problem: 1) more effective mapping rule of UEs to radio resources; 2) more reasonable assignment of retransmission that avoids twice colliding when retransmitting. The determination of the access region mentioned in [10] is driven by the UE identity(ID) , which may result in unfairness between UEs.

In this paper, we propose an Acknowledgement (ACK) feedback based UE-to-CTU mapping rule, in which the UE who has retransmission data can get an exclusive radio resource to ensure a successful retransmission. The rest of this paper is organized as follows. Section II introduces the procedure of grant-free transmission and

radio resource assignment. Section III gives the proposed mapping rule and the theoretical derivation. Section IV evaluates the performance of the proposed method compared with the conventional methods. Section V concludes this paper.

II. GRANT-FREE TRANSMISSION AND RADIO RESOURCE ASSIGNMENT

In a mMTC contention based network as shown in Fig. 1, the devices are connected to the same base station (BS) with star topology. Each UE has an ID assigned by BS when they are accessed in the cell, where UEi denotes the UE with the IDi[1].
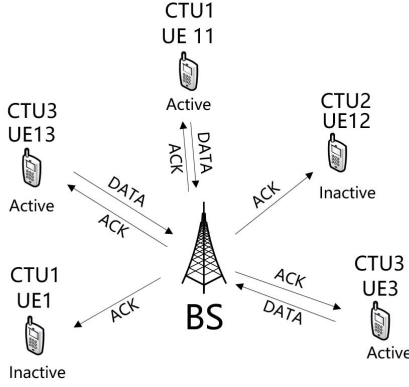

Fig. 1: Contention based grant-free network in mMTC

A grant-free transmission procedure is shown in Fig. 2. Once data arrives, UE transmits data immediately in the preconfigured resources at transmission opportunity without any signaling required. At the receiver, BS blindly detects active UEs, decodes the uplink data with no active codebook knowledge and then sends back an ACK sequence to all the UEs for transmission confirmation. UEs check the ACK sequence to see whether they need to take the random back-off mechanism to conduct retransmission or not.

The structure of ACK feedback is described below. Assume that the number of UEs is N. An $N \times 1$ vector $\boldsymbol{x} = (x_1, x_2, ..., x_N)^T$ represents the ACK feedback sequence, where $x_n = 1$ means that the BS decodes the data successfully from the nth UE, and $x_n = 0$ indicates that the nth UE has not been decoded or doesn't transmit data at all.

In order to support SCMA-based uplink grant-free multiple access, preconfigured resource is known in advance by UEs as the basic unit for contention. In SCMA, codebooks may go through different channels, and the receiver can estimate the channel impulse response of different UEs with different pilots. Thus, it can still decode the UEs' data by applying MPA algorithm even if they are carried over identical codebook as long as different

---

[1]To simplify the figure, only a part of UEs are shown. The mapping rule of UE-to-CTU is given in (1).
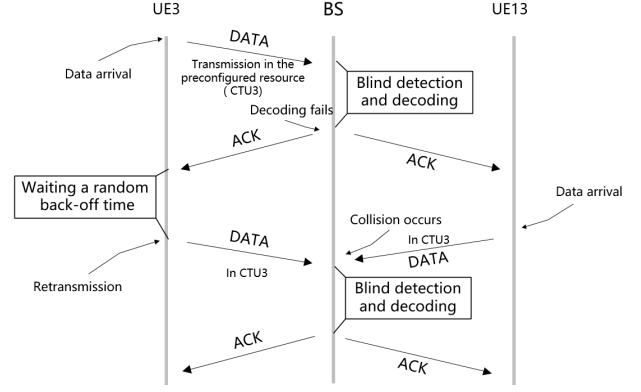

Fig. 2: Procedure of grant-free transmission

pilots are used. Hence, the preconfigured resource unit for contention can be defined as a combination of time, frequency, codebooks and pilots, which is called CTU. Assume that J codebooks are defined over one time-frequency resource and each codebook associates L pilot sequences. Thus, the number of CTUs is $J \times L$ in the given time-frequency region [10]. In mMTC, UEs can be far more than CTUs, so that one CTU may be shared by several UEs according to a certain UE-to-CTU mapping rule. Then, the collision may occur when the UEs who map to the same CTU transmit data at the same time. As shown in Fig. 1, a collision occurs between UE3 and UE13.

Therefore, a reasonable mapping rule of UE-to-CTU is necessary to reduce the collision probability. In [10], the author introduces a mapping rule with the mapping relationship driven from the UE ID. If the total number of CTU is N, the CTU index that the UE chooses to transmit data can be defined as:

$$CTU\,index = UE\,ID\,mod\,N. \tag{1}$$

The collisions are resolved via the random back-off procedure as illustrated in Fig. 2. Retransmission UE picks a random back-off time from a back-off window and retransmits data on the preconfigured CTU.

However, this mapping rule may result in unfairness between the UEs. Some CTUs are allocated to more UEs compared with the others since the distribution of UEs' ID may not be continuous. As illustrated in Fig. 3, the CTU3 and CTU4 are allocated to three UEs while CTU2 and CTU5 only have one UE, so that CTU3 and CTU4 are more likely to have collisions. Furthermore, if the traffic among UEs is not balanced, this fixed mapping rule may exacerbate the unfairness. Under some conditions, the data arrival time of part UEs may have certain correlation. As the forest fire alarm system, the sensors which are distributed around the fire area will report temperature data over the same period. And if some of them are associated to the same CTU, serious collisions occur constantly. Although the unfairness can be alleviated

by changing the mapping rule in a predefined manner from time to time, the UE may collide again when it conducts the retransmission in the picked random back-off time slot. Excessive retransmissions increase the energy consumption. To solve this problem, a new mapping rule
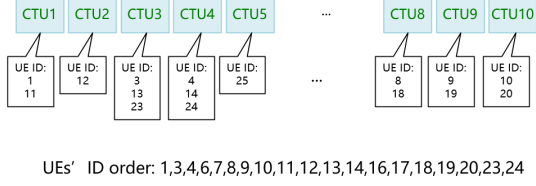

Fig. 3: UE-to-CTU mapping relationship in conventional rule

of UE-to-CTU is proposed, in which the UE who has retransmission requirement can get a special allocation.

III. ACK FEEDBACK BASED UE-TO-CTU MAPPING RULE

In this section, the mapping rule of UE-to-CTU is described firstly. Then, theoretical derivation of the collision probability is given. Finally, the grant-free transmission procedure with the proposed mechanism will be presented.

A. Mapping Rule of UE-to-CTU

As mentioned in the previous section, the fixed mapping relationship can cause excessive collisions in the retransmission procedure. To avoid multiple collisions, the UE collided in the last time slot can retransmit data in an exclusive CTU in the proposed mapping rule.

In Fig. 4, a collision happens between UE2 and UE12, so they are assigned to an exclusive CTU for retransmission respectively in the next time slot. Due to the grant-free mechanism, the reallocation process of UE-to-CTU is done directly by UEs according to the transmission state information contained in the ACK.

The structure of ACK sequence has been described in the previous section. Each UE has one bit in the ACK sequence, where 1 denotes successful decoding and 0 indicates the failed decoding or no-received data. This representation does not distinguish the conditions of retransmission from non-transmission in last slot. Therefore, we add a bit to divide this two states. Then $x_n = 00$ means the BS has not received data from the nth UE, $x_n = 11$ denotes that the data is successfully decoded and $x_n = 01$ indicates that the nth UE needs to retransmit.

The ACK structure infers that the UEs know which one needs to retransmit, in other words, they know both their own state and the others' states. Hence, UEs marked by 01 can be allocated in an exclusive CTUs by the order of the corresponding position in ACK sequence. The other UEs will be evenly distributed in the rest of the CTUs in a order, as illustrated in Fig. 4.

One thing to note is that, when decoding fails, the BS is unable to know the corresponding UEs' ID. The only

thing it knows is the pilots/CTUs they transmitted on. All the UEs who belong to this CTU will be marked to 01, since we can't distinguish the failed one exactly. In this case, the UE who doesn't have retransmission requirement may be allocated to an exclusive CTU, such as UE14 in Fig. 4, thus resulting in resource waste. It can be predicted that the scheme performance will degrade when CTUs are shared by excessive UEs. However, the impact is still less in the proposed scheme than in the conventional one. This matter will be discussed in detail in the following subsection and section IV.

In Fig. 4, UEs' ID is continuous in order to make it easier to understand. In practice, the UEs' ID may be discontinuous. Nevertheless, it will not affect the implementation of the proposed mapping rule whose allocation method is based on the order in the ACK rather than ID.
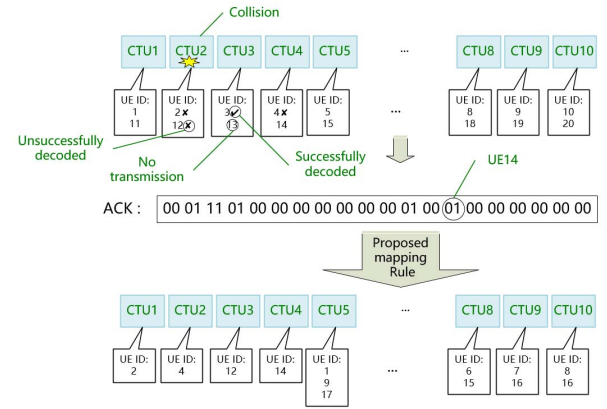

Fig. 4: ACK feedback based mapping rule

B. Theoretical Analysis

We consider a grant-free system with N CTUs. The average number of UEs associated with the same CTU is n, so there are $N \times n$ UEs in total. N_m is the number of CTUs who are allocated with m UEs. p represents each UE's average packet arrival rate per transmission time interval (TTI), and it is also a measure of traffic frequency. The number of UEs which transmit data simultaneously within the same CTU is k. Apparently, there is a collision in the CTU when $k \geq 2$.

[13] takes a Poisson distribution to build the traffic model with average packet arrival rate of λ per TTI. As we know, the Poisson distribution $X \sim P(\lambda)$ is an approximation of the binomial distribution $X \sim B(n, p)$ when n is large and p is small. In this paper, we analyze the collision situations within the category of CTU, where the number of UEs within one CTU is not large enough. So we take the binomial distribution to analyze here. The probability that k UEs send data simultaneously follows the binomial distribution, which can be written as

$$P(X = k) = b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}. \quad (2)$$

$$P_{con} = \frac{N \sum_{k=2}^{n} p_{n,k,p} k \sum_{k=1}^{n-1} p_{n-1,k,p}(k+1) + N(1 - \sum_{k=2}^{n} p_{n,k,p}) \sum_{k=2}^{n} p_{n,k,p} k}{Nnp}. \tag{7}$$

Then the collision probability within one CTU can be written as

$$P(X \geq 2) = \sum_{k=2}^{n} \binom{n}{k} p^k (1-p)^{n-k}. \tag{3}$$

In this paper, we denote $p_{n,k,p}$ as $b(k; n, p)$. P_{con} and P_{pro} represent the conventional method and the proposed method respectively, which are the collision probability of UEs that conduct transmissions. In order to provide the theoretical analysis for the proposed mapping rule, the collision probability of two rules will be discussed below.

1) The conventional mapping rule
   In one transmission process, CTUs can be divided into two categories: with retransmission UEs and without retransmission UEs. The number of retransmission UEs in total CTUs is

$$N \sum_{k=2}^{n} p_{n,k,p} k = N \sum_{k=2}^{n} \binom{n}{k} p^k (1-p)^{n-k} k. \tag{4}$$

The probability of retransmission UE can be defined as

$$\sum_{k=1}^{n-1} p_{n-1,k,p}(k+1) = \sum_{k=1}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-k-1}(k+1). \tag{5}$$

(5) is the probability that one or more UEs (except for the retransmission UE) conduct transmission in the CTU. The number of CTUs without retransmission UEs is

$$N(1 - \sum_{k=2}^{n} p_{n,k,p}) = N(1 - \sum_{k=2}^{n} \binom{n}{k} p^k (1-p)^{n-k}). \tag{6}$$

Since the number of collision UEs in one CTU is $\sum_{k=2}^{n} p_{n,k,p} k$, and Nnp is the number of active UEs in total, P_{con} can be written as (7).

2) The proposed mapping rule
   The number of CTUs which have collisions is $N \sum_{k=2}^{n} p_{n,k,p}$. So the number of CTUs which associates only one UE can be written as

$$N_{single} = Nn \sum_{k=2}^{n} p_{n,k,p}. \tag{8}$$

Then the number of CTUs that associates multiple UEs is

$$N_{multi} = N - N_{single}. \tag{9}$$

Because the no-collision UEs will be evenly distributed in the N_{multi} CTUs in order, the number of UEs in each CTU may be different. The probability

need to be counted respectively. N_k and N_{k+1} can be obtained by the following equations

$$N_k + N_{k+1} = N_{multi},$$
$$kN_k + (k+1)N_{k+1} = Nn(1 - \sum_{k=2}^{n} p_{n,k,p}). \tag{10}$$

So P_{pro} can be written as

$$P_{pro} = \frac{N_k \sum_{k=2}^{n} p_{n,k,p} k + N_{k+1} \sum_{k=2}^{n+1} p_{n+1,k,p} k}{Nnp}. \tag{11}$$

3) Probability comparison
   Since $\sum_{k=2}^{n} p_{n,k,p} k$ can be simplified as

$$\sum_{k=2}^{n} p_{n,k,p} k = np[1 - (1-p)^{n-1}], \tag{12}$$

the difference $P_{con} - P_{pro}$ can be written as

$$P_{diff} = (1 - q^{n-1})(1 - q^{n-1} + (n-1)p)$$
$$+ (1 - q^n - npq^{n-1})(n^2 - 1)(q^n - q^{n-1}), \tag{13}$$

where $q = 1 - p$. We can infer that P_{diff} must be greater than zero by derivation and it is a convex function. Fig. 5 shows P_{diff} with respective to p and n. In the conventional mapping rule, we have not considered the probability that the collided UEs pick the same back-off time slot and then collide again. Even so, the proposed mapping rule still has lower collision probability than the conventional one.
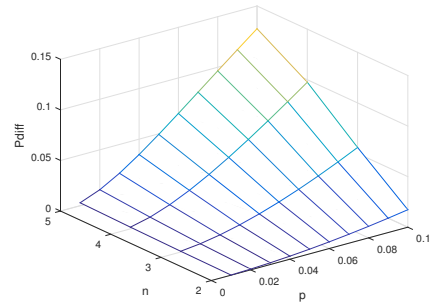
Fig. 5: P_{diff} as a function of p and n

C. Procedure with The Proposed Mapping Rule

Here we assume that the transmission delay is 3 time slots. MAPi[2] denotes the mapping rule generated in time slot i by UEs, and ACKi denotes the ACK sequence sent in time slot i by BS. The grant-free transmission procedure with the proposed UE-to-CTU mapping rule is shown in Fig 6. And the details are described as following.

[2]For $i = 1,2,...,5,6$, MAPi:$CTU index = UE ID \bmod N$.

- *Time slot 1*:
  UE: Transmit data according to MAP1.
- *Time slot 4*:
  BS: Generate ACK1 according to the decoding result and send it out. All the UEs, who stay in the same CTU along with the failed UEs, will be marked as 01. The UE-to-CTU mapping relationship can be learned from MAP1. Generate the MAP4 according to ACK1. Generation method is described in *Time slot 7*.
- *Time slot 7*:
  UE: Once data arrives, start to receive the ACK1.[3] Generate a new mapping relationship MAP4 according to ACK1. Generation method: Find their own corresponding signals x_n in the ACK1. If $x_n = 01$, count the number N_{single} of all the UEs whose $x_n = 01$ and get the order index j among them, i.e., $CTUindex = j \bmod N_{single}$.
  If not, count the number N_{multi} of all the UEs whose $x_n \neq 01$ and get the order index j among them, i.e., $CTUindex = (j \bmod N_{multi}) + (N - N_{multi})$. Finally, transmit data.
- *Time slot 10*:
  BS: Generate ACK7 according to the decoding result and send it out. Mark x_n referring the mapping relationship known from MAP4. Generate the MAP10 according to ACK7.
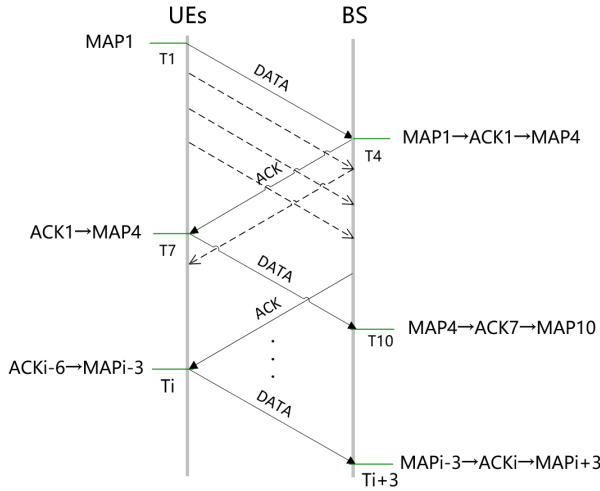

Fig. 6: Procedure of grant-free transmission with the proposed method

IV. PERFORMANCE EVALUATION

In this section, we compare the proposed mapping rule with the two conventional methods that employ two different back-off mechanism. The simulation parameters, results and analysis are given below.

[3]Note that UEs who do not have transmission requirements do not need to receive ACK and stay in the Energy Conserved Operation (ECO) state for the energy saving.

A. Schemes for Comparison

1) CMR with FB: Conventional mapping rule with fixed contention window back-off algorithm.
2) CMR with BEB: Conventional mapping rule with binary exponential back-off(BEB) algorithm.
3) The proposed mapping rule.

B. Simulation Parameters

In the simulation, we assume that the UEs' ID is continuous. the number of CTUs is 1000. In CMR with FB, the back-off window is 15. In CMR with BEB, the minimize back-off window is 4, the maximum is 32 and the basic back-off time is one time slot. Here, we assume that the data must be successfully decoded as long as no collision occurs.

C. Simulation Results

Fig. 7 and Fig. 8 shows the simulation results in two conditions. In Fig. 7, n is 3, then the total number of UEs associated with BS is $N \times n = 3000$. Fig. 7 shows the average collision rate with respective to traffic frequency (p). In Fig. 8, we consider $p = 0.01$ and give the average collision rate versus the average number of UEs in one CTU (n). As we can see, the proposed mapping rule have a better performance than CMR with FB and CMR with BEB. Comparing with these two conventional methods, the collision probability is reduced by around 20% and 35% respectively, when n is around 3 and p is around 0.01.
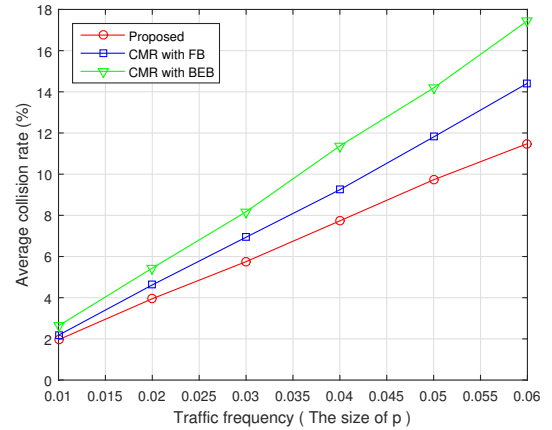

Fig. 7: Collision possibility vs. traffic frequency (p)

Fig. 9 gives the distribution of collision rates as a function of traffic frequency (p). It shows clearly that the proposed mapping rule has a lower collision rates distribution than the other two for all p.

Fig. 10 is the distribution statistics of each UE's retransmission times. Here, we set $p = 0.01$, $n = 3$ and set the run-time of simulation to be 5000. Apparently, the proposed rule has lower retransmission times than the conventional methods. Furthermore, some collision data have
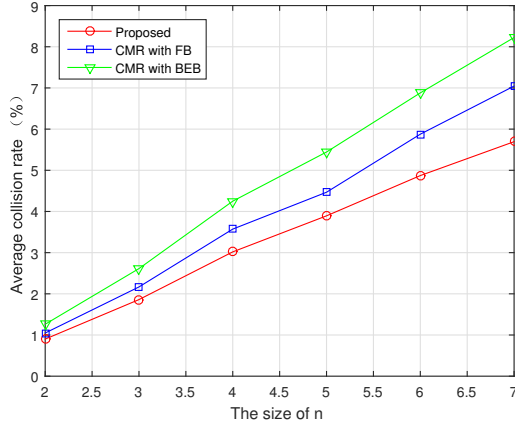
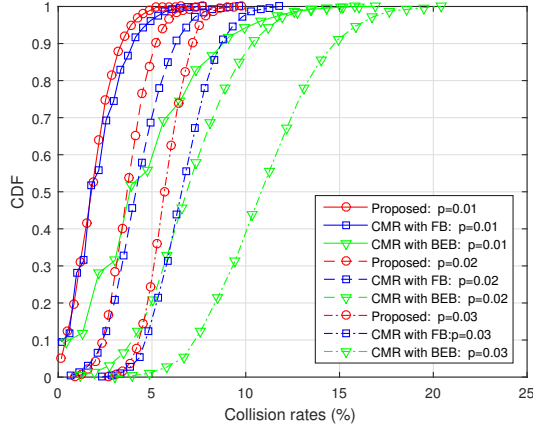Fig. 8: Collision possibility vs. average number(n) of UEs associated with the same CTU

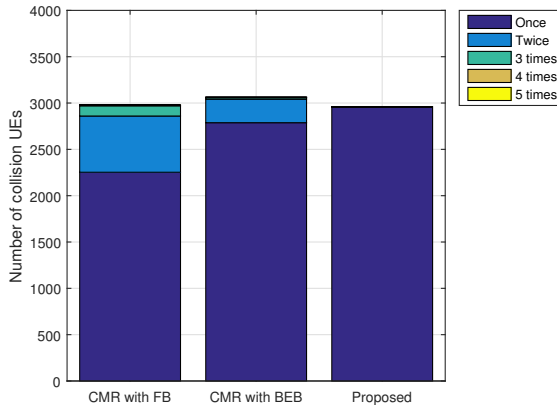Fig. 9: CDF of collision rates as a function of traffic frequency

Fig. 10: Distribution statistics of each UE's retransmission times

to retransmit two or more times in conventional mapping rule, thus resulting in higher energy consumption.

V. CONCLUSIONS

In order to support mMTC applications, in this paper, we propose an ACK feedback based UE-to-CTU mapping rule for SCMA uplink grant-free transmission. By reformulating the ACK structure and exploiting its information to indicate resource reallocation, our rule provides the UE which has retransmission data with an exclusive transmission resource to avoid twice collision. The theoretical derivation and simulation results confirm that the proposed method can decrease the collision probability and reduce the number of transmission. Thus, it can save the energy, which can be adapted to mMTC.

REFERENCES

[1] "NGMN 5G white paper," *https://www.ngmn.org/uploads/media/NGMN 5G White Paper.pdf.*

[2] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, 2016.

[3] T. Yunzheng, L. Long, L. Shang, and Z. Zhi, "A survey: Several technologies of non-orthogonal transmission for 5G," *China Commun*, vol. 12, no. 10, pp. 1–15, 2015.

[4] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "Scma codebook design," in *IEEE Vehicular Technology Conference*, 2014, pp. 1–5.

[5] H. Nikopour and H. Baligh, "Sparse code multiple access," *IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, pp. 332–336, 2013.

[6] F. Wei and W. Chen, "Low complexity iterative receiver design for sparse code multiple access," *IEEE Trans. Commun*, vol. 65, no. 2, pp. 621–634, 2017.

[7] Z. Li, W. Chen, F. Wei, F. Wang, X. Xu, and Y. Chen, "Joint codebook assignment and power allocation for scma based on capacity with gaussian input," in *Communications in China (ICCC), 2016 IEEE/CIC International Conference on*, 2016, pp. 1–6.

[8] K. Zhou, N. Nikaein, R. Knopp, and C. Bonnet, "Contention based access for machine-type communications over LTE," in *IEEE Vehicular Technology Conference*, 2015, pp. 1–5.

[9] A. Bayesteh, E. Yi, H. Nikopour, and H. Baligh, "Blind detection of SCMA for uplink grant-free multiple-access," *International Symposium on Wireless Communications Systems*, pp. 853–857, 2014.

[10] K. Au, L. Zhang, H. Nikopour, E. Yi, A. Bayesteh, U. Vilaipornsawai, J. Ma, and P. Zhu, "Uplink contention based scma for 5G radio access," in *Globecom Workshops*, 2014, pp. 900–905.

[11] B. Wang, L. Dai, Y. Yuan, and Z. Wang, "Compressive Sensing Based Multi-User Detection for Uplink Grant-Free Non-Orthogonal Multiple Access," *IEEE Vehicular Technology Conference*, pp. 1–5, 2015.

[12] D. Lin, G. Charbit, and I. K. Fu, "Uplink contention based multiple access for 5G cellular IoT," *IEEE Vehicular Technology Conference*, 2016.

[13] J. Zhang, L. Lu, Y. Sun, Y. Chen, J. Liang, J. Liu, H. Yang, S. Xing, Y. Wu, and J. Ma, "Poc of scma-based uplink grant-free transmission in ucnc for 5G," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1353–1362, 2017.