Hybrid S-ALOHA/TDMA Protocol for LTE/LTE-A Networks with Coexistence of H2H and M2M Traffic

Nannan Sui[1], Cong Wang[1], Wei Xie[1], Youyun Xu[1,2]
[1] College of Communications Engineering, PLA University of Science and Technology, China
[e-mails: nannansui@foxmail.com, lgdx_wangcong@sina.com, 2446440@qq.com, yyxu@vip.sina.com]
[2] National Engineering Research Center of Communication and Network Technologies,
Nanjing University of Posts and Telecommunications, China
*Corresponding author: Cong Wang

Received July 7, 2016; revised November 7, 2016; accepted December 1, 2016; published February 28, 2017

Abstract

The machine-to-machine (M2M) communication is featured by tremendous number of devices, small data transmission, and large uplink to downlink traffic ratio. The massive access requests generated by M2M devices would result in the current medium access control (MAC) protocol in LTE/LTE-A networks suffering from physical random access channel (PRACH) overload, high signaling overhead, and resource underutilization. As such, fairness should be carefully considered when M2M traffic coexists with human-to-human (H2H) traffic. To tackle these problems, we propose an adaptive Slotted ALOHA (S-ALOHA) and time division multiple access (TDMA) hybrid protocol. In particular, the proposed hybrid protocol divides the reserved uplink resource blocks (RBs) in a transmission cycle into the S-ALOHA part for M2M traffic with small-size packets and the TDMA part for H2H traffic with large-size packets. Adaptive resource allocation and access class barring (ACB) are exploited and optimized to maximize the channel utility with fairness constraint. Moreover, an upper performance bound for the proposed hybrid protocol is provided by performing the system equilibrium analysis. Simulation results demonstrate that, compared with pure S-ALOHA and pure TDMA protocol under a target fairness constraint of 0.9, our proposed hybrid protocol can improve the capacity by at least 9.44% when $\lambda_1:\lambda_2=1:1$ and by at least 20.53% when $\lambda_1:\lambda_2=10:1$, where λ_1,λ_2 are traffic arrival rates of M2M and H2H traffic, respectively.

Keywords: Machine-to-machine (M2M), human-to-human (H2H), physical random access channel (PRACH), LTE/LTE-A, hybrid medium access control protocol.

This work is supported by Major Research Plan of National Natural Science Foundation of China (No. 91438115), National Natural Science Foundation of China (No. 61371123, No. 61301165).

https://doi.org/10.3837/tiis.2017.02.005 **ISSN : 1976-7277**

1. Introduction

Machine to machine (M2M) communications are emerging service paradigm for future internet of things (IoT) and show an exponentially increase in the fields of pervasive applications such as smart industries, intelligent transportation, intelligent home, and e-health [1,2]. The long term evolution (LTE) and long term evolution advanced (LTE-A) networks can accelerate M2M implementation with its ready to use infrastructure and high capacity [3,4]. However, M2M traffic is usually characterized by massive devices, small data transmission, and large uplink-to-downlink traffic ratio, which is very different from human to human (H2H) traffic [5-7]. The coexistence of H2H and M2M communications poses great challenges to LTE/LTE-A networks such as physical random access channel (PRACH) overload, high signaling overhead, resource underutilization, and fairness [8-10].

The current random access procedure in LTE/LTE-A networks consists of four signaling messages, which are preamble transmission (Message 1), random access response (Message 2), scheduling request (Message 3), and contention resolution (Message 4). Just Message 1 occupies 6 resource blocks (RBs) in frequency domain and 2 time slots in time domain (that is 12 RBs in total). On one hand, PRACH would be easily congested when massive M2M devices send access requests simultaneously. On the other hand, signaling overhead on the PRACH are heavy to many M2M applications with small-size packets (the effective payload of many emergency alarming applications and environment monitoring applications would be several bits or bytes). Moreover, the retransmissions caused by preamble collision would exacerbate the overload problem and increase the signaling overhead. Therefore, in order to accommodate the M2M communications in LTE/LTE-A networks, we need to reconsider both the random access procedure and data scheduling thereafter.

In [11], the authors demonstrated that one stage protocols (like Slotted ALOHA (S-ALOHA)) could serve more M2M devices than two stage protocols (like time division multiple access (TDMA)) if the effective payload size is smaller than a threshold. Considering the pros and cons of different solutions presented for M2M communications [9,10], it's reasonable to combine the strength of different MAC protocols to make the LTE/LTE-A networks more suitable for the M2M traffic. However, the work in [11] has not explicitly presented a MAC protocol, and the analysis cannot be applied to the LTE/LTE-A networks directly.

Inspired by [11], we derive a packet size threshold $\bar{D}$ based on capacity analysis and prove that: if packet size $D < \bar{D}$, S-ALOHA is superior to TDMA; if packet size $D > \bar{D}$, TDMA is superior to S-ALOHA. Based on this result, we propose an S-ALOHA and TDMA hybrid MAC protocol to address these problems. The reserved resource blocks for uplink transmission in a cycle are divided into S-ALOHA part for M2M traffic and TDMA part for H2H traffic. The resource granularity of the S-ALOHA part is finer, e.g., one orthogonal frequency division multiplexing (OFDM) symbol in time domain for one data channel. Thus one resource block can bear multiple small-size M2M packets. As to the TDMA part, we optimize the tradeoff between PRACH and data channel allocation. We also exploit dynamic access class barring (ACB) to alleviate network overload problem.

In order to maximize the channel utility while ensure the fairness in terms of success rate, we formulate the hybrid MAC protocol as a joint resource allocation and ACB parameters optimization problem. We first prove that the problem is convex without fairness constraint.

Then considering the fairness, we derive the optimal solution in a cycle-by-cycle manner according to traffic load.

The main contributions of this paper are summarized as follows:

- In order to solve the problems of signaling overhead and resource underutilization, we propose a hybrid S-ALOHA and TDMA protocol. The resource granularity of S-ALOHA part is fined for M2M traffic with small-size packets. The tradeoff between PRACH and data channel in the TDMA part is optimized for H2H traffic.
- With the objective of maximizing the channel utility, we first prove that the joint resource allocation and ACB optimization problem is convex without fairness constraint. Then considering the fairness, we derive the optimal solution for each cycle according to traffic load estimation.
- We derive the upper performance bound for the proposed hybrid MAC protocol for large transmission attempt limit $(W \geq 9)$ based on system equilibrium analysis.
- We validate our analysis through numerical and simulation results. The excellent performance of the proposed hybrid MAC protocol is demonstrated through comparisons with pure S-ALOHA and pure TDMA protocol.

The rest of the paper is organized as follows. We review the related works in Section 2. The proposed hybrid protocol is presented in Section 3. The joint resource allocation and ACB optimization problem is formulated in Section 4. In Section 5, we derive the performance bound for the proposed hybrid protocol based on equilibrium analysis. The numerical and simulation results are presented in Section 6. We conclude this paper in Section 7.

2. Related Work

Although system enhancements have been applied to LTE/LTE-A to accommodate emerging M2M traffic, system overload and signaling overhead in the random access procedure are still key problems [12-15]. In [16], an online algorithm is proposed to adjust the access grant time interval periods for M2M clusters to satisfy delay requirements. In [17], a prioritized random access with dynamic access barring framework is applied to avoid congestion. In [18], the authors proposed a tree splitting based collision resolution method. In [19], the authors used a special preamble to estimate the traffic load and then adjust the allocation of random access resources. The authors in [20] try to address the system overload problem by designing a device-to-divice based heterogeneous mobile cloud computing architecture.

The M2M and H2H coexisting scenario has been investigated in [21-24]. In [21], the joint allocation and disjoint allocation schemes are studied and the results showed that the former is better. The performance of three different overload control schemes are investigated in [22]. A utility function based random access resource allocation scheme is proposed in [23] for M2M traffic coexisting with H2H traffic. In [24], the authors analyzed the throughput-delay performance when M2M sensors coexist with H2H traffic in the fiber-wireless smart grid.

However, the above works mainly focus on delay and success rate without considering signaling overhead. In [11], the authors derived that one stage strategies (like S-ALOHA) are better than two stage strategies (like TDMA) for small payload sizes. Therefore, the strength of different MAC protocols can be combined to solve the signaling overhead problem [9,10]. However, the work in [11] has not explicitly present a MAC protocol, and the analysis cannot be applied to the LTE/LTE-A networks directly. In [25], each frame is divided into contention only period and transmission only period. M2M devices contend for transmission slots in contention only period based on carrier sense multiple access (CSMA) scheme. Only granted devices can send data in transmission only period based on TDMA. In [26], the authors proposed an IEEE 802.15.4 based hybrid CSMA-TDMA scheme. The difference from [25] is that M2M devices in [26] can also transmit data in the contention period based on slotted CSMA/CA. However, both the works in [25] and [26] are not based on LTE/LTE-A specifications. Moreover, the objective of [26] is to optimize the node action strategy, not the resource allocation. In [27], M2M devices can transmit data immediately upon receiving Message 2 to reduce signaling overhead. However, all the proposed hybrid protocols in [25-27] are exclusively engineered for M2M traffic without considering the H2H traffic requirements.

In this paper, we design a hybrid S-ALOHA and TDMA protocol for the LTE/LTE-A networks with coexistence of M2M and H2H traffic. Not only the overload problem but also the signaling overhead and fairness are considered and solved through joint optimization of resource allocation and ACB. Additionally, we extend the work in [28] with ACB to provide an upper performance bound for the proposed hybrid protocol based on equilibrium analysis.

3. System Model and the Hybrid S-ALOHA/TDMA Protocol

3.1 System Model

We consider the uplink transmission in a single TDD LTE/LTE-A cell, in which massive M2M devices coexist with H2H UEs. The packet arrival rates of M2M and H2H traffic follow the Poisson distribution with mean λ_1, λ_2 respectively, where usually $\lambda_1 \gg \lambda_2$. We assume that the inter-arrival time between two packets is long enough that each terminal can only generate one new packet in a cycle until the previous packet is transmitted successfully or dropped due to failure. Neither packet loss on the wireless channel nor the backoff scheme is considered in this paper. Hence, a packet transmission attempt may fail due to ACB check failure, contention collision, or insufficient data channels. All the failed terminals would start the retransmission attempts immediately in the next cycle. A packet will be dropped if it has not been sent successfully within the W transmission attempts.

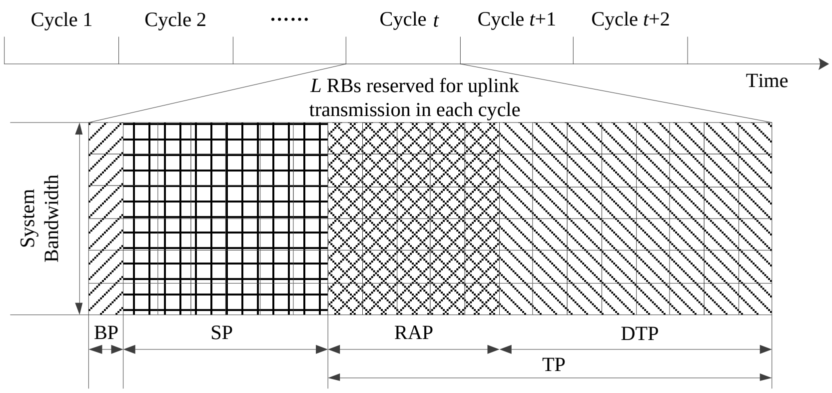


Fig. 1. Frame structure of the proposed hybrid protocol.

As shown in **Fig. 1**, BS schedules the radio resources on a cycle-by-cycle basis in time domain. In each cycle, L RBs are reserved for uplink transmission, which include the essential downlink signaling messages. The proposed hybrid S-ALOHA/TDMA protocol divides these reserved resource bolocks into three parts: broadcast period (BP), S-ALOHA period (SP) for M2M traffic, and TDMA period (TP) for H2H traffic. The TDMA period is further divided into random access period (RAP) for PRACH and data transmission period (DTP) for contention-free packet transmission. For the tractability of analysis, we assume that the four random access Messages would all complete in the random access period. Let $L_{BP,t}, L_{SP,t}, L_{TP,t}, L_{RAP,t}, L_{DTP,t}$ denote the number of RBs allocated to each period in cycle t, then:

$$\begin{aligned} L &= L_{BP,t} + L_{SP,t} + L_{TP,t} \\ &= L_{BP,t} + L_{SP,t} + L_{RAP,t} + L_{DTP,t} \end{aligned} \tag{1}$$

We assume that a random access opportunity would consume $1/\gamma$ RBs. Let $\alpha_t = L_{RAP,t} / L_{TP,t}$, then the number of available preambles in random access period is $L_{TP,t}\alpha_t\gamma$. Let D_1, D_2 denote the size of M2M and H2H packet in the unit of RB. We assume that one data channel is exactly fit for one packet, let $M_{1,t}, M_{2,t}$ denote the number of data channels in S-ALOHA period and TDMA period:

$$\begin{aligned} M_{1,t} &= \left\lfloor \frac{L_{SP,t}}{D_1} \right\rfloor \\ M_{2,t} &= \left\lfloor \frac{L_{DTP,t}}{D_2} \right\rfloor \end{aligned} \tag{2}$$

3.2 Frame Structure of the Hybrid S-ALOHA/TDMA Protocol

The details of each period of the proposed hybrid protocol are stated as follows:

1) Broadcast Period: At the beginning of cycle t, the BS estimates the number of active terminals based on the statistics obtained until cycle $t-1$. Considering the fairness, BS calculates the optimal resource allocation (include the start and end of each period) and ACB parameters $p_{1,t}, p_{2,t}$ for M2M and H2H traffic respectively. Then the BS broadcasts a notification message containing these parameters to all the terminals for synchronization.

Upon receiving this notification message, each active terminal would execute ACB check procedure [9,17]. A terminal would generate a random variable $x \in [0,1]$. If $x \le p$, this terminal can continue the transmission attempt. Otherwise, this terminal has to wait for the next cycle. Let $N_{1,t}, N_{2,t}$ denote the number of active M2M devices and H2H UEs, and let $Y_{1,t}, Y_{2,t}$ denote the number of contending M2M devices and H2H UEs. We can obtain:

$$E\left[Y_{i,t}\right] = E\left[N_{i,t}\right] p_{i,t}, \quad i = 1,2 \tag{3}$$

2) S-ALOHA Period: The contending M2M devices send packets based on S-ALOHA protocol in this period. Let $S_{1,t}^n$ denote the number of M2M devices succeeding in the nth transmission attempt in cycle t, and let $S_{1,t}$ denote the total number of successful devices.

3) Random Access Period: The contending UEs will select a preamble randomly and send it on the PRACH. The failed UEs due to collision would start retransmission in the next cycle immediately.

Let $S_{RAP,t}$ denote the total number of successful UEs in the random access period. If $S_{RAP,t} \le M_{2,t}$, each successful UE will be allocated one data channel. However, if $S_{RAP,t} > M_{2,t}$, the BS can only schedule $M_{2,t}$ UEs out of $S_{RAP,t}$ successful UEs randomly. The BS will notify the corresponding UEs of the resource allocation results via Message 4. The unscheduled UEs have to start retransmission requests in the next cycle. For clarity of this paper, we refer to successful UEs as the scheduled UEs.

4) Data Transmission Period: In this period, the scheduled UE would send packet in the reserved data channel without collision. Let $S_{2,t}^{n}$ denote the number of H2H UEs succeeding in the nth transmission attempt in cycle t and let $S_{2,t}$ denote the total number of successful UEs.

Because the number of RBs in broadcast period is usually small and relatively constant, without losing the generality, we just ignore the $L_{BP,t}$ in the following analysis. Actually, the joint optimization of resource allocation and ACB is an integer programming problem, which makes the analysis difficult. For the tractability of analysis, we relax the integer constraint and assume that all the variables used in this paper are real. As shown in Sec. 6, the applicability of this assumption to the system is validated by numerical and simulation results.

3.3 Performance Metrics

Four performance metrics are introduced in this paper, which are channel utility, success rate, delay, and fairness.

1) Channel Utility: Channel utility R_t is defined as the number of resource blocks that have been used to transmit data successfully in cycle t. Let $R_{1,t}, R_{2,t}$ denote the channel utility of M2M and H2H traffic respectively, we can derive that:

$$\begin{aligned} &R_{i,t} = S_{i,t} D_i, \qquad i = 1,2 \\ &R_t = R_{1,t} + R_{2,t} \end{aligned} \tag{4}$$

2) Success Rate: Success rate is defined as the ratio of the number of successful terminals to the number of active terminals. Let $P_{1,t}, P_{2,t}$ denote the success rate of M2M and H2H traffic at the end of cycle t respectively. Let $X_{1,t}, X_{2,t}$ denote the number of new arrivals of M2M and H2H packets in cycle t respectively. We can derive that:

$$P_{i,t} = \frac{\sum_{n=1}^{t} S_{i,n}}{\sum_{n=1}^{t} X_{i,n}}, \quad i = 1,2 \tag{5}$$

3) Delay: Delay is defined as the average number of transmission cycles that a successful terminal has experienced. Let $d_{1,t}, d_{2,t}$ denote the delay of M2M and H2H traffic at the end of cycle t respectively, we can compute:

$$d_{i,t} = \frac{\sum_{n=1}^{t}\sum_{m=1}^{W} m S_{i,n}^{m}}{\sum_{n=1}^{t} S_{i,n}}, \quad i = 1,2 \tag{6}$$

4) Fairness: In this paper, we exploit Jain's fairness index $P_{F,t}$ to evaluate the fairness of the proposed hybrid protocol in terms of success rate [27]:

$$P_{F,t} = \frac{\left(\mathrm{P}_{1,t} + P_{2,t}\right)^2}{2\left(\mathrm{P}_{1,t}{}^2 + P_{2,t}{}^2\right)} \tag{7}$$

4. Joint Resource Allocation and ACB Optimization

4.1 Packet Size Threshold $\bar{D}$

In this paper, the concept of *capacity* is defined as the maximum channel utility that a particular MAC protocol can achieve. We model the random access procedure as a multi-channel S-ALOHA system [30]. For brevity, we reuse the variables introduced in Sec. 3 just omit the subscript t.

1) Capacity of S-ALOHA: Known the number of reserved RBs L, the expectation number of successful devices and channel utility can be calculated as:

$$\begin{aligned} E\left[S_1\right] &= Y_1 e^{-\frac{Y_1}{L/D}} \\ E\left[R_1\right] &= D\cdot E\left[S_1\right] = D\cdot Y_1 e^{-\frac{Y_1}{L/D}} \end{aligned} \tag{8}$$

where L/D is the number of data channels. Let $Y_1^* = L/D$, it's easy to prove that the capacity of S-ALOHA C_1 can be obtained when $Y_1 = Y_1^*$:

$$C_1 = \max_{Y_1} E\left[R_1\right] = Le^{-1} \tag{9}$$

2) Capacity of TDMA: According to the analysis in Sec. 3.2, the expectation value of S_2 and R_2 in the TDMA system can be obtained:

$$\begin{aligned} E\left[S_2\right] &= E\left[\min\left\{S_{RAP}, M_2\right\}\right] \\ &\le \min\left\{E\left[S_{RAP}\right], M_2\right\} \\ E\left[R_2\right] &= DE\left[S_2\right] \\ &\le \min\left\{DY_2 e^{-\frac{Y_2}{L\alpha\gamma}}, L\left(1-\alpha\right)\right\} \end{aligned} \tag{10}$$

Then the upper bound of TDMA capacity C_2 and optimal α can be derived as:

$$\begin{aligned} C_2 &= \max_{Y_2} DY_2 e^{-\frac{Y_2}{L\alpha\gamma}} \\ s.t.\quad DY_2 e^{-\frac{Y_2}{L\alpha\gamma}} &= L\left(1-\alpha\right) \end{aligned} \tag{11}$$

Similar to Eq. (9), the optimal Y_2^*, α^* should satisfy $Y_2^* = L\alpha^*\gamma$, so we can have:

$$\begin{aligned} C_2 &= DL\alpha^*\gamma e^{-1} = \frac{LD\gamma}{e+D\gamma} \\ \alpha^* &= \frac{1}{1+D\gamma/e} \end{aligned} \tag{12}$$

3) Packet Size Threshold $\bar{D}$: Let $C_1 = C_2$, the packet size threshold can be derived as:

$$\bar{D} = \frac{e}{e-1}\cdot\frac{1}{\gamma} \tag{13}$$

We can summarize as follows: if packet size $D < \bar{D}$ (typical for M2M applications), then $C_1 > C_2$, S-ALOHA protocol is superior to TDMA protocol; if packet size $D > \bar{D}$ (typical for H2H applications), then $C_1 < C_2$, TDMA protocol is superior to S-ALOHA protocol.

In this paper, only probabilistic statistics makes sense. For brevity, we omit the expectation operator $E[\cdot]$ and let random variable X itself denote its expectation value $E[X]$.

4.2 Access Class Barring

Given the number of reserved RBs L, based on the capacity analysis in Sec. 4.1, we should let all the terminals pass the ACB check if $N_i < Y_i^*$ and should let only Y_i^* terminals pass the

ACB check if $N_i \geq Y_i^*$. Hence, we can regard Y_i^* as the maximum number that the system can accommodate. Therefore, we define that the system is *saturate* if $N_i \geq Y_i^*$. The optimal ACB parameters can be derived as follows:

$$p_i = \min\left(Y_i^* / N_i, 1\right), \quad i = 1,2 \tag{14}$$

We define *saturation threshold* as the minimum number of RBs that allow all the N_i active terminals participate in resource contention. Let $Y_i^* = N_i$, we can obtain the saturation threshold $L_{1,th}, L_{2,th}$ for S-ALOHA and TDMA system respectively:

$$\begin{aligned} L_{1,th} &= N_1 D_1 \\ L_{2,th} &= \frac{N_2}{\alpha^* \gamma} \end{aligned} \tag{15}$$

Accordingly, we can say a system is saturate if $L \leq L_{i,th}$ and is non-saturate if $L > L_{i,th}$. Then the channel utility of S-ALOHA and TDMA system can be determined as follows:

$$R_1 = \begin{cases} Le^{-1}, & if \quad L \leq L_{1,th} \\ N_1 e^{-\frac{N_1}{L/D_1}} D_1, & if \quad L > L_{1,th} \end{cases} \tag{16}$$

$$\begin{aligned} &R_2 = \begin{cases} L(1-\alpha^*), & if \quad L \leq L_{2,th} \\ N_2 e^{-\frac{N_2}{L\alpha\gamma}} D_2, & if \quad L > L_{2,th} \end{cases} \\ &s.t. \quad N_2 e^{-\frac{N_2}{L\alpha\gamma}} D_2 = L(1-\alpha) \end{aligned} \tag{17}$$

4.3 Joint optimization of resource allocation and ACB

In cycle t, the active terminals comprise new arrivals and backlogged terminals that failed in cycle $t-1$. Therefore, if we let $N_{i,t}^n$ denote the number of active terminals that transmit the nth requests in cycle t, we can have [31]:

$$\begin{aligned} N_{i,t}^n &= \begin{cases} \lambda_i, & n = 1 \\ N_{i,t-1}^{n-1} - S_{i,t-1}^{n-1}, & n = 2,\ldots,W \end{cases} \\ N_{i,t} &= \sum_{n=1}^{W} N_{i,t}^n \end{aligned} \tag{18}$$

where $i = 1,2$. Because each terminal has the same success probability, so the number of successful terminals in different transmission stage can be computed as:

$$S_{i,t}^n = S_{i,t} \frac{N_{i,t}^n}{N_{i,t}}, \quad i = 1,2 \tag{19}$$

According to the analysis in Sec. 4.2, in order to maximize channel utility R_t of the proposed hybrid protocol with fairness constraint, the joint resource allocation and ACB optimization problem is formulated as follows:

$$\max_{L_{SP,t},L_{TP,t},p_{1,t},p_{2,t}} R_t = \begin{cases} L_{SP,t}/e + L_{TP,t}(1-\alpha^*), & if\ L_{SP,t} \le L_{1,th}, L_{TP,t} \le L_{2,th} \\ L_{SP,t}/e + N_{2,t} e^{-\frac{N_{2,t}}{L_{TP,t}\alpha_t\gamma}} D_2, & if\ L_{SP,t} \le L_{1,th}, L_{TP,t} > L_{2,th} \\ N_{1,t} e^{-\frac{N_{1,t}}{L_{SP,t}/D_1}} D_1 + L_{TP,t}(1-\alpha^*), & if\ L_{SP,t} > L_{1,th}, L_{TP,t} \le L_{2,th} \\ N_{1,t} e^{-\frac{N_{1,t}}{L_{SP,t}/D_1}} D_1 + N_{2,t} e^{-\frac{N_{2,t}}{L_{TP,t}\alpha_t\gamma}} D_2, & if\ L_{SP,t} > L_{1,th}, L_{TP,t} > L_{2,th} \end{cases}$$

$$\begin{aligned} s.t.\quad (a)\quad & N_{2,t} e^{-\frac{N_{2,t}}{L_{TP,t}\alpha_t\gamma}} D_2 = L_{TP,t}(1-\alpha_t) \\ (b)\quad & L = L_{SP,t} + L_{TP,t},\quad L_{SP,t}, L_{TP,t} \in [0, L] \\ (c)\quad & p_{1,t} = \min\left(\frac{L_{SP,t}/D_1}{N_{1,t}}, 1\right), p_{2,t} = \min\left(\frac{L_{TP,t}\alpha_t\gamma}{N_{2,t}}, 1\right) \\ (d)\quad & P_{F,t} \ge P_{F,th} \end{aligned} \tag{20}$$

where $P_{F,t}$ could be obtained by Eqs. (5)(7) and $P_{F,th}$ is a predefined fairness threshold.

Firstly, we simplify the problem of (20) by ignoring the fairness constraint (20d). Hence, we can derive the following theorem.

Theorem 1: The channel utility R_t of the proposed hybrid protocol without fairness constraint is a convex function with respect to $L_{SP,t}, L_{TP,t}$.

Proof: According to Eqs. (16)(17), if $L_{SP,t} \le L_{1,th}, L_{TP,t} \le L_{2,th}$, it's easy to derive that:

$$\begin{aligned} \frac{\partial^2 R_{1,t}}{\partial L_{SP,t}^2} = 0 \\ \frac{\partial^2 R_{2,t}}{\partial L_{TP,t}^2} = 0 \end{aligned} \tag{21}$$

Because $R_t = R_{1,t} + R_{2,t}$ and $L = L_{SP,t} + L_{TP,t}$, then we just need to prove that $R_{1,t}$ is convex with respect to $L_{SP,t}$ if $L_{SP,t} > L_{1,th}$ and that $R_{2,t}$ is convex with respect to $L_{TP,t}$ if $L_{TP,t} > L_{2,th}$.

According to Eq. (16), if $L_{SP,t} > L_{1,th}$, we can obtain the second derivative of $R_{1,t}$ with respect to $L_{SP,t}$ as:

$$\frac{\partial^2 R_{1,t}}{\partial L_{SP,t}^2} = \left(N_{1,t} D_1\right)^2 e^{-\frac{N_{1,t}D_1}{L_{SP,t}}} \cdot \frac{1}{L_{SP,t}^3}\left[\frac{N_{1,t}D_1}{L_{SP,t}} - 2\right] < 0 \tag{22}$$

where $L_{1,th} = N_{1,t} D_1$. So the convexity of $R_{1,t}$ is proved.

According to Eq. (17), if $L_{TP,t} > L_{2,th}$, we can obtain the second derivative of $R_{2,t}$ with respect to $L_{TP,t}$ as:

$$\begin{aligned} \frac{\partial^2 R_{2,t}}{\partial L_{TP,t}^2} = {} & N_{2,t}^3 D_2 e^{-\frac{N_{2,t}}{L_{TP,t}\alpha_t\gamma}} \cdot \left(\frac{1}{L_{TP,t}^2\alpha_t\gamma} + \frac{1}{L_{TP,t}\alpha_t^2\gamma}\frac{\partial\alpha_t}{\partial L_{TP,t}}\right)^2 - N_{2,t}^2 D_2 e^{-\frac{N_{2,t}}{L_{TP,t}\alpha_t\gamma}} \\ & \cdot\left[\frac{2}{L_{TP,t}^3\alpha_t\gamma} + \frac{2}{L_{TP,t}^2\alpha_t^2\gamma}\frac{\partial\alpha_t}{\partial L_{TP,t}} + \frac{2}{L_{TP,t}\alpha_t^3\gamma}\left(\frac{\partial\alpha_t}{\partial L_{TP,t}}\right)^2 - \frac{1}{L_{TP,t}\alpha_t^2\gamma}\frac{\partial^2\alpha_t}{\partial L_{TP,t}^2}\right] \end{aligned} \tag{23}$$

According to constraint Eq. (20a), we can derive that:

$$\frac{\partial\alpha_t}{\partial L_{TP,t}}=\frac{1-\alpha_t-\dfrac{N_{2,t}(1-\alpha_t)}{L_{TP,t}\alpha_t\gamma}}{L_{TP,t}+\dfrac{N_{2,t}(1-\alpha_t)}{\alpha_t^2\gamma}}$$

$$\frac{\partial^2\alpha_t}{\partial L_{TP,t}^2}=\frac{1}{L_{TP,t}+\dfrac{N_{2,t}L_{TP,t}(1-\alpha_t)}{L_{TP,t}\alpha_t^2\gamma}}\cdot\left\{-2\frac{\partial\alpha_t}{\partial L_{TP,t}}-N_{2,t}^2L_{TP,t}(1-\alpha_t)\left(\frac{1}{L_{TP,t}^2\alpha_t\gamma}+\frac{1}{L_{TP,t}\alpha_t^2\gamma}\frac{\partial\alpha_t}{\partial L_{TP,t}}\right)^2\right. \tag{24}$$
$$\left.+N_{2,t}L_{TP,t}(1-\alpha_t)\left[\frac{2}{L_{TP,t}^3\alpha_t\gamma}+\frac{2}{L_{TP,t}^2\alpha_t^2\gamma}\frac{\partial\alpha_t}{\partial L_{TP,t}}+\frac{2}{L_{TP,t}\alpha_t^3\gamma}\left(\frac{\partial\alpha_t}{\partial L_{TP,t}}\right)^2\right]\right\}$$

Substitute Eq. (24) into Eq. (23), we can have:

$$\frac{d^2R_{2,t}}{dL_{TP,t}^2}=\frac{N_{2,t}^2D_2e^{-\frac{N_{2,t}}{L_{TP,t}\alpha_t\gamma}}}{\alpha_t^3\gamma}\left(L_{TP,t}+N_{2,t}^2D_2e^{-\frac{N_{2,t}}{L_{TP,t}\alpha_t\gamma}}\frac{1}{L_{TP,t}\alpha_t^2\gamma}\right)^{-3}\cdot\left[\frac{N_{2,t}}{L_{TP,t}\alpha_t\gamma}-2\right]<0 \tag{25}$$

So the convexity of $R_{2,t}$ is proved.

Therefore, the **Theorem 1** is proved.

Let $\Pi=\{L_{SP,t}^*,L_{TP,t}^*,p_{1,t}^*,p_{2,t}^*\}$ denote the optimal solution to the channel utility maximization problem without fairness constraint. Due to the assumption $D_2>\bar{D}>D_1$, we can derive that $1-\alpha^*>e^{-1}$. So the optimal solution is $L_{TP,t}^*\in[L_{2,th},L]$, we then propose Algorithm 1 to solve the presented problem.

Algorithm 1: Channel utility maximization without fairness constraint

1: Estimate the traffic load in cycle t: $N_{1,t},N_{2,t}$
2: Compute the saturation threshold: $L_{1,th},L_{2,th}$
3: **if** $\left.\frac{\partial R_t}{\partial L_{TP,t}}\right|_{L_{TP,t}=L}\geq 0$ **then**
4: $\quad L_{TP,t}^*=L$
5: **else**
6: $\quad$ Obtain $L_{TP,t}^*$ by solving the equation $\frac{\partial R_t}{\partial L_{TP,t}}=0,\ \ L_{TP,t}\in[L_{2,th},L]$
7: **end if**
8: According to Eqs. (20a)(20c), we can obtain the optimal solution $\Pi=\{L_{SP,t}^*,L_{TP,t}^*,p_{1,t}^*,p_{2,t}^*\}$

Reconsider the fairness constraint in Eq. (20d), let $\Pi_F=\{L_{SP,t}^F,L_{TP,t}^F,p_{1,t}^F,p_{2,t}^F\}$ denote the optimal solution, then Π_F can be determined by Algorithm 2 in two steps. In step 1, we can compute fairness index $P_{F,t}^*$ that correspond to the solution $\Pi=\{L_{SP,t}^*,L_{TP,t}^*,p_{1,t}^*,p_{2,t}^*\}$. Then in step 2, we discuss the optimal solution Π_F in three cases.

Case 1: It means that the solution Π satisfy the fairness constraint, so we have $\Pi_F=\Pi$.

Case 2: If $P_{F,t}^*<P_{F,th}$ and $\mathrm{P}_{1,t}>P_{2,t}$, we need to allocate more RBs to H2H traffic to improve the fairness. Because R_t is monotonically decreasing with respect to $L_{TP,t}\in[L_{TP,t}^*,L]$, the optimal solution is $L_{TP,t}^{F} \in \left[L_{TP,t}^{*}, L\right]$ which is the nearest point to $L_{TP,t}^{*}$ and satisfy $P_{F,t} = P_{F,th}$.

Case 3: If $P_{F,t}^{*} < P_{F,th}$ and $\mathrm{P}_{1,t} < P_{2,t}$, we need to allocate more RBs to M2M traffic to improve the fairness. Because R_t is monotonically increasing with respect to $L_{TP,t} \in \left[0, L_{TP,t}^{*}\right]$, the optimal solution is $L_{TP,t}^{F} \in \left[0, L_{TP,t}^{*}\right]$ which is the nearest point to $L_{TP,t}^{*}$ and satisfy $P_{F,t} = P_{F,th}$.

Algorithm 2: Channel utility maximization with fairness constraint

Step 1: Obtain the solution $\Pi = \left\{L_{SP,t}^{*}, L_{TP,t}^{*}, p_{1,t}^{*}, p_{2,t}^{*}\right\}$ according to Algorithm 1 and compute the fairness index $P_{F,t}^{*}$.

Step 2: Obtain the optimal solution that satisfy the fairness constraint in different cases.

Case 1: if $P_{F,t}^{*} \geq P_{F,th}$ **then**

$\Pi_F = \Pi$

Case 2: else if $P_{F,t}^{*} < P_{F,th}$ && $\mathrm{P}_{1,t} > P_{2,t}$ **then**

Compute $L_{TP,t}^{F}$ which is nearest to $L_{TP,t}^{*}$ by solving equation $P_{F,t} = P_{F,th}$ on $L_{TP,t} \in \left[L_{TP,t}^{*}, L\right]$;

Case 3: else if $P_{F,t}^{*} < P_{F,th}$ && $\mathrm{P}_{1,t} < P_{2,t}$ **then**

Compute $L_{TP,t}^{F}$ which is nearest to $L_{TP,t}^{*}$ by solving equation $P_{F,t} = P_{F,th}$ on $L_{TP,t} \in \left[0, L_{TP,t}^{*}\right]$;

end if

Output: According to Eqs. (20a)(20c), we can obtain the optimal solution $\Pi_F = \left\{L_{SP,t}^{F}, L_{TP,t}^{F}, p_{1,t}^{F}, p_{2,t}^{F}\right\}$

5. System Equilibrium Analysis with ACB

5.1 Preliminaries

The system would finally evolve to a steady state. However, there may exist more than one equilibrium point for different system load. Therefore, the aim of this section is to analyze the optimal equilibrium point to maximize the channel utility of the system when ACB is enabled. Similar to [28], we define $f_i, i = 1,2$ as one-shot success probability that a terminal send a packet successfully in a given cycle. We define $\varepsilon_i, i = 1,2$ as failure probability that a terminal fails in all the W transmission attempts. According to Eqs. (4)(5) in [28], the system balance equation can be rewritten as follows:

$$\varepsilon_i = \left(1 - f_i\right)^{W} \tag{26}$$

$$N_i f_i = \lambda_i \left(1 - \varepsilon_i\right) \tag{27}$$

Then the channel utility of the proposed hybrid protocol in steady state can be computed as:

$$\begin{aligned} R &= R_1 + R_2 \\ &= \lambda_1 D_1 \left(1 - \varepsilon_1\right) + \lambda_2 D_2 \left(1 - \varepsilon_2\right) \end{aligned} \tag{28}$$

As to the delay, we can compute as:

$$\begin{aligned} d_i &= \sum_{n=1}^{W} n \cdot \frac{\left(1 - f_i\right)^{n-1} f_i}{1 - \left(1 - f_i\right)^{W}} \\ &= \left(\frac{1}{f_i} - 1\right) \frac{1 + \left(W - 1\right)\left(1 - f_i\right)^{W} - W\left(1 - f_i\right)^{W-1}}{1 - \left(1 - f_i\right)^{W}} + 1 \end{aligned} \tag{29}$$

where $i = 1,2$ means M2M and H2H traffic respectively.

Readers may refer to [28] for more details.

5.2 Equilibrium Analysis for S-ALOHA System with ACB

For consistency with [28], we define $O = L_1 / D_1$ as the number of data channels reserved for M2M traffic. We also define the system load as $\rho := \lambda_1 / O$ and the average number of contending devices on a data channel as $z := N_1 p_1 / O$.

1) Saturate Case: If the S-ALOHA system is saturate when it reaches the steady state, we would have $z = 1$ according to Sec. 4.2. A successful terminal should first pass the ACB check and then succeed in the data channel contention. Therefore, we can have the system balance equation in the saturate case as follows:

$$\begin{aligned} f_1 &= p_1 e^{-\frac{N_1 p_1}{L_1/D_1}} = p_1 e^{-1} \\ N_1 p_1 e^{-1} &= \frac{L_1 / D_1}{e} \\ &= \lambda_1 \left[1 - (1 - f_1)^W\right] \end{aligned} \tag{30}$$

From Eq. (30), we can derive that:

$$\begin{aligned} p_1 &= e\left[1 - \left(1 - \frac{L_1 / D_1}{e\lambda_1}\right)^{1/W}\right] \\ f_1 &= 1 - \left(1 - \frac{L_1 / D_1}{e\lambda_1}\right)^{1/W} \\ \varepsilon_1 &= 1 - \frac{L_1 / D_1}{e\lambda_1} \end{aligned} \tag{31}$$

Because $0 \le p_1 \le 1$, then for large W ($W \ge 9$), we can have:

$$0 \le L_1 / D_1 \le e\lambda_1 \left[1 - \left(1 - e^{-1}\right)^W\right] \sim e\lambda_1 \tag{32}$$

2) Non-saturate Case: If the S-ALOHA system is non-saturate when it reaches the steady state, there must have $p_1 = 1, z < 1$. So we can derive the system balance equation as:

$$\begin{aligned} f_1 &= e^{-\frac{N_1}{L_1/D_1}} = e^{-z} \\ \frac{N_1}{\lambda_1} &= \frac{z}{\rho} = \frac{1 - \left(1 - e^{-z}\right)^W}{e^{-z}} \end{aligned} \tag{33}$$

We define function $h(z)$ as:

$$h(\mathrm{z}) := \mathrm{g}(\mathrm{e}^{-z}) := \frac{1 - \left(1 - e^{-z}\right)^W}{e^{-z}} \tag{34}$$

In **Fig. 2**, we depict the function $h(z)$ and its corresponding tangents. According to the analysis in Sec. V-C in [28], $h(z)$ has exactly one convex piece on $[0, z_0]$ and one concave piece on $[z_0, \infty)$. For $W \ge 9$, there are two load boundaries $0 < \rho_1 < \rho_2 < \infty$. Actually, z / ρ_1 is the tangent of $h(z)$ on $[z_0, \infty)$, and z / ρ_2 is the tangent of $h(z)$ on $[0, z_0]$. The asymptotic solution of ρ_2 is $\rho_2 \sim e^{-1}$ at $z \sim 1$:

(i) For $\rho > \rho_2$, i.e., $L_1 / D_1 < e\lambda_1$, the S-ALOHA system has only one equilibrium point on $(1, +\infty)$, which is conflict to the non-saturate assumption $z < 1$. It means that the ACB scheme would ensure that the S-ALOHA system cannot evolve to this equilibrium point on $(1,+\infty)$;

(ii) For $\rho \le \rho_2$, i.e., $L_1/D_1 \ge e\lambda_1$, in addition to the only solution on $(0,1]$, the S-ALOHA system may have more than one equilibrium point on $(1,+\infty)$. However, the ACB scheme would ensure that the S-ALOHA system can only evolve to the only equilibrium on $(0,1]$.

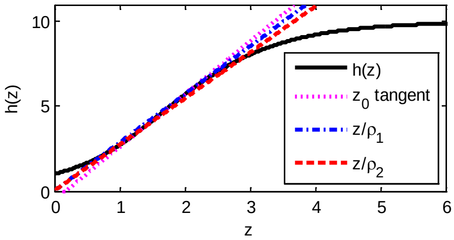


Fig. 2. Function $h(z)$ for $W=10$, z_0 is the inflexion point of $h(z)$.

Therefore, we can define the saturation threshold for S-ALOHA in steady state as $L_{1,th}^S := e\lambda_1 D_1$. We can draw a conclusion that the ACB scheme would ensure that the S-ALOHA system would finally evolve to the optimal equilibrium point:

(i) For $L_1 \ge L_{1,th}^S$, the ACB scheme could ensure that the S-ALOHA system is non-saturate when it reaches steady state. There is only one optimal equilibrium point that satisfies $z \le 1$. For the non-saturate case, we can derive that:

$$\begin{aligned} 1-\varepsilon_1 &= 1-\left(1-e^{-z}\right)^W \sim 1 \\ R_1 &= \lambda_1 D_1\left[1-\left(1-e^{-z}\right)^W\right] \sim \lambda_1 D_1 \end{aligned} \tag{35}$$

where $f_1 = e^{-z}$ can be derived from Eq. (32).

(ii) For $L_1 < L_{1,th}^S$, the S-ALOHA system will be saturate when it reaches steady state. For the saturate case, we can derive that:

$$\begin{aligned} 1-\varepsilon_1 &= \frac{L_1/D_1}{e\lambda_1} \\ R_1 &= \lambda_1 D_1\left(1-\varepsilon_1\right) = L_1/e \end{aligned} \tag{36}$$

5.3 Equilibrium Analysis for TDMA System with ACB

For TDMA system, we define $O = L_2\alpha\gamma$ as the number of preambles reserved for H2H traffic. We also define the system load as $\rho := \lambda_2/O$ and $z := N_2 p_2/O$. Following the analysis in Sec. 5.2, we can define the saturation threshold for TDMA system in steady state as $L_{2,th}^S := \frac{e\lambda_2}{\alpha^*\gamma}$. Then we can have a similar conclusion as follows:

(i) For $L_2 \ge L_{2,th}^S$, the ACB scheme can ensure that the TDMA system would evolve to the only equilibrium point on $z \in (0,1]$, where the system is non-saturate. Then we can derive that:

$$\begin{aligned} 1-\varepsilon_2 &= 1-\left(1-e^{-z}\right)^W \sim 1 \\ R_2 &= \lambda_2 D_2\left[1-\left(1-e^{-z}\right)^W\right] \sim \lambda_2 D_2 \end{aligned} \tag{37}$$

(ii) For $L_2 < L_{2,th}^S$, the TDMA system will be saturate when it reaches steady state. For the saturate case, we can derive that:

$$
\begin{aligned}
& p_2 = e\left[1-\left(1-\frac{L_2\alpha^*\gamma}{e\lambda_2}\right)^{1/W}\right] \\
& f_2 = 1-\left(1-\frac{L_2\alpha^*\gamma}{e\lambda_2}\right)^{1/W} \\
& 1-\varepsilon_2 = \frac{L_2\alpha^*\gamma}{e\lambda_2} \\
& R_2 = \lambda_2 D_2\left(1-\varepsilon_2\right) = L_2\left(1-\alpha^*\right)
\end{aligned} \tag{38}
$$

5.4 Equilibrium Analysis for the Proposed Hybrid Protocol

Now we can formulate the channel utility maximization problem for the proposed hybrid protocol in steady state as :

$$
\begin{aligned}
& \max_{L_{SP}, L_{TP}, p_1, p_2} R = \begin{cases} L_{SP}/e + L_{TP}(1-\alpha^*), & if\ L_{SP} < L_{1,th}^S, L_{TP} < L_{2,th}^S \\ L_{SP}/e + \lambda_2 D_2, & if\ L_{SP} < L_{1,th}^S, L_{TP} \ge L_{2,th}^S \\ \lambda_1 D_1 + L_{TP}(1-\alpha^*), & if\ L_{SP} \ge L_{1,th}^S, L_{TP} < L_{2,th}^S \\ \lambda_1 D_1 + \lambda_2 D_2, & if\ L_{SP} \ge L_{1,th}^S, L_{TP} \ge L_{2,th}^S \end{cases} \\
& s.t. \quad (a)\quad L = L_{SP} + L_{TP}, \quad L_{SP}, L_{TP} \in \left[0, L\right] \\
& \qquad\quad \text{(b)}\quad P_F \ge P_{F,th}
\end{aligned} \tag{39}
$$

where P_F can be computed according to Eqs. (7)(35)(36)(37)(38).

Obviously, the channel utility R is monotonically increasing on $L_{TP} \in \left[0, L_{2,th}^S\right]$, and is monotonically decreasing on $L_{TP} \in \left[L - L_{1,th}^S, L\right]$, and would be stable on $L_{TP} \in \left[L_{2,th}^S, L - L_{1,th}^S\right]$ if $L_{1,th}^S + L_{2,th}^S < L$. Hence, in this paper, we define $\Pi^S = \left\{L_{SP}^*, L_{TP}^*, p_1^*, p_2^*\right\}$ as optimal solution to Eq. (39) without fairness constraint as follows:

$$
L_{TP}^* = \begin{cases} L_{2,th}^S, & if\ L_{1,th}^S + L_{2,th}^S \ge L \\ L_{2,th}^S + \dfrac{\lambda_2 D_2}{\lambda_1 D_1 + \lambda_2 D_2}\left(L - L_{1,th}^S - L_{2,th}^S\right), & if\ L_{1,th}^S + L_{2,th}^S < L \end{cases} \tag{40}
$$

Now let's consider the fairness constraint. Similar to Algorithm 2, the optimal solution $\Pi_F^S = \left\{L_{SP}^F, L_{TP}^F, p_1^F, p_2^F\right\}$ to Eq. (39) can be determined by Algorithm 3:

Algorithm 3: Optimal equilibrium point for the proposed hybrid protocol

---

Step 1: Obtain the solution $\Pi^S = \left\{L_{SP}^*, L_{TP}^*, p_1^*, p_2^*\right\}$ according to Eq. (40) and compute the fairness index P_F^*.

Step 2: Obtain the optimal solution that satisfy the fairness constraint in different cases.

Case 1: if $P_F^* \ge P_{F,th}$ **then**

$\Pi_F^S = \Pi^S$

Case 2: else if $P_F^* < P_{F,th}$ && $1-\varepsilon_1 > 1-\varepsilon_2$ **then**

Compute L_{TP}^F which is nearest to L_{TP}^* by solving equation $P_F = P_{F,th}$ on $L_{TP} \in \left[L_{TP}^*, L\right]$;

Case 3: else if $P_F^* < P_{F,th}$ && $1-\varepsilon_1 < 1-\varepsilon_2$ **then**

Compute L_{TP}^F which is nearest to L_{TP}^* by solving equation $P_F = P_{F,th}$ on $L_{TP} \in \left[0, L_{TP}^*\right]$;

---

end if

Output: According to Eqs. (31)(38) (39a), we can obtain the optimal solution $\Pi_F^S = \left\{L_{SP}^F, L_{TP}^F, p_1^F, p_2^F\right\}$

6. Numerical and Simulation Results

6.1 Parameters Setting

We consider a TDD LTE/LTE-A cell with 3MHz bandwidth and 10ms transmission cycle. In each cylce, the reserved RBs for uplink transmission is $L=120$. In order to obtain the performance results of the system in steady state, the simulation time is 500 cycles (i.e., 5s). We set the transmission attempts limit $W=10$.

We assume that the resources in pure S-ALOHA and pure TDMA protocol are also divided into two parts for M2M and H2H traffic respectively. The optimal resource allocation and ACB parameters can be derived according to Sec. 4 easily. The optimal number of resource allocation in simulation is the integer nearest to the continuous solution obtained in Sec. 4.

For brevity, in the following subsections, we will let 'EA' denote numerical results obtained by equilibrium analysis, let 'num' denote numerical results obtained by analytical analysis, and let 'sim' denote results obtained by simulation.

6.2 Performance of the Proposed Hybrid Protocol without Fairness Constraint

Let traffic ratio $\lambda_1:\lambda_2=1:1$, the channel utility of the proposed hybrid protocol without fairness constraint are shown in **Fig. 3**. The EA curve is divided into three stages by two inflexion points which correspond to λ_B,λ_D in X axis. While the numerical curve is divided into four stages by three inflexion points which correspond to $\lambda_A,\lambda_C,\lambda_D$.

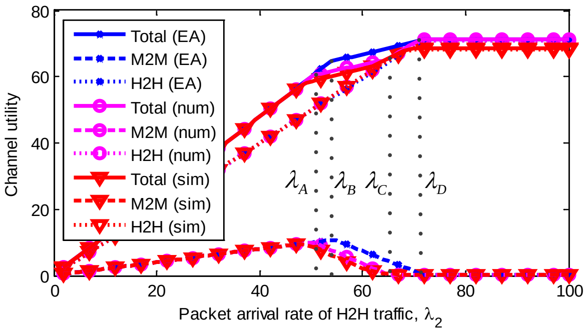


Fig. 3. Channel utility of the hybrid protocol without fairness constraint ($D_1=0.2$, $D_2=1$, $\gamma=4$).

Obviously, the total channel utility in the EA scenario is higher than that in the numerical scenario when $\lambda_2\in(\lambda_A,\lambda_D)$. The reason can be explained as follows. The TDMA part in the numerical scenario is non-saturate when $\lambda_2\in(\lambda_A,\lambda_D)$, while the TDMA part in the EA scenario is non-saturate when $\lambda_2\in(\lambda_A,\lambda_B)$ and is exactly saturate when $\lambda_2\in(\lambda_B,\lambda_D)$. Thus, according to Eq. (37), almost all the H2H UEs would succeed in both the numerical and EA scenarios, as shown in **Fig. 4(a)**. So the channel utility of the H2H traffic are the same as shown in **Fig. 3**. On the other hand, due to the proposed cycle-by-cycle resource allocation scheme, the number of RBs occupied by H2H traffic in the numerical scenario is larger than that in the EA scenario, while the number of RBs occupied by M2M traffic is just the reverse as shown in **Fig. 5**. Hence, the channel utility of the M2M traffic in the EA scenario is higher than that in the numerical scenario when $\lambda_2 \in (\lambda_A, \lambda_D)$. As a result, the total channel utility in the EA scenario is higher than that in the numerical scenario.

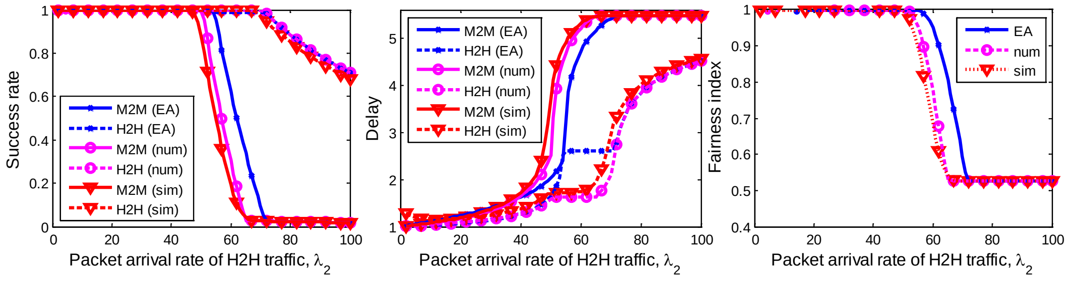


Fig. 4. (a) Success rate (b) Delay (c) Fairness index ($D_1 = 0.2$, $D_2 = 1$, $\gamma = 4$).

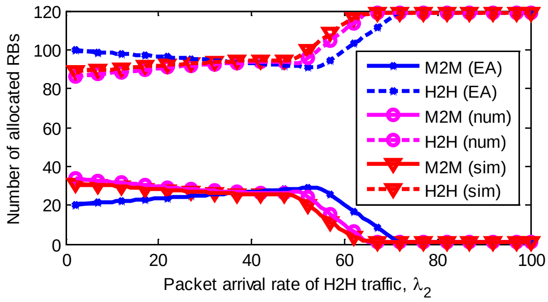


Fig. 5. Resource allocation for M2M and H2H traffic ($D_1 = 0.2$, $D_2 = 1$, $\gamma = 4$).

When $\lambda_2 < \lambda_A$, both the S-ALOHA part and the TDMA part in these two scenarios are non-saturate. Hence, as shown in **Fig. 4(a)**, almost all the terminals would succeed at last. Thus, the channel utility in the EA scenario and numerical scenario are the same when $\lambda_2 < \lambda_A$. When $\lambda_A > \lambda_D$, all the RBs would be allocated to the H2H traffic and the TDMA part is saturate in both the EA scenario and numerical scenario. So the total channel utility in these two scenarios are also the same when $\lambda_A > \lambda_D$.

In summary, the equilibrium analysis in Sec. 5 provides an upper bound on channel utility. As shown in **Fig. 4**, the success rate of M2M traffic decreases dramatically with λ_2 when the system load is heavy. And the fairness index decreases dramatically too. Therefore, in order to satisfy the service requirements of M2M and H2H traffic simultaneously, fairness has to be considered in the MAC protocol design.

As we can see from these figures, the simulation results match with the numerical analysis and the difference between them could be acceptable. The main reason for the difference is that the data channels reserved in TDMA part is either surplus or insufficient due to the randomness in the random access procedure, i.e., the approximation error in Eq. (11).

6.3 Performance of the Proposed Hybrid Protocol with Fairness Constraint

In this subsection, we would compare the performance of the proposed hybrid protocol with and without fairness constraint through simulation results. We let 'w/o' denote the results without fairness constraint and let 'w' denote the results with fairness constraint.

As shown in **Fig. 6**, the channel utility curves with fairness constraint are also divided into four stages by corresponding inflexion points. In the first two stages, the system load is relatively low, we can find in **Fig. 7(c)** that $P_F > P_{F,th}$. Therefore, the optimal solutions and performance results for these two cases with and without fairness constraint are the same.

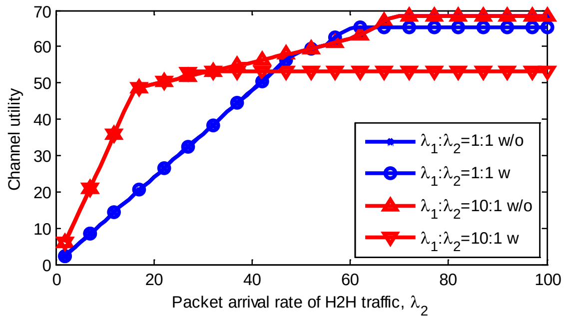


Fig. 6. Channel utility for the proposed hybrid protocol ($D_1 = 0.2$, $D_2 = 1$, $\gamma = 4$).

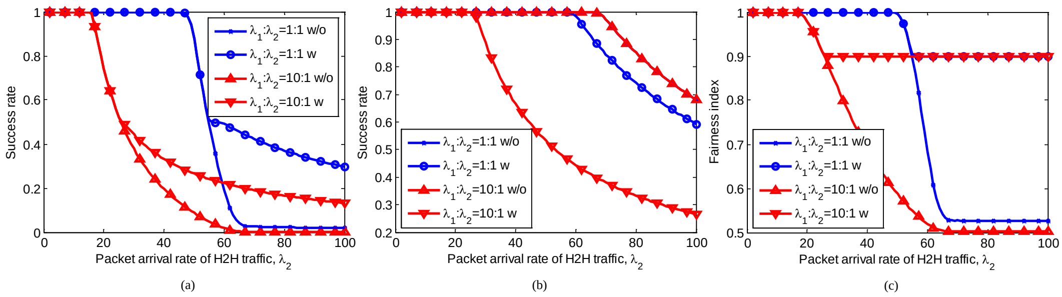


Fig. 7. (a) Success rate of M2M traffic (b) Success rate of H2H traffic (c) Fairness index ($D_1 = 0.2$, $D_2 = 1$, $\gamma = 4$).

In the third stage shown in **Fig. 6**, comparing with the case without fairness constraint, the channel utility with fairness constraint is higher. The reason can be explained as follows. In order to satisfy the fairness constraint, the hybrid protocol would allocate more RBs to M2M traffic. Whereas the TDMA part is still non-saturate. Hence, comparing with the case without fairness constraint, in the case with fairness constraint, the channel utility of the M2M traffic is higher while the channel utility of H2H traffic is the same.

In the last stage, both the S-ALOHA part and TDMA part are saturate. As shown in **Fig. 6**, the channel utility without fairness constraint is higher than that with fairness constraint.

As shown in **Fig. 7**, comparing with the case without fairness constraint, the performance of M2M traffic has been improved and the fairness has been guaranteed when the fairness constraint is considered. However, the cost is the performance decline of the H2H traffic.

We also increase the traffic ratio $\lambda_1 : \lambda_2$ from 1:1 to 10:1, where $\lambda_1 : \lambda_2 = 10:1$ denote the ultra dense deployment of M2M devices in future. In **Fig. 7(b)**, we can find that the increasement of M2M traffic load has no influence on the performance of H2H traffic in the case without fairness constraint. Then considering fairness, we can find that the performance of both the M2M and H2H traffic has degraded. However, the performance loss is much more prominent for H2H traffic.

6.4 Performance Comparison

In this subsection, we would compare the performance of the proposed hybrid protocol with pure TDMA protocol and pure S-ALOHA protocol. All the results are obtained by simulations in the case with fairness constraint.

Fig. 8 shows that the capacity of the hybrid protocol is higher than that of pure S-ALOHA and pure TDMA protocols. At the point $\lambda_2 = 100$, the channel utility in hybrid protocol is 9.44% higher than that in pure TDMA protocol and 47.57% higher than that in pure S-ALOHA protocol when $\lambda_1 : \lambda_2 = 1:1$. And the percentage is 40.33% and 20.53% respectively when $\lambda_1 : \lambda_2 = 10:1$. With the growth of traffic ratio $\lambda_1 : \lambda_2$, the capacity of all these three protocols would decrease. However, the capacity of pure TDMA protocol is even lower than that of pure S-ALOHA protocol when $\lambda_1 : \lambda_2 = 10:1$. It shows that current pure TDMA protocol is not suitable for future dense deployment of M2M devices.

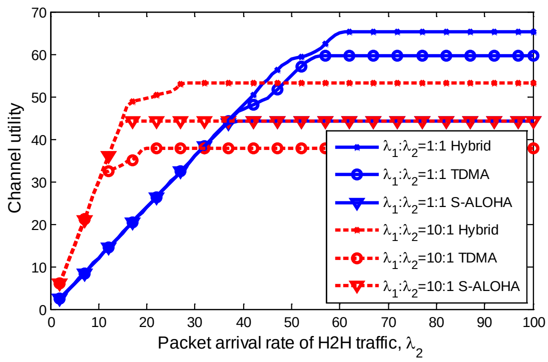


Fig. 8. Channel utility for different MAC protocols ($D_1 = 0.2$, $D_2 = 1$, $\gamma = 4$).

The delay and success rate of M2M traffic are shown in **Fig. 9**. We can find that the hybrid protocol is superior to both the pure S-ALOHA and pure TDMA protocol for small λ_2. On the other hand, when λ_2 is large, the hybrid protocol is better than pure TDMA protocol but is worse than the pure S-ALOHA protocol. When traffic ratio $\lambda_1 : \lambda_2$ increases from 1:1 to 10:1, the performance of hybrid protocol is closer to that of pure S-ALOHA protocol.

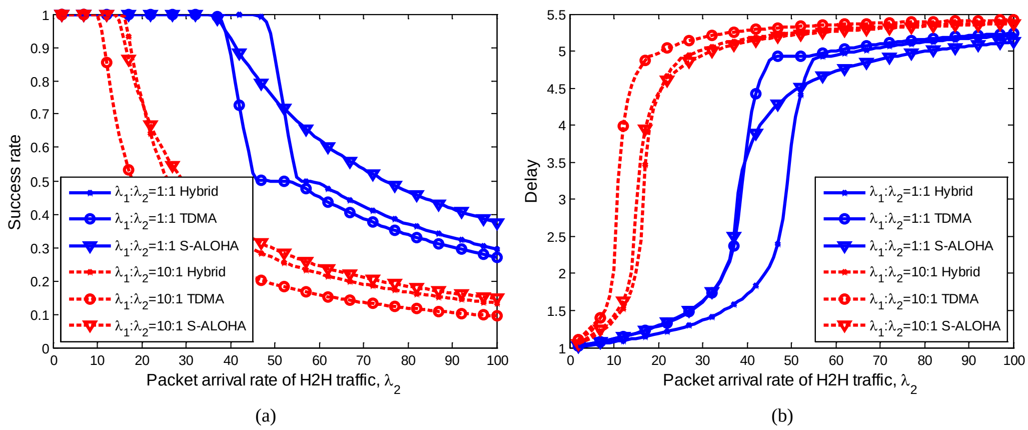
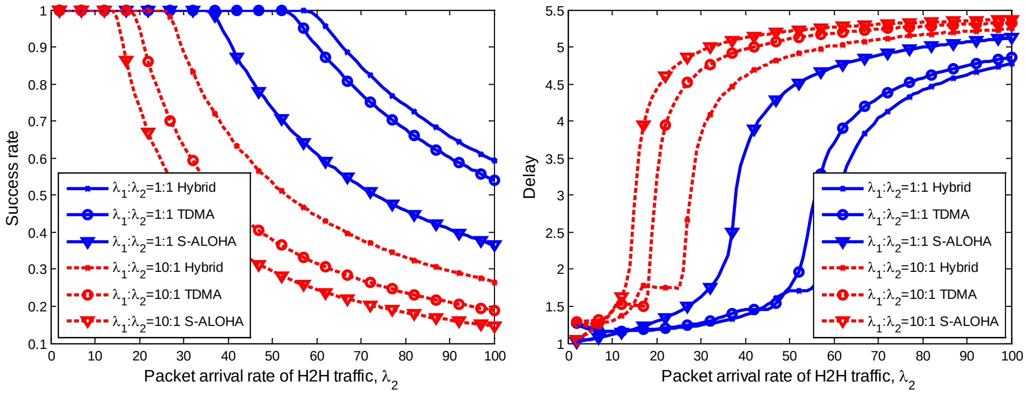


Fig. 9. (a) Success rate of M2M traffic (b) Delay of M2M traffic ($D_1 = 0.2$, $D_2 = 1$, $\gamma = 4$).

As to the delay and success rate of H2H traffic shown in **Fig. 10**, we can find that the hybrid protocol is almost always superior to both the pure S-ALOHA protocol and pure TDMA protocol. Though TDMA protocol is more suitable for lager-size H2H traffic than S-ALOHA protocol, due to the approximation error in Eq. (11), the delay of H2H traffic in the hybrid protocol is higher than that in the pure S-ALOHA protocol for very small λ_2.



(a) (b)

Fig. 10. (a) Success rate of H2H traffic (b) Delay of H2H traffic ($D_1 = 0.2$, $D_2 = 1$, $\gamma = 4$).

As to the fairness index, **Fig. 11** demonstrates that all these three protocols satisfy the fairness constraint, which proves the validity of Algorithm 2. Because the packet size has no impact on capacity of the S-ALOHA system, so the fairness index of the pure S-ALOHA protocol is the highest.

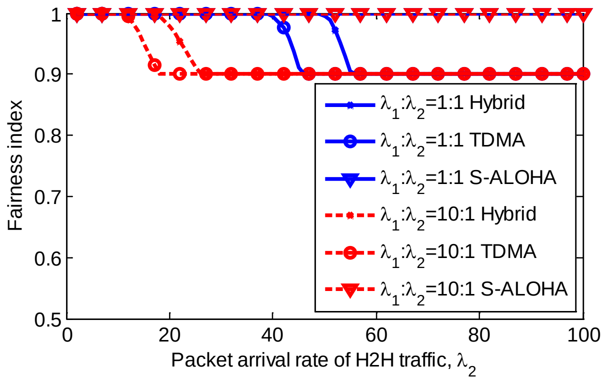


Fig. 11. Fairness index for different MAC protocols ($D_1 = 0.2$, $D_2 = 1$, $\gamma = 4$).

As shown in the above figures, we can conclude that the proposed hybrid protocol is always better than pure TDMA protocol for both M2M and H2H traffic. Comparing with pure S-ALOHA protocol, we can find that the proposed hybrid protocol is always better for H2H traffic, but would result in relatively low performance loss in M2M traffic when system load is high. When the traffic ratio $\lambda_1 : \lambda_2$ increases, we can find that both the delay and success rate metrics of all these three protocols would deteriorate. In summary, the proposed hybrid protocol would behave more like pure TDMA protocol when M2M traffic load is low and more like pure S-ALOHA protocol when M2M traffic load is high.

7. Conclusion

In this paper, we propose a novel S-ALOHA and TDMA hybrid MAC protocol to address the signaling overhead and resource underutilization problems when massive M2M devices coexist with H2H UEs. Without considering the backoff scheme, we derive a joint resource allocation and ACB optimization algorithm to maximize the channel utility and satisfy the fairness constraint. The simulation results demonstrate that the capacity of the proposed hybrid protocol is 9.44% higher than that of pure TDMA protocol and 47.57% higher than that of pure S-ALOHA protocol when $\lambda_1 : \lambda_2 = 1:1$. And the percentages are 40.33% and 20.53% respectively when $\lambda_1 : \lambda_2 = 10:1$. Moreover, a comprehensive performance analysis in terms of success rate, delay, and fairness has been conducted. We can draw a conclusion that the proposed hybrid protocol is always superior to the pure TDMA protocol, especially in the future dense M2M deployment scenario. However, comparing with the results of equilibrium analysis, it shows that the proposed cycle-by-cycle resource allocation and ACB control algorithm is suboptimal. Therefore, in future work, we aim at deriving a more sophisticated algorithm to achieve the upper performance bound, which is obtained by the equilibrium analysis. We would also incorporate the backoff scheme into our proposed hybrid protocol.

References

[1] F. Boccardi, R. W. Heath Jr., A. Lozano, et al., "Five Disruptive Technology Directions for 5G," *IEEE Communications Magazine*, pp. 74-80, Feb. 2014. Article (CrossRef Link)

[2] W. H. Chin, Z. Fan, R. Haines, "Emerging Technologies and Research Challenges for 5G Wireless Networks," *IEEE Wireless Communications*, pp. 106-112, Apr. 2014. Article (CrossRef Link)

[3] K. Chang, A. Soong, M. Tseng, et al., "Global Wireless Machine-to-Machine Standardization," *IEEE Internet Computing*, vol. 15, no. 2, pp. 64-69, March/April 2011. Article (CrossRef Link)

[4] Fayezeh Ghavimi, Hsiao-Hwa Chen, "M2M Communications in 3GPP LTE/LTE-A Networks: Architectures, Service Requirements, Challenges, and Applications," *IEEE Communication Surveys and Tutorials*, vol. 17, no. 2, pp. 525-549, Second Quarter, 2015. Article (CrossRef Link)

[5] 3GPP TS 22.368 V11.0.1, "Service Requirements for Machine-Type Communications," Feb. 2011. Article (CrossRef Link)

[6] M. Zubair Shafiq, Lusheng Ji, Alex X. Liu, et al., "Large-Scale Measurement and Characterization of Cellular Machine-to-Machine Traffic," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1960-1973, Dec. 2013. Article (CrossRef Link)

[7] Ericsson, "More than 50 billion connected devices," Feb. 2011. Article (CrossRef Link)

[8] Kan Zheng, Suling Qu, Jesus Alonso-Zarate, et al., "Challenges of Massive Access in Highly Dense LTE-Advanced Networks With Machine-To-Machine Communications," *IEEE Wireless Communications*, pp. 12-18, June 2014. Article (CrossRef Link)

[9] Andres Laya, Luis Alonso, Jesus Alonso-Zarate, "Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 4-16, First Quarter 2014. Article (CrossRef Link)

[10] Ajinkya Rajandekar, Biplab Sikdar, "A Survey of MAC Layer Issues and Protocols for Machine-to-Machine Communications," *IEEE Internet of Things Journal*, vol. 2, no. 2, pp. 175-186, Apr. 2015. Article (CrossRef Link)

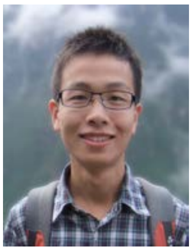
[11] Harpreet S. Dhillon, Howard Huang, Harish Viswanathan, et al., "Fundamentals of Throughput Maximization with Random Arrivals for M2M Communications," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 4094-4109, Nov. 2014. Article (CrossRef Link)

[12] Rapeepat Ratasuk, Athul Prasad, Zexian Li, et al., "Recent Advancements in M2M Communications in 4G Networks and Evolution Towards 5G," in *Proc. of 18th Int. Conf. on Intelligence in Next Generation Networks*, pp. 52-57, 2015. Article (CrossRef Link)

[13] Min Chen, Jiafu Wan, Sergio Gonz´alez, et al., "A Survey of Recent Developments in Home M2M Networks," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 98-114, First Quarter 2014. Article (CrossRef Link)

[14] Jaewoo Kim, Jaiyong Lee, Jaeho Kim, et al., "M2M Service Platforms: Survey, Issues, and Enabling Technologies," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 61-76, First Quarter 2014. Article (CrossRef Link)

[15] Mohammad Tauhidul Islam, Abd-Elhamid M. Taha, Selim Akl, "A Survey of Access Management Techniques in Machine Type Communications," *IEEE Communications Magazine*, pp. 74-81, April 2014. Article (CrossRef Link)
[16] Peng Si, Jian Yang, Shuangwu Chen, et al., "Adaptive Massive Access Management for QoS Guarantees in M2M Communications," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 7, pp. 3152-3166, July 2015. Article (CrossRef Link)
[17] Tzu-Ming Lin, Chia-Han Lee, Jen-Po Cheng, et al., "PRADA: Prioritized Random Access With Dynamic Access Barring for MTC in 3GPP LTE-A Networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2467-2472, June 2014. Article (CrossRef Link)
[18] German Corrales Madueno, Cedomir Stefanovic, Petar Popovski, "Efficient LTE Access with Collision Resolution for Massive M2M Communications," in *Proc. of Globecom Workshop*, pp. 1433-1438, 2014. Article (CrossRef Link)
[19] Taehoon Kim, Kab Seok Ko, Inkyu Bang, et al., "A Random Access Scheme Based on a Special Preamble for Supporting Emergency Alarms," in *Proc. of IEEE WCNC*, pp. 1768-1773, 2014. Article (CrossRef Link)
[20] Minho Jo, Taras Maksymyuk, Bohdan Strykhalyuk, et. al, "Device-to-Device Based Heterogeneous Radio Access Network Architecture for Mobile Cloud Computing," *IEEE Wireless Communications*, pp. 50-58, June 2015. Article (CrossRef Link)
[21] A. Pourmoghadas, P. G. Poonacha, "Performance Analysis of a Machine-to-Machine Friendly MAC Algorithm in LTE-Advanced," in *Proc. of ICACCI*, pp. 99-105, 2014. Article (CrossRef Link)
[22] T. P. C. de Andrade, C. A. Astudillo, N. L. S. da Fonseca, "The Impact of Massive Machine Type Communication Devices on the Access Probability of Human-to-Human Users in LTE Networks," in *Proc. of IEEE Latin-America Conference*, pp. 1-6, 2014. Article (CrossRef Link)
[23] Kan Zheng, Fanglong Hu, Wenbo Wang, et al., "Radio Resource Allocation in LTE-Advanced Cellular Networks with M2M Communications," *IEEE Communications Magazine*, pp. 184-192, July 2012. Article (CrossRef Link)
[24] Martin Levesque, Frank Aurzada, Martin Maier, et al., "Coexistence Analysis of H2H and M2M Traffic in FiWi Smart Grid Communications Infrastructures Based on Multi-Tier Business Models," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 3931-3942, Nov. 2014. Article (CrossRef Link)
[25] Yi Liu, Chau Yuen, Xianghui Cao, et al., "Design of a Scalable Hybrid MAC Protocol for Heterogeneous M2M Networks," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 98-111, Feb. 2014. Article (CrossRef Link)
[26] Bharat shrestha, Ekram Hossain, Kae Won Choi, "Distributed and Centralized Hybrid CSMA/CA-TDMA Schemes for Single-Hop Wireless Networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 4050-4065, July 2014. Article (CrossRef Link)
[27] Dimas Tribudi Wiriaatmadja, Kae Won Choi, "Hybrid Random Access and Data Transmission Protocol for Machine-to-Machine Communications in Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 33-46, Jan. 2015. Article (CrossRef Link)
[28] Revak R. Tyagi, Frank Aurzada, Ki-Dong Lee, Sang G. Kim, Martin Reisslein "Impact of Retransmission Limit on Preamble Contention in LTE-Advanced Network," *IEEE Systems Journal*, vol. 9, no. 3, pp. 752-765, Sep. 2015. Article (CrossRef Link)
[29] Feilong Lin, Cailian Chen, Ning Zhang, et. al, "Autonomous Channel Switching: Towards Efficient Spectrum Sharing for Industrial Wireless Sensor Networks," *IEEE Internet of Things Journal*, vol. 3, no. 2, pp. 231-243, April 2016. Article (CrossRef Link)
[30] Chia-Hung Wei, Ray-Guang Cheng, Shiao-Li Tsao, "Modeling and Estimation of One-Shot Random Access for Finite-User Multichannel Slotted ALOHA Systems," *IEEE Communications Letters*, vol. 16, no. 8, pp. 1196-1199, Aug. 2012. Article (CrossRef Link)
[31] Chia-Hung Wei, Ray-Guang Cheng, Shiao-Li Tsao, "Performance Analysis of Group Paging for Machine-Type Communications in LTE Networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3371-3382, Sept. 2013. Article (CrossRef Link)



Nannan Sui received his B.S. degree in communications engineering from Tianjin University (TJU), Tianjin, China, in 2011 and M.S. degree in signal and information processing from college of communications engineering, PLA University of Science and Technology (PLAUST), Nanjing, China, in 2014, respectively. He is currently a PhD candidate in college of communications engineering, PLAUST. His research interests include machine type communications, ultra dense networking, heterogeneous networks, and radio resource management.

Cong Wang received his Ph.D. degree from PLA University of Science and Technology (PLAUST), China in 2004. Now he is an associate professor at PLAUST. His research interests include computer networks and wireless communications.

Wei Xie received B.S. degree in mobile communication engineering in 1999 from Nanjing Institute of Communication Engineering (NICE) and M.S. degree in communication and information systems in 2002 from PLA University of Science and Technology (PLAUST). He is now an associate Professor of PLAUST. His research interests include mobile communications and systems, wireless communication networks, cooperative communications， cognitive networks and wireless sensor networks.

Youyun Xu graduated from Shanghai Jiao Tong University with a Ph.D. degree in information and communication engineering in 1999. He is currently a professor with Nanjing University of Posts and Telecommunications, China. Now, his research interests are focusing on new generation wireless mobile communication systems (LTE-A and 5G), advanced channel coding and modulation techniques, multi-user information theory and radio resource management, wireless sensor networks, cognitive radio networks, etc.