Fair Configuration Scheme for Random Access in NB-IoT with Multiple Coverage Enhancement Levels

Ruki Harwahyu, *Member, IEEE*, Ray-Guang Cheng, *Senior Member, IEEE*, Da-Hao Liu, and Riri Fitri Sari, *Senior Member, IEEE*

Abstract—Narrowband Internet of Things (NB-IoT) is a new technology developed to support low-power wide area networks (LPWAN) services. To extend its coverage and decrease its transmission power, devices in one NB-IoT cell are divided into several coverage enhancement (CE) levels with different random access configuration. This potentially results in unfair access, especially when massive number of devices in different CE levels simultaneously accessing the network. This work presents an effective strategy to configure the random access in NB-IoT to yield fair performance across CE levels. An analytical model is used to estimate the performance of each CE level and overall system performance in term of normalized throughput and average access delay. Simulation is incorporated to verify the accuracy of the model. Different practical assumptions of fair system are explored and examined in the experiment. The result shows that the analytical model is accurate under various loads. Additionally, the proposed search strategy is proven to be able to obtain the configuration which yield acceptable throughput fairly for all CE levels.

Index Terms—NB-IoT, random access, access fairness

1 Introduction

Narrowband Internet of Things is a relatively new 3GPP standard aiming to support low-cost, low-power, and low data rate devices. It is designed for wider or deeper underground coverage, which is 20dB improvement over GPRS. NB-IoT adopts the existing LTE-A OFDMA technology while only requiring 180 kHz bandwidth [1], [2]. Due to its extended coverage, IoT devices in one NB-IoT cell have a wide variation of signal quality. To deal with it, NB-IoT divides those devices in one cell into several coverage enhancement levels. Each CE level corresponds to a certain number of repetitions with frequency hopping to ensure good reception. This is one of the main causes for unfair access which will be discussed later.

Similar to LTE, NB-IoT devices that need to transmit their data needs to conduct random access (RA) procedure, which is a 4-message handshake. First, an 'anonymous' signal sequence, called preamble, is sent by the backlogged device (s). Base station then 'announces' the detected preamble(s) in a message called RA reply (RAR). Device(s) whose preamble is detected transmit their unique connection request in Msg3. Base station then replies the correctly-received Msg3 with Msg4, allowing the respective device to conduct data transmission. More details of RA procedure are easily obtained in existing literature such as [3], [4].

In NB-IoT, preamble is transmitted in special time-frequency resources called Narrowband Physical Random Access CHannel (NPRACH) which is allocated periodically by base station. The interval and duration of NPRACH are set according to the duration of the repeated transmission, and are known by the devices via periodically-broadcasted Message Information Block – 2 (MIB2).

An NB-IoT cell can be configured with one to three CE levels. Each CE level is associated with a specific range of reference signal received power (RSRP). Devices are divided into the three CE levels based on their RSRP, where devices in the lower CE level have higher RSRP while devices in the higher CE level have the lower RSRP. Base station allocates distinct NPRACH for each CE level. Consequently, NPRACH of the higher CE level has longer duration than that of the lower one. To serve up to three NPRACH for the three CE levels, base station divides the 180 kHz bandwidth into 48 sub-carriers, each with a sub-carrier spacing of 3.75 kHz. The basic unit of the sub-carrier allocation for an NPRACH of one CE level is 12 sub-carriers. Therefore, an NPRACH for a CE level can have 12, 24, 36, or 48 sub-carriers [2].

A device starts a random access procedure by transmitting a preamble in an NPRACH at its initial CE level where it resides. In NB-IoT, there is only one sequence as the preamble, which is to be used by all devices. A preamble consists of four symbol groups. Each symbol group contains one cyclic prefix plus five symbols. For a transmission attempt, the same preamble is repeatedly transmitted within the same NPRACH. A base station chooses a proper number of repetitions for different CE levels to ensure good signal quality at

- *R. Harwahyu and R. F. Sari are with the Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia, Depok, Java 16424, Indonesia. E-mail: {ruki.h, riri}@ui.ac.id.*
- *R.-G. Cheng and D.-H. Liu are with the Department of Electronics and Computer Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan. E-mail: {crg, m10402267}@mail.ntust.edu.tw.*

Manuscript received 29 Jan. 2019; revised 17 Oct. 2019; accepted 21 Dec. 2019. Date of publication 25 Dec. 2019; date of current version 4 Mar. 2021. (Corresponding author: R. Harwahyu.)
Digital Object Identifier no. 10.1109/TMC.2019.2962422

the receiver. The first repetition is transmitted at a sub-carrier randomly chosen from the sub-carriers allocated to the CE level. Collision occurs if two or more devices transmitting at the same NPRACH choose the same initial sub-carrier. Collided devices must retransmit at a newly chosen initial sub-carrier in the next available NPRACH after backoff. The process repeats until the maximum number of attempts in this CE level is reached. The failed device can restart the whole RA procedure in the next higher CE level. The device will declare a RA failure if the RA procedure fails at the highest CE level [2].

1.1 Related Works

In principle, RA in NB-IoT is similar to Multi-channel Slotted Aloha. The existence of multiple CE levels in NPRACH can be modeled as multiple Multi-channel Slotted Aloha protocols that operate in parallel. In a Multi-Channel Slotted Aloha, concurrent access from a massive number of devices may significantly decrease the performance [5]. Massive access by IoT devices is likely to happen in various IoT applications due to synchronized periodic reporting [6], synchronized paging [7], or simultaneous emergency reporting [8]. Combined with a limited number of sub-carriers in NB-IoT, the devices need to retransmit more attempts in order to be successful (causing longer access delay), more energy are wasted in the collisions and collision resolutions, and the actual number of devices that can utilize the resource for data transmissions is very limited. Additionally, due to different RA configuration in each CE level, which is made to ensure good signal quality at different RSRP, devices in different CE level may have unequal chances to access the network successfully. This unfair performance among CE levels is contra productive to NB-IoT's support for extended coverage.

Study about access performance of the Multi-channel Slotted Aloha network and its application in the RA procedure has been extensively conducted for its normalized throughput [9], [10], [11] and mean access delay [11], [12], [13], [14], [15]. The trade-off between the normalized throughput and mean access delay was studied in [15]. A Poisson approximation model to estimate the number of success and collided devices in the first RA slot was presented in [5]. The congestion of the RA can be relieved by reducing the instantaneous load of the requests [16] at the expense of a degraded quality-of-service for some lower priority services [17], [18], or reduce their probability to transmit the preamble by dynamically updating the relevant access parameters [19].

Another challenge of serving IoT application with cellular system is efficiency. Although it does not directly impacting the contention during medium access, it is also worth to mention that the signaling process normally used in cellular system is considerably large for small IoT data. A study in [20] has summarized four approaches for this inefficiency problem in LTE systems and its derivatives such as NB-IoT. In this regard, NB-IoT can operate in connectionless mode, where the data is sent short after the completion of RA procedure on the active default data bearer.

To further decrease the RA contention, coded RA is introduced in [21]. It mandates each device to randomly choose several preambles to be sent in series in one attempt. This virtually increases the number of preambles and thus increase the chance of devices to be successful since collision is occurs when more than one device choose exactly the same preambles consecutively in one attempt. However it also increases transmission duration and power. A non-contention approach was introduced in [22] which favor queueing over random backoff for the retransmissions. With such queueing mechanism, the collided devices are split into smaller independent queues with its own RA resource (hence, collision is impossible). Each device has to maintain its local counter to know their position in the queue. This mechanism is adopted by [23] to provide its LTE implementation. With this approach, the queue may grow quickly under heavy loading of contending device, which yield higher access delay for successful devices.

From MAC layer perspective, i.e., where the contention performance is studied, the existence of multiple CE levels is comparable to multiple RACHs operating in parallel. Although the working principle of its RA procedure is not much different from that of LTE (again from MAC layer perspective), the fairness issue arises among devices in different CE levels, calling more research and exploration.

A fairness-aware algorithm was proposed to maximize the throughput in 4G OFDMA systems [24], in which the trade-off among system throughput and fairness maximization was relaxed by virtual physical resource block. This trade-off was explored in [25] and a sub channel allocation algorithm was subsequently proposed to mitigate interference for OFDMA cell. Small cells densification [26] and drone-aided cells [27] are studied, which are rather straight forward solution for improving fairness among subscribers, but less practical for NB-IoT whose each cell is designed with CE in mind. Fairness is also concerned in resource allocation and interference management for Device-to-Device (D2D) communication [28]. A study in [29] enhances the fairness of Non-Orthogonal Multiple Access (NOMA) system with improved transmit power allocation.

Our previous work in [4] pointed that excessive repetitions of preamble transmission in NB-IoT may not be necessary. We demonstrated that we can trade extra time-diversity gain obtained from retransmissions to reduce the number of repetitions [4]. In [3], we presented a heuristic tool to configure multiple RA parameters to maximize the access probability for the whole cell (all CE levels) under a given access delay constraint in NB-IoT with multiple CE levels [3]. In this paper, however, we focus on the fairness among multiple CE levels in NB-IoT. The following subsection outlines our contribution in this regard.

1.2 Our Contribution

Normally, subscribers expect to receive similar service quality regardless of their signal quality. Devices in cell edge always experience lower access success probability. Additionally, when they do succeed, they need more time to transmit the same amount of data compared to devices with better signal quality (due to repetition). Please also keep in mind that to be successful, they need to resolve the collision by several times of retransmission. Finally, when more devices are in poorer signal quality, the resource efficiency is decreased (e.g., fewer byte of data can be transmitted in a resource block).

In this paper, we study the fairness of the normalized throughput for the devices attempting to access an NB-IoT

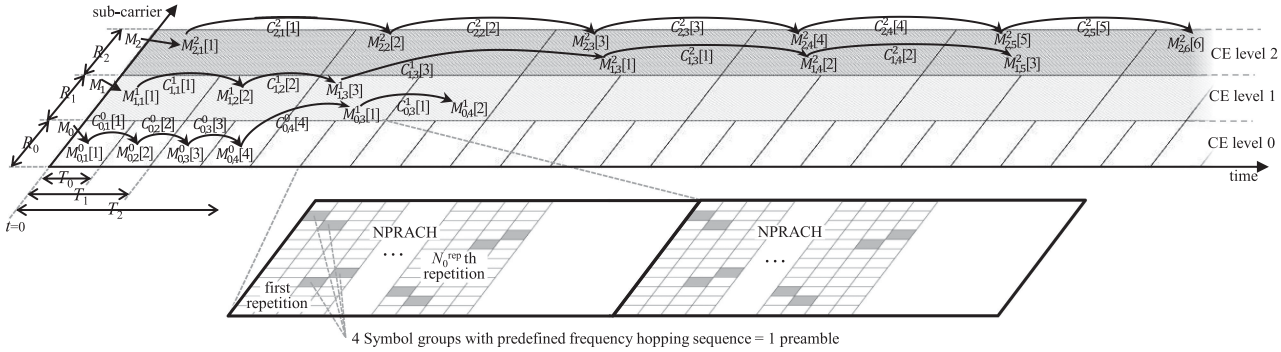

Fig. 1. Propagation of preamble transmission attempts (numerologies are for simplicity and the ease of observation only, which is not always the case in the real world).

network supporting multiple CE levels. Additionally, the performance in term of average access delay is also evaluated. Normalized throughput is calculated as the total amount of data packets successfully transmitted divided by the amount of all backlogged data packets. This is effectively similar to RA access success probability [2] when each data transmission requires RA procedure and the data sent by participating IoT devices having the same size (which is reasonable for fairness analysis). Meanwhile, average access delay is calculated as total time spent by all data packet that is successfully transmitted divided by number of data packets successfully transmitted. First we derive an analytical model to estimate the normalized throughput and average access delay of each CE levels by considering the maximum number of preamble transmissions in each CE level, backoff window in each CE level, and number of sub-carriers in each CE level. These parameters provide flexible yet not-straightforward and rather complex behavior of NPRACH [3]. We then adopt the Jain's fairness index [30] to quantify the fairness of multiple configurations. Finally, we propose a strategy in configuring the RA parameters to maximize the normalized throughput while ensuring certain fairness index. Impact of error when estimating backlogged load toward the effectiveness of the strategy is subsequently discussed.

The rest of the paper is organized as follow. Section 2 delivers the general RA model considered in this work. In this section, several key RA parameters exploited in the later section is described. This model is then elaborated mathematically in Section 3. In this section, the analytical model is constructed in a flow to obtain the formula to estimate the performance metrics of normalized throughput and average access delay. Subsequently, Section 4 elaborates how a fair configuration can be obtained following our framework which utilizes the analytical model discussed in Section 3. Section 5 presents several motivating cases to evaluate the system with and without our improved configurations. Finally, Section 6 delivers the conclusion of this work.

2 System Model

This work considers an NB-IoT cell with M stationary IoT devices. Base station activates B CE levels. Subsequently, let us denote the number of devices residing in CE level $b\,(0 \leq b < B)$ as M_b, and $M_0 + \ldots + M_{B-1} = M$. Let us denote the number of (distinct) sub-carriers reserved by the base station for CE level b as R_b. Following the specification in [31], $R_0 + \ldots + R_{B-1} \leq 48$. Let T_b be the NPRACH periodicity (i.e., interval between the beginnings of two successive NPRACHs) of CE level b; $N_{\mathrm{PT},b}$ be the maximum number of transmission attempts of a device residing at CE level b; N_{G} be the maximum number of transmission attempts of a device in all CE levels; and $W_{\mathrm{BO},b}$ be the backoff window in CE level b. For simplicity, it is assumed that a device can transmit preamble and receive acknowledgement within T_b ms. The notations used in this and subsequent sections are summarized in Appendix.

Fig. 1 illustrates high-level view of preamble transmissions in the time-frequency resource of NB-IoT with $B = 3$. For simplicity and readability, it assumes no backoff $(W_{\mathrm{BO},0} = W_{\mathrm{BO},1} = W_{\mathrm{BO},2} = 0)$; similar number of sub-carriers in all CE levels $(R_0 = R_1 = R_2)$; and higher CE level has twice periodicity as the lower one $(T_2 = 2T_1 = 4T_0)$ which makes sense since higher CE level normally employ more repetitions. To illustrate the propagation of transmission attempts, let us consider that all of the first preamble transmission attempts are conducted at the beginning of time. It is assumed in this figure that each device can conduct up to $N_{\mathrm{G}} = 6$ preamble transmissions, but CE level 0 can only "accept" 4 attempts and CE level 1 can only "accept" 3 attempts $(N_{\mathrm{PT},0} = 4,\ N_{\mathrm{PT},1} = 3)$. Hence, devices from CE level 0 can transmit 4 attempts in CE level 0 and 2 attempts at CE level 1; devices from CE level 1 can transmit 3 attempts in CE level 1 and 3 attempts at CE level 2; and devices from CE level 2 can transmit 6 attempts in CE level 2. The notation of $M_{c,i}^b[n]$ and $C_{c,i}^b[n]$ found in Fig. 1 are discussed in Section 3.

In reality, base station can configure NPRACH in each CE level to start at different time in respect to the beginning of the radio frame. Let us denote with d_b the delay (in ms) of the first NPRACH from the beginning of the radio frame where it resides. RA procedure is completed when Msg4's ACK is completely received by base station, which in CE level b happens $d_{\mathrm{ok},b}$ ms after the beginning of successful preamble transmission attempt. Meanwhile, collision is known when Msg3's ACK is not received (timeout), which in CE level b happens $d_{\mathrm{col},b}$ ms after the beginning of the collided preamble transmission attempt. For simpler observation in the later discussion, $d_{\mathrm{ok},b}$ and $d_{\mathrm{col},b}$ are assumed to be constant in CE level b. For easier observation, it is assumed that within an RAR window base station can acknowledge all of the preambles available in each NPRACH. Additionally, let us assume that the probability density function of devices' arrival over

time in the cell is generalized as $A(t)$ for $t = [0, t_{\mathrm{A}}^{\max}]$, which is practical to represent any arrival model.

3 Performance Estimation

In this section, we discuss the derivation of our analytical model which can be used to estimate the performance of NPRACH in term of normalized throughput normalized throughput and average access delay. This performance metric is used in the next section for fairness measurement.

Our proposed analytical model is constructed by exploiting the estimation

$$s \cong me^{-m/r}, \tag{1}$$

which has an acceptable accuracy in estimating s number of successful devices out from m devices contending for r preambles [5]. In our system, contention may happen in each NPRACH of each CE level. Let us denote the number of devices transmitting preamble at the ith NPRACH in CE level b as M_i^b. M_i^b is a summation of $M_{c,i}^b$ for $0 \le c \le b$, where $M_{c,i}^b$ denotes the number of devices from CE level c which transmit at slot i in CE level b. $M_{c,i}^b$ is a summation of $M_{c,i}^b[n]$ for $1 \le n \le N_{b,c}$, where $M_{c,i}^b[n]$ denotes the number of devices from CE level c which transmit its nth attempt at slot i in CE level b, while $N_{b,c}$ denotes the number of attempts that is allowed to be conducted by devices from CE level c in CE level b. Let us also denote by $S_{c,i}^b[n]$ and $C_{c,i}^b[n]$ the number of devices from CE level c which success and collide, respectively, when transmitting their nth attempt slot i in CE level b. In this case,

$$M_{c,i}^b[n] = S_{c,i}^b[n] + C_{c,i}^b[n]. \tag{2}$$

In principle, our analytical model uses (1) recursively for all transmission attempts in each slot in each CE level. To simplify this concept, Fig. 1 illustrates the 'propagation' of $M_{c,i}^b[n]$ and $C_{c,i}^b[n]$, while $S_{c,i}^b[n]$ is not shown since successful devices do not contending the system anymore, although basically non-zero $S_{c,i}^b[n]$ can emerge in any slot of any CE level. For the ease of observation, Fig. 1 considers that retransmissions are conducted without backoff (i.e., in the next immediate slot).

In the lowest CE level of NB-IoT where fewer repetitions are configured, retransmission is conducted with power ramping to increase its detection probability [2]. This is done by increasing the transmission power at the next attempt. Due to this mechanism, the detection probability of the nth preamble transmission attempt is modeled as $1 - e^{-n}$ herein [5].

By applying $M_{c,i}^b[n]$, power ramping, R_b, and $N_{b,c}$ into (1), we can estimate $S_{c,i}^b[n]$ as

$$S_{c,i}^b[n] \cong \begin{cases} (1-e^{-n})M_{c,i}^b[n]e^{\frac{\sum_{c=0}^{b}\sum_{n=1}^{N_{b,c}}(1-e^{-n})M_{c,i}^b[n]}{-R_b}}, & \text{for } b = 0; \\ M_{c,i}^b[n]e^{\frac{\sum_{c=0}^{b}\sum_{n=1}^{N_{b,c}}M_{c,i}^b[n]}{-R_b}}, & \text{otherwise.} \end{cases} \tag{3}$$

According to 3GPP [2], power ramping is only applied in the lowest CE level (i.e., non extended coverage). Hence, in CE level 0, $S_{c,i}^b[n]$ is estimated using the first clause of (3). Meanwhile without power ramping, $S_{c,i}^b[n]$ can be estimated

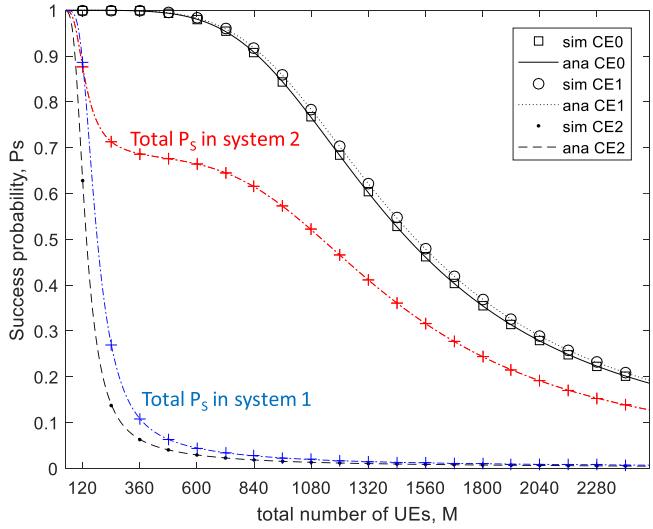

Fig. 2. Comparison of cell with 1 CE level (system 1) and 3 CE levels (system 2).

using the second clause of (3), which can be obtained by simply removing $(1 - e^{-n})$ from the first clause. For estimating total number of successful devices from CE level c which transmit at slot i of CE level b, $\sum_{n=1}^{N_{b,c}} S_{c,i}^b[n]$, we can take (1) and replace s with $\sum_{n=1}^{N_{b,c}} S_{c,i}^b[n]$, m with $\sum_{c=0}^{b}\sum_{n=1}^{N_{b,c}} M_{c,i}^b[n]$, and r with R_b. However for (3), we need to segregate $\sum_{n=1}^{N_{b,c}} S_{c,i}^b[n]$ into $S_{c,i}^b[n]$ because the detection probability is varied by n. Thus, in (3), e is multiplied by $M_{c,i}^b[n]$ instead of $\sum_{c=0}^{b}\sum_{n=1}^{N_{b,c}} M_{c,i}^b[n]$.

When number of resources in RAR equal to the number of available preamble, we denote that the expression in (1) is more precise than the similar expression in [3] since [3] applies power ramping for all CE levels, which is not too accurate when applied to for NB-IoT.

As illustrated in Fig. 2, $M_{c,i}^b[n]$ may 'emerge' from a first attempt, retransmission at the same CE level, and retransmission at the higher CE level. *First attempts* only exist at the slots where the arrival occurs. Without loss of generality, when considering a general arrival function of $A(t)$ and NPRACH periodicity of CE level b, T_b, $M_{c,i}^b[n]$ for $n = 1$ can be calculated as

$$M_{c,i}^b[1] = M_c \int_{\max(0,(i-2)T_b+1)}^{(i-1)T_b} A(t)dt, \text{for } b = c. \tag{4}$$

Notice that the constraint in (4), i.e., $b = c$, emphasis that the first arrival is always conducted in device's initial CE level where it resides.

Retransmission in the same CE level denotes an attempt which is conducted at the same CE level as the previous collided attempt. Devices from CE level c may conduct up $N_{b,c} - 1$ retransmissions in CE level b (-1 to exclude the first attempt). Hence, excluding the first attempt, $M_{c,i}^b[n]$ for $1 < n \le N_{b,c}$ can be calculated as

$$M_{c,i}^b[n] = \sum_{k=K_{\min}}^{K_{\max}} \alpha_{k,i} C_{c,k}^b[n-1], \tag{5}$$

with k being the index of slots at which the previous attempt was collided, which is bounded by $K_{\min}$ and $K_{\max}$. In (5), $\alpha_{k,i}$ denotes the probability that a device which collided at slot k retransmits at slot i. The calculation of this probability is necessary since backoff is applied to the fixed-duration NPRACHs. To allow the reader to focus more on the high-level concept of our proposed analytical model, the more technical derivation of $K_{\min}$, $K_{\max}$, and $\alpha_{k,i}$ are given in Appendix.

Retransmission in the higher CE level, for example CE level b, is possible for a devices from CE level c after it collided $N_{\mathrm{PT},b-1}$ times in CE level $b-1$ and if $N_{b,c} > 0$ for $b > c$. Hence, when $N_{b,c} = 0$, $M^b_{c,i}[1] = 0$. When $N_{b,c} > 0, M^b_{c,i}[1]$ with $b > c$ can be calculated as

$$M^b_{c,i}[1] = \begin{cases} C^{b-1}_{c,J_{\max}}[N_{\mathrm{PT},b-1}], & \text{if} T_b \le T_{b-1}; \\ \sum_{j=J_{\min}}^{J_{\max}} C^{b-1}_{c,j}[N_{\mathrm{PT},b-1}], & \text{otherwise.} \end{cases} \tag{6}$$

where j is the index of slot(s) in CE level $b-1$ where the $N_{\mathrm{PT},b-1}$th attempt was collided. Since retransmission in the higher CE level is conducted immediately without backoff, the upper and lower bounds of j, $J_{\min}$ and $J_{\max}$, respectively, can be calculated by considering the timing differences and NPRACH periodicity across the CE levels. The derivation of $J_{\min}$ and $J_{\max}$ are presented in more detail in Appendix.

To use our analytical model, we need to summarize the iteration to obtain the normalized throughput. As said in the beginning of this section, the iteration is conducted for each NPRACH in each CE level. While number of CE level is known (i.e., B), number of NPRACH in each CE level is unknown at this point. Thus, let us denote with $I_{b,c}$ the number of NPRACHs in CE level b that can be used by devices reside CE level c. Corollary, $I_{b,c} > 0$ when $N_{b,c} > 0$. The first among the $N_{b,c}$ attempts is conducted immediately after arrival if $b = c$ (i.e., if it is in its initial CE level), or being conducted after finishing the attempts in the previous CE level (s) if $b > c$. Meanwhile, the 2nd and the next attempts are each conducted after collision and backoff. Hence, for $c \le b \le B-1$, $I_{b,c}$ can be calculated as

$$I_{b,c} = \begin{cases} \left\lceil \frac{t^{\max}_{\mathrm{A}}}{T_b} \right\rceil + 1 + (N_{b,c} - 1) \left\lceil \frac{d_{\mathrm{col,b}} + W_b - 1}{T_b} \right\rceil, & \text{if } b = c; \\ \left\lceil \frac{I_{b-1,c} T_{b-1} - d_b}{T_b} \right\rceil + 1 + (N_{b,c} - 1) \left\lceil \frac{d_{\mathrm{col},b} + W_b - 1}{T_b} \right\rceil, & \text{if } b > c. \end{cases} \tag{7}$$

Finally, let us derive the formula to estimate the normalized throughput of devices in CE level c, H_c, and the overall normalized throughput of the system, H. H_c is calculated by totaling the number of devices residing CE level c which is successful in their attempt (any attempt number) in any NPRACH in CE level c or higher, which is then normalized/divided by total number of devices residing in CE level c, M_c. Thus, we can express H_c as

$$H_c = \frac{\sum_{b=c}^{B-1} \sum_{i=1}^{I_{b,c}} \sum_{n=1}^{N_{b,c}} S^b_{c,i}[n]}{M_c}. \tag{8}$$

Meanwhile, H as the overall performance metric for the whole system, including all CE levels, can be obtained as

$$H = \frac{\sum_{c=0}^{B-1} H_c M_c}{M}. \tag{9}$$

TABLE 1
Parameters Used in Fig. 2

Case	Parameters				
	B	T_b (ms)	N_{G}	$W_{\mathrm{BO},b}$ (ms)	R_b (sub-carriers)
System 1	1	640	5	1024	48
System 2	3	80, 160, 640	5	2048, 2048, 1024	12, 24, 12

In addition to normalized throughput, it is also important to understand the delay performance of the system. Let us derive the average access delay for successful devices from CE level c, D_c, as

$$D_c = \frac{\sum_{b=c}^{B-1} \sum_{i=1}^{I_{b,c}} \left(\left((i-1)T_b + d_{\mathrm{ok},b} - \overline{t_{\mathrm{A}}} \right) \sum_{n=1}^{N_{b,c}} S^b_{c,i}[n] \right)}{\sum_{b=c}^{B-1} \sum_{i=1}^{I_{b,c}} \sum_{n=1}^{N_{b,c}} S^b_{c,i}[n]}. \tag{10}$$

This indicates that the average access delay for successful devices from CE level c can be obtained by accumulating total time used by each successful attempt that is conducted by devices from CE level c. These comprises the attempts conducted in any slot of CE level $[c, B-1]$. Subsequently, this amount is then normalized by the number of successful attempts for devices from CE level c.

4 Fair Configuration

A recent study in [3] presents an optimization strategy to maximize the system throughput, which is the averaged throughput among CE levels configured in the cell. However, there is a problem since it does not pay attention in the fairness among the three CE levels. For this regard, let us observe Fig. 2 which compares the success probability (i.e., normalized throughput) of two systems. We can think of them as two distinct cells with identical condition (similar service area, channel condition, number of devices, packet generation rate, etc.). System 1 uses 1 CE level to serve all the devices while System 2 uses 3 CE levels for the same purpose. Detailed configurations of the two systems are listed in Table 1 where the configuration for System 2 is the optimal configuration obtained from [6] to maximize the access success probability of the cell. Fig. 2 plots the success probability of each CE levels in System 2, total success probability of System 2, and total success probability of System 1 (which is the same as the success probability in CE level 1 of System 1 since it only has 1 CE level). The total success probability of System 2 is significantly higher than that of System 1. However, in System 2, success probability of CE level 0 is very low compared to CE levels 1 and 2. In System 2, devices do not have equal chance to be successfully access the network. This is a clear sign that the optimal configuration in [6] may result in huge performance difference across CE levels.

The purpose of this paper is to determine a fair resource allocation strategy such that all devices in the system have similar or almost similar normalized throughputs. The Jain's fairness index [30], denoted with F herein, is adopted to evaluate the fairness of normalized throughputs among the three CE levels, which gives

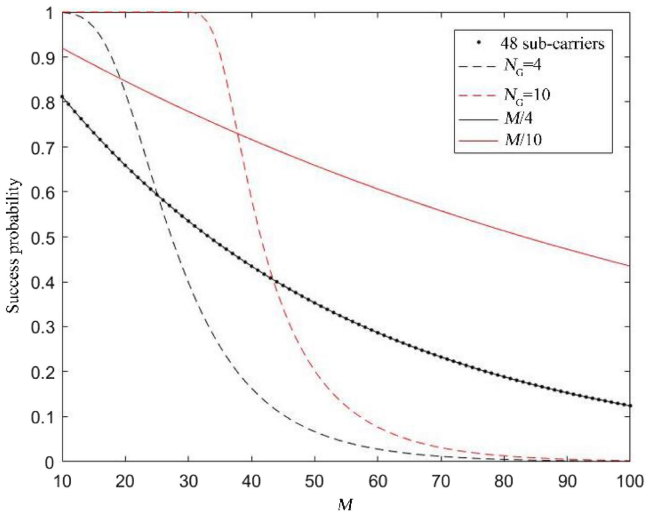

Fig. 3. Influence of each parameter toward success probability.

$$F = \frac{\left[\sum_{c=0}^{B-1} H_c\right]^2}{B \cdot \sum_{c=0}^{B-1} H_c^2}. \tag{11}$$

Max-min fairness is not adopted since the system may operate in overload condition, i.e., not every flow is satisfied. On the other hand, fairly shared spectrum efficiency (FSEE) is not applied since the spectral efficiency between CE levels may not be the same. We deal with a search problem whose objective is to find the configurations which can yield a fair performance in term of normalized throughput among CE levels. However in practice, 'fair' only is not enough. An NB-IoT system should also be able to effectively operate with certain throughput. Hence, 'fair' system can be translated as a system whose configuration is able to yields

- the highest overall normalized throughput as long as the normalized throughput in each CE level is above a certain threshold, or
- the highest overall normalized throughput with fairness index among CE levels above a certain threshold.

Notice that the overall throughput obtained by the proposed fair configuration may be lower than the one obtained by throughput maximization approach [3]. This is because without additional preamble or allowed RA delay, realizing fair access chance among devices in different network environment dynamics (represented by different CE levels) sometimes can only be done by (slightly) worsen the performance of the lower CE level to improve the performance of the higher CE level, which otherwise will be very poor.

Random access configuration in NB-IoT consists of several parameters, namely number of CE level, maximal number of attempt that device from each CE level can perform in each CE level, backoff window in each CE level, and number of sub-carrier or preamble in each CE level. Finding a combination of those parameters which favors our objective not an easy task since each parameter has different influence toward the system's performance. To illustrate this problem, simple exemplary cases were simulated and the results are shown in Fig. 3.

Fig. 3 shows simulation result of 5 exemplary cases for serving M devices with only 1 CE level and no backoff. The configurations of these simplified settings are detailed in Table 2 and are chosen to demonstrate the effect of different parameters toward the success probability. The case "48 sub-carriers" represents the situation when base allocates 48 sub-carriers without retransmission. The cases $N_G = 4$ and $N_G = 10$ represent the situations when base station allows up to 4 and 10 transmission attempts, respectively, with only 12 sub-carriers. The cases $M/4$ and $M/10$ represent the situation when only $\frac{1}{4}$ and $\frac{1}{10}$ of the devices, respectively, are transmitting their preamble when base station allocates 12 sub-carriers without retransmission.

TABLE 2
Parameters Used in Fig. 3

Case	Parameters			
	M	N_G	$W_{BO,b}$ (ms)	R_b (sub-carriers)
48 sub-carriers	M	1	0	48
$N_G = 4$	M	4	0	12
$N_G = 10$	M	10	0	12
$M/4$	$M/4$	1	0	12
$M/10$	$M/10$	1	0	12

This figure shows that that cases "48 sub-carriers" and $M/4$ yield the same performance. This is because they have the same ratio of number of sub-carriers toward number of devices. Meanwhile, case $M/10$ yields higher normalized throughput that the previous two because it has higher ratio of number of sub-carriers toward number of devices, causing it to has less contention. Cases $N_G = 4$ and $N_G = 10$ achieve high normalized throughput for lower M, but quickly drop to very low normalized throughput when M grows. This indicates that increasing number of retransmission without increasing the ratio of number of sub-carriers toward number of devices is not too effective in handling large number of devices. Overall from the result of Fig. 3, we can find that in the small M, increasing N_G plays important role in increasing the throughput. Meanwhile for larger M, increasing the ratio of number of sub-carriers toward number of devices is necessary. In practice, this ratio is also affected by backoff window as it spreads the load in time domain. We adopt the problem definition in [3] to find the configurations which maximize system's throughput and use it to narrow down our search space. In addition to fairness and throughput, in practice an NB-IoT system may also include certain delay requirement. Let us denote by D_{max} the maximum tolerable access delay. Hence, the solution to our search problem must also fulfill

$$\begin{aligned} &(I_{b,c} - 1)T_b + d_{ok,b} \leq D_{max} \\ &\forall 0 \leq c \leq b \leq B - 1. \end{aligned} \tag{12}$$

Overall, the flow of our algorithm to obtain fair access for multiple CE levels is depicted in Fig. 4. In this flowchart, D_{max} denotes the maximum tolerable access delay which normally decided by the operator, P denotes the number of possible connections before performances degradation (which happens at $M = M_{max,P}$), and M_{max} denotes the maximum number of the devices in the cell that the operator hope to support. From this flowchart, we can immediately observe that the main contribution of our algorithm is the configuration sequence of the available parameters. This overall forms an iterative search with specified well-thought-out heuristic

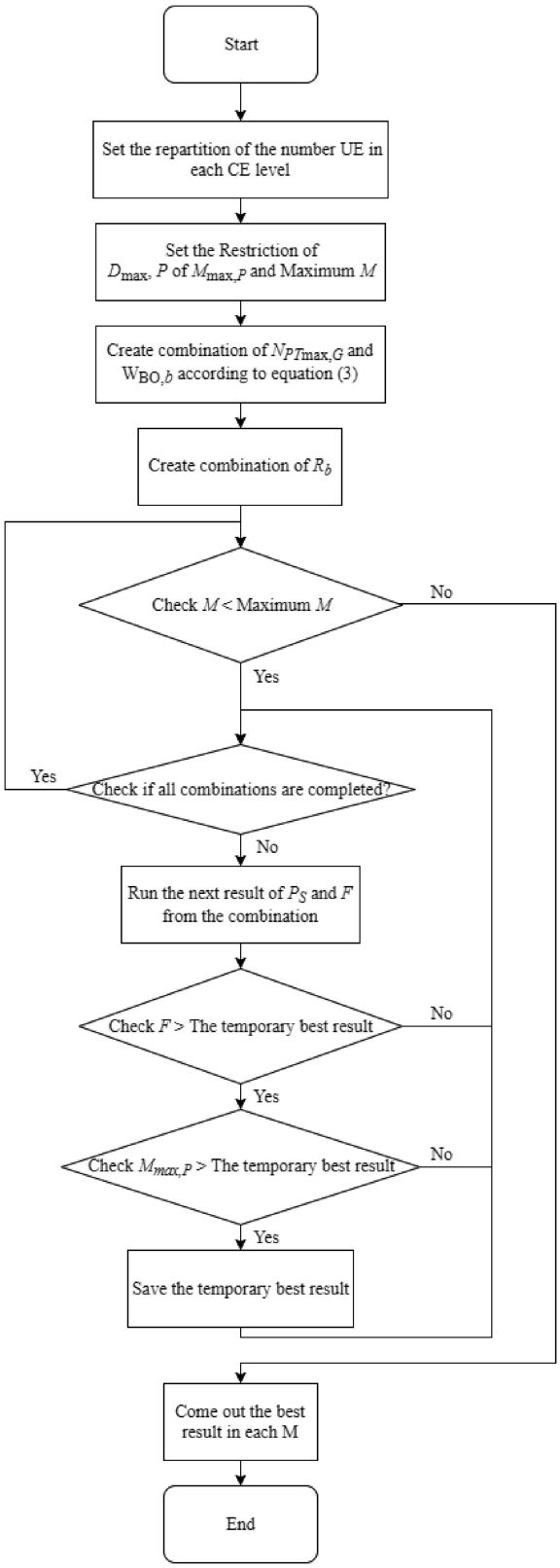

Fig. 4. Flowchart of the proposed algorithm.

thresholds. First, our algorithm takes the input from the setting of CE level division, which defines number of devices in each CE level. Our algorithm also takes $D_{\max}$, P and $M_{\max}$. Subsequently, the algorithm constructs possible combination of maximum number of attempts, N_{G}, and backoff window, $W_{\mathrm{BO},b}$ since both of this parameters affect the maximum

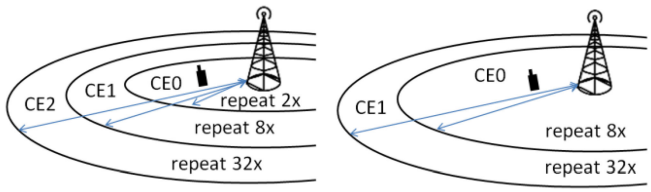

Fig. 5. Serving the same cell area with 3 (left) and 2 (right) CE levels.

delay. Finally, the algorithm then checks and searches for the optimal configuration based on the success probability, P_{S} and fairness index F. Theoretical complexity of this algorithm is $\mathrm{O}(\mathrm{N}^2)$ as it contains 2 loops which outweigh the other calculations. The 2 loops are for the two combinations constructed at the 4th and 5th steps in Fig. 4, which are trimmed down only for feasible combinations according to (12). Meanwhile, the cyclomatic complexity [32] of the proposed algorithm is 5, which remains the same for any networks consisting of multiple CE levels, such as NB-IoT, LTE-M and possible 5G derivatives.

5 Evaluation

A series of experiments are performed to evaluate the correctness and the effectiveness of the algorithm in various configuration strategies. A fixed cell area with varying number of fixed IoT devices, M, of 120 to 1080 is observed. This represents varying real-world load condition. Base station enables $B \in \{1, 2, 3\}$ CE level(s), and sets their RSRP threshold such that $M_b = M/B$ for $0 \leq b < B$. While $M_b = M/B$ may not always applicable, it is in fact the most efficient allocation of CE level's resource. When one of the CE levels has fewer devices, having more resources allocated to it may lead to resource starvation of the other CE levels with more devices. Additionally, since we are seeking for fairness, our attempt is easier when we pre-configure the system to be fair while we can since it can be made without sacrificing cell coverage. Notice that based on (4), our proposed algorithm can still operate regardless of M_b value, for $0 \leq b < B$.

As for the load, simultaneous arrival of M_b devices in CE level $0 \leq b < B$ is taken into account. This is to represent the extreme cases which may occur in IoT environment such as during emergency in safety application. In such extreme condition, the access fairness is most likely to be neglected and thus realizing access fairness is more challenging. Some IoT applications may have various arrival patterns and rarely encounters simultaneous arrival. Owing to (4), in this evaluation, performance of the proposed algorithm under lower load (also those with longer arrival duration) can roughly be represented by lower M.

When there is only 1 CE level configured by base station, it is assumed that NPRACH repetition of 32 is required to cover all of the area. For our study on fairness among CE levels, multiple CE levels are required. Hence, in this evaluation, base station is assumed to configure 2 or 3 CE levels in the cell, illustrated in Fig. 5. When 3 CE levels are configured, NPRACH repetitions for CE level 0, 1, and 2 are set to be 2, 8 and 32, respectively [2]. When there are 2 CE levels, NPRACH repetitions for CE level 0 and 1 are 8 and 32, respectively. The NPRACH periodicity for NPRACH repetition of 2, 8, and 32 are 80, 160, and 640ms, respectively. Additionally,

TABLE 3
Search Space of The Parameters

Parameter	Search space
B	2, 3
N_{G}	3, 4, 5, 6, 7, 8, 10, 20, 50, 100, 200
$N_{\mathrm{PT},b}$ for $0 \leq b \leq B-1$	3, 4, 5, 6, 7, 8, 10
$W_{\mathrm{BO},b}$ for $0 \leq b \leq B-1$	0, 256, 512, 1024, 2048, 4096, 8192, 16384, 32768, 65536, 131072, 262144, 524288 (ms)
R_b for $0 \leq b \leq B-1$	0, 12, 24, 36, 48

it is assumed that the RA must not took longer than $D_{\max} = 10\mathrm{s}$ to complete [6].

In this evaluation, search will be conducted to find the optimal configuration. Different strategies with different objective are examined. These objectives represent different point of views when defining and applying fairness into NB-IoT system with multiple CE levels. The search space of the parameters is listed in Table 3. A custom-made computer simulation is performed to conduct an exhaustive search throughout all combinations with 10^5 iterations. Our proposed analytical model is used to perform more 'clever' and faster search following [3] and the details elaborated in Section 3. The result from simulation is used to evaluate the correctness of our analytical model.

We examine four search strategies which represent different paradigm/definition of access fairness.

- Strategy 1 is taken from [3] and is included here as baseline for comparison. In Strategy 1, the base station choses the configuration which maximize normalized throughput in the cell regardless of the fairness among CE levels.
- Strategy 2 defines that the system is fair as long as the normalized throughput in each CE level is at least 80 percent. Hence, in Strategy 2, base station choses the configuration which yields the highest overall normalized throughput as long as the normalized throughput in each CE level is ≥ 80 percent.
- Strategy 3 defines that the system is fair if the fairness metric F is at least 80 percent. Hence, in Strategy 3, base station choses the configuration which yields the highest overall normalized throughput with $F \geq 80$ percent.

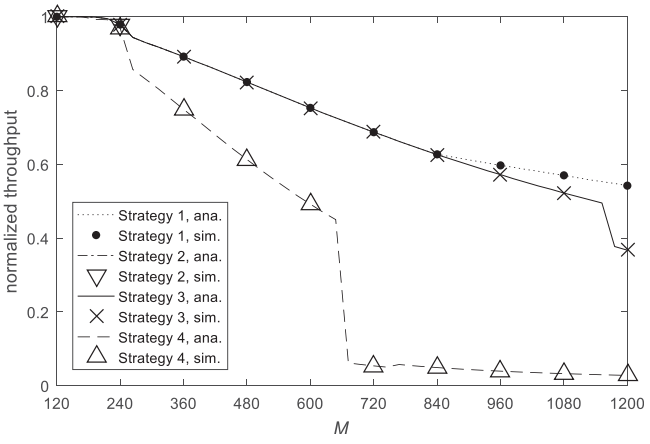

Fig. 6. Overall normalized throughput for $B = 3$.

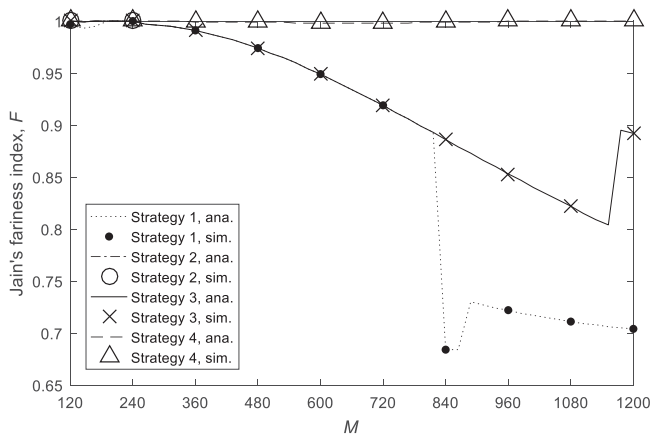

Fig. 7. Jain's fairness index F among the three CE levels for $B = 3$.

- Strategy 4 defines that the system is fair only when the maximum fairness metric F is obtained. Hence, in Strategy 4, base station choses the configuration which maximize F regardless of the obtained overall or CE level's normalized throughput.

Note that due to fewer granularities in the search space of each parameter listed in Table 3, the strategies may not yield smooth line plots.

5.1 Three CE Levels

For the ease of our discussion, let us first discuss the result for $B = 3$ as it is proven to be more optimal than $B = 2$ in [3]. The results for normalized throughput and fairness index are depicted in Figs. 6 and 7, respectively. In these figures, the result from our analytical model is also compared with the corresponding simulation results.

Fig. 6 shows the normalized throughput of the four strategies. In this figure, we can observe that the result from our analytical model is consistent with the result from simulation. This denotes that the analytical model is accurate. Furthermore, performance of each strategy is depicted. Strategy 1 returns the highest normalized throughput compared to the other strategies. This is as expected since the search in this strategy is focused to obtain the highest access success probability, as stated in [3]. Strategy 2 only has some results at $M < 360$. This is because when $M \geq 360$, there is no configuration that can meet the requirement, i.e., one or more CE levels has its normalized throughput below 80 percent. Strategy 3 obtains similar normalized throughput as Strategy 1 for $M \leq 816$. Meanwhile for $M > 816$, Strategy 3 obtains lower normalized throughput than Strategy 1. This is because when $M > 816$, $F \geq 80$ percent can only be obtained when the normalized throughput is lower than that of Strategy 1. Strategy 4 obtains the lowest normalized throughput compared to the other strategies for $M \geq 264$. It even drops sharply to a very low normalized throughput for $M > 648$.

Fig. 7 shows the fairness index of the compared strategies. Similar to the result in Fig. 6, the result in Fig. 7 also shows consistency between the result obtained by analytical and simulation, denoting the high accuracy of our analytical model. In this figure, Strategy 1 yields the lowest F among the compared strategies, especially when $M > 816$. This indicates that to maximize the overall normalized throughput, it may 'sacrifice' some CE levels (e.g., deliberately sets

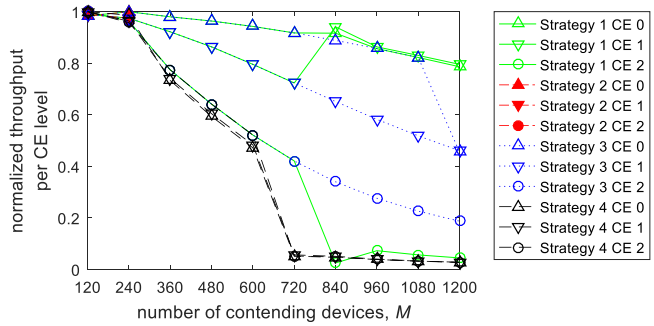

Fig. 8. Comparison of normalized throughput in the three CE levels.

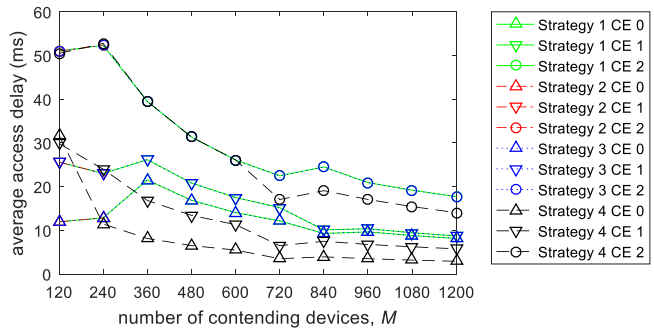

Fig. 9. Comparison of average access delay in the three CE levels.

one or two CE levels to have heavier contention and letting the other CE level harvest higher throughput), which is not fair. Strategy 2 only has some results at $M < 360$ due to similar reason as in Fig. 6. Strategy 3 has similar result as Strategy 1 for $M \leq 816$, which is also the case in Fig. 6. For $M > 816$, it maintains higher F compared to Strategy 1 since Strategy 3 choses the configuration to maintain $F \geq 80$ percent even if it obtains slightly lower normalized throughput than Strategy 1. Strategy 4 obtains the highest F compared to the other scenarios since it always seeks for the highest F regardless of the throughput, which is the complete opposite of Strategy 1.

As an additional side note, notice that the inflection points found in Figs. 6 and 7 are simply due to the search. The search aims to find the configurations that meet the requirements. Hence, the linearity of the obtained normalized throughput and fairness index in respect to the load M is not its concern. To give more detailed elaboration of the fairness index measured in Fig. 7, Fig. 8 is provided to show the normalized throughput of each CE level. From Fig. 8, we can better understand the meaning behind the plots in Fig. 7.

Fig. 9 shows the average access delay for successful devices from each CE level. In this figure, we can observe the delay performance of the system. It is observed that with Strategy 1, overall, $D_0 < D_1 < D_2$, since in Fig. 8 overall $H_0 > H_1 > H_2$. It is also occur in Strategy 3 and 4. Meanwhile for Strategy 2, it is still consistent to what are shown in Figs. 6, 7 and 8, that it only have some results for $M < 360$. Obviously, when two strategies yield similar result in Fig. 8, it will also be the same in Fig. 9.

5.2 Two CE Levels

The results for normalized throughput and fairness index are depicted in Figs. 10 and 11, respectively. Overall, the findings are similar with those of Figs. 6 and 7. Simulation results in Figs. 10 and 11 similarly confirm the accuracy of

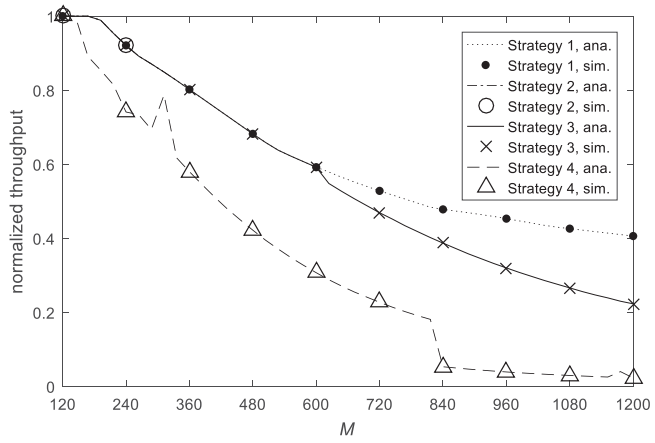

Fig. 10. Overall normalized throughput for $B = 2$.

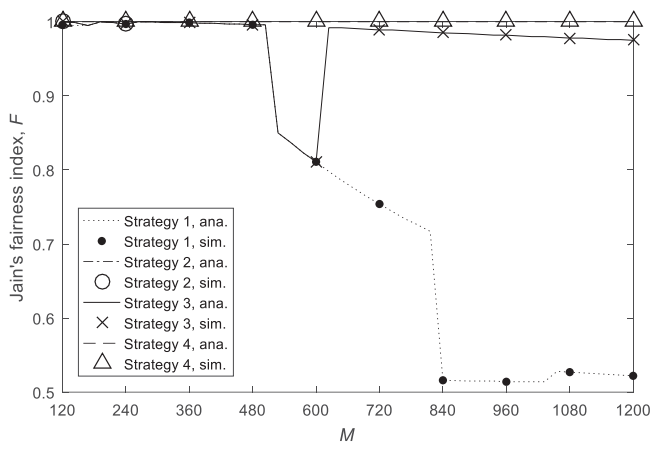

Fig. 11. Jain's fairness index F among the three CE levels for $B = 2$.

our analytical model against the provided simulation. Strategy 1 has the highest normalized throughput with the lowest F compared to the other strategies; especially in the larger M. Strategy 2 can only find its optimal configuration for $M \leq 336$. Strategy 3 yields similar normalized throughput as Strategy 1 when $M \leq 600$, and yields lower normalized throughput compared to Strategy 1 when $M > 600$. Strategy 4 yields the lowest normalized throughput with the highest F.

When comparing the result of each strategy between Figs. 10 and 6, we recognize that all strategies yield higher normalized throughput under $B = 3$ instead of $B = 2$, except Strategy 4 when $672 \leq M \leq 936$. In addition to supporting the finding in [3], this phenomenon indicates that there exist fairer configuration(s) when we decrease the number of CE levels from 3 to 2. However, from business point of view these configurations are not practical as it yields very low normalized throughput.

The result for fairness is shown in Fig. 11, which is calculated using (11) based on the data which is plotted in Fig. 10. In Fig. 11, we can observe that Strategy 4 is the fairest one. However, reflecting this result to Fig. 12, we can understand that although it is fair, all CE levels under Strategy 4 suffer from very low normalized throughput under higher load. Meanwhile, Strategy 3 yields almost similar fairness index as Strategy 4 in several values of M (see Fig. 11) while fairly higher normalized throughput for all CE levels (see Fig. 12).

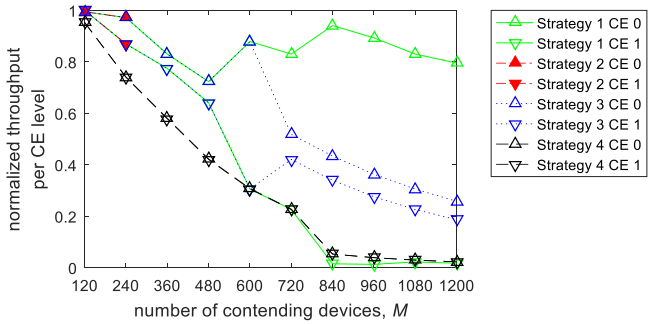

Fig. 12. Comparison of normalized throughput in the two CE levels.

This shows a trade-off between fairness and normalized throughput when the contention is very high (i.e., $M \gg R$).

In addition to normalized throughput, Fig. 13 shows the average access delay for successful devices from each CE level. Normally in a random access system, normalized throughput is inversely proportional to the average access delay. Hence, when two strategies yield similar result in Fig. 12, it is also similar in Fig. 13. With only 2 CE levels in the cell, Strategy 1 yields $D_0 < D_1$, since $H_0 > H_1$ as depicted in Fig. 12. This situation is similar to Strategy 3 and 4. Meanwhile Strategy 2 only have some results for $M < 336$.

5.3 Impact of Estimation Error

Our algorithm utilizes our iterative model to estimate the number of contending and successful user in each PRACH. The model uses an assumed load, $A(t)$, which may be hard to predict in real scenario. In this regard, Fig. 14 exhibits the scenarios when the assumed $A(t)$ has no error, 10 percent error, and 20 percent error. In Fig. 14, only result from Strategy 3 is shown since from the previous experiments it exhibits the best trade-off between normalized throughput and fairness. In these figures, $B = 2$ is chosen such that error of E, $0 < E < 1$, represents the situation when there are $(1-E)M/2$ actual users in CE level 0 (and respectively $(1+E)M/2$ actual users in CE level 1), but our algorithm configures the system to handle $M/2$ devices in CE level 0 and $M/2$ devices in CE level 1.

Fig. 14 shows the normalized throughput for each CE level (top), for the whole system (middle) and the obtained fairness index (bottom) for different error rate. Note that when there is no error, the result is similar to Strategy 3 in Figs. 10, 11 and 12 (but with different resolution in x-axis). From Fig. 14, we can observe that the error in estimating load in each CE level, in general, decreases the effectiveness of our algorithm in maximizing overall normalized throughput

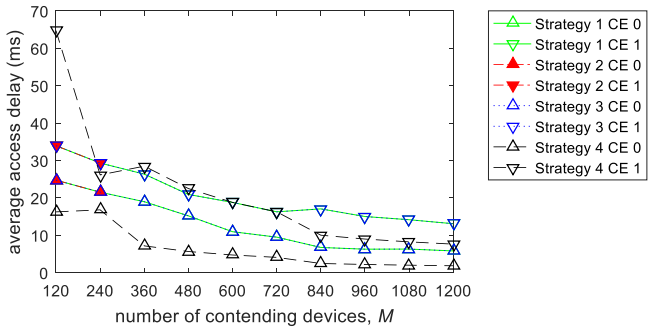

Fig. 13. Comparison of average access delay in the two CE levels.

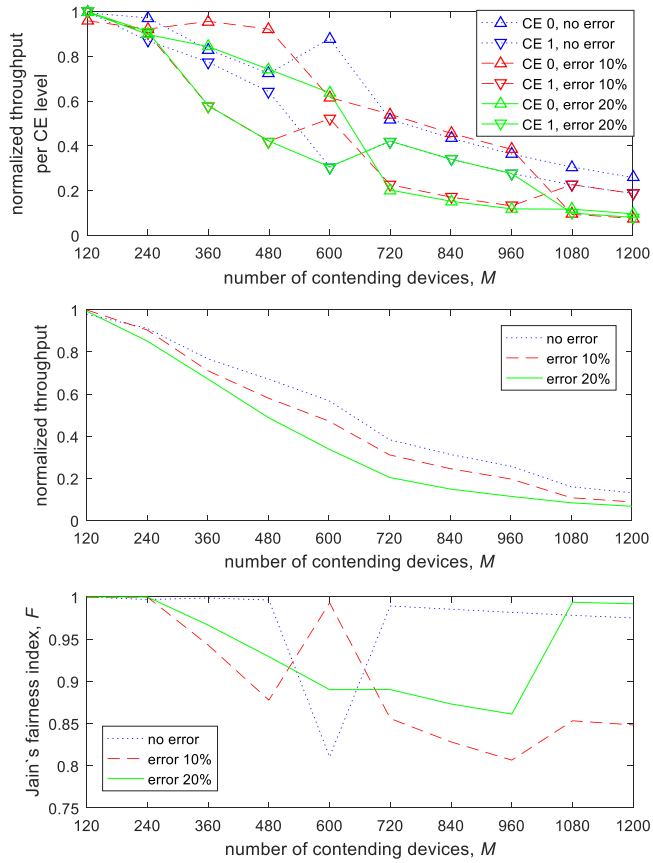

Fig. 14. Normalized throughput of each CE level (top), for all CE levels (middle) and fairness index (bottom).

and fairness index. Overall, as the error increases, normalized throughput for the system is decreased. However, due to search criteria of Scenario 3, the fairness can still be managed to be higher or equal to 80 percent, with varying normalized throughput in each CE level. The fairness seems to be uncorrelated with the error rate since the algorithm has fewer granularities when tuning the parameters (see Table 3). Notice that although our algorithm yields low F for $M = 600$ when having no estimation error, the obtained F is 0.81, which is still above 0.8. More importantly, in such situation, it obtains higher overall throughput compared to those with estimation error.

5.4 Robustness Under Different Arrival Models

Fig. 15 exhibits the overall normalized throughput (top) and Jain's fairness index (bottom) obtained by the three strategies under various arrival pattern in the three CE levels. In this comparison, similar settings as the previous evaluation are used except that CE level 0 has uniform arrival pattern from the beginning until the 30th slots; CE level 1 has beta arrival pattern with $\alpha = 2$ and $\beta = 5$ from the 15th until the 30th slots; while CE level 2 has beta arrival pattern with $\alpha = 5$ and $\beta = 2$ from the beginning until the 50th slots. This configuration is taken to demonstrate the robustness of the proposed algorithm under various arrival models. In Fig. 15, only result from Strategy 3 is shown since from the previous experiments it exhibits the best trade-off between normalized throughput and fairness.

The exhaustive search result is included to validate the proposed result. Overall, the result obtained by our strategy

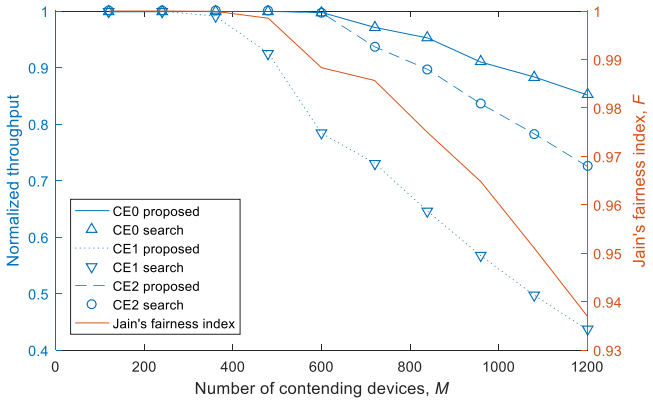

Fig. 15. Normalized throughput and fairness index under various arrival patterns.

is exactly the same as the one obtained by exhaustive search. This indicates that the proposed strategy is robust to be implemented for various traffic arrival patterns.

6 Conclusion

In this paper, an effective strategy to obtain the configuration which yield fairer throughput across CE level in an NB-IoT system is presented. The effectiveness of the strategy is verified by a brute-force search using our analytical and computer simulations. During evaluation, different practical assumptions of a fair system are explored and examined. The result shows that the analytical model is accurate under various loads. Additionally, the proposed search strategy is proven to be able to obtain the configuration which yield acceptable throughput fairly for all CE levels. The delay performance is also evaluated and consistently demonstrates that higher normalized throughput yields lower average access delay. The proposed search strategy is also evaluated under multiple estimation error scenarios. It is worth to be noticed that in NB-IoT, having only one CE level can surely be fair for all devices. However, it has a very low normalized throughput. Therefore, we propose the usage of multiple CE level, but with proper configuration to achieve fairness with acceptable throughput without compromising the coverage.

Acknowledgment

This work was supported by the United States Agency for International Development (USAID) through the Sustainable Higher Education Research Alliance (SHERA) Program for Universitas Indonesia's Scientific Modeling, Application, Research and Training for City- centered Innovation and Technology (SMART CITY) Project, Grant #AID-497-A-1600004, Sub Grant #IIE- 00000078-UI-1.

References

[1] Y. P. E. Wang *et al.*, "A primer on 3GPP narrowband internet of things," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 117–123, Mar. 2017.
[2] 3GPP, TS 36.321, "Medium access control (MAC) protocol specification, V13.2.0," Jun. 2016.
[3] R. Harwahyu and R. G. Cheng, "Optimization of random access channel in NB-IoT," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 391–402, Jan. 2018.
[4] R. Harwahyu, R. G. Cheng, W. J. Tsai, J. K Hwang, and G. Bianchi, "Repetitions vs retransmissions: Tradeoff in configuring NB-IoT random access channels," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3796–3805, Jan. 2019.
[5] C. H. Wei, R. G. Cheng, and S.L. Tsao, "Modeling and estimation of one-shot random access for finite-user multichannel slotted ALOHA systems," *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1196–1199, Aug. 2012.
[6] 3GPP, TR 45.820, "Cellular system support for ultra-low com-plexity and low throughput internet of things (CIoT), V13.1.0," Nov. 2015.
[7] 3GPP, TR 37.868, "RAN improvements for machine-type communications, V1.0.0," Aug. 2011.
[8] 3GPP, R2-104870, "Pull based RAN overload control," Aug. 2010.
[9] I. Rubin, "Group random-access discipline for multi-access broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 5, pp. 578–592, Sep. 1978.
[10] H. H. Tan and H. Wang, "Performance of multiple parallel slotted ALOHA channels," in *Proc. INFOCOM*, 1987, pp. 931–940.
[11] Y.-J. Choi, S. Park, and S. Bahk, "Multichannel random access in OFDMA wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 603–613, Mar. 2006.
[12] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted ALOHA with multipacket reception," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2636–2656, Jul. 2005.
[13] P. R. Jelenkovic and J. Tan, "Stability of finite population ALOHA with variable packets," Dept. Elect. Eng., Columbia Univ., NY, USA, Rep. EE2009-02-20, 2009.
[14] P. Zhou, H. Hu, H. Wang, and H.-H. Chen, "An efficient random access scheme for OFDMA systems with implicit message transmission," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2790–2797, Jul. 2008.
[15] L. Kleinrock and F. Tobagi, "Packet switching in radio channels: Part I—Carrier sense multiple-access modes and their throughput— Delay characteristics," *IEEE Trans. Commun.*, vol. COM-23, no. 12, pp. 1400–1416, Dec. 1975.
[16] R. Harwahyu, X. Wang, R. F. Sari, and R. G. Cheng, "Analysis of group paging with pre-backoff," *EURASIP J. Wireless Commun. Netw.*, vol. 1, no. 34, pp. 1–9, Feb. 2015.
[17] Z. Dawy *et al.*, "Toward massive machine type cellular communications," *IEEE Wireless Commun .*, vol. 24, no. 1, pp. 120–128, Feb. 2017.
[18] T. M. Lin *et al.*, "PRADA: Prioritized random access with dynamic access barring for MTC in 3GPP LTE-A networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2467–2472, Jun. 2014.
[19] L. Ferdouse, A. Anpalagan, and M. Nadarajah, "Performance of dynamic access class barring method in cellular M2M networks," *Wireless Personal Commun.*, vol. 91, no. 3, pp. 1471–1487, Dec. 2016.
[20] D. Sijabat, R. Harwahyu, and R. G. Cheng, "Energy-efficiency of RACH-based small data transmission scheme in LTE networks," in *Proc. IEEE Int. Conf. Telecommun. Signal Process.*, 2017, pp. 106–109.
[21] E. Paolini *et al.*, "Coded random access: Applying codes on graphs to design random access protocols," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 144–150, Jun. 2015.
[22] A. Laya, L. Alonso, and J. Alonso-Zarate, "Contention resolution queues for massive machine type communications in LTE," in *Proc. IEEE Personal Indoor Mobile Radio*, 2015, pp. 2341–2318.
[23] A. Samir, M. M Elmesalawy, A. S. Ali, and I. Ali, "An improved LTE RACH protocol for M2M applications," *Mobile Inf. Syst.*, vol. 2016, Aug. 2016, Art. no. 3758507.
[24] M. Afif, W. B. Hassen, and S. Tabbane, "A resource allocation algorithm for throughput maximization with fairness increase based on virtual PRB in MIMO-OFDMA systems," *Wireless Netw.*, vol. 25, no. 3, pp. 1083–1097, Feb 2018.
[25] V. K. Gupta and G. S. Kasbekar, "Achieving arbitrary throughput–fairness trade-offs in the inter-cell interference coordination with fixed transmit power problem," in *Proc. Netw. Games Control Optim.*, 2019, pp. 17–35.
[26] F. Mohammadnia *et al.*, "Mobile small cells for adaptive ran densification: Preliminary throughput results," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2019, pp. 1–7.
[27] E. Arribas, V. Mancuso, and V. Cholvi, "Fair cellular throughput optimization with the aid of coordinated drones," in *Proc. IEEE Conf. Comput. Commun. Workshops*, 2019, pp. 295–300.
[28] Z. Y. Sun and D. N. Yang, "A D2D wireless resource allocation scheme based on overall fairness," *3D Res.*, vol. 2019, no. 10(2), pp. 1–10, May 2019.
[29] M. M. Al-Wani *et al.*, "On short term fairness and throughput of user clustering for downlink non-orthogonal multiple access system," in *Proc. IEEE 89th Veh. Technol. Conf.*, 2019, pp. 1–6.

[30] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared systems, digital equipment corporation," Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA., 1984.
[31] 3GPP, TS 36.331, "Radio resource control (RRC) protocol specification, V14.1.0," Dec. 2016.
[32] A. H. Watson *et al.*, "Structured testing: A testing methodology using the cyclomatic complexity metric," US Department of Commerce, Technology Administration, National Institute of Standards and Technology, vol. 500, no. 235, Sep. 1996.

Ruki Harwahyu received the BE degree in computer engineering from Universitas Indonesia (UI), Depok, Indonesia, in 2011. He received the ME degree from UI, and the MSc degree from the National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan, both in computer and electronic engineering, in 2013. He received PhD degree in electronic and computer engineering from NTUST, in 2018. Currently, he is an assistant professor with the Department of Electrical Engineering, Faculty of Engineering, UI. He serves as a member of editorial board in two international journals. He has been involved in organizing several IEEE international conferences. He received the Honorary Award from CTCI Foundation, Taiwan, in 2017. His current research interests include computer and communication networks and Internet of Things. He is a member of the IEEE.

Ray-Guang Cheng (S'94–M'97–SM'07) received the BE, ME, and PhD degrees from National Chiao Tung University, Hsinchu, Taiwan, in 1991, 1993, and 1996, respectively, all in communication engineering. From 1997 to 2000, he was a researcher and a project leader with the Advance Technology Center, Computer and Communication Laboratories, Industrial Technology Research Institute (ITRI), Hsinchu. He led the 3G Protocol project and his team was named Top Research Team of the Year by ITRI in 2000. From 2000 to 2003, he was a senior manager of the R&D division at BenQ Mobile System Inc., Hsinchu. He is currently a professor with the Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan. He a member of Phi Tau Phi scholastic honor society, and the holder of IEEE Wireless Communication Professional Certification. His research interests include 5G machine-type communications and IoT management platform. He is a senior member of the IEEE.

Da-Hao Liu received the graduate degree in electrical engineering from National Taiwan Ocean University, in 2015. He received the master degree in electronic and computer engineering from the National Taiwan University of Science and Technology, Taiwan, in 2018. He is currently a engineer with Foxconn FIH Mobile Limited, Taipei, Taiwan. His current research interests include baseband circuit design of cell phone and router.

Riri Fitri Sari received the bachelor's degree in electrical engineering from Universitas Indonesia. She received the MSc degree in computer science and parallel processing from the University of Sheffeld, UK, and the PhD degree from the School of Computing, University of Leeds, UK. She is currently a professor of computer engineering with the Department of Electrical Engineering, Universitas Indonesia. Her teaching and research interests include IoT, vanet, communications network, and implementation of ICT for humanities. She is a senior member of the IEEE.