SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •
Special Focus on Machine-Type Communications

October 2017, Vol. 60 100304:1–100304:13
doi: 10.1007/s11432-017-9162-9

Group-based joint signaling and data resource allocation in MTC-underlaid cellular networks

Xuefei ZHANG[1*], Yue WANG[1], Ruiqi WANG[1], Yanzhao HOU[1] & Da WANG[2]

[1]*School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing* 100876*, China;*
[2]*China Telecom Corporation Limited Technology Innovation Center, Beijing* 100031*, China*

Received April 21, 2017; accepted June 7, 2017; published online September 5, 2017

Abstract Machine-type communications (MTC) are gaining significant research attention as one of the most promising technologies for the fifth generation (5G) mobile networks. A critical issue handled by MTC is support for massive numbers of connections, which is a growing problem that will become increasingly challenging as MTC share spectrum resources with cellular communication. Here, not only the number of connections but also the data rate requirements of cellular users (CUEs) need to be considered. Given these issues, in this paper, we formulate a group-based joint signaling and data resource optimization model constrained by network resource and data rate requirements in order to maximize the number of connections. We also note that this problem is nonconvex and that obtaining an optimal solution is computationally complex for MTC with massive numbers of users (UEs). Therefore, we decompose the problem into group-based data aggregation and resource allocation subproblems. To solve these two subproblems, we develop an adaptive group head selection algorithm and a joint signaling and data resource allocation algorithm that satisfy both the data rate requirement and resource constraints, respectively. Our simulation results show that our proposed algorithms significantly improve the number of connections when compared with other classic methods. Furthermore, our results reveal that the limiting factor on the number of connections changes with the ratio of the number of MTC UEs to that of CUEs and the ratio of data requirement of MTC UEs to that of CUEs. Finally, we note that our proposed group-based resource allocation algorithm can effectively improve the number of connections, especially when more MTC UEs and a small amount of MTC data are present.

Keywords data aggregation, matrix analysis, machine-type communications, signaling and data resource allocation

Citation Zhang X F, Wang Y, Wang R Q, et al. Group-based joint signaling and data resource allocation in MTC-underlaid cellular networks. Sci China Inf Sci, 2017, 60(10): 100304, doi: 10.1007/s11432-017-9162-9

1 Introduction

Machine-type communications (MTC) or machine-to-machine (M2M) communications enable direct communication between multiple devices. In recent years, the number of MTC devices has grown tremendously. As one of the most promising technologies for fifth-generation (5G) mobile networks, MTC will enable innumerable applications in areas such as smart homes, healthcare, automotive communications, and intelligent cities. Several forecasts predict significant market growth over the next few years for both

* Corresponding author (email: zhangxuefei@bupt.edu.cn)

MTC devices and MTC connectivity segments, with compound annual growth rates estimated to exceed 25% [1–3]. More specifically, according to these forecasts, billions of machines and industrial devices will potentially be able to benefit from MTC [4].

MTC are primarily characterized by a high device density, a low data rate, an acceptable level of delay tolerance, and high connection/communication frequency, all of which represent a very different set of requirements when compared with human-to-human (H2H) communications. The cellular infrastructure is the most mature mobile communications infrastructure designed for H2H communications. It can be a good choice as the primary bearer network for MTC owing to its ability to provide wide coverage and effectively manage MTC connections. Moreover, recent predictions indicate that MTC over cellular mobile networks could serve as leverage to provide much-needed mobile operator revenue [5]. Further, 3rd Generation Partnership Project (3GPP) and IEEE 802.16p both proclaim that MTC devices should be attached to existing cellular infrastructures (e.g., Long Term Evolution-Advanced (LTE-A) or IEEE Wi.MAX2.0) [6]. However, the current cellular infrastructure communication mechanism, which primarily aims to promote high data rates for H2H communications, is liable to cause frequent signaling setup and updates or an overloaded Physical Random Access Channel (PRACH), i.e., signaling congestion, because of the massive number of connections inherent in MTC. The overloaded PRACH is further aggravated if MTC terminals repeat their access attempts without realizing that the unsuccessful random access attempts are the result of an overload. Therefore, a critical issue here is the handling of massive numbers of accesses from a large number of MTC devices while simultaneously satisfying data rate requirements of H2H communications [7, 8].

To solve severe overload problems in the LTE-A PRACH, mobile network operators can minimize the frequency of attempts of MTC devices to implement a particular procedure without having to throttle them or even disallow the concerned MTC devices from connecting to the network and/or executing the intended procedure [4]. Within 3GPP, six possible remedies have been identified to combat the PRACH overload problem: the backoff scheme, slotted access scheme, access-class barring scheme, pull-based scheme, PRACH resource separation scheme, and dynamic PRACH resource allocation scheme [9]. Owing to the traffic load dynamics on the PRACH, these schemes alone cannot fully alleviate the overload problem.

Most of the current research focuses on access mechanism design to achieve a low signaling overhead for small data transmissions in cellular networks. In [10], Zheng et al. provided an overview of the 3GPP LTE-A MTC standard and random access mechanism, investigating the limits of MTC coexisting with H2H communication in terms of congestion probabilities and access delays. Further, a survey of the random access and alternatives has been proposed to improve the operation of the random access channel of LTE and LTE-A [8]. In [8], Laya et al. provided a comprehensive discussion of the different alternatives and identified the strengths and weaknesses of each alternative.

In [5], Ksentini et al. remedied the weaknesses of typical random access approaches through a congestion-aware admission control solution that selectively rejects signaling messages from MTC devices based on a probability established by a proportional integrative derivative controller that reflects the congestion level of a relevant core network node, the effectiveness of their approach was verified via computer simulation. Further, in [11], Li et al. proposed a parallel and distributed admission control algorithm that improves access fairness for M2M devices with different tolerant delays, their approach utilizes the concavity of the payoff function.

In addition to random-access-based, congestion-based, and data-rate-based access rules [12–14], group- or cluster-based multistep access mechanisms have proven to be effective in further reducing signaling overhead by sacrificing access delay [15, 16]. A group-based access mechanism delegates the random access procedure of the devices to an access group controlled by a designated device, this device is referred to as the group head. MTC devices construct a group based on their location [15] or subscription features [17]. However, generally, the number of MTC devices in a group is fixed [15]. Further, as described by Huq et al., simulation results are analyzed to establish a reasonable group size [16]. In [18], Hossain et al. introduced the notion of a flexible admission/connection through which devices can be grouped dynamically for a certain period of time. This grouping leads to a hierarchical network framework that can provide efficient transmission via data aggregation [19]. Here, data aggregation implies that devices transmit messages to their corresponding group head, which handles a bulk of messages in the same group. If data transmissions are small in size, data aggregation can serve as a potential approach to effectively mitigate overhead and congestion.

Resource allocation is also a common approach for reducing signaling congestion during an access procedure. In [14, 20, 21], the access problem with data rate requirements can be represented using a queuing model to perform access management. Further, this problem, together with power allocation, is the focus of [22–25], in which power allocation under time-division multiple access (TDMA) and frequency-division multiple access (FDMA) are explored. Analytical results show that FDMA supports a higher load compared with TDMA under the peak power constraint. Moreover, cognitive radio technology, which can contribute to resource exploration, is introduced into access control in [26–30].

Despite some encouraging results obtained via MTC access mechanism design or resource allocation, most previous studies primarily assume the following two independence constraints: (1) independent signaling and data optimization, which, to the best of our knowledge, involves little work on joint resource optimization or resource competition between signaling and data transmission, and (2) independent MTC and cellular communication, i.e., typically specific resource blocks (RBs) or time slots are allocated to MTC. Clearly, it is excessively demanding to provide exclusive resources to MTC under the supervision of a cellular network. Therefore, sharing (or reusing) will likely be a trend in MTC-underlaid cellular networks to remedy the impracticalities resulting from the two independence constraints described above.

On the basis of the abovementioned aspects, in this paper, we propose a group-based joint signaling and data resource optimization model. Our specific contributions are enumerated as follows.

(1) We formulate a group-based joint signaling and data resource optimization model constrained on network resources and data rate requirements by grouping joint signaling and data resource allocation. Unfortunately, the problem is nonconvex and obtaining an optimal solution is computationally complex for MTC with massive numbers of users (UEs). Therefore, we decompose the problem into group-based data aggregation and resource allocation subproblems.

(2) We generate a hierarchical MTC framework using grouping and data aggregation, which can thereby increase the number of connections. Unlike existing group-based solutions that are highly dependent on specific cases, we provide a general group and data aggregation solution that adaptively performs group construction and selects a group head.

(3) Based on MTC grouping and data aggregation, we further maximize the number of MTC connections under the given data rate requirements for both cellular and MTC UEs. Considering that we are signaling-constrained in the control plane and interference-constrained in the data plane, we propose a joint signaling and data resource allocation algorithm. The essence of our algorithm is adaptive RBs (or bandwidth) and power allocation that simultaneously considers signaling and data transmission for both MTC and cellular communication; to achieve this, we use the eigenvalue analysis of the interference matrix.

In addition to this introductory section, we organized our paper as follows. In Section 2, we describe our system model and provide problem formulation. Next, in Section 3, we propose a joint signaling and data resource optimization model for MTC-underlaid cellular networks that maximizes the number of connections under the given data rate requirements. In Section 4, we provide our simulation results and discussion. Finally, in Section 5, we conclude our paper and provide avenues for future work.

2 System model and problem formulation

In this section, we consider uplink transmission in MTC-underlaid cellular networks, as depicted in Figure 1, where one base station (BS) is located at the center of the cell and M MTC UEs and N cellular UEs are uniformly distributed within the cell. Further, we denote the sets of cellular UEs (CUEs) and MTC UEs as lie and UM, respectively. Without loss of generality, we adopt a contention-free manner of allocating the preambles [31], whereby all UEs initiate the connection to the BS by allocating a

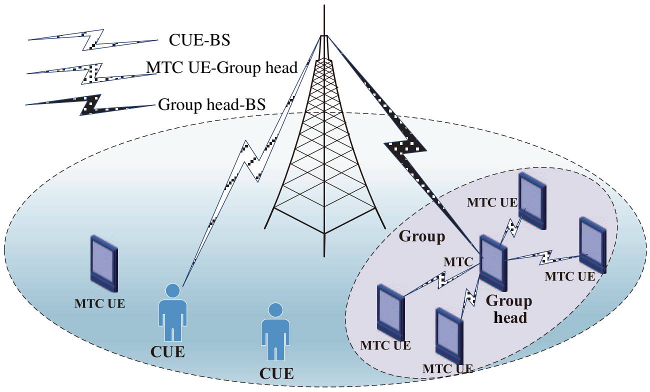

Figure 1 (Color online) System model.

dedicated preamble before data transmission. Note that we can also adopt the contention-based method by applying the concept of random access opportunities (RAOs) in a probabilistic manner [7]. Moreover, it is reasonable to assume that compared with MTC UEs, CUEs are given a higher priority for obtaining a preamble.

2.1 Group-based transmission for MTC UEs

To relieve signaling congestion, we adopt group-based MTC data transmission for MTC UEs. More specifically, MTC UEs are assigned into if disjoint groups (indexed by $k \in \boldsymbol{K} = \{1, \ldots, K\}$) based on the transmission environment (e.g., channel quality and user locations). In each group, one MTC UE is selected as the group head, and the set of group heads is denoted as $\mathcal{A}_a = \{A_1, A_2, \ldots, A_K\}$. The set of MTC UEs in the k-th group is denoted as $\mathcal{M}_k$. In our model, MTC UEs in different groups, MTC UEs not belonging to any group, and one of the CUEs may reuse the same spectrum. Given this, group heads bulk all received data and forward them to the BS. The number of MTC UEs in each group is constrained by the maximum value Z. Finally, MTC UEs can reuse the cellular spectrum and produce mutual interference between the cellular link and MTC UEs to the group head link.

2.2 Resource allocation

Regarding the data rate and signaling requirements of UEs, we address the dynamic resource allocation algorithm in this subsection. Here, PRACH resources can be dynamically adjusted by the BS based on PRACH load and overall network load. When a resource is used for PRACH, it cannot be used for data transmission. Therefore, we divide the available maximum bandwidth, denoted by B, into bandwidth B_{d} for data transmission and bandwidth B_{s} for signaling transmission. For data transmission, CUEs and MTC group heads are assigned portions of the orthogonal spectrum with a total bandwidth of Bd. The data transmission for CUEs is equally allocated to each UE. The transmit power profile vector is denoted by $\boldsymbol{P} = [P(1), P(2), \ldots, P(M+N)]^{\mathrm{T}}$, where $P(i)$ is the transmit power of the i-th user. Further, ΔB represents the unit bandwidth for signaling transmission. Therefore, the total available resources assigned to signaling transmission are determined by the number of unit bandwidths for signaling transmission. We assume here that MTC UEs and CUEs have minimum data rate requirements, denoted by R_T^m and R_T^c, respectively.

2.3 Problem formulation

To formulate the problem, MTC UEs are assigned to K disjoint groups. Only some MTC UEs form a group, while others are independent points not belonging to any group. The set of MTC UEs transmits data to their nearest group heads; then, the group heads become transmitters that transmit aggregated data to the BS. The uplink signal-to-interference-plus-noise ratio for the UE j receiving signal from UE

i is expressed as

$$\gamma\left(i,j\right)=\frac{P_i d_{i,j}^{-\alpha} h_{i,j}}{\sum_{l\neq i,l\in\chi_i} P_l d_{l,j}^{-\alpha} h_{l,j}+\sigma^2}, \tag{1}$$

where χ_i denotes the set of transmitters sharing the same spectrum with the transmitter i, $h_{i,j}$ represents the distance-independent fading of the $i-j$ link following a Gaussian distribution, $d_{i,j}$ is the distance of the uplink between user i and j, α is the path-loss exponent, and σ^2 is the variance of the additive noise. In the uplink period, the BS receiving the signal from the served CUE suffers interference from the MTC UEs to its group head. Therefore, the uplink data rate of the CUE $c\in\mathcal{U}_c$ is formulated as

$$R(c)=B_w\log\left(1+\frac{P_c d_{c,B}^{-\alpha} h_{c,B}}{\sum_{l\neq c,l\in\chi_c} P_l d_{l,B}^{-\alpha} h_{l,B}+\sigma^2}\right),\quad c\in\mathcal{U}_c, \tag{2}$$

where $B_w=\frac{B_{\rm d}}{K+N}$ and χ_c denotes the set of transmitters sharing the same spectrum with the cellular transmitter c. Similarly, the interference to MTC UE m originates from the CUEs and other MTC UEs that share the same spectrum resources. The uplink data rate for the MTC UE $m\in\mathcal{U}_M$ is represented by

$$R\left(m\right)=B_w\log\left(1+\frac{P_m d_{m,r}^{-\alpha} h_{m,r}}{\sum_{l\neq m,l\in\chi_m} P_l d_{l,m}^{-\alpha} h_{l,m}+\sigma^2}\right),\quad m\in\mathcal{U}_M, \tag{3}$$

where χ_m denotes the set of transmitters sharing the same spectrum with the MTC transmitter m and $d_{m,r}$ denotes the distance of the uplink between user m and r (r denotes the group head if UE m belongs to a group or denotes BS if UE m does not belong to any group).

Instead of the data rate, the number of connections serves as an important metric for MTC. To maximize the number of connections, group head selection, resource and power allocation, and interference distribution must be designed properly. To maximize the number of connections while satisfying the data rate requirement and the bandwidth and power constraints, if we denote the sets of served CUEs and MTC UEs as $\mathcal{K}_C$ and $\mathcal{K}_M$, respectively, the optimization problem can be formulated as

$$\begin{aligned}\mathbf{P1}:\quad &\max_{\mathcal{A}_a,\mathcal{M}_k,\boldsymbol{P},B_{\rm d}}\quad |\mathcal{K}_C|+|\mathcal{K}_M|\\ \text{s.t.}\quad &\begin{cases}\text{C1}: 0\leqslant B_{\rm d}+B_{\rm s}\leqslant B,\\ \text{C2}: |\mathcal{K}_C|+|\mathcal{K}_M|\leqslant B_{\rm s}/\Delta B,\\ \text{C3}: R\left(c\right)\geqslant R_T^c,\ c\in\mathcal{K}_C,\\ \text{C4}: R\left(m\right)\geqslant R_T^m,\ m\in\mathcal{K}_M,\\ \text{C5}: |\mathcal{K}_M|=\sum\limits_{k\in\mathcal{A}_a}|\mathcal{M}_k|,\\ \text{C6}: |\mathcal{M}_k|\leqslant Z,\\ \text{C7}: 0\leqslant\boldsymbol{P}\leqslant\boldsymbol{P}_{\max}.\end{cases}\end{aligned} \tag{4}$$

The goal of the above optimization problem is to maximize the number of connections by designing a set of group heads $\mathcal{A}_a$, the power allocation vector $\boldsymbol{P}$, and the data transmission bandwidth $B_{\rm d}$. Here, C1 is the constraint on the allocated bandwidth for data and signaling transmission, while constraint C2 ensures that the number of connections is limited by the allocated bandwidth for signaling transmission. Constraints C3 and C4 ensure that the data rate requirements of CUEs and MTC UEs are met, respectively. Constraint C5 ensures that the number of MTC connections is equal to the total number of served MTC UEs. Further, constraint C6 ensures that the number of MTC UEs in each group is limited. Finally, C7 is the transmit power constraint.

3 Group-based joint signaling and data resource allocation

The optimization problem **P1** is non-convex, and its variables are tightly coupled by the constraints. Obtaining the joint optimal group head selection and resource allocation requires an exhaustive search, which very quickly reaches an extremely high level of computational complexity. As an alternative, we seek a low-complexity suboptimal solution of **P1**. We intend to construct the optimal group dynamically in order to achieve proper data aggregation and allocate bandwidth and power to serve as many UEs as possible. Therefore, we decompose problem (**P1**) into two subproblems, i.e., a group head selection subproblem and an optimal bandwidth and power allocation subproblem. For the first subproblem, we intend to optimize the selection of the group head and decide the set corresponding to each group, thereby minimizing intragroup distances for all groups while satisfying the size of each group. For the second subproblem, we optimize the power and bandwidth using optimized access topologies in order to maximize the the number of connections. In the subsections that follow, we formulate and analyze these two subproblems.

3.1 Group head selection

To solve the group head selection subproblem, we aim to minimize the intra-group cooperative distance while adhering to the constraint defining the maximum number of connections for each group. We denote $\boldsymbol{E}_{m\times m}$ as the distance matrix, in which the element $e_{i,k}\in\boldsymbol{E}_{m\times m}$ represents the distance between the MTC UE i and the MTC UE k. This matrix is clearly a nonnegative symmetric matrix. To reduce the path loss between MTC UE and the group head, the group head selection subproblem aims to minimize the intragroup distances for all groups as follows:

$$\begin{aligned}\mathbf{P2}:\quad &\min_{\mathcal{A}_a,\mathcal{M}_k}\ \sum_{k=1}^{K}e_{i,k}\quad i\in\mathcal{M}_k,\ k\in\mathcal{A}_a\\ &C5:|\mathcal{M}_k|\leqslant Z,\ k\in\boldsymbol{K},\end{aligned}\tag{5}$$

where constraint C5 ensures that the number of UEs in each group is limited. Therefore, to solve the subproblem **P2**, we must not only identify the set of group heads $\mathcal{A}_a$ but also need to adjust the members in each group $\mathcal{M}_k$ to satisfy the condition C5.

K-means is a classical algorithm for solving the clustering or grouping problems; it is widely used because of its high computational efficiency [32]. We propose a modified K-means method to solve subproblem **P2**. Given the number of desired groups K, we select a set of group heads to minimize the sum of the squared Euclidean distance between each point and its nearest group heads using an iterative approach. Therefore, using our modified K-means algorithm, we can obtain the set of group heads.

Next, we must justify whether the constraint is satisfied. If it is satisfied, we have obtained a solution to subproblem **P2**. If it is not satisfied, we further develop an adjustment algorithm to move MTC UEs from their current groups to other groups. Our adjustment algorithm is defined as follows.

Adjustment algorithm. For arbitrary MTC UEs, the loss experienced by transferring user $i\in\mathcal{U}_M$ from the current group g to another groups k ($g\neq k$) is given as

$$L_{i,k}=e_{i,k}-e_{i,g}.\tag{6}$$

Next, we denote $\boldsymbol{x}_{m\times K}$ as the association matrix in which the element $x_{i,k}$ is an indicator variable. Here, $x_{i,k}=1$ represents the i-th MTC UE $i\in\mathcal{U}_M$ associated with the k-th group, otherwise, $x_{i,k}=0$. To minimize transfer loss, we have

$$\begin{aligned}\min_{\boldsymbol{x}}\quad &\sum_{i=1}^{M}\sum_{k=1}^{K}L_{i,k}x_{i,k}\\ \text{s.t.}\quad &\sum_{i=1}^{M}x_{i,k}\leqslant Z,\ \forall k\in\mathcal{A}_a,\\ &\sum_{k=1}^{K}x_{i,k}=1,\ \forall i\in\mathcal{U}_M,\\ &x_{i,k}\in\{0,1\},\ \forall i\in\mathcal{U}_M,\ \forall k\in\mathcal{A}_a.\end{aligned}\tag{7}$$

Algorithm 1 Group head selection

Require:

The number of group heads K, the maximum number of MTC UEs in a group Z and the distance matrix $\boldsymbol{E}$.

Ensure:

The set of group heads is selected by adopting the K-means method.

if $\exists |\mathcal{M}_j| > Z,\ \forall j \in \boldsymbol{K}$ **then**

solve the transfer loss minimization problem and move MTC UEs into corresponding group heads using the Hungarian method.

end if

Eq. (7) can be considered as an assignment problem in which each group head can be viewed as a machine with a processing capacity Z and each MTC UE can be viewed as a job requiring one unit of processing. When MTC UE i is assigned to the k-th group head, it incurs cost $L_{i,k}$. Because the total processing capacity of all machines is not equal to the number of jobs, we add $KZ-M$ virtual jobs to the job sets with costs of zero. Now, we can use the Hungarian method [20] to solve the assignment problem. Algorithm 1 is our group head selection algorithm.

3.2 Resource allocation

In the previous subsection, we solved subproblem **P2** and obtained a set of group heads denoted as $\mathcal{A}_a$. In this subsection, we turn to the second subproblem, i.e., maximizing the number of connections by adjusting bandwidth and transmit power. Given set $\mathcal{A}_a$, we transform the primary problem **P1** into a resource allocation subproblem that includes power and bandwidth allocation, i.e., subproblem **P3**, which we present as

$$\mathbf{P3}: \quad \max_{\boldsymbol{P}, B_{\mathrm{d}}} \quad |\mathcal{K}_C| + |\mathcal{K}_M|$$
$$\text{s.t.} \quad \begin{cases} \mathrm{C1}: 0 \leqslant B_{\mathrm{d}} + B_{\mathrm{s}} \leqslant B, \\ \mathrm{C2}: |\mathcal{K}_C| + |\mathcal{K}_M| \leqslant B_{\mathrm{s}}/\Delta B, \\ \mathrm{C3}: R\left(c\right) \geqslant R_T^c,\ c \in \mathcal{K}_C, \\ \mathrm{C4}: R\left(m\right) \geqslant R_T^m,\ m \in \mathcal{K}_M, \\ \mathrm{C5}: \mathcal{K}_M = \sum\limits_{k \in \mathcal{A}_a} \mathcal{M}_k, \\ \mathrm{C6}: 0 \leqslant \boldsymbol{P} \leqslant \boldsymbol{P}_{\max}. \end{cases} \tag{8}$$

From (8) above, we first reformulate constraints C3 and C4 in vector form as

$$\left[\boldsymbol{I} - a\boldsymbol{A}\right] \boldsymbol{P} \geqslant \boldsymbol{b}, \tag{9}$$

where $\boldsymbol{I}$ is the identity matrix, $a = 10^{\frac{R_T}{B_{\mathrm{d}}/(K+N)}} - 1$, $A\left(k,l\right) = h_{k,l} d_{k,l}^{-\alpha}$ is the normalized channel gain matrix

$$\boldsymbol{A}_{k,l} = \begin{cases} 0, & l = k \ \text{ or } \ l \notin \chi_k, \\ \dfrac{A\left(k,l\right)}{A\left(k,k\right)}, & l \neq k,\ l \in \chi_k, \end{cases} \tag{10}$$

and the target vector

$$\boldsymbol{b} = a\sigma^2 \left[\frac{1}{A(1,1)}, \frac{1}{A(2,2)}, \ldots, \frac{1}{A(M+N, M+N)}\right]^{\mathrm{T}}.$$

We observe here that the objective function is linear and the constraint set is convex with respect to the power profile vector $\boldsymbol{P}$ when the bandwidth B_{d} is given. Therefore, under B_{d}, we can obtain the optimal power allocation using a standard convex programming method when the feasible set is nonempty. We define the matrix $\boldsymbol{F} = a\boldsymbol{A}$; matrix $\boldsymbol{F}$ comprises nonnegative elements and is irreducible. Based on the Perron-Frobenius theorem, the following well-known lemma (proved in [33, 34]) provides a necessary and sufficient condition for the feasibility of **P3**.

Algorithm 2 Power control and bandwidth allocation algorithm

Require:

Assume that all UEs are active, $m_a \leftarrow M$, $n_a \leftarrow K$, $\mathcal{U}_{TX} = \mathcal{U}_C \cup \mathcal{U}_M$ and $B_{\mathrm{d}} = B - \mathrm{ceil}\left((m_a + n_a)\right)\Delta B$.

Ensure:

1: Set $\boldsymbol{F} = a\boldsymbol{A}$ and calculate $\rho(\boldsymbol{F})$. If $\rho(\boldsymbol{F}) < 1$, i.e., the feasibility condition is satisfied, go to Step 4; otherwise, go to Step 2.

2: If $\mathcal{U}_M \neq \emptyset$, choose the column of $\mathcal{U}_M$ such that $k = \arg\max_{k\in\mathcal{U}_M} \|\boldsymbol{f}_k\|_2$; otherwise choose the column of $\mathcal{U}_C$ such that $k = \arg\max_{k\in\mathcal{U}_c} \|\boldsymbol{f}_k\|_2$.

3: Delete the k-th row and the k-th column of $\boldsymbol{F}$ and generate a new and reduced matrix $\boldsymbol{F}$.

if $k \in \mathcal{M}_l$ **then**

  $|\mathcal{M}_l| \leftarrow |\mathcal{M}_l| - 1$.

  if $|\mathcal{M}_l| = 0$ **then**

    Set $n_a \leftarrow n_a - 1$.

  end if

else

  $m_a \leftarrow m_a - 1$.

end if

$B_{\mathrm{d}} = B - \mathrm{ceil}\left((m_a + n_a)\right)\Delta B$. Return to Step 1.

4: Solve the optimization problem and obtain the feasible solution $|\mathcal{K}_C|$ and $|\mathcal{K}_M|$.

Lemma 1. From [26], the constraint set in **P1** is nonempty if and only if the maximum modulus eigenvalue of $\boldsymbol{F}$ is less than one, i.e., $\rho(\boldsymbol{F}) < 1$, where $\rho(\boldsymbol{F})$ denotes the spectral radius of the matrix $\boldsymbol{F}$.

When the feasibility condition is satisfied, we are able to obtain $\boldsymbol{P}$ by solving **P3**. Otherwise, we obtain a feasible solution for **P3** by removing some UEs that cause maximum interference to other UEs until the feasibility condition is satisfied [26]. More specifically, the k-th row and the k-th column of $\boldsymbol{F}$ are deleted until the updated matrix satisfies $\rho(\boldsymbol{F}) < 1$ where $k = \arg\max_{k\in\mathcal{U}_M} \|\boldsymbol{f}_k\|_2$. In most of cases, CUEs have higher priority in transmission than MTC UEs; therefore, CUEs may be rejected only when all MTC UEs are refused access. We describe the complete power control and bandwidth allocation procedure in Algorithm 2. Note that ceil(x) denotes the ceiling function that obtains the minimum integer larger than x.

3.3 Computational complexity

As described in the subsections above, our primal problem **P1** is decomposed into two subproblems, which are then separately solved using Algorithms 1 and 2. Using Algorithm 1, we identify the set of group heads, which serves as the input to Algorithm 2. Using Algorithm 2, we obtain the power allocation vector and data transmission bandwidth as output.

Algorithm 1 uses the modified K-means method and the Hungarian algorithm to implement an adjustment group head selection algorithm. The computational complexity of the K-means method is $T1 \times O(M \times K)$, where $T1$, M and K denotes the number of iterations of the K-means method, the number of MTC UEs and the number of groups for all MTC UEs, respectively. The computational complexity of the Hungarian algorithm is $T2 \times O(M \times K)$ where $T2$ denotes the number of iterations of the Hungarian algorithm. Therefore, the overall complexity of our approach here is $(T1 + T2) \times O(M \times K)$.

The complexity of Algorithm 2 is $T3 \times O(M + N)$, where $T3$ denotes the number of iterations of Algorithm 2 and N denotes the number of CUEs. Combining these computational complexities, the overall computational complexity of our proposed algorithm is $(T1 + T2) \times O(M \times K) + T3 \times O(M + N)$.

4 Simulation results

In this section, we describe the simulation results of our proposed algorithms for MTC underlaid cellular networks. Unlike classic methods, our method shows the number of connections in terms of the ratio of the data rate of MTC UEs to that of CUEs, the ratio of the number of CUEs to that of MTC UEs, and the number of preambles. One of the comparable classic methods allocates the fixed signaling resource, while

Table 1 Parameter configuration

Item	Value
Total number of users	70
Cell radius	400 m
Data rate requirement of CUEs	0.45 Mb/s
Number of groups	10 (normal)
Bandwidth	20 MHz
Fading	Rayleigh fading

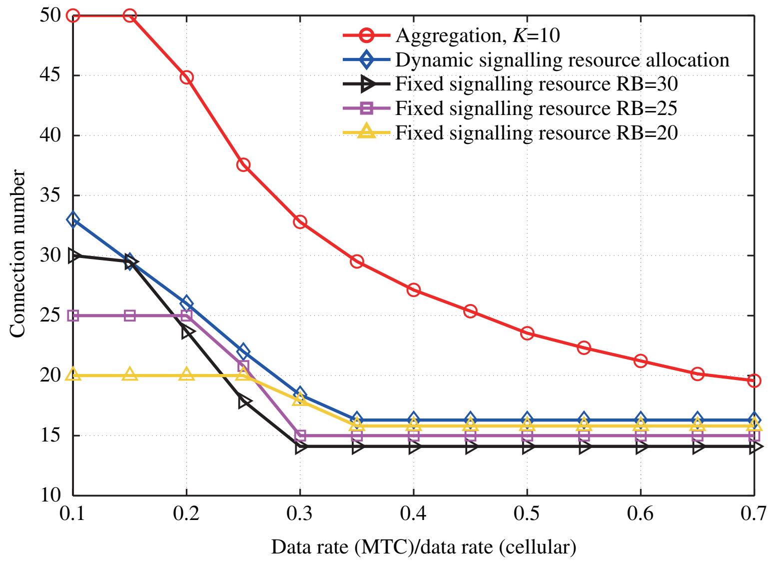

Figure 2 (Color online) Numbers of connections versus the ratio of data rate of MTC to that of cellular communications for our proposed algorithm (i.e., aggregation $K = 10$) and other classic approaches.

a second classic method is our proposed dynamic signaling resource allocation approach but without the group-based data aggregation. In our simulations, we use three different RB values for signaling transmission, i.e., 20 RBs, 25 RBs, 30 RBs. Moreover, we explore the effect the number of groups on the number of connections. The simulation parameters are listed in Table 1.

In Figure 2, we show how our group-based proposed algorithm performs with a significant gain in terms of the number of connections when compared with the fixed and dynamic signaling resource allocation schemes; these data verify that aggregation is indeed an effective approach to increase the number of connections. Moreover, such advantages our proposed group-based algorithm are more obvious when the data-rate differences between MTC and cellular communications are large, in other words, our proposed algorithm is appropriate for the small data transmissions inherent to MTC. Furthermore, the number of connections decreases as the data rate ratio of MTC to cellular communications increases. It then reduces to a stable value, as shown in Figure 2. The intersections of the curves representing different fixed signaling resources indicate that the limiting factor on the the number of connections changes from the signaling resource constraint to the data resource constraint as the MTC rate requirement increases.

In Figure 3, we present results that verify that a smaller ratio of the number of CUEs to that of MTC UEs gives a more distinct advantages to our group-based proposed algorithm over the fixed and dynamic signaling resource allocation schemes. The intersections of the curves indicate that the limiting factor on the the number of connections changes from the signaling resource constraint to the data resource constraint as upon simulation with a lower number of MTC UEs.

As shown in Figure 4, the number of connections increases to a stable value with the number of preambles, where the stable points results from the limited data transmission resources. The dynamic

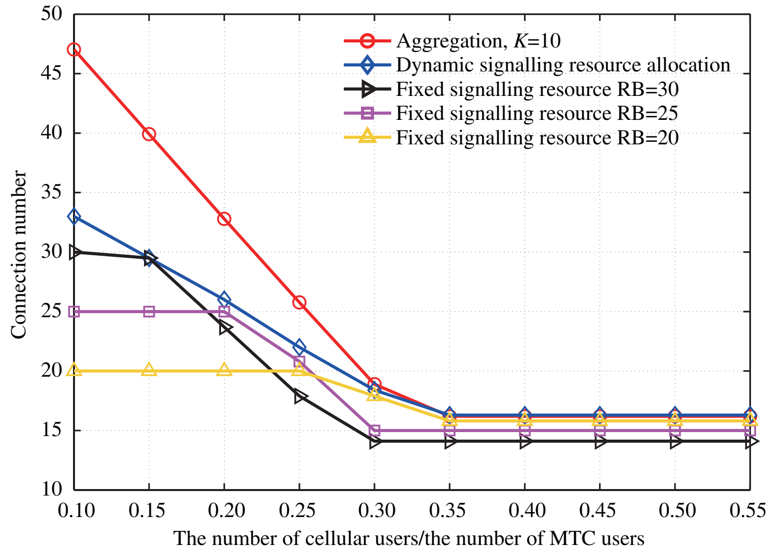

Figure 3 (Color online) Numbers of connections versus the number of CUEs to that of MTC UEs for our proposed algorithm (i.e., aggregation $K = 10$) and other classic approaches.

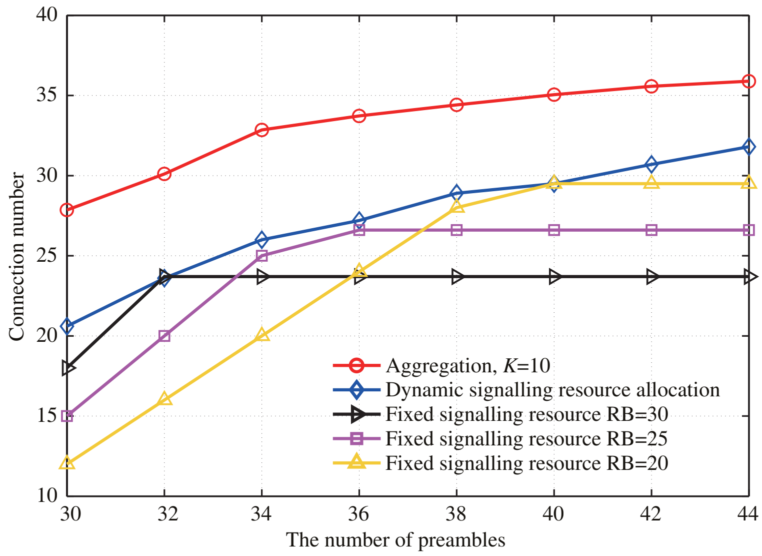

Figure 4 (Color online) Numbers of connections versus the number of preambles for our proposed algorithm (i.e., aggregation $K = 10$) and other class approaches.

resource allocation schemes perform better than the fixed ones here, with our group-based approach delivering a high performance. Another remarkable result is that the lower the RB values, the higher the number of connections as the number of preambles increase. While the limiting factor on the number of connections changes from the signaling resource constraint to the data resource constraint with an increasing number of preambles, the system may save more RBs for data transmission to support more connections.

In Figure 5, we present the number if groups versus the number of connections. From Figure 5, we conclude that varying the number of groups did not contribute to changes in the number of connections because increases in the number of groups reduce the RBs allocated to each group, which then limits the number of connections while guaranteeing the data rate requirements (i.e., C3 and C4 in (8)). Therefore, the number of connections remains stable even though the number of groups increases.

Finally, in Figure 6, we show the convergence of our proposed algorithm versus the number of iterations. Our proposed algorithm converges very rapidly after only approximately seven iterations. Note that the

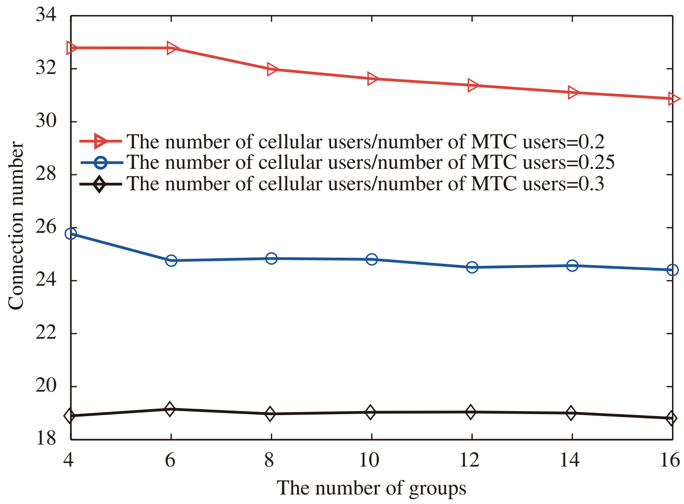

Figure 5 (Color online) Numbers of connections versus the number of groups in our proposed algorithm.

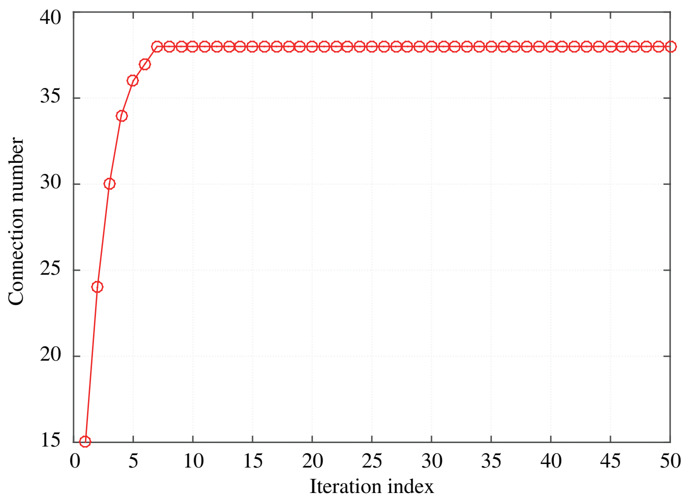

Figure 6 (Color online) Numbers of connections in terms of the number of iterations in our proposed algorithm.

computational complexity of our proposed algorithm has been analyzed thoroughly in Subsection 3.3.

5 Conclusion

In this paper, we formulated a group-based joint signaling and data resource optimization model to maximize the number of connections constrained by network resources and data rate requirements. Given the non-convexity of this problem, we decomposed it into two subproblems, i.e., a group-based data aggregation subproblem and a resource allocation subproblem. Our simulation results show that our proposed algorithm significantly increased the number of connections when compared with other classic methods. Furthermore, we deduced that the limiting factor on the number of connections changes with the ratio of the number of MTC UEs to that of CUEs and with the ratio of data requirements of MTC UEs to that of CUEs.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61461136002, 61631005) and Fundamental Research Funds for the Central Universities.

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Boswarthick D, Elloumi O, Hersent O. M2M Communications: A Systems Approach. Hoboken: John Wiley & Sons, 2012

2 Lucero S. Maximizing Mobile Operator Opportunities in M2M: the Benefits of an M2M-Optimized Network. ABI Research. 2010

3 Malm A, Tobias R. Wireless M2M and Mobile Broadband Services. Berg Insight. 2007

4 Taleb T, Kunz A. Machine type communications in 3GPP networks: potential, challenges, and solutions. IEEE Commun Mag, 2012, 50: 178–184

5 Ksentini A, Hadjadj-Aoul Y, Taleb T. Cellular-based machine-to-machine (M2M): overload control. IEEE Netw, 2012, 26: 54–60

6 Lien S Y, Liau T H, Kao C Y, et al. Cooperative access class barring for machine-to-machine communications. IEEE Trans Wirel Commun, 2012, 11: 27–32

7 Oh C Y, Hwang D, Lee T J. Joint access control and resource allocation for concurrent and massive access of M2M devices. IEEE Trans Wirel Commun, 2015, 14: 4182–4192

8 Laya A, Alonso L, Alonso-Zarate J. Is the random access channel of LTE and LTE-A suitable for M2M communications? A survey of alternatives. IEEE Commun Surv Tut, 2014, 16: 4–16

9 3GPP. Study on RAN improvements for machine type communications. TR 37.868 V0.5.1. 2010. http://www.3gpp.org/ftp/Specs/archive/37_series/37.868/

10 Zheng K, Ou S, Alonso-Zarate J, et al. Challenges of massive access in highly dense LTE-advanced networks with machine-to-machine communications. IEEE Wirel Commun, 2014, 21: 12–18

11 Li X, Li D, Bai Y. Distributed differential admission control algorithm for delay-tolerant machine-to-machine devices. IET Commun, 2015, 9: 1230–1239

12 Utkovski Z, Eftimov T, Popovski P. Random access protocols with collision resolution in a noncoherent setting. IEEE Wirel Commun Lett, 2014, 4: 445–448

13 Ksentini A, Taleb T, Ge X, et al. Congestion-aware MTC device triggering. In: Proceedings of IEEE International Conference on Communications, Sydney, 2014. 294–298

14 Si P, Yang J, Chen S, et al. Adaptive massive access management for QoS guarantees in M2M communications. IEEE Trans Veh Technol, 2015, 64: 3152–3166

15 Farhadi G, Ito A. Group-based signaling and access control for cellular machine-to-machine communication. In: Proceedings of the 78th Vehicular Technology Conference, Las Vegas, 2013

16 Huq R M, Moreno K P, Zhu H, et al. On the benefits of clustered capillary networks for congestion control in machine type communications over LTE. In: Proceedings of the 24th International Conference on Computer Communication and Networks, Las Vegas, 2016

17 Taleb T, Ksentini A. An efficient scheme for MTC overload control based on signaling message compression. In: Proceedings of IEEE Global Communications Conference, Atlanta, 2013. 342–346

18 Hossain M I, Laya A, Militano F, et al. Reducing signaling overload: flexible capillary admission control for dense MTC over LTE networks. In: Proceedings of the 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, Hong Kong, 2015. 1305–1310

19 Kwon T, Cioffi J M. Random deployment of data collectors for serving randomly-located sensors. IEEE Trans Wirel Commun, 2011, 12: 2556–2565

20 Niyato D, Wang P, Dong I K. Performance modeling and analysis of heterogeneous machine type communications. IEEE Trans Wirel Commun, 2014, 13: 2836–2849

21 Ge X H, Chen J Q, Wang C X, et al. 5G green cellular networks considering power allocation schemes. Sci China Inf Sci, 2016, 59: 022308

22 Zhang H, Huang S, Jiang C, et al. Energy efficient user association and power allocation in millimeter wave based ultra dense networks with energy harvesting base stations. IEEE J Sel Area Commun, 2017. doi: 10.1109/JSAC.2017.2720898

23 Abbas R, Shirvanimoghaddam M, Li Y, et al. Random multiple access for M2M communications with QoS guarantees. IEEE Trans Wirel Commun, 2017, 65: 2889–2903

24 Dhillon H S, Huang H C, Viswanathan H, et al. On resource allocation for machine-to-machine (M2M) communications in cellular networks. In: Proceedings of IEEE Globalcom Workshops, Anaheim, 2012. 1638–1643

25 Pang Y C, Chao S L, Lin G Y, et al. Network access for M2M/H2H hybrid systems: a game theoretic approach. IEEE Commun Lett, 2014, 18: 845–848

26 Zhang Z, Long K, Wang J, et al. On swarm intelligence inspired self-organized networking: its bionic mechanisms, designing principles and optimization approaches. IEEE Commun Surv Tut, 2014, 16: 513–537

27 Zhang Z, Long K, Wang J. Self-organization paradigms and optimization approaches for cognitive radio technologies: a survey. IEEE Wirel Commun, 2013, 20: 36–42

28 Zhang H, Jiang C, Mao X, et al. Interference-limited resource optimization in cognitive femtocells with fairness and imperfect spectrum sensing. IEEE Trans Veh Tech, 2016, 65: 1761–1771

29 Zhang G, Wei H, Liang T, et al. A novel two-stage dynamic spectrum sharing scheme in cognitive radio networks. China Commun, 2016, 13: 236–248

30 Zhang H, Jiang C, Beaulieu N C, et al. Resource allocation for cognitive small cell networks: a cooperative bargaining game theoretic approach. IEEE Trans Wirel Commun, 2015, 14: 3481–3493

31 Lo A, Law Y W, Jacobsson M, et al. Enhanced LTE-advanced random-access mechanism for massive machine-to-machine (M2M) communications. In: Proceedings of the 27th Meeting of Wireless World Research Form (WWRF), Doha, 2011

32 Mitchell T M. Machine Learning. New York: McGraw-Hill Education Ltd, 1997

33 Foschini G J, Miljanic Z. A simple distributed autonomous power control algorithm and its convergence. IEEE Trans Veh Tech, 1993, 42: 641–646

34 Lee N, Lin X, Andrews J G, et al. Power control for D2D underlaid cellular networks: modeling, algorithms, and analysis. IEEE J Sel Area Commun, 2013, 33: 1–13