User Association for Load-Balance in Heterogeneous M2M Networks

Wen-feng MA¹, Hui TIAN¹*, Cong WANG¹ and Liu-bin SONG²

¹College of Field Engineering, Army Engineering University, Nanjin, Jiangsu, China
²China Communication Technology (Nanjing) Co., Ltd, Nanjin, Jiangsu, China

*Corresponding author

Keywords: Energy efficiency, Load balance, User association, M2M communication, Heterogeneous networks.

Abstract. In this paper, we propose an energy efficient user association scheme for uplink heterogeneous networks with machine-to-machine (M2M) and human-to-human (H2H) coexistence. A group based random access protocol is designed for massive number of machine-type-communications (MTC) user equipment’s (UEs) transmissions. A user association problem is formulated considering the load balance among multiple BSs. A distributed iterative algorithm is developed to solve the optimization problem. In addition, the convergence of the proposed algorithm is proved. Simulation results show that our proposed scheme outperforms other schemes in terms of energy efficiency and load balance.

Introduction

Recently, machine-to-machine (M2M) applications are playing a more and more important role in our daily life, such as, smart grid, vehicular telematics, healthcare, public safety and so on [1]. The characteristics of M2M communications can be summarized as massive number of devices, extremely low power consumption, small burst transmission, group control, one-way data traffic (uplink), etc. [1]. Heterogeneous networks (HetNets) integrating multiple wireless access technologies can provide a promising architecture for supporting M2M communications. In a heterogeneous cell, the problem of user association, i.e., the problem of determining which base station (BS) serves a particular user by certain rules, is one of major challenges. In addition, energy efficiency optimization has become an important research area in both M2M communications and user association. Therefore, this paper investigates optimization problem of user association aiming at maximizing the network energy efficiency in a wireless uplink heterogeneous network, in which the characteristics of the M2M communications are considered.

In the literature, the user association problem has been explored extensively [2]-[13]. In [2], a user association scheme for load balancing was proposed in HetNets. A distributed pricing-based user association scheme was developed for load balancing in downlink heterogeneous cellular networks [3]. In [4], a unified framework was designed for QoS-driven user association. An opportunistic user association scheme was investigated in [5]. An uplink user association scheme based on college admissions game was proposed in [6]. [8] proposed a user association scheme aiming at maximizing BSs’ energy efficiency. In [9], authors developed a joint BS operation and user association algorithm in HetNets. However, the schemes proposed in [2]-[10] are not suitable for massive number of MTC UEs’ transmissions due to allocating fixed resources to each UE. Accordingly, the above mentioned schemes [2]-[10] can not be applied to the user association for M2M communications directly.

In this paper, we propose an energy efficient user association scheme for uplink HetNets. A group based random access protocol is designed for massive number of MTC UEs’ transmissions. A user association problem for UEs’ energy efficiency maximization is formulated considering load balance among multiple BSs. A distributed algorithm based on dual theory is developed to solve the proposed problem. Finally, simulation results show that our proposed algorithm can achieve the same performance as the exhaustive search algorithm when the number of UEs is small. Regardless of the
number of UEs, the proposed scheme outperforms the existing schemes in terms of energy efficiency and load balance.

System Model
Consider an uplink HetNet consisting of N BSs and D MTC UEs, as shown in Figure 1, where there are multiple tiers of BSs, such as macro, pico and femto. The set of all BSs is denoted by \( \mathcal{B} \triangleq \{1, 2, \ldots, N\} \) and the set of UEs associated with BS \( j \) is \( \mathcal{B}_j \). The \( D \) MTC UEs are indexed by the set \( \mathcal{D} \triangleq \{1, 2, \ldots, D\} \). In the HetNet, the smallest wireless resources structure is denoted by a resource block (RB), which contains both time and frequency domains resources. The number of available RBs at a given BS is related to scheduling interval duration and system bandwidth. We assume that the total number of available RBs at BS \( j \) (\( j \in \mathcal{B} \)) is indexed by the set \( \mathcal{L}_j \triangleq \{1, 2, \ldots, L\} \).

The channel model between UEs and BSs is assumed to be flat-fading and slow-fading. The coherence time and bandwidth of channel are larger than the time and frequency domains interval of an RB, respectively. In this paper, we suppose that a certain interference cancellation technique is adopted by all BSs, such as uplink coordinated scheduling with MU-MIMO [11], the interference caused by UEs belonged to adjacent cells can be ignored. Consequently, the received Signal-to-Noise Ratio (SNR) from MTC UE \( i \) (\( i \in \mathcal{D} \)) to BS \( j \) can be expressed as \( \gamma_{ij} = \frac{P_i G_{ij} |h_{ij}|^2}{N_0} \), where \( P_i \) denotes the transmit-tted power of MTC UE \( i \), \( h_{ij} \) is the Rayleigh channel fading coefficient between MTC UE \( i \) and BS \( j \), \( G_{ij} \) denotes the effect of large scale fading factor, such as path loss, log normal shadowing and antenna gains, and \( N_0 \) denotes the power of Additive White Gaussian Noise. \( h_{ij} \) is modeled as zero-mean, independent and circularly-symmetric complex Gaussian random variables with variance \( \sigma^2_{h_{ij}} \).

![Figure 1. Illustration of a three-tier heterogeneous cellular network.](image)

In M2M communications, massive MTC UEs’ transmissions are one of key M2M features. An extremely large number of MTC UEs simultaneously or nearly simultaneously attempt to access BS in many application scenarios, such as public safety, healthcare, metering, etc. For MTC UEs, due to the constrained resources, it is impossible for the BS to allocate fixed wireless resources to every MTC UE. Consequently, a random access (RA) mechanism [15]-[16], in which each MTC UE competes for the left RBs at the BS, is adopted to solve the problem of massive access requests and obtain load balancing among all MTC UEs.

With the number of MTC UEs increasing, the network will be congested and overloaded due to competition among MTC UEs despite the small burst size of M2M traffic data. Therefore, a group based RA mechanism is designed to improve the performance of RA protocol and avoid all available RBs used by one MTC UEs at a BS. Let \( \beta_j \) denote the maximum ratio of available RBs which are allocated to MTC UE \( d \) (\( d \in \mathcal{B}_j \)) for contending at BS \( j \). We divide the all available RBs into \( 1/ \beta_j \) group. The MTC UEs competing for the same RBs form a group. Given \( \bar{L}_j \) available RBs at BS \( j \), the probability of the MTC UE \( d \) successfully capturing a RB is \( p_{d,j} = \frac{\bar{L}_j}{\beta_j} \left( 1 - p_{d,j} \right)^{\beta_j - 1} \), where \( p_{d,j} \) denotes the probability of the MTC UE \( d \) selecting a certain RB in its group at BS \( j \). We assume that
the probability of each MTC UE selecting a certain RB at the same group is equal, i.e., \( p_{sj} = p_{dj} = 1/(\beta_j L_j) \), \( \forall s, d \in B, s \neq d, \forall j \in B \). For simplicity, let \( p_j \) describe the probability of a certain RB selected by a MTC UE at BS \( j \), \( \forall j \in B \), i.e., \( p_j = 1/(\beta_j L_j) \).

The contention-based RA mechanisms in LTE is used for applying for transmission resources and re-establishing a connection upon failure. BS will allocate a RB to a MTC UE if the MTC UE competes successfully during access stage. If a collision occurs, the UE performs a retransmission with the uniform backoff algorithm. Therefore, the interference caused by multiple MTC UEs during the data transmission phase will not considered in this paper.

**Problem Formulation**

First, we define the energy efficiency of a UE as the number of transmitted bits per unit joule, i.e., bits-per-joule. The energy efficiency of MTC UE \( d \) is the ratio of the expected value of MTC UE \( d \)'s maximum rate and MTC UE \( d \)'s transmitted power, and can be computed by

\[
E_{e}^{M}_{dj} = \frac{E_{f}^{L}_{dj}}{P_{d}} = \beta_j L_j p_{dj}^{g} \text{M} f \log_2(1+\gamma_{dj})/P_{d}
\]

where \( E_{f}^{L}_{dj} \) is the expected value of the MTC UE \( d \)'s maximum rate.

Then, we define a indicator matrix \( X = [x_{ij}]_{M \times D}, i \in D, j \in B \), where each entry \( x_{ij} \) is a binary variable, \( x_{ij} \in \{0,1\} \). When the MTC UEs \( i \) is associated with BS \( j \), \( x_{ij} = 1 \); otherwise, \( x_{ij} = 0 \). In practice, not every UE can access any BSs due to the limitation of BSs' coverage. Let \( S' \) denote the set of available accessing BSs for MTC UE \( i \) (\( i \in D \)). Then, the overall energy efficiency is expressed as

\[
E_{overall} = \sum_{d=1}^{D} \sum_{j=1}^{\left|S'\right|} x_{dj} E_{e}^{M}_{dj}
\]

(2)

where \( |A| \) denotes the cardinality of set \( A \). Consequently, the problem of user association for energy efficient is how to determine the indicator matrix \( X \) for maximizing the overall energy efficiency.

User association scheme for MTC UEs is designed for maximizing the overall MTC UEs' energy efficiency and balancing load among all BSs. Note that, although using linear utility function can achieve an energy-efficient optimal solution, we adopt logarithmic utility function because it can achieve load balancing and some level of fairness among all MTC UEs, which can also reduce the collision probability at BSs. Then, similar to the first sub-problem, the MTC UEs association problem is also to determine a binary indicator matrix \( XD \) to maximize the sum of MTC UEs' utility of energy efficiency, which can be written as

\[
\text{maximize} \ \sum_{d=1}^{D} \sum_{j=1}^{\left|S'\right|} x_{dj} \log \left(E_{e}^{M}_{dj}(x_{dj})\right)
\]

subject to (a) \( \sum_{j=1}^{\left|S'\right|} x_{dj} = 1, \ \forall d \in D \)

(b) \( \sum_{d=1}^{D} x_{dj} = B_j, \ \forall j \in B \)

(c) \( \sum_{j=1}^{N} B_j = D \)

(d) \( x_{dj} \in \{0,1\}, \ \forall (d, j) \in D \times B \)

(3)

where constraint (3a) ensures that each MTC UE can only associate with one BS, and constraint (3c) reflects that the all MTC UEs are served.

Note that the combination optimization problem in (3) is NP-hard. To overcome this, similar to [2]-[4], the constraint \( x_{dj} \in \{0,1\} \) is relaxed to continuous values with \( 0 \leq x_{dj} \leq 1 \), i.e., the MTC UEs can be associated with more than one BSs. After relaxation of parameter \( x_{dj} \), the optimization (3) can be rewrite as
maximize \( \sum_{d=1}^{D} \sum_{j=1}^{|\mathcal{S}|} x_{dj} \log \left( E e^{x_{dj}}(x_{dj}) \right) \)

subject to
(a) \( \sum_{j=1}^{|\mathcal{S}|} x_{dj} = 1, \quad \forall d \in \mathcal{D} \)
(b) \( x_{dj} = B_j, \quad \forall j \in \mathcal{B} \)
(c) \( \sum_{j=1}^{N} B_j = D \)
(d) \( 0 \leq x_{dj} \leq 1, \quad \forall (d, j) \in \mathcal{D} \times \mathcal{B} \)

Since the optimization of (4) is no longer combinatorial problem, the complexity can be relatively reduced compared to the optimization (3). However, in practice, it is more difficult to perform resources scheduling among multiple BSs while considering the multiple-BS association. In the following section, to solve the optimization (3), we propose a distributed algorithm that also takes into account the single-BS association constraint.

Algorithm

In this section, we design a distributed algorithm to obtain the solutions of the user association problems. We will use a dual decomposition method to solve the dynamical MTC UEs' association problem (3).

Similar to the methods in [2] and [3], we take the parameter \( B_j \) as an optimization variable in following analysis. Then, we define the Lagrangian function associated with the problem (4) as

\[
L(X^D, B, \mu, \nu) = \sum_{d=1}^{D} \sum_{j=1}^{|\mathcal{S}|} x_{dj} (a_{dj} - \mu_j) + \sum_{j=1}^{N} B_j \left[ (B_j - 1) \log (1 - p_j) \right] \\
+ \sum_{j=1}^{N} \mu_j B_j - \nu \left( \sum_{j=1}^{N} B_j - D \right)
\]

where \( a_{dj} = \beta_j L_n \Delta f \log_2 (1 + \gamma_{dj}) / P_d \), \( B = (B_1, ..., B_N) \), \( X^D = (x_1, ..., x_d, ..., x_D) \), thereinto, \( x_d \) denotes the MTC UE \( d \)'s association results. \( \mu = (\mu_1, ..., \mu_N) \) \((\mu_j \geq 0, \forall j \in \mathcal{B})\) denotes the Lagrange multipliers associated with the constraints in (4b), similarly \( \nu \) \((\nu \geq 0)\) represents Lagrange multiplier associated with the constraint in (4c).

We define the Lagrange dual function as the maximum value of the Lagrangian function over \( X^D \) and \( B \):

\[
g(\mu, \nu) = \sup_{X^D, B} L(X^D, B, \mu, \nu)
\]

Considering the single-BS association constraint, the solution of optimization in (4) with respect to \( X^D \) is

\[
x_{dj}^* = \begin{cases} 
1 & j = \arg \max_j (a_{dj} - \mu_j) \\
0 & j \neq \arg \max_j (a_{dj} - \mu_j)
\end{cases}
\]

Note that if there are multiple maximizers in (7), MTC UE can choose any one of them.

Then, taking derivative to \( L(X^D, B, \mu, \nu) \) with respect to \( B_j \), we have
\[ B_j^* = \frac{\nu - \mu_j}{2\log(1-p_j)} + \frac{1}{2} \]  

(8)

Substituting (7) and (8) into (6), the Lagrange dual function can be rewritten as

\[ g(\mu, \nu) = \sum_{d=1}^{D} \max_{j \in S_d} \left\{ \log(a_{ij}) - \mu_j \right\} + \frac{1}{4} \sum_{j=1}^{N} \frac{\nu - \mu_j + \log(1-p_j)}{\log(1-p_j)} \]

\[ \times \left[ \mu_j - \nu \frac{1}{2} \log(1-p_j) \right] + \nu D \]

(9)

The Lagrangian dual problem of (4) is to minimize \( g(\cdot) \) over the dual function over \( \mu \) and \( \nu \),

\[ \minimize g(\mu, \nu) \]

subject to \( \nu > 0, \mu_j > 0, \forall j \in B \)

(10)

The dual optimization problem (10) can be solved by the subgradient projection method [17]. The Lagrange multipliers \( \mu \) and \( \nu \) are updated iteratively in the opposite direction to the gradient. First, it can be seen from (9) that dual function is a differentiable function of \( \nu \), by setting its gradient to be 0, the optimal \( \nu \) can be computed by

\[ \nu(t+1) = \frac{\sum_{j=1}^{N} \frac{\mu_j(t)}{2\log(1-p_j)} - \frac{1}{2} N + D}{\sum_{j=1}^{N} \frac{1}{2\log(1-p_j)}} \]

(11)

where \( t \) denotes the iteration time. However, since that dual function is not a differentiable function of \( \mu_j \), by subgradient method, the Lagrange multiplier \( \mu_j \) is updated by

\[ \mu_j(t+1) = \left[ \mu_j(t) - \alpha(t) \left( \frac{\nu(t) - \mu_j(t)}{2\log(1-p_j)} + \frac{1}{2} - \sum_{d=1}^{D} x_{ij}(t) \right) \right] \]

(12)

where \( \alpha(t) \) represents the step size, \( [x]^+ \) denotes the maximum of the argument of \( x \) and 0, and \( x_{ij}(t) \) can be obtained according to \( \mu_j(t) \).

In summary, the distributed iterative algorithm can be briefly described as: each BS broadcasts its price \( \mu_j \). MTC UEs make decision to associate which BS according to \( \mu_j \). Then, BSs update their prices \( \mu \) and \( \nu \) based on MTC UEs’ decisions. This procedure repeats and converges to a final solution, i.e., an indicator matrix for MTC UEs \( XD \) if \( |g(\mu(t), \nu(t)) - g(\mu(t-1), \nu(t-1))| < \epsilon \), where \( \epsilon \) is an arbitrary number. The detail of the algorithm is summarized as the Algorithm 1.
Algorithm 1: Distributed Iterative Algorithm

While \[ |g(\mu(t),\nu(t)) - g(\mu(t-1),\nu(t-1))| > \varepsilon \]

For each MTC UE \( d, d=1:D \)

Based on \( \mu(t) \), MTC UE \( d \) calculates \( x_d(t) \) according to (7) and reports its association results to the BSs;

End For

For each BS \( j, j=1:N \)

Based on \( \mu(t) \) and \( \nu \), BS \( j \) updates \( B_j(t+1) \) according to (8).

Based on \( X^D(t) \), BS \( j \) updates \( p_j(t+1) \) according to (12).

End For

Based on \( \mu(t) \), all BSs update \( \nu(t+1) \) according to (11).

\( t = t+1; \)

End While

Simulation Results and Discussion

In this section, numerical simulations are presented to confirm our analysis developed in the previous sections. A three-tier HetNet is considered in the simulations. The transmitted powers of MTC UE are set to 18 dBm, respectively. All BSs and MTC UEs are located within a square area with dimensions of 1000m×1000m. One macro BS is fixed at the center of the square area, the rest of BSs and all UEs are randomly located in this area. Note that \( N_{macro}, N_{pico}, \) and \( N_{femto} \) denote the numbers of macro BSs, pico BSs and femto BSs, respectively. Small scale Rayleigh fading with \( \sigma^2 \), and large scale path loss and log-normal shadowing are considered for the channel model. We use path loss \( 128.1+37.6 \log_{10}(r), 140.7+36 \log_{10}(r), \) and \( 127+30 \log_{10}(r) \) for macro BS, pico BSs and femto BSs, respectively. Power spectral density of noise is -174dBm/Hz. System bandwidth is 10 MHz. Bandwidth of a RB is 180 KHz.

Meanings of the terms used in Figs. 2-3 are as follows. UAEE scheme refers to our proposed user association schemes for energy-efficient. UALB scheme refers to the user association scheme for load balancing in [2]. Max SNR scheme means that each MTC UEs will be assigned to the BS with largest received SNR [18]. ES stands for exhaustive search algorithm. RA denotes random selection scheme that means that the MTC UEs associate with a BS randomly.

Figure 2 shows the convergence of our designed algorithm under different numbers of MTC UEs. It can be seen from Figure 2 that the convergence speed drops with the number of MTC UEs \( D \) increasing. Even so, after less than 150 iterations, our algorithm converges to final solution when \( D=1000 \), which demonstrates that the proposed iterative algorithm is suitable for the scenarios where there is larger number of MTC UEs. Note that the step size \( \alpha(t) \) is set to \( 0.5/t \) and \( \varepsilon = 0.001 \) in this and following simulations.

![Figure 2](image-url)

Figure 2. The convergence speed of Algorithm 1 under different numbers of MTC UEs.
Figure 3 compares the number of MTC UEs at each BS among different schemes. The Max SNR scheme results in very unbalance loads: femto BSs are overloaded because of the better channel quality between UEs and femto BSs, while macro BS that possesses the largest number of BSs serves few UEs. Since the user association results of the UALB scheme are more related to the channel condition between UE and BS, there is still slight overload at the femto BSs because of their limited radio resources. While the UAEE scheme can adaptively adjust the user association results according to the number of both UEs and available RBs at BSs.

![Figure 3. The average number of MTC UEs at each BS for different schemes.](image)

**Conclusions**

In this paper, we propose a user association scheme for energy efficiency maximization in the uplink HetNet. A group based random access user association problem is formulated as a maximization of the overall UEs’ energy efficiency while considering the load balance among multiple BSs. Then, a distributed iterative balance among multiple BSs. Then, a distributed iterative algorithm based on the dual decomposition method is designed for obtaining the user association results. Numerical results show that our proposed scheme outperforms the existing methods in terms of the load balance and the average overall energy efficiency.

**Acknowledgement**

This work is supported by National Natural Science Foundation of China (No. 61771486, No. 61671472), Jiangsu Province Natural Science Foundation (No. BK20160079).

**References**


