A Belief Propagation Approach for Distributed User Association in Heterogeneous Networks

Youjia Chen*[†], Jun Li*, He (Henry) Chen*, Zihuai Lin*, Guoqiang Mao^{‡§}, Jianyong Cai[†]

*School of Electrical and Information Engineering, The University of Sydney, Sydney, Australia

[†]College of Photonic and Electronic Engineering, Fujian Normal University, Fuzhou, China

[‡]School of Computing and Communications, The University of Technology Sydney, Sydney, Australia

[§]National ICT Australia (NICTA), Australia

Email: {youjia.chen, ju.li1, he.chen, zihuai.lin}@sydney.edu.au, g.mao@ieee.org, cjy@fjnu.edu.cn

Abstract-In heterogeneous networks (HetNets), the load between macro-cell base stations (MBSs) and small-cell BSs (SBSs) is imbalanced due to transmit power disparities and ad-hoc deployment of SBSs. This significantly impacts the system performance and user experience. Associating more users to the SBSs is an effective way to solve this problem. In this paper, we formulate the user-BS association problem as a distributed optimization problem with proportional fairness as the objective. Specifically, we propose a novel distribute algorithm based on the belief propagation (BP) method to solve the user-BS association problem via iteratively message passing between the users and BSs. Also, we develop an approximation calculation in the BP method to reduce the computational complexity and transmission overhead of message passing. Simulation results show that the proposed algorithm well approaches the optimal system performance (by exhausting search) with low complexity and fast convergence.

I. INTRODUCTION

The concept of LTE-Advanced-based HetNets has been proposed as a promising solution to meeting the explosive growth of data demand. It improves spectral efficiency per unit area by increasing the cell density and provides a uniform experience to users anywhere inside the cell. HetNets are composed of regularly deployed MBSs and overlapped SBSs (like picos, femtos and relays). These low-cost and flexibly deployed SBSs eliminate the coverage holes and increase the capacity in hot-spots. In HetNets without appropriate user association, a majority of users are normally connected to MBSs since they have larger transmit power compared to SBSs. This unbalanced load results in sub-optimal system performance and highly limits the benefits of cell splitting.

There have been many efforts in open literature toward this load-balancing problem. One of the effective solutions is adjusting the coverage of small cells to make more users connect to SBSs. Cell range expansion studied for LTE Advanced [1] expands the coverage of small cells by adding a bias to the received signal from SBSs. In [2], the authors propose a 'Picocell First' scheme to associate users to a SBS as long as the signal-to-interference-plus-noise ratio (SINR) is larger than a tuning parameter.

Meanwhile, many works formulate the user association problem as an optimization problem. Different kinds of objective functions are adopted in related literature. For instance, [3] chooses max-min fairness, [4] utilizes global proportional fairness and proves it to be NP-hard, [2] and [5] transform global proportional fairness to local proportional fairness, and [6] encompasses several different utility functions.

In this paper, we formulate the user association problem as a distributed optimization problem by utilizing proportional fairness as the objective. Since the brief propagation algorithm has showed its advantages to solve distributed optimization problems [7], [8], we aim at optimizing user association via BP in an effective and efficient manner. We first decompose the global proportional fairness to local ones. Then we develop a novel BP algorithm to address the formulated optimization problem in a distributed manner.

The main contributions of this paper are summarized as follows. (1) We develop a factor graph to model the user-BS association problem, which decomposes the centralized network-wide optimization into the maximization of local functions at individual BSs. (2) We propose a novel distribute algorithm based on the BP method via iteratively message passing between the users and BSs. (3) We develop an approximation scheme to reduce the computational complexity and transmission overhead of message passing in the BP algorithm.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this paper, we consider the downlink in an LTE-Advanced HetNet composed of MBSs, pico-cell and femto-cell BSs. The pico-cell BSs are assumed to transmit at a much lower power than MBSs, while the transmit power of femto-cell BSs is the lowest. Some femto-cell BSs are considered be closed femtos, that is, only allowing access to their closed subscriber group (CSG) members. We make the following two assumptions:

1) Each user can only associate with one of the BSs.

2) Channels between users and BSs are considered to be static during the optimization process of the association.

Let B denote the set of BSs and U denote the set of users. Then the received SINR of user $i \in U$ from BS $j \in B$ can be written as

$$SINR_{ij} = \frac{P_j g_{ij}}{N_0 + \sum_{h \in B, h \neq j} P_h g_{ih}},\tag{1}$$

where P_j is the transmit power of BS j, N_0 is the power of additive white Gaussian noise, and g_{ij} is the channel gain between the user i and BS j that includes the path loss, shadow

fading and antenna gain. Then, the spectral efficiency γ_{ij} can be written as $\gamma_{ij} = \log (1 + SINR_{ij})$.

We denote the user association variable by x_{ij} , where $x_{ij} = 1$ if user *i* is associated with BS *j*, and $x_{ij} = 0$ otherwise. Also, we denote by *W* the bandwidth of BSs, by t_{ij} the serving time and by R_{ij} the rate that the user *i* obtained for BS *j*. So, the effective transmission rate of user *i* can be given by:

$$R_i = \sum_{j \in B} x_{ij} R_{ij} = \sum_{j \in B} x_{ij} W_j \gamma_{ij} t_{ij}.$$
 (2)

We assume that all users have equal priority. Then the BSs offer the same service to each user. Let the number of users associated with the BS j be N_j , and we have $t_{ij} = \frac{1}{N_j}$. We select proportional fairness as our global objective function, which has been proved to be associated with the logarithmic utility function [9]. We can get the problem formulation as

$$\max_{\{x_{ij}\}} \sum_{i \in U} \log(R_i) = \max_{\{x_{ij}\}} \sum_{j \in Bi \in U} \log(\frac{W}{N_j} x_{ij} \gamma_{ij})$$
(3)

$$s.t.\sum_{j\in B} x_{ij} = 1, \forall i \in U$$
(4a)

$$N_j = \sum_{i \in U} x_{ij}, \forall j \in B$$
(4b)

$$x_{ij} \in \{0,1\}, \forall i \in U, \forall j \in B.$$

$$(4c)$$

III. FACTOR GRAPH MODEL AND BP ALGORITHM

In this section, we introduce the BP algorithm to solve this optimization problem. In order to adopt this message passing method, we first develop a factor graph model. Then we transform the optimization problem into a marginal distribution estimation problem, which can be solved by BP.

A. Factor Graph Model for Association Problem

We develop a factor graph model G = (V, E) to represent the user association problem, as shown in Fig. 1. The vertices V consist of factor nodes (each BS's local utility function) and variable nodes (each user's association variable). If the received signal at the user i from the BS j is no less than the required signal strength, and BS j allows user i to access it, then an edge $(i, j) \in E$ exists between them. We use H(v) to denote the set of neighboring nodes of a node v.

1) Factor Nodes: Due to the one-user-one-BS constraint, the optimization of a network-wide system utility can be decomposed to the maximization of the local function in each BS. Mathematically, we have

$$\max_{\{x_{ij}\}} \sum_{i \in U} \log(R_i) = \max_{\{x_{ij}\}} \sum_{j \in B} \sum_{i \in H(j)} x_{ij} \log(R_{ij}) = \max_{\{x_{ij}\}} \sum_{j \in B} f_j,$$
(5)

where f_j is the local utility function, and

$$f_j = \sum_{i \in H(j)} x_{ij} \log(\frac{W}{\sum\limits_{i \in H(j)} x_{ij}} \gamma_{ij}).$$
(6)

Therefore, in our model, the factor node corresponding to the BS j represents the local utility function f_j .



Fig. 1. A Factor Graph Model of User Association.

2) Variable Nodes: The variable node, denoted by the set $\mathbf{x}_i \triangleq \{x_{ih}\}, h \in H(i)$, represents the association variable for the user *i*. This is because the optimization problem can be viewed as the optimization of \mathbf{x}_i from the perspective of the user *i*. We take USER3 in Fig. 1 as an example. USER3 can receive signals from BS2 and BS3. Due to the constraints (4a) and (4c), the association variable \mathbf{x}_3 for USER3 has two possible values, i.e., 1) $\mathbf{x}_3 = [1, 0]$, that is, USER3 connects to BS2 instead of BS3; 2) $\mathbf{x}_3 = [0, 1]$, that is, USER3 chooses BS3. We define the set of the variable node in the factor graph as $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_{N_U}\}$.

Based on the above analysis on the factor graph model, the network-wide optimization problem in (5) can be rewritten as:

$$\max_{\mathbf{x}} F(\mathbf{x}), \ F(\mathbf{x}) \triangleq \sum_{j \in B} f_j(\mathbf{x}_{H(j)}),$$
(7)

where $\mathbf{x}_{H(j)}$ represents set of all the user association variables of those potential users connected with BS *j*.

B. Transform of Optimal Problem

To transform the optimization problem into a marginal distribution estimation that can be solved by the BP, we follow the approach in [7] and define a probability distribution function based on the utility function in (7). That is

$$p(\mathbf{x}) = \frac{1}{Z} \exp(\mu F(\mathbf{x})), \qquad (8)$$

where μ is a positive number, and Z is a function of μ that is used to normalize this expression. According to [7], the result of large deviations shows that when $u \to \infty$, $p(\mathbf{x})$ concentrates around the maxima of $F(\mathbf{x})$, that is,

$$\lim_{u \to \infty} \mathbb{E}\left(\mathbf{x}\right) = \arg\max_{\mathbf{x}} F\left(\mathbf{x}\right) \tag{9}$$

where $\mathbb{E}(\mathbf{x})$ denotes the expectation of variable \mathbf{x} . In our optimal problem, $\mathbb{E}(\mathbf{x}) = \{\mathbb{E}(\mathbf{x}_1), \mathbb{E}(\mathbf{x}_2), \cdots, \mathbb{E}(\mathbf{x}_{N_U})\}.$

From Eq. (9), once we obtain $\mathbb{E}(\mathbf{x})$, we will have a good estimation for the maximization of $F(\mathbf{x})$. Furthermore, to obtain the elemnt $\mathbb{E}(\mathbf{x}_i)$ in $\mathbb{E}(\mathbf{x})$, we need to calculate the probability mass function (PMF) with respect to each \mathbf{x}_i , denoted by $p(\mathbf{x}_i)$, which can be solved by the BP algorithm.

C. Message Iterations in BP algorithm

Generally in the BP algorithm, the belief messages represent the estimation of marginal distribution of variables. Specifically in our case, factor node f_j only cares about one association parameter x_{ij} in the variable \mathbf{x}_i . That is, factor node f_j only updates the probability of the value $\mathbf{x}_i = [x_{ih} = 0, x_{ij} = 1], h \in H(i) \setminus j$, which is denoted by $\Pr(x_{ij})$. In other word, the message passing along the edge (i, j) represents the probability that the user *i* selects the BS *j*.

1) message from user to BS: In the iteration t, user i tells its possible serving BS j the probability of choosing it:

$$m_{i \to j}^{t} \left(x_{ij} = 1 \right) = \varphi_i \left(\mathbf{x}_i \right) \prod_{l \in H(i) \setminus j} m_{l \to i}^{t-1} \left(x_{il} = 0 \right).$$
(10)

This message is based on the messages sending from other possible serving BSs except BS *j*. Here, $\varphi_i(\mathbf{x}_i)$ is the normalization function which make sure $\sum_{l \in H(i)} P(x_{il} = 1) = 1$. Obviously, we have $m_{i \to j}^t(x_{ij} = 0) = 1 - m_{i \to j}^t(x_{ij} = 1)$.

Before the iteration begins, the user does not have any information from the network. Therefore the initial message can be set uniformly. For example, the user i has 3 serving BSs, the initial messages can be set as 1/3 for each BS.

2) message from BS to user:

$$m_{j \to i}^{t}(x_{ij}) = \mathbb{E}[p(\mathbf{x}_{H(j)}) \mid_{x_{ij}}]$$
$$= \sum_{k \in H(j) \setminus i} \left(\left[\exp\left(\mu f_{j}(x_{kj}, x_{ij})\right) \right] \mid_{x_{ij}} \prod_{k} m_{k \to j}^{t}(x_{kj}) \right),$$
(11)

where the second equation is from Eq. (8).

From Eq. (11), we can see that message is connected to the objective function of BS j when user i makes certain association choice on whether x_{ij} equals to 1 or). Based on the two values of x_{ij} , we have

$$m_{j \to i}^{t} (x_{ij} = 1)$$

$$= \sum_{k \in H(j) \setminus i} \left\{ \exp \left[\mu \left(\sum_{k} x_{kj} \log \left(R_{kj} \right) + \log \left(R_{ij} \right) \right) \right] \prod_{k} m_{k \to j}^{t} \left(x_{kj} \right) \right\}$$

$$= \sum_{k \in H(j) \setminus i} \left\{ \left[\prod_{k} \left(\frac{W \gamma_{kj}}{\sum_{l \in H(j) \setminus i} x_{lj} + 1} \right)^{x_{kj}} \frac{W \gamma_{ij}}{\sum_{l} x_{lj} + 1} \right]^{\mu} \prod_{k} m_{k \to j}^{t} \left(x_{kj} \right) \right\}$$
(12)

$$m_{j \to i} (x_{ij} = 0)$$

$$= \sum_{k \in H(j) \setminus i} \left\{ \left[\prod_{k} \left(\frac{W \gamma_{kj}}{\sum_{l \in H(j) \setminus i} x_{lj}} \right)^{x_{kj}} \right]^{\mu} \prod_{k} m_{k \to j} (x_{kj}) \right\}.$$
(13)

D. User Association Decisions

We assume there are T iterations in our BP alogrithm. After T iterations, the probability that the user i associates with BS j can be calculated as

$$P(x_{ij} = 1) = \varphi(\mathbf{x}_i) m_{j \to i}^T (x_{ij} = 1) \prod_{l \in H(i) \setminus j} m_{j \to i}^T (x_{il} = 0).$$

$$(14)$$

Based on Eq. (14), association decision can be made, i.e., user *i* selects the BS which has the maximum posterior probability among all the association options as its serving BS. After the association choices for all the users in the entire network are determined, the maximization of system utility function can be achieved.

IV. PRACTICAL IMPLEMENTATION DISCUSSION

A. Approximation

In the message iterations above, the BS should send two messages (Eq. (12) and (13)) to each of its potential users. The calculation of them needs $2^{|H(j)|-1}$ combination cases of its possible users, which leads to large amount of computation. Besides, the message overhead is heavy if the BS sends 2 messages to each of it potential users.

Here we propose an approximation process to reduce the computational complexity and make the messages transmit in a broadcast manner. The likelihood ratio $\frac{m_{j \to i}(x_{ij}=1)}{m_{j \to i}(x_{ij}=0)}$ can be used to replace the two messages form BS j to user i, because Eq. (10) only depends on this ratio after normalization. To reduce the complexity, we make approximations on the likelihood ratio as

$$\frac{m_{j \to i} (= 1)}{m_{j \to i} (= 0)} = \frac{\mathbb{E}\left(\exp\left(f_{j} \mid x_{ij}=1\right)\right)}{\mathbb{E}\left(\exp\left(f_{j} \mid x_{ij}=0\right)\right)} \approx \frac{\exp\left(\mathbb{E}\left(f_{j} \mid x_{ij}=0\right)\right)}{\exp\left(\mathbb{E}\left(f_{j} \mid x_{ij}=0\right)\right)}$$

$$= \frac{\exp\left(\mathbb{E}\left(\sum_{k \in H(j) \setminus i} x_{kj} \log \frac{W\gamma_{kj}}{\sum_{\substack{\sum \\ k \in H(j) \setminus i}} x_{lj} + 1\right)\right) \exp\left(\mathbb{E}\left(\log \frac{W\gamma_{ij}}{\sum_{\substack{l \in H(j) \setminus i}} x_{lj}}\right)\right)$$

$$\approx \exp\left[\log W\gamma_{ij} - \mathbb{E}\left(\log\left(\sum_{l \in H(j) \setminus i} x_{lj} + 1\right)\right)\right]$$

$$\approx W\gamma_{ij} + \mathbb{E}\left(\sum_{l \in H(j) \setminus i} x_{lj} + 1\right) = \mathbb{E}\left(\sum_{l \in H(j)} x_{lj}\right) - \mathbb{E}(x_{ij}) + 1 + W\gamma_{ij}$$
(15)

Here, we assume $\mu = 1$ here, since it is a constant and does not affect the approximations. And we use the approximation that $\mathbb{E}(\exp(x)) \approx \exp(\mathbb{E}(x))$ and $\log\left(\sum_{l \in H(j) \setminus i} x_{lj} + 1\right) \approx \log\left(\sum_{l \in H(j) \setminus i} x_{lj}\right)$. From Eq. (15), we can see that the likelihood ratio sent from

BS j to user i can be replaced by $\mathbb{E}\left(\sum_{l \in H(j)} x_{lj}\right)$, since the

parameters W, γ_{ij} and x_{ij} are all known to user *i*. Thus, after approximation, the belief message transmitted by BS *j* is written as $m_j^t = \sum_{l \in H(j)} m_{l \to j}^t (x_{lj} = 1)$, which can be transmitted in a broadcast manner. The likelihood can be easily calculated by use as Eq. (15).

B. Discussion for the Dynamic Case

The wireless cellular network is dynamic, i.e. channel gains vary, users arrive in, hand off and so on. In these cases, new user-BS association decisions are required by user equipments. We call the users who need to make or renew their association decisions as 'new users', oppositely, other users as 'existing users'.

Compared with the static BP algorithm that involves all users, it is practical to consider the dynamic case with only assigns new users to certain BS without changing the association of existing users. However, the information about existing users are very important, because they contribute to the current load and utility of this BS. For simplicity, in our dynamic algorithm, the number of existing users in BS j, denoted by K_j .

In this case, the real-time factor graph only consists of these new users and their optional BSs. And all steps is the same except the broadcasting value of E(Y), which should be updated as $E(Y) = K_j + \sum_{k} p_{kj}$.

V. NUMERICAL AND SIMULATION RESULTS

A. Parameters

We consider a HetNet which consists of a macro cell, and several pico cells, with a square area of 700m. The total users in this system is 100. We analyze the uniform layout, i.e., the pico-cell BSs and users are uniformly distributed in the HetNet. Detail physical layer parameters using in the simulation are listed in the following Table I [10].

Network Parameter		
Parameter	Macro Cell	Pico Cell
Power	43dBm	30dBm
Carrier Frequency	2GHz	
Antenna Gain	15dBi/macro cell	5dBi/pico cell
Minimum Distance	35m	10m
Noise Power	-104dBm	
Path loss	128.1+37.6lg(d/1000)	140.7+36.7lg(d/1000)
Shadowing s.d	8dB	10dB
Minimum SINR	-10dB	

TABLE I

B. Other User Association Rules for Comparison

We also simulate some existing association rules and compare them with the proposed BP association algorithm. A brief description of them is given as follows:

1) Max-SINR (Max-Received Signal Power): User chooses to associate with the BS j which provide the strongest down-link received signal power.

2) Range Extension (RE) in [11]: User associates with the BS that has minimum path loss between them. It can be viewed as a special case of CRE.

3) Near-Optimal results: The exhausting search is used to find out the optimal user-association solution. Due to the limit of MATLAB's computation capability, we adopt a preprocessing method. That is, the users who receive stronger signal from any SBS than the MBS are pre-judged to associated with the SBS and excluded from the exhausting search. From numerous tests, the result obtained by this lossy preprocessing is proved to extremely close to optimal result.

C. Performance Metrics for Simulation

The performance metrics studied in our paper are as follows. 1) Percentage of users offloaded to small-cells: this value reflects the load-balancing performance.

1) Geometric Mean Spectral Efficiency: it is proved that maximizing our object function is equivalent to maximizing the geometric mean rate of users, that is, $\sqrt[N]{\prod_{i=1}^{N} R_i}$, where N denote the total number of all users in this network.

2) Minimum Spectral Efficiency: the minimum spectral efficiency among all users, that is, $\min_i \{R_i\}$. It reflects the fairness among users from another perspective.

D. Simulation Results



Fig. 2. Comparison of small-cell User Association Statistics in Two Layouts.

Fig.2 shows the percentage of users in the system which associate with the pico-cell BSs based on three different association algorithms. We consider both 4 pico-cell and 8 pico-cell cases. As we know, RE encourages users near the pico-cell BSs to connect with them, because its criterion is related to the distance between users and BSs. From the figure, our BP algorithm achieves the best effect among the three algorithms. In the following, we will show the the advantages of the BP in other performance metrics in a 8 pico-cell case.

Fig. 3 depicts the CDF curves of the geometric mean spectral efficiency for different user association schemes. Both the RE and BP greatly increase the geometric mean spectral efficiency in the HetNet, i.e, the proportional fairness of the system. Also, the BP algorithm has much better performance than the RE. This is because the BP algorithm is the solution for the optimization problem, while RE only focuses



Fig. 3. Geometric Mean Spectral Efficiency With Three Association Rules.

on increasing the number of users associated with SBSs. Most importantly, the result of the BP algorithm extremely approaches the near optimal result achieved by exhausting search with pre-process.



Fig. 4. Minimum Spectral Efficiency With Three Association Rules.

Fig.4 compares the minimum spectral efficiency. RE does not perform better than Max-SINR, while our BP algorithm has almost three times gain compared with them and is close to the near optimal result. Based on the above simulations, we can see that the BP algorithm not only improves the average rate of users, but also guarantees the baseline of users' experience.

Fig. 5 shows the relative error $\frac{U_s - U_d}{U_s}$, where U_s and U_d denote the utility using static algorithm and dynamic algorithm respectively. This relative error reveals the gap between the dynamic and static algorithms. We can see the gap between them increases with the increasing number of new users. However, this gap is still very small even if the percentage of new users increases to 40 percent.

VI. CONCLUSION

In this paper, we developed a framework based on the BP algorithm to solve the user association problem in HetNets.



Fig. 5. Relative Error Caused by Dynamic Algorithm.

This distributed algorithm achieves a good result in loadbalancing, approaches the optimal result of maximizing the system's proportional fairness, and improves the minimum of user rate. Meanwhile, it is proved to be low-complexity and quick convergence, and it can easily satisfy the one-user-one-BS constraint. The effective dynamic algorithm makes our association strategy more practical.

REFERENCES

- NTT DOCOMO, R1-103264, "Performance of eICIC with Control Channel Coverage Limitation," 3GPP Std., Montreal, Canada, May 2010.
- [2] D. Fooladivanda and C. Rosenberg, "Joint Resource Allocation and User Association for Heterogeneous Wireless Cellular Networks," in *IEEE Transactions on Wireless Communications*, vol. 12, pp. 248-257, Jan. 2013.
- [3] Y. Bejerano, S. J. Han, and L. Li, "Fairness and load balancing in wireless LANs using association control," in *IEEE/ACM Transactions* on Networking, vol. 15, pp. 560573, June 2007.
- [4] T. Bu, L. Li, and R. Ramjee, "Generalized proportional fair scheduling in third generation wireless data networks," in *Proc., IEEE INFOCOM*, pp. 112, Apr. 2006.
- [5] Q. Ye, B. Rong, Y. Chen, M. AL-Shalash, C. Caramanis and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," in *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2706 - 2716, Jun. 2013.
- [6] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Distributed optimal user association and cell load balancing in wireless networks," in *IEEE/ACM Transactions on Networking*, vol. 20, pp. 177-190, Feb. 2012.
- [7] S. Rangan and R. Madan, "Belief Propagation Methods for Intercell Interference Coordination in Femtocell Networks," in *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, pp. 631–640. Apr. 2012.
- [8] Y. Chen, Z. Lin, B. Vucetic and J. Cai, "Inter-cell interference management for heterogenous networks based on belief propagation algorithms," in *Proc., IEEE WCNC*, pp. 1056-1061, Apr. 2013.
- [9] F. P. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, pp.33-37, 1997.
- [10] 3GPP TR 36.814, "Further advancements for E-UTRA physical layer aspects," v.9.0.0, Mar. 2010, www.3gpp.org.
- [11] Qualcomm, "LTE advanced: heterogeneous networks." in white paper, Jan. 2011.