Utilizing Loss Tolerance and Bandwidth Expansion for Energy Efficient User Association in HetNets

Muhammad Umar Bin Farooq*, Junaid Qadir†, M. Majid Butt‡, Muhammad Naeem§ and Ali Imran*
*School of Electrical and Computer Engineering, University of Oklahoma, Tulsa, OK, USA.
†Information Technology University, Punjab, Pakistan.
‡Nokia Bell Labs, France.
§COMSATS Universty Islamabad, Wah Campus, Pakistan.
Email: *(umar.farooq, ali.imran)@ou.edu, †junaid.qadir@itu.edu.pk, ‡majid.butt@nokia-bell-labs.com,
§muhammadnaeem@gmail.com

***Abstract*—5G is expected to serve diverse applications and users due to the popularity of Internet of Things (IoT), big data and industrial applications. Many of these IoT and industrial applications have inherent loss tolerance that can be used to enable energy efficient uplink communication. The uplink energy efficient system will increase the battery life of devices enabling new use cases in industrial IoT. In this paper, we map the effects of application loss tolerance to the rate requirements of the user. We then mathematically model an energy minimization problem for the uplink user association and resource allocation in heterogeneous networks. We aim to provide acceptable quality of service (QoS) with improved energy efficiency by exploiting the loss tolerance and bandwidth expansion simultaneously. A distributed uplink joint user association and resource allocation strategy for uplink energy per bit minimization is presented. We conduct extensive simulation based study for a heterogeneous network to evaluate the performance of our proposed schemes. Average energy per bit consumption in the proposed scheme is -74 dB compared to -53 dB in state-of-the-art channel individual offset (CIO) scheme.**

***Index Terms*—Industrial IoT, Energy Efficiency, Loss Tolerant Applications**

I. Introduction

The rise of smartphones, running diverse bandwidth and energy-hungry applications, has heralded a paradigm shift in modern cellular communications with a greater emphasis on energy efficiency and efficient bandwidth utilization. It is expected that future 5G networks will have to support a large number of users with extremely diverse requirements and a wide variety of applications such as big data, industrial IoT and multimedia communication to name a few. A *one-model-fits-all* approach can no longer meet the ever varying user demands with limited cellular resources. In the light of these requirements, recent research efforts [1], [2] have emphasized the importance of context awareness, loss tolerance, and energy efficiency, for effectively providing 5G services while meeting the expected QoS.

The loss tolerance of many modern applications, and devices motivates the case for loss tolerance and context awareness in 5G. Many industrial IoT devices send large amount of data periodically and have constraints on the uplink transmit power and battery life of the device. These devices might tolerate errors for lower power consumption [3]. Similarly modern big data and machine learning applications have inherently built-in loss tolerance [2] providing an extra cushion of loss tolerance. The loss tolerance of each device is dictated by the target industrial application—i.e., an application might tolerate a certain error rate for lower energy consumption while another user might not want to compromise the quality.

To meet the energy efficiency and bandwidth requirements of 5G, network densification has emerged as a prominent solution [4]. A number of small base stations (BS) in addition to traditional BS are deployed during network densification. This densification creates a heterogeneous network (HetNet) that brings BS closer to users with improved links. However, HetNet creates new challenges for the research community such as uplink-downlink asymmetry, load imbalance, backhaul bottleneck and mobility management [5]. An optimal association for uplink becomes sub-optimal for downlink due to the uplink-downlink asymmetry. Mohamed et al. [6] proposed uplink-downlink separation where a user can associate itself with different BS for uplink and downlink. It is shown that the separated architecture provides a better uplink-downlink rate coverage [7].

In this work, we exploit decoupled uplink-downlink architecture and present a distributed energy efficient uplink user association and resource allocation scheme. We consider a network scenario in which the network has enterprise small cells and there is multi-user demand diversity (time-dependent disparity in the user load). The network is designed to meet the user QoS requirements at peak load scenarios. Our solution uses application level loss tolerance and *bandwidth expansion* strategy, i.e., allocating more bandwidth where available (e.g., at off-peak times) in a bid to lower energy per bit consumption [8], [9]. We define *residual bandwidth* at a BS as the bandwidth that is not assigned to a user and is available for allocation to new users. Our association scheme considers residual bandwidth in addition to the path loss to the serving BS. The BS with more residual bandwidth can provide more spectrum to the user and hence the user can reduce the transmit power for the same QoS requirements. We also model the effects of application and user specific loss tolerance to the rate requirements of the users. It is shown that by exploiting both loss tolerance and bandwidth expansion, user association and resource allocation can be performed more efficiently.

Major contributions of this paper are:

- A model to map the effects of application and user level loss tolerance on effective rate requirements of the user is presented. This model helps to achieve higher uplink energy efficiency by incorporating loss tolerance in user association and resource allocation decisions.

- We present, to the best of our knowledge, a first of its kind work that uses both *bandwidth expansion* and *loss tolerance* in user association and resource allocation decisions. An energy per bit minimization problem is formulated and a distributed user association and resource allocation algorithm that exploits bandwidth expansion and loss tolerance is presented.
- Performance evaluation shows that our proposed scheme yields a significant improvement in uplink energy per bit consumption compared to the state-of-the-art maximum reference signal received power (RSRP) and channel individual offset (CIO) schemes.

II. Related Work

The loss tolerance of applications and users has already been used for energy efficiency in literature. Butt et al. [10] presented energy efficient scheduling for loss tolerant applications. The scheduling scheme allowed to drop a certain number of packets with constraints on average packet drop and successive packet loss. The energy-performance tradeoff without perfect channel state information at the transmitter and receiver was explored in [11].

Bandwidth expansion has been used for energy efficient resource allocation strategy in one-tier network that trades bandwidth for energy efficiency at off-peak times [8]. The authors in [9] quantified the limits of bandwidth expansion factor and associated gains in downlink energy efficiency. These efforts on bandwidth expansion consider traditional one-tier network. HetNets, with more than one option for association, provide another degree of freedom in bandwidth expansion. Bandwidth expansion was extended to HetNets model in [12].

Zhou et al. [13] formulated a mixed-integer non-linear programming optimization problem for energy efficient user association and presented a three layer iterative algorithm to solve the problem. Zhang et al. [14] studied the problem of user association and power allocation in millimeter based networks. They convert the mixed integer programming optimization problem into a convex problem by relaxing the user association indicator and solve it by Lagrangian dual decomposition. Wahedi et al. [15] presented a distributed user association scheme for mobile and IoT devices in HetNets. They proposed a multi-class user driven algorithm based on multi-armed bandit game with improvements in throughput, signalling and energy efficiency. A matching game based algorithm for user association was presented in [16]. The authors aim to maximize user throughput with fairness and minimize uplink transmit power consumption. In contrast to the previous work in user association and resource allocation, we use loss tolerance for the first time in joint uplink user association and resource allocation. We exploit both bandwidth expansion and loss tolerance for uplink energy per bit minimization.

III. System Model

A. Diversity and Loss Tolerance of Applications

We consider that the rate requirement of a user results from the rate requirements emerging from a set of applications at the user. Let us assume that n_u represents the number of applications running on a user u. A vector $\vec{A_u}$ of dimension $1{\times}\mathrm{n}_u$ at each user u of length n_u represents the rate requirement of each application. Each element A_{ui} of the vector $\vec{A_u}$ represent the rate required by each application. We also incorporate the inherent loss tolerance of different applications in our model. Loss tolerance is defined as an average data loss an application tolerates, i.e., the QoS of the application is acceptable at the receiver. The loss tolerance of applications can help to reduce its rate requirements.

We define another vector $\vec{E_u}$ of dimension $1{\times}\mathrm{n}_u$ to capture the loss tolerance at each user u for each application on the user. The average tolerable rate requirement R_u of a user u is calculated using the actual rate requirements A_u and E_u as follows,

$$R_u = \vec{A_u}(1 - \vec{E_u}^{\,T}) \tag{1}$$

Both $\vec{A_u}$ and $\vec{E_u}$ have real numbers and the transpose of $\vec{E_u}$ converts it from a row vector to a column vector making R_u a scalar number. Each element of $\vec{E_u}$, E_{ui} captures the loss tolerance of application i running at a user u. The value of E_{ui} can vary from one user to another for the same application to model different acceptable QoS for different users. For example, one user might want to save cost for a low quality video on a video based application but another user is willing to pay more but does not want to compromise the video quality for the same application. Although the same application is running on both users, the value of E_{ui} for the first user is higher than the latter.

We define a matrix of loss tolerance Φ_u at a user that has multiple levels of loss tolerance for each application i. Φ_u is a $n_u \times l_u$ matrix where l_u is the number of levels for loss tolerance for a specific user u. So, Φ_u has rows equal to the number of applications and columns representing each level of loss tolerance. Each entry of the matrix Φ_u, $\Phi_u(i,j)$ is the j^{th} level of tolerance for the application i. The first entry of each row, $\Phi_u(i,1)$ is the highest level of loss tolerance for application i. E_{ui}, the chosen level of loss tolerance for application i on user u, can take values from the i^{th} row of the matrix Φ_u. The value of tolerance E_{ui} for application i at user u should be less than or equal to $\Phi_u(i,1)$ of matrix Φ_u.

$$\Phi_u = \begin{bmatrix} \Phi_u(1,1) & \Phi_u(1,2) & ... & ... & \Phi_u(1,l_u) \\ \Phi_u(2,1) & \Phi_u(2,2) & ... & ... & \Phi_u(2,l_u) \\ . & . & & & . \\ . & . & & & . \\ . & . & & & . \\ \Phi_u(n_u,1) & \Phi_u(n_u,2) & ... & ... & \Phi_u(n_u,l_u) \end{bmatrix} \tag{2}$$

B. Loss Tolerance Illustrated Through An Example

An instance of loss tolerance matrix at user u, Φ_u is described below for 3 applications and maximum tolerance factor 3.

$$\Phi_u = \begin{bmatrix} 0.19 & 0.12 & 0.06 \\ 0.11 & 0.03 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3}$$

The user has three applications running on it and the maximum number of levels of loss tolerance for any application at the user is three. So, Φ_u comes out to be a 3×3 matrix with i^{th} row containing levels of loss tolerance for application i. The first applications at the user has three levels of loss tolerance: 19%, 12% and 6%. This shows that the user cannot tolerate more than 19% errors for the first application. The second application has only two tolerance levels of 11% and 3%. So, the user can tolerate a maximum of 11% errors for second application. The last entry of second row is zero because the user is willing to pay more for second application if errors are zero. The third application at the user is non-tolerant and has a loss tolerance of zero. Hence, the third row of Φ_u contains all zero elements.

C. Network Topology

We consider a two-tier HetNet topology in which a macro-cell network is overlaid with small cells. Each macro cell has three sectors with directed antennas while each small cell has omni-directional antennas. There is at least one small cell under the coverage area of a macro cell. Macro and small cells use the same frequency spectrum for transmission and the frequency reuse factor is 1. The total bandwidth at a BS c is ϵ_c.

The minimum bandwidth $\eta_{u,c}$ required by user u from a BS c to meet its rate requirements is given by the following equation where $\gamma'_{u,c}$ is the SINR with maximum power level $\max P_{u,c}$ of user,

$$\eta_{u,c} = \frac{R_u}{\log_2(1+\gamma'_{u,c})} \tag{4}$$

We use $\gamma_{u,c}$ to denote SINR of user u associated with BS c and is given by,

$$\gamma_{u,c} = \frac{P_{u,c} G_u \delta a (d_{u,c})^{-\beta}}{K + \sum\limits_{\forall j \in U_u} P_{t,j} G_j \delta a (d_{j,c})^{-\beta}} \tag{5}$$

where $P_{u,c}$ is the transmit power of user u communicating to BS c, G_u and G_j are the UE gains, δ is signal shadowing, a is path loss constant, $P_{t,j}$ is the transmit power of the interfering user j, $d_{u,c}$ and $d_{j,c}$ are the distances from user u and interfering user j to BS c respectively, β is the path loss exponent, and K is the thermal noise power. The set U_u contains all the uplink users in the neighboring cells which have been allocated the same bandwidth resources as user u and hence become the interferers for u. If a BS can provide bandwidth greater than $\eta_{u,c}$ to a user, the user can achieve the same rate R_u even with a decreased SINR. This newly available margin in SINR can be used to reduce the transmit power. We define the available bandwidth $B_{u,c}$ as resources that a BS c decides to allocate for a user u. The minimum value of $B_{u,c}$ is $\eta_{u,c}$ but a BS can decide to provide $B_{u,c}$ greater than $\eta_{u,c}$ to reduce the power consumption. The rate R_u for user u from BS c is given by,

$$R_u = B_{u,c} \log_2(1+\gamma_{u,c}) \tag{6}$$

The transmit power of the user u to communicate with BS c can be found by replacing the value of SINR in equation (6),

$$P_{u,c} = \left(2^{\frac{R_u}{B_{u,c}}} - 1\right) \frac{K + \sum\limits_{\forall j \in U_u} P_{t,j} G_j \delta a (d_{j,c})^{-\beta}}{G_u \delta a (d_{u,c})^{-\beta}} \tag{7}$$

It is assumed in the above equation that BS has complete interference information. The energy per bit in uplink $(E_b)_{u,c}$ for a user u communicating with BS c can be computed using (7) is given by,

$$(E_b)_{u,c} = \frac{P_{u,c}}{R_u} \tag{8}$$

A set U_c contains all the users connected to BS c and set C contains all the BS in the network. Now we want to choose the optimal bandwidth and BS that minimizes energy per bit for all the users. The optimization problem is formulated as,

$$\begin{array}{llll} \min\limits_{B_{u,c}, d_{u,c}} & \sum\limits_{C} \sum\limits_{U_c} (E_b)_{u,c} & & \\ \text{s.t.} \quad \sum_{u=1}^{U_c} B_{u,c} & \leq & \epsilon_c & \forall c \\ \gamma_{u,c} & \geq & 0 & \forall u, c \\ E_{ui} & \leq & \Phi_u(i,1) & \forall u, i \\ B_{u,c} & \geq & 0 & \forall c, u \end{array} \tag{9}$$

The goal of eq. (9) is to choose the best BS and resources for minimizing the energy per bit consumption. $B_{u,c}$ is the bandwidth assigned to the user u from BS c and $d_{u,c}$ is the distance between the user and the BS. The first constraint describes that for all C the sum of allocated bandwidth for all the users associated to a BS c cannot be greater than the total bandwidth at the BS ϵ_c. ϵ_c is already fixed for each BS c during the network deployment. The second constraint restricts the SINR for all U_c which cannot be less than 0 dB. This makes sure that SINR is good enough to decode the signal even with lower transmit power. The third constraint states that for all U_c the chosen value of the loss tolerance should be less than or equal to the maximum loss tolerance for an application i of the user u. This ensures minimum QoS requirement for all the applications at a user. The forth constraint describes that for all C and U_c the assigned bandwidth is always greater than or equal to zero.

The optimization variables of eq. (9) are the bandwidth allocated to the users and the distance between the user and the BS. Bandwidth assigned to a user can vary from minimum bandwidth required by the user to the total bandwidth at the BS and is a continuous variable. However, there are a finite number of candidate base stations a user can communicate to and hence the distance between the user and the BS can take a finite number of discrete values. So, the distance is a discrete variable. The first constraint is dependent on the continuous variable, bandwidth and the second constraint is dependent on SINR $\gamma_{u,c}$. $\gamma_{u,c}$ in turn depends on the discrete variable, distance as shown in eq. (5). The problem is a mixed integer non-linear programming problem. These kind of problems are computationally hard to solve and are considered computationally intractable. The search space of eq. (9) grows exponentially with the number of users in the network. An optimal solution can be obtained using an exhaustive search on all the users in the network which is computationally very expensive and is not suitable for practical systems. Furthermore, a new global solution needs to be computed whenever a user enters or leaves the network. Hence, we present a distributed solution to solve the problem.

IV. Resource Allocation and User Association Methodology

We present a two-fold user association and resource allocation scheme where both the BS and the user are involved in the association decision. The simple distributed model proposed in [12] is used to compute the $AssociationScore$ in the following way,

$$AssociationScore = \left(2^{\frac{R_u}{B_{u,c}}} - 1\right)^{\alpha} \left(\frac{1}{G_u \delta a(d_{u,c})^{-\beta}}\right)^{1-\alpha} \quad (10)$$

where α is the association exponent that can be used to vary the importance of residual bandwidth in the association decision. Association decision will give less importance to residual bandwidth with lower values of α. In addition, the user association will be similar to Max RSRP when α is set to zero. $AssociationScore$ considers the impact of the two optimization variables and it is based on uplink transmit power described in eq. (7). Hence, associating to the BS with minimum $AssociationScore$ minimizes the uplink power and in turn the energy per bit of the user. In this work, we jointly use loss tolerance and bandwidth expansion for energy efficient user association and resource allocation. We present a simple distributed algorithm for user association and resource allocation. Two entities take part in the user association and resource allocation decision—the BS and the user.

A. Algorithm at the Base Station

Base stations receive the rate requirements of a new user u. Each BS calculates the minimum bandwidth $\eta_{u,c}$ required to meet the rate requirements of the user according to eq. (4). The base stations for which minimum bandwidth is less than the total bandwidth are chosen as candidates for association. Each candidate BS then performs an exhaustive search to find optimal resources for all the already connected users and the new user. This exhaustive search exploits bandwidth expansion at off-peak times and finds an optimal multiple $x_{u,c}$ of the minimum bandwidth for all the connected users to the BS c. The exhaustive search increases the value of $x_{u,c}$ for each connected user until either all the resources of the BS are allocated or the second constraint of eq. 9 is met. This ensures that the BS attempt to allocate 100% of the resources in such a way which minimizes the average power consumption of all the connected users while meeting all the constraints of eq. (9). However, second constraints of eq. (9) limits the value of $x_{u,c}$ making sure that the signal can be decoded even after the power reduction due to the bandwidth expansion. The bandwidth $B_{u,c}$ that the BS c decides to allocate to the user u is given by,

$$B_{u,c} = x_{u,c}\eta_{u,c} \quad (11)$$

Each candidate BS transmits this $B_{u,c}$ to the user u along with a signal to indicate the presence of residual bandwidth at the BS even after allocating $B_{u,c}$ to the user. This indicates that the BS has enough bandwidth to improve services for some applications at the user. The algorithm at the BS is similar to the distributed algorithm of [12] with the indication of residual bandwidth as an extra step.

B. Algorithm at the user

Each new user calculates the minimum rate requirements to meet the needs of all the applications. The minimum rate requirements are calculated using the maximum value of loss tolerance for all the applications at the user. The first element of each row $\Phi_u(i,1)$ from matrix Φ_u is the maximum tolerance of each application i as described in eq. (2). One element from i^{th} row of Φ_u, indicating all levels of tolerance for application i , can be selected as the value of E_{ui} . E_{ui} is given the first value of the row i from the matrix Φ_u i.e. $\Phi_u(i,1)$ to compute minimum rate requirement for user u. This minimum rate requirement is then transmitted to all the base stations. Each BS then decides the optimal bandwidth for allocation to the user as described in subsection IV-A. The user receives the optimal allocated bandwidth $B_{u,c}$ from each candidate BS c. The user then calculates the $AssociationScore$ for each BS as described in eq. (10). The user associates to the BS with minimum $AssociationScore$. The BS also sends a signal indicating the availability of residual bandwidth at the BS. The applications at the user cannot be given lower levels of loss tolerance if there is no residual bandwidth at the BS. However, some applications at the user can operate at lower tolerance if there is residual bandwidth at the BS. In case of residual bandwidth at the BS, the user decreases the tolerance level of an application and computes the new rate requirements. The application is selected based on a pre-defined importance ranking at each user. The user then receives a new bandwidth from the BS based on the new rate requirements. This step is repeated until there is no residual bandwidth at the BS or all the application at the user are operating at the lowest level of loss tolerance.

Algorithm 1 User Assocation and Resource Allocation Algorithm at the user u

(a) Find the rate requirements R_u of the user u as described in equation 1;
(b) Publish the rate requirements R_u to all the BS;
(c) Use $B_{u,c}$ from each BS to compute the $AssociationScore$;
(d) Associate user u to the BS c with minimum $AssociationScore$;
while *residual bandwidth at the BS and there are lower levels of loss tolerance for one or more application* **do**
  (e) Decrease the value of E_{ui} for the most important application i at the user;
  (f) Repeat Step (a) i.e. compute the new rate requirements;
  (g) Get new bandwidth $B_{u,c}$ from the BS for the new rate requirements of step f;
end

C. An Illustrative Toy Example

An illustrative example of the algorithm at the BS and at the user is given in Fig. 1. An instance is shown where a new user wants to join the network and both the user and the BS interact for optimal user association and resource allocation. The user can associate with any one of the three candidate base stations shown. The user calculates rate requirements R_u of 18.7 kbps. The first entry from each row of loss tolerance matrix Φ_u is used to calculate R_u. All BS calculate the minimum bandwidth required $\eta_{u,c}$ which comes out to be 2 kHz, 5.7 kHz and 6.3 kHz for BS 1, 2 and 3 respectively. Each BS then exploits *bandwidth expansion* and calculate the optimal bandwidth for

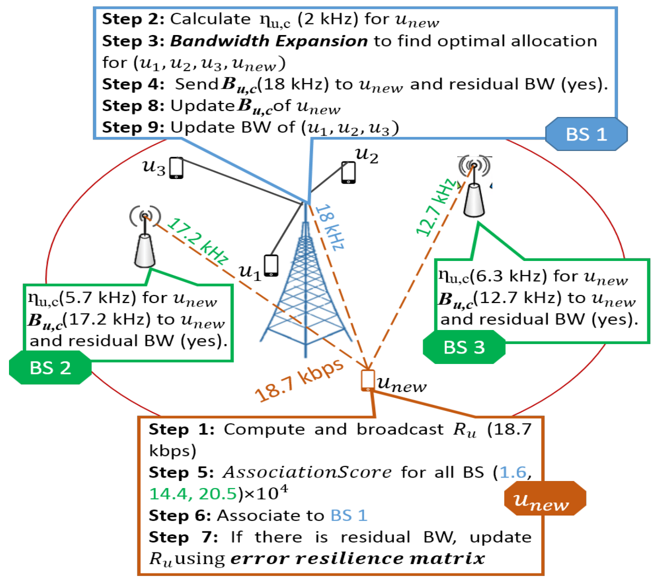

Fig. 1. An illustrative toy example in which we enlist the steps taken during the interaction of the algorithms at the user and at the BS.

allocation to all the user. The bandwidth $B_{u,c}$ that each BS decides to allocate to the user for BS 1, 2 and 3 is 18 kHz, 17.2 kHz and 12.7 kHz respectively. The user calculates the $AssociationScore$ for each BS and associates to BS 1 with minimum score of 1.6×10^4. If there is residual bandwidth at the BS, the user uses the *loss tolerance matrix* to provide better rates to as many applications as possible. All the steps of the algorithm at the BS are shown for BS 1 while brief highlights of main steps are shown for BS 2 and BS 3. The user association strategy exploits bandwidth expansion and loss tolerance to minimize the energy per bit consumption of the user while providing the best possible rates to as many applications as possible.

V. Simulation Setup and Results

We present the simulation setup and simulation results in this section. We perform Monte Carlo simulations to compare the proposed user association and resource allocation algorithm with the state of the art Max RSRP and CIO based schemes [17]. User is associated to BS with highest RSRP in Max RSRP scheme while a constant offset is added in RSRP of small cells for CIO based scheme to shift the users towards small cells.

A. Simulation Setup

We employ a 3GPP-compliant LTE network topology [18] with macro cells overlaid with small cells. We deploy and simulate a two-tier HetNet with 7 macro BS. Small BS are distributed in each sector of macrocell with uniform density. A fraction of both indoor and outdoor UEs are concentrated near small base stations to model hotspot scenarios. The ratio of indoor to outdoor users is 4:1. The number of applications at a user follows a uniform distribution with range $[1, \max n_u]$ for the simulation analysis. $\max n_u$ is the maximum number of applications that can run on any user. The rate requirement of an application i at a user A_{ui} also follows a uniform distribution with range $[1, \max A]$. $\max A$ is the maximum rate requirement of an application at any user. The simulation parameters are summarized in Table I.

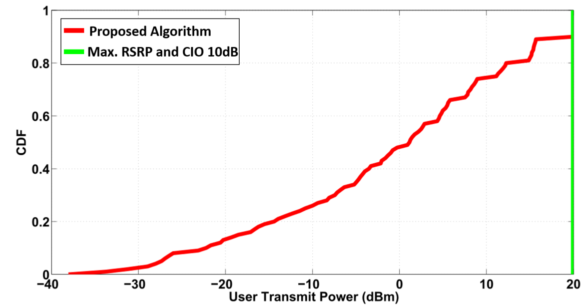

Fig. 2. Power comparison of our scheme compared to Max RSRP and 10 dB CIO based schemes. *The transmit power of all users is 20 dBm in CIO based and Max RSRP schemes while our algorithm uses bandwidth expansion to reduce the uplink transmit power of the users.*

TABLE I
Description of Simulation Parameters

Parameter Description	Value
Number of Macro BS	7
Number of Sectors per Macro BS	3
Number of small BS per sector	1
Number of Users per sector	25
System Bandwidth	10 MHz
Maximum User Transmit Power ($\max P_{u,c}$)	20 dBm
Transmission Frequency	2 GHz
Inter-site Distance of Macro BS	500m
Macro BS and small BS Height	25m and 10m
Network Topology	Hexagonal
Association Exponent	0.5
User and BS Noise Figure	7 dB and 5 dB
$\max n_u$ and $\max A$	10 and 10kbps

B. Simulation Results

In this subsection, we provide an analysis of the proposed user association and resource allocation algorithm from the described simulation setup. A comparison of uplink transmit power is presented in Fig. 2. Both Max RSRP and CIO based schemes transmit at the maximum power threshold of the user (20 dBm). Our scheme exploits off-peak times and uses bandwidth expansion to reduce the transmit power of the users. At off-peak times, there is excess bandwidth at the BS and hence can be assigned to the communicating users. We use this residual bandwidth to allocate more resources to the users. Hence, the user can transmit at a lower power and still achieves the same rate requirements due to bandwidth expansion. The average uplink power consumption in our algorithm is -1 dBm compared to 20 dBm in both Max RSRP and CIO based schemes. This reduction in transmit power at off-peak times can decrease the overall power consumption of users and will expedite the inclusion of battery limited devices in 5G architecture.

Both algorithms at the BS and at the user ensure lower transmit power with the highest possible rates to as much application as possible. However, the algorithm also guarantees rates above the minimum rates for all applications incorporating the highest level of loss tolerance. A comparison of energy per bit consumption defined in eq. (8) is presented

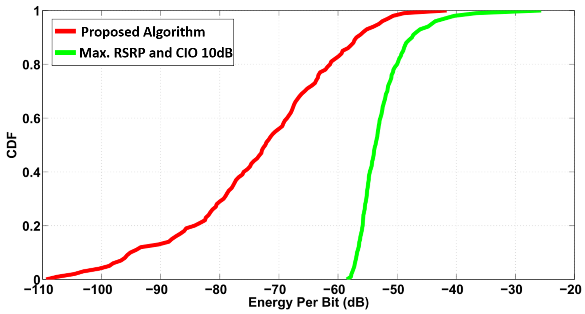

Fig. 3. Energy per bit consumption of our algorithm compared to the Max RSRP and CIO based association. *Our solution exploits both bandwidth expansion and loss tolerance and achieves better energy per bit compared to Max RSRP and CIO based scheme.*

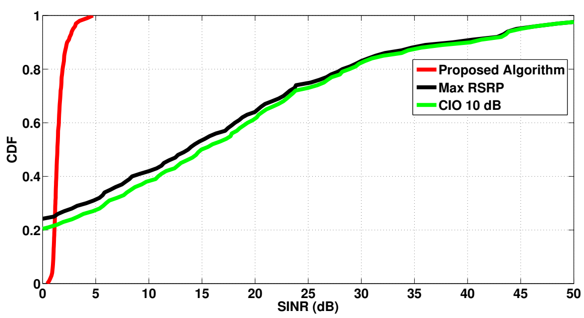

Fig. 4. SINR comparison of our scheme compared to Max RSRP and 10 dB CIO based schemes. Lower uplink transmit power decreases SINR in the proposed algorithm and bandwidth expansion ensures same rate to users despite lower SINR.

in Fig. 3. The proposed algorithm outperforms Max RSRP and CIO based association. The average uplink energy per bit is -74 dB in our algorithm compared to -53 dB in Max RSRP and CIO based scheme. Hence, the user can transmit the same number of bits in our algorithm with less energy consumption.

The proposed scheme in this paper aims to lower the uplink energy per bit consumption and uplink transmit power. SINR in our scheme is expected to decrease due to lower transmit power. Fig. 4 compares the SINR of the proposed algorithm with Max RSRP and CIO based association schemes. As expected the SINR in our scheme is lower than both Max RSRP and CIO based association. However, the rates to the user remains the same due to bandwidth expansion where user can get bandwidth more than the minimum requirements.

VI. Conclusion

In this paper, we exploit bandwidth expansion and loss tolerance of application to design a user association and resource allocation schemes. We present two algorithms, one at the BS and the other at the user, that work hand in hand to reduce the uplink energy per bit consumption of the users. In our scheme, each user starts with the lowest but acceptable rate for all applications running on it. The user then deploys an iterative algorithm to provide better rates to as much applications as possible. Our numerical study shows that the proposed scheme achieves an average reduction of 21 dB in energy per bit consumption as compared to Max RSRP and CIO based association scheme. This reduction in energy per bit of user can help to include more battery constrained in 5G networks opening new research avenues in industrial IoT.

Acknowledgments

This work is partially supported by the National Science Foundation under Grant Numbers 1923669 and 1718956. The statements made herein are solely the responsibility of the authors. For more details about these projects please visit: http://www.ai4networks.com

References

[1] O. Onireti, J. Qadir, M. A. Imran, and A. Sathiaseelan, “Will 5G see its blind side? Evolving 5G for universal internet access,” in *Proceedings of the 2016 workshop on Global Access to the Internet for All*, pp. 1–6, ACM, 2016.

[2] J. Qadir, A. Sathiaseelan, U. B. Farooq, M. Usama, M. A. Imran, and M. Shafique, “Approximate networking for universal internet access,” *Future Internet*, vol. 9, no. 4, p. 94, 2017.

[3] F. Samie, L. Bauer, and J. Henkel, “An approximate compressor for wearable biomedical healthcare monitoring systems,” in *2015 International Conference on Hardware/Software Codesign and System Synthesis*, pp. 133–142, IEEE, 2015.

[4] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer, “Network densification: The dominant theme for wireless evolution into 5G,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, 2014.

[5] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, “User association in 5G networks: A survey and an outlook,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1018–1044, 2016.

[6] A. Mohamed, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, “Control-data separation architecture for cellular radio access networks: A survey and outlook,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 446–465, 2016.

[7] S. Singh, X. Zhang, and J. G. Andrews, “Joint rate and SINR coverage analysis for decoupled uplink-downlink biased cell associations in HetNets,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5360–5373, 2015.

[8] S. Videv and H. Haas, “Energy-efficient scheduling and bandwidth-energy efficiency trade-off with low load,” in *IEEE International Conference on Communications (ICC)*, pp. 1–5, June 2011.

[9] M. M. Butt, B. Schubert, M. Kurras, K. Börner, T. Haustein, and L. Thiele, “On the energy-bandwidth trade-off in green wireless networks: System level results,” in *1st IEEE International Conference on Communications in China Workshops (ICCC)*, pp. 91–95, Aug 2012.

[10] M. M. Butt and E. A. Jorswieck, “Maximizing system energy efficiency by exploiting multiuser diversity and loss tolerance of the applications,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4392–4401, 2013.

[11] M. M. Butt, E. A. Jorswieck, and A. Mohamed, “Energy and bursty packet loss tradeoff over fading channels: A system-level model,” *IEEE Systems Journal*, 2016.

[12] U. B. Farooq, U. S. Hashmi, J. Qadir, A. Imran, and A. N. Mian, “User transmit power minimization through uplink resource allocation and user association in HetNets,” in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2018.

[13] T. Zhou, Y. Huang, and L. Yang, “Energy-efficient user association in downlink heterogeneous cellular networks,” *IET Communications*, vol. 10, no. 13, pp. 1553–1561, 2016.

[14] H. Zhang, S. Huang, C. Jiang, K. Long, V. C. Leung, and H. V. Poor, “Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1936–1947, 2017.

[15] Y. M. Waheidi, M. Jubran, and M. Hussein, “User driven multiclass cell association in 5G HetNets for mobile & IoT devices,” *IEEE Access*, vol. 7, pp. 82991–83000, 2019.

[16] M. Anany, M. M. Elmesalawy, and A. M. A. El-Haleem, “Matching game-based cell association in multi-rat HetNet considering device requirements,” *IEEE Internet of Things Journal*, 2019.

[17] 3rd Generation Partnership Project, “Radio resource control (rrc); protocol specification,” in *3GPP TS 36.331 version 8.6.0 Release 8*, 2009.

[18] 3rd Generation Partnership Project, “Physical layer aspects for evolved universal terrestrial radio access (e-utra),” in *3GPP TR 25.814 version 7.1.0 Release 7*, 2006.