

Received August 25, 2020, accepted September 2, 2020, date of publication September 7, 2020, date of current version September 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3022350

Energy Efficient Resource Allocation and Scheduling for Delay-Constrained Multi-Cell Broadcast Networks

LIANG LU¹, LINYU HUANG¹, (Member, IEEE), AND WEI HUA¹

College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China

Corresponding author: Linyu Huang (lyhuang@scu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61801318.

ABSTRACT Resource allocation in wireless cellular networks has been an important and challenging issue. With the rapid development of wireless communications, more and more mobile multimedia and innovative applications (e.g. augmented reality and video conferences) have emerged, which are suitable for wireless broadcast and have requirements on transmission delay. In this paper, we focus on delay-constrained multi-cell broadcast networks, taking user mobility into consideration for future channel estimation. The promising network coding technique is also adopted to improve transmission efficiency. The objective is to minimize total transmission energy consumption via joint resource allocation and scheduling optimization. Both optimal and heuristic solutions are proposed. Simulation results show that the proposed schemes perform better than other benchmark schemes, in terms of transmission energy consumption and success rate.

INDEX TERMS Wireless broadcast, energy efficiency, context-awareness, resource allocation, network coding.

I. INTRODUCTION

With the rapid development of information and communication technology (ICT) industry, mobile data traffic has been growing tremendously in recent years [1]. Such huge traffic data brings more energy consumption, which is unfriendly for network operators and natural environment. Therefore, energy efficient green communication has attracted significant attention. In addition, 60% to 80% of wireless cellular network energy is consumed by base stations (BSs) [2]. Hence, it is crucial to improve resource utilization and energy efficiency for BSs.

Resource allocation, as a possible solution, has become an important and challenging issue. Since the problem of maximizing network utility was proposed in [3], resource allocation optimization has been a hot topic and a lot of schemes have been suggested [4], [5]. To further improve system performance via resource allocation optimization, one possible way is to design resource allocation schemes based on the emerging characteristics of cellular networks. Nowadays, more and more smart wireless terminals are used. These smart terminals are usually equipped with various

sensors and can obtain a lot of context information, such as users' current location, moving speed and direction, etc. These context information can be used to predict users' future location, channel conditions, etc. [6], [7], thereby improving the accuracy and efficiency of resource allocation.

Another characteristic of modern mobile communication is that more and more multimedia and innovative applications (such as video conference and augmented reality) have appeared. These applications generally have the feature that multiple users request the same file and have requirements on delivery time. For instance, multiple users may access some popular audio or video files on mobile phones at the same time, with quality of service (QoS) requirements on download time. Research [8] shows that most of the downstream traffic comes from downloading the same files. Research [9] shows that on YouTube, the world's leading video site, 80% of the visits come from the top 10% popular video resources. For such applications, wireless broadcast via a common channel to those users can greatly reduce the burden of BSs and improve bandwidth utilization [10]. Besides, network coding (NC) [11] has been shown to be a promising technique to improve network throughput and bandwidth utilization, especially for broadcast transmissions. By taking linear or nonlinear processing of packets, NC can effectively reduce

The associate editor coordinating the review of this manuscript and approving it for publication was Francisco Rafael Marques Lima¹.

the number of retransmissions, so as to decrease total energy consumption [12]. For example, in Fig. 1, a BS broadcasts three packets (i.e. a, b, and c) to three users (i.e. $user_1$, $user_2$, and $user_3$). However, due to unsatisfied channel conditions, after some transmissions, $user_1$ lost packet a, $user_2$ and $user_3$ lost packets b and c, respectively. Traditionally, BS needs to retransmit each lost packet sequentially. Whereas, in NC based broadcast, the BS can encode a new packet $d = a \oplus b \oplus c$ and retransmits it once to all users via broadcast. Upon receiving d, $user_1$ can recover the lost a by $a = d \oplus b \oplus c$. $user_2$ and $user_3$ can recover their own lost packets in a similar way. By doing so, the minimum number of retransmissions can be reduced from three to one.

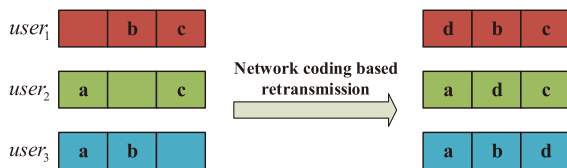


FIGURE 1. Network coding based retransmission.

Considering the above issues, we take the users' moving context into consideration, to estimate the future position and channel state information for users. Network coding technology also applies to improve broadcast efficiency. Then we make the joint resource allocation and scheduling optimization for multi-cell networks, aiming to minimize the transmission energy consumption of all BSs.

The rest of this paper is organized as follows: We review related work and describe the main contributions of this paper in Section II. System model is presented in Section III. Section IV presents the problem formulation and the optimal solution. In Section V, we propose a heuristic algorithm with polynomial complexity to solve the optimization problem. In Section VI, we carry out simulations to evaluate the performance of the proposed schemes. Realistic considerations are discussed in Section VII. Finally, conclusion is drawn in Section VIII.

II. RELATED WORKS AND PAPER CONTRIBUTIONS

Many efforts have been made for resource allocation on energy efficient broadcast. Radio resources such as bandwidth, time slots, transmit antennas, transmit power, etc., can be assigned based on the channel state [13].

In the case of energy efficient broadcast, a well-known problem of looking for a broadcast tree to minimize the total energy consumption, called minimum energy broadcast (MEB), has already been intensively studied. The MEB problem has been proved to be NP-hard and a flurry of heuristic algorithms have been proposed [14]–[19]. Local search is usually employed to integrate with other algorithms for solving the MEB problem, such as nested partitioning and linear programming [16], swarm optimization (PSO) [17] and genetic algorithm (GA) [18]. GA based approach may drop into local optimal solutions or may find the optimal solution

by slow convergence speed. To overcome these problems, an enhanced scheme by using memetic algorithm (MA) was proposed [19]. However, these works did not consider user mobility.

For wireless mobile networks, one intricacy lies in the fact that the channel conditions might be changing in time due to environmental factors and user mobility [20]. During movements, the channel degradation may occur due to path loss, fading, or noise. For mobility management, numerous studies have been made in recent years. By analyzing the entropy of each individual's trajectory, a 93% potential predictability in user mobility across the whole user base was found in [21]. Mobility prediction approaches such as Markov chain, Bayesian network, artificial neural network based approaches, and so on, have been suggested [22]. Obviously, it is possible to predict mobile users' future trajectories based on their historical movements. If BSs know a user's future context information (e.g. trajectory) integrated with channel state information (CSI) fed back from the user, they can schedule a corresponding broadcast plan under the delay constraint in advance for energy saving. Vehicle mobility prediction based rebroadcast nodes selecting algorithm was adopted in [23], [24]. The work in [23] proposed a reliable broadcast routing scheme based on mobility prediction (RB-MP), to select rebroadcast nodes in vehicular ad hoc network (VANET). Specifically, RB-MP divides the neighbors into several sets according to the movement direction, firstly; then it utilizes the position and velocity to predict the maintenance time of all neighbors; finally, several rebroadcast nodes are selected. Similarly, the work of [24] selected rebroadcast vehicle by considering link available time which is acquired by the relative speed, radio transmission range, and the inter-vehicle distance between vehicles moving in the same direction and on the same road. Whereas, the work of [25], [26] considered channel conditions for broadcast/multicast. Combining duty-cycle with context-awareness, the work of [25] studied efficient context-aware multi-hop broadcasting protocol, which allows nodes to adapt their behavior according to the changing network dynamics and context. The authors [26] considered two state channel conditions, either good or bad, and investigated an opportunistic multicast scheduling algorithm. The algorithm estimates optimum data rate and optimum number of users having appropriate channel conditions to select user group for multicast. Because when the channel condition is good, less power is required for the transmission of data, therefore energy can be saved.

Nevertheless, the aforementioned works are broadcast/multicast routing optimization approaches and did not consider reducing redundant retransmissions to improve energy efficiency. To process packets and eliminate too much number of retransmission, network coding (NC) has been widely implemented. The work of [27] showed the energy efficiency of network coding over routing, and proposed a minimum energy multicast approach based on network coding. In multicast, two-stage cooperative multicast is part of

the most promising schemes to increase system throughput. In the first stage, users, which have successfully received data from BS, can serve as relays to forward the data to other users in the second stage. However, total energy consumption increased. To improve energy efficiency, the authors of [28] applied dynamic network coding at the selected relays. Their simulation showed better energy efficiency. For a lossy network, random linear network coding (RLNC) can achieve a reliable and efficient transmission [29], [30], and enhance other technologies, such as device-to-device (D2D) communications [31], [32]. Additionally, network coding can integrate other ways to further improve energy efficiency. By exploiting network coding and node sleeping scheme, energy efficiency of multicast is improved in [33].

Despite these literature, to the best of our knowledge, no prior work considered the resource allocation optimization for transmission energy minimization, based on both users' future context information and network coding. In our previous work [34], single cell networks were considered. But in real scenarios, users tend to move among multiple cells.

In this paper, we study the joint resource allocation and scheduling problem for delay-constrained multi-cell networks, by utilizing users' moving context information for upcoming channel state information estimation. Network coding technique is also adopted to improve retransmission efficiency. The objective is to minimize the total transmission energy consumed by BSs. The main contributions of this paper are summarized as follows:

- A network coding based multi-cell broadcast system model is constructed in the paper;
- We formulate the broadcast energy minimization problem as a context-aware network coded broadcast scheduling optimization problem. To get the optimal solution, we transform the optimization problem into an MIP problem;
- We propose a polynomial-time heuristic scheme to solve the optimization problem;
- Simulation results show that both the MIP based solution and heuristic scheme can significantly reduce transmission energy consumption and improve broadcast success rate, compared with two conventional benchmark schemes.

III. SYSTEM MODEL

A. NETWORK OVERVIEW

We consider a time slotted network with M BSs and K users, which are labeled as BS_j , $j = 1, 2, \dots, M$ and U_i , $i = 1, 2, \dots, K$, respectively. Each BS is situated in the center of a cell and connected to the base station control center (BSC) via a core network (wired backbone network). BSC acts as routing and BSs management. Frequency reuse is also adopted among cells. Neighboring BSs use different frequencies to avoid co-channel interference. The K users are moving among the cells in a certain mobility pattern. At a particular time slot, we assume that multiple users are requesting a same source file from the network. In this paper,

we concentrate on the process that M BSs broadcast the requested file to K users, subject to a delay constraint that the delivery process must be completed in T time slots. The network architecture is illustrated in Fig. 2.

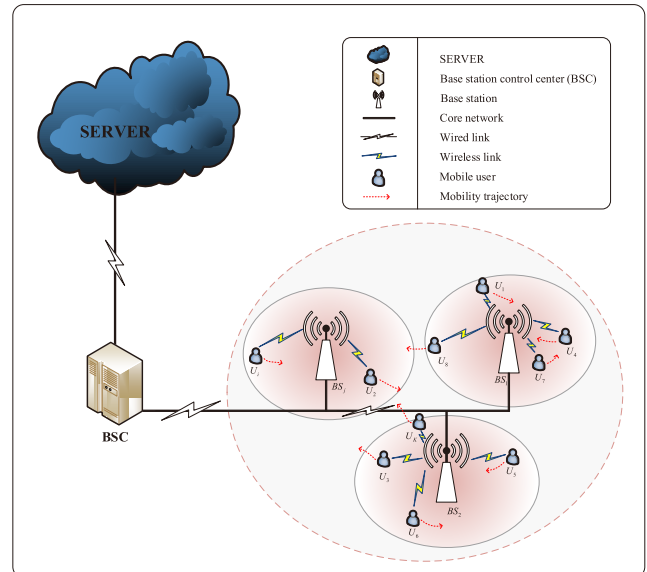


FIGURE 2. Network Architecture.

The requested file is transmitted from the server to the BSC, then arrives at all BSs, and finally from BSs to users via wireless channels. In each BS, the file is divided into N equal length packets and only one packet can be transmitted per time slot. The transmitted contents and schedule of all BSs are controlled by the BSC. When a user is moving among the cells, it will be served by different BSs and receives the required packets from the assigned BS. However, at a particular time slot, a user can only be served by one BS at most. It is assumed that each BS can obtain the served users' context information, including the current position, moving direction, and speed. Based on a user's historical context information, BSs are assumed to be able to predict the user's channel gain in the next T time slots [7], and feed the information back to the BSC. We will discuss the value of T and its impact in a practical system (see Section VI-C) later. BSC can schedule the broadcast of BSs by exploiting the users' future channel gain.

B. NETWORK CODING

In order to improve broadcast reliability and reduce the number of retransmissions, NC technology is introduced to encode the transmitted packets. In this paper, we use the property of random linear network coding (RLNC) [35] or deterministic linear network coding (DLNC) [36] for broadcast, and do not specify a specific network coding algorithm. It is assumed that appropriate NC parameters are adopted, so that the RLNC or DLNC based broadcast can ensure that all transmitted encoded packets are useful to all users. Therefore, to decode a source file that is split to N equal-sized

packets, a user only needs to receive N encoded packets. Note that this assumption is achievable by both RLNC or DLNC with some constraints. The detail of network coding is out of the scope of this paper, the interested readers might refer to [35], [36].

In this paper, each BS adopts NC to encode the transmitted packets based on the receiving status of the served users, to ensure that the transmitted packets are useful to all the served users. In other words, a user U_i can decode the source file successfully as long as it has received N encoded packets, which may be sent from different assigned BSs. In this case, U_i is said to be *complete*. Otherwise, it is *incomplete*.

C. LINK MODEL AND ENERGY CONSUMPTION

In a transmission, it is assumed that U_i is able to successfully receive a transmitted packet from the assigned BS if the received signal-to-interference-plus-noise ratio (SINR) at the user side is no less than a certain fixed threshold γ . Otherwise, U_i cannot receive that packet. At a particular time slot t , received SINR of U_i relates to BS_j is denoted as Γ_{ijt} , and can be written as

$$\Gamma_{ijt} = \frac{G_{ijt}P_{jt}}{\eta_0}, \quad (1)$$

where G_{ijt} is the channel gain, P_{jt} is the transmit power of BS_j at the t -th time slot. The maximum transmit power of all BSs is set as P_{max} . η_0 is the sum of noise and interference from neighbor cells to U_i , and modeled as a constant for simplicity.

An $M \times T$ matrix $\mathbf{P} = \{P_{jt}\}$ is defined to denote the transmit power of BSs at all T time slots. In this paper, we are aiming to minimize the total transmission energy consumed by all BSs of the whole broadcast process, by determining optimal \mathbf{P} . For a transmission with P_{jt} , transmission energy consumption by BS_j is P_{jt} times the duration of a time slot. For the convenience of calculation, it is assumed that each time slot is a unit time, and will be ignored throughout the paper. Therefore, the total transmission energy consumption of all BSs is denoted as E , and can be written as

$$E = \sum_{j=1}^M \sum_{t=1}^T P_{jt} \quad (2)$$

IV. CONTEXT-AWARE NETWORK CODED BROADCAST SCHEDULING

As stated in Section III, each BS can broadcast one packet at most each time slot, and the broadcast must be finished within delay constraint T . T must be no less than N , since otherwise the broadcast cannot be completed. When $T > N$, BSs do not need to transmit at every time slot in general. Hence, the BSC needs to specify the transmission time slots for each BS. Similarly, users do not have to successfully receive a packet in each time slot, and BSC also needs to determine the users to which a BS needs to transmit. If a user U_i is assigned to receive a packet from a BS at a slot, we call that U_i is a *target user*. When $T = N$, each user must be a target user at every time slot. At a particular slot, once the target users of

BS_j are determined, based on the packet reception status and CSI of its target users, the transmission contents and required transmit power of BS_j at that time slot can be calculated as well.

We define a $K \times M \times T$ matrix \mathbf{X} to indicate whether a user is a target user. If $X_{ijt} = 1$, U_i is a target user of BS_j at time slot t . Otherwise, U_i is not a target user. That is

$$X_{ijt} = \begin{cases} 1, & \text{if } U_i \text{ is a target user of } BS_j \text{ at time slot } t \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Now, we will introduce how to determine the optimal transmit power of BSs to cover the target users. In general, it shows a positive correlation between broadcast coverage and transmit power under the constraint of maximum power P_{max} . We say that the minimum power of BS_j used to cover a particular user U_i at the t -th slot is the *required power*, and denote it as Λ_{ijt} . Λ_{ijt} can be calculated by (1) when the SINR Γ_{ijt} equals to threshold γ , and can be written as

$$\Lambda_{ijt} = \frac{\gamma \eta_0}{G_{ijt}} \quad (4)$$

A $K \times M \times T$ matrix $\mathbf{\Lambda} = \{\Lambda_{ijt}\}$ is defined to denote all the required power. For a particular time slot, the minimum transmit power of a BS is generally determined by the target user with the worst channel gain. As long as the 'worst' target user can be covered, the others can be covered as well. Therefore, a BS can determine the minimum transmit power at a time slot by first calculating the required power of all its target users, and then chooses the maximum one as the transmit power. The broadcast with such selected transmit power can ensure that the expected received SINR of target users meets the successful reception SINR requirement with minimum transmit power.

Since BSs already know the channel gains of each user at the next T time slots, they can obtain the required power $\mathbf{\Lambda}$. The primary job of the paper is to determine \mathbf{X} based on $\mathbf{\Lambda}$, and further determine the optimal \mathbf{P} to complete broadcast with minimum energy consumption. It is worth noting that when a user U_i is specified as a target user and the assigned BS uses the corresponding required power to transmit a packet, the received SINR is no less than the threshold γ and that packet is treated as successfully received by U_i . Therefore, in such a scenario, the number of times that a user is specified as a target user represents the number of packets it will receive. The major notions used in this paper are given in Table 1. The optimization problem of context-aware network coded broadcast scheduling can be formulated as:

$$\begin{aligned} & \text{Minimize} && \sum_{j=1}^M \sum_{t=1}^T P_{jt} \\ & \text{Subject to: C1: } && P_{jt} = \max_{i=1}^K \{X_{ijt} \Lambda_{ijt}\}, \quad \text{for } j \in [1, M], \\ & && t \in [1, T]; \\ & && \text{C2: } \sum_{j=1}^M \sum_{t=1}^T X_{ijt} \geq N, \quad \text{for } i \in [1, K]; \end{aligned}$$

TABLE 1. Major notations.

| Notation | Definition |
|-----------------|--|
| K | Number of users |
| T | Available time slots for transmission |
| M | Number of base stations (BSs) |
| U_i | User index |
| t | Time slot index |
| BS_j | Base station index |
| N | Total number of packets of the source file |
| G_{ijt} | Channel gain from BS_j to U_i at t -th slot |
| P_{jt} | Transmit power of BS_j at t -th slot |
| Γ_{ijt} | The SINR of U_i relates to BS_j at t -th slot |
| η_0 | The sum of noise and interference from neighbor cells to U_i |
| γ | Threshold of SINR |
| E | Total transmission energy consumption of all BSs |
| X_{ijt} | Indicator of whether a user is a target user |
| Λ_{ijt} | The required power of BS_j to U_i at t -th slot |

$$\begin{aligned}
 \text{C3: } & \sum_{j=1}^M X_{ijt} \leq 1, \quad \text{for } i \in [1, K], t \in [1, T]; \\
 \text{C4: } & X_{ijt} = 0, \text{ if } \Lambda_{ijt} \geq P_{max}, \quad \text{for } i \in [1, K], \\
 & j \in [1, M], t \in [1, T]. \quad (5)
 \end{aligned}$$

where C1 limits the transmit power of a BS to the maximum required power of all its target users in the cell. C2 ensures that each user can be covered by N times at least, such that it can receive enough packets for decoding the source file. C3 guarantees that at each time slot, each user can only be specified as the target user of one BS at most. C4 limits the transmit power of each BS to the maximum power P_{max} . If the required power of a user is higher than P_{max} , the user cannot be specified as a target user.

Context-aware network coded broadcast scheduling in single cell networks has been proved to be an NP-hard problem in our previous work [34]. Single cell networks can be regarded as a special case of multi-cell networks when the number of BSs M is equal to 1. Accordingly, the problem in [34] is reducible to the problem (5) stated above. Thus the problem (5) is also NP-hard.

To obtain the optimal solution of problem (5), we transform C1 into $P_{jt} \geq X_{ijt} \Lambda_{ijt}$, for $i \in [1, K], j \in [1, M], t \in [1, T]$. Consequently, the optimization problem becomes an mixed integer programming (MIP) problem which can be solved by some standard mathematical tools, such as branch and bound and cutting planes [37]. We call this method context-aware with mixed integer programming (CAMIP). Due to the high computational complexity of the optimal solution, in the next section, we present a polynomial-time heuristic algorithm to solve the minimum energy problem.

V. HEURISTIC ALGORITHM

Now, let us check the optimization problem again. For network coding based broadcast, as long as a user has been covered by N times, it can decode the source file and is said to be *complete*. However, to a user who has received all the required packets, extra transmitted packets are useless and

may cause extra energy waste. Therefore, the ideal broadcast is that each user is exactly covered by N times. Based on this idea, in this section, we propose a polynomial-time heuristic algorithm for the optimization problem.

We first define a $K \times M \times T$ dynamic cost matrix $C = \{C_{ijt}\}$, where C_{ijt} denotes the energy cost of choosing U_i as a target user of BS_j at j -th time slot. Since the duration of a slot is assumed to be a time unit, the energy cost C_{ijt} is set to be the extra increased transmit power in order to cover U_i , based on previously selected target users.

Note that C is dynamically updated once a new target user is selected. Consider the scenario that BS_j may need to serve two user U_i and U_{i+1} at t -th slot, where the required power $\Lambda_{ijt} < \Lambda_{(i+1)jt} \leq P_{max}$. If BS_j only wants to cover U_{i+1} , the energy cost $C_{(i+1)jt} = \Lambda_{(i+1)jt}$. However, if U_i is already selected as a target user, the minimum transmit power of BS_j is $P_{jt} = \Lambda_{ijt}$. On that basis, if BS_j wants to serve U_{i+1} further, the minimum transmit power of BS_j is $P_{jt} = \Lambda_{(i+1)jt}$. Therefore, the extra transmit energy cost used to cover U_{i+1} is $C_{(i+1)jt} = (\Lambda_{(i+1)jt} - \Lambda_{ijt})$.

Recall that the number of times that a user is specified as a target user represents the number of packets it will receive. The basic idea of the proposed heuristic algorithm is to greedily specify target users by finding the minimum value from the valid elements in cost matrix C , until every user has been selected as a target user for N times. Initially, set X to be zero and C_{ijt} be Λ_{ijt} , since if U_i is chosen as the target user of BS_j at t -th slot, BS_j must increase the transmit power to Λ_{ijt} at t -th slot and will cause extra energy cost $C_{ijt} = \Lambda_{ijt}$. Then, find the minimal element $C_{i'j't'}$ from C and set $X_{i'j't'} = 1$, subject to the following constraints: 1) the transmit power of each BS must be no greater than P_{max} , 2) at each time slot, each user can only be specified as the target user of one BS at most, and 3) every user cannot be covered more than N times. After that, update C by setting $C_{i'j't'}$ to be $(\Lambda_{i'j't'} - \Lambda_{i'j't'})$, where $i \neq i'$ and $X_{i'j't'} \neq 1$. Since $U_{i'}$ has been specified as the target user of $BS_{j'}$ at the t' -th slot, it cannot be selected as a target user of current BS and any other BSs at the t' -th slot again. Thus we increase the costs to be greater than P_{max} to avoid being selected again. Similarly, for all complete users, they should not be covered anymore to avoid unnecessary energy waste. Hence, the related costs of these users need to be greater than P_{max} . Then, specify the target user corresponding to the minimum valid element in C . Repeat the process of updating C and specifying target users greedily, until every user has been selected as a target user for N times. When the process is complete, X indicates the target users' assignments. Transmit power P_{jt} can be determined by $\max_{i=1}^K \{X_{ijt} \Lambda_{ijt}\}$ and total energy consumption of broadcasting can be computed by (2). This algorithm is called Context-Aware Heuristic Scheme (CAHS) and described in Algorithm 1.

Now, we are going to analyze the computational complexity of CAHS. It is assumed that the complexity of finding the minimum or maximum value from n numbers is typical $O(n)$. In STEP 1, it takes KTM operations to initialize X and takes

Algorithm 1 Context-Aware Heuristic Scheme (CAHS)**Require:** Required power matrix Λ .**Ensure:** Transmit power \mathbf{P} and total energy consumption E .**STEP 1:** Set $\mathbf{X}=\mathbf{0}$, $\mathbf{C}=\Lambda$.**STEP 2:** Count the number for U_i where $\Lambda_{ijt} > P_{max}$, for $i \in [1, K]$, $j \in [1, M]$, $t \in [1, T]$. If counts of any user is less than N , the broadcast cannot be completed and the algorithm halts. Otherwise continue.**STEP 3:** Let $C_{i'j't'} = \min(C_{ijt})$, for $i \in [1, K]$, $j \in [1, M]$, $t \in [1, T]$.**STEP 4:** Update matrix \mathbf{X} by setting $X_{i'j't'} = 1$.**STEP 5:** Set $C_{i'j't'} = \Lambda_{i'j't'} - \Lambda_{i'j't'}$, for $i \in [1, K]$, $i \neq i'$, $\Lambda_{i'j't'} \leq P_{max}$ and $X_{i'j't'} \neq 1$.**STEP 6:** Set $C_{i'j't'} = 2P_{max}$, for $j \in [1, M]$. Since $U_{i'}$ has been specified as the target user of BS_j at t' -th slot, it cannot be selected as a target user of any other BSs at t' -th slot again. Thus we increase the costs to be greater than P_{max} to avoid being selected again.**STEP 7:** Set $C_{ijt} = 2P_{max}$, for all complete users U_i , $t \in [1, T]$, $j \in [1, M]$. If a user is complete, it should not be selected as a target user anymore, so we increase the costs C_{ijt} related to it to be greater than P_{max} to ensure that each user is exactly covered by N times.**STEP 8:** If any user is incomplete, proceed to STEP 3, otherwise continue.**STEP 9:** Compute transmit power of BSs $P_{jt} = \max_{i=1}^K \{X_{ijt} \Lambda_{ijt}\}$, for $j \in [1, M]$, $t \in [1, T]$; energy consumption $E = \sum_{j=1}^M \sum_{t=1}^T P_{jt}$, and return.

KTM operations to initialize \mathbf{C} . The overall complexity of STEP 1 is $O(KTM)$. In STEP 2, it first takes KTM operations to finish counting for each user, then takes K operations to compare the number of counts with N . The overall complexity of STEP 2 is $O(KTM)$. In STEP 3, the complexity is $O(KTM)$ to find the minimum cost from \mathbf{C} . In STEP 4, the complexity is $O(1)$ to specify a target user. After a target user is determined, STEP 5 will take maximum $(K - 1)$ operations to update the cost of other users in the same BS at that time slot. So the complexity of STEP 5 is $O(K)$. In STEP 6, it takes M operations to guarantee the target user cannot be selected as a target user at the time slot again. The maximum complexity is $O(M)$. In STEP 7, it takes one operation to check whether a user is complete firstly. If so, it takes at most TM operations to update the value of C_{ijt} associated with that user. This procedure will repeat at most K times for all users. Therefore, the maximum complexity is $O(KTM)$. STEP8 will take at most K operations to check if any user is incomplete, hence the complexity is $O(K)$. Therefore, for the loop from STEP 3 to STEP 8, the total complexity of each loop is $O(KTM)$. It may repeat at most NK loops to cover K users for N times. Hence the overall complexity from STEP 3 to STEP 8 is $O(K^2TMN)$. Finally, In STEP 9, it first takes K operations to calculate the required power and takes K operations to find the maximum value

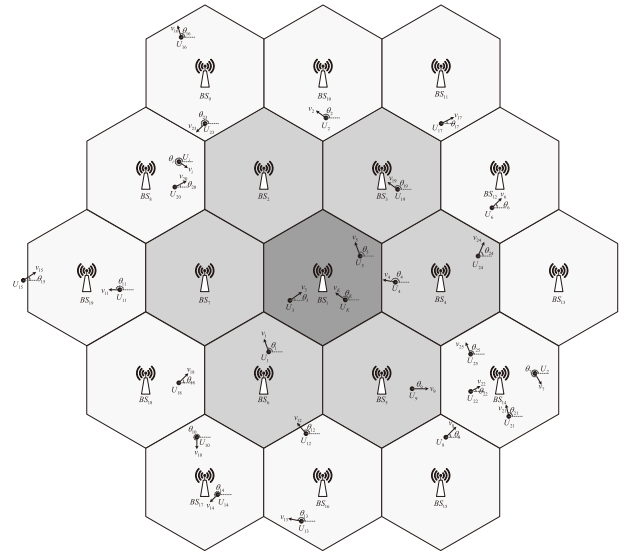
as the transmit power. These operations will be repeated TM times at most to determine \mathbf{P} . Besides, it takes TM operations to compute the energy summation E . The total complexity of STEP 9 is $O(KTM)$. Consequently, the overall complexity of CAHS is $O(K^2TMN)$.

VI. SIMULATION AND PERFORMANCE EVALUATION

In this section, we evaluate the performance of proposed CAMIP and CAHS by simulation. The setting and parameters are introduced firstly. In subsequent part, we introduce two benchmark schemes for comparison and present performance metrics. Lastly, simulation results are discussed.

A. SIMULATION SETUP

A three-tier multi-cell network with 19 BSs is used in the simulation, as shown in Fig. 3. Suppose that K users are randomly distributed in the cells at the beginning. U_i , $i = 1, 2, \dots, K$, moves with an initial angle θ_i and a constant speed v_i between 80-100 km/h which is the typical speed of vehicles on highway. Note that the assumption that users move at a constant speed and angle is to simplify the calculation. In real scenarios, this assumption is not needed. Meanwhile, to ensure that all users can possibly decode all the source packets, the number of slots that a user can be covered by at least one cell cannot be less than N in the simulation.

**FIGURE 3.** Network used in simulation.

Next, we will introduce how to determine the channel gain G_{ijt} . Suppose that BSs already know the users' context information, including current location and future moving trajectories. At a particular time slot t , the path loss between U_i and BS_j can be determined by

$$PL = 128.1 + 37.6 \log_{10}(d_{ijt}) + L_s \quad (6)$$

where d_{ijt} denotes the distance (in km) between U_i and BS_j , L_s indicates shadow fading. Let A_{ijt} be the attenuation (in dB)

caused by shadow fading, and α be the path loss exponent. Then the channel gain G_{ijt} can be rewritten as

$$G_{ijt} = \frac{10^{-A_{ijt}/10}}{d_{ijt}^\alpha} \quad (7)$$

Note that A_{ijt} is usually modeled as a zero-mean Gaussian random variable with standard deviation σ . In general, the empirical value for σ is between 6 dB and 12 dB, and α is between 4 and 6. In this paper, we assume that network operators can collect shadow fading data, which can be used by BSs. The correlation between shadow fading at different times can be expressed by Gauss Markov process as follows

$$A_{ij(t+1)} = \rho_{ijt} A_{ijt} + \sqrt{1 - \rho_{ijt}^2} W_t \quad (8)$$

where W_t is zero-mean Gaussian random variable with 10dB standard deviation, and ρ_{ijt} is the correction coefficient which can be determined by

$$\rho_{ijt} = e^{-\frac{|\Delta d|}{d_{cor}}} \quad (9)$$

where $|\Delta d|$ is the position change of the mobile user from t - th slot to $(t + 1)$ - th slot, and d_{cor} is the correlation distance and set to 50 meters in our simulation according to 3GPP [38]. For more details of this shadow fading model, we refer the readers to [39]. In practice, in a particular time slot t , for U_i , the BS_j can know the G_{ijt} from CSI. By using (7), A_{ijt} can be determined. Due to the fading correlation, BSs can use (8) to estimate $A_{ijt'}$ for the follow T slots, where $t' > t$.

The duration of a time slot is assumed to be 0.167 second, which is same as [40], for wireless video streaming. Additional parameters of the simulation are shown in Table 2. The simulation environment is established by MATLAB, and Gurobi 8.1 [41] is utilized to solve the CAMIP. In the simulation, for each set of parameters, we present the average performance over 3000 realizations.

TABLE 2. Simulation parameters.

| Parameter | Value |
|----------------------------------|-------------|
| Cell radius | 500 m |
| Bandwidth | 10 MHz |
| Number of resource block (RB) | 50 |
| Maximum transmit power P_{max} | 46 dBm |
| SINR threshold γ | 3 dB |
| Minimum distance | 35 m |
| Noise density | -174 dBm/Hz |
| Noise figure | 9 dB |
| Correlation distance | 50 m |

B. BASELINE APPROACHES AND PERFORMANCE METRICS

This paper mainly focuses on minimizing energy consumption of broadcast in multi-cell networks through power allocation. To the best of our knowledge, there is no reference

approach that can be directly applied to our system model for performance evaluation. Therefore, based on related works, we introduce a non-context-aware non-network coding based baseline broadcast approach. Furthermore, a non-context-aware network coding based baseline approach is also introduced to reveal the influence of network coding on broadcast process. Because of non-context-awareness, the two conventional schemes keep broadcasting in chronological order until all users are complete or all time slots are consumed. At the same time, in order to achieve minimum energy consumption, each user will be specified as the target user of BS with minimum required power.

In the case of the first non-network coding based algorithm, BSs transmit source packets and retransmit the lost source packets, which is called general scheme (GS). GS has two phases: initial phase and retransmission phase. In the initial phase, all BSs synchronously transmit the N source packets sequentially, which will consume N time slots.

After sequentially broadcast the N packets once, if any user is still incomplete, the algorithm steps into retransmission phase. BSs first check the lost packets and denote the collection as R . Then, the BSs will retransmit the packets in R to the users. To retransmit a packet $r \in R$, BSs first find all the incomplete users which require r , denoted as $users^*$. Then just like the initial phase, $users^*$ will be specified as the target users of the BSs with minimum required power, provided the required power is no more than P_{max} . However, in the retransmission phase, users may need different packets at the same slot, so we can specify different BSs to retransmit different packets at a slot. When all users are complete or all time slots are used, the algorithm halts.

The other benchmark scheme adopts the same transmission pattern as GS, with the only difference that network coding technique is applied, which is referred to as general scheme with network coding (GSNC). GSNC broadcasts network coded packets, and retransmits encoded packets which are beneficial to all users. Therefore, in the retransmission phase, GSNC does not consider that users request different packets at the same time. The pseudo code of GS and GSNC is shown in Algorithm 2 and Algorithm 3 respectively in Appendix I.

Here, metrics used to evaluate the performance of these schemes are described as follows:

- Energy consumption (normalized): We use energy consumption (normalized) to indicate the transmission energy consumed by all BSs for broadcasting the file to all users.
- Success rate: In a realization of broadcast, we call that the broadcast succeeds only if all users are complete after T slots. The success rate is defined as the number of successful realizations over the total number of realizations, which can indicate the reliability of transmission.
- Average broadcast latency (slots): Broadcast latency of a realization is defined as the total number of time slots consumed when all users are complete. The average broadcast latency is defined as the averaged value of broadcast latency over all realizations (in time slots).

C. NUMERICAL RESULTS

We evaluate the performance of proposed CAMIP and CAHS by comparing to two conventional schemes GS and GSNC using performance metrics with various K , T , and M .

Firstly, let $T = 64$, $N = 16$, K varies from 5 to 100, results are shown in Fig. 4. From Fig. 4a, we can find that the energy consumption of CAMIP and CAHS is always far less than GS and GSNC. In particular, when $K = 100$, CAMIP and CAHS can save energy 75% and 71% respectively, compared

to GS. As can be observed from Fig. 4b, three network coding based schemes, CAMIP, CAHS, and GSNC always maintain a success rate of 100%, whereas the success rate of non-network coding based scheme GS gradually decreases with K increasing. When $K = 100$, the success rate of GS is 68% only. In Fig. 4c, the average broadcast latency of context-aware schemes CAMIP and CAHS are far higher than GS and GSNC. When $K = 100$, the average broadcast latency of CAMIP and CAHS is 130% and 121% higher than GS, respectively.

Next, let $K = 50$, $N = 16$, T varies from 32 to 128, results are shown in Fig. 5. Results illustrated in Fig. 5a indicates that the energy consumption of CAMIP and CAHS decreases when T increasing. On the contrary, it increases in GS and GSNC. That means that when the delay constraint is relaxed, the performance of context-aware schemes gets better. Whereas, GS and GSNC broadcast in chronological order, more relaxed constraint brings more random simulated network, which leads to more energy consumption. Note that the energy consumption of GS and GSNC should show a nearly constant when T increasing. In Fig. 5b, with the increase of T , the success rate of GS increases gradually. When $T = 32$, the success rate of GS is 54%. While $T = 128$, the success rate is 99%. This happens because more relaxed time constraint means more time slots for retransmission. Broadcast failure caused by insufficient time slots will reduce, hence the success rate of GS increases. In Fig. 5c, the average broadcast latency of CAMIP and CAHS becomes higher when T increasing, whereas GS and GSNC have little changes. The reason is that CAMIP and CAHS may choose later slots that have better channel gain to reduce energy consumption.

Lastly, let $K = 50$, $T = 64$, N varies from 8 to 32, simulation results are depicted in Fig. 6. The energy consumption of the four schemes increases when N increasing in Fig. 6a. Meanwhile, CAMIP and CAHS always outperform GS and GSNC. In Fig. 6b, with the increase of N , the success rate of GS decreases gradually. When $N = 8$, the success rate is 99%, When $N = 32$, the success rate is 36%. The reason is same as T increasing. Fig. 6c depicts the increasing average broadcast latency of the four schemes. As N increases, all schemes need more time slots to finish broadcasting.

D. SUMMARY

From Fig. 4 to Fig. 6, compared to the benchmark schemes, we can conclude that the context-aware network coding based schemes, CAMIP and CAHS, can significantly save energy in most cases, and improve success rate as well. However, the payment for less energy consumption is higher broadcast latency, because context-aware schemes would take full advantage of all time slots for energy saving, therefore takes a longer time to complete broadcast. The analysis of the delay constraints shows that when the delay constraint is relaxed, the performance of context-aware schemes can be improved. Besides, from the comparison between GS and GSNC, it is revealed that network coding can significantly improve the

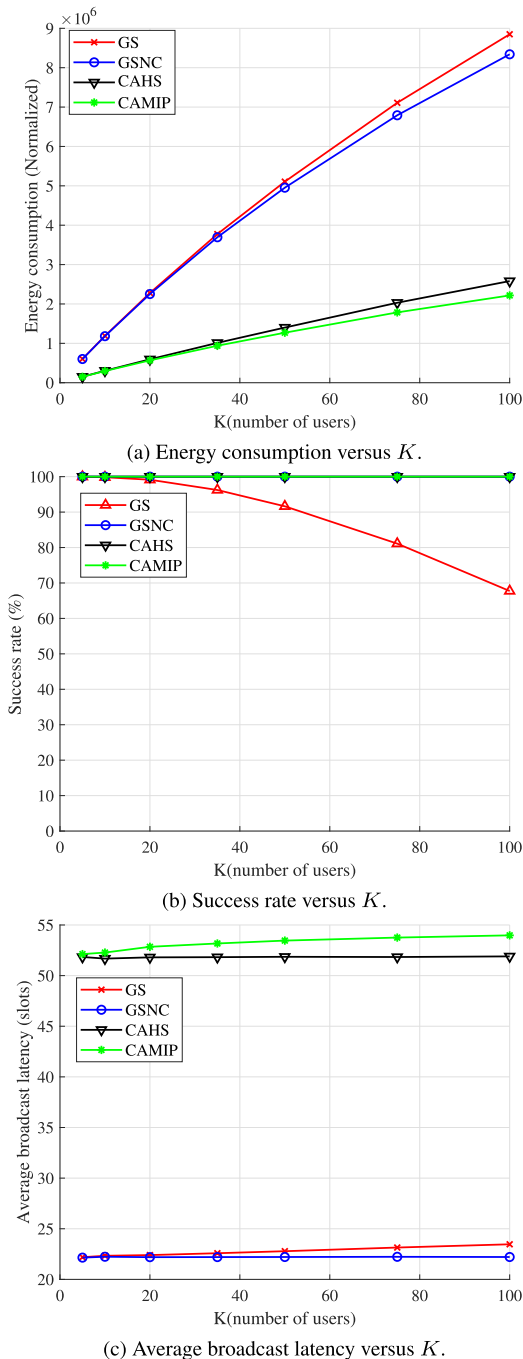


FIGURE 4. Results versus K .

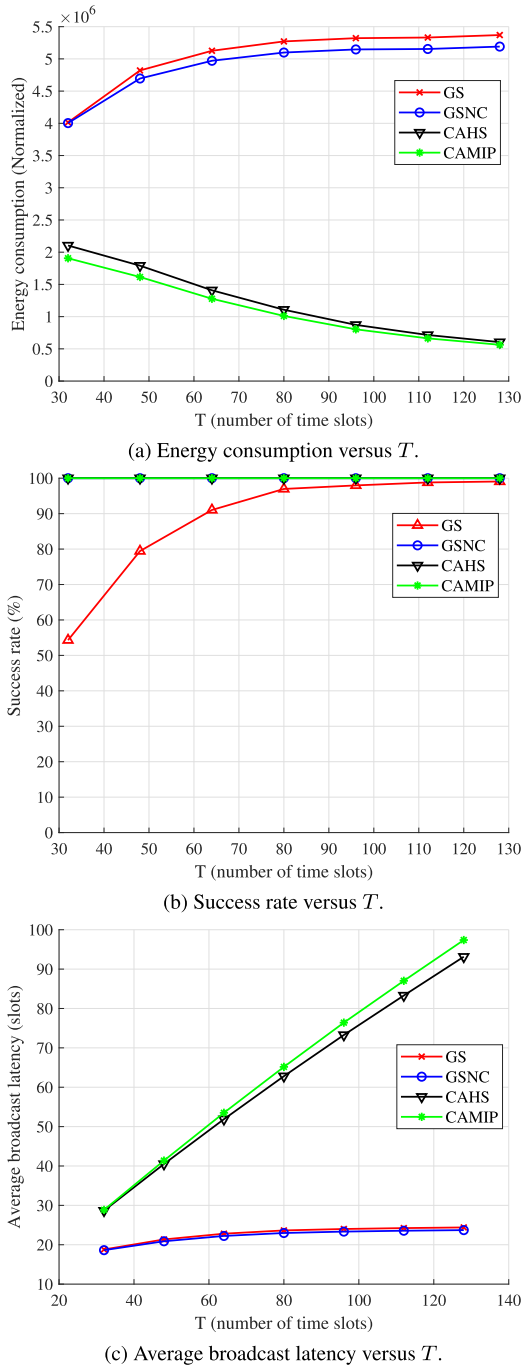


FIGURE 5. Results versus T .

success rate of broadcast and save transmission energy for BSs.

VII. REALISTIC CONSIDERATION

In this section, we discuss some problems which may arise in practice and present corresponding solutions. Since all the above parts are finished under the assumption that the channel gain of mobile users can be perfectly predicted, the proposed CAMIP and CAHS exhibit remarkable performance compared with conventional schemes. In practice, because

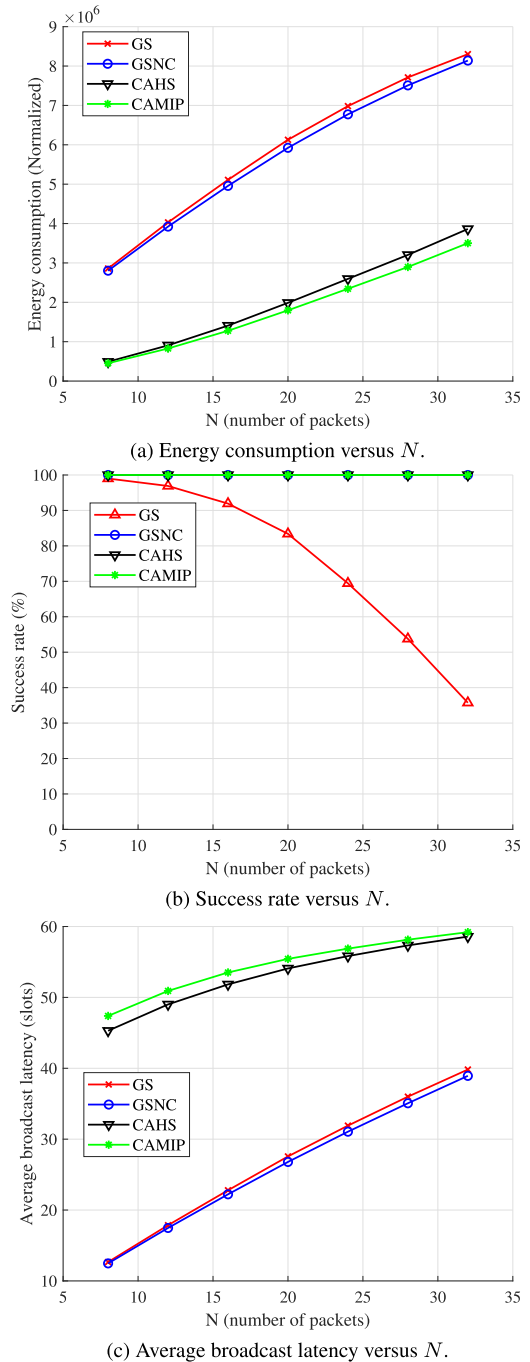


FIGURE 6. Results versus N .

of high variability of users' channel condition or inaccurate prediction, BSs cannot perfectly obtain users' future channel information in general, which will be expected to result in some performance degradation. Meanwhile, to simplify the problem and reduce computation complexity, we only consider noise and omit co-channel interference. However, there does exist many systems suffering co-channel interference from neighbor cells.

For the proposed schemes, both imperfect channel prediction or co-channel interference may cause performance

degradation, mainly in terms of packet loss during transmissions. To evaluate how much prediction errors or co-channel interference will affect the proposed schemes, we investigate packet loss ratio under various situations. We define a new metric named valid transmission rate to denote the ratio of the total number of received packets to the total number of transmitted packets for all users.

A. IMPERFECT PREDICTION

In the proposed schemes, the transmit power of a BS is determined based on the predicted channel gain of its target users. To evaluate the impact of prediction errors, we assume the real channel gain of a transmission is the sum of corresponding predicted channel gain plus a zero-mean Gaussian random variable with different standard deviations. Let $K = 50$, $T = 64$, $N = 16$, and the standard deviation varies from 0 to 2 dB, we evaluate the packet loss ratio via simulation. The results are presented in Fig. 7. From the figure, we can find that the valid transmission rate drops down to below 75% when the standard deviation is 0.25 dB. Besides, the valid transmission rate decreases as the standard deviation increases. In other words, the influence of imperfect prediction is more serious when the standard deviation is large. In particular, when the standard deviation is 0.25 dB, the valid transmission rates of CAMIP and CAHS is 73% and 64%, respectively. When the standard deviation is 2 dB, the valid transmission rates of CAMIP and CAHS reduce to 69% and 63%, respectively.

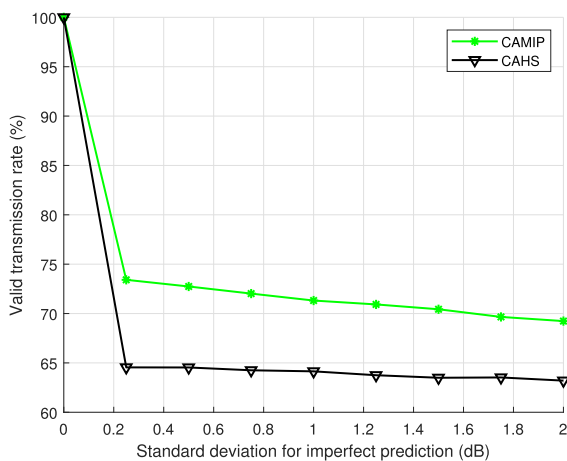


FIGURE 7. Valid transmission rate versus imperfect prediction.

Under this situation, to ensure that more packets can be received by users, one possible way is to increase the transmit power of BSs. To investigate whether it is possible to alleviate or overcome the packet loss by increasing transmit power, we choose a specific standard deviation and raise transmit power to different levels. Considering the high degree of mobility prediction [21] and channel prediction [7], we select 0.5 dB standard deviation as the level of imperfect prediction. Note that the real SINR threshold of successful reception, i.e. γ , is still 3 dB. But we increase the target SINR when

calculating the minimum required power Λ , instead of using γ directly. Let $K = 50$, $T = 64$, $N = 16$, standard deviation is equal to 0.5 dB, the target SINR varies from 3 to 5 dB, the results are illustrated in Fig. 8.

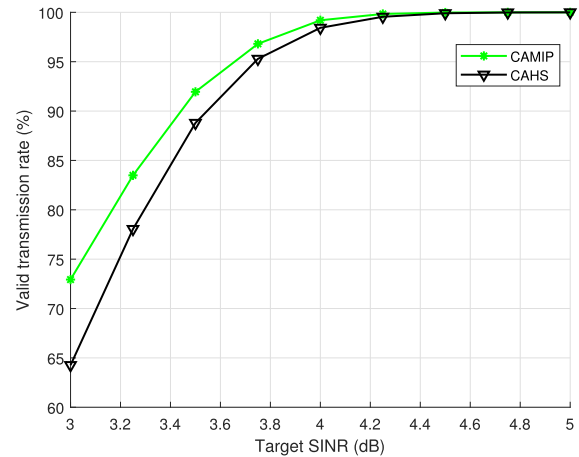


FIGURE 8. Valid transmission rate versus target SINR.

From Fig. 8, we can see that the valid transmission rate increases as the used target SINR increases. The valid transmission rate can get a huge improvement when the target SINR is increased by 0.5 dB. When the target SINR equals to 3.5 dB, the valid transmission rate of CAMIP and CAHS is 92% and 89%, respectively. If the target SINR increases continuously, the valid transmission rate approaches to 100% gradually. As expected, the results show that increasing the target SINR is an efficient way to alleviate the impact of imperfect prediction. It is worth noting that we should choose a proper threshold according to the degree of imperfect prediction to balance the energy consumption and valid transmission rate since the increase of target SINR will also lead to higher energy consumption.

Actually, BSs can dynamically adjust the transmit power, and the problem can be further alleviated based on real time CSI feedback in practice. Specifically, BSs can adjust the transmit power of BSs for the next time slot properly by comparing the predicted channel gain with the current real channel gain. For example, if the real channel gain is lower than the predicted one used in scheduling, the BS can raise the transmit power appropriately, and ensure that the target users can also receive packets at the next time slot. Similarly, if the real channel gain is higher than the predicted one, the BS can reduce the transmit power appropriately to reduce energy consumption.

B. CO-CHANNEL INTERFERENCE

Recall that the noise is modeled as a constant and co-channel interference is omitted when calculating the minimum required power Λ . This assumption holds when the inter-cell interference is negligible, which can be achieved by some inter-cell interference mitigation methods, such as conventional hard frequency reuse with large frequency

reuse factors, fractional frequency reuse with multiple cell regions [42], and soft frequency reuse [43]. For these frequency reuse techniques, the co-channel interference is mitigated at the expense of lower spatial bandwidth efficiency [44]. However, it is known that all BSs are simultaneously reusing the same frequency resources can maximize system-wide spectral efficiency [45], which leads to non-negligible co-channel interference. Since co-channel interference is omitted when making resource pre-allocation, the proposed schemes will suffer packet loss when applied to those systems with co-channel interference.

Simulations are carried out to investigate how much performance degradation will cause by co-channel interference. In our simulation, in the resource allocation phase, we still assume the frequency reuse factor is seven and no co-channel interference is considered. While in the transmission phase, the frequency reuse factor is regarded as one and inter-cell interference is counted when calculating the real SINR at user side. Let $T = 64$, $N = 16$, K varies from 5 to 100. The results are illustrated in Fig. 9. As the number of users increases, the valid transmission rate decreases first and then increases. Since the number of BSs is 19, when K is small, there exists some spare BSs at each time slot with high probabilities, which will not cause inter-cell interference. As the number of users increases, more and more BSs will work, and the co-channel interference will become more serious, which will make the valid transmission decrease.

When the number of users is larger than 19, each BS may sever multiple users simultaneously. For a given cell, the transmit power is allocated to make the ‘worst’ user achieve the SINR threshold γ , which means the received SINR of the other users in this cell is larger than γ without considering co-channel interference. In other words, the co-channel interference mainly affects those users located at cell borders, with a negligible impact on those users located close to BS. As the number of users increases, the ratio of border users to total users decreases, which makes the valid transmission increase. Therefore, when the number of users

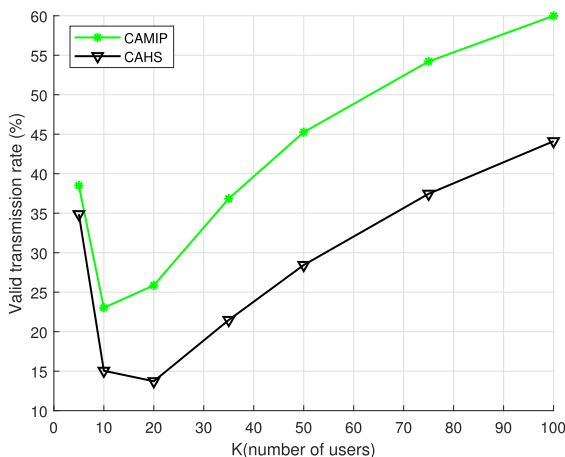


FIGURE 9. Valid transmission rate versus K when considering co-channel interference.

is large, the influence of co-channel interference becomes smaller. Nevertheless, co-channel interference still causes heavy performance degradation. In particular, when $K = 50$, the valid transmission rates of CAMIP and CAHS is only 45% and 28%, respectively.

To alleviate the impact of co-channel interference, BSs can increase the transmit power to reduce packet loss. Therefore, we can treat co-channel interference as ‘noise’ and set a larger constant value for η_0 in equation (1). Here, we evaluate how the valid transmission rate changes by setting different constant values for η_0 . Note that the transmission from neighboring cells is regarded as co-channel interference when determine real received SINR at user side. Similarly, we set $K = 50$, $T = 64$, $N = 16$, and test the ratio of successfully transmitted packets by varying η_0 from -130 to -50 dBm. The results are presented in Fig. 10. It is easy to conclude that the valid transmission rate becomes higher with the increase of constant η_0 . When the constant is equal to -50 dBm, both CAMIP and CAHS can achieve 100% valid transmission rate, which means the packet loss caused by co-channel interference can be eliminated when an appropriate value is selected for η_0 .

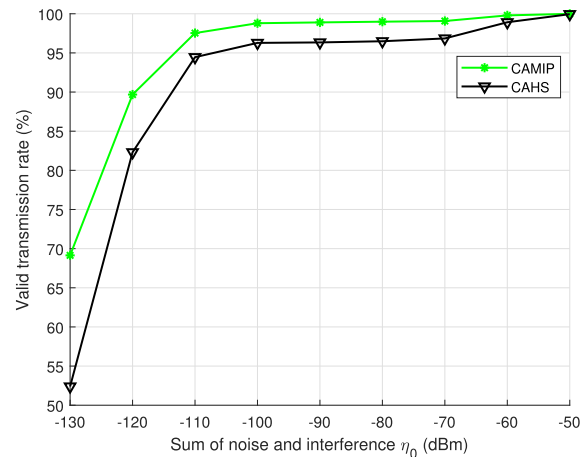


FIGURE 10. Valid transmission rate versus η_0 .

C. INFEASIBLE SCENARIOS

In our simulation assessment, all users are assumed to be feasible, i.e. all users must be within transmission range of BSs for at least N slots. This requirement is unnecessary in real scenarios. Due to signal attenuation or interference, there may exist some users which cannot be covered for N times by BSs within the latency constraint. For these users, they can never successfully receive all source packets in any case. In this scenario, we can simply skip these users and consider the remaining feasible users only. Take the users who can satisfy the latency constraint as a group, and then use the proposed schemes to allocate resources.

Besides, infeasible situation may also happen during the transmission process. As resource pre-allocation schemes, they are designed to run once before transmission. For a user

who is required to receive N packets, the proposed schemes will pre-allocate N transmissions for that user. However, as explained in the previous subsections, both imperfect prediction and co-channel interference may cause packet loss, which will make that user infeasible. To address this issue, on one hand, we can appropriately raise or dynamically adjust transmit power to reduce the probability of packet loss. On the other hand, we can run the proposed resource allocation schemes multiple times during the transmission process, by taking the receiving status of all users as dynamic input. By doing so, extra transmissions can be dynamically allocated for those users which have suffered packet loss, so as to increase the valid transmission rate.

We evaluate the performance of running multiple allocations during broadcasting to cope with potential infeasible situations caused by imperfect prediction and co-channel interference. Instead of making resource allocation before transmission, we run the proposed schemes before every time slot, taking the most recent receiving status as input. To evaluate the performance improvement under imperfect prediction scenarios, we run simulations with $K = 50$, $T = 64$, $N = 16$ and standard deviation is equivalent to 0.5 dB. If we run the resource allocation schemes once before transmission, the valid transmission rate of CAMIP and CAHS is 73% and 65%, respectively. While for this multiple running method, the valid transmission rate of CAMIP and CAHS can separately achieve 93% and 90%. To evaluate the performance improvement under co-channel interference scenarios, we run simulations with $K = 50$, $T = 64$, $N = 16$, and count co-channel interference at the user side. If the resource allocation schemes run once before transmission, the valid transmission rate of CAMIP and CAHS is only 45% and 28%, respectively. While for this multiple running method, the valid transmission rate of CAMIP and CAHS increases to 76% and 63%, respectively. From the results, we can find that the multiple allocation method shows very considerable improvements.

VIII. CONCLUSION

This paper considers the problem of energy minimization for delay-constrained multi-cell broadcast networks. Network coding integrated with context-awareness which exploits context information to predict users' channel gain in certain future time periods for transmission scheduling is considered for energy saving. We formulate the energy minimization problem as a context-aware network coded broadcast scheduling optimization problem and prove it is NP-hard. In order to get the optimal solution, we transform the optimization problem into an MIP problem which can be solved by some standard tools, called CAMIP. Due to the high complexity of CAMIP, a polynomial-time heuristic algorithm CAHS is proposed. In the simulation, we set a network with 19 BSs, and consider three performance metrics (i.e. energy consumption, success rate, and average broadcast latency) for performance evaluation. Meanwhile, two conventional schemes, GS and GSNC, are introduced for comparison.

Algorithm 2 General Scheme (GS)

Require: Required power matrix Λ .

Ensure: Transmit power \mathbf{P} and total energy consumption E .

```

1: Define a set  $R$  as the packet collection, define  $users^*$  to
   save the incomplete users, and  $BSs^*$  to save BSs which
   can be used to transmit packets at a particular time slot
   in retransmission phase.
2: Initialize target user indication matrix  $\mathbf{X}=\mathbf{0}$ .
3: // Initial phase, broadcast  $N$  packets sequentially
4: for  $t \in [1, N]$  do
5:   for  $i \in [1, K]$  do
6:      $\Lambda_{ij't'} \leftarrow \min(\Lambda_{ijt})$ , for all  $BS_j$  which can cover  $U_i$ 
       at time slot  $t$ 
7:      $X_{ij't'} \leftarrow 1$ 
8:   end for
9: end for
10: // Find the collection of packets need to be retransmitted
11: for  $i \in [1, N]$  do
12:   if the  $i$ -th packet is required by any user then
13:     put  $i$ -th packet into  $\mathbf{R}$ 
14:   end if
15: end for
16: // Retransmission phase, retransmit lost packets
17: while there exists user is incomplete do
18:    $t = t + 1$ 
19:   if  $t > T$  then
20:     return
21:   end if
22:    $users^* \leftarrow$  all incomplete users
23:    $BSs^* \leftarrow$  all BSs
24:   for  $r \in R$  do
25:     for all user  $U_i$  who need packet  $r$  of  $users^*$  do
26:        $\Lambda_{ij't'} \leftarrow \min(\Lambda_{ijt})$ , for all  $BS_j$  which can cover
          $U_i$  at time slot  $t$ 
27:        $X_{ij't'} \leftarrow 1$ 
28:       delete  $U_i$  and  $BS_j$  from  $users^*$  and  $BSs^*$  respec-
         tively
29:     end for
30:     if  $users^*$  or  $BSs^*$  is empty then
31:       break
32:     end if
33:   end for
34: end while
35:  $P_{jt} = \max_{i=1}^K \{X_{ijt} \Lambda_{ijt}\}$ , for  $j \in [1, M]$ ,  $t \in [1, T]$ 
36:  $E = \sum_{j=1}^M \sum_{t=1}^T P_{jt}$ 
37: return  $\mathbf{P}$  and  $E$ 

```

Simulation results show that both CAMIP and CAHS could significantly decrease transmission energy consumption of BSs and improve broadcast success rate as well, compared with conventional schemes. However, every coin consists of two sides. The proposed schemes get better energy consumption performance, at the expense of higher average broadcast latency and decoding complexity. Finally, we discuss some

Algorithm 3 General Scheme With Network Coding (GSNC)**Require:** Required power matrix Λ .**Ensure:** Transmit power \mathbf{P} and total energy consumption E .

```

1: Initialize target user indication matrix  $\mathbf{X}=\mathbf{0}$ 
2: // Initial phase, broadcast  $N$  packets sequentially
3: for  $t \in [1, N]$  do
4:   for  $i \in [1, K]$  do
5:      $\Lambda_{i'j't'} \leftarrow \min(\Lambda_{ijt})$ , for all  $BS_j$  which can cover  $U_i$ 
       at time slot  $t$ 
6:      $X_{i'j't'} \leftarrow 1$ 
7:   end for
8: end for
9: // Retransmission phase, retransmit lost packets
10: while there exists user is incomplete do
11:    $t = t + 1$ 
12:   if  $t > T$  then
13:     return
14:   end if
15:   for all incomplete users do
16:      $\Lambda_{i'j't'} \leftarrow \min(\Lambda_{ijt})$ , for all  $BS_j$  which can cover  $U_i$ 
       at time slot  $t$ 
17:      $X_{i'j't'} \leftarrow 1$ 
18:   end for
19: end while
20:  $P_{jt} = \max_{i=1}^K \{X_{ijt} \Lambda_{ijt}\}$ , for  $j \in [1, M]$ ,  $t \in [1, T]$ 
21:  $E = \sum_{j=1}^M \sum_{t=1}^T P_{jt}$ 
22: return  $\mathbf{P}$  and  $E$ 

```

problems which may arise in practice and present the corresponding solutions. We can appropriately raise or dynamically adjust transmit power, or execute multiple allocations according to the receiving status of all users to reduce the probability of packet loss.

In future work, we are going to study a system model considering co-channel interference and to investigate more application scenarios based on the proposed approaches, such as edge computing.

APPENDIX I. CONVENTIONAL SCHEMES

See Algorithms 2 and 3.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions that helped improve the quality of this paper.

REFERENCES

- [1] Ericsson. (Jun. 2020). *Ericsson Mobility Report*. [Online]. Available: <https://www.ericsson.com/49da93/assets/local/mobility-report/documents/2020/june2020-ericsson-mobility-report.pdf>
- [2] A. De Domenico, E. Calvanese Strinati, and A. Capone, "Enabling green cellular networks: A survey and outlook," *Comput. Commun.*, vol. 37, pp. 5–24, Jan. 2014.
- [3] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, Apr. 1998.
- [4] J. R. Marasevic, "Resource allocation in wireless networks: Theory and applications," Ph.D. dissertation, Dept. Elect. Eng., Columbia Univ., New York, NY, USA, 2016.
- [5] S. Stanczak, M. Wiczanowski, and H. Boche, *Resource Allocation in Wireless Networks: Theory and Algorithms*, vol. 4000. Berlin, Germany: Springer, 2006.
- [6] A. Nadembega, A. Hafid, and T. Taleb, "A destination and mobility path prediction scheme for mobile networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2577–2590, Jun. 2015.
- [7] Q. Liao, S. Valentin, and S. Starićzak, "Channel gain prediction in wireless networks based on spatial-temporal correlation," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2015, pp. 400–404.
- [8] M. Proebster, M. Kaschub, T. Werthmann, and S. Valentin, "Context-aware resource allocation for cellular wireless networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, no. 1, p. 216, Dec. 2012.
- [9] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system," in *Proc. 7th ACM SIGCOMM Conf. Internet Meas. (IMC)*, 2007, pp. 1–14.
- [10] J. J. Gimenez, D. Gomez-Barquero, J. Morgade, and E. Stare, "Wideband broadcasting: A power-efficient approach to 5G broadcasting," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 119–125, Mar. 2018.
- [11] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
- [12] C. Zhan, Y. Xu, J. Wang, and V. Lee, "Reliable multicast in wireless networks using network coding," in *Proc. IEEE 6th Int. Conf. Mobile Adhoc Sensor Syst. (MASS)*, Oct. 2009, pp. 506–515.
- [13] A. Ahmad, S. Ahmad, M. H. Rehmani, and N. U. Hassan, "A survey on radio resource allocation in cognitive radio sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 888–917, 2nd Quart., 2015.
- [14] M. Čagalj, J.-P. Hubaux, and C. Enz, "Minimum-energy broadcast in all-wireless networks: NP-completeness and distribution issues," in *Proc. 8th Annu. Int. Conf. Mobile Comput. Netw.*, 2002, pp. 172–182.
- [15] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides, "Energy-efficient broadcast and multicast trees in wireless networks," *Mobile Netw. Appl.*, vol. 7, no. 6, pp. 481–492, Dec. 2002.
- [16] S. Al-Shihabi, P. Merz, and S. Wolf, "Nested partitioning for the minimum energy broadcast problem," in *Learning and Intelligent Optimization (Lecture Notes in Computer Science)*, vol. 5313. Berlin, Germany: Springer, 2007, pp. 1–11.
- [17] P.-C. Hsiao, T.-C. Chiang, and L.-C. Fu, "Static and dynamic minimum energy broadcast problem in wireless ad-hoc networks: A PSO-based approach and analysis," *Appl. Soft Comput.*, vol. 13, no. 12, pp. 4786–4801, Dec. 2013.
- [18] A. Singh and W. N. Bhukya, "A hybrid genetic algorithm for the minimum energy broadcast problem in wireless ad hoc networks," *Appl. Soft Comput.*, vol. 11, no. 1, pp. 667–674, Jan. 2011.
- [19] D. Arivudainambi and D. Rekha, "Memetic algorithm for minimum energy broadcast problem in wireless ad hoc networks," *Swarm Evol. Comput.*, vol. 12, pp. 57–64, Oct. 2013.
- [20] L. Georgiadis, M. J. Neely, and L. Tassioulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends Netw.*, vol. 1, no. 1, pp. 1–144, 2005, doi: 10.1561/13000000001.
- [21] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, Feb. 2010.
- [22] H. Zhang and L. Dai, "Mobility prediction: A survey on state-of-the-art schemes and future applications," *IEEE Access*, vol. 7, pp. 802–822, 2019.
- [23] P. Lai, X. Wang, N. Lu, and F. Liu, "A reliable broadcast routing scheme based on mobility prediction for VANET," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2009, pp. 1083–1087.
- [24] S.-H. Cha, J.-E. Lee, and M. Ryu, "Directed broadcasting with mobility prediction for vehicular sensor networks," *J. Distrib. Sensor Netw.*, vol. 12, no. 7, pp. 1–9, 2016.
- [25] I. Jemili, D. Ghrab, A. Belghith, and M. Mosbah, "Context-aware broadcast in duty-cycled wireless sensor networks," *Int. J. Semantic Web Inf. Syst.*, vol. 13, no. 3, pp. 48–67, Jul. 2017.
- [26] A. Shrivastava, P. Bansod, K. Gupta, and S. N. Merchant, "An improved multicast based energy efficient opportunistic data scheduling algorithm for VANET," *AEU-Int. J. Electron. Commun.*, vol. 83, pp. 407–415, Jan. 2018.
- [27] Y. Wu, P. A. Chou, and S.-Y. Kung, "Minimum-energy multicast in mobile ad hoc networks using network coding," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1906–1918, Nov. 2005.

- [28] Z. Guoyan and L. Yonghua, "Cooperative multicast transmission strategy for energy-efficient dynamic network coding," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, Jun. 2013, pp. 479–483.
- [29] A. Tassi, F. Chiti, R. Fantacci, and F. Schoen, "An energy-efficient resource allocation scheme for RLNC-based heterogeneous multicast communications," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1399–1402, Aug. 2014.
- [30] P. Li, S. Guo, S. Yu, and A. V. Vasilakos, "Reliable multicast with pipelined network coding using opportunistic feeding and routing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 12, pp. 3264–3273, Dec. 2014.
- [31] Y. Li, J. Zhou, J. Wang, Z. Bao, T. Q. S. Quek, and J. Wang, "On data dissemination enhanced by network coded device-to-device communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3963–3976, Jun. 2020.
- [32] L. Wang, Y. Li, B. Pan, Q. Wu, J. Yin, and L. Xu, "Network coding for efficient video multicast in device-to-device communications," *Sensors*, vol. 20, no. 8, p. 2254, Apr. 2020.
- [33] D. Jiang, Z. Xu, W. Li, C. Yao, Z. Lv, and T. Li, "An energy-efficient multicast algorithm with maximum network throughput in multi-hop wireless networks," *J. Commun. Netw.*, vol. 18, no. 5, pp. 713–724, 2016.
- [34] L. Huang, C. W. Sung, and C. S. Chen, "Context-aware wireless broadcast for next generation cellular networks," in *Proc. IEEE Int. Conf. Commun. Syst. (ICCS)*, Nov. 2014, pp. 349–353.
- [35] C. W. Sung, K. W. Shum, L. Huang, and H. Y. Kwan, "Linear network coding for erasure broadcast channel with feedback: Complexity and algorithms," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2493–2503, May 2016.
- [36] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.
- [37] Y. Pochet and L. A. Wolsey, *Production Planning By Mixed Integer Programming*. Berlin, Germany: Springer, 2006.
- [38] *Further Advancements for E-UTRA Physical Layer Aspects*, document TR 36.814, 3GPP, 2010.
- [39] H. L. Bertoni, *Radio Propagation for Modern Wireless Systems*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2000.
- [40] M. Dräxler, J. Blobel, P. Dreimann, S. Valentin, and H. Karl, "Smarter-Phones: Anticipatory download scheduling for wireless video streaming," in *Proc. Int. Conf. Workshops Networked Syst. (NetSys)*, Mar. 2015, pp. 1–8.
- [41] L. Gurobi Optimization. (2020). *Gurobi Optimizer Reference Manual*. [Online]. Available: <http://www.gurobi.com>
- [42] H. Zhang, X.-D. Xu, J.-Y. Li, X.-F. Tao, S. Tommy, and B. Carmen, "Performance of power control in inter-cell interference coordination for frequency reuse," *J. China Univ. Posts Telecommun.*, vol. 17, no. 1, pp. 37–43, Feb. 2010.
- [43] G. Giambene and T. A. Yahiya, "LTE planning for soft frequency reuse," in *Proc. IFIP Wireless Days (WD)*, Nov. 2013, pp. 1–7.
- [44] M. S. Obaidat, F. Zarai, and P. Nicopolitidis, *Modeling and Simulation of Computer Networks and Systems: Methodologies and Applications*. Burlington, VT, USA: Morgan Kaufmann, 2015.
- [45] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Found. Trends Commun. Inf. Theory*, vol. 9, no. 2, pp. 113–381, 2013, doi: 10.1561/01000000069.



LIANG LU received the B.Sc. degree from the School of Information Engineering, Shaoyang University, Shaoyang, China, in 2018. He is currently pursuing the M.Sc. degree with the College of Electronics and Information Engineering, Sichuan University, Chengdu, China. His research interests include mobility prediction and resource allocation in wireless networks.



LINYU HUANG (Member, IEEE) received the Ph.D. degree in electronics engineering from the City University of Hong Kong, in 2014. He joined the Faculty as a Lecturer with the College of Electronics and Information Engineering, Sichuan University, Chengdu, China, in 2014. His research interests include wireless communications, resource allocation, power control, and network coding.



WEI HUA received the B.Sc., M.Sc. and Ph.D. degrees in biomedical engineering from Sichuan University, Chengdu, China, in 1990, 1997, and 2006, respectively. He is currently an Associate Professor with the College of Electronics and Information Engineering, Sichuan University. His research interests include multilateration, power line communication, and wireless communications.

...