Wireless Backhaul Optimization Algorithm Considering Time Delay in 5G Dynamic Heterogeneous Scenarios

Hui Zhang[1,2,3], Fengle Li[1]

[1] Jiangsu Key Lab of Wireless Communications, Nanjing University of Posts and Telecommunications, China

[2] National Engineering Research Center of Communications and Networking,
Nanjing University of Posts and Telecommunications, China

[3] Provincial Key Laboratory for Computer Information Processing Technology, Soochow University, China

zhhjoice@126.com, lifengle1991@126.com

Abstract

In the 5G ultra-dense scenario, the deployment of the base station has massiveness, randomness, and other characteristics. This makes wireline backhaul costs too high and difficult to implement, and thus wireless backhaul is a better choice to solve the problem. This paper proposes a wireless backhaul optimization algorithm by considering the time delay for 5G dynamic heterogeneous scenarios. First, a systematic analysis of the delay problem in the 5G dynamic heterogeneous scenario is performed, and an accurate backhaul optimization index is established. Furthermore, a basic backhaul model is built. On this basis, from the two perspectives of the user demand difference and the number of user overload, the improved model 1 and the improved model 2 are respectively constructed, and a hierarchical algorithm is constructed to quickly solve the problem. Simulations verify the effectiveness of the algorithm.

Keywords: 5G environment, Dynamic heterogeneous network, Wireless backhaul optimization, Delay

1 Introduction

With the rapid development of wireless communication technologies and the increasingly diversified needs of users, the fifth-generation mobile communication [1] technology (5G) supporting high-rate, low-latency, and massive-device connectivity has emerged. In the future 5G networks, in order to address the explosive growth of data traffic in local hotspots, ultra-dense networking technology will be commonly used. The network density makes the access point's access capability greatly enhanced, but the deployment of Small Cell Base Stations (SCBSs), such as Femtocell, Picocell, and Microcell, has ultra-dense and random features, which will provide low costs for small base stations. High-reliability backhaul becomes a challenge [2], and the backhaul link is liable to become the bottleneck for the entire network.

To date, many feasible solutions have been proposed to address the return problem [3-4]. The wired backhaul solution represented by optical fibre has the advantages of high speed, high reliability, etc. but has high costs and is time-consuming to deploy [5]. To reduce the increase in return costs caused by the network density, the use of wireless backhaul transmission technology is already a consensus in the field of 5G backhaul research. The most typical architecture currently adopted is the wireless backhaul aggregator node (BAN) architecture, and the use of millimetre-wave wireless transmission technology is currently the focus of research. In addition, the potential and application of millimetre waves [6] to 5G networks has been confirmed in [7], thus making it possible to use high-quality wireless backhaul for non-line-of-sight transmissions.

The current research on wireless backhaul optimization in the 5G environment focuses on reducing deployment costs, maximizing network throughput, and optimizing energy efficiency. In [9], the joint deployment costs of the best aggregation node, power control, channel arrangement, and routing are chosen to minimize the deployment costs of the network. In [10], the use of a hierarchical network structure to shape the small-cell network is considered. The wireless backhaul model studies the impacts of backhaul delay and deployment costs. In [11], the backhaul link and the access link in the network are considered comprehensively, and a joint optimization model is constructed to maximize the network throughput. In [12], taking into account the power consumption during the handover between the sleep state and the active state of a small base station, a heuristic base station dormancy scheme on two time scales is proposed. In [13], the base station is considered to be dormant at the same time. In addition

to the issue of power consumption in the backhaul network, a study was conducted to maximize the system's energy efficiency.

In an ultra-dense network with wireless backhaul, the user experience mainly depends on the end-to-end performance of the access link and the backhaul link, and the user experience is mainly limited by the network's backhaul capability. If the traffic of the backhaul SCBS needs to be balanced, the services that meet the user's needs can be provided through a rationally planned backhaul route. However, in an actual network, the services in each SCBS are always changing. On the one hand, this traffic is due to the turbulence of access services caused by users, and on the other hand, it is due to the complex interference environment and the channel state. If fixed-planned backhaul routing is still used, then users with large backhaul traffic will experience relatively large delays due to network congestion [8], thus resulting in degraded service quality. In backhaul networks, latency is a key factor affecting the user's business experience. Therefore, flexible backhaul technologies need to be adopted to rationally reuse the backhaul resources to match the changes in the business. Some studies [14] show that if the backhaul link delay problem is taken into account and the overall performance of the network is re-examined, it is not always advantageous to increase the small base station to limit the network densification without limitation. The current study takes very little of this into account. In addition, there is basically no research on optimization indicators. According to the above analysis, the user experience cannot be guaranteed. In view of the above problems, this paper proposes a wireless backhaul optimization algorithm that considers the delay of 5G dynamic heterogeneous scenarios. The algorithm incorporates the time delay into the optimization model, systematically analyses the time delay problem, establishes the time delay optimization index accurately, and constructs a basic backhaul model. On this basis, we construct the improved model 1 and improved model 2, respectively. The model is based on the differences in user requirements and the number of users overloaded, and the establishment of the above models takes into account the channel dynamics. Finally, a hierarchical algorithm is used to quickly solve the problem.

2 Network Scenarios and System Assumptions

This paper studies a two-layer heterogeneous network scenario in which a BAN is located in the upper layer and many SCBSs are covered by the BAN in the lower layer. The BAN is connected to the core network through a dedicated, high-speed fibre. The BAN uses a millimetre wave to communicate with the

SCBSs, and users communicate with the SCBSs with 6 GHz or less microwave frequency bands. Each SCBS user is covered by various communication needs, as shown in Figure 1. We assume that the BAN can perform both the aggregator and base station functions, has certain access capabilities, and can cover a certain area. This article considers mixed backhaul scenarios. There are two types of backhaul options. First, users can directly connect to the BAN through a one-hop wireless backhaul link, and then BAN performs the subsequent wired backhaul to the core network. Second, taking the corresponding SCBS as the relay, the user can connect to the BAN through a two-hop wireless backhaul link, and then BAN similarly performs the subsequent wired backhaul to the core network.
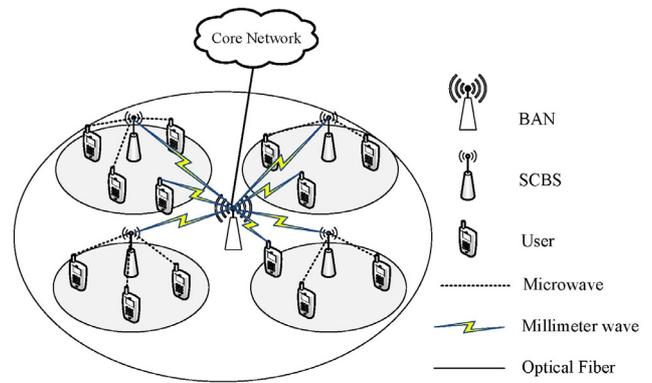
Figure 1. 5G two-tier heterogeneous network scenarios

Furthermore, we only consider the uplink transmission of the network. The set of SCBSs is defined as $SC = \{SC_1, SC_2, ..., SC_l, ..., SC_L\}$, and the set of users covered by SCBSs is defined as $UE_l = \{UE_{l1}, UE_{l2}, ..., UE_{ln}, ..., UE_{lN_l}\}$, in which UE_{ln} represents user n in the cell SC_l, and N_l denotes the number of users in the cell SC_l. It is assumed here that all SCBSs implement seamless coverage of the entire network and they do not intersect with each other, which means that $UE_i \bigcap UE_j = \varnothing$ ($i \neq j$).

Therefore, the number of users is $N = \sum_{l=1}^{L} N_l$. The access selection vector is defined as $A = \{a_{11}, a_{12}, ..., a_{1N_1}, a_{21}, a_{22}, ..., a_{2N_2}, ..., a_{L1}, a_{L2}, ..., a_{LN_L}\}$, where $a_{ln} = 1$ ($n \leq N_l$) indicates that user UE_{ln} directly connects to the BAN through a one-hop wireless backhaul link, and $a_{ln} = 0$ ($n \leq N_l$) indicates that user UE_{ln} takes the cell SC_l as the relay and connects to the BAN through a two-hop wireless backhaul link. The SCBS channel allocation matrix is defined as

$$B = \begin{bmatrix} b_{11,1} & b_{11,2} & ... & b_{11,M_1} \\ b_{12,1} & b_{12,2} & ... & b_{12,M_1} \\ ... & ... & b_{ln,m} & ... \\ b_{1N_1,1} & b_{1N_1,2} & ... & b_{1N_1,M_1} \\ b_{21,1} & b_{21,2} & ... & b_{21,M_1} \\ b_{22,1} & b_{22,2} & ... & b_{22,M_1} \\ ... & ... & ... & ... \\ b_{1N_2,1} & b_{2N_2,2} & ... & b_{2N_2,M_1} \\ ... & ... & ... & ... \\ b_{LN_L,1} & b_{LN_L,2} & ... & b_{LN_L,M_1} \end{bmatrix}$$

where $b_{ln,m} = 1$ indicates that the bandwidth m is assigned to user UE_{ln}, and $b_{ln,m} = 0$ indicates that the bandwidth m is not assigned to user UE_{ln}.

Here, $BW = \{BW_1, BW_2, ..., BW_{M_1}\}$ represents the channel vector that can be allocated by the SCBS, and M_1 is the number of channels. Define the BAN channel number allocation vector as $C = \{c_{11}, c_{12}, ..., c_{1N_1}, ..., c_{l1}, c_{l2}, ..., c_{ln}, ..., c_{lN_l}, ..., c_{L1}, c_{L2}, ..., c_{LN_L}\}$, where c_{ln} denotes the number of BAN channels assigned to it, and $c_{ln} \geq 1$ ($\forall l, n$). It should be noted that regardless of whether the user accesses the BAN backhaul or accesses the SCBS backhaul, a certain amount of BAN bandwidth must be allocated to the user. If $a_{ln} = 1$, then c_{ln} represents the number of channels allocated to the $UE_{ln} \rightarrow BAN$ link for the transmission of UE_{ln} data packets. If $a_{ln} = 0$, then c_{ln} represents the number of channels allocated to the $SC_l \rightarrow BAN$ link in the $UE_{ln} \rightarrow SC_l \rightarrow BAN$ path for the transmission of UE_{ln} packets.

The following definitions are made as follows. Theoretically, the allocated channel size should be continuous (big or small) so that it best meets the nature of the optimization and finds the true optimal solution. In practice, the allocated channel size is discrete because it is easy to implement in practice and is simple and convenient. This paper assumes that the channels are discrete, with the difference being that the BAN allocates multiple channels (i.e., the number of channels) to one SCBS, while the SCBS allocates only one channel to a certain user. The BAN allocates different numbers of channels to different users, which is in line with reality, is convenient for theoretical optimization and is flexible and optimal. The SCBS only allocates one channel to one user, and this is based on the following considerations. During a simple period, interference analysis for the user is facilitated.

The algorithm proposed in this paper is based on the following assumptions.

(1) The service optimization index is based on the packet granularity for the service uplink transmission scenario.

(2) To facilitate interference analysis, it is assumed that the bandwidths of the channels are equal. That is, the channel sizes on the BAN side are equal, and the channel sizes on the SCBS side are also equal. The total channel bandwidth of the BAN should be greater than the SCBS, and the bandwidth of each channel is equal to the SCBS channel.

(3) Assume that the three links of $UE_{ln} \rightarrow BAN$, $UE_{ln} \rightarrow SC_l$, and $SC_l \rightarrow BAN$ each transmit uplink traffic at a fixed, identical transmission rate.

(4) Assuming that the link channel capacity is less than its transmission rate, packet loss will occur, and each SCBS will have an infinite queue length so that packet loss will not occur due to queuing (or congestion). If the link generates packet loss, the system will start heavy packets. The send mechanism will then resend the package.

3 Delay Analysis

In the wireless communications environment, the link delay includes the propagation delay and transmission delay. Among them, the propagation delay depends on the signal propagation distance and the propagation speed of the electromagnetic wave, which is the speed of light. In the 5G communication scenario, the signal propagation distance is very short, and thus the propagation delay is very small and can be ignored. The transmission delay in this paper can be defined as the delay caused by the transmission of data packets from the sending end to the receiving end. In the following, system analysis will be performed for the delay problem in heterogeneous wireless scenarios.

3.1 SINR Analysis

In this scenario, the UE_{ln}-oriented wireless links a total of three types: link ℓ_{ln}^{UB} where UE_{ln} accesses the BAN, link ℓ_{ln}^{US} where UE_{ln} accesses the SC_l, and ℓ_{ln}^{SB} where SC_l accesses the BAN. The transmission powers of the three types of links are P_{UE}, P_{UE}, and P_{SC}. Here, it should be noted that, for the ease of the description, it is assumed that the transmission power of all users is equal, the transmission power of all cell base stations is equal, and the distances are d_{ln}^{UB}, d_{ln}^{US}, and d_l^{SB}, respectively. Assuming that all link channels obey small-scale Rayleigh fading, the received power is exponentially distributed as

$$f(p_{r_{ln}}^i) = \begin{cases} \dfrac{1}{\theta_{ln}^i} e^{-\frac{p_{r_{ln}}^i}{\theta_{ln}^i}}, & p_{r_{ln}}^i > 0 \\ 0, & p_{r_{ln}}^i = 0 \end{cases} \tag{1}$$

Here, the average received powers are,

$$\theta_{ln}^{UB} = p_{UE}\left(d_{ln}^{UB}\right)^{-\alpha} \tag{2}$$

$$\theta_{ln}^{US} = p_{UE}\left(d_{ln}^{US}\right)^{-\alpha} \tag{3}$$

$$\theta_{ln}^{SB} = p_{SC}\left(d_{l}^{SB}\right)^{-\alpha} \tag{4}$$

Here, α is the free space decline index, which is generally $\alpha = 2 \sim 4$.

Next, the signal to interference plus noise ratio (SINR) of the three types of links is further analysed. For the link ℓ_{ln}^{UB}, ℓ_{l}^{SB} connecting the BAN, there is no interference in the BAN because the same channel is not assigned in the BAN. Moreover, since the BAN adopts millimetre wave access, the BAN has a long distance, and it can be considered that there is no mutual interference between the BANs (an approximate on-off model). Therefore, the signal-to-noise-ratio of ℓ_{ln}^{UB} and ℓ_{l}^{SB} can be expressed as follows:

$$SINR_{ln}^{UB} = \frac{p_{r_{ln}}^{UB}}{n_0^{BAN} \cdot c_{ln}\Delta_{BAN}} \tag{5}$$

Here, $p_{r_{ln}}^{UB}$ represents the power received from user UE_{ln} at the BAN end and n_0^{BAN} represents the noise power spectral density of the BAN. Δ_{BAN} indicates the single channel bandwidth of the BAN. $c_{ln}\Delta_{BAN}$ indicates the total bandwidth allocated to UE_{ln}.

$$SINR_{ln}^{SB} = \frac{p_{r_{ln}}^{SB}}{n_0^{BAN} \cdot c_{ln}\Delta_{BAN}} \tag{6}$$

Here, $p_{r_{ln}}^{SB}$ denotes the power received by the BAN end from the user SC_l.

For the link ℓ_{ln}^{US} that accesses SC_l, since the same channel is not duplicated in the cell, there is no mutual interference in the cell SC_l, but there is mutual interference between the cells.

$$SINR_{ln}^{US} = \frac{p_{r_{ln}}^{US}}{n_0^{SC_l}\Delta_{SC} + \sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n' \\ (l' \neq l)}} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}\left(l'n'\right)} \tag{7}$$

Here, $p_{r_{ln}}^{US}$ represents the power received by user UE_{ln} from SC_l and $n_0^{SC_l}$ represents the noise power spectral density of SC_l. Δ_{SC} indicates the single channel bandwidth of the SCBS. $p_{r_{ln}}^{SU}\left(l'n'\right)$ represents the SC_l power received on the channel assigned to UE_{ln} by $UE_{l'n'}$. $\sum\limits_{\forall m}\sum\limits_{\forall l'n'} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}\left(l'n'\right)$ represents the total

power of interference from neighbouring cells. Here, the received power $p_{r_{ln}}^{SU}\left(l'n'\right)$ obeys the exponential distribution.

$$f\left(p_{r_{ln}}^{SU}\left(l'n'\right)\right) = \begin{cases} \dfrac{1}{\theta_{ln}^{SU}\left(l'n'\right)} e^{-\frac{p_{r_{ln}}^{SU}\left(l'n'\right)}{\theta_{ln}^{SU}\left(l'n'\right)}}, & p_{r_{ln}}^{SU}\left(l'n'\right) > 0 \\ 0 & , p_{r_{ln}}^{SU}\left(l'n'\right) = 0 \end{cases} \tag{8}$$

where $\theta_{ln}^{SU}\left(l'n'\right) = p_{UE}\left(d_{ln}^{SU}\left(l'n'\right)\right)^{-\alpha}$, and $d_{ln}^{SU}\left(l'n'\right)$ indicates the distance from $UE_{l'n'}$ to SC_l.

3.2 Packet loss Analysis

To calculate the packet loss rate, we first analyse the probability of successful uplink transmission. If $SINR \geq SINR_{th}$, we can obtain the corresponding uplink successful transmission probability. The thresholds for defining three types of links are

$$SINR_{ln}^{UB}\left(th\right) = 2^{\frac{H}{c_{ln}\Delta_{BAN}}} - 1 \tag{9}$$

$$SINR_{ln}^{US}\left(th\right) = 2^{\frac{H}{\Delta_{SC}}} - 1 \tag{10}$$

$$SINR_{ln}^{SB}\left(th\right) = SINR_{ln}^{UB}\left(th\right) \tag{11}$$

Among them, H is the transmission rate, and according to the assumption, the sending rates of the three links are the same.

We first derive the probability of successful uplink transmission from user UE_{ln} to the BAN:

$$\begin{aligned} p_{UE_{ln} \to BAN} &= P\left\{SINR_{UE_{ln} \to BAN} \geq SINR_{ln}^{UB}\left(th\right)\right\} \\ &= P\left\{\frac{p_{r_{ln}}^{1}}{n_0^{BAN} \cdot c_{ln}\Delta_{BAN}} \geq SINR_{ln}^{UB}\left(th\right)\right\} \\ &= \int_{n_0^{BAN} \cdot c_{ln}\Delta_{BAN} \cdot SINR_{ln}^{UB}\left(th\right)}^{+\infty} \frac{1}{\theta_{ln}^{1}} e^{-\frac{p_{r_{ln}}^{1}}{\theta_{ln}^{1}}} dp_{r_{ln}}^{1} \\ &= \exp\left(-\frac{n_0^{BAN} \cdot c_{ln}\Delta_{BAN} \cdot SINR_{ln}^{UB}\left(th\right)}{\theta_{ln}^{UB}}\right) \end{aligned} \tag{12}$$

In the same way, we can get the uplink's successful transmission probability from SC_l to the BAN for the small base station:

$$\begin{aligned} p_{SC_l \to BAN} &= P\left\{SINR_{SC_l \to BAN} \geq SINR_{ln}^{SB}\left(th\right)\right\} \\ &= \exp\left(-\frac{n_0^{BAN} \cdot c_{ln}\Delta_{BAN} \cdot SINR_{ln}^{SB}\left(th\right)}{\theta_{ln}^{SB}}\right) \end{aligned} \tag{13}$$

Derive the user's uplink success transmission probability from UE_{ln} to SC_l:

$$p_{UE_{ln} \to SC_l} = P\{SINR_{UE_{ln} \to SC_l} \geq SINR_{ln}^{US}(th)\}$$

$$= P\left\{ \frac{p_{r_{ln}}^{US}}{n_0^{SC_l}\Delta_{SC} + \sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n'\\(l'\neq l)}} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}(l'n')} \geq SINR_{ln}^{US}(th) \right\} \quad \textbf{(14)}$$

$$= P\left\{ p_{r_{ln}}^{US} \geq \left[n_0^{SC_l}\Delta_{SC} + \sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n'\\(l'\neq l)}} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}(l'n') \right] \cdot SINR_{ln}^{US}(th) \right\}$$

We assume that the distribution functions of the received power from different users are independent of each other. It is assumed from the assumption that $\left\{ p_{r_{ln}}^{SU}(l'n') \right\}$ is a set of mutually independent random variables that are all subject to an exponential distribution with a parameter of $\frac{1}{\theta_{ln}^{SU}(l'n')}$, where $l'n'$ denotes the user $UE_{n'}$ in $SC_{l'}$. According to the assumption that it is easy to know that $p_{r_{ln}}^{US}$ and $\sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n'\\(l'\neq l)\\(l'n'\neq l''n'')}} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}(l'n')$ are independent of each other, we note $X_{ln} = p_{r_{ln}}^{US}$, and then X_{ln} follows an exponential distribution with a parameter of $\frac{1}{\theta_{ln}^{US}}$. In the formula $\sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n'\\(l'\neq l)}} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}(l'n')$, $b_{ln,m}b_{l'n',m}$ is a Boolean variable, and then $\sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n'\\(l'\neq l)}} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}(l'n')$ is a linear combination of random variables $p_{r_{ln}}^{SU}(l'n')$. Note that $Y_s = \sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n'\\(l'\neq l)}} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}(l'n')$ and $s = \sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n'\\(l'\neq l)}} b_{ln,m}b_{l'n',m}$, where s is a range of $[0, L-1]$. Furthermore, in order to facilitate the derivation of the probability density function, we record the parameter $\frac{1}{\theta_{ln}^{SU}(l'n')}$ of the variable $p_{r_{ln}}^{SU}(l'n')$ corresponding to $b_{ln,m}b_{l'n',m}$ in $\{b_{ln,m}b_{l'n',m}\}$ in turn as $\lambda_i \, (i=1,2,...,s)$. According to [15], the probability density function of the random variable is

$$f(y_s) = \begin{cases} \prod\limits_{i=1}^{s}\lambda_i \left[\dfrac{e^{-\lambda_1 y_s}}{\prod\limits_{i=2}^{s}(\lambda_i - \lambda_1)} - \dfrac{e^{-\lambda_2 y_s}}{\prod\limits_{i=3}^{s}(\lambda_i - \lambda_2)(\lambda_2 - \lambda_1)} \right. \\ \quad + \dfrac{e^{-\lambda_3 y_s}}{\prod\limits_{i=4}^{s}(\lambda_i - \lambda_3)(\lambda_3 - \lambda_2)(\lambda_2 - \lambda_1)} + ... \\ \left. \quad + (-1)^{s-1}\dfrac{e^{-\lambda_s y_s}}{\prod\limits_{j=1}^{s-1}(\lambda_s - \lambda_j)} \right], \quad y_s > 0 \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad , \, y_s \leq 0 \end{cases} \quad \textbf{(15)}$$

Define $\xi_1, \xi_2, ..., \xi_n$ as the temporary random variables with independent identical distributions, which follow the exponential distributions with the parameters $\lambda_i \, (\lambda_i > 0)(i=1,2,...,n)$ respectively. Let $Z_n = \sum\limits_{i=1}^{n}\xi_i$. Then, the probability density function of the sum of n random variables is

$$f(z_n) = \begin{cases} \prod\limits_{i=1}^{n}\lambda_i \left[\dfrac{e^{-\lambda_1 z_n}}{\prod\limits_{i=2}^{n}(\lambda_i - \lambda_1)} - \dfrac{e^{-\lambda_2 z_n}}{\prod\limits_{i=3}^{n}(\lambda_i - \lambda_2)(\lambda_2 - \lambda_1)} \right. \\ \quad + \dfrac{e^{-\lambda_3 z_n}}{\prod\limits_{i=4}^{n}(\lambda_i - \lambda_3)(\lambda_3 - \lambda_2)(\lambda_2 - \lambda_1)} + ... \\ \left. \quad + (-1)^{n-1}\dfrac{e^{-\lambda_n z_n}}{\prod\limits_{j=1}^{n-1}(\lambda_n - \lambda_j)} \right], \quad z_n > 0 \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad z_n \leq 0 \end{cases} \quad \textbf{(16)}$$

As is known from the independence of $p_{r_{ln}}^{US}$ and $\sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n'\\(l'\neq l)\\(l'n'\neq l''n'')}} b_{ln,m}b_{l'n',m}p_{r_{ln}}^{SU}(l'n')$, X_{ln} and Y_s are two independent random variables. It is known from the definition of the independence of random variables that the joint probability density function of two mutually independent random variables is equal to the product of their respective edge probability density functions. Moreover, the joint probability density of X_{ln} and Y_s is

$$f(x_{ln}, y_s) = \begin{cases} \lambda_{ln}e^{-\lambda_{ln}x_{ln}} \cdot \prod\limits_{i=1}^{s}\lambda_i \left[\dfrac{e^{-\lambda_1 y_s}}{\prod\limits_{i=2}^{s}(\lambda_i - \lambda_1)} - \dfrac{e^{-\lambda_2 y_s}}{\prod\limits_{i=3}^{s}(\lambda_i - \lambda_2)(\lambda_2 - \lambda_1)} \right. \\ \quad + \dfrac{e^{-\lambda_3 y_s}}{\prod\limits_{i=4}^{s}(\lambda_i - \lambda_3)(\lambda_3 - \lambda_2)(\lambda_2 - \lambda_1)} + ... \quad \begin{matrix} x_{ln} > 0, \\ y_s > 0 \end{matrix} \\ \left. \quad + (-1)^{s-1}\dfrac{e^{-\lambda_s y_s}}{\prod\limits_{j=1}^{s-1}(\lambda_s - \lambda_j)} \right], \\ 0, \qquad\qquad\qquad\qquad\qquad\qquad\quad other \end{cases} \quad \textbf{(17)}$$

Then, the UE_{ln} to SC_l uplink's success transmission probability is as follows:

$$p_{UE_{ln} \to SC_l} = P\left\{ SINR_{UE_{ln} \to SC_l} \geq SINR_{ln}^{US}(th) \right\}$$

$$= P\left\{ \frac{p_{r_{ln}}^{US}}{n_0^{SC_l}\Delta_{SC} + \sum\limits_{\forall m}\sum\limits_{\substack{\forall l'n' \\ (l' \neq l)}} b_{ln,m} b_{l'n',m} p_{r_{ln}}^{SU}(l'n')} \geq SINR_{ln}^{US}(th) \right\}$$

$$= P\left\{ x_{ln} \geq \left(n_0^{SC_l}\Delta_{SC} + y_s \right) \cdot SINR_{ln}^{US}(th) \right\}$$

$$= \int_0^{+\infty} \int_{\left(n_0^{SC_l}\Delta_{SC}+y_s\right)\cdot SINR_{ln}^{US}(th)}^{+\infty} \lambda_{ln} e^{-\lambda_{ln,}x_{ln}} \cdot \prod_{i=1}^{s}\lambda_i \left[\frac{e^{-\lambda_1 y_s}}{\prod\limits_{i=2}^{s}(\lambda_i - \lambda_1)} - \frac{e^{-\lambda_2 y_s}}{\prod\limits_{i=3}^{s}(\lambda_i - \lambda_2)(\lambda_2 - \lambda_1)} \right.$$

$$\left. + \frac{e^{-\lambda_3 y_s}}{\prod\limits_{i=4}^{s}(\lambda_i - \lambda_3)(\lambda_3 - \lambda_2)(\lambda_2 - \lambda_1)} + ... + (-1)^{s-1}\frac{e^{-\lambda_s y_s}}{\prod\limits_{j=1}^{s-1}(\lambda_s - \lambda_j)} \right] dx_{ln} dy_s$$

$$= \frac{e^{-\lambda_{ln} \cdot n_0^{SC_l}\Delta_{SC}\cdot SINR_{ln}^{US}(th)} \cdot \prod\limits_{i=1}^{s}\lambda_i}{\lambda_{ln}\cdot SINR_{ln}^{US}(th)} \cdot \left[\frac{\lambda_1 + \lambda_{ln}\cdot SINR_{ln}^{US}(th) + 1}{\left[\lambda_1 + \lambda_{ln}\cdot SINR_{ln}^{US}(th)\right]\cdot \prod\limits_{i=2}^{s}(\lambda_i - \lambda_1)} \right.$$

$$- \frac{\lambda_2 + \lambda_{ln}\cdot SINR_{ln}^{US}(th) + 1}{\left[\lambda_2 + \lambda_{ln}\cdot SINR_{ln}^{US}(th)\right]\cdot \prod\limits_{i=3}^{s}(\lambda_i - \lambda_2)(\lambda_2 - \lambda_1)}$$

$$+ \frac{\lambda_3 + \lambda_{ln}\cdot SINR_{ln}^{US}(th) + 1}{\left[\lambda_3 + \lambda_{ln}\cdot SINR_{ln}^{US}(th)\right]\cdot \prod\limits_{i=4}^{s}(\lambda_i - \lambda_3)(\lambda_3 - \lambda_2)(\lambda_2 - \lambda_1)}$$

$$\left. + ... + (-1)^{s+1}\frac{\lambda_s + \lambda_{ln}\cdot SINR_{ln}^{US}(th) + 1}{\left[\lambda_s + \lambda_{ln}\cdot SINR_{ln}^{US}(th)\right]\cdot \prod\limits_{j=1}^{s-1}(\lambda_s - \lambda_j)} \right] \tag{18}$$

After obtaining the probability of successful uplink transmission on each link, we further analyse the packet loss rate of each link. Let the bit error rates for link ℓ_{ln}^{UB}, link ℓ_{ln}^{US}, and link ℓ_{ln}^{SB} be BER_{ln}^{UB}, BER_{ln}^{US}, and BER_{ln}^{SB}, respectively.

$$BER_{ln}^{UB} = 1 - p_{UE_{ln} \to BAN} \tag{19}$$

$$BER_{ln}^{US} = 1 - p_{UE_l \to SC_l} \tag{20}$$

$$BER_{ln}^{SB} = 1 - p_{SC_l \to BAN} \tag{21}$$

Define the length of each packet as PL ($PL \geq 3$). The packet loss rate is related to the error correction coding algorithms. For an error correction coding algorithm, its error-correction capability is positively related to the redundancy of transmitted information. By adjusting the parameter settings of the given error correction coding algorithm, its error-correction capability and information redundancy can be changed accordingly. For an actual system, the error-correction capability and information redundancy are

contradictory. In order to achieve the excellent channel utilization, it is assumed that there are three bit errors in the data packet, and then the packet is lost. At this time, the data packet needs to be retransmitted. The packet loss rate is

$$PER_{ln}^i = 1 - \left(1 - BER_{ln}^i\right)^{PL} - C_{PL}^1\left(1 - BER_{ln}^i\right)^{PL-1}\cdot BER_{ln}^i - C_{PL}^2\left(1 - BER_{ln}^i\right)^{PL-2}\left(BER_{ln}^i\right)^2 \tag{22}$$

Here, $C_{PL}^1 \frac{PL!}{1!(PL\ 1)!}$ and $C_{PL}^2 \frac{PL!}{2!(PL\ 2)!}$, which both are combination number formulas

3.3 Delay Analysis

Here, the transmission delay is used to describe the time occupied by transmitting a packet successfully between any two communication nodes. The transmission delay can be categorized as the instant transmission delay and the average transmission delay (i.e. the mathematical expectation of instant transmission delay). Due to the randomness of the instant transmission delay, it cannot be optimized

directly. Therefore, we deduce the average transmission delay and then use it as the optimization objective.

The average transmission delay of link ℓ^i_{ln} from the packet loss rate is deduced as

$$
\begin{aligned}
\tau^i_{ln} = & \left(1 - PER^i_{ln}\right)T + PER^i_{ln}\left(1 - PER^i_{ln}\right)2T + ... \\
& + \left(PER^i_{ln}\right)^{RT}\left(1 - PER^i_{ln}\right)(RT+1)T \\
& + \left(PER^i_{ln}\right)^{RT+1}(RT+1)T
\end{aligned} \tag{23}
$$

Among them, RT is the maximum number of retransmissions, T represents the transmission delay of one trip. In the actual transmission process, the transmission process is as follows. After the sender sends a data packet, it will enter the waiting state. If the data packet is received by the receiver, the receiver will send a confirmation message to the sender and send it. The same packet is sent to the sender. After confirmation from the sender, a confirmation message is sent to the receiver, and the packet is sent. Therefore, we let $T = 3\dfrac{PL}{H}$ denote the transmission delay of one trip.

τ^i_{ln} will be further organized as

$$
\tau^i_{ln} = \frac{1 - \left(PER^i_{ln}\right)^{RT+1}}{1 - PER^i_{ln}} \cdot T \tag{24}
$$

In particular, when $RT = \infty$, $\tau^i_{ln} = \dfrac{1}{1 - PER^i_{ln}}T$.

When $RT = 0$, $\tau^i_{ln} = T$.

Extending from the link to the path, the delay analysis for the $UE \to BAN$ path (that is, the path UE_{ln} directly back from the BAN) is relatively simple. A one-hop wireless link is directly connected to the wired core network, and the delay of the cable is very small. This documents that the studied wireless backhaul problem that is studied is irrelevant and can be ignored. Therefore, the initial delay and the service arrival delay of the radio access side in this path are both τ^{UB}_{ln}. For the $UE \to SC \to BAN$ path (that is, the backhaul path of UE_{ln} to BAN via SC_l), to facilitate analysis, note that $\tau^{US}_{ln} = \tau^{US}$ and $\tau^{SB}_{ln} = \tau^{SB}$. The analysis process is as follows. When $\tau^{US}_{ln} \geq \tau^{SB}_{ln}$,
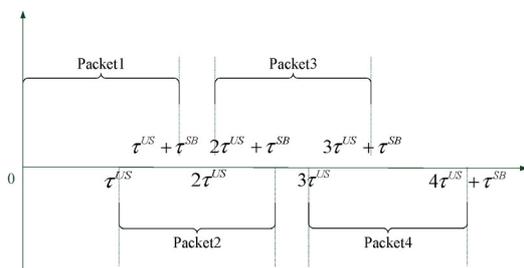
Figure 2. Arrival delay analysis I

From Figure 2, we can see that Packet1, Packet2, Packet3, ..., Packet k arrive at the BAN in the order of $\tau^{US} + \tau^{SB}$, $2\tau^{US} + \tau^{SB}$, $3\tau^{US} + \tau^{SB}$, ..., $k\tau^{US} + \tau^{SB}$. The time of arrival for SC_l is τ^{US}, $2\tau^{US}$, $3\tau^{US}$, ..., $k\tau^{US}$. Since $\tau^{US} \geq \tau^{SB}$, there is no queue congestion at SC_l. Therefore, from the expected average point of view, the service arrival delays (packet-to-packet time interval) for each packet to reach the BAN (radio access end) is $\tau^{US} = ma\left\{\tau^{US}, \tau^{SB}\right\}$, and the initial delay is $\tau^{US} + \tau^{SB}$.

When $\tau^{US}_{ln} < \tau^{SB}_{ln}$,
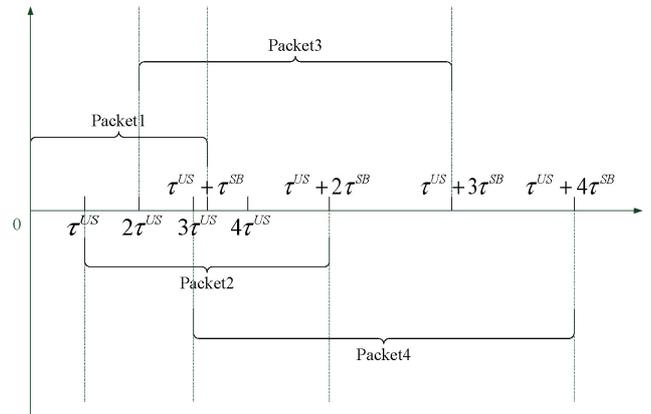
Figure 3. Arrival delay analysis II

From Figure 3, we can see that Packet1, Packet2, Packet3, ..., Packet k arrive at the BAN in the order of $\tau^{US} + \tau^{SB}$, $\tau^{US} + 2\tau^{SB}$, $\tau^{US} + 3\tau^{SB}$, ..., $\tau^{US} + k\tau^{SB}$. The time of arrival of SC_l is τ^{US}, $2\tau^{US}$, $3\tau^{US}$, ..., $k\tau^{US}$. Since $\tau^{US} < \tau^{SB}$, the queuing delays are τ^{US}, $2\tau^{US}$, $3\tau^{US}$, ..., $k\tau^{US}$ at SC_l. Here, assume that the queue can be infinitely long. Therefore, from the expected average point of view, the service arrival delays for each packet to reach the BAN (radio access end) is $\tau^{SB} = ma\left\{\tau^{US}, \tau^{SB}\right\}$, and the initial delay is $\tau^{US} + \tau^{SB}$.

4 Optimization Model Establishment

4.1 Basic Backhaul Model

Based on the above time delay analysis, the optimization goal can be expressed as follows:

$$
\begin{aligned}
\left[A^*, B^*, C^*\right] &= \underset{\forall a_{ln}, b_{ln,m}, c_{ln}}{\arg\min} \{U\} = \underset{\forall a_{ln}, b_{ln,m}, c_{ln}}{\arg\min} \left\{ \sum_l \sum_n U_{ln} \right\} \\
&= \underset{\forall a_{ln}, b_{ln,m}, c_{ln}}{\arg\min} \left\{ \sum_l \sum_n (1 - a_{ln})\left[r \cdot \max\left(\tau^{US}_{ln}, \tau^{SB}_{ln}\right)\right] + a_{ln} \cdot \tau^{UB}_{ln} \right\}
\end{aligned} \tag{25}
$$

Here, U_{ln} represents the optimization objective of

UE_{ln}, and U represents total optimization objective.

It should be specifically stated here that r is the initial delay compensation factor, and $r \geq 1$ reflects the effect of the initial delay in the $UE \rightarrow SC \rightarrow BAN$ path. In particular, when the arrival delays of two different paths (i.e. $UE \rightarrow SC \rightarrow BAN$ and $UE \rightarrow BAN$) are the same, the path with smaller initial delay (i.e. $UE \rightarrow BAN$) is chosen as the optimal path.

In turn, we can establish a basic optimization model:

$$\underset{\forall a_{ln}, b_{ln,m}, c_{ln}}{\arg\min} \left\{ \sum_l \sum_n (1 - a_{ln}) \left[r \cdot \max\left(\tau_{ln}^{US}, \tau_{ln}^{SB}\right) \right] + a_{ln} \cdot \tau_{ln}^{UB} \right\} \quad (26)$$

$$s.t. \quad \sum_{m=1}^{M_1} b_{ln,m} + a_{ln} = 1 \ \forall l, n \quad (27)$$

$$\sum_{n=1}^{N_l} b_{ln,m} \leq 1 \ \forall l, m \quad (28)$$

$$\sum_{\forall l, n} c_{ln} \leq M_2 \quad (29)$$

$$a_{ln} \in \{0, 1\} \ \forall l, n \quad (30)$$

$$b_{ln,m} \in \{0, 1\} \ \forall l, n, m \quad (31)$$

$$c_{ln} \geq 1 \ \forall l, n \quad (32)$$

Among them, constraint (27) indicates that the UE_{ln} can directly access the BAN backhaul, and then the SCBS does not allocate channels. Conversely, it may access the SCBS, the SCBS performs the backhaul, and then the SCBS allocates one bandwidth. Constraint (28) indicates that the same channel of the same cell is allocated to a certain user at most in order to avoid intra-cell interference. Constraint (29) defines that the number of channels allocated by the BAN cannot exceed its maximum value, where M_2 is the number of channels that the BAN can allocate. Constraints (30) and (31) limit a_{ln} and $b_{ln,m}$ to binary numbers, respectively. Constraint (32) indicates that regardless of whether the user accesses the BAN backhaul or accesses the SCBS backhaul, a certain amount of BAN bandwidth must be allocated for the user.

4.2 Improved Model 1

As this paper assumes that each user (access service) does not have an absolute requirement for time delay, the optimization model system is a soft guarantee for the arrival delay of each user's service. The optimization of the overall goal leads to the unbalanced optimization of individual goals (single user delays). The basic backhaul optimization model assumes that the user requirements are consistent and does not conform to the actual situation of the system. Therefore,

we consider introducing delay weights according to the access service type. The optimization goal is re-stated as follows:

$$\left[A^*, B^*, C^* \right] = \underset{\forall a_{ln}, b_{ln,m}, c_{ln}}{\arg\min} \{U\} = \underset{\forall a_{ln}, b_{ln,m}, c_{ln}}{\arg\min} \left\{ \sum_l \sum_n U_{ln} \right\}$$
$$= \underset{\forall a_{ln}, b_{ln,m}, c_{ln}}{\arg\min} \left\{ \sum_l \sum_n w_{ln} \left\{ (1 - a_{ln}) \left[r \cdot \max\left(\tau_{ln}^{US}, \tau_{ln}^{SB}\right) \right] + a_{ln} \cdot \tau_{ln}^{UB} \right\} \right\} \quad (33)$$

Among them, w_{ln} is weighted, and the weight of the service that requires high real time is greater.

Therefore, the improved model 1 is as follows:

$$\underset{\forall a_{ln}, b_{ln,m}, c_{ln}}{\arg\min} \left\{ \sum_l \sum_n w_{ln} \left\{ (1 - a_{ln}) \left[r \cdot \max\left(\tau_{ln}^{US}, \tau_{ln}^{SB}\right) \right] + a_{ln} \cdot \tau_{ln}^{UB} \right\} \right\}$$

$$s.t. \quad \sum_{m=1}^{M_1} b_{ln,m} + a_{ln} = 1 \ \forall l, n \quad (34)$$

$$\sum_{n=1}^{N_l} b_{ln,m} \leq 1 \ \forall l, m \quad (35)$$

$$\sum_{\forall l, n} c_{ln} \leq M_2 \quad (36)$$

$$a_{ln} \in \{0, 1\} \ \forall l, n \quad (37)$$

$$b_{ln,m} \in \{0, 1\} \ \forall l, n, m \quad (38)$$

$$c_{ln} \geq 1 \ \forall l, n \quad (39)$$

4.3 Improved Model 2

The network load is used to characterize the resource utilization of the network under certain conditions. It is generally measured by the number of users that can be carried or the amount of available channel resources. There is no problem of overloading the number of users in the establishment of the basic backhaul model and the improved model 1. Furthermore, we discuss the special case where the number of users (load) exceeds the network carrying capacity, which is

$$N = \sum_{l=1}^{L} N_l > M_2 \quad (40)$$

At this point, even if you set $c_{ln} = 1 (\forall l, n)$, there is still $\sum_{\forall l, n} c_{ln} = N > M_2$. That is, the channels of the BAN are not allocated to all users. At this time, an Admission Control (AC) mechanism needs to be activated to reject the requests of some users. As for which users are rejected, the following improved model 2 needs to be established for accepting how users allocate resources:

$$\arg\min\left\{U + \omega\left(M_2 - \sum_{\forall l,n} c_{ln}\right)\right\} \qquad (41)$$

$$s.t. \sum_{m=1}^{M_1} b_{ln,m} + a_{ln} \leq 1 \ \forall l,n \qquad (42)$$

$$\sum_{n=1}^{N_l} b_{ln,m} \leq 1 \ \forall l,m \qquad (43)$$

$$\sum_{\forall l,n} c_{ln} \leq M_2 \qquad (44)$$

$$a_{ln} \in \{0,1\} \ \forall l,n \qquad (45)$$

$$b_{ln,m} \in \{0,1\} \ \forall l,n,m \qquad (46)$$

$$c_{ln} \in \{0,1\} \ \forall l,n \qquad (47)$$

Among them, ω is the adjustment factor. If ω is smaller, it is apt to not assign the BAN channel to the user. If ω is bigger, it tends to assign the channel to the user. Here, ω is set as large so that resources are allocated to users as much as possible, and then U is optimized (making the delay as small as possible). Note that if $c_{ln} = 0$, the user's request is rejected.

5 Improved Model Solution

The improved model 1 and improved model 2 are essentially integer programming problems. In order to obtain the corresponding optimal solutions quickly, the branch and bound algorithms are adopted to solve the above two optimization models.

5.1 Improved Model 1 Solution

In the improved model 1, the solution vector A is a Boolean vector, the solution matrix B is a Boolean matrix, and the solution vector C is an integer vector, which stores the decision variables for the user to connect to the SCBS or BAN, the channel allocation for the SCBS, and the channel allocation for the BAN. Based on the characteristics of solution, we divide the original optimization problem into three layers. First, the initialization of the solution is performed, which is as follows.

Initialization: Generate $A = \text{zeros}(1, N)$,

$\qquad\qquad B = \text{zeros}(N, M_1)$, $C = \text{ones}(1, N)$

Input: BW, SC, L, N_l, M_1

for $b_{ln} \in B$

    for $BW_m \in BW$

        if

$$\left(m = \left(\sum_{l=1}^{l-1} N_l + n\right) \bmod M_1\right) \| \left(\left(\sum_{l=1}^{l-1} N_l + n\right)\Big/ M_1 \in N^+\right)$$

then $b_{ln,m} \leftarrow 1$ **end if**

        end for

end for

for $BW_m \in BW$

for $SC_l \in SC$

    if $\sum_{n=1}^{N_l} b_{ln,m} \geq 2$ **then**

$b_{ln,m'} = find\left(b_{ln,m} = 1\right), b_{ln,m'} \leftarrow 0, a_{ln} \leftarrow 1$ **end if**

    end for

end for

Output: Initial solution A**, initial solution** B**, initial solution** C

The first layer solves the matrix A. After initialization, we branch at each component of the vector A according to constraint (37). Then, we iterate the process. In addition, a feasible solution of the matrix B is obtained according to constraint (34) while solving the vector A. The solution process is shown in Algorithm 1.

Algorithm 1. Algorithm for Solving A Vector

Input: A, B, C, w_{ln}, r_1, RT, PL, H, P_{UE}, P_{SC}, d_{ln}^{UB}, d_{ln}^{US}, d_l^{SB}

Calculate f_0；

for $a_{ln} \in A$

    if $a_{ln} = 0$ **then**

$a_{ln} \leftarrow 1, b_{ln,m'} = find\left(b_{ln,m} = 1\right), b_{ln,m'} \leftarrow 0$, Calculate f_{ln},

        if $f_{ln} < f_0$ **then** $f_0 \leftarrow f_{ln}$ **else** $a_{ln} \leftarrow 0, b_{ln,m'} \leftarrow 1$

        end if

    end if

end for

Output: optimized solution A^***, feasible solution** B**, value** f_0

In the second layer, solve the matrix B. Based on the solution of the first layer, we further optimize the channel allocation of users connected to the SCBS. According to constraints (34) and (35), the available channels in the cell in which the user is located are traversed, respectively, and each binary component in the B matrix is branched. The solution process is as described in Algorithm 2.

Algorithm 2. Algorithm for Solving B Matrix

Input: A^*, B, C, BW, f_0, N_l, M_1, w_{ln}, r_1, RT, PL, H, P_{UE}, P_{SC}, d_{ln}^{UB}, d_{ln}^{US}, d_l^{SB}

for $a_{ln} \in A$

if $\left(a_{ln} = 0\right) \&\& \left(\exists b_{ln,m'} = 1\right)$ **then**

    for $BW_m \in BW$

        if $\left(BW_m \neq BW_{m'}\right) \&\& \left(\sum_{n'=1}^{N_l} b_{ln',m} = 0\right)$

        then $b_{ln,m} \leftarrow 1$, $b_{ln,m'} \leftarrow 0$, Calculate f_m,

        if $f_m < f_0$ **then** $f_0 \leftarrow f_m$

        else $b_{ln,m} \leftarrow 0$, $b_{ln,m'} \leftarrow 1$ **end if**

          end if

            end for

    end if

end for

Output: optimized solution B^*

In the third layer, solve the vector C. In the third layer's solution, we assign the channels in the BAN according to constraints (36) and (39). The solution process is as described in Algorithm 3.

Algorithm 3. Algorithm for Solving C Vector

Input: A^*, B^*, C, f_0, L, N_l, M_2, w_{ln}, r_1, RT, PL, H, P_{UE}, P_{SC}, d_{ln}^{UB}, d_{ln}^{US}, d_l^{SB}

while $\sum_{l=1}^{L} \sum_{n=1}^{N_l} c_{ln} < M_2$

    for $c_{ln} \in C$

        $C' \leftarrow C$, $c_{ln} = c_{ln} + 1$, Calculate f_{ln}, $C \leftarrow C'$,

    end for

    $\left[f_{l^*n^*}\right] = \arg\min\left(f_{ln}\right)$,

    if $f_{l^*n^*} < f_0$ **then** $f_0 \leftarrow f_{l^*n^*}$, $c_{l^*n^*} = c_{l^*n^*} + 1$

    end if

end while

Output: optimized solution C^*

5.2 Improved Model 2 Solution

Solve the following improvement model 2 here. The solution is initialized according to constraints (45), (46), and (47) of the improved model 2, and constraint (42) is transformed into the solution process. At this time, the original optimization model is transformed into constraints (43) and (44). We also solve the optimization model in three layers. The vector, the matrix, and the vector are still initialized according to the initialization method in the improved model 1 of this paper. Constraint (43) is used to limit the elements in the vector to facilitate the solution of the first layer and the second layer.

The first layer solves matrix A. The current optimization problem is first solved as if the number of users does not exceed the load capacity of the network. That is, it ignores constraint (44) and sets ω to 0. The solution process is shown in the solution algorithm 1 of

the improved algorithm 1 in this paper.

The second layer solves matrix B. Based on the solution of the first layer, the user channel allocation to the SCBS is further optimized. According to constraint (43), the search is performed on the free channels in each cell in turn. The solution process is as described in Solving Algorithm 2 in Improved Algorithm 1 herein.

The third layer solves vector C. We must focus on the load capacity of the network. Based on the solution of the first and second layers, $\left(N - M_2\right)$ iterations are performed. Each iteration rejects the user with the longest delay in the current optimization result until $N = M_2$. The solution process is shown in Algorithm 4.

Algorithm 4. Algorithm for Solving C Matrix

Input: A^*, B^*, C, f_0, L, N_l, M_1, M_2, ω, w_{ln}, r_1, RT, PL, H, P_{UE}, P_{SC}, d_{ln}^{UB}, d_{ln}^{US}, d_l^{SB}

while $\sum_{l=1}^{L} \sum_{n=1}^{N_l} a_{ln} + \sum_{l=1}^{L} \sum_{n=1}^{N_l} \sum_{m=1}^{M_1} b_{ln,m} > M_2$ **do**

    for $a_{ln} \in A$

        if $a_{ln} = 1$ **then**

      $a_{ln} \leftarrow 0$, **Algorithm 1**, **Algorithm 2**,

      Calculate f_{ln}, $a_{ln} \leftarrow 1$

        else for $b_{ln,m} \in B$

        if $\exists b_{ln,m'} = 1$ **then**

      $b_{ln,m'} \leftarrow 0$, **Algorithm 1**, **Algorithm 2**,

      Calculate f_{ln}, $b_{ln,m'} \leftarrow 1$

          end if

          end for

          end if

    end for

    $\left[f_{l^*n^*}\right] = \arg\min\left(f_{ln}\right)$, $a_{l^*n^*} \leftarrow 0$, $b_{l^*n^*,1} \leftarrow 0$, $b_{l^*n^*,2} \leftarrow 0$, ..., $b_{l^*n^*,M_1} \leftarrow 0$,

end while

for $a_{ln} \in A$

    if $\left(a_{ln} = 0\right) \&\& \left(\sum_{m=1}^{M_1} b_{ln,m} = 0\right)$ **then** $c_{ln} \leftarrow 0$

    end if

end for

Output: optimized solution C^*

---

6 Simulations

In this section, in order to verify the effectiveness of the proposed optimization algorithm, MATLAB simulation software is used to evaluate the performance of the algorithm. In the $500m \times 500m$ area, the BAN is set at the centre, and the four SCBSs

are evenly distributed. The average communication radius of the SCBS is 175 m in order to achieve seamless coverage of the entire area. N users are evenly distributed and have no mobility (the topology is fixed). Each user belongs to one cell. Each user accesses the same service category. Therefore, the weighted value is $w_{ln} = 1(\forall l, n)$. Each uplink channel is simulated as a Rayleigh fading channel. In addition, the number of retransmissions has a great influence on our optimization results. In theory, the greater the number of retransmissions, the greater the average transmission delay, but the smaller the probability of packet loss events at the receiving end. Conversely, the smaller the number of retransmissions, the smaller the average transmission delay, but the larger the packet loss at the receiving end, and the greater the possibility of an incident. In the actual system, we define the number of retransmissions RT as 5 times to achieve a compromise between the two. Other simulation parameters are shown in Table 1.

Table 1. Simulation parameters

Parameter	Description	Value
α	Free space transmission coefficient	3
n_0^{BAN}	Thermal noise of BAN	-174 dBm/Hz
n_0^{SCBS}	SCBS thermal noise	-174 dBm/Hz
Δ_{BAN}	BAN allocated channel bandwidth	20 MHz
Δ_{SC}	SCBS allocated channel bandwidth	20 MHz
M_2	Number of channels that the BAN can allocate	250
P_{UE}	User's transmission power	7 dBm
P_{SC}	SCBS transmission power	27 dBm
H	Packet sending rate	100 Mbit/s
PL	The size of the package	65536 bit
r	Initial delay compensation factor	1.1
ω	Regulatory factor	1000

To verify the effectiveness of the improved model 1 based solved algorithm (IM1SA) and the improved model 2 based solved algorithm (IM2SA) presented in this paper, the proposed algorithms are compared. Based on the 5G two-tier heterogeneous network environment, the following three types of comparison algorithms are constructed. The first one is a wireless backhaul optimization algorithm for Single Backhaul Scenarios (WBOASBS) [16]. This algorithm is intended for a single backhaul scenario. Only the backhaul mode of accessing the BAN through the SCBS is supported. Correspondingly, a comprehensive delay index considering channel dynamics is established to optimize the channel allocation. Second is a wireless backhaul optimization algorithm for a static channel scenario WBOASCS (Wireless Backhaul Optimization Algorithm). For Static Channel Scenarios [17], this algorithm supports two types of backhaul approaches for hybrid backhaul scenarios. Correspondingly, it builds an integrated delay index

that does not consider channel dynamics to optimize the channel allocation. The third is based on the initial delay. The Wireless Backhaul Optimization Algorithm Based on Initial Delay (WBOABID) is a wireless backhaul optimization algorithm (WBOABID) [14]. This algorithm is designed for hybrid backhaul scenarios and supports two types of backhaul methods. The initial delay indicators that take the channel dynamics into account are optimized accordingly (i.e., the channel allocation). The rest of the above algorithm is similar to this algorithm. The above algorithm is compared using the two performance indicators of average transmission delay and average throughput. Among them, the average transmission delay calculates the channel allocation result based on different algorithms, substitutes the actual network environment to count the time delay of all the users' services, and then takes the average of the number. The average throughput calculates the channel allocation result based on different algorithms. By substituting in the actual network environment to count the uplink transmission throughput of all users, the corresponding average number is available.

To prove the effectiveness of the IM1SA proposed in this paper, this algorithm is compared with the arrival delays of the above three algorithms. Figure 4 shows the comparison results of the average transmission delay between the IM1SA and the other three algorithms under the condition of the low traffic load situation (that is, when the number of users is less than M_2). It can be seen that four algorithms are used under different M_1 values. The overall delay increases as the number of users M_1 increases. When the number of users is small, regardless of whether the M_1 value is 100 or 150, the interference between users is small. Therefore, the transmission delays of all algorithms are equal. At a relatively low level, the delays of various algorithms under different M_1s are relatively close. If the number of given users is high, the interference will be significantly reduced with the increase of M_1. As M_1 increases, the transmission delay of various algorithms also decreases. In addition, it can be seen from the figure that, in any case, the transmission delay performance of the IM1SA is always better.

To further verify the performance of the IM1SA, the above four types of algorithms are compared from the perspective of throughput. Here, throughput represents the amount of data successfully transmitted in a unit of time. Figure 5 shows the comparison results of the average throughput of the IM1SA and the other three algorithms under the low traffic load situation. It can be seen that under different M_1 values, the average throughput of the four algorithms follows the number of users N when there is an increase or decrease. When a given number of users is small, no matter
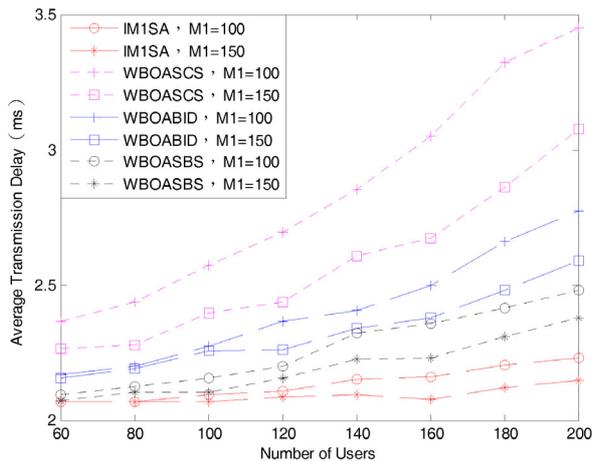
Figure 4. Comparison of average transmission delays of different algorithms (Low traffic load situation)

whether the M_1 value is 100 or 150, the interference between users is small at this time. Therefore, the average throughput of various algorithms is at a relatively high level, and thus it is different. The throughputs of various algorithms under M_1 are relatively close. If the number of given users is high, the interference will be significantly reduced with the increase of M_1. Therefore, with the increase of M_1, the average throughput of various algorithms will increase. In addition, as seen from the figure, the average throughput performance of the IM1SA is always better.

Figure 4 and Figure 5 show the simulation results of the four types of algorithms under low traffic loads (that is, the number of users is less than M_2). With the increase of the number of users, once M_2 is exceeded, the above four types of algorithms will face the situation of no solution, which means that the above algorithm cannot be applied to the situation that the number of users is overloaded. Therefore, this paper proposes the IM2SA.
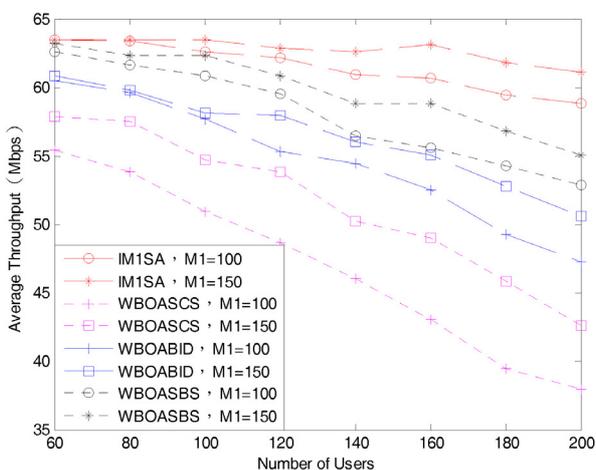
Figure 5. Comparison of average throughput of different algorithms (Low traffic load situation)

Figure 6 and Figure 7 respectively show the average delay and average throughput of the IM2SA with the number of users under high traffic load conditions (that is, the number of users is greater than M_2). Because the IM2SA will reject access to some users, the transmission delay of the solid rejected user will be averaged according to the upper limit of the integrated delay, and the corresponding average transmission delay indicator will be obtained. The throughput of the rejected user will be taken as 0 to calculate the statistics of the mean average throughput indicator. From Figure 6 and Figure 7, it can be seen that as the number of users increases, the average transmission delay of the IM2SA increases and the average throughput of the IM2SA decreases. In the case of a given number of users, with the increase of M_1, the average transmission delay of the IM2SA also decreases, and its average throughput increases. Obviously, the IM2SA can effectively accept control users and reasonably allocate channel resources. With the rapid increase in the number of users, the network performance (average delay and average throughput) deteriorate only slightly.
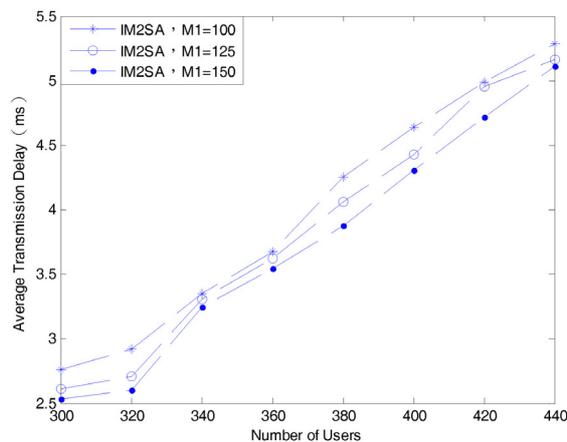
Figure 6. Average delay of IM2SA varies with the number of users under different M_1s
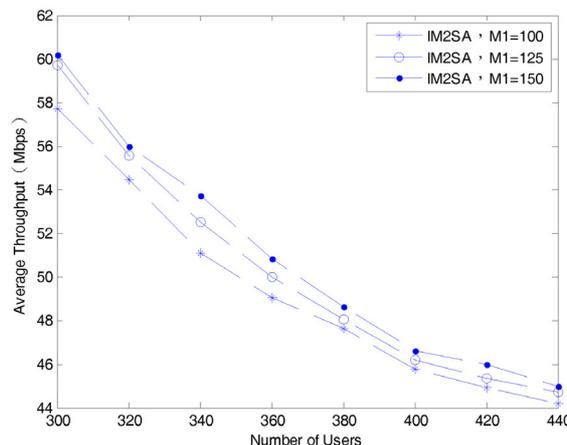
Figure 7. Average throughput of IM2SA varies with the number of users under different M_1s

7 Conclusion

This paper proposes a wireless backhaul optimization algorithm that considers the time delay for 5G dynamic heterogeneous scenarios. Based on the systematic analysis of the time delay problem under this scenario, an accurate return optimization index is established and a basic backhaul model is constructed. Furthermore, from the two perspectives of the user demand difference and the number of overloaded users, the improved model 1 and the improved model 2 are respectively constructed, and a hierarchical algorithm is constructed to quickly solve the problem. Simulations verify the effectiveness of the proposed algorithm.

Acknowledgments

References

[1] D. Soldani, A. Manzalini, Horizon 2020 and Beyond: On the 5G Operating System for a True Digital Society, *IEEE Vehicular Technology Magazine*, Vol. 10, No. 1, pp. 32-42, March, 2015.

[2] S. Chia, M. Gasparroni, P. Brick, The Next Challenge for Cellular Networks: Backhaul, *IEEE Microwave Magazine*, Vol. 10, No. 5, pp. 54-66, August, 2009.

[3] O. Tipmongkolsilp, S. Zaghloul, A. Jukan, The Evolution of Cellular Backhaul Technologies: Current Issues and Future Trends, *IEEE Communications Surveys & Tutorials*, Vol. 13, No. 1, pp. 97-113, May, 2011.

[4] K.-C. Ting, A Comparison Model for WLAN Technologies, 802.11n and HeNB in LTE and the Future 5G Networks, *Journal of Internet Technology*, Vol. 19, No. 1, pp. 257-268, January, 2018.

[5] S. Rangan, T. S. Rappaport, E. Erkip, Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges, *Proceedings of the IEEE*, Vol. 102, No. 3, pp. 366-385, March, 2014.

[6] M. Jamalzadeh, L.-D. Ong, M. N. B. M. Nor, 5G Technologies: A New Network Architectures and Design, *Journal of Internet Technology*, Vol. 19, No. 7, pp. 1983-1991, December, 2018.

[7] J. Kim, A. F. Molisch, Quality-Aware Millimeter-Wave Device-to-Device Multi-Hop Routing for 5G Cellular Networks, *IEEE International Conference on Communications (ICC)*, Sydney, Australia, 2014, pp. 5251-5256.

[8] A. F. Ribadeneira, *An Analysis of the MOS under Conditions of Delay, Jitter and Packet Loss and an Analysis of the Impact of Introducing Piggybacking and Reed Solomon FEC for VOIP*, Master Thesis, Georgia State University, Atlanta, USA, 2007.

[9] M. N. Islam, A. Sampath, A. Maharshi, O. Koymen, N. B. Mandayam, Wireless Backhaul Node Placement for Small Cell Networks, *48th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, USA, 2014, pp. 1-6.

[10] M. Shariat, E. Pateromichelakis, A. U. Quddus, R. Tafazolli, Joint TDD Backhaul and Access Optimization in Dense Small-Cell Networks, *IEEE Transactions on Vehicular Technology*, Vol. 64, No. 11, pp. 5288-5299, November, 2015.

[11] X. X. Xu, W. Saad, X. J. Zhang, X. B. Xu, S. D. Zhou, Joint Deployment of Small Cells and Wireless Backhaul Links in Next-Generation Networks, *IEEE Communications Letters*, Vol. 19, No. 12, pp. 2250-2253, December, 2015.

[12] Y. Chen, X. Wen, Z. Lu, H. Shao, J. Cheng, Cooperative Base Station Operation for Green Dense Heterogeneous Networks, *IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Bandung, Indonesia, 2015, pp. 31-35.

[13] S. Han, C. Yang, A. F. Molisch, Spectrum and Energy Efficient Cooperative Base Station Doze, *IEEE Journal on Selected Areas in Communications*, Vol. 32, No. 2, pp. 285-296, February, 2014.

[14] D. C. Chen, T. Q. S. Quek, M. Kountouris, Wireless Backhaul in Small Cell Networks: Modelling and Analysis, *IEEE 79th Vehicular Technology Conference (VTC Spring)*, Seoul, South Korea, 2014, pp. 1-6.

[15] S. L. Song, On Distribution of Sums of n Independent Random Variables Subject to Exponential Distribution, *Journal of Liaoning Normal University (Natural Science Edition)*, No. 4, pp. 51-58, December, 1990.

[16] W. C. Xia, J. Zhang, S. Jin, H. B. Zhu, Delay-Based User Association in Heterogeneous Networks with Backhaul, *China communications*, Vol. 14, No. 10, pp. 130-141, October, 2017.

[17] H. X. Wang, R. P. Li, L. Fan, H. G. Zhang, Joint Computation Offloading and Data Caching with Delay Optimization in Mobile-Edge Computing Systems, *9th International Conference on Wireless Communications and Signal Processing (WCSP)*, Nanjing, China, 2017, pp. 1-6.

Biographies
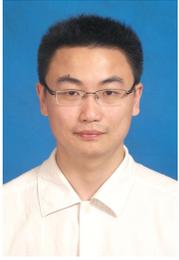
Hui Zhang received his Ph.D. degree in Information and Communication Engineering from Nanjing University of Posts and Telecommunications (NJUPT), China, in 2008. He is a Professor and Master Supervisor in the Jiangsu Provincial Key Lab of Wireless Communications, NJUPT. His research direction is future wireless network.

Fengle Li received his bachelor degree in Information and Computing Science from Ludong University, China, in 2014. He is working towards his master degree at College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, China. His research direction is wireless backhaul technology.