Study of cell clustering optimization algorithm for load balancing in 5G scenario

Yuhan Wang

1 School of Electronic Science and Engineering, Xiamen University, Xiamen, Fujian, 361005, China

Abstract. In the context of 5G networks at this stage, massive data transmission and highly dense deployment of base stations make severe interference and huge load volumes an inevitable new challenge. The traditional interference coordinated cell-clustering scheme only considers the inter-micro base station path loss without jointly considering load balancing, an important factor affecting system performance. This paper first introduces the existing cell-clustering schemes, and then proposes a cell-clustering optimization algorithm that combines path loss and load balancing on the basis of 5G cellular network architecture, and then simulate the optimization algorithm and compare the results with the traditional schemes to verify its ability to balance the load, increase the amount of transceiver data, and optimize the resource allocation and performance effects.

1. Introduction

Since 2018, the fifth-generation communication technology has gradually developed into the latest generation of cellular mobile communication technology. In order to achieve the 1000x increase in data flow density and 10-100x increase in the number of devices that 5G networks are intended to achieve, one of the most effective ways is through ultra-dense network technology. However, in densely deployed base station networks, severe interference and huge load volumes become important issues. Therefore, the reasonable allocation of macro base station resources for micro base stations based on clustering centralized control technology in 5G macro-micro base station scenario to achieve load balancing has become one of the important ways to effectively improve the system performance [1].

At this stage, relevant researchers at home and abroad have explored the problems of interference coordination and load balancing in mobile communication networks [2-3].

Gong B M’s work [2] has great reference value for studying the path loss in 5G macro-micro scenario, which proposed an interference coordination scheme based on subdivision clustering in dynamic TDD LTE systems, aiming to eliminate serial interference in LTE systems configured with dynamic TDD technology. However, this scheme only considered the path loss and ignored the performance requirements of load balancing, and did not make a clear definition of the applicable scenarios. Zhu X R et al [3] provided a reference value in the selection of performance check parameters and dynamic clustering method. However, this scheme only set the interference level as a dynamic clustering parameter and did not consider the intra-cluster load.

To address the situation that the system performance may be degraded due to overload and uneven resource allocation in a certain cluster, this paper proposes a resource allocation optimization algorithm based on the 5G cellular network architecture with low path loss and load balancing in order to optimize resource allocation, improve throughput, and reduce interference.
2. Load Balancing Optimization Algorithm Based on Cell-clustering

2.1. Scene architecture
This paper focuses on the interference in the downlink of a cellular network composed of macro and micro base stations in 5G scenario. In a cellular network, there are M macro base stations represented by blue dots, and the coverage area of macro base stations is distributed in a cellular shape; there are N micro base stations represented by black dots, and the micro base stations are scattered randomly in the coverage area of macro base stations. The structure schematic is shown in Figure 1.

![Figure 1. Structure of cellular network composed of macro base stations and micro base stations.](image)

2.2. Algorithm description
The traditional small area clustering scheme usually only considers the path loss between macro and micro base stations. In this paper, we propose a new scheme that jointly considers the path loss and the load situation of macro base stations, and an optimization algorithm that jointly equalizes the load on the basis of considering the path loss.

In the above network architecture, let the coordinates of the i-th micro base station be \((p_x_i, p_y_i), i \in [1, N]\) and the coordinates of the j-th macro base station be \((m_x_j, m_y_j), j \in [1, M]\).

Then the distance between the i-th micro base station and the j-th macro base station is shown in equation (1).

\[
R_{ij} = \sqrt{(p_x_i - m_x_j)^2 + (p_y_i - m_y_j)^2}
\]  (1)

The path loss is measured using macro and micro base station coordinate distance, and the calculation formula is shown in equation (2).

\[
PL_{ij} = 10^{3.8 + 20.9 \log_{10} R_{ij}}
\]  (2)

Set each micro base station with a flag whether it is at the edge or not, as shown in equation (3).

\[
Flag_i = \begin{cases} 
0 & \text{when the } i \text{-th micro base station is at the edge of the cell} \\
1 & \text{when the } i \text{-th micro base station is not at the edge of the cell}
\end{cases}
\]  (3)

Set the cluster head load threshold, and calculate the formula as shown in equation (4).

\[
\text{average} = \frac{N}{M}
\]  (4)
Let the number of micro base stations in the j-th macro base station cluster be mj, and the weighting value is set in the load balancing optimization process as shown in equation (5).

\[ G_{ij} = 0.8 \times PL_{ij} + 0.2 \times m_j \]  

(5)

2.2.1. Micro base station clustering scheme based on path loss
The algorithm is shown in Figure 2.
1. The number of micro base stations in each cluster is initialized, which is generally 0.
2. The i-th micro base station \((0 \leq i < N)\) is selected, and the path loss \(PL_{ij} (0 \leq j < M)\) with each macro base station is calculated cyclically according to equation (1) and (2).
3. The smallest \(PL_{ij}\) corresponding to the macro base station number \(j\) is selected and saved as \(ID = j\).
4. Cluster the i-th micro base station into the ID cluster.
5. Continue to select the next micro base station until all micro base stations are clustered.
The micro base station clustering based on path loss is completed.

2.2.2. Micro base station clustering optimization algorithm based on balanced load
In this paper, we propose an optimization algorithm that jointly considers path loss and load balancing and add the optimization part of balanced load after the micro base station clustering scheme based on path loss algorithm. This part of the optimization algorithm is shown in Figure 3. Here we assume that the data traffic within each micro base station is equal.
1. The flag \(Flag_i \in \{0,1\}\) of whether each micro base station is at the edge is set according to equation (3).
2. Set the cluster head load threshold average according to equation (4).
3. The number of micro base stations \(m_j\) in each macro base station cluster is detected, and the macro base stations whose number of micro base stations in the cluster exceeds the threshold \(average\) after the first clustering are selected, and their cluster head numbers are put into an array \(macroarray\) for subsequent optimization adjustment.
4. The k-th number of macro base stations \((0 \leq k < length(macroarray))\) in the \(macroarray\) is selected, and its number is \(n = macroarray[k]\), and the numbers of all micro base stations in the cluster are put into the \(picoarray\).
5. The first micro base station with \(Flag_i = 1\) is retrieved one by one from the \(picoarray\) and numbered \(i\). The weighted value \(G_{ij}\) \((m_j < average; j \neq n)\) of the j-th macro base station is calculated according to equation (5).
6. The smallest \(G_{ij}\) corresponding to the macro base station number \(j\) is selected and saved as \(ID = j\).
7. Cluster the i-th micro base station into the i-th cluster of \(ID\).
8. Calculate the number of micro base stations \(m_n\) within the n-th macro base station cluster again, and if the threshold \(average\) is still exceeded, i.e., \(m_n > average\), repeat steps 5~7 of the cycle. If \(m_n \leq average\), the internal load optimization of the cluster numbered by the next element in the array \(macroarray\) is performed (repeat steps 4~7) until the micro base stations in the cluster numbered by all elements in the array \(macroarray\) are re-clustered and the optimization process is completed.
Figure 2. Algorithm flow of path loss based clustering scheme.

Figure 3. Optimization algorithm flow with joint consideration of path loss and load balancing.
3. Experiment and Result Analysis

In this paper, two sets of simulation experiments are used to verify the performance of the above algorithm for comparison, and the simulation scenario used is referred to Figure 1, and the specific simulation parameters are described in Section 3.1. In the first experiment, the number of micro base stations is adjusted, the distribution of micro base stations and the bd (see equation (5) in 3.2) are obtained before and after the optimization, as well as the bd decrease percentage to verify that the optimization scheme can achieve the purpose of load balancing; in the second experiment, the UE and packet sending and receiving functions are added to test the optimization, and the number of UEs is adjusted during the test to obtain multiple sets of received data and performance improvement percentages of terminals before and after optimization to verify that the optimized solution can achieve the purpose of improving system performance.

NS3 is an open-source discrete-event-based network simulation simulator, and the entire project source code is available for free on the NS3 website. Layer, application layer, transport layer, data link layer, and other multi-layer protocols and applications in NS3, making it suitable for simulation experiments on small-distinct cluster-based resource allocation optimization algorithms in 5G scenarios [4].

3.1. Simulation scenarios and simulation parameters

The simulation scenario of the first group of experiments is shown in Figure 1, and the simulation parameters are shown in Table 1.

| Parameter                          | Value                  |
|------------------------------------|------------------------|
| Number of macro base stations      | 7                      |
| Number of micro base stations      | 40, 80                 |
| Macro base station coverage radius | 200                    |
| Macro base station spacing         | 400                    |
| Number of UEs in micro base station| 5                      |
| Total number of UEs               | 200                    |

To verify the validity of the performance test, in the second set of experiments, the data transceiver simulation scenario is simplified by using a scenario with two macro base stations and six micro base stations, and the simulation parameters are shown in Table 2.

| Parameter                          | Value                  |
|------------------------------------|------------------------|
| Number of macro base stations      | 2                      |
| Macro base station coordinates     | (0,0) and (400,0)      |
| Number of Micro Base Stations      | 6                      |
| Micro base station coordinates     | (-75,100),(-75,-100),(200,100),(200,-100),(475,100),(475,-100) |
| Macro base station Coverage Radius | 200                    |
| Macro base station spacing         | 400                    |
| Number of UEs in micro base station| 5, 10, 15              |
| Total number of UEs               | 30, 60, 90             |
| Single packet data volume         | 1024Bytes              |
3.2. Simulation Performance Metrics

Two performance metrics are taken in the simulation of this paper, which are the balance degree $bd$ and the amount of data received at the terminal.

The concept of balance degree is proposed here, and the variance is used to characterize the balance degree. Balance degree is defined in equation (6). $bd$ is used to measure the degree of load balancing of the clustering scheme, and the smaller the $bd$, the better the load balancing of the clustering result.

$$bd = \frac{\sum_{i=1}^{M}(m_i-average)^2}{M}$$

In addition, the size of the data received by the terminal is used to measure the performance of the clustering scheme, which is obtained experimentally by multiplying the number of outgoing packets by the size of each packet. The larger the amount of data received by the terminal, the better the performance of the clustering scheme.

3.3. Analysis of Simulation Results

3.3.1. Clustering results and equilibrium

In the simulation scenario of Experiment 1 ($M=7$), the simulation experiments are conducted with 80 and 90 micro base stations ($N=80$ and 90), respectively, marking Scheme 1 as the traditional scheme considering only path loss (algorithm of 2.2.1) and Scheme 2 as the proposed new optimized scheme considering path loss and load balancing (combining algorithm 2.2.1 and 2.2.2).

Figure 4 gives the distribution of results for different numbers of micro base stations ($N=80$ and $N=90$) clustered using scheme I and scheme II, respectively.

![Figure 4.a) The clustering results of scheme I at $N=40$.](image)

![Figure 4.b) The clustering results of scheme II at $N=40$.](image)

![Figure 4.c) The clustering result of scheme I at $N=80$.](image)

![Figure 4.d) The clustering result of scheme II at $N=80$.](image)
From Figure 4 a) b) and Figure 1 c) and d), we can see that the clustering results are different before and after the optimization for two different numbers of micro base stations. The optimized clustering results show a more balanced load on the macro base stations, as the lowest path loss is no longer the only basis for clustering, but the load on the macro base stations is considered jointly.

Tables 3 and 4 give the balance degree $bd$ (calculated according to equation (6)) for different number of micro base stations ($N=80$ and $N=90$) clustered using scheme I and scheme II, respectively. Three random distributions of micro base stations represented by three random numbers were taken in the experiments.

| Table 3. Balance degree $bd$ when $N=40$. |
|-----------------------------------------|
| Random number 1 | Random number 2 | Random number 3 |
| Scenario I $bd$ | 1.28571 | 4.42857 | 3.57143 |
| Scenario II $bd$ | 0.714286 | 1.28571 | 1.85714 |
| $bd$ decrease percentage | 44.4442% | 70.9678% | 48.0001% |

| Table 4. Balance degree $bd$ when $N=80$ |
|-----------------------------------------|
| Random number 1 | Random number 2 | Random number 3 |
| Scenario I $bd$ | 3.85714 | 11.2857 | 9.57143 |
| Scenario II $bd$ | 1.28571 | 2.42857 | 5.28571 |
| $bd$ decrease percentage | 66.6668% | 78.4810% | 44.7762% |

From Tables 3 and 4, we can see that in all three cases of random assignment of micro base station coordinates, the $bd$ of scenario 2 is significantly smaller than that of scenario 1, and the $bd$ after optimization decreases by a certain percentage. This indicates that the number of micro base stations within each macro base station cluster is more evenly distributed after adding the optimization algorithm that jointly considers load balancing, and the optimized clustering results show a more balanced load situation.

3.3.2. Amount of data received at the terminal
The simulation experiments are conducted in the data sending and receiving test scenario ($M=2$, $N=6$), and marking scheme I as the traditional scheme considering only path loss (algorithm of 2.2.1) and scheme II as the proposed new optimized scheme considering both path loss and load balancing (combining algorithm 2.2.1 and 2.2.2).

Table 5 gives the data sending and receiving results after clustering with different number of UEs using scheme I and scheme II, respectively, to obtain multiple comparisons of the amount of data received by terminals before and after optimization, as well as the performance improvement percentage.

| Table 5. Amount of data received by the terminal for the two schemes |
|---------------------------------------------------------------|
| The amount of data received by the terminal of Scenario 1 |
| 30 UEs | 60 UEs | 90 UEs |
| 1024×24 Bytes | 1024×47 Bytes | 1024×73 Bytes |
| Amount of data received by the terminal in Scenario 2 |
| 1024×27 Bytes | 1024×51 Bytes | 1024×78 Bytes |
| Performance Improvement Percentage | 12.5% | 8.51% | 6.85% |
It can be seen from Table 5 that in the data transceiver test scenario, the amount of data received by the terminals is different before and after the optimization of the three UE number cases. The amount of data received by the terminals in Scheme 2 is obviously larger than that in Scheme 1, and the amount of data received by the terminals after optimization increases by a certain percentage compared with that before optimization. This indicates that the clustering of the micro base station is more reasonable after the optimization algorithm of joint load balancing is added, which improves the performance of sending and receiving data. The optimized clustering results show the ability to send and receive larger amounts of data.

4. Conclusion

In this paper, an optimization algorithm is proposed for the 5G macro-micro scenario based on the combined consideration of path loss and load balancing for cell-clustering, by analyzing the system performance degradation due to overload and uneven resource allocation in a certain cluster. Through simulations in a multi-cell cellular network scenario, it is concluded that the new scheme based on the joint consideration of path loss and load balancing of subdivision clusters can effectively balance the load of macro base stations and improve the data reception compared with the traditional scheme that only considers path loss. In the 5G cellular network scenario with large-scale deployment of micro base stations, the application of the load balancing scheme based on cell-clustering is a promising development direction, and its future development will focus on optimizing the resource allocation of macro base stations to further reduce interference and improve throughput.

References:
[1] Zhang, J.M., Xie, W.L., Yang, F.Y. (2016) Network architecture and implementation of 5G super-dense networking. Telecommunications Science, 32(06):36-43.
[2] Gong, B.M., Liu, X.F., Gao, Y.H., et al. (2013) Research on interference coordination scheme based on small differentiated cluster in dynamic TDD LTE system. Modern Telecommunication Technology, 43(08):31-35.
[3] Zhu, X.R., Zhu, W.R. (2016) Small differential cluster and power allocation algorithm based on interference coordination in ultra-dense small peak and nest networks. Journal of Electronics and Information Technology, 38(05):1173-1178.
[4] NS-3 project. (2019) NS-3 manual. https://www.nsnam.org/docs/manual/html/index.html