

Joint Optimization for Visual Data Transmission in the Resource Constraint 5G IoT

Jingce Xu^{1*}, Xinsheng Zhang¹ and Jianfei Lu¹

¹Management Consulting Institute, State Grid Energy Research Institute, Beijing, China

Abstract. With the integration of 5G and Internet of Things (IoT), the application of large-scale IoT devices networking is increasingly extensive, such as building management, property maintenance, autonomous vehicles, healthcare, and shopping to tourism. In these scenes, the volume of data transmission is quite large, especially visual data (image, video, et al). However, due to the limited resource of IoT devices, such as battery, power, bandwidth, visual data transmission is complex to optimize, single objective optimization is difficult to insure the optimal latency, throughput and power at the same time. In this paper, we propose a method to jointly optimize the resource allocation of visual data transmission in resource constraint 5G IoT. Instead of single objective optimization, we combine the bandwidth, power consumption and latency into a hybrid model, then propose a low-complexity algorithm to solve this multiple objective optimization problem. The simulation results demonstrate that the proposed method increases the comprehensive utility of visual data transmission in the resource constraint 5G-IoT by 35%-48% compared with existing approaches.

1 Introduction

In recent years, with the progress of the standardization and utilization of 5G, the application of 5G has been greatly developed. International Telecommunication Union (ITU) reports that the global 5G connection will reach 1.1 billion in 2020, 12% of global wireless network connections[1]. Driven by 5G, IoT has entered a new stage of rapid development, the number of IoT devices will overcome 50 billion by the end of 2020[2]. 5G and IoT are promoting the development of smart city, healthcare, smart building and other emerging applications, but the expansion of these applications makes the scale and variety of IoT in mobile networks continue to increase, challenging the data transmission in 5G IoT. The challenges are elaborated as follows.

The large volume of data generated by IoT devices, especially visual data. Data is the core of all smart application. However, in the current mode, data collection and data processing are separated, sensors collect data and transmit to data center, then the data center analyze data to get result. This leads to a large number of data need to be transmitted through the

channel between IoT devices and data center. In 2020, the data generated by IoT devices reaches to 4.4ZB, of which 80% is visual data[3]. This is because visual sensors are commonly embedded in IoT devices, and therefore they can be widely used in various applications, such as mobile surveillance, entertainment, and UAVs (Unmanned Aerial Vehicles).

Heterogeneity and resource scarcity of IoT devices.

The heterogeneity of IoT devices is particularly reflected in the differences among performance of devices. As TABLE 1 shows, various types of devices have great differences in calculation speed, storage capacity, connection bandwidth, battery and so on. As a result, large-scale visual data transmission between different devices and data centers presents severe compatibility problems. On the other side, compared with the cable power supply equipment such as personal computer, the available resources of IoT devices are scarce. This greatly affects the working time of the IoT devices, and its continuous working time is often less than 10 hours, so the considerable resource allocation is important and necessary. It implies that the heterogeneity and resource scarcity of IoT devices pose a natural challenge for visual transmission over IoT devices.

Table 1. The performance criteria of different IoT devices

Device Type	CPU Speed	Available Bandwidth	Memory	Power Consumption	Battery
PC	2-4GHz	≤ 1Gbps	≤ 64GB	300W	Cable Power
Intelligent Watch	0.3-1.2GHz	512Kbps	≤ 1GB	≤ 100mW	≈ 1Wh
Cellphone	≤ 2.8GHz	≤ 867Mbps	≤ 8 GB	≤ 10W	7-16 Wh

*xujingce@sgeri.sgcc.com.cn

UGV	1-2GHz	1Mbps	-	≈ 30W	≈ 48Wh
UAV	-	14.4Mbps	-	≤ 600W	≈ 97Wh
Surveillance Camera	≤ 1GHz	100Mbps	-	48W	Solar Energy

The development of scenarios require higher trans-mission performance. With the development of intelligent applications, the demand for high throughput, low latency and high quality visual data transmission is becoming more and more intense. For example, in smart building scenario, the latency of data transmission needs to be less than 100 milliseconds to ensure that users have a good experience. But as the resource of IoT devices is always constraint, It is difficult to guarantee high latency, high throughput and low power consumption at the same time. This brings challenges to the optimization of visual data transmission in 5G IoT.

To address the challenges mentioned above, we propose a method to jointly optimize the resource allocation of visual data transmission in resource constraint 5G IoT. By con-sidering throughput, latency and power consumption at the same time, the comprehensive utility of visual data trans-mission in 5G IoT is optimized. The contributions of this paper are listed as follows:

- 1) we explore the factors that affect the transmission performance and the relationship between them. Then, we formulate the hybrid optimization model for 5G IoT visual data transmission system, by considering throughput, latency and power consumption of the IoT devices.
- 2) Due to the high complexity when directly solve the proposed model, we transforms and decomposes the model and propose an algorithm with low complexity to get the optimum from the hybrid optimization model efficiently.
- 3) we conduct experiments in real scenario dataset to test and verify proposed method. The results

demonstrate the proposed method increase the comprehensive utility of visual data transmission in in the resource constraint 5G IoT in 35%-48% compared with existing approaches.

The rest of this paper is organized as follows. In Section 2, we give the system model and problem formulation. Then the optimal resource allocation strategy and low-complexity solution algorithm is presented in Section 3. In Section 4, we present and discuss the performance evaluation results. Finally, this study is concluded in Section 5.

2 System Model and Problem Formulation

2.1 Overview of Visual Data Transmission System in 5G IoT

In this paper, we consider the visual data transmission system shown in Fig. 1. The transmission system is consist of a 5G station and multiple IoT devices with heterogeneous performance. We assume that the 5G station can recognize the throughput, power consumption and transmission latency of IoT devices, and allocate bandwidth for them. However, the existing methods for bandwidth allocation are incapable in 5G IoT transmission system, the reason is the confliction among power consumption, bandwidth and latency. The IoT devices transmit data through wireless network, higher bandwidth means higher power consumption. Otherwise, the lower power consumption leads to higher latency.

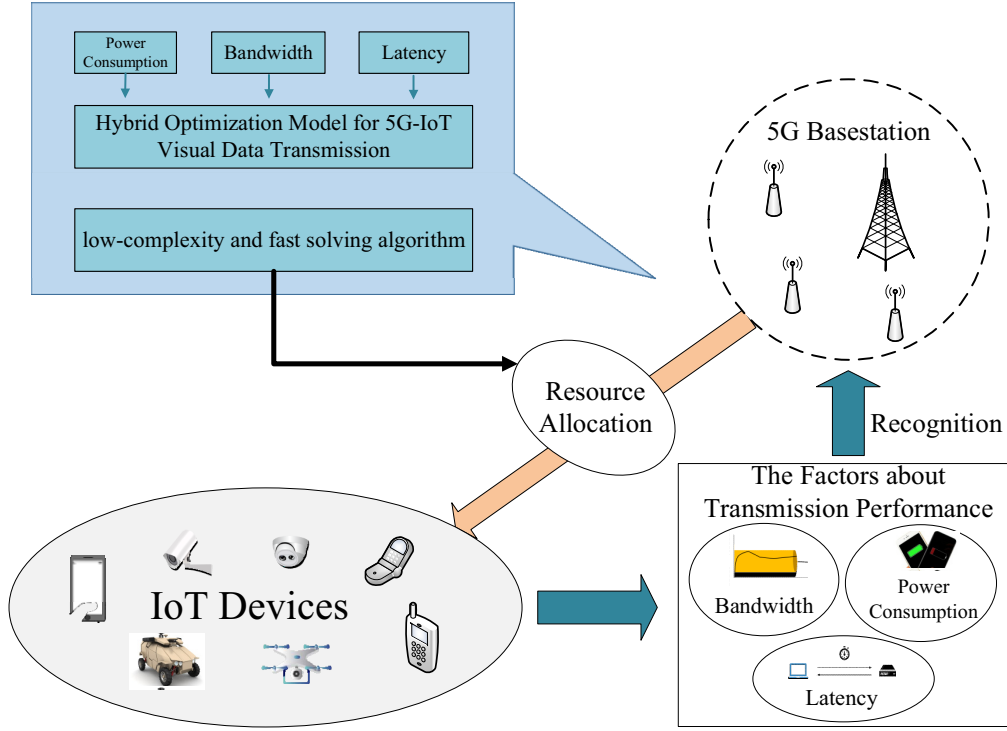


Figure 1. Overview of the 5G IoT visual data transmission network.

As Fig. 1 shows, instead of single objective optimization for power consumption, bandwidth or latency, we comprehensively consider bandwidth, power consumption and transmission latency, and use mixed indicators to establish a hybrid optimization model. By optimizing the model, the resource allocation strategy that can ensure optimal system efficiency is obtained.

2.2 System Model

In order to comprehensively consider the impact of the power consumption, bandwidth and transmission delay of the 5G IoT visual data transmission system on the overall transmission performance, we use the mixed utility to measure the performance. We use the comprehensive utility q as the definition in [4], as follows:

$$q = \frac{U_R(B)}{U_C(B)} \quad (1)$$

where B is a vector with the dimension of n , representing the bandwidth allocated to devices in the system. $U_R(B)$ is the weighted total bandwidth, $U_C(B)$ is the cost of data transmission. For unifying the unit, we normalize the bandwidth, latency and power consumption of each device in the system. Let w be the weight of devices in the system, $U_R(B)$ is represented as follow:

$$U_R(B) = \frac{\sum_{i=1}^n w_i B_i}{B_{max}} \quad (2)$$

where B_{max} is the maximum system bandwidth limit. the system transmission cost is as follows:

$$U_C(B) = \sum_{i=1}^n w_i \left(\frac{L_i}{L_{max}} + \frac{P_i}{P_{max}} \right) \quad (3)$$

L_i represents the transmission delay of devices, L_{max} represents the maximum transmission delay that the

system can tolerate, P_i represents the power consumption of devices, and P_{max} represents the maximum power consumption.

In the scenario of visual data transmission, the original data is divided into data packets with the same size. Let the size of the data packet be S , then the transmission delay can be represented as follows:

$$L_i = \frac{S}{B_i} \quad (4)$$

In this paper, the devices access to basestation through 5G wireless network, the relationship between bandwidth and power consumption is shown as Shannon theorem[5]:

$$B_i = F_i \log \left(1 + \frac{g_i P_i}{N_i} \right) \quad (5)$$

In equation (5), F_i represents the frequency of wireless signal, g_i represents the signal gain, P_i represents the transmission power of the signal, that is, the power consumption of device. N_i is the noise of the wireless channel. In the process of wireless signal propagation, the wireless signal will be affected by path loss, that is, as the transmission distance increases, the signal will become weaker, thereby affecting the transmission bandwidth. For quantifying the loss caused by transmission distance, we use the equation mentioned in [6] to represent the path loss as below:

$$PL_i = PL_0 + 10\alpha \log_{10} \left(\frac{d_i}{d_0} \right) \quad (6)$$

In equation (6), α is a parameter related to the transmission environment, d_i is the distance between the device and the 5G basestation, and d_0 is the distance between the device and basestation when the path loss is PL_0 . Combining equation (5) and (6), the transmission bandwidth influenced by distance and transmission power can be defined as follows:

$$B_i = F_i \log \left(1 + \frac{g_i P_i - PL_i}{N_i} \right) \quad (7)$$

By transforming equation (7), we can also obtain the power consumption:

$$P_i = \frac{N_i(2^{B_i/F_i-1})+PL_i}{g_i} \quad (8)$$

The objective of this paper is to obtain a bandwidth allocation strategy that maximizes the comprehensive utility q of transmission system. The objective can be formulated as follows:

$$\max_{B_i} \frac{\sum_{i=1}^n w_i B_i / B_{max}}{\sum_{i=1}^n w_i \left(\frac{L_i}{L_{max}} + \frac{P_i}{P_{max}} \right)} \quad (9)$$

$$\text{s.t. } 0 < B \leq B_{max} \quad (10)$$

$$0 < P \leq P_{max} \quad (11)$$

$$0 < L \leq L_{max} \quad (12)$$

The optimization problem defined in equation (9)-(12) is a non-linear fractional programming problem. There are two main difficulties for solving this problem. The one is that the problem requires a very large amount of calculation in the solution process, another one is that multiple local optimums are harmful to obtain global optimum. These difficulties will consume too much computing resources and also increase the transmission delay. A low-complexity solution algorithm is necessary to solving this problem.

3 Joint Optimization Algorithm for Comprehensive Utility

In this section, we propose a theorem to transform the non-linear fractional programming problem above to convex optimization problem. Then we construct the solution space and design the search algorithm to find optimum.

3.1 Monotonic Equivalent Transformation Theorem

By observing $U_R(B)$ and $U_C(B)$ defined in (2) and (3), Obviously, $U_R(B) \geq 0$ and $U_C(B) > 0$, let $q^* = \max_{B_i} \frac{U_R(B)}{U_C(B)}$, according to [7], we can get the following theorem:

Theorem 1. When $U_R(B) \geq 0$ and $U_C(B) > 0$, for any bandwidth resource allocation strategy that satisfies the constraints in equation (10)-(12), the following two equations are equivalent:

$$q^* = \max_{B_i} \frac{U_R(B)}{U_C(B)} \quad (13)$$

$$G(q^*) = \max_{B_i} (U_R(B) - q^* U_C(B)) = 0 \quad (14)$$

And $G(q)$ has the following properties:

- $G(q)$ is a continuous convex function.
- $G(q)$ strictly decreases monotonically.
- The solution with $G(q) = 0$ is unique.

Proof. Let $B_0 = \arg \max_{B_i} \frac{U_R(B)}{U_C(B)}$, proof (13) \Rightarrow (14) firstly.

According to $U_R(B) \geq 0$ and $U_C(B) > 0$, we can obtain that

$$q^* = \frac{U_R(B_0)}{U_C(B_0)} = \max_{B_i} \frac{U_R(B)}{U_C(B)} \geq 0 \quad (15)$$

For $\forall B$, equation (15) implies to

$$q^* = \max_{B_i} \frac{U_R(B)}{U_C(B)} \Rightarrow q^* \geq \frac{U_R(B)}{U_C(B)} \quad (16)$$

That is

$$U_R(B) - q^* U_C(B) \leq 0 \quad (17)$$

$$q^* = \frac{U_R(B_0)}{U_C(B_0)} \Rightarrow U_R(B_0) - q^* U_C(B_0) = 0 \quad (18)$$

Combining equation (17) and (18), we can obtain $G(q^*) = \max_{B_i} (U_R(B) - q^* U_C(B)) = 0$. (13) \Rightarrow (14) is

proven. Then proof (14) \Rightarrow (13). Similarly, there is a equation

$$\begin{aligned} G(q^*) &= \max_{B_i} (U_R(B) - q^* U_C(B)) \\ &= U_R(B_0) - q^* U_C(B_0) \\ &= 0 \\ &\geq U_R(B) - q^* U_C(B) \end{aligned} \quad (19)$$

We can obtain that:

$$U_R(B_0) - q^* U_C(B_0) = 0 \Rightarrow q^* = \frac{U_R(B_0)}{U_C(B_0)} \quad (20)$$

$$U_R(B) - q^* U_C(B) \leq 0 \Rightarrow \frac{U_R(B)}{U_C(B)} \leq q^* \quad (21)$$

(14) \Rightarrow (13) is proven.

Observing the form of $U_R(B)$ and $U_C(B)$, we can see that $U_R(B)$ and $U_C(B)$ are both consecutive continuous convex functions. $G(q)$ is the convex combination of $U_R(B)$ and $U_C(B)$, so $G(q)$ is also a continuous convex function.

Prove up.

According to theorem.1, the optimization problem (9)-(12) can be transformed as follows:

$$\max_{B_i} G(q) \quad (22)$$

$$\text{s.t. } 0 < B \leq B_{max} \quad (23)$$

$$0 < P \leq P_{max} \quad (24)$$

$$0 < L \leq L_{max} \quad (25)$$

Equation (22)-(25) is a nonlinear convex optimization problem. For any q , there is only an optimal allocation strategy B_i which can maximize $G(q)$. As $G(q)$ strictly decreases monotonically, we can find the optimum q^* in the solution space.

3.2 Construct Solution Space

For constructing the solution space of equation (22)-(25), we need to find the upper bound and lower bound of optimum q^* .

Obviously, the lower bound $q_{min} = 0$, as equation (15) shows.

For the upper bound q_{max} ,

$$q_{max} = \frac{\max U_R(B)}{\min U_C(B)} \quad (26)$$

$$\max U_R(B) = \sum_{i=1}^n w_i \quad (27)$$

$$\begin{aligned} U_C(B) &= \sum_{i=1}^n w_i \left(\frac{L_i}{L_{max}} + \frac{P_i}{P_{max}} \right) \\ &= \sum_{i=1}^n w_i \left(\frac{S}{B_i L_{max}} + \frac{N_i(2^{B_i/F_i-1})}{g_i P_{max}} \right) \end{aligned} \quad (28)$$

$$\geq \sum_{i=1}^n w_i \sqrt{\frac{SN_i}{g_i B_i L_{max} P_{max}}} = \min U_C(B)$$

If and only if $B_i L_{\max} = N_i (2^{B_i/F_i} - 1)$, $\min U_C(B) = \sum_{i=1}^n w_i \sqrt{\frac{SN_i}{g_i B_i L_{\max} P_{\max}}}$. So,

$$q_{\max} = \frac{\sum_{i=1}^n w_i}{\sum_{i=1}^n w_i \sqrt{\frac{SN_i}{g_i B_i L_{\max} P_{\max}}}} \quad (29)$$

In summary, the solution space of equation (22)-(25) is

$$q^* \in \left[0, \frac{\sum_{i=1}^n w_i}{\sum_{i=1}^n w_i \sqrt{\frac{SN_i}{g_i B_i L_{\max} P_{\max}}}} \right] \quad (30)$$

Algorithm 1 Joint Optimization Algorithm for Comprehensive Utility

Input: Device performance $B_{\max} L_{\max} P_{\max} F_i g_i N_i S$
Output: Allocation strategy B

```

0: Initialize  $q_{\min} = 0, q_{\max}$ 
0: Set the accuracy  $\varepsilon$ 
0: while  $|q_{\max} - q_{\min}| > \varepsilon$  do
0:    $q = \frac{q_{\max} + q_{\min}}{2}$ 
0:   Solve the optimization problem (22)-(25), obtain  $B$ 
0:   if  $G(q) \leq 0$  then
0:      $q_{\max} = q$ 
0:   else
0:      $q_{\min} = q$ 
0:   end if
0: end while
0: return  $B = 0$ 

```

4 Performance Evaluation

In order to evaluate the performance of proposed method, we design the simulation experiment. This experiment will evaluate the proposed method from four aspects: bandwidth, latency, power consumption and comprehensive utility.

4.1 Dataset

In this paper, we use GO!Track[8] as the dataset to simulate the distribution of IoT devices. GO!Track dataset is collected from GPS devices located on urban road vehicles in the city. It contains the location information of 117 devices and their moving paths, which can well simulate the distribution of devices in cities. For simplicity, the position of the basestation is artificially set at the origin, as Fig.2 shows.

3.3 Optimization Algorithm

After the solution space is determined, we can design a search algorithm to obtain the optimum. The steps of the algorithm are as follows:

1) Select the midpoint of the solution space $q = (q_{\max} - q_{\min})/2$ as the optimum in the current iteration. Solve the model (22)-(25) and obtain the allocation strategy B' and $G(q)$.

2) If $G(q) \leq 0$, according to equation (14), we should decrease q_{\max} . Let $q_{\max} = q$ and restart the iteration.

3) If $G(q) > 0$, according to equation (14), we should increase q_{\min} . Let $q_{\min} = q$ and restart the iteration.

4) After the optimum q^* satisfies accuracy, the algorithm will stop and return the optimal allocation strategy B .

The whole steps are shows in Algorithm.1.

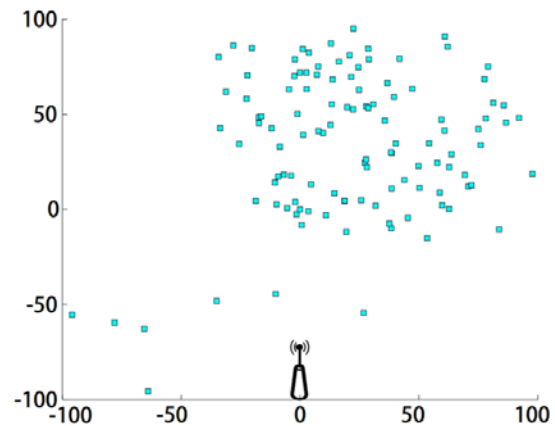


Figure.2. The distribution of devices in GO!Track dataset.

4.2 Experiment Setup

In this paper, we initialize the parameters in the algorithm with 5G network characteristics and wireless path loss mentioned in [9] and [10], as Table.2 shows.

Table 2. The Table of Values and parameters used in the experiment.

Parameter Name	Value
B_{\max}	1Gbps
L_{\max}	10s
P_{\max}	1W
F_i	3.5GHz

g_i	20db
N_i	10db
S	2MB
α	1
PL_0	3.4db
d_0	1m

For better experiment, we select three state-of-art method as the reference substance:

- VIRTUS[11], a bandwidth-first method. The idea of this method is to establish a probability model about the transmission bandwidth based on the throughput of data packets, and use the maximum likelihood method to obtain the optimal bandwidth allocation strategy.
- nDCs[12], a latency-first method. The idea of this method is to establish a time series model related to the transmission delay based on the transmission power consumption of different devices, and to reduce the transmission delay by adjusting the power consumption on the time series model.
- NOMA[13], a power-efficiency-first method. The idea is to take the power consumption per bit as the optimization index, establish the transmission optimization model, and use the greedy algorithm to search the solution space to obtain the optimal bandwidth resource allocation strategy.

4.3 Performance Analysis on Sole Index

In order to evaluate the performance of the proposed method on a single index, we conduct this experiment and measure the transmission bandwidth, latency and power consumption. The results are recorded and displayed in the form of cumulative distribution function (CDF).

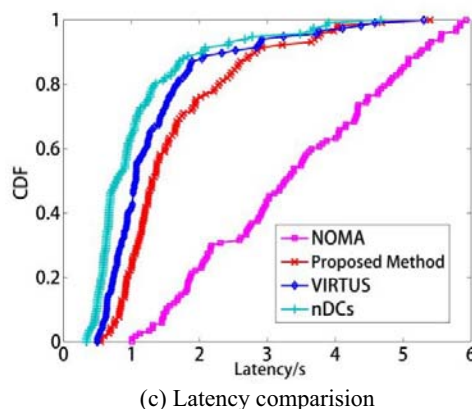
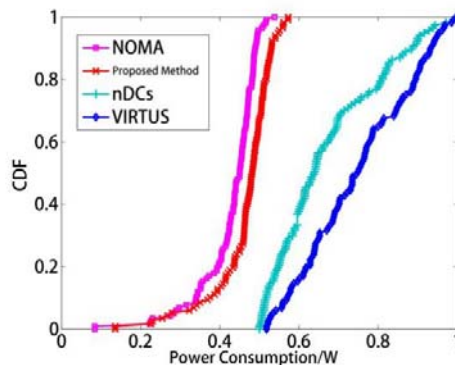
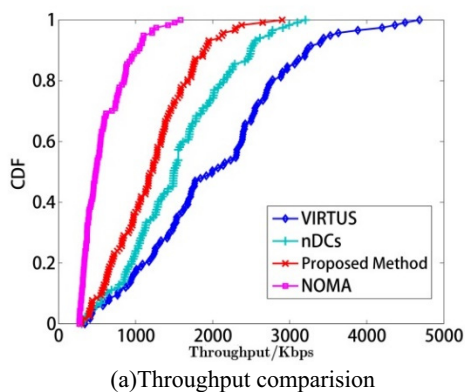


Figure.3. The comparison between the proposed method and others methods in sole index.

As fig.3 shows, in the aspect of sole index, the proposed method is not optimal. Fig.3 (a) shows that the proposed method is only better than the NOMA in terms of transmission throughput. Compared with VIRTUS and nDCs, the transmission throughput of proposed method is lower than both. Fig.3 (b) shows that the performance of proposed method in power consumption is lower than NOMA, and better than VIRTUS and nDCs. Fig.3 (c) shows that the latency of proposed method is close to VIRTUS and nDCs.

4.4 Performance Analysis on Comprehensive Utility

Due to the complexity of the visual data transmission, the performance of a sole index can not reflect the actual performance of the algorithm. Therefore, this experiment evaluates proposed method from the aspect of comprehensive utility defined as follows.

$$U_i = \frac{B_i/B_{max}}{L_i/L_{max} + P_i/P_{max}} \quad (31)$$

In equation (31), B_i , P_i , L_i denotes the transmission bandwidth, power consumption and transmission latency of the devices. B_{max} , P_{max} , L_{max} is shown in Table 2. The comprehensive utility takes consideration of the influence of transmission bandwidth, power consumption and latency on the transmission performance, so it can better reflect the performance of the proposed method. The experiment results show in Fig. 4.

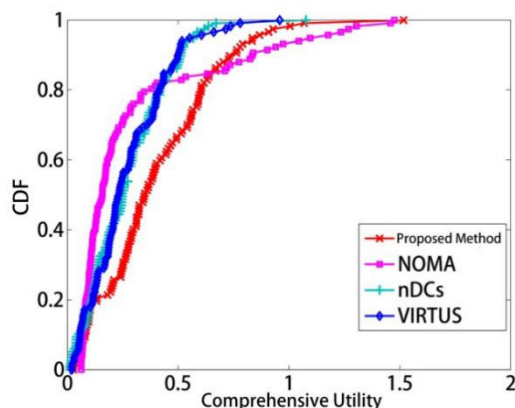


Figure.4. The comparison between the proposed method and others methods in comprehensive utility.

As shown in Fig.4, because the transmission bandwidth, power consumption and latency are considered comprehensively, the proposed method achieves the best effect in the comprehensive utility. Specifically, the proposed method is 35.4% higher than NOMA, 45.3% higher than nDCs and 47.8% higher than VIRTUS.

5 Conclusion

In this paper, we propose a method to jointly optimize the resource allocation of visual data transmission in resource constraint 5G IoT. By building and solving the hybrid objective model with throughput, power consumption and latency, the optimal resource allocation strategy can be obtained. Through numerical simulation experiments on GO!Track data set, the experimental results show that the proposed method improves comprehensive utility of transmission system by 35.4% - 47.8%.

Acknowledgments

This work is supported by the Science and Technology Project of State Grid Corporation of China (SGJSYF00KFJS2000088) and State Grid Energy Research Institute Technology Development Project (526700190004).

References

1. Husenovic K, Bedi I, Maddens S, et al. Setting the Scene for 5G: Opportunities & Challenges[J]. International Telecommunication Union, 2018, 56.
2. Forecast, Global Mobile Data Traffic. "Cisco visual networking index: global mobile data traffic forecast update, 2017–2022." Update 2017 (2019): 2022.
3. Lund D, MacGillivray C, Turner V, et al. Worldwide and regional internet of things (IoT) 2014–2020 forecast: A virtuous circle of proven value and demand[J]. International Data Corporation (IDC), Tech. Rep, 2014, 1: 9.

4. Rao, Jaya B., and Abraham O. Fapojuwo. "On the tradeoff between spectral efficiency and energy efficiency of homogeneous cellular networks with outage constraint." *IEEE transactions on vehicular technology* 62.4 (2012): 1801-1814.
5. Jerri, Abdul J. "The Shannon sampling theorem—Its various extensions and applications: A tutorial review." *Proceedings of the IEEE* 65.11 (1977): 1565-1596.
6. George R Maccartney, Theodore S Rappaport, Mathew K Samimi, and Shu Sun. Millimeter-wave omnidirectional path loss data for small cell 5G channel modeling. *IEEE Access*, 3:1573–1580, 2015.
7. Konglin Zhu, Wenting Zhi, Lin Zhang, Xu Chen, and Xiaoming Fu. Social-aware incentivized caching for D2D communications. *IEEE Access*, 4:7585–7593, 2016.
8. Michael O Cruz, Hendrik Macedo, and Adolfo Guimaraes. Grouping similar trajectories for carpooling purposes. In *2015 Brazilian Conference on Intelligent Systems (BRACIS)*, pages 234–239. IEEE, 2015.
9. NGMN Alliance. 5g white paper. Next generation mobile networks, white paper, pages 1–125, 2015.
10. George R Maccartney, Theodore S Rappaport, Mathew K Samimi, and Shu Sun. Millimeter-wave omnidirectional path loss data for small cell 5G channel modeling. *IEEE Access*, 3:1573–1580, 2015.
11. Rubem Pereira and Ella Grishikashvili Pereira. Video streaming considerations for Internet of Things. In *2014 International Conference on Future Internet of Things and Cloud*, pages 48–52. IEEE, 2014.
12. Tiago Gama Rodrigues, Katsuya Suto, Hiroki Nishiyama, and Nei Kato. Hybrid method for minimizing service delay in edge cloud computing through VM migration and transmission power control. *IEEE Transactions on Computers*, 66(5):810–819, 2017.
13. Farzad Samie, Lars Bauer, and Jörg Henkel. IoT technologies for embedded computing: A survey. In *Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*, page 8. ACM, 2016.