Hindawi
Security and Communication Networks
Volume 2022, Article ID 4126494, 11 pages
https://doi.org/10.1155/2022/4126494

Research Article

QoE-Aware Video Delivery in Multimedia IoT Network with Multiple Eavesdroppers

Dapeng Wu,[1,2,3] **Ruixin Xu**,[1,2,3] **Hong Zhang**,[1,2,3] **Zhidu Li**,[1,2,3] **Ruyan Wang**,[1,2,3] **Alexander Fedotov**,[4] **and Vladimir Badenko**[4]

[1]*School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China*
[2]*Advanced Network and Intelligent Connection Technology Key Laboratory of Chongqing Education Commission of China, Chongqing, China*
[3]*Chongqing Key Laboratory of Ubiquitous Sensing and Networking, Chongqing, China*
[4]*Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia*

Correspondence should be addressed to Hong Zhang; hongzhang@cqupt.edu.cn

Received 19 February 2022; Accepted 15 June 2022; Published 8 August 2022

Academic Editor: Anwar Ghani

How to deal with the increasing video traffic and diverse service demands while ensuring the security of transmission is an open issue in the multimedia Internet of Things (IoT). This paper addresses this issue and studies a secure delivery scheme under a multicast scenario in the presence of multiple eavesdroppers where small base stations (SBSs) can send videos to users cooperatively. Aiming at potential eavesdroppers, a channel model including artificial noise is introduced to reduce the harm of illegal data acquisition. A network quality of experience (QoE) optimization problem is first formulated to account for video quality and delivery delay. In order to solve the nonconvex problem, the successive convex approximation (SCA) technique is applied to optimize multicast group beamforming, reduce the possibility of multicast video eavesdropping, and select video quality where a heuristic scheme is proposed to maximize the network QoE. The effectiveness of the proposed scheme is finally validated by extensive simulations in terms of algorithm convergence performance and network QoE-enhanced performance.

1. Introduction

Recently, with the rapid development of mobile communication networks, the multimedia-oriented Internet of Things (IoT) network has shown a strong development momentum. Diversified multimedia services, such as video surveillance, make video data occupy a large proportion of multimedia IoT [1–3]. According to the existing data, the mobile network traffic data has increased by 42% from 2020 to 2021 [4]. By the end of 2022, the multimedia content data will account for 82% of the global mobile traffic [5]. The integration of various mobile terminals and multimedia IoT will promote the rapid growth of this number. In particular, the continuous upgrading of the multimedia IoT industry will have higher requirements for services, resulting in more video data, such as automatic driving and smart city. The resulting multimedia data brings more tremendous pressure to the uplink and downlink transmission of IoT [6]. In this regard, deploying MEC at the small base station (SBS) can bring services to the edge of the network and enable IoT users to obtain better experience [7, 8]. Due to the large number of users and the diversity of video requirements, it is worth studying how to efficiently deliver videos to IoT users with limited resources.

With layered technology such as H.264/moving Picture Experts Group-4 (MPEG-4) scalable video coding (SVC), a video stream can be divided into different quality levels to provide users with more personalized services [9]. Generally, more multimedia data layers can bring users a better experience, but it also needs to pay more communication costs. When faced with many requests, some additional video enhancement data may squeeze the resources of other

vulnerable users and introduce additional interference, resulting in unfair resource allocation [10]. On the other hand, users with similar needs in the network can be divided into the same multicast group through multicast transmission. Multicast transmission of SVC video can make full use of its hierarchical structure, effectively reducing the system energy consumption and significantly reducing the reception delay in the face of many IoT users [11]. With a reasonable beamforming design, SBS can use limited communication resources to serve more users on the premise of reducing resource loss [9]. However, the nature of multicast transmission makes the communication over this medium vulnerable to eavesdropping [12]. In network service, transmission security and privacy issues will directly bring terrible experiences to users. A part of the literature has studied this problem from the perspective of security protocols and encryption algorithms [11, 13, 14]. In addition, there are often eavesdroppers in the process of wireless transmission. Eavesdroppers trying to access multicast services without authorization will cause economic losses to operators. Using beamforming technology and SVC technology [15–18], through more accurate beamforming design, artificial noise can be introduced to reduce the channel quality of potential eavesdroppers as much as possible, making it difficult for them to obtain complete transmitted video [19]. On the other hand, if the eavesdropper cannot obtain the basic layer data, the transmission security of the complete video can be guaranteed.

In the literature, researchers usually design optimization strategies to improve the performance in the network from two directions, that is, secure transmission and cached video delivery. On the security of transmission studies, it is often assumed that there are potential eavesdroppers in the network. The transmission signals of actual users are designed from the perspective of the physical layer to improve confidentiality and prevent data leakage [20, 21]. However, most studies regard security as the main goal rather than a prerequisite to optimizing network performance. An effective active caching strategy can filter out popular data from a large amount of data for caching to reduce the delay of users obtaining content and improve the user experience. However, this often depends on the screening of a large number of historical data, and the improvement of performance depends too much on the accuracy of cache. Using reasonable resource allocation strategy, the improvement of network performance will often have greater guarantee [22, 23]. Many efforts have been devoted to the efficient transmission of cached content at the edge of the network. However, how to jointly consider the cost of content cache location and limited network resources to deliver video while ensuring transmission security still lacks understanding.

Motivated by this, we study a video delivery scheme to maximize the weighted sum of QoE in a multimedia IoT network. The multicast transmission model with eavesdroppers, cache cost model, and QoE model are first introduced, based on which a network QoE optimization problem is formulated. By applying the SCA technique, cooperative beamforming and video quality selection are jointly optimized to guide how to deliver videos to different IoT users. Furthermore, extensive simulation experiments are carried out to verify the QoE enhancement performance of our proposed scheme.

The contributions of this paper are summarized as follows:

(i) In the multimedia IoT network scenario, the corresponding transmission model and the user QoE model are designed according to the existence of eavesdroppers. The design of the beamforming model is used to prevent insecurity in the multicast process, which is reflected in the design of the optimization problem.

(ii) An alternating iterative scheme considering system beamforming design and user acquired video quality selection strategy is designed with SCA technology. The QoE optimal solution under the condition of fixed video quality is found step by step through alternating updating.

(iii) A user selection scheme for video quality enhancement is designed by adding penalty parameters. The simulation results show that, combined with an iterative alternating algorithm, this mechanism can ensure that the system can use limited resources to obtain the best weighted QoE.

The rest of the paper is organized as follows. The related works are reviewed in Section 2. Section 3 introduces the transmission model and the optimization problem. The algorithm is designed in Section 4. Section 5 provides and discusses the simulation results. Section 6 concludes the paper.

2. Related Work

In the literature, the service experience enhancement is usually studied through proactive caching and content delivery. When seeking the best experience, the security of transmission also needs to be considered [24]. Hence, the related works are reviewed based on the lines of security of multicast transmission and cached-enabled video delivery, respectively.

On the security of transmission study, the focus is mainly on the impact of the insecure channel model and preventing content from being eavesdropped on during multicast transmission. In [20], two cognitive single-group multicast secure beamforming (SGMC-S-BF) schemes were proposed for a scenario where there exists one eavesdropper who is actually a regular user of the legitimate communication system. However, it attempts to access unauthorized multicast services. In [25], a task-oriented user selection incentive mechanism was proposed, which clusters the similarity of users by jointly considering the security and fairness of served users to realize efficient user dynamic selection. In [26], the constant modulus (CM) signaling was studied, and beamforming is designed using the semidefinite relaxation (SDR) technique and a custom-build nonconvex alternating direction method of multipliers (ADMM) algorithm, respectively. In [27], power minimization and

secrecy rate maximization were considered in the secrecy network. A closed-form solution of transmit beamforming is given by exploiting the Bernstein-type inequality and the S-Procedure to convert the probabilistic secrecy rate constraint into the determined constraint. Literatures [20–27] focus on improving security performance but ignore the improvement of user performance indicators. In [21], a SDP-based secure layered video transmission scheme was proposed, but it is not suitable for the case of multiple base stations. Hence, the improvement of video delivery performance considering a nonsecure environment still needs further study in multigroup multicast scenarios.

Research on video delivery focuses on limited resource allocation and delay optimization. In [22], a dynamic interest capture model in the Industrial Internet was proposed to mine the individual user interest, based on which a group interest aggregation algorithm is then studied to determine the content caching strategies for edge nodes. In [28], an adaptive active cache scheme combined with reinforcement learning was proposed to improve user QoE of content-centered edge cache IoT and reduce cache cost. However, [22–28] mainly focus on the active cache scheme, ignoring system performance optimization by transmission strategy. In [29], the authors predicted video content requests from the perspective of the traces collected over a big city, and a joint active caching, power allocation, and user association scheme was designed to optimize the QoE of user content delivery. In [30], an active cache and user association scheme based on belief propagation was proposed to maximize the revenue of operators on the premise of ensuring the quality of service (QoS). In [31], a social-aware spectrum sharing and caching helper selection (SSC) strategy was proposed to share and cache downlink spectrum resources of multicast storage resources and unload multimedia content. In [32], the authors addressed the impact of cache on the transmission design, and a robust joint optimization strategy is designed for the case of incomplete channel information to minimize the transmission cost. In [33], a multiquality video transmission scheme based on SBS clusters division was proposed to maximize the economic efficiency (ECE) of network transmission. In summary, [29–33] optimized video delivery based on only unicast or multicast transmissions under the scenario with single base station. In addition, the above literature ignores the personalized demands of users, and multiquality video transmission is often more accurate to serve each user and save limited network resources.

Due to the increasingly personalized user demands and the complexity of the network environment, SVC technology is usually applied to meet the diverse requirements of video delivery in IoT network. In [34], the author used the stochastic geometry tool to increase the probability of successful transmission of multiquality video and proposed a two-stage transmission optimization algorithm based on the convex optimization and the packing problem. In [35], a matching sensing scheme considering relay selection was proposed to optimize the cooperative transmission performance of SVC video and improve the QoE of users. In [36], a multicast video transmission strategy based on the multicast subgrouping and SVC technology is proposed, and the transmission efficiency and fairness between users were discussed. In [37], the author addressed the application of multiquality video in multicast transmission and described the optimal beamforming as a power minimization problem and user experience maximization problem, respectively. In [23], cache-assisted data rate (CADR) was proposed as a performance index to measure the SVC video transmission performance in nonorthogonal channel D2D scene, and a two-stage joint optimization scheme is proposed. However, [34–37] mainly discussed the perspective of multiquality video transmission, but the cost of cache location and transmission interference in the case of multigroup multicast has not been considered. The D2D auxiliary transmission strategy proposed by [23] is difficult to apply directly in multicast transmission based on multi-SBS.

In summary, on the premise of the existence of eavesdroppers, there is still lack of deep understanding on how to meet the personalized user demands through jointly optimizing the multiquality video and multicast transmission beamforming and ensure transmission security, which motivates this paper.

3. System Model

In this section, we present a multimedia IoT network architecture, transmission model with eavesdroppers, cache model, and delay cost and QoE model, and the modeling of the optimization problem is introduced.

3.1. Transmission Model. As depicted in Figure 1, we focus on a cached-enabled radio access network that includes multiple densely deployed SBSs and multiple IoT users and eavesdroppers indicated by $\mathcal{K} = \{1, \ldots, k, \ldots, K\}$, $\mathcal{U} = \{1, \ldots, u, \ldots, U\}$, and $\mathcal{U}_E = \{1, \ldots, u_E, \ldots, U_E\}$, respectively, where videos can be transmitted through nonorthogonal multicast mode. In addition, each of SBS is configured with A_k antennas and can cache popular content in the equipped MEC storage unit at the network edge, so that IoT user requests can be responded to quickly. Based on the content and the quality level of a video request, the IoT users are divided into several multicast groups $\mathcal{G} = \{1, \ldots, g, \ldots, G\}$. In this regard, wireless resources can be saved by multicasting the videos to users with identical content and quality request in a group.

In this paper, SBS k can be associated with multiple groups k_g, and the user in such group is represented by g_u. The channel power gain between IoT user u, eavesdroppers j, and SBS k is denoted by $\mathbf{h}_u = [\mathbf{h}_{1,u}^H, \ldots, \mathbf{h}_{K,u}^H]^H \in \mathbb{C}^{A_n K \times 1}$ and $\mathbf{e}_u = [\mathbf{e}_{1,u_E}^H, \ldots, \mathbf{e}_{K,u_E}^H]^H \in \mathbb{C}^{A_n U_E \times 1}$, respectively. The beamforming vector of SBS k to multicast group g is denoted by $\mathcal{W}_k = \{\mathbf{w}_{k,g} \in \mathbb{C}^{A_k \times 1}\}$. On the other hand, $\mathbf{v}_k = [\mathbf{v}_1^H, \ldots, \mathbf{v}_K^H]^H \in \mathbb{C}^{A_n K \times 1}$ represents the artificial noise generated by SBS by beamforming, which is used to further reduce the channel conditions of eavesdroppers. Besides, $P_k > 0$ is used to denote the maximum transmission power of SBS k; there holds

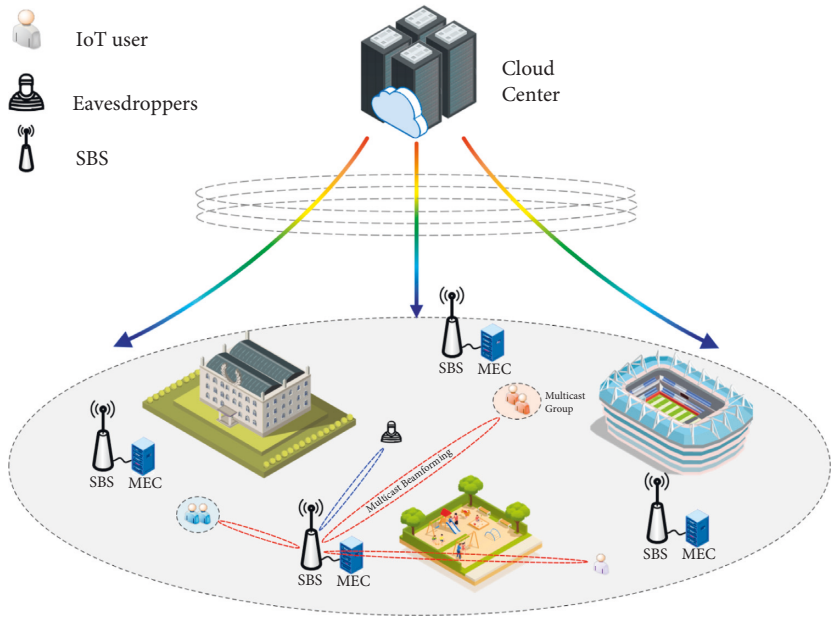

Figure 1: A video streaming multicast transmission scenario with eavesdroppers.

$$\sum_{g\in\mathcal{G}} w_{k,g2}^2 + \mathbf{v}_{k2}^2 \le P_k, \quad \forall k \in \mathcal{K}. \tag{1}$$

For multicast group g, we use $\mathbf{w}_g = [\mathbf{w}_{1,g}^H, \ldots, \mathbf{w}_{K,g}^H]^H$ to represent the beamforming vector group it receives. For simplicity, the SVC technology encodes a video into two layers, and SBSs will store different quality files of a video in the cache space. To deal with the quality level of each video, the base layer (BL) and the enhancement layer (EL) are used to characterize the basic video quality and the enhancement video quality, respectively. The multimedia contents quality requested by IoT user u is represented by $l_u = \{1, 2\}$, where $l_u = 1$ represents that IoT user requests low-quality multimedia contents and $l_u = 2$ represents that IoT user requests high-quality multimedia contents. Besides, the requested video content and the video quality of the multicast group g are represented as g_f and g_l, respectively. Note that high-quality multimedia contents include both BL data and EL data at the same time. A video content cannot be decoded with only EL data. Let l_u^{req} denote the quality level requested by user u; there holds

$$l_u \ge l_u^{\text{req}}, \quad \forall u \in \mathcal{U}. \tag{2}$$

In order to improve the multimedio content delivery efficiency, a user can be assigned to at most two groups where one group only receives BL data and the other receives EL data, as depicted in Figure 2. The network only needs to allocate a multicast beamforming vector to provide BL data to users requesting low-quality video and users requesting high-quality video simultaneously. On the other hand, when SBS transmits data to multicast group users, eavesdroppers can receive broadcast information and obtain unauthorized video content. By reducing its signal-to-interference-to-noise ratio (SINR), it cannot restore the complete video, so as to prevent the content from being intercepted. In this sense, the network only needs to assign only one multicast beamforming vector to provide BL version video to the users requesting low-quality videos and those requesting high-quality videos at the same time. The received signals of IoT user u and eavesdropper j are given by, respectively,

$$\begin{aligned} y_u &= \sum_{g\in\mathcal{G}_u} \mathbf{h}_u^H \mathbf{w}_g s_g + \sum_{i\in G\setminus\mathcal{G}_u} \mathbf{h}_u^H \mathbf{w}_i s_i + n_u, \quad \forall u \in \mathcal{U}, \\ y_{u_E} &= \sum_{g\in\mathcal{G}} \mathbf{e}_{u_E}^H \mathbf{w}_g s_g + \sum_{k\in\mathcal{K}} \mathbf{e}_{u_E}^H \mathbf{v}_k s_g + n_j, \quad \forall u_E \in \mathcal{U}_E, \end{aligned} \tag{3}$$

where $\mathcal{G}_u$ represents the multicast group set assigned by the user u. s_g with $\mathbb{E}|s_g|^2 = 1$ represents the precoding binary symbols of the video requested by group g, and $n_u \sim CN(0, \sigma_u^2)$ denotes the additive white Gaussian noise (AWGN). Applying Shannon's theory, the transmission rate of u when requesting a video f with quality level l holds as

$$R_{u,f_l} = B\log_2\left(1 + \Gamma_{u,f_l}\right). \tag{4}$$

The SINR is denoted by

$$\Gamma_{u,f_l} = \frac{\left|\mathbf{h}_u^H \mathbf{w}_{g,f_l}\right|^2}{I_{u,f_l}}, \quad \forall u \in g,\ g \in \mathcal{G}, \tag{5}$$

where

$$\begin{aligned} I_{u,f_l} = \sum_{l>l_{\text{req}}} \left|\mathbf{h}_u^H \mathbf{w}_{g,f_l}\right|^2 + \sum_{i\in\setminus\{g\}} \left|\mathbf{h}_u^H \mathbf{w}_i\right|^2 + \sigma^2, \\ \forall u \in g,\ g \in \mathcal{G}. \end{aligned} \tag{6}$$

In (5), $\mathbf{w}_{g,f_l}$ denotes the multicast group beamforming vector which provides content f with quality level l. In (6), the first term denotes higher-quality version signal

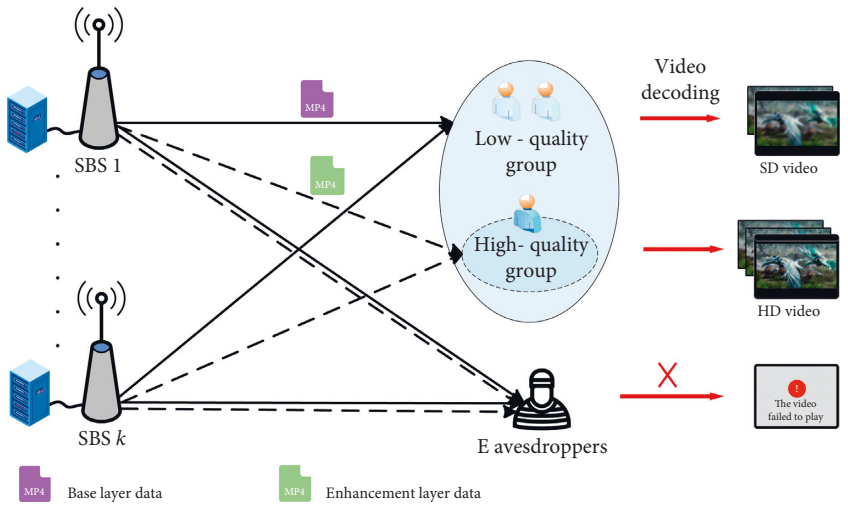

Figure 2: Cooperative video multicast transmission model.

interference to the same video content, and the second term denotes the other group signal interference. Since SBS actively generates artificial noise, the coded information can be transmitted to the user in advance, so it will not be regarded as noise. In addition, to ensure the user fairness in group g, the transmission rate should depend on the user with the worst channel condition in the group, and the group transmission rate thus holds

$$R_{g,f_l} = \min_{u \in g} R_{u,f_l}. \tag{7}$$

On the other hand, the transmission rate of eavesdropper j can be obtained as

$$\begin{aligned} R_{u_E,f_l} &= B\log_2\left(1 + \Gamma_{u_E,f}\right), \\ \Gamma_{u_E,f} &= \frac{\left|\mathbf{e}_{u_E}^H \mathbf{w}_{g,f_l}\right|^2}{I_{u_E}}, \forall u_E \in \mathcal{U}_E, \end{aligned} \tag{8}$$

where

$$I_{u_E} = \sum_{i>1} \left|\mathbf{e}_u^H \mathbf{w}_i\right|^2 + 4 \sum_{k \in \mathcal{K}} \left|\mathbf{e}_u^H \mathbf{v}_k\right|^2 + \sigma^2, \forall u_E \in \mathcal{U}_E. \tag{9}$$

In (9), the noise can inform the authorized user in advance, and it will not cause additional interference to the target user. Due to the characteristics of SVC, we can ensure the security of multicast video transmission only by ensuring that the video of BL layer is not completely eavesdropped. Therefore, we give a maximum tolerance $\Gamma_{u_E,f} < \Gamma_{\text{toler}}$ of the eavesdropper, so that it cannot obtain the complete video.

3.2. Cache Model. The limited video content library is denoted by $\mathcal{F} = \{1, \ldots, f, \ldots, F\}$, where F is the library size. The size of a unit video block with quality l of different content f is denoted by $S_{f,l}$. In order to facilitate the experiment, different video blocks of the same quality are assumed to have the same size S_l, which can be realized by adaptive slicing of video. The popularity of all contents is assumed to follow the Zipf distribution [38], that is, the probability that content $f - th$ being requested is given by

$$\begin{aligned} P(f) &= \frac{f^{-\alpha}}{\sum_{i=1}^{F} i^{-\alpha}}, \\ f &= 1, 2, \ldots, F, \end{aligned} \tag{10}$$

where α denotes the Zipf parameter. The larger α, the more significant the popularity skewness between different contents. Suppose that each SBS will cache contents according to the popularity ranking until the cache space is full, and the cache state of content f at SBS k can be obtained as

$$c_{k,f} = \begin{cases} 1, & \text{if } f \text{ is cached by SBS } k, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

3.3. Delay Cost and QoE Model. As the cache capacity of each MEC is finite, only some popular videos can be precached. If a video is not cached in the MEC of the local SBS, the content must be obtained from the neighbor SBS or cloud center, which will introduce the extra delay in the video transmission process. Since multiple SBSs cooperate to transmit multimedia content to IoT users simultaneously, let D_{f_u} denote the extra delay; there holds

$$D_{f_u} = \begin{cases} 0, & \prod_{k \in \mathcal{K}} c_{k,f} = 1, \\ d_f, & \prod_{k \in \mathcal{K}} c_{k,f} = 0, \exists c_{k,f} = 1, \forall k \in K, \\ d_0, & \sum_{k \in \mathcal{K}} c_{k,f} = 0. \end{cases} \tag{12}$$

Equation (12) indicates that there is no extra delay if the video is stored in all MEC cache equipped with SBS. If the video is only cached in the MEC of a part of SBSs, the extra delay equals a fixed value d_f. It is caused by the video content sharing between SBSs using Xn interface in the 5G networks [39]. A fixed extra delay $d_0\,(d_0 \gg d_f)$ is needed if the video can only be obtained in the cloud center [40]. Furthermore, taking the video transmission delay from the local SBS to the user into account, the total video delivery delay is given by

$$D_u = \frac{S_l}{R_{g,f_l}} + D_{f_u}, \quad \forall u \in g, \forall g \in \mathcal{G}. \tag{13}$$

Compared with other factors, the startup delay of the video often directly affects the willingness of the user to play the video. In addition, from the practical experience, the benefits of reducing the startup delay meet the feature of diminishing margins rather than a fixed value. The lower the delay is, the lower the QoE improvement that can be achieved. Hence, we apply a logarithm to characterize the impact of startup delay of the video on the QoE in this paper [41]. On the other hand, a higher-quality level of video is able to improve the QoE compared to the one with low-quality level. Hence, the QoE of user u can be modelled by

$$QoE_u = (1-\eta)\frac{l_u}{l_{\mathrm{EL}}} + \eta\left(\log_2\left(1 + \frac{\beta}{\max\limits_{l \le l_u}\{D_u\}}\right)\right), \tag{14}$$

where $\eta \in [0,1]$ is a weight factor that characterizes the importance of the video quality and that of the delay on the QoE. Besides, β is a positive parameter used to control the marginal benefit of the startup delay. The first term indicates the impact of the definition of the user u-requested video on the overall QoE. The second term indicates the impact of the transmission delay on the overall QoE, where $\max_{l \le l_u}\{D_u\}$ represents the maximum transmission delay of all the video data received by user u.

3.4. Problem Formulation. In this paper, the aim is to maximize the sum of the QoE of each user that is called network QoE through jointly optimizing the video version selection and multicast group beamforming; there holds

$$Q = \sum_{u \in \mathcal{U}} QoE_u. \tag{15}$$

In order to improve the QoE of the IoT user, an efficient beamforming strategy can effectively reduce the transmission delay in the wireless transmission process. At the same time, the eavesdropper should provide the poor channel conditions designed to hinder its action. On the premise of ensuring the basic demands of users, providing users with video quality enhancement services at a low enhancement cost also helps to improve the QoE of the network. However, the improvement of video quality will bring additional interference to the same frequency transmission. Therefore, video quality selection and multicast group beamforming are jointly optimized. The optimization problem is formulated as follows:

$$\mathscr{P}_1 \max_{\mathbf{w},\mathbf{l}} Q, \tag{16a}$$

$$s.t.\ R_{g,f_l} \le R_{u,f_l},\ \forall u \in g,\ \forall g \in \mathcal{G}, \tag{16b}$$

$$R_l^{\mathrm{req}} \le R_{g,f_l}, \quad \forall g \in \mathcal{G}, \tag{16c}$$

$$\sum_{g \in \mathcal{G}} \left\|\mathbf{w}_{k,g}\right\|_2^2 + \mathbf{v}_{k2}^2 \le P_k, \quad \forall k \in \mathcal{K}, \tag{16d}$$

$$\Gamma_{u_E,f} < \Gamma_{\mathrm{toler}}, \forall u_E \in \mathcal{U}_E, \tag{16e}$$

$$l_u \ge l_u^{\mathrm{req}}, \quad \forall u \in \mathcal{U}. \tag{16f}$$

In problem $\mathscr{P}_1$, (16b) is the fairness constraint within the multicast group, (16c) represents the video bit rate constraint to ensure the QoS requirements of each, where R_l^{req} is the bit rate threshold corresponding to the video quality, (16d) is to the power constraint of each SBS, (16e) is SINR constraint of eavesdropper, (16f) means that the quality level of the video obtained by each user should be not lower than the corresponding requested one.

4. Algorithm Design

Since $\mathscr{P}_1$ is an MINLP problem, it is difficult to solve by convex technique directly. Therefore, it is promising to decouple the process of video quality selection and that of multicast group beamforming. First, the video quality selection parameter is fixed as $\mathbf{l}$. The optimization problem $\mathscr{P}_1$ can then be transformed as

$$\mathscr{P}_{1,1} \max_{\mathbf{w},\mathbf{v}} \sum_{u \in \mathcal{U}} \left(\log_2\left(\left(\frac{1+\beta}{\max\limits_{l \le l_u}}\right)\{D_u\}\right)\right), \tag{17a}$$

$$\text{s.t. } (16b), (16c), (16d), (16e). \tag{17b}$$

However, the quadratic term of $\mathbf{w}$ in constraint (16b) appears in both the numerator and the denominator, resulting in its nonconvex nature. On the other hand, since the right side of (16e) is a constant term, it is still a convex term, and the artificial noise part $\mathbf{v}$ can be solved directly. To extract the fractional part, we introduce auxiliary variables in the SINR part of (16b) and modify the original constraint into

$$\Gamma_u \le \frac{\left|\mathbf{h}_u^H \mathbf{w}_{g_u}\right|^2}{I_u}, \quad \forall u \in g, \forall g \in \mathcal{G}, \tag{18}$$

$$\sum_{l > l_{req}} \left|\mathbf{h}_u^H \mathbf{w}_{g,f_l}\right|^2 + \sum_{i \in G \setminus \{g\}} \left|\mathbf{h}_u^H \mathbf{w}_i\right|^2 + \sigma^2 \le I_u, \quad \forall u \in g, \forall g \in \mathcal{G}. \tag{19}$$

Consequently, the nonconvex constraint (16b) of the original problem is replaced by (18) and (21). Nevertheless, constraint (18) is still a nonconvex constraint. In this regard, we design an approximate convex lower bound to relax constraint (18). Let $f(\mathbf{w}_{g_u}, I_u) = \|\mathbf{h}_u^H \mathbf{w}_{g_u}\|_2^2 / I_u$ and perform

Taylor's first-order expansion of $f(\mathbf{w}_{g_u}, \gamma_u)$ at feasible points $\mathbf{w}_{g_u}^{(t)}$ and $\gamma_u^{(t)}$; constraint (18) can then be replaced by an auxiliary function as

$$\psi^{(t)}\left(\mathbf{w}_{g_u}, \gamma_u\right) \triangleq \frac{2\mathscr{R}e\left(\left(\mathbf{w}_{g_u}^{(t)}\right)^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{w}_{g_u}\right)}{I_u^{(t)}} - \left(\frac{\left|\mathbf{h}_u^H \mathbf{w}_{g_u}^{(t)}\right|}{I_u^{(t)}}\right)^2 \gamma_u, \tag{20}$$

where $\mathscr{R}e(\cdot)$ represents the real part and t represents the number of iterations. For any $(\mathbf{w}_{g_u}^{(t)}, I_u^{(t)})$ that satisfies the constraint, there are $f(\mathbf{w}_{g_u}^{(t)}, I_u^{(t)}) = \psi(\mathbf{w}_{g_u}^{(t)}, I_u^{(t)})$ and $\nabla f(\mathbf{w}_{g_u}, I_u) = \nabla \psi(\mathbf{w}_{g_u}, I_u)$. Therefore, constraint (18) is transformed into a second-order cone (SOC) constraint as

$$\Gamma_u \le \psi^{(t)}\left(\mathbf{w}_{g_u}, I_u\right). \tag{21}$$

Since the second derivative of (21) is more than zero, the transformed constraint has the convex property. Through iteration, the problem can be solved using the CVX tool and MATLAB. To substitute $\mathbf{w}$ with $\mathbf{w}^{(t)}$ in problem $\mathscr{P}_{1,1}$, parameters $I^{(t)}$, $\Gamma^{(t)}$, and $\psi^{(t)}$ can be obtained. In this way, all constraints are still satisfied since $\mathbf{w}^{(t)}$ is a feasible point. Further applying $I^{(t)}$, $\Gamma^{(t)}$, and $\psi^{(t)}$ to solve problem $\mathscr{P}_{1,1}$, the solution $\mathbf{w}^{(t+1)}$ can be further regarded as the input of the $(t+1)$-th iteration. In the iterative process, the construction of $\psi^{(t)}$ requires the same gradient value as the original function, the objective function value of the $t-th$ iteration must not be greater than that of the $(t+1)-th$ iteration, there holds $\sum_{u\in\mathscr{U}} Q_u^{(t+1)} \ge \sum_{u\in\mathscr{U}} Q_u^{(t)}$, and the iterative process is monotonically decreasing. In addition, considering that the system power is limited, the convergence of the iterative solution can be guaranteed according to the monotone boundedness theorem. When $\mathbf{w}^{(t)} = \mathbf{w}^{(t+1)}$ holds, the iteration converges and the optimal multicast group beamforming can be achieved. Generally, the convergence speed of the optimization problem is strongly related to the initial beamforming value in the first iteration, that is, $\mathbf{w}^{(0)}$. In what follows, an auxiliary problem is further formulated to guide how to determine $\mathbf{w}^{(0)}$:

$$\mathscr{P}_{1,2} \max_{\mathbf{w},\mathbf{v}} \delta, \tag{22a}$$

$$\text{s.t. } R_l^{\text{req}}\delta \le B\log_2\left(1+\varphi_u\right), \quad \forall u \in g, \forall g \in \mathscr{G}, \tag{22b}$$

$$(16d), (16e), (19), (21), \tag{22c}$$

where $\delta = \min_{u\in\mathscr{U}}\left\{R_{u,l}/R_l^{\text{req}}\right\}$, which is called the rate satisfaction. $\delta \ge 1$ indicates that the corresponding solution $\mathbf{w}^*$ satisfies (16c) and (16d). The solution approach of problem $\mathscr{P}_{1,2}$ is summarized in Algorithm 1. By setting $\mathbf{w}^0 = \mathbf{w}^*$, problem $\mathscr{P}_{1,1}$ can be solved iteratively by convex tools.

In step 3, Algorithm 1 solves a second-order cone programming problem, calculated by the interior point method, and its computational complexity can be expressed as $\mathcal{O}(((G+3)(A_kK+4))^{3.5})$. In addition, the computation cost in each iteration is bounded by $\mathcal{O}(T_1((G+3)(A_kK+4))^{3.5})$, where T_1 is the number of iterations in Algorithm 1.

In the video quality selection process, in order to degrade the interference from the signal of EL video to other groups, a heuristic algorithm is designed as follows. For a given user u, an auxiliary variable z_u is introduced to measure the cost of quality enhancement, which can further help to select the video quality level. The QoS constraint (16c) is then transformed into $(R_{g,l} - R_{u,l}) \le z_u$, where larger z_u results in more costs to enhance the video quality for user u, and vice versa. Therefore, we further formulate the following optimization problem to minimize the sum of z_u of each user:

$$\mathscr{P}_{1,3} \min_{\mathbf{w},\mathbf{v}} \sum_{u\in\mathscr{U}_{\text{BL}}} z_u, \tag{23a}$$

$$\text{s.t. } R_{g,f_l} \le B\log_2\frac{\left(1+\varphi_u\right)}{R_l^{\text{req}}} + z_u, \quad \forall u \in g, \forall g \in \mathscr{G}, \tag{23b}$$

$$0 \le z_u, \quad \forall u \in \mathscr{U}, \tag{23c}$$

$$(16b), (16c), (16e), (19), (21), \tag{23d}$$

where $\mathscr{U}_{\text{BL}}$ represents the user who initially requested the low-quality multimedia content. In (23c), the corresponding enhancement cost of IoT users requesting high-quality multimedia content are all set to 0. By solving problem $\mathscr{P}_{1,3}$, the video quality enhancement cost set $\mathbf{z}$ can be obtained, which is further used to select the user in $\mathscr{U}_{BL}$ with the smallest $\mathbf{z}$ to provide EL videos to improve the QoE performance.

Based on Algorithm 1, the complexity of Algorithm 2 is mainly determined by the user selection process in step 5 and the maximization process of QoE in step 10. T_2 and T_3 represent the iteration times of the above two steps, respectively, and the computational complexity of Algorithm 2 can be expressed as $\mathcal{O}((T_1 + U_{BL}(T_2+T_3))((G+3)(A_kK+4))^{3.5})$, where U_{BL} is the number of IoT users requesting low-quality video.

5. Simulation Results

In this section, simulation results are presented and discussed. Without other highlights, the simulation parameters are set as follows. The simulation scenario includes 7 SBSs with $A_k = 2$ that are distributed on the vertex and center of a regular hexagon with a side length of 100 m, and 40 IoT users and 3 eavesdroppers that are randomly distributed on a circle with a radius of 200 m. The channel gain from SBS k to the user u is defined as $\mathbf{h}_{k,u} = \sqrt{1/(1+d_{k,u}/d_0)^{\rho}}\,\tilde{h}_{k,u}$, where the path loss factor is set as $\rho = 3$, the standard distance is set as $d_0 = 50m$, and the noise power is set as 1. According to the transmission unit size requirements in IEEE 802.11, we set the EL data to 2 Mbit, corresponding to 1080 p definition video, and set the BL data to 1 Mbit, corresponding to 720 p video. The remaining simulation parameters are shown in Table 1. The network performance is evaluated in terms of average network rate, multicast group rate, and network QoE.

Figure 3 shows the convergence performance of Algorithm 1 under four different random channel realizations

Input: Channel condition and QoS threshold of multi-quality video.
Output: Optimal beamforming vector and rate satisfaction.
Step:
(1) Set the $t = 0$, beamforming vector $\mathbf{w}^{(0)}, \mathbf{v}^{(0)}$ satisfies constraint (16d);
(2) Calculate $I_u^{(t)}$ of each user as follows:

$$I_u^{(t)} = \sum_{l>l_{\text{req}}} |\mathbf{h}_u^H \mathbf{w}_{\mathbf{g},f_l}^{(\mathbf{t})}|^2 + \sum_{i \in G \setminus \{g\}} |\mathbf{h}_u^H \mathbf{w}_i^{(\mathbf{t})}|^2 + \sigma^2, \ \forall u \in g, \ \forall g \in \mathcal{G}$$

(3) Solve $\mathcal{P}_{1,2}$;
(4) Update $\mathbf{w}^{(\mathbf{t}+1)}, \mathbf{v}^{(\mathbf{t}+1)}, \mathbf{r}^{(\mathbf{t}+1)}, \Gamma^{(\mathbf{t}+1)}$;
(5) $\delta^* \leftarrow \delta^{(t)}$;
(6) $\leftarrow \mathbf{v}^{(\mathbf{t})}$;
(7) $\mathbf{w}^* \leftarrow \mathbf{w}^{(\mathbf{t})}$;
(8) If $\delta^{(\mathbf{t}+1)} - \delta^{(\mathbf{t})} \le \epsilon$ iteration stop. Otherwise, set $t \leftarrow t + 1$ and go to Step 2.
(9) Output $\mathbf{w}^*$, $\mathbf{v}^*$ and δ^*.

Algorithm 1: Maximize rate satisfaction δ algorithm.

Input: Channel condition, cache transmission delay, delay satisfaction factor and QoS threshold of multi-quality video.
Output: Optimal beamforming vector, optimal version selection set and maximum network QoE.
Step:
(1) Set the $t = 0$, feasible beamforming vector $\mathbf{w}^{(0)}, \mathbf{v}^{(0)}$, only require low-quality video user set $\mathcal{U}_{BL}$;
(2) Calculate the optimal user rate satisfaction δ^* and optimal beamforming vector $\mathbf{w}^*, \mathbf{v}^*$ by solving $\mathcal{P}_{1,2}$.
(3) If $\delta^* > 1$
(4) Calculate current optimal QoE Q^* by solve $\mathcal{P}_{1,1}$;
(5) Calculate the enhancement cost set $\mathbf{z}$ by solve $\mathcal{P}_{1,3}$;
(6) Obtain the user index $\mathbf{u}^*$ corresponding to the minimum value in $\mathbf{z}$;
(7) $\mathcal{U}_{BL} \leftarrow \mathcal{U}_{BL} \setminus u^*$;
(8) $l_{u^*} \leftarrow 2$;
(9) Calculate network QoE $Q\prime$ and beamforming vector $\mathbf{w}\prime$ by solve $\mathcal{P}_{1,1}$;
(10) If the problem is unsolved or the network QoE drops $(Q' \le Q^*)$
(11) Break.
(12) Else
(13) Update $Q^* \leftarrow Q'$, $\mathbf{w}^* \leftarrow \mathbf{w}\prime$, $\mathbf{v}^* \leftarrow \mathbf{v}\prime$, $\mathbf{l}^* \leftarrow \mathbf{l}'$;
(14) End if
(15) If $\mathcal{U}_{BL} = \varnothing$ iteration stop. Otherwise, set $t \leftarrow t + 1$ and go to Step 5.
(16) Else
(17) The solution is not feasible, the current resources cannot satisfy all users.
(18) End if
(19) Output $\mathbf{w}^*$, $\mathbf{v}^*$, $\mathbf{l}^*$ and Q^*.

Algorithm 2: Maximize QoE algorithm.

Table 1: Simulation parameters.

Symbol	Value
Transmission power of SBS	40 dBm
Bandwidth	10 MHz
Number of video contents	100
Cache ratio in each MEC	0.6
Video block size	1 Mbit (BL), 2 Mbit (EL)
Maximum tolerance SINR of eavesdropper	−10 dB
Delay satisfaction factor	0.2
Marginal effect factor	5
Transmission delay	0.2 (core network)/0.05 (MEC)
Zipf coefficient	1

(RCR) from the average transmission rate. All the simulation results are under the condition of secure communication. Under the interference of artificial noise, the eavesdropper cannot obtain the content at the lowest decoding rate. The beamforming of the multicast group is initialized to $v_{k,a}^H, w_{k,g,a}^H = \text{rand} \times \sqrt{P_k / ((G+1) \times A_k)}$, where $ran\,d$ is a random factor within [0, 1]. It is observed that, after the first 5 iterations, the average rate in four different RCR can reach the 96.72% level obtained in the 20th iteration, which verifies the convergence performance of Algorithm 1. Moreover, the convergence performance of Algorithm 1 proposed is insensitive to the channel conditions, which verifies the robustness of the proposed scheme.

In Figure 4, four multicast groups are randomly selected, and the rate performance of multicast groups with different delay costs and intra group video request quality is compared.

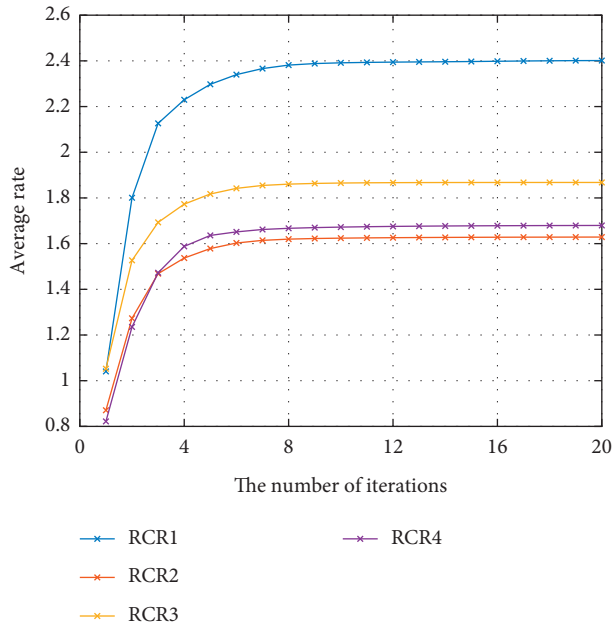

Figure 3: Average rate under different RCR.

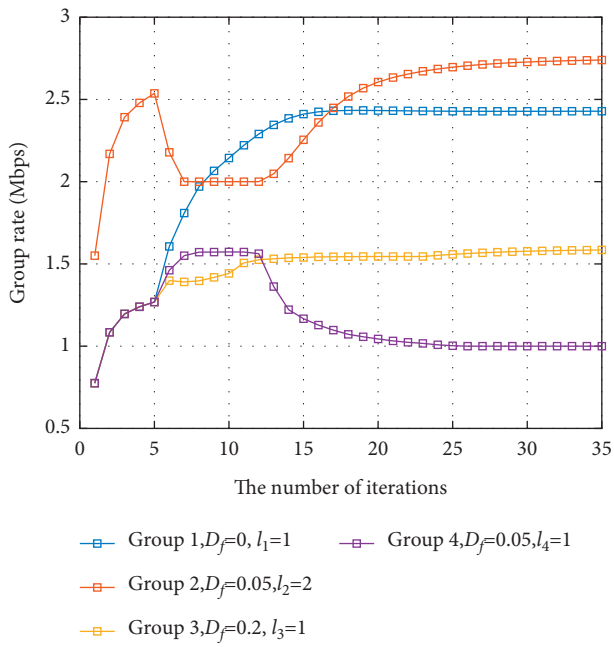

Figure 4: Multicast group rate performance in different groups.

In the first five iterations, the proposed scheme tries to find out the appropriate initial parameters for subsequent optimization. After that, the QoE is maximized iteratively. First, group 2 and group 4 changed significantly near the fifth and 12th iterations. This is because they have high video quality requirements. Allocating additional resources helps improve the previous item of QoE, which is more advantageous than the optimization of delay. For group 4, because there are few users in the group and the channel conditions are poor, it is only necessary to maintain a minimum QoS requirement at the end to save limited resources. Group 1 selects the

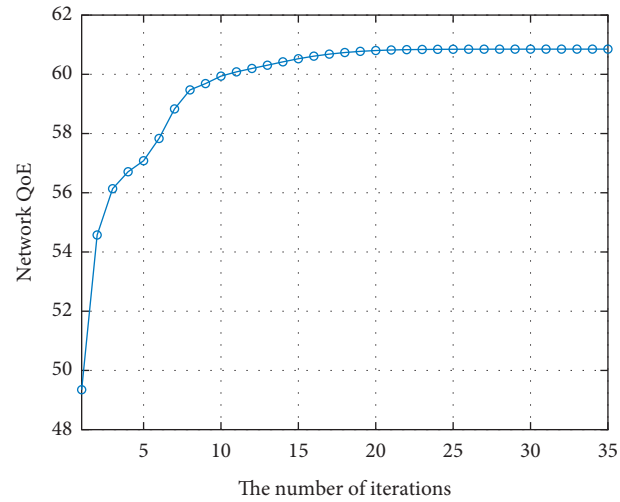

Figure 5: Trajectory of network QoE weighted sum.

multicast group with the most members in the group and allocates resources to it, helping to improve the efficiency of beamforming and optimize network QoE. For group 3, although the delay cost $D_f = 0.2$, this makes the QoE delay part have a large room for improvement, so it is still in a slow rising state. The above shows that, with the optimization of network QoE, the resource scheduling of different groups will still be adjusted adaptively according to iteration.

Figure 5 shows the trend of network QoE weighted sum. Although the allocation of resources between different multicast groups fluctuates in Figure 4, the QoE of the whole network increases monotonically. The allocation of resources will only change the growth rate of network QoE, which shows the effectiveness of the proposed algorithm for network QoE optimization.

In Figure 6, the network QoE performance was depicted under different cache capabilities of SBSs. In order to validate the effectiveness of the proposed scheme, three baseline schemes, namely, QoE-Max, QoS-Only, and Unicast, are introduced. The QoE-Max scheme maximizes the network QoE based on multicast group beamforming optimization only, without considering the enhancement of video quality. The QoS-Only scheme only guarantees the basic QoS requirement of video quality of each user. The Unicast-QoE algorithm will transmit multimedia content to users through unicast cooperative transmission, producing additional interference. The cache ratio in Figure 5 indicates the proportion of the content library that an MEC can cache. The network QoE of the four schemes shows an increasing trend, but there are still some fluctuations due to the uncertainty of users' requests for popular content. In addition, as the MEC cache ratio increases, more space is used to store infrequently used content. Furthermore, compared with the Unicast-QoE algorithm with the worst performance, the proposed algorithm improves the network QoE by 23.59% on average. And compared with the QoE-Max algorithm, which ignores video quality enhancement, the proposed algorithm improves by 6.68% because the additional video quality brings users a better experience. During the experiment, the Unicast-QoE algorithm needs to design

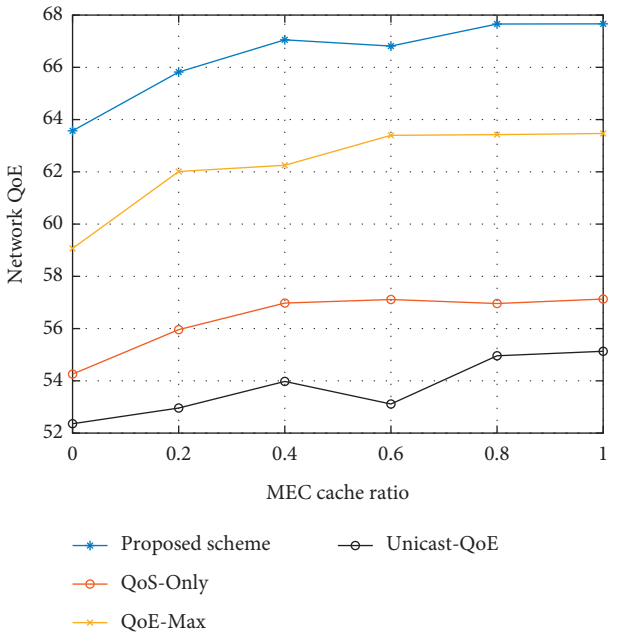

Figure 6: Network QoE performance under different cache capabilities of SBSs.

beamforming separately for all users, which will produce additional resource consumption and interference. In addition, among the four algorithms, the proposed algorithm also has the best performance because the proposed algorithm is optimized for video quality and beamforming simultaneously.

6. Conclusions

In this paper, a QoE-aware video delivery scheme was studied in multimedia IoT network with potential eavesdroppers. A joint video quality selection and multicast group beamforming scheme were proposed to maximize the network QoE while preventing data eavesdropping as much as possible. Extensive simulation results validated the effectiveness of the proposed scheme in terms of user satisfaction, multicast group rate, and network QoE [42].

Data Availability

The relevant data used to support the results of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (61871062, 61771082, 61901071, and U20A20157), the Science and Technology Research Program of Chongqing Municipal Education Commission (KJQN202000626), Natural Science Foundation of Chongqing, China (cstc2020jcyj-zdxmX0024), and University Innovation Research Group of Chongqing (CXQT20017).

References

[1] B. Jedari, G. Premsankar, G. Illahi, M. D. Francesco, A. Mehrabi, and A. Ylä-Jääski, "Video caching, analytics, and delivery at the wireless edge: a survey and future directions," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 431–471, 2021.

[2] C. Long, Y. Cao, T. Jiang, and Q. Zhang, "Edge computing framework for cooperative video processing in multimedia IoT systems," *IEEE Transactions on Multimedia*, vol. 20, no. 5, pp. 1126–1139, 2018.

[3] A. Nauman, Y. A. Qadri, M. Amjad, Y. B. Zikria, M. K. Afzal, and S. W. Kim, "Multimedia Internet of things: a comprehensive survey," *IEEE Access*, vol. 8, pp. 8202–8250, 2020.

[4] Ericsson, "Ericsson mobility report," https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/november-2021.

[5] V. Cisco, "networking index: Global mobile Data Traffic Forecast Update," https://www.cisco.com.

[6] S. He and W. Wang, "Multimedia upstreaming cournot game in non-orthogonal multiple access Internet of things," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 398–408, 2020.

[7] T. X. Tran, A. Hajisami, P. Pandey, and D. Pompili, "Collaborative mobile edge computing in 5G networks: new paradigms, scenarios, and challenges," *IEEE Communications Magazine*, vol. 55, no. 4, pp. 54–61, 2017.

[8] Z. Li, Y. Zhou, D. Wu, T. Tang, and R. Wang, "Fairness-aware federated learning with unreliable links in resource-constrained Internet of things," *IEEE Internet of Things Journal*, p. 1, 2022.

[9] Y. Fallah, H. Mansour, S. Khan, P. Nasiopoulos, and H. Alnuweiri, "A link adaptation scheme for efficient transmission of H.264 scalable video over multirate WLANs," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 7, pp. 875–887, 2008.

[10] H. Zhu, Y. Cao, T. Jiang, and Q. Zhang, "Scalable NOMA multicast for SVC streams in cellular networks," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6339–6352, 2018.

[11] S. Pizzi, C. Suraci, A. Iera, A. Molinaro, and G. Araniti, "A sidelink-aided approach for secure multicast service delivery: from human-oriented multimedia traffic to machine type communications," *IEEE Transactions on Broadcasting*, vol. 67, no. 1, pp. 313–323, 2021.

[12] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, 2008.

[13] J. Xiong, R. Bi, M. Zhao, J. Guo, and Q. Yang, "Edge-assisted privacy-preserving raw data sharing framework for connected autonomous vehicles," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 24–30, 2020.

[14] Y. Tian, T. Li, J. Xiong, M. Z. A. Bhuiyan, J. Ma, and C. Peng, "A blockchain-based machine learning framework for edge services in IIoT," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1918–1929, 2022.

[15] D. Hwang, J. Yang, K. Kwon, J. Joung, and H.-K. Song, "Equalization-based beamforming for secure multicasting in multicast wiretap channels," *IEEE Access*, vol. 9, pp. 33826–33835, 2021.

[16] H. Luo, Q. Li, L. Yang, and J. Qin, "A fast algorithm for fractional QCQP and applications to secure beamforming in cognitive nonorthogonal multiple access networks," *IEEE Transactions on Signal Processing*, vol. 69, pp. 6237–6250, 2021.

[17] P. Huang, Y. Hao, T. Lv, J. Xing, J. Yang, and P. T. Mathiopoulos, "Secure beamforming design in relay-assisted Internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6453–6464, 2019.

[18] P. V. Tuan, T. Trung Duy, and I. Koo, *Multiuser MISO Beamforming Design for Balancing the Received Powers in Secure Cognitive Radio Networks*, in *Proceesings of the in 2018 IEEE Seventh International Conference on Communications and Electronics (ICCE)*, pp. 39–43, Hue, Vietnam, July2018.

[19] M. R. A. Khandaker and K. K. Wong, "Masked beamforming in the presence of energy-harvesting eavesdroppers," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 1, pp. 40–54, 2015.

[20] S. Fan and J. Xu, *Single-Group Multicast Secure Beamforming via Learning the Eavesdropper's Channel Correlation*, in *Proceedings of the in 2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, Dublin, Ireland, July2020.

[21] D. W. K. Ng, R. Schober, and H. Alnuweiri, "Power efficient MISO beamforming for secure layered transmission," in *Proceedings of the in 2014 IEEE Wireless Communications and Networking Conference*, pp. 422–427, Istanbul, Turkey, April2014.

[22] Z. Li, X. Gao, Q. Li, J. Guo, and B. Yang, "Edge caching enhancement for industrial Internet: a recommendation-aided approach," *IEEE Internet of Things Journal*, p. 1, 2022.

[23] J. Ma, L. Liu, H. Song, R. Shafin, B. Shang, and P. Fan, "Scalable video transmission in cache-aided device-to-device networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 4247–4261, 2020.

[24] Y. Hong, X. Jing, and H. Gao, "Programmable weight phased-array transmission for secure millimeter-wave wireless communications," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 399–413, 2018.

[25] J. Xiong, X. Chen, Q. Yang, L. Chen, and Z. Yao, "A task-oriented user selection incentive mechanism in edge-aided mobile crowdsensing," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2347–2360, 2020.

[26] Q. Li, C. Li, and J. Lin, "Constant modulus secure beamforming for multicast massive MIMO wiretap channels," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 264–275, 2020.

[27] Z. Chu, H. Xing, M. Johnston, and S. Le Goff, "Secrecy rate optimizations for a MISO secrecy channel with multiple multiantenna eavesdroppers," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 283–297, 2016.

[28] D. Huang, X. Tao, C. Jiang, S. Cui, and J. Lu, "Trace-driven QoE-aware proactive caching for mobile video streaming in metropolis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 62–76, 2020.

[29] X. He, K. Wang, and W. Xu, "QoE-driven content-centric caching with deep reinforcement learning in edge-enabled IoT," *IEEE Computational Intelligence Magazine*, vol. 14, no. 4, pp. 12–20, 2019.

[30] L. Liu, Y. Zhou, J. Yuan, W. Zhuang, and Y. Wang, "Economically optimal MS association for multimedia content delivery in cache-enabled heterogeneous cloud radio access networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 7, pp. 1584–1593, 2019.

[31] N.-S. Vo, T.-M. Phan, M.-P. Bui, X.-K. Dang, N. T. Viet, and C. Yin, "Social-aware spectrum sharing and caching helper selection strategy optimized multicast video streaming in dense D2D 5G networks," *IEEE Systems Journal*, vol. 15, no. 3, pp. 3480–3491, Sept, 2021.

[32] Y. Zhou, F. R. Yu, J. Chen, and Y. Kuo, "Cache-aware multicast beamforming design for multicell multigroup multicast," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 11681–11693, 2018.

[33] X. Zhang, T. Lv, Y. Ren, W. Ni, N. C. Beaulieu, and Y. J. Guo, "Economical caching for scalable videos in cache-enabled heterogeneous networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 7, pp. 1608–1621, 2019.

[34] D. Jiang and Y. Cui, "Analysis and optimization of caching and multicasting for multi-quality videos in large-scale wireless networks," *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 4913–4927, 2019.

[35] Z. Xu, Y. Cao, W. Wang, T. Jiang, and Q. Zhang, "Incentive mechanism for cooperative scalable video coding (SVC) multicast based on contract theory," *IEEE Transactions on Multimedia*, vol. 22, no. 2, pp. 445–458, 2020.

[36] A. De La Fuente, J. J. Escudero-Garzás, and A. García-Armada, "Radio resource allocation for multicast services based on multiple video layers," *IEEE Transactions on Broadcasting*, vol. 64, no. 3, pp. 695–708, 2018.

[37] C. Guo, Y. Cui, D. W. K. Ng, and Z. Liu, "Multi-quality multicast beamforming with scalable video coding," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5662–5677, 2018.

[38] Y. Jin, Y. Wen, and C. Westphal, "Optimal transcoding and caching for adaptive streaming in media cloud: an analytical approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 12, pp. 1914–1925, 2015.

[39] S. Chen, Z. Yao, X. Jiang, J. Yang, and L. Hanzo, "Multi-agent deep reinforcement learning-based cooperative edge caching for ultra-dense next-generation networks," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2441–2456, 2021.

[40] S.-E. Elayoubi and J. Roberts, "Performance and cost effectiveness of caching in mobile access networks," in *Proceedings of the 2nd ACM Conference on Information-Centric Networking*, pp. 79–88, 2015.

[41] D. Zegarra Rodríguez, R. Lopes Rosa, E. Costa Alfaia, J. Issy Abrahão, and G. Bressan, "Video quality metric for streaming service using DASH standard," *IEEE Transactions on Broadcasting*, vol. 62, no. 3, pp. 628–639, 2016.

[42] R. O. Afolabi, A. Dadlani, and K. Kim, "Multicast scheduling and resource allocation algorithms for OFDMA-based systems: a survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 240–254, 2013.