# EVB360: Efficient 360-Degree Video Broadcast in Next-Generation Cellular Networks

Chetna Singhal and Jacob Chakareski

***Abstract*—We propose EVB360, an efficient 360° video broadcast scheme for next-generation NR-MBMS (5G New Radio MBMS) cellular networks. EVB360 employs a dynamic broadcast area definition, adaptive 360-degree video encoding, and tiled transmission using optimized transmission parameters. We define broadcast areas by using K-means clustering and then proportionally assign resources to broadcast adaptively the encoded tiles. EVB360 ensures high level of immersive video broadcast streaming quality to a higher number of served users. Adaptive encoding of 360° video tiles and radio resource (modulation-coding, subframe) allocation is governed by user viewing angle, content request, position, and received signal strength. Our experiments show that our solution enables considerable gains in viewport PSNR (57.44%) and number of served users (72.19%), over a recent state-of-the-art method VRCAST.**

***Index Terms*—New-Radio Multimedia Broadcast Multicast Services (NR-MBMS), Multicast-Broadcast Single Frequency Networks (MBSFN), Immersive video broadcast.**

## I. Introduction

Immersive 360° video streaming is increasingly used in diverse applications such as virtual reality, gaming, and entertainment [1]. In the immersive environment, a viewer changes her viewing direction and the content is accordingly rendered. However, streaming such content requires very high bandwidth and is challenging [1]. A 360-degree video can be spatially divided into small portions called tiles that can be encoded at different quality levels. This has enabled tiling-based viewport-adaptive 360° video streaming, where tiles are delivered to clients based on their viewing direction and network conditions. Concretely, the viewport tiles are delivered at high quality and the rest are sent at lower quality [2].

Mobile (and digital) television broadcast over wireless network is gaining popularity and it comprises of on-demand content streaming and multimedia broadcast to heterogeneous customers on their smart devices like TVs, phones, and car-infotainment systems [3]. Substantial network resource are used for streaming on-demand content to mobile users using unicast connections. Such a system does not scale to scenarios with high user density. Multicast/broadcast offers a scalable solution with challenges like user-centric adaptation, ensuring smooth quality, and incorporating user interactivity. This motivates the investigation of resource efficient 360° video broadcast solutions for future large/dense cellular networks.

The Further evolved Multimedia Broadcast Multicast Services (FeMBMS) standard, defined in 3GPP Release 14, provides mechanism to support multimedia streaming over cellular network [4]. With the advent of next generation communication standards, efforts are being made to define NR-MBMS (5G New Radio MBMS) and advance FeMBMS [5]. Herein, *synchronization area* consists of 3GPP 5G Next Generation base stations (gNodeB) that are synchronized in time. According to the standard MBSFN (Multicast Broadcast Single Frequency Network) consists of a set of gNodeBs that broadcast the same content simultaneously over the same frequency resource. The user equipment (UE) thereby receives a better quality by combining the received signal from the gNBs in the given synchronization area. The unicast and broadcast transmissions happen together in every cell sharing the cell capacity.

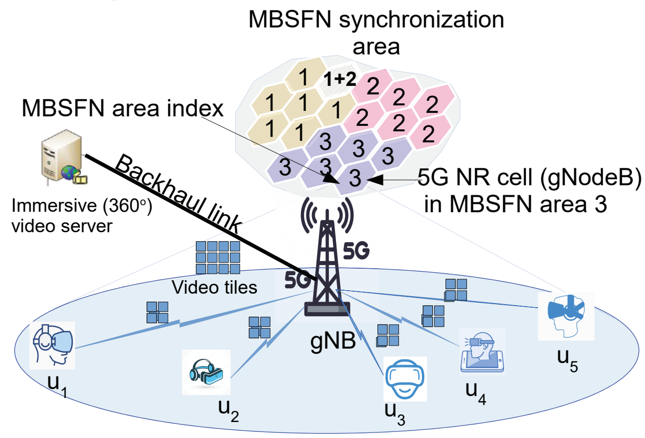


Fig. 1. Sample scenario of EVB360 in 5G NR networks.

We investigate an efficient 360° video broadcast solution, EVB360, that can be deployed in future cellular networks and is suitable for large and dense network scenarios. Fig. 1 shows a sample scenario of EVB360 in 5G-NR cellular networks. Some gNodeBs can be a part of multiple MBSFNs resulting in an overlap. We define a conjoin set as MBSFN areas with at least one gNodeB in common. These MBSFNs in such an arrangement (in conjoin set) need to use disjoint frequency allocation for broadcast. Fig. 1, the example shows $\{1, 2\}$ is a conjoin set. The tiles of an immersive video are sent from the multimedia server to gNodeB (gNB) using the backhaul link. The gNB thereafter broadcast the tiles to the heterogeneous users based on their viewing angle. In the given example, users $u_1 - u_6$ are all accessing the $360^o$ video at different viewing angles and are receiving the corresponding tiles.

Various characteristics like demographic, socio-economic, and instanceof the day, impact the pattern of TV content consumption by users [6]. Hence, clustering is a logical method to dynamically define MBSFN areas and optimally perform NR-MBMS radio resource allocation for varied viewing angles and content popularity in a given MBSFN area. EVB360 is a novel scheme that forms MBSFN areas based on users' requests, position, received signal strength, and viewing angle. Then, given the set of MBSFNs, we aim to maximize the immersive quality delivered to the users by adaptive 360° video tile encoding and efficient radio resource allocation. Our experimental results demonstrate considerable gains in viewport PSNR and number of served users, over a recent state-of-the-art method VRCAST.

## II. Related Work

Viewport-adaptive 360° video streaming has been studied in [1, 2] via the design of efficient 360° video representations and resource allocation methods. Live scalable 360° video network multicast has been investigated in [7] via rate-distortion optimization and user viewport prediction. The reference method we consider in our experiments is known as VRCAST and has been studied in [8] for mobile streaming of live 360° videos. It considers user grouping, adaptive resource allocation, and tile-quality selection. However, it focuses on live multicast and does not address multiple parallel broadcast sessions for larger and denser cellular networks, where a unilateral channel condition based grouping for a limited number of users fails. Our approach aims to fill this gap.

Clustering based allocation of MBMS resource to heterogeneous users [9, 10] and dynamic definition of MBSFN areas for optimized multicast [11] has been studied previously. Efficient encoding of immersive multimedia based on multi-criteria clustering for efficient 360° videobroadcast is yet to be addressed. Adaptive immersive tile-based multimedia encoding, and optimized NR-MBMS resource allocation, as considered in this paper, represents a novel topic that has not been studied before.

## III. EVB360 System Architecture and Components

EVB360 system architecture is illustrated in Fig. 2. The gNodeBs serve the respective User Equipments (UEs) based on the notified device capability, and immersive content request, and viewing angle. This information is forwarded by gNB gateway (NR-MBMS) and MCE (Multicell/multicast Coordination Entity). The MCE coordinates transmission between multiple cells in MBSFN area. The MCE and Gateway adaptively and dynamically define MBSFN areas and efficiently allocate the broadcast resources. The server for immersive multimedia content encodes 360° video tiles using efficient selection of quantization parameter based on content requests, viewing angles, and radio resource constraints.

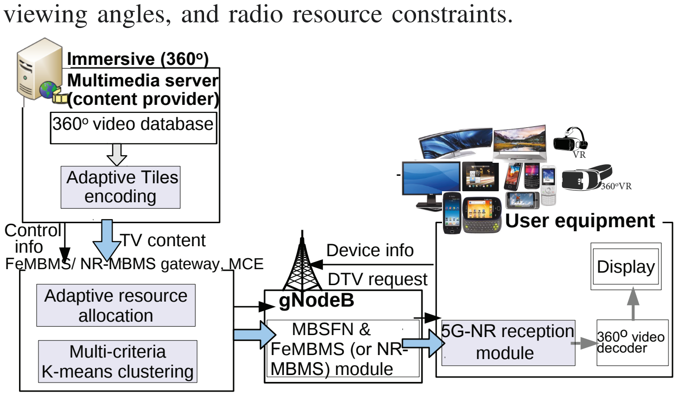


Fig. 2. Proposed EVB360 system architecture.

### A. NR-MBMS resource allocation

The gNodeB dynamically selects the MCS (modulation and coding scheme) based on CQI (channel quality indicator) of the users being served by it. The gNodeB broadcasts the tiles to the users using the selected MCS. The radio resource in 5G downlink are arranged into resource blocks (RBs). Each RB includes twelve frequency subcarriers consecutively with a duration of 0.0625-1 ms governed by the bandwidth. A periodic scheduling of radio resource supports the dynamic allocation in a cellular network consisting of mobile users. The latest broadcast standard, NR-MBMS, supports allocating all subframes (100% radio resource) for broadcasting content [4, 5]. The poorest user (with least CQI) governs the resource allocation usually in the broadcast network. In order to broadcast content on a portion of resource (in terms of subframes), i.e. $\sigma$, $0 < \sigma \leq 1$, the overall capacity is, $C(\sigma, m) = B \cdot \sigma \cdot e_m$. Here, $B$ MHZ is the bandwidth of the channel, $m$ is the dynamically selected MCS, and $e_m$ is the spectral efficiency according the NR-MBMS standard technology.

## IV. EVB360: MBSFN Area Formation, Tile Quality Adaptation, and Resource Allocation

### A. MBSFN area formation: Multi-criteria K-means clustering

The MBSFN areas and the content to be broadcasted in it are defined by the MCE. This is governed by the CQI, device capabilities, and 360° content request by UEs in the area that report these factors in a periodical manner. Hence, only the popular content is broadcast on request in an area instead of broadcasting everything. The multi-criteria clustering used in EVB360 is based on user $u_i$ information: SINR ($\gamma_i$) correlated to CQI, 360° content requested $p_i$, and gNodeB location ($\vartheta_i$) that the user is associated with. The result of clustering is formation of groups of UEs (in nearby cells) with similar 360° video content interest and spatial position. These groups are then mapped to gNodeB sets defining the MBSFN areas. This approach ensures MBSFN area formation that dynamically serves the diverse user requests in an efficient manner.

We begin with a set of 360° content set, $\mathcal{P}_k \subseteq \mathcal{P}$, to be broadcast in every MBSFN area $k$. This contains all content requested by users in a cluster. The efficient content set to be broadcast in the MBSFN area is determined by the approach given in Section IV-B which combinedly considers the availability of broadcast radio resource and quantization parameter options for various 360° video tiles.

Multi-dimensional clustering is an NP hard problem [12]. K-Means is a well known and a popular algorithm for clustering. Its implementation to classify data is simple and easy [13]. Additionally it is fast with few computations and has linear complexity O(n). EVB360 uses heuristic Lloyds K-means algorithm [14] for formation of $K$ MBSFN areas by performing multi-criteria clustering of diverse users. We begin with $K = 1$ and increase $K$ by one in each iteration. When the performance is lower than previous iteration that gives us the optimal value of $K$. It is important to also consider that the NR-MBMS standard limits the participation of gNodeB in at max eight MBSFNs. In order to prevent violation of this constraint (i.e. for higher values of $K$) we limit gNodeB association to a maximum of eight clusters that has the highest proportion of its users. The Algorithm 1 summarizes the procedure of MBSFN area formation for a given $K$. K-means++ method selects the cluster centroids initially [15]. The algorithm gives the $K$ MBSFN areas, set of 360° video $\mathcal{P}_k$, the overlap conjoined sets for each 360° broadcast content. The overlap conjoined set of 360° video program $p$ is a set of adjacent MBSFN areas in the spatially overlapping (partial or complete) region that broadcast $p$.

---

**Algorithm 1** MBSFN Area formation, given $K$

---

**Input:** $u_i$, $1 \leq i \leq N$, $K$

1) Uniformly select first centroid $c_1$ from $u_i$, $1 \leq i \leq N$.
2) **for each** $i = 1$ to $N$
   Calculate $d^2(u_i, c_1) = (\vartheta_i - \vartheta_{c_1})^2 + (\gamma_i - \gamma_{c_1})^2$
3) **Select the second centroid** $c_2 = u_m$, with a probability, $\frac{d^2(u_m,c_1)}{\sum_{i=1}^{N} d^2(u_i,c_1)}$
4) **Select remaining centroids,** $c_k$**:**
   **for each** $k = 3$ to $K$
   Select centroids $c_k = u_m$ at random with probability $\frac{d^2(u_m,c_k)}{\sum_{x_i \in C_k} d^2(u_i,c_k)}$, where $C_k$ is a set of all $u_i$ closest to $c_k$
5) **Formation of** $K$ **clusters**
   **for each** $i = 1$ to $N$
   $k^\star = \min_k d^2(u_i, c_k)$
   Assign user $i$ to cluster $k^\star$
6) **Cluster update**
   **for each** $k = 1$ to $K$
   Average measure: $\left(\sum_{i=1}^{N} d^2(u_i, c_k)\right)/N$
   Reassign centroid $c_k$ to decrease average measure
7) Reiterate step 5-6 until cluster assignments are unchanged
8) A gNB belongs to MBSFN area $k$ if more than two of its UEs are in $k$
9) $\mathcal{P}_k$ is the set of360° video programs requested by users in $k$ and will be broadcast in $k$
10) The overlap conjoined set for $p$ include the MBSFN areas that are broadcasting $p$ and overlap spatially

**Output:** $\{\mathcal{P}_k\}_k$, Overlap conjoined sets, $\mathcal{O}$, $K$ MBSFN areas

---

*B. Adaptive immersive tile-based 360° multimedia encoding and resource allocation*

After obtaining the $K$ MBSFN areas, 360° program set $\{\mathcal{P}_k\}_k$, and the conjoined overlap sets, we further refine the broadcasting framework by outline the following in our proposed EVB360 scheme: (i) the 360° programs and their tiles to be broadcast, given the broadcast radio resource constraint, (ii) quantization parameter to adaptively encode each tile of the 360° program, and (iii) resource blocks and MCS level that to broadcast each tile of each 360° video.

Fig. 3 depicts the subframe allocation, $\sigma^k_{p,\tau}$, and MCS selection, $m^k_{p,\tau}$, for broadcasting the immersive video tiles $(1 \leq \tau \leq T_p)$ of the 360° programs $(1 \leq p \leq P_k)$ in MBSFN area $k$ $(1 \leq k \leq K)$. Our proposed adaptive immersive tile-based multimedia encoding and efficient radio resource

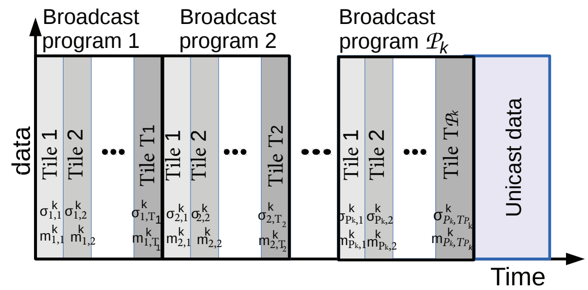


Fig. 3. Logical representation of resource allocation for EVB360.

allocation aims to maximize the overall video quality and the number of served users by taking into account the diverse user requests, their viewing angles, and channel conditions. We consider a single content request by each user at any given time for clarity in our framework discussion. However the approach can be extended to multiple requests as well. We consider $T_p$ denotes the total number of tiles of immersive program $p$. The user 360° multimedia broadcast quality is defined as follows.

**Definition 1.** *360° immersive multimedia broadcast service quality $\mathcal{Q}_i$ of a user $u_i$ that requests program p, is the effective quality of the tiles $\tau_i$ it receives based on its viewing angle and its channel conditions. It is defined as:*

$$\mathcal{Q}_i = \begin{cases} Q(q_{p,\tau_i}), & \text{if } \gamma_i \geq SINR\_thr(m^p_{\tau_i}) \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

*where $q_{p,\tau_i}$ is the quantization parameter used for encoding tile $\tau_i$ of program p, $SINR\_thr$ is the SINR threshold of the MCS $m^k_{p,\tau_i}$ allocated for transmission of corresponding tile.*

During the resource scheduling period, the channel characteristics for a user is nearly stationary and its CQI is unchanged [9]. During this interval the broadcast of all tiles $(T_p)$ in a group of frames of the requested $360^o$ program takes place. Hence, if the SINR of a user in MBSFN area $k$ is greater than the threshold for MCS level $m^k_{p,\tau_i}$ assigned to immersive video tile $\tau_i$ of $p$ in the MBSFN area, it can receive $\tau_i$ successfully. The broadcast service quality experienced by user $u_i$ is non-zero only if receives acceptable video quality levels corresponding to rate supported by MCS $m^k_{p,\tau_i}$.

We formulate the optimization problem to be performed by each MBSFN area $k$ $(1 \leq k \leq K)$ for each 360° program $p$ in $\mathcal{P}$, considering the quantization parameter $q^k_{p,\tau}$, $1 \leq \tau \leq T_p$, for adaptive 360° multimedia encoding and optimal broadcast radio resource allocation. We formulate two dual objectives: (i) EVB360($Q_{max}$) maximizes the immersive multimedia broadcast service quality for the served users in the system, and (ii) EVB360($N_{max}$) maximizes the number of served users $N$ in the system with a guaranteed minimum acceptable video quality, subject to multiple system constraints. Specifically, the following must hold valid:

(a) the total broadcast transmission rate requirement should be bound by the resource capacity at each gNB;
(b) the broadcasting resource (time-frequency) allocated is upper-bound to 1, at each gNodeB;
(c) same quantization level $(q^{(o)}_{p,\tau})$, fraction of subframes

($\sigma_{p,\tau}^{(o)}$) to each video tile ($\tau$, $1 \le \tau \le T_p$), and MCS ($m_{p,\tau}^{(o)}$) should be assigned by all MBSFNs in the same overlap conjoined set $\mathcal{O}$ for program $p$;

(d) the dynamically selected MCS for broadcasting the 360° video tiles should be from the set of values allowed by the NR-MBMS standard technology.

We formulate the respective optimization problems below, where the constraints (2a)-(2d) capture the above conditions.

$$\textbf{EVB360(Q}_{\textbf{max}}\textbf{):} \max_{[\mathbf{q}^1,\ldots,\mathbf{q}^K]} \sum_{i=1}^{N} \mathcal{Q}_i, \tag{2}$$

$$\text{s.t.:} \sum_{p=1}^{|\mathcal{P}|}\sum_{\tau=1}^{T_p} \mathbb{1}_{p,j} R(q_{p,\tau}^{(o)}) \le \sum_{p=1}^{|\mathcal{P}|}\sum_{\tau=1}^{T_p} \mathbb{1}_{p,j} C(\sigma_{p,\tau}^{(o)}, m_{p,\tau}^{(o)}), \tag{2a}$$

$$\sum_{p=1}^{|\mathcal{P}|}\sum_{\tau=1}^{T_p} \mathbb{1}_{p,j}\sigma_{p,\tau}^{(o)} \le 1, \ \forall \text{gNB } j \tag{2b}$$

$$q_{p,\tau}^{k} = q_{p,\tau}^{(o)};\ \sigma_{p,\tau}^{k} = \sigma_{p,\tau}^{(o)};\ m_{p,\tau}^{k} = m_{p,\tau}^{(o)}, \quad \forall \tau,\ p \text{ broadcast in } k \in \mathcal{O} \tag{2c}$$

$$1 \le m_{p,\tau}^{k} \le 15, \forall \tau, k. \tag{2d}$$

$$\textbf{EVB360(N}_{\textbf{max}}\textbf{):} \max_{[\mathbf{q}^1,\ldots,\mathbf{q}^K]} \sum_{i=1}^{N} \phi_i \quad \text{s.t.: (2a)} - \text{(2d)}, \tag{3}$$

where $\mathbb{1}_{p,j}$ and $\phi_i$ are indicator functions that take a value 1 to indicate broadcast of program $p$ by gNodeB $j$, and the condition $Q(q_{p,\tau_i}) \ge Q_{min}$ and $\gamma_i \ge SINR\_thr(m_{\tau_i}^{p})$ (i.e. the user is served with an acceptable minimum quality level) being satisfied, respectively. The vector of quantization parameters selected for encoding the 360° video tiles of programs being broadcast in MBSFN area $k$ ($p \in \mathcal{P}_k$, $1 \le k \le K$) is $\mathbf{q}^k = [q_{1,1}^k, \ldots, q_{1,T_1}^k, \ldots, q_{P_k,1}^k, \ldots, q_{P_k,T_{P_k}}^k]$. Note that only if atleast some users enjoy an acceptable service quality $\mathcal{Q}_i > 0$ for receiving $p$, it will be broadcast by $j$ when $j$ is part of at least one MBSFN area $k$ and $p \in \mathcal{P}_k$. Thus, the resource requirement constraints in (2b), (2c) are applicable for program $p$ only if it is broadcast by $j$. These constraints (2b)-(2c) are applied at each gNB as it can be a part of more than one MBSFN area and needs more stringent conditions.

The auxiliary variables $q_{p,\tau}^{(o)}$, $\sigma_{p,\tau}^{(o)}$ and $m_{p,\tau}^{(o)}$, pertaining to program $p$ and its tile $\tau$, in constraints (2b), (2c) are correspond to the overlap conjoined set $\mathcal{O}$ to which gNodeB $j$ belongs. The set $\mathcal{O}$ is updated based on the 360° programs that are actually being broadcast in the respective MBSFN areas ny means of Algorithm 1. The data rate $R(q_{p,\tau}^{(o)})$ required for transmission of tile $\tau$ of program $p$ is set by the encoder.

The optimization problem EVB360(Q$_{max}$) is solved based on the following proposition that enables formulating a low-complexity solution to the problem described below:

**Proposition 1.** *The objective* (2) *is a strictly decreasing function of the quantization levels* $[\mathbf{q}^1, \ldots, \mathbf{q}^K]$.

**Proof:** It is evident from Fig. 4 that both quality and rate are strictly decreasing functions of $q$. Analytically, we model

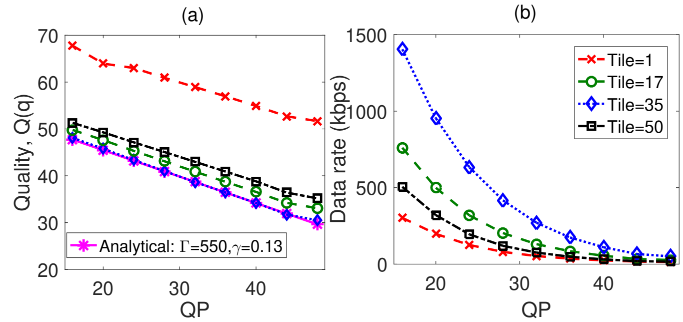


Fig. 4. (a) Quality (with analytical model) and (b) Rate variation with QP level for a 360-degree video.

$Q(q) = 10 \cdot \log_{10}(\Gamma \cdot e^{\gamma \cdot (q - q_{min} + 1)})$. It is shown in Fig. 4(a) that this video quality model for Tile 35 of a 360-degree video corresponds to $\Gamma = 550$ and $\gamma = 0.13$ with RMSE =0.0113. The non-negative weighted linear sum of strictly decreasing functions are decreasing [16]. Hence, by proving a generic $\mathcal{Q}_i$ is strictly decreasing function of the quantization levels, we prove that the same for our objective function. The first derivative of $\mathcal{Q}_i$ with respect to $q_{p,\tau}^k$ (i.e., the quantization level of the program requested by $u_i$ in the MBSFN(s) to which the UE belongs) is of the form $c \cdot q$ ($c < 0$ and a constant). This is negative thus proving the assertion. □

The objective function (2) with constraints (2a)-(2d) selects the highest possible quantization parameter level for the group of users requesting tiles of a program such that the resource constraint in the network are met. This is ensured in Function II of Algorithm 2.

The optimization problem for EVB360(N$_{max}$) is solved by selecting the lowest possible $m_{p,\tau}$ (in Function I of Algorithm 2) and the highest possible $q_{p,\tau}$ $\forall \tau, p$ such that $Q(q_{p,\tau}) > Q_{min}$. This ensures that the maximum number of users in the system have $\phi$=1 which is in accordance with objective (3). Our optimization problem solution in EVB360 framework is implemented using Algorithm 1-2 and provide:

*(i)* immersive 360° content set to be broadcast in MBSFN areas (indicator function value) as output of Algorithm 1;

*(ii)* radio resource allocation in terms of subframes $\sigma_{p,\tau}^k$, dynamic MCS selected $m_{p,\tau}^{(o)}$ for each tile of the 360° video broadcast in MBSFN areas as output of Function I in Algorithm 2, and

*(iii)* optimal quantization parameter for immersive video tiles' as output of Function II in Algorithm 2.

## V. Performance Evaluation

We have used five 360° videos from Youtube with diverse content types (Office, City, Sports, Jungle, and Sunrise) to assess the performance of our proposed scheme. They have 4K spatial resolution, 1800 frames, and 30fps frame rate. Each video is divided into M = 64 tiles (8x8 tiling) that are compressed using High Efficiency Video Coding (HEVC) [17] into 9 quality versions corresponding to 9 QP values of 16, 20, 24, 28, 32, 36, 40, 44, and 48.

Fig. 5 shows a sample network scenario we have considered. It includes 200 uniformly randomly distributed users

**Algorithm 2** Adaptive immersive 360° multimedia encoding and resource allocation

**Input:** $\{\mathcal{P}_k\}_k$, Overlap conjoined sets, $K$ MBSFN areas

**Function I: Resource_alloc** $(\{\mathcal{P}_k\}_k,\ u_i,\ i \in cluster\ k)$ **:**

1) **Proportion of resource allocation to program tiles**
   **for each program** $p \in \mathcal{P}_k$
   **for each tile** $\tau = 1$ to $T_p$
   Compute no. of users requesting $\tau$ of $p$, $n_k^p$
   Compute no. of users in cluster $k$, $N_k$
   $\sigma_{p,\tau}^{(o)} = \frac{n_\tau^p}{N_k}$
2) **MCS allocation to program tiles**
   **for each program** $p \in \mathcal{P}_k$
   **for each tile** $\tau = 1$ to $T_p$
   **EVB(Q$_{\max}$)**:$\overline{\gamma}$ =SINR that maximizes average quality in the system for all users requesting $\tau$ of $p$
   **EVB(N$_{\max}$)**:$\overline{\gamma}$=least SINR, user subset requesting $\tau$ of $p$
   Select $m_{p,\tau}^{(o)} = \min_m (SINR_{thr}(m) \geq \overline{\gamma})$

**return** $m_{p,\tau}^{(o)},\ \sigma_{p,\tau}^{(o)}$

**Function II: Tile_encoding** $(\{\mathcal{P}_k\}_k,\ m_{p,\tau}^{(o)},\ \sigma_{p,\tau}^{(o)},)$ **:**

1) **Quantization parameter selection for video tiles**
   **for each program** $p \in \mathcal{P}_k$
   **for each tile** $\tau = 1$ to $T_p$
   Select $q_{p,\tau}^{(o)} = \max_q (C(\sigma_{p,\tau}^{(o)}, m_{p,\tau}^{(o)}) \geq R(q))$

**return** $q_{p,\tau}^{(o)}$

**Output:**$m_{p,\tau}^{(o)},\ \sigma_{p,\tau}^{(o)},\ q_{p,\tau}^{(o)}$

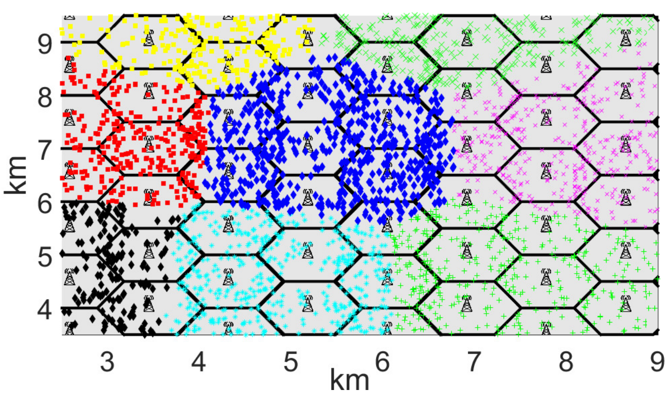


Fig. 5. Multi-criteria K-means area formation (8 clusters).

per cell and 52 gNodeB cells. Each user randomly views the broadcast program at a particular viewing angle selected at random. The 5G NR simulation parameters are listed in Table I. The SINR for each user in an MBSFN area with a given number of interfering cells is computed according to [5]. We evaluate the performance of our system by averaging the results over 100 iterations (95% confidence interval) with uniformly random distribution of users in the cell. We also examine the impact of the number of users.

Our approach leads to the formation of 8 MBSFN areas in the given scenario, as shown in Fig. 5, indicated by different color markers. The k-means based heuristic clustering approach in EVB360 results in Mahalonobis and Euclidean distance [18] values as 1.76 and 7.94, respectively. The corresponding measure values using another clustering method like nearest neighbor [18] for the same scenario are 2.29 and 9.74, respectively. This effectively shows the efficacy of the

TABLE I
SIMULATION PARAMETERS

| Simulation Parameter | Value |
|---|---|
| Channel bandwidth | 50 MHz |
| Frequency | 3.4 GHz |
| Number of data carriers | 1200 |
| Receiver noise figure | 7 dB |
| Maximum transmitter output power | 46 dBm |
| Transmitter (Receiver) antenna gain | 18 (0) dBi |
| Building loss | 14.0 dB |
| Receiver sensitivity | −106.4 dBm |
| Shadowing standard deviation | 8 dB |
| Average cell radius | 700 m |

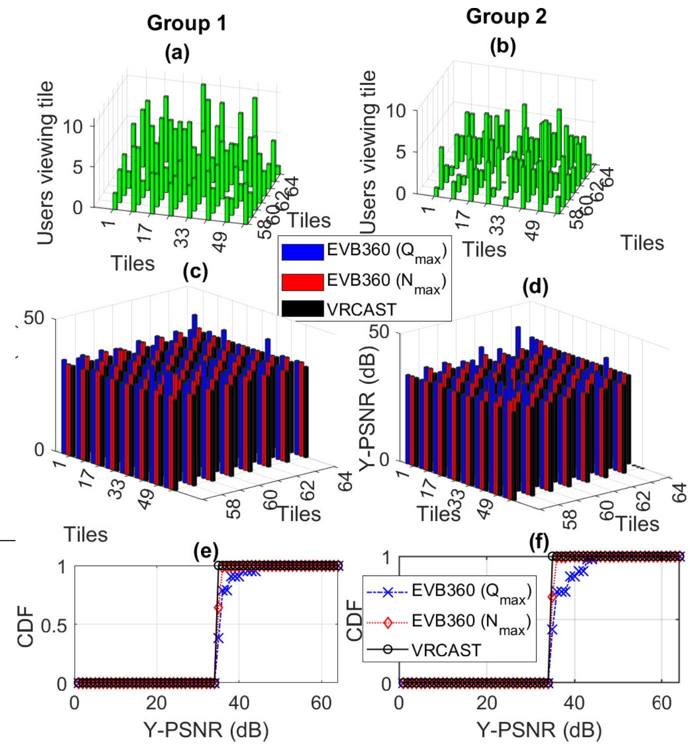


Fig. 6. (a),(b) Number of users requesting tiles (based on viewing angle); (c),(d) corresponding Y-PSNR (dB); (e),(f) CDF of Y-PSNR, of Program 1 and 2, respectively.

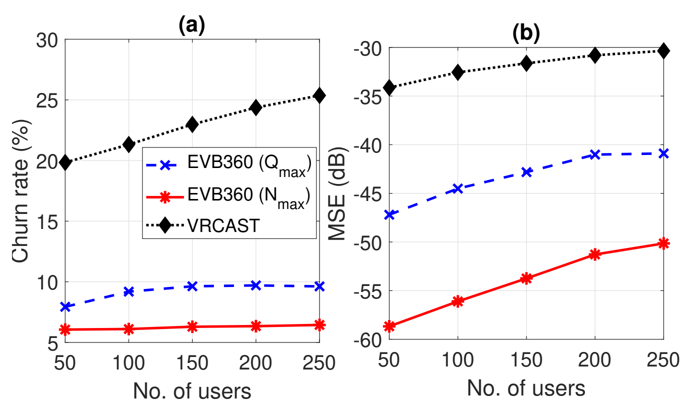


Fig. 7. (a) Churn rate and (b) MSE for EVB360 with increasing number of users per cell.

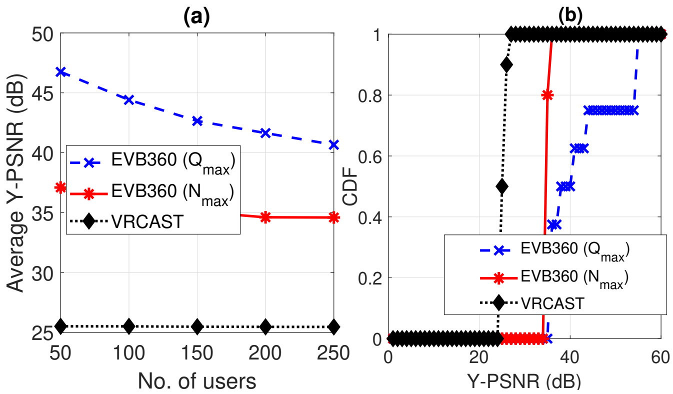


Fig. 8. (a) Average and (b) CDF of Y-PSNR for EVB360 with increasing number of users per cell

k-means based clustering approach used in EVB360.

The significance of tile based immersive video broadcast in EVB360 is evident from Fig. 6. Fig. 6(a)-(b) shows the number of users viewing tiles of program 1 and 2, based on their viewing angle, respectively. The two MBSFN groups employ adaptive tile encoding based on the resource constraint and user distribution within the MBSFN area. EVB360 (both $Q_{max}$ and $N_{max}$ schemes) selects efficient QP and MCS level as compared to existing scheme (VRCAST [8]) in dense network scenario resulting in better quality. The broadcast quality for the MBSFN areas (1-2) corresponding to the immersive video tiles of Program 1 and 2 is shown in Fig. 6(c)-(d), respectively. It is also evident from the CDF in Fig. 6(e)-(f) in that our proposed scheme, EVB360, provides a higher quality (or atleast equivalent to VRCAST) to all users (including the poorest channel quality user) in the system.

We also examine the performance of EVB360 (both $Q_{max}$ and $N_{max}$ schemes) in regard to churn rate (i.e., proportion of unserved users) and immersive video quality (in terms of viewport PSNR and MSE). It is evident from Fig. 7(a) that the churn rate increases as the number of users per cell increases. The churn rate of EVB360($Q_{max}$) and EVB($N_{max}$) is 40.34% and 67.72% (on average) lower than VRCAST. The MSE increases with increase in number of users and the performance of EVB360($Q_{max}$) and EVB($N_{max}$) is 30.78% and 52.48% better, respectively, than the existing scheme [8], as shown in Fig. 7(b). The Viewport PSNR (V-PSNR) reduces with an increase in number of users per cell but is maintained above 36 dB for the two EVB360 methods unlike VRCAST that has 25 dB V-PSNR, as shown in Fig. 8.

## VI. Conclusions

We have developed a novel and efficient immersive multimedia broadcast scheme, EVB360, for next generation cellular networks, that improves the overall $360^o$ video quality significantly while serving increased number of users. It considers the users' channel conditions, viewing angle, and service requests for adaptively broadcasting the $360^o$ immersive video content. The EVB360 scheme dynamically defines the MBSFN areas using multi-criteria K-means++ clustering. Thereafter, optimal tile encoding quality with efficient radio resources allocation is performed to either maximize the immersive broadcast video quality (i.e. EVB($Q_{max}$) scheme) or maximize the number of served users (i.e. EVB($N_{max}$) scheme). We have shown that our proposed framework, EVB360, outperforms other recent scheme by providing 35.82% better quality while serving 72.19% higher number of users in the a large-dense cellular network. In future work, we will extend our framework to include more efficient user clustering for energy-efficient prioritized broadcast services.

## Acknowledgement

This work has been supported in part by NSF awards CCF-2031881, ECCS-2032387, CNS-2040088, CNS-2032033, and CNS-2106150.

## References

[1] X. Corbillon, A. Devlic, G. Simon, and J. Chakareski, "Viewport-adaptive navigable 360-degree video delivery," in *Proc. Int'l Conf. Communications*. Paris, France: IEEE, May 2017.

[2] J. Chakareski, R. Aksu, X. Corbillon, G. Simon, and V. Swaminathan, "Viewport-driven rate-distortion optimized 360° video streaming," in *Proc. Int'l Conf. Communications*. Kansas City, MO, USA: IEEE, May 2018.

[3] C. Singhal and S. De, "Energy-efficient and Qoe-aware TV broadcast in next-generation heterogeneous networks," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 142–150, 2016.

[4] L. Richter, M. Hoyer, and S. Ilsen, "A software defined radio based fembms measurement receiver: Test results," in *Proc. IEEE International Symp. on Broadband Multimedia Sys. and Broadcasting (BMSB)*, Jeju, Korea, June 2019, pp. 1–9.

[5] J. J. Gimenez, J. L. Carcel, M. Fuentes, E. Garro, S. Elliott, D. Vargas, C. Menzel, and D. Gomez-Barquero, "5g new radio for terrestrial broadcast: A forward-looking approach for nr-mbms," *IEEE Transactions on Broadcasting*, vol. 65, no. 2, pp. 356–368, 2019.

[6] M. Tavakoli and M. Cave, "Modelling television viewing patterns," *Journal of Advertising*, vol. 25, no. 4, pp. 71–86, 1996.

[7] R. Aksu, J. Chakareski, and V. Swaminathan, "Viewport-driven rate-distortion optimized scalable live 360° video network multicast," in *Proc. ICME Int'l Workshop on Hot Topics in 3D (Hot3D)*. San Diego, CA, USA: IEEE, Jul. 2018.

[8] O. Eltobgy, O. Arafa, and M. Hefeeda, "Mobile streaming of live 360-degree videos," *IEEE Transactions on Multimedia*, vol. 22, no. 12, pp. 3139–3152, 2020.

[9] C. Singhal, C. F. Chiasserini, and C. E. Casetti, "Emb: Efficient multimedia broadcast in multi-tier mobile networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 11 186–11 199, 2019.

[10] J. Chen, M. Chiang, J. Erman, G. Li, K. K. Ramakrishnan, and R. K. Sinha, "Fair and optimal resource allocation for LTE multicast (eMBMS): Group partitioning and dynamics," in *IEEE INFOCOM*, Kowloon, Hong Kong, 2015, pp. 1–6.

[11] A. de la Fuente Iglesias, R. P. Leal, and A. G. Armada, "Performance analysis of eMBMS in LTE: Dynamic MBSFN areas," in *OPNETWORK*, Washington DC, Aug. 2013, pp. 1–6.

[12] D. Aloise, A. Deshpande, P. Hansen, and P. Popat, "NP-hardness of Euclidean sum-of-squares clustering," *Machine Learning*, vol. 75, no. 2, pp. 245–248, 2009.

[13] G. Wilkin and X. Huang, "A practical comparison of two K-Means clustering algorithms," *BMC Bioinformatics*, vol. 9, no. 6, 2008.

[14] L. Galluccio, O. Michel, and P. Comon, "Unsupervised clustering on multi-component datasets: Applications on images and astrophysics data," in *European Signal Processing Conference*, Lausanne, Switzerland, Aug 2008, pp. 1–5.

[15] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *ACM SIAM*, Philadelphia, PA, Jan. 2007, pp. 1027–1035.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[17] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[18] C. Shen, J. Kim, and L. Wang, "Scalable large-margin mahalanobis distance metric learning," *IEEE Transactions on Neural Networks*, vol. 21, no. 9, pp. 1524–1530, 2010.