This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



# Socially Aware Energy Efficient Mobile Edge Collaboration for Video Distribution

Dapeng Wu, Qianru Liu, Honggang Wang, Dalei Wu, and Ruyan Wang

***Abstract*—To relieve the current overload of cellular networks caused by the continuously growing multimedia service, mobile edge collaboration, which exploits edge users to distribute videos for base station (BS), provides an effective way to share the heavy BS load. With the emergence of mobile edge technologies for Internet of Things (IoT) applications such as device-to-device (D2D), and machine to machine (M2M), how to exploit users' social characteristics and mobility to minimize the number of transmissions of BS and how to improve the quality of experience (QoE) of users have become the key challenges. In this paper, we study two aspects that are critical to these issues. One is the two-step detection mechanism, namely the establishment of virtual communities and collaborative clusters. Specifically, we take into consideration user preference for content and location. First of all, a virtual community is established, which exploits the coalition game based on the user's preference list to dynamically divide users into multiple communities. Then, to take full advantage of the temporary link established between users, a grid-based clustering method is proposed to manage the video requesting users. On the other hand, we propose a Scalable Video Coding (SVC) sharing scheme based on user's social attributes. This approach makes video distribution more flexible at the edge of mobile network through collaboration among users, and effectively reduces transmission energy consumption of transmitters. Numerical results show that the proposed mechanism can not only effectively alleviate the BS load, but also dramatically improve the reliability and adaptability of video distribution.**

***Index Terms*—Mobile edge collaboration, scalable video coding, video distribution, multicast, energy-efficient, user attributes.**

## I. Introduction

WITH the rapidly increasing number of mobile users and wireless multimedia applications, the limited spectrum resources and ever-growing service requirements have become the main problems in mobile communications. According to Cisco's latest Visual Networking Index (VNI) report, the global mobile data traffic will increase nearly 11-fold from 2014 to 2019, of which mobile video services will account for nearly 75% [1]. In order to cope with this ongoing increasing mobile data demand, mobile edge computing (MEC) is introduced by the ETSI Industry Specification Group (ISG), which provides an information technology service environment and cloud computing capabilities within the radio access network (RAN) at the edge of mobile network [2]- [5]. It has the potential to significantly reduce latency, ensure highly efficient network operation, enhance the utilization of network resources, and optimize user experience.

In addition, the growing demand for video services from users which cause serious BS overloads eventually. Nevertheless, an enormous amount of duplicate content requirements exist in the local service [6]. Many studies suggest the close relationship between click-though rate (CTR) and popularity of a given video and follow the Zipf distribution [7]. According to the architecture of the current mobile communication network, BS needs to allocate resources in time and frequency domain for each user demand respectively, and sends the same content to individual users repeatedly [8]. Obviously, for multimedia service with bulk of data, such a service model is bound to cause a great waste of network energy [9]. Meanwhile, redundant content transmission makes the available spectrum resources extremely limited, which leads to a rapid decline in the quality of experience (QoE).

Currently, there are two main methods to solve the above-mentioned problems: 1) Significant improvement of the physical-layer link capacity between transmitter and receiver. However, the limited spectrum resources are insufficient to satisfy the explosive growth of multimedia services; 2) Decrease of the cell size to improve the efficiency of radio resources. Although this approach can enhance the capacity of the cellular capacity, the cost of establishing new cell sites and providing associated backhaul will increase accordingly.

Device-to-device (D2D) communications, as defined in 3GPP, enable the physically proximal device to transmit data directly over a short range under the control of BS [10]. By using D2D technology to distribute multimedia data at the edge of network in a collaborative manner, the scarcity of network resources caused by the large-scale concurrent multimedia data transmission can be resolved fundamentally. Based on the conception of MEC, we propose a next-generation collaboration architecture that lies at the edge of the cellular network-Mobile Edge Collaboration technology, which transfers some functions from the core network to edge network and relieves the burden of the network by using temporary links established between edge users. Compared with the traditional cellular communication, the edge computing mode based on D2D communications has the following advantages [11], [12]: 1) Due to the short range characteristic of the

Dapeng Wu, Qianru Liu, Ruyan Wang are with the Chongqing University of Posts and Telecommunications, CQ, 400065, China. (E-mail: wudapengphd@gmail.com, 2370094121@qq.com, wangry@cqupt.edu.cn)

Honggang Wang and Dapeng Wu are with the University of Massachusetts Dartmouth, MA, 02747, USA. (Email: hwang1@umassd.edu, dwu2@umassd.edu)

Dalei Wu is with the University of Tennessee at Chattanooga, TN, 37403, USA. (E-mail:dalei-wu@utc.edu)

This work is supported in part by the Natural Science Foundation of China (61371097). National Science Foundation (NSF) 1429120 and 1451629, and Program for Innovation Team Building at Institutions of Higher Education in Chongqing (CXTDX201601020). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Shiwen Mao (Corresponding author: Honggang Wang).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



D2D communications, the energy consumption and video transmission delay can be effectively reduced; 2) Through the cooperation between D2D users, not only a large number of transmission resources can be reduced, but also the BS load can be alleviated; and 3) Owing to the D2D users, the limited spectrum resources of the cellular system can be multiplexed, which can effectively improve resource utilization, and then the system capacity can be greatly enhanced.

Though D2D communications can offload BS traffic and further enhance the spectrum utilization, applying D2D-based edge computing method to solve the problem of massive video still faces following challenges [13]: 1) Video sharing typically introduces additional energy overheads for the transmitter, how to trade off the consumption and willingness of video sharing; 2) Due to the different preferences of users for the quality of multimedia playback, how to adapt to the needs of different video services; 3) The network environment is complex and dynamical, how to overcome the impact of channel errors and packet loss during network transmission; and 4) The connection between users has dynamic characteristics, and the movement will make the D2D link exhibits discontinuous features in time and space domain. In addition, users in cellular communication systems have significant social attributes [14]-[16]. The difference of social relationships among users lead to various transmission preferences for them, which will provide different service priorities for receiving user.

To solve the above problems, a video distribution mechanism based on dynamic social mobile edge collaboration is proposed in this paper. According to the similarity degree of user's interest, the logical relationship among users are analyzed and virtual communities are constructed in a distributed manner. Besides, by using SVC technology, such video stream is divided into a multi-layer nested bitstream, which consists of a base layer and multiple enhancement layers [17]. Specifically, users send requests to BS in a group or virtual community manner according to their physical location to reduce redundant transmission. During a subsequent video transmission, in order to avoid packet loss caused by channel differences, multiple levels of bitstream can be provided for respective users according to the energy state of each user's device and its current location. Subsequently, to meet the user experience requirements, additional enhancement layer data is further forwarded to the user based on the logical relationship between users in the best effort manner, namely user attributes are used to control encoded data forwarding. The main contributions are as follows:

- By analyzing the similarity degree of user interest in different contents, a distributed logical relation detection method is proposed. Using coalition games, the users are dynamically divided into different virtual communities, which makes the analysis results have dynamic evolution characteristics. Furthermore, according to the geographical distribution and the residual energy of users, the set of cooperative transmission users can be reasonably established for video requesting users.
- The SVC sharing mechanism based on the social characteristics of users is proposed in this paper, which can transmit the video data between users at various bitrates. The user can decide what to share in the local cache according to the similarity of proximity ones, which not only avoids redundant transmission, but also effectively reduces the transmission energy consumption of transmitters.
- A multicast mechanism based on the energy consumption of BS is employed in this paper. Under this scenario, the highest link rate in the cluster is selected as the multicast rate to obtain the optimal cooperative cluster number $\lambda$, which is combined with the SVC for selectively receiving the enhancement layers among users. The reasonable value of $\lambda$ balances the relationship between user's QoE and BS load; moreover, video distribution can be achieved more flexible and adapt to dynamic features of edge distribution.

The rest of this paper is organized as follows. Section II presents an overview of the previous work related to user attributes and energy efficiency in D2D communications. A grid-based clustering method based on game theory is presented in Section III. Section IV, a user attribute aware video distribution mechanism using SVC is proposed. Finally, simulation results are analyzed in Section V, and the conclusions are reached in Section VI.

## II. Related Work

Recently, there has been much effort on the video distribution over wireless networks. In general, previous works on this problem have typically assumed that users in each community can communicate with each other locally, and users are distributed uniformly in the cell. The assumption, however, is impractical for real D2D application scenarios. The relative works about data transmission based on user attributes and energy efficiency D2D communications are introduced below.

The random mobility of cellular users and dynamic device trajectories make the multicast structure dynamical, the connections between users unstable and the data transmission process frequently interrupted [18]. In [19], a reliable D2D multicast algorithm is proposed to ensure that every user receives the data correctly. The works in [20], [21] pointed out that the cellular data distribution can be achieved through D2D opportunistic deliveries, which could reduce the BS load and enhance the data transmission efficiency. The work in [22], D2D communication was used to improve the performance of multicast transmission in cellular networks. The authors proposed to use D2D communications inside the clusters to enhance the multicast performance. The work in [23], a heuristic algorithm was proposed, that used the history of user mobility of the previous day to identify a target set of users for the cellular deliveries. In [24] developed the incentive compatible contracts to encourage users to participate in data acquisition and distributed computing programs. The authors in [25] exploited social community to design efficient data forwarding in delay tolerant networks. Local users are very likely to request same video services and the resulting repeated multimedia data transmission consumes massive network resources in the traditional cellular network, which notably

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



affects delay and efficiency of the transmission process [26], [27]. The common assumption of these previous studies for cooperative D2D communications is that all users are willing to help any other user. However, D2D communications would incur significant energy consumption, there is no sufficient reason to assume that all users would cooperate with each other, especially for video service with bulk of data.

Much effort has been made in the literature to improve energy efficiency. The work in [22] proposed a buffering mechanism to offload and alleviate the heavy (BS) load. When a user device receives a content request, it directly checks the buffer of itself or its neighbors. To reduce of the energy consumption of BS in a video transmission and sharing scenario, in [28] proposed to construct in-band underlaying D2D clusters where the total energy consumption of each cluster is constrained. The works in [29]- [31] proposed several resource allocation algorithms based on game theory to improve the energy efficiency of D2D communication system. In [32], the author proposed a user group partition algorithm based on the coalitional game, which saves the total energy of the system by cooperation among users. However, this approach is largely limited by the geographical location of user device.

To summarize, it is different from our work, since we carried the comprehensive consideration on the energy consumption of BS, social attribute and physical location of users, as well as cooperation between edge users. In addition, we propose a video data distribution mechanism based on SVC. On one hand, SVC based multicast scheme provides differentiated services to community users with varying channel conditions, which improves the adaptability of system. On the other hand, user attributes are used to control SVC encoding and forwarding so as to meet the needs of different edge users.

## III. Node Relationship Analysis

### A. Virtual Community

The social attributes, which describe the long-term interaction behavior between users, are stable [33]. Using parameters associated with social attributes, users can make the best use of temporary link to complete the distribution of video data. In addition, virtual community is a cooperative organization composed of user devices, which can opportunistically share resources and dynamically connect nodes. As shown in Fig. 1, virtual communities are marked with different colors and users are going to or have arrived at their interest points. As mentioned in the introduction, in traditional video distribution mode, each user obtains desired content from BS independently. While multiple users access the same content, BS will transmit the content repeatedly. Such transmission mode not only increases the burden of BS, but also decreases the resource utilization of the network. Therefore, we propose a virtual community model which can alleviate the load of BS through collaboration among edge users. Different from the usual communities based on the location, the virtual communities can be established in a self-organized manner according to the interest preference of users [34]. Users in units of clusters send request to BS based on their physical location at current time, which greatly avoids redundant transmission of video, so as to achieve the purpose of saving energy for BS. Finally, users share data with each other by exploiting the regularity of their activities.

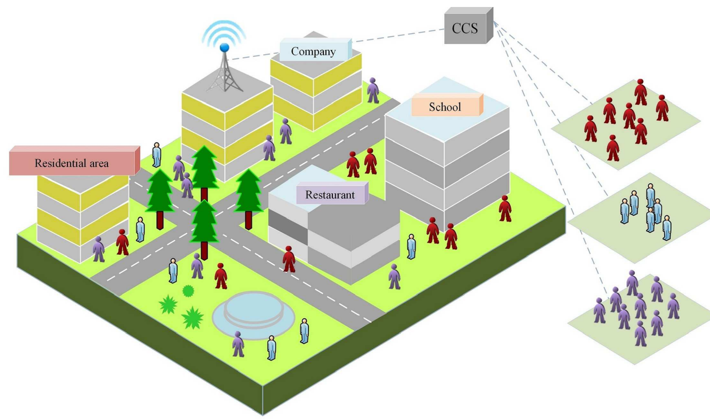



Fig. 1. The distribution of users in virtual communities

Virtual communities have dynamic characteristics, so it is necessary to update the node information in real-time. In this paper, Mobile Management Entity (MME) receives, updates and manages the node information through the Cooperative Control Server (CCS). The function of CCS is similar to the cluster-head or cloud-head, which accesses the mobile node within the community in a centralized way. In addition, this architecture can take advantage of both centralized and distributed architectures to deliver video in a ”best effort” manner at the edge of the network.

In addition, to meet the requirements of different users for video services, as well as the dynamically changing network environment, H.264/SVC standards was developed by International Telecommunication Union (ITU) in 2007 [35]. The basic idea is that the video is encoded into one base layer (BL) and one or more enhancement layers (ELs), and each layer containing the information relative to a quality level. The basic quality of the video stream is guaranteed by providing the user with the base layer. The quality of video can be improved as the number of received enhancement layers increases. An enhancement layer can be decoded if and only if all enhancement layers below it are received. Therefore, SVC-based sharing mechanism provides a degree of freedom by matching video stream characteristics to user device capability and the social characteristics of different users. For ease of searching, Table 1 summarizes the notations used in this paper.

### B. Virtual Communities Establishment

To improve the utilization of network resources, we introduce interest similarity to distinguish users, that is, users determine their affiliation with community by comparing the interest similarity with others. In this paper, the similarity of interest between users is obtained by collaborative filtering method, and this approach can be used to analyze and predict acquisition of user groups who own similar behavior or preference, namely, the users have similar interests primarily because they are interested in the same contents.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



TABLE I. Notations

| Notation | Destination | Notation | Destination |
|---|---|---|---|
| $N(u_i)$ | The preference set of user $u_i$ | $S_{u_i,u_j}$ | The similarity degree between user $u_i$ and $u_j$ |
| $A_t$, $C_t$ | The users set and coalition set after iteration $t$ times | $x_{u_i}^{p_k}(t)$ | The interest degree for point $p_k$ of user $u_i$ at time $t$ |
| $\mathcal{D}(u_i,u_j)$ | The difference of interest position between $u_i$ and $u_j$ | $\lambda$ | The number of collaborative cluster |
| $M_l^x$ | The number of users in the cooperative cluster $P_l^x$ | $R_{\max}^x$, $r_{\min}^x$ | The highest and low link rate in cooperative cluster |
| $r_{\min}$ | The lowest link rate for all interested users | $T_0$ | The base layer after SVC |
| $T_1,T_2,\cdots T_{\gamma-1}$ | The enhancement layer, respectively | $\overline{E}_{u_i}$ | The residual energy of user $u_i$ |
| $\mu_{u_i}$ | The enhancement layers received by user $u_i$ via BS | $\overline{\mu}_{u_i}$ | The enhancement layers received by user $u_i$ via D2D |
| $t_D$ | The user tolerance delay | $g_\tau(x)$ | User $u_i$ preferences for video $\tau$ playback quality |

For a given user $u_i$ and $u_j$ , let $N(u_i)$ and $N(u_j)$ denote their preference set respectively, which can be obtain by analyzing the history records. Assuming that the total number of videos is $M$ in a certain period. User access to the video $k$ can be represented by a binary random variable $a_k$, which is equal to 1 if the user caches the video $m$, otherwise $a_k = 0$. Thus $N(u_i) = \{a_1^i, a_2^i, \cdots a_M^i\}$ and $N(u_j) = \{a_1^j, a_2^j, \cdots a_M^j\}$ define the interest vectors of user $u_i$ and $u_j$ for video set, respectively. Thereby, the interest similarity between user $u_i$ and $u_j$ can be derived from the Jaccard coefficient [36], as shown in Eq. (1)

$$S_{u_i,u_j} = \frac{|N(u_i) \cap N(u_j)|}{|N(u_i) \cup N(u_j)|} \tag{1}$$

where $0 \le S_{u_i,u_j} \le 1$. Obviously, the lager value of $S_{u_i,u_j}$, the more similar the user's interest preference. Subsequently, the similarity degree between user $u_i$ and the others is obtained according to the Eq. (1), and arranged in descending order to constitute the user's $u_i$ preference sequence.

In order to achieve better experience through collaboration between users, the user invites others with the highest interest similarity to join the virtual community, and users in the same community can achieve higher video quality by sharing the cache. In the process, since user invites the most preferred one from current sequence to join the coalition, they can not be out of the current coalition and construct a new coalition to collect more data. Therefore, users are divided into different virtual communities dynamically according to their preference sequences. To further exploit the stable social relationship, and improve data forwarding efficiency accordingly, the coalition game model is employed to describe the relationship among users to minimize the virtual community difference. As can be seen, dividing users into multiple virtual communities according to the users' preference sequence can be transformed into a coalition game problem, and the optimal virtual community solution can be achieved by acquiring the core solution of the coalition game.

In this paper, each user can only belong to a virtual community, who holds the interest in a short time, so the virtual communities formed via interest similarity between users with strong stability, and all the coalitions are found via iteration manner. The procedures are as follows: For the first iteration $t = 1$, the number of users $|A_1| = N$ and coalition set $C_1 = \emptyset$, user $u_i$ is selected from $A_1$ arbitrarily. Then user $u_i$ sends an invitation message to one who have highest interest similarity. If it has not been involved in any coalitions, user $u_i$ will receive "accept" message, and then marks the user as $u_i^{opt}$ and make it join the coalition. However, if the user has involved in other coalition or replied "reject" message, user $u_i$ will continue to invite the next user in preference sequence. User $u_i^{opt}$ who accepts the invitation starts to invite the next user with highest interest similarity based on its preference sequence, and the invitation process is repeated, until the original user $u_i$ is invited to join the coalition, and then the first iteration ends. At the second iteration $t = 2$, the users who have already belonged to the coalition $C_1$ are removed, that is, $A_2 = A_1/C_1$. Then coalition $C_2$ is built from users set $A_2$. In the iteration $t$, users who belong to the previous coalition $\sum_{k=1}^{t-1} C_k$ is removed, namely, $A_t = A_1/\sum_{k=1}^{t-1} C_k$. This process is repeated until the set of users $A_t = \emptyset$, the iteration ends.

### C. Cluster Cooperation

According to the above method, the logical relationship between users is obtained based on the similarity of the preferred content. However, users with strong mobility, and their daily activities have certain regularity, thus the geographical distribution of users should also be analyzed. In addition, users are allowed to receive the corresponding video within the tolerance time $t_{\mathrm{D}}$ to enhance the user experience.

Users with mobility will access multiple geographic points of interest on their trajectories. Hence, the Point of Interest (PoI) tracks the location of particular users in chronological order. $P_i = \{p_i^{t_1}, p_i^{t_2}, \cdots p_i^{t_k}\}$ is used to represent the location of a user in different epochs. In most cases, the daily activities of users are with strong regularity. Fig. 2 is a two-dimensional graph of users position-time regularity obtained by statistical analysis over a period of time. Obviously, there is a greater chance of encountering between user 1 and 3 in the period of $t_2 \sim t_4$, while the probability of encountering between user 1 and 2 is larger in the period of $t_4$ and $t_9$.

Assuming that users in the virtual community formed by the coalition game are recorded as $\mathcal{M} = \{u_1, u_2, \cdots, u_M\}$; the PoI of user $u_i$ is $U_i = \{u_i^{p_1}, u_i^{p_2}, \cdots, u_i^{p_{k_i}}\}$. All users' PoIs are counted in the community and expressed as $\Psi_M = U_1 \cup U_2 \cup \cdots \cup U_M$ and $|\Psi_M| = K$, respectively.

The preference of various users to different PoI can be known according to the users' PoI for a given period. The larger appearance times clearly indicates the higher interest degree for a certain interest point. Thus, the preference degree of user $u_i$ for $p_k$ is obtained by the ratio between the number

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



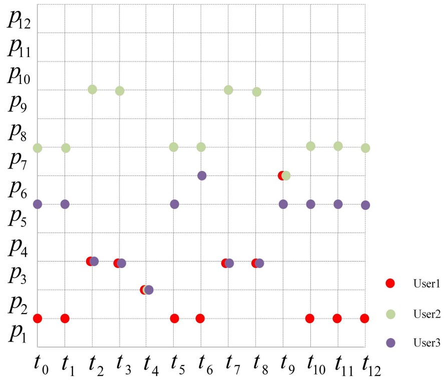



Fig. 2. Users location-timing diagram

of times user reaching the PoI and the total number of times. Therefore, $x_{u_i}^{p_k}$ is given by

$$x_{u_i}^{p_k} = \frac{m_{actul}}{\theta} \tag{2}$$

In Eq. (2), let $m_{actual}$ denotes the number of times that user $u_i$ at $p_k$, and $\theta$ indicates the total number of records. where $x_{u_i}^{p_k}(t)$ denotes the interest degree of user $u_i$ for point $p_k$ in time $t$, and $0 \leq x_{u_i}^{p_k} \leq 1$, $\sum\limits_{k=1}^{K} x_{u_i}^{p_k} = 1$, respectively. Furthermore, the interest degree for all points of user $u_i$ as shown in Eq. (3):

$$X_{u_i}(t) = (x_{u_i}^{p_1}(t), x_{u_i}^{p_2}(t), \cdots, x_{u_i}^{p_K}(t)) \tag{3}$$

At time $t$, the interest of user $u_i$ and $u_j$ for given point $P_k$, namely $x_{u_i}^{p_k}(t)$ and $x_{u_j}^{p_k}(t)$ are different [37]. The difference of interest between users is smaller, the similarity of trajectories and behaviors is more higher. The average of the interest difference for each location between user $u_i$ and $u_j$ defines as their degree of interest difference, as shown in Eq. (4):

$$\mathcal{D}(u_i, u_j) = \frac{\sum\limits_{k=1}^{K} |x_{u_i}^{p_k} - x_{u_j}^{p_k}|}{K} \tag{4}$$

where $\mathcal{D}(u_i, u_j)$ denotes the difference of interest position between user $u_i$ and $u_j$, and $0 < \mathcal{D}(u_i, u_j) \leq 1$. In particular, when $\mathcal{D}(u_i, u_j) = 1$ indicates that two users share no interest point.

According to users location in the virtual community at time $t$, PoIs are divide into $k \times k$ grid areas, namely cluster as shown in Fig. 3. The maximum distance in the same cluster is the communication range $R$ between D2D users, where the $k$ equals to $k = (\text{INT}\sqrt{L}) + 1$, and the $k \times k$ grid areas are denoted as $P = \{p_1, p_2, \cdots p_l \cdots, p_{k^2}\}$. BS selects optimal matching user according to the degree of preference for each location of interest at time $t$. The optimal matching user means the user with the highest degree of preference for $p_l$. Let $U_{optimal}^{p_l}$ indicates the users who satisfy the constraint $P_l = \{P | \max\{X_{u_i}(t)\}\}$. In other words, the cluster contains users with the lowest $\mathcal{D}(u_i, u_j)$ at time $t$.

If BS multicasts video data to all clusters, huge resources will be consumed. In this paper, BS selects $\lambda$ clusters as collaborative ones according to distances between it and the users with the highest link rate in each cluster. If there are equal distances, a cluster with a higher number of users is preferred. Obviously, the value of $\lambda$ is closely related to the user experience and BS energy consumption. If $\lambda$ is small, the number of transmission from BS to clusters reduces; at the same time, the number of users increases who do not receive desired video data in the duration of $t_{\text{D}}$. These users should obtain contents from BS directly, thus, it has little effects on offload BS traffic. On the contrary, users obtain the full video at a high probability in limited time and gain better QoE rapidly, but that raises the number of transmissions of BS, leading to a significant increase in BS energy consumption. Therefore, the optimal value of $\lambda$ is critical to the overall system.

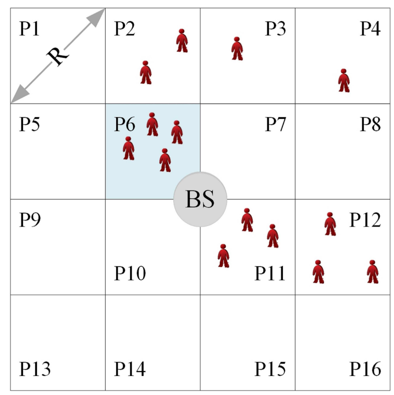



Fig. 3. Network coverage division

Without loss of generality, it assumes that the set of cooperative cluster is $P_{co} = \{p_l^1, \cdots, p_l^x, \cdots p_l^\lambda\}, 1 \leq \lambda \leq L$. The number of users in the cooperative cluster $p_l^x$ is $M_l^x (M_l^x \geq 3)$. In order to reduce energy consumption of BS, while ensuring that all users can receive basic video service in real time, let $R_{\max}^x$ denotes the highest link rate in the cooperative cluster $p_l^x$, and $r_{\min}$ indicates the lowest link rate for all interested users. All users can be guaranteed to obtain low quality videos via receiving and decoding the base layer with the multicast rate $r_{\min}$. Meanwhile, in order to get the optimal cooperative cluster number $\lambda$, BS selects the maximum transmission rate in the cluster as the multicast rate, and users with "poor" channel quality selectively receive the enhancement layers according to their capabilities. Therefore, when $\lambda_x$ satisfies the constraint condition is given by

$$\frac{T_0}{r_{\min}} \cdot P_{\text{BS}} + \sum_{x=1}^{\lambda_x} \frac{\sum\limits_{i=1}^{\gamma-1} T_i}{R_{\max}^x} \cdot P_{\text{BS}} \leq \frac{T_0 + \sum\limits_{i=1}^{\gamma-1} T_i}{r_{\min}} \cdot P_{\text{BS}} \tag{5}$$

It is concluded that $\lambda = \lambda_x$, namely there are $\lambda_x$ clusters that receive videos via BS multicast. Where the right side of the inequation represents the energy consumption which BS directly multicasts the video data to all interested users; $(T_0/r_{\min}) \cdot P_{\text{BS}}$ represents the energy consumption which BS multicasts the base layer to all interested users, and $\sum\limits_{i=1}^{\gamma-1} T_i/R_{\max}$ indicates the shortest transmission time required

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



for BS multicasting all enhancement layers to the cluster $p_l^x$. In addition, there are $\gamma(\gamma \geq 1)$ video layers, where $T_0$ denotes the base layer and $T_1, T_2, \cdots T_{\gamma-1}$ represent all the enhancement layers.

By simplifying the above inequation, the constraint for $\lambda_x$ can be denoted by

$$\sum_{x=1}^{\lambda_x} \frac{1}{R_{\max}^x} \leq \frac{1}{r_{\min}} \tag{6}$$

Since BS multicasts the video data to the cooperative cluster, the priority order of the multicast is decided by the distance. Therefore, from the above inequation, it is obvious that the inequation $R_{\max}^x \geq r_{\min}$, namely $\lambda_x \geq 1$, showing that at least one cooperative cluster receives all enhancement layers.

## IV. Edge Collaborative Sharing

Many studies have shown that the functions associated with wireless communication will consume more energy compared with any other component on the mobile device [38]. As mentioned in the introduction, the data sharing typically incurs additional energy overheads [39]- [40], which reduces device operation time, and ultimately results in reducing the user's QoE [41], [42]. In order to reduce the energy consumption, an edge collaborative sharing mechanism based on SVC is proposed. Users have different social attributes and social relationships, which will lead to different transmission preferences, and provides different service priorities for receiving user. Therefore, user attributes are used to control SVC encoding and video data forwarding, which not only effectively reduces energy consumption of transmitters, but also makes the video distribution process more flexible and easier to control [43], [44]. It also balances the relationship between users' QoE with BS load by above Eq. (5). In this paper, the process of edge collaborative sharing follows a two phase cooperative communication: in phase 1, BS transmits the video data to users; in phase 2, the users who successfully received the video data in phase 1 shared the received data to other edge users.

### A. Videos Distribution Model

In this subsection, a video distribution scenario is constructed. As shown in Fig. 4, to provide the basic video service for all users, the base layer is always multicast when BS multicasts the layered bitstream [45]. Apparently, the video quality improves as the number of received enhancement layers increases. Besides, the user with the highest link rate at the current time is denoted as $u_w$, whose rate is selected as the multicast rate, meanwhile, $u_w$ receives all enhancement layers from BS. Other users selectively receive video data according to location and residual energy. If users are not satisfied with the video quality, they would receive more enhancement layers from other edge users. Assume that residual energy of user $u_i$ is $\overline{E}_{u_i}$, and the distance between BS and user $u_i$ is denoted by $d_{c,u_i}$. Specifically, the number of enhancement layers that user $u_i$ can receive is proportional to $\overline{E}_{u_i}$, but inversely proportional to $d_{c,u_i}$. In addition, Take user $u_w$ as a reference, other users' reception rate $\eta_i$ is given by

$$\eta_i = (\frac{\overline{E}_{u_i}}{d_{c,u_i}})/(\frac{\overline{E}_{u_m}}{d_{c,u_m}}) \tag{7}$$

In addition, the enhancement layers received by user $u_i$ can be calculated by

$$\mu_i = \mathrm{INT}\left(\eta_i \cdot (\gamma - 1)\right) \tag{8}$$

where $(\gamma - 1)$ indicates the total number of enhancement layers, and $\eta_i$ denotes the reception rate of user $u_i$ relative to user $u_w$.

As illustration in Fig. 4, all users receive the base layer from BS, while two user receive all enhancement layers in cooperative cluster 1. user 1 with the highest link rate, whose rate is selected as multicast rate, so user 1 receives all enhancement layers. For user 3, $\eta_{u_3} \geq 1$ is calculated by Eq. (7), so user 3 also obtains all enhancement layers. Similarly, $\eta_{u_2} < 1$, so user 2 obtains $T_1$ and $T_2$.

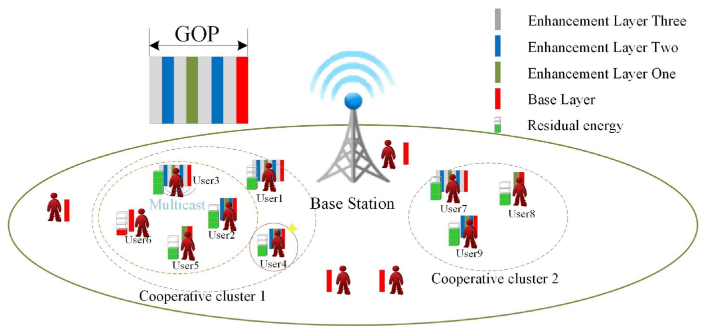



Fig. 4. Video distribution model

### B. Video Distribution from Base Station

As mentioned above, BS multicasts the base layer to all interested users and multicasts the enhancement layers to each cooperative cluster. Besides, in order to ensure that all users can receive the minimum quality video, the lowest transmission rate $r_{\min}$ is selected for multicast the base layer. Meanwhile, to obtain the optimal cooperative cluster number, we choose the highest link rate as multicast enhancement layers rate.

The user with "poor" channel quality selectively receives enhancement layers according to own location and energy. In this paper, assuming that $P_{\mathrm{BS}}$ is the transmitting power of BS, the channel gain between BS and user $u_i$ is $h_{c,u_i}$, so the received signal of user $u_i$ can be denoted by $P_{\mathrm{BS}} h_{c,u_i}$. In general, channels are assumed to follow Rayleigh distribution, the distance between BS and user $u_i$ is denoted by $d_{c,u_i}$, and the channel fading factor and Gaussian channel coefficient are denoted by $\partial$ and $h_0$, respectively. Because the communication between BS and user is under the impact of their distance and channel fading, the channel gain between them can be denoted by

$$h_{c,u_i} = d_{c,u_i}^{-\partial} h_0 \tag{9}$$

In addition, the noise of the received signal is composed of the additive Gaussian white noise $n_0 B$, and the same

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



frequency interference noise by D2D user $u_x$. In this paper, the link channel of user is assumed to be shared by only a single pair of D2D users. Therefore, the noise of received signal by user $u_i$ can be denoted by $P_{u_x} h_{c,u_x} + n_0 B$. Furthermore, according to the Shannon equation, the transmission rate between BS and user can be denoted by Eq. (10):

$$R_{c,u_i} = B\log_2(1 + \frac{P_{\rm BS} h_{c,u_i}}{P_{u_x} h_{c,u_x} + n_0 B}) \tag{10}$$

It assume $E_{u_i}$ is the available energy when user $u_i$ joins the virtual community. $T_0/r_{min} + \sum_{k=1}^{\mu_i} T_k/R_{\max}^x$ represents the required time to receive the video data from BS, and $P_{u_i,r}$ is the receiving power of user $u_i$, so the energy consumption of user $u_i$ is $(T_0/r_{\min} + \sum_{k=1}^{\mu_i} T_k/R_{\max}^x) \cdot P_{u_i,r}$. Therefore, the residual energy of user $u_i$ can be calculated by Eq. (11):

$$\overline{E}_{u_i} = E_{u_i} - (\frac{T_0}{r_{\min}} + \sum_{k=1}^{\mu_i} \frac{T_k}{R_{\max}^x}) \cdot P_{u_i,r} \tag{11}$$

In this phase, $c_{c,u_i} = \sum_{\gamma=1}^{\mu_i} T_\gamma$ denotes enhancement layers received by user $u_i$, where $\mu_i = \text{INT}\left(\eta_i \cdot (\gamma - 1)\right)$ indicates the number of enhancement layers. Let $T_0/r_{\min}$ represent the required time to transmit the base layer to all interested users, and $T_0/r_{\min} \cdot P_{c,s}$ indicates the energy consumption which BS directly multicasts the base layer to all interested users. Similarly, $c_{c,u_i}^x/R_{\max}^x$ denotes the time at which the enhancement layer is transferred to cooperative cluster. Therefore, BS's energy consumption can be denoted by

$$E_{multicast} = (\frac{T_0}{r_{\min}} + \sum_{x=1}^{\lambda} \sum_{l=1}^{M_l^x} \frac{c_{c,u_i}^x}{R_{\max}^x}) \cdot P_{c,s} \tag{12}$$

In this section, BS multicasts all enhancement layers to clusters in order until the energy consumption of BS does not satisfy Eq. (5).

### *C. Video Sharing under Virtual Communities Collaboration*

In this section, we utilize social characteristics and mobility of users to transfer some core network functions to the network edge; meanwhile, make full use of temporary link established among users instead of BS to complete the transmission of video. Hence, the number of transmissions through BS should be minimized by the cooperation of edge users.

*1) Intra-Cluster Data Forwarding:* In order to enhance the quality of video, users selectively receive enhancement layers according to their own position and energy [46]. There are some video coding strategies [52] [47] [48], which could be applied before the data forwarding. In [52], an efficient algorithm for reducing the computational complexity of multiple view coding has been proposed. [47] investigated a fast motion estimation (ME) method to reduce the encoding complexity of the H.265/HEVC encoder. In [48], a content similarity based fast reference frame selection algorithm has been proposed for reducing the computational complexity of the multiple reference frames based interframe prediction. However, users have different preferences for multimedia playback quality. Some users may not satisfied with the current video quality, so they expect to receive more enhancement layers to recover higher quality. In addiction, residual energy of user will have a serious impact on the edge communication process. In view of the above situation, we allow the user to multicast enhancement layers to other users, and multicast user $u_w$ is selected according to residual energy of the device.

In the D2D communication based mobile edge collaboration phase, assuming that $P_{u_w}$ is transmitting power of D2D user $u_w$. According to the Shannon Eq. (9), the transmission rate between D2D users can be denoted by $r_{u_w,u_x}$. Traditionally, if user multicasts a video stream to other users according to the "good" channel, user with a "poor" channel may fail to decode packets. To avoid the situation above, in this paper, the multicast rate is determined by "worst" channel quality as shown in Eq. (13):

$$r_{\min}^x = min\{r_{u_w,u_x}\}, x \neq w \,, x \subset p_l^x \tag{13}$$

where the multicast rate between user $u_w$ and other users in the cluster is $r_{\min}^x$, and $p_l^x$ denotes the x-th cooperative cluster, which has $M_l^x$ users.

User $u_w$ multicasts the received enhancement layers $\sum_{k\in[1,\gamma-1]} T_k$ to others in the cluster. The transmission delay is $t = \sum_{k\in[1,\gamma-1]} T_k/r_{\min}^x$. Therefore, the energy consumed by user $u_w$ to send the data can be calculated by

$$E_{co,s}^{(w)} = \frac{\sum_{k\in[1,\gamma-1]} T_k}{r_{\min}^x} P_{u_w,s} \tag{14}$$

In addition, define user's preference for multimedia quality as $g_\tau(x)$, where $0 \leq g_\tau(x) \leq 1$ is a generalized function that can be obtained by fitting the actual data. $g_\tau(x)$ denotes the personalized video services provided to users. Therefore, the energy consumed for receiving data on the D2D communication link can be given by

$$E_{co,r}^{(x)} = \sum_{x=1}^{M_{\min}^x} \frac{T_{\mu_x} + T_{\mu_{x+1}} \cdots + T_{\text{INT}(g_\tau(x)\cdot(\gamma-1))}}{r_{\min}^x} \cdot P_{u_x,r} \tag{15}$$

where $x$ denotes users except $u_w$ in the cluster $P_l^x$; $\text{INT}\left(g_\tau(x) \cdot (\gamma - 1)\right) - \mu_x$ indicates the enhancement layer that user $u_x$ failed to receive.

With the above mentioned edge collaboration structure, the multimedia distribution within the cluster is divided into the following two cases:

Case 1: If the residual energy of user $u_w$ is sufficient to multicast some or all of enhancement layers to other users in the cluster, the situation satisfies Eq. (16):

$$\overline{E}_{u_m} \geq \frac{\sum_{k\in[1,\gamma-1]} T_k}{r_{\min}^x} \cdot P_{u_m} \tag{16}$$

where $\sum_{k\in[1,\gamma-1]} T_k$ represents the demand enhancement layers of other users, and $r_{\min}^x$ indicates the lowest link rate between

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



user $u_w$ and other users in the cluster. In this case, the multicast is performed by user $u_w$.

Case 2: If the residual energy of user $u_w$ does not satisfy the constraint in Eq. (16). In this case, $\eta_i > 1$ indicates that user $u_i$ is superior to user $u_w$ in the residual energy and location at this current time, so user $u_i$ receives all enhancement layers. The condition satisfies $\overline{E}_{u_i} > \overline{E}_{u_w}$. Thus the multicast is performed by user $u_i$, so users in the cluster are able to receive the desired video data. If the residual energy of user $u_i$ still does not satisfy the condition (15), user $u_i$ performs the multicasting according to the value of the enhancement layer in order, which is determined by users requirements, so it can not meet the needs of all users. Similarly, if $\eta_i \le 1$, user $u_w$ multicasts enhancement layers to meet the requirements of most users.

As depicted in Fig. 4, the residual energy of user 1 is insufficient to multicast the required enhancement layer to other users. In this case, the conclusion $\eta_{u_3} > 1$ can be drawn from Eq. (7), thus the multicast is performed by user 3. Based on the different preferences for video quality, users can selectively receive the enhancement layer multicasted by user 3 according to their own preference. As shown in Fig. 4, user 2, 5 and 6 will continue receive the enhancement layers, while user 4 is no longer receives any video data.

To summarize, there exists at least one user in collaborative cluster receiving all enhancement layers. In particular, if the residual energy of user $u_w$ does not satisfy the constraint of Eq. (16), it multicasts in order by counting the enhancement layers of all users' requests. User $u_w$ finishes the multicasting until the residual energy is insufficient to complete the transmission of a certain enhancement layer, where the enhancement layers that user $u_i$ desires to obtain is $T_{\mu_i+1}, \cdots, T_{g_\tau(i)(\gamma-1)}$. Since the decoding of the enhancement layer $T_k$ relies on the former $T_{k-1}$ layer, some users can not decode immediately after receiving it, that is, the enhancement layers have potential value. Subsequently, the adjacent user transmits the remaining enhancement layers according to their similarity, thus the enhancement layers with potential value transformed into valuable ones, to provide the higher quality video services after decoding.

*2) Inter-Cluster Data Forwarding:* To make full use of the temporary link established between edge users. In this subsection, a SVC sharing mechanism based on social attributes of users is proposed. Users decide what to share in the local cache according to the similarity of proximity ones. Here, $\mu_{u_i} = \mathrm{INT}\left((\gamma-1)\cdot g_\tau(u_i)\right)$ indicates the number of enhancement layers that user $u_i$ prefers; while $\overline{\mu}_{u_i} = \mathrm{INT}\left(S_{u_x,u_i}\cdot \mu_x\right)$ denotes the number of enhancement layers actually received via D2D communications. So the transmission can be divided into three cases:

Case 1: If $\mu_{u_i} \ge \overline{\mu}_{u_i}$, user $u_i$ has finished receiving;

Case 2: If $\mu_{u_i} < \overline{\mu}_{u_i}$ and $t < t_\mathrm{D}$, user $u_i$ will continue to receive enhancement layers via other users until $\mu_{u_i} = \overline{\mu}_{u_i}$;

Case 3: If $\mu_{u_i} < \overline{\mu}_{u_i}$ and $t > t_\mathrm{D}$, user $u_i$ will continue to receive enhancement layers via BS.

According to user mobility analysis results given by Ref. [49], the encounter interval between users can be described by the Pareto distribution. It assumes that $T_{u_i,u_j}$ denotes the encounter interval between user $u_i$ and $u_j$. The complementary cumulative distribution function of $T_{u_i,u_j}$ can be calculated by

$$Pr_{u_i,u_j}^{CDF}\left\{T_{u_i,u_j} > t\right\} = \left(\frac{\tau_{u_i,u_j}^{\min}}{t}\right)^{\alpha_{u_i,u_j}}, t \ge \tau_{u_i,u_j}^{\min} \quad (17)$$

where the parameter $\alpha_{u_i,u_j} > 0$ determines the shape of the cumulative distribution function of encounter interval $T_{u_i,u_j}$, and $\tau_{u_i,u_j}^{\min}$ denotes the minimum possible value of $T_{u_i,u_j}$. Then the probability that user $u_i$ and $u_j$ meet each other at least once during time period $t_\mathrm{D}$ is given by

$$Pr_{u_i,u_j}^{meet}(t_\mathrm{D}) = 1 - Pr_{u_i,u_j}^{CDF}\left\{T_{u_i,u_j} > t_\mathrm{D}\right\} \quad (18)$$

At the same time, the number of enhancement layers that user $u_x$ has cached is $\lambda_x$. The content to be transmitted by BS is difference between the number of enhancement layers desired by user $u_i$ and actually obtained $\lambda_x$ in time $t_\mathrm{D}$, which is given by

$$c_\mathrm{BS} = \sum_{x=1,x\ne i}^{M_l^x}\left[\sum_{p=1}^{\mu_{u_i}} T_p - \mathrm{Pr}_{u_i,u_x}^{meet}(t_D)\sum_{q=1}^{\overline{\mu}_{u_i}} T_q\right] \quad (19)$$

Through above analysis, BS needs to directly transmit $c_\mathrm{BS}(u_i)$ to user $u_i$, the transmission time is $c_\mathrm{BS}(u_i)/R_{c,u_i}$. Hence, the energy consumption of BS at this phase is shown in Eq. (20):

$$E_\mathrm{cellular} = \sum_{u_i\in M}\frac{c_\mathrm{BS}(u_i)}{R_{c,u_i}}\cdot P_{c,s} \quad (20)$$

The total energy consumption of BS is $E_\mathrm{total} = E_\mathrm{cellular} + E_\mathrm{multicast}$. With the above mentioned edge collaboration structure, the multimedia distribution among clusters is further divided into the following two cases:

Case 1: If $E_\mathrm{total} \le (T_0 + \sum_{i=1}^{\gamma-1} T_i)/r_{\min}\cdot P_\mathrm{BS}$, it indicates that the total energy consumption of BS still satisfies constraint of Eq. (5), In this case, BS sends the required enhancement layer to each user individually.

Case 2: If $E_\mathrm{total} > (T_0 + \sum_{i=1}^{\gamma-1} T_i)/r_{\min}\cdot P_\mathrm{BS}$, by counting the enhancement layer of users' content and requests, BS multicasts them to meet the requirements of majority users.

By the above process, the highest link rate in the cluster is selected as the multicast rate based on energy consumption model to obtain the optimal cooperative cluster number. Then, BS multicasts the encoded video data to cooperative clusters, allowing users to selectively receive enhancement layers. In the intra-cluster communication phase, the multicast users are selected according to location and residual energy of devices at the current time, which cooperates with BS to distribute the video in the network edge. In the inter-cluster communication phase, some core network functions are transferred to network edge by utilizing user's social characteristics and mobility, and the burden on BS is minimized by the cooperation of edge users. The pseudo code of algorithm 1, namely the edge collaborative sharing mechanism based on energy consumption is shown below.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



**Algorithm 1** Edge collaboration sharing mechanism

**Initialization:** Select cooperative cluster $p_l^x$;
User tolerance delay $t_{\mathrm{D}}$;

**Intra-cluster sharing phase**
The maximum link rate in the cluster is used as the multicast rate, user $u_m$ receives all enhancement layers, Other users in the cluster receives part or all enhancement layers according to Eq.(7)
**if** $\mu_x \geq \mathrm{INT}\left(g_\tau(x) \cdot (\gamma-1)\right), u_x \in P_l^x$ **then**
  Intra-cluster communication phase ends
**else**
  **if** $\overline{E}_{u_m} \geq (\sum\limits_{k \in [1,\gamma-1]} T_k / r_{\min}^x) \cdot P_{u_m}$ or $\eta_i < 1$ **then**
    User $u_m$ multicasts the desired enhancement layers to other users
  **else**
    User $u_i$ multicasts the desired enhancement layers to others
  **end if**
**end if**
END
**Inter-cluster sharing phase**
**if** $\mathrm{INT}\left((\gamma-1) \cdot g_\tau(u_i)\right) \geq \mathrm{INT}\left(S_{u_x,u_i} \cdot \mu_x\right)$ **then**
  User $u_i$ has finished receiving
**else**
  **if** $t \leq t_{\mathrm{D}}$ **then**
    User $u_i$ continues to receive enhancement layers, Until $\mathrm{INT}\left((\gamma-1) \cdot g_\tau(u_i)\right) \geq \mathrm{INT}\left(S_{u_x,u_i} \cdot \mu_x\right)$
  **else**
    **if** $E_{\mathrm{total}} \leq (T_0 + \sum\limits_{i=1}^{\gamma-1} T_i)/r_{\min} \cdot P_{\mathrm{BS}}$ **then**
      BS sends the required enhancement layer to each user individually
    **else**
      BS multicasts according to the value of enhancement layers in order
    **end if**
  **end if**
**end if**
END

## V. Performance Analysis

In this paper we evaluate the performance of the proposed mechanism by using MATLAB. By exploiting user attributes and SVC, the intra-cluster and inter-cluster video data forwarding can be reasonably achieved to reduce redundant transmissions of BS. In this section, we provide various simulation results to show the performance of the proposed algorithms in terms of energy consumption, video transmission rate, and video quality.

Without loss of generality, we use human mobility trace datasets Infocom 06, which is real mobility trace and encounter of users [50]. The datasets records 98 iMotes' contact during the conference IEEE Infocom 2006 among which there are 20 static iMotes and 78 mobile iMotes. In our simulation, we use the 20 static iMotes as user' interest points and 78 mobile iMotes as mobile users; Moreover, a measured video standard scene based on SVC is utilized in this paper [51]. The video sequence is encoded using a Group of Picture (GOP) structure, each GOP is formed by a number of encoded video frames, which include I frame, P frame and B frame respectively. We assume that maximum transmission range of D2D is 50m, and the coverage radius of the base station is 500m. The bandwidth for wireless transmission is 1MHz. The power density of the Gaussian noise is -174 dB/Hz. The multicast distance between BS and user with the highest link rate in the cluster is denote by $D_{\mathrm{B}-c_x}$. In addition, to measuring the quality of received video and transmission strategy, we consider the following three parameters:

PSNR: Peak Signal to Noise Ratio (PSNR) is an standard objective evaluation of video quality. It is a function related to the mean square error between the original video frames and processed video frames.

VDR: Video Delivery Rate (VDR) is defined as the ratio between the number of users receiving the desired enhancement layers and the total number of video requesting users.

MOS: Mean Option Score (MOS) is a subjective evaluation of the video quality by users.

### A. Quantitative Analysis

In this paper, we analyze network performance from two aspects: the computational complexity and the overhead of energy.

*1) Computational Complexity:* In this paper, the analysis of the computational complexity is mainly focused on the establishment of the virtual community, and the complexity of other processes is approximately linear. The establishment of virtual community is to explore the paths beginning from each user. Since each user will be visited once, the computational complexity of the process is $O(|C_t|)$. Therefore, the overall computational complexity of all users is $O(\sum_{t=1}^{T}|C_t|)$, where $T$ is the number of iterations. More specifically, if a path forms a self-cycle with a path length of 1 and it is required to visit the remaining users who do not belong to any virtual community, the upper bound of computational overhead is $O(\sum_{t=1}^{T}|C_t| = \sum_{i=1}^{N} i = \frac{N(N+1)}{2})$. If all the users form a virtual community, that is, the path length is $N$, in this case, the lower bound of computational overhead is $O(N)$. Thus, the mechanism overall has a computational complexity of $O(N^2)$.

*2) Energy Overhead:* In this section, we analyze the upper and lower limits of energy consumption at both the user's side and BS side. As mentioned above, the total energy consumption of BS is $E_{\mathrm{total}}$, if the user has received the required video during time period $t_{\mathrm{D}}$, then no additional transmission is required. In this scenario, the lower bound of energy consumption is obtained, which is given by Eq. (21):

$$\frac{T_0}{r_{\min}} \cdot P_{\mathrm{BS}} + \sum_{x=1}^{\lambda_x} \frac{\sum\limits_{i=1}^{\gamma-1} T_i}{R_{\max}^x} \cdot P_{\mathrm{BS}} \leq E_{\mathrm{total}} \tag{21}$$

On the contrary, as in the above case 2, the upper limit of energy consumption can be calculated by Eq. (22):

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



$$E_{\text{total}} \leq \frac{T_0 + \sum_{i=1}^{\gamma\text{-}1} T_i}{r_{\min}} \cdot P_{\text{BS}} + \sum_{u' \subset \mathbb{N}} \frac{T_{u'}}{r'_{\min}} \cdot P_{\text{BS}} \tag{22}$$

where $u'$ denotes users who have not received the enhancement layer $T_{u'}$, and the number of these users is less than or equal to $N - \lambda$; $r'$ denotes the lowest rate of these users.

When the equality holds in Eq. (21), namely, the left side of the inequation represents the total energy consumption of BS. Therefore, the total energy consumption of the edge user can be denoted by

$$E_u = \frac{\sum_{k \in [1, \gamma-1]} T_k}{r^x_{\min}} P_{u_s} + \sum_{u' \subset \mathbb{N}} \frac{T_{u'}}{r_{u_s,x}} P_{u_s,x} \tag{23}$$

In particular, when the number of users $u'$ is equal to $N - \lambda$ and $T_{u'} = T_1$ are satisfied, at which point the total energy consumption of the edge user is minimized, and the value is 0, because in this scenario, there is no communication between the edge users.

## B. Energy Consumption Analysis

To evaluate whether the proposed algorithms can help to reduce the energy consumption of BS, we conduct series of simulations, and two distribution mechanisms without considerations of user attribute and SVC method are compared. The offloading rate under various tolerant delay is shown in Fig. 5. The offloading rate increases along with the growing tolerant delay. When the multicast distance between BS and user with the highest link rate in the cluster is small, the offloading rate for BS is high, and the offloading rate tends to be flat with the increase in the tolerant delay. This is because that the probability of users encounter increases when the tolerant delay increases. It can be seen from Eq. (10), when other conditions keep the same, the transmission rate decreases as the distance from BS to user increases. Under the constraint condition Eq. (5), the number of cooperating clusters is determined by the maximum transmission rate in the cluster, that means the number of cooperative clusters increases as the distance between BS and the user decreases. Obviously, when the tolerant delay is small, the maximum offloading ratio achieves up to 72.34%, and with the increase of tolerant delay, the offloading rate achieves up to 92.04% eventually. The varying tendencies of the curve of the $D_{\text{B-c}} = 100m$ consistent with $D_{\text{B-c}} = 200m$, this is because that the number of cooperative clusters is the same in both cases.

The energy efficiency under various numbers of users is shown in Fig. 6. Where the multicast distance between BS and user with the highest link rate is about 200m, and the tolerant delay is 1500s. Obviously, the energy efficiency of all three mechanisms will increase as the number of users in the virtual community increases. This is because there are more opportunities for collaboration among edge users when the number of users increases. For our proposed mechanism, which employs the reasonable virtual community detection

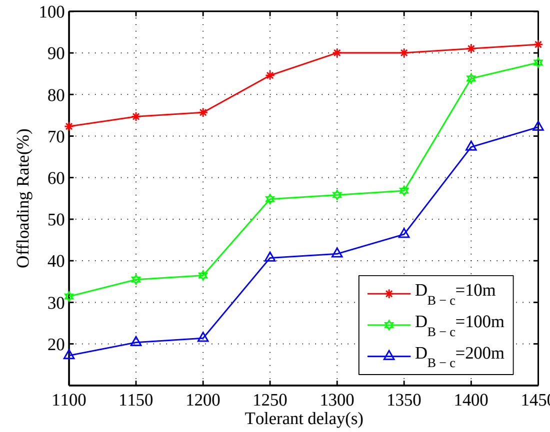



Fig. 5. The offloading rate under various tolerant delays

and SVC sharing mechanism based on the social characteristics of users, so the energy efficiency is higher than other two cases. For no user attribute mechanism, since the data sharing typically incurs additional energy overheads, thus some users are reluctant to provide their caches via D2D communication. But compared with the third mechanism, the application of SVC effectively reduces the energy consumption of transmitters, so the transmitters are more willing to share the cache with other users. According to numerical results, our proposed mechanism can achieve 20.3% and 49.6% higher energy efficiency respectively, when compared with the other two cases.

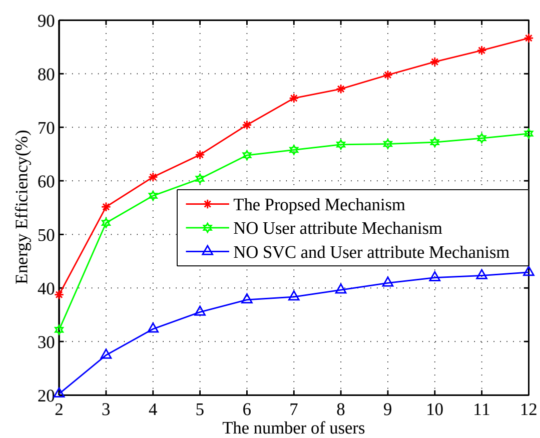



Fig. 6. The energy efficiency under various numbers of users

## C. Video Transmission Strategy Analysis

The video delivery ratio under various distances from BS to user $u_w$ is shown in Fig. 7. Apparently, the decrease of the video delivery ratio along with the growing distance from BS to user $u_w$. This is due to the fact that the number of cooperative clusters decreases as the distance from BS to user $u_w$ increases, and hence the number of users receiving the video data from the BS decreases dramatically. As can be seen, the ratio of delivery using proposed mechanism is 77.42% when $t = 0$, and 96.77% when $t = 2000$. The reason is that there are more cooperation opportunities among users when the tolerant delay increases, so users can improve video delivery ratio through shared caching.

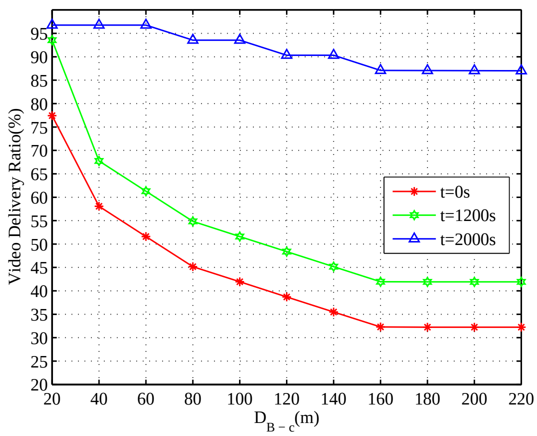



Fig. 7. The video delivery ratio under various distances from BS to user $u_m$

The video delivery rate under various numbers of users is shown in Fig. 8. As can be seen, as the number of users in the virtual community increases, the video delivery rate of our proposed mechanism presents the trend declines at the beginning and then promotes. This is because as the number of users increases, more cooperation opportunities raised among users, and more benefits of utilizing user attribute becomes. On the contrary, the video delivery rate of other two mechanisms have been decreasing because of the data sharing typically incurs additional energy overheads, and users are always resource limited, so most users are reluctant to provide their caches via D2D communication. Furthermore, the application of SVC effectively reduces energy consumption of transmitters, thus transmitters are more willing to share the cache with other users. Results show that the proposed mechanism can make up to around 21% performance gain over the solution that no user attribute mechanism.

### D. Video Quality Analysis

Fig. 9 illustrates the relationship between the video quality PSNR and the average bit rate. Apparently, the video quality of the proposed mechanism is higher than those of the video distribution mechanism without using user attribute and SVC method. This is because that the data sharing incurs additional energy overheads, some users are reluctant to share their caches via D2D communication; moreover, as the average bit rate increases, the compressing and decoding rate increases rapidly, and higher quality videos can be transmitted to users. Fig. 10 shows the MOS of all three mechanisms. The different color segments represent different experience qualities. When the PSNR range is about $20 \sim 25$, user's subjective experience quality is "poor"; $25 \sim 31$, $31 \sim 37$, $37 \sim 40$ indicate "Fair", "Good" and "Excellent", respectively. As can be seen from the Fig. 10, the subjective evaluation of the video quality by users

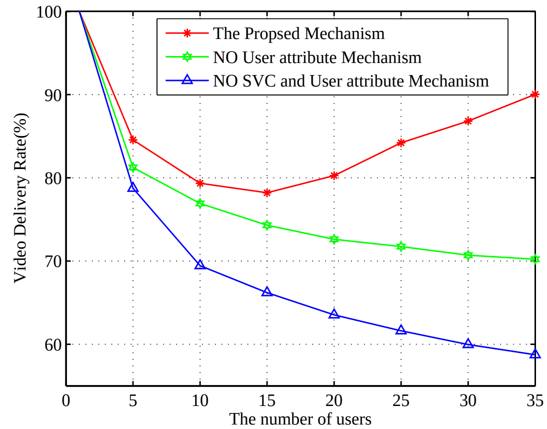



Fig. 8. The video delivery rate under various number of users

grows as the PSNR increases. Results show that the proposed mechanism improves the user experience.

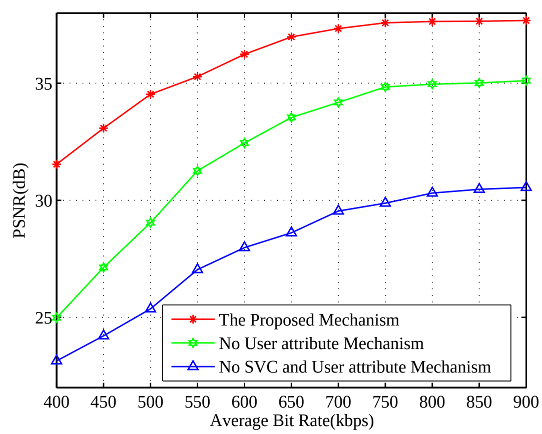



Fig. 9. The relationship between the video quality and the average bit rate

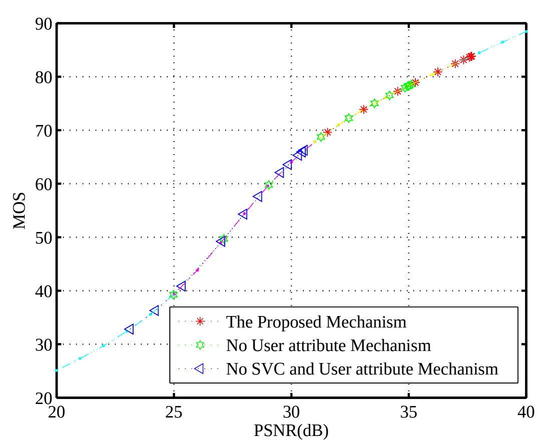



Fig. 10. The relationship between the PSNR and MOS

Fig. 11 reflects the video recovery performance of three

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



mechanisms with Foreman and Aspen sequences. Obviously, the quality of video frames of the proposed mechanism is better than those of the video distribution mechanism without using user attribute and SVC. According to Fig. 10, the average PSNR of the proposed mechanism is higher than 35dB and users' MOS is "good" or "excellent". But the other two mechanisms are between 25dB to 35 dB. In these cases, the MOS is "Fair" or "Good". So the mechanism proposed in this paper is better than the other two, making user's experience improve significantly.

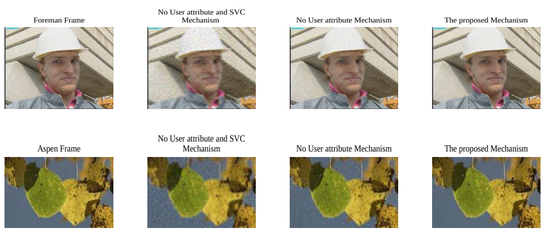



Fig. 11. The video recovery performance of three mechanisms on Foreman and Aspen video sequence

## VI. Conclusion

In this paper, we have proposed a social attributes-based scalable video coding mechanism for video data distribution at the edge of mobile network. Using the proposed scheme, user's social characteristics and mobility are exploited to transfer some core network function to the network edge. So the number of transmission of BS is minimized by the cooperation of edge users. Besides, the use of SVC provides users with different quality video services, reduces the packet loss caused by multicast, enhances the QoE of users. Finally, a multicast mechanism based on the energy consumption of BS can substantially enhance the data delivery rate and network resource utilization. Results have shown that BS load can be dramatically reduced. In this paper, we investigated the H.264/SVC in the proposed approach, which mainly considers the scalability and adaptability of video streaming. The H.265/SHVC can also be applied in the proposed approach. In our future work, we will also investigate how to exploit the spectrum resources at off-peak time.

## References

[1] Cisco Visual Networking Index: Forecast and Methodology, 2014-2019, Cisco Syst., San Jose, CA, USA, Feb, 2015.

[2] T. X. Tran, A. Hajisami, P. Pandey, and D. Pompili, "Collaborative Mobile Edge Computing in 5G Networks: New Paradigms, Scenarios, and Challenges[J]," *IEEE Communications Magazine*, vol. 55, no. 4, pp. 54-61, 2016.

[3] A. R. Khan, M. Othman, S. A. Madani, and S. U. Khan, "A survey of mobile cloud computing application models," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 393-413, Feb. 2014.

[4] P. Hu, H. Ning, T. Qiu, H. Song, Y. Wang, and X. Yao, "Security and Privacy Preservation Scheme of Face Identification and Resolution Framework Using Fog Computing in Internet of Things," *in IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1-1, 2017. doi: 10.1109/JIOT.2017.2659783

[5] M. Patel, B. Naughton, C. Chan, N. Sprecher, S. Abeta, and A. Neal, "Mobile-edge computing," *Mobile-Edge Comput.-Introductory*, Tech. Rep., Sep. 2014.

[6] J. Chen, M. Chiang, J. Erman, G. Li, K. K. Ramakrishnan, and R. Sinha, "Fair and optimal resource allocation for LTE multicast (eMBMS): group partitioning and dynamics," in *Proc. IEEE INFOCOM*, May. 2015, pp. 1266-1274.

[7] H. Yu, D. Zheng, B. Y. Zhao, and W. Zheng, "Understanding user behavior in large-scale video-on-demand systems," ACM SIGOPS Oper. Syst. Rev., vol. 40, no. 4, pp. 333-344, Oct. 2006.

[8] L. Zhou, "Mobile Device-to-Device Video Distribution: Theory and Application," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 12, no.3, pp. 1253-1271, 2015.

[9] A. Habbal, S. Goudar, S. Hassan, "Context-aware Radio Access Technology Selection Approach in 5G Ultra Dense Networks," *IEEE Access*, vol. 5, pp. 6636-6648, 2017. doi: 10.1109/ACCESS.2017.2689725

[10] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42C49, Dec. 2009.

[11] F. Malandrino, C. Casetti, and C,-F. Chiasserini, "Toward D2D-enhanced heterogeneous networks," *IEEE Commun. Mag.*, vol. 52, no. 11, 2014, pp. 94-100.

[12] L. Melki, S. Najeh, and H. Besbes, "Radio resource management scheme and outage analysis for network assisted multi-hop D2D communications," *Digital Communications and Networks*, vol. 2, no. 4, 2016, pp. 225-232.

[13] L. Zhou, "On Data-Driven Delay Estimation for Media Cloud," *IEEE Trans. Multimedia*, vol. 18, no. 5, 2016, pp. 905-915.

[14] Y. Cao, C. Long, T. Jiang, and S. Mao, "Share communication and computation resources on mobile devices: a social awareness perspective," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 52-59, 2016.

[15] D. Wu, J. Yan, H. Wang, D. Wu, and R. Wang, "Socially incentive mechanisms for video distribution in device-to-device communications," *IEEE Trans. Multimedia*, vol. PP, no. 99, pp. 1-1. doi: 10.1109/TMM.2017.2692648

[16] D. Wu, L. Zhou, and Y. Cai, "Social-Aware Rate Based Content Sharing Mode Selection for D2D Content Sharing Scenarios," *IEEE Trans. Multimedia*, vol. PP, no. 99, pp. 1-1. doi: 10.1109/TMM.2017.2700621

[17] D. Hu, and S. Mao, "On Medium Grain Scalable Video Streaming over Femtocell Cognitive Radio Networks," *IEEE J. Sel. Areas Commun.*I, vol. 30, no. 3, 2012, pp. 641-651.

[18] D. P. Wu, Y. Y. Wang, H. G. Wang, B. R. Yang, C. G. Wang, and R. Y. Wang, "Dynamic Coding Control in Social Intermittent Connectivity Wireless Networks," *IEEE Trans. Veh. Technol.*, vol. pp, no. 99, 2015, pp. 1-13.

[19] J. Seppala, T. Koskela, and T. Chen, "Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks[C]," *IEEE Wireless Communications and Networks Conference(WCNC)*, Cancun Quintana Roo: IEEE, Mar. 2011, pp. 986-991.

[20] B. Han, P. Hui, V. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan, "Cellular Traffic Offloading through Opportunistic Communications: A Case Study," in *Proceedings of ACM 5th Workshop on Challenged Networks (CHANTS)*, Chicago, IL, Sep. 2010, pp. 31-38.

[21] O. Helgason, F. Legendre, V. Lenders, M. May, and G. Karlsson, "Performance of opportunistic content distribution under different levels of cooperation," in *Proc. Eur. Wireless*, 2010, pp. 903-910.

[22] N. Golrezaei, P. Mansourifard, A. F. Molisch, and A. G. Dimakis, "Base-Station Assisted Device-to-Device Communications for HighThroughput Wireless Video Networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, 2014, pp. 3665-3676.

[23] B. Han et al., "Cellular traffic offloading through opportunistic communications: A case study," in *Proc. ACM CHANTS*, 2010, pp. 31-38.

[24] L. Duan et al., "Motivating smartphone collaboration in data acquisition and distributed computing," *IEEE Trans. Mobile Comput.*, vol. 13, no. 10, pp. 2320-2333, Oct. 2014.

[25] R. Cabaniss, S. Madria, G. Rush, A. Trotta, and S.Vulli, "Dynamic social grouping based routing in a mobile ad-hoc networks," in *Proc. HiPC*, 2010, pp. 1-8.

[26] D. P. Wu, J. He, H. G. Wang, C. G. Wang, and R. Y. Wang, "A hierarchical packet forwarding mechanism for energy harvesting wireless sensor networks," *IEEE Commun. Mag.*, vol. 53, no. 8, 2015, pp. 92-98.

[27] C. Luo, S. Guo, S. Guo, L. T. Yang, G. Min, and X. Xie, "Green Communication in Energy Renewable Wireless Mesh Networks: Routing, Rate Control, and Power Allocation," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 12, 2014, pp. 3211-3220.

[28] Y. Shen, C. Jiang, T. Q. S. Quek, H. Zhang, and Y. Ren, "Deviceto-device cluster assisted downlink video sharing-A base station energy saving approach," in *Proc. IEEE GlobalSIP*, Atlanta, GA, USA, Dec. 2014, pp. 108-112.

[29] C. Xu, L. Y. Song, Z. Han, Q. Zhao, X. L. Wang, X. Cheng, and B. L. Jiao, "Efficiency resource allocation for device-to-device underlay communication systems: A reverse iterative combinatorial auction based approach," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, 2013, pp. 348-358.

[30] D. Wu, J. Wang, R. Hu, Y. M. Cai, and L. Zhou, "Energy-efficient resource sharing for mobile device-to-device multimedia communications," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, 2014, pp. 2093-2103.

[31] S. Mumtaz, H. Lundqvist, K. M. S. Huq, and A. Radwan, "Smart Direct-LTE communication: An energy saving perspective," *Ad Hoc Networks*, 2014, 13, Part B(0): 296-311.

[32] Y. Zhang, F. M. Li, X. L. Ma, K. H. Wang, and X. H. Liu, "Cooperative Energy-Efficient Content Dissemination Using Coalition Formation Game Over Device-to-Device Communications," *Canadian Journal of Electrical and Computer Engineering*, vol. 39, no. 1, 2016, pp. 2-10.

[33] G. S. Zhao, X. M. Qian, and X. Xie, "User-Service Rating Prediction by Exploring Social Users Rating Behaviors," *IEEE Trans. Multimedia*, vol. 18, no. 3, 2016, pp. 496-506.

[34] L. F. Sun, X. Y. Wang, Z. Wang, H. V. Zhao, and W. W. Zhu, "Social-Aware Video Recommendation for Online Social Groups," *IEEE Trans. Multimedia*, 2016, DOI, 10.1109/TMM.2016.2635589.

[35] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the h.264/avc standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103-1120, Sep. 2007.

[36] X. F. Zhang, and D. Q. Dai, "A Framework for Incorporating Functional Interrelationships into Protein Function Prediction Algorithms," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 9, no. 3, 2012, pp. 740-753.

[37] X. Cheng, J. C. Liu, H. Y. Wang and C. G. Wang, "Coordinate Live Streaming and Storage Sharing for Social Media Content Distribution," *IEEE Trans. Multimedia*, vol. 14, no. 6, 2012, pp. 1558-1565.

[38] G.P. Perrucci, F.H.P. Fitzek, G.Sasso, W. Kellerer and J. Widmer, "On the Impact of 2G and 3G Network Usage for Mobile Phones Battery Life," In *European wireless 2009*, Aaborg, Denmark, May. 2009.

[39] N. Golrezaei, A. F. Molisch, A. G. Dimakis, and G. Caire, "Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution," *IEEE Commun. Mag.*, vol. 51, no. 4, 2013, pp. 142-149.

[40] L. Duan et al., "Motivating smartphone collaboration in data acquisition and distributed computing," *IEEE Trans. Mobile Comput.*, vol. 13, no. 10, 2014, pp. 2320-2333.

[41] Z. Jiang and S. Mao, "Energy delay trade-off in cloud offloading for multicore mobile devices," in *Proc. IEEE GLOBECOM*, San Diego, CA, USA, Dec. 2015, pp. 1-6.

[42] Y. Xu, and S. Mao, "A survey of mobile cloud computing for rich media applications," *IEEE Wireless Commun.*, vol. 20, no. 3, 2013, pp. 46-53.

[43] S. Mao, Y. T. Hou, and M. W, "Exploiting edge capability for wireless sensor networking," *IEEE Wireless Commun.*, vol. 15, no. 4, 2008, pp. 67-73.

[44] J. Song, P. Shen, K. Wang, L. Zhang and H. Song, "Can Gray Code Improve the Performance of Distributed Video Coding," *in IEEE Access*, vol. 4, pp. 4431-4441, 2016. doi: 10.1109/ACCESS.2016.2604358

[45] S. Hua, Y. Guo, Y. Liu, H. Liu, and S. S. Panwar, "Scalable video multicast in hybrid 3G/ad-hoc networks," *IEEE Trans. Multimedia*, vol. 13, no. 2, pp. 402-413, Apr. 2011.

[46] D. Wu, J. Wang, R. Q. Hu, Y. Cai, and L. Zhou, "Energy-Efficient Resource Sharing for Mobile Device-to-Device Multimedia Communications," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, 2014, pp. 2093-2103.

[47] Z. Pan, Y. Zhang, and S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," IEEE Transactions on Broadcasting, vol. 61, no. 2, pp. 166-176, June 2015.

[48] Z. Pan, P. Jin, J. Lei, Y. Zhang, X. Sun, and S. Kwong, "Fast Reference Frame Selection Based on Content Similarity for Low Complexity HEVC Encoder", Journal of Visual Communication and Image Representation, vol. 40, Part B, pp. 516-524, Oct. 2016.

[49] T. Karagiannis , L. Boudec, and J. Vojnovic, "Power law and exponential decay of intercontact times between mobile devices," *IEEE Trans. Mob.Comput.*, vol. 9, no. 10, 2010, pp. 1377-1390.

[50] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD trace cambridge/haggle/imote/infocom2006 (v. 2009-05-29)," Downloaded from http://crawdad.cs.dartmouth.edu /cambridge /haggle/imote/infocom2006, 2009.

[51] "Video traces for network performance evaluation," [Online] Available: http://trace.eas.asu.edu/tracemain.html.

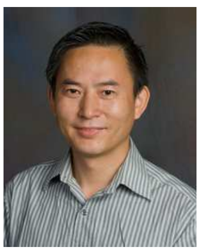
[52] Z. Pan, J. Lei, Y. Zhang, X. Sun, and S. Kwong, "Fast motion estimation based on content property for low-complexity H.265/HEVC encoder," IEEE Transactions on Broadcasting, vol. 62, no. 3, pp. 675-684, Sept. 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMM.2017.2733300, IEEE Transactions on Multimedia



**Dapeng Wu** received the M.S. degree in communication and information systems in June 2006 from Chongqing University of Posts and Telecommunications, and his Ph.D. degree from Beijing University of Posts and Telecommunications in 2009. His research interests include ubiquitous networks, IP QoS architecture, network reliability, and performance evaluation in communication systems.

**Qianru Liu** received the M.S. degree candidate in information and communication engineering at Chongqing University of Posts and Telecommunications. Her research interests include D2D communication.



**Honggang Wang** received his Ph.D. in computer engineering from the University of Nebraska-Lincoln in 2009. He is an associate professor at the University of Massachusetts (UMass) Dartmouth. His research interests include wireless health, body area networks (BANs), cyber security, mobile multimedia and cloud, wireless networks and cyber-physical systems, and big data in mHealth.

**Dalei Wu** (M10) received his Ph.D. degree in computer engineering from the University of Nebraska-Lincoln, Lincoln, NE, USA, in 2010. His research interests include networking, intelligent systems, cyber-physical systems, and complex dynamic system modeling and optimization.

**Ruyan Wang** received the Ph.D. degree in 2007 from the University of Electronic and Science Technology of China (UESTC) and his M.S. degree from Chongqing University of Posts and Telecommunications (CQUPT), China, in 1997. Since December 2002, he has been a professor with the Special Research Centre for Optical Internet and Wireless Information Networks at CQUPT. His research interests include network performance analysis and multimedia information processing.