Cooperative Live Video Multicast for Small Cell Base Stations with Overlapping Coverage

Özgür Alay † and Antonios Argyriou ‡
†Simula Research Laboratory, Oslo, Norway
‡Department of Electrical and Computer Engineering, University of Thessaly, Volos, Greece
E-mails: ozgu@simula.no and anargyr@uth.gr

Abstract—In this paper, we present a cooperative video multicast protocol for densely deployed small cells. We assume that the small cell Base Stations (BS) have overlapping coverages and they are allowed to have interfering transmissions. The proposed cooperative multicast protocol realizes a two-hop transmission system where each hop has its own time slot. In the first hop, the small cell base stations simultaneously transmit the packets to the relays who are also the users of multicast transmission. The relays then follow the optimal strategy under interference conditions and apply successive interference cancellation (SIC) decoding. In the second hop, the relays apply a Distributed Space Time Code (DSTC) that dynamically adapts itself based on the result of SIC and broadcast the result. This cooperative transmission scheme is complemented with application layer Forward Error Correction (FEC) in order to handle the remaining packet level losses. Through extensive simulations, we investigate the performance of the proposed scheme and show that the proposed protocol can efficiently handle inter-user interference, leading to superior performance over the state-of-the-art purely orthogonal cooperative transmission schemes.

I. INTRODUCTION

Propelled by powerful smart phones and tablets, mobile video is more popular today than ever [1]. According to Cisco’s Global Mobile Data Traffic Forecast, mobile data traffic grew 69% in 2014 and is expected to grow almost tenfold by 2019 [2]. Among all mobile content types, mobile video generates 55% of the mobile traffic today and is expected to generate 72% of the mobile traffic in 2019.

The greedy behavior of the video traffic imposes higher communication data rate requirements on the network. As a response to this demand, mobile network operators (MNOs) build a denser wireless infrastructure by deploying small cells in order to increase wireless capacity [3], [4]. On the other hand, video traffic has certain characteristics that can be exploited for efficient video delivery. For example, for on-demand video delivery, the asynchronous re-use of the same video by several users can be optimized by caching [5]. Applying such mechanisms to the topology considered in this paper (Figure 1) means that the two BSs with overlapping coverage can store several videos, and when a user desires a certain video it has higher probability of receiving it locally since it can reach several BSs [5].

For applications such as live video streaming, caching is not applicable and it is currently unclear how a MNO can leverage the presence of multiple BSs optimally in order to provide a high quality live video stream to multiple users. One promising technique for efficient delivery of live video is to use wireless multicast where popular video events are delivered to many wireless nodes simultaneously in a bandwidth-efficient manner. Several studies focus on video multicast in wireless networks from different perspectives. For example, in order to provide differentiated quality to multiple users, the use of scalable video for wireless multicast has been studied in [6],[7]. Robust and flexible wireless video multicast with the use of network coding has been explored in [8]. Video multicast that can take advantage of heterogeneous networks has been studied in [9]. Finally, energy efficiency of video multicast for 4G networks has been explored in [10].

In our dense small cell setup in Figure 1, one can take advantage of cooperative communications [11]. Cooperative communications is especially attractive for multicast since the relays are part of the multicast user population. Hence, any performance improvements will directly impact the relays as well. Such a cooperative multicast transmission scheme has been first described in [12] using orthogonal relay transmission and it has been shown to provide significant performance improvements for the scenario in Figure 1 with only a single BS. The central concept behind the cooperative multicast is that a BS transmits a packet and the users who decode the packet of interest become the first hop users. Selected first hop users then transmit the packet to receivers who did not receive it during the first transmission attempt. This study has

Fig. 1: We consider a topology where two small cell BSs (BS1 and BS2) deliver the same live video to multicast users (S1, S2 and S3). In the first hop, instead of time/frequency sharing of a given set of resources, BSs are allowed to have interfering transmissions. In the second hop, the users (S1 and S2) decode the packets using SIC and re-transmit to the remaining users (S3) using DSTC to provide reliability.
been extended in [13] with the use of Randomized DSTC (R-DSTC) in the second hop. Since the signal is decoded at more than one node, the use of R-DSTC ensures a diversity gain for the transmission to the remaining users. This protocol makes very good use of the fact that information is available at several nodes. All these schemes, including the cooperative ones, only consider a single base station and assumes BSs have orthogonal channel access in case of multiple BSs. However, orthogonal channel access is fundamentally sub-optimal in the high SNR regime for the multiple input single output (MISO) uplink AWGN and fading channels [14]. This means that in the high SNR regime, more than one BS can transmit simultaneously towards a single destination leading to higher spectral efficiency or multiplexing gain. Simultaneous multi-source transmission for cooperative video delivery was first reported in [15]. In that work two sources and one relay were used for forwarding independent video streams to two different users. The results showed that higher video quality is possible for both end users if they jointly and simultaneously use an intermediate low complexity relay.

In this paper, we investigate a cooperative wireless live video multicast scheme where multicast users lay in the overlapping coverage areas of two BSs. We assume the two BSs simultaneously transmit while being loosely coordinated in the first hop. The relays, following the optimal strategy under interference conditions, apply SIC decoding. Then, in the second hop, the relays employ an adaptive DSTC [16] that dynamically adapts itself based on the result of SIC and forward the packets to other multicast users. This cooperative transmission scheme is complemented with application layer FEC in order to handle the remaining packet level losses. We compare our scheme with the protocol reported in [13] that represents a state-of-the-art cooperative video multicast solution where BSs only support orthogonal transmission.

The main contributions and results of this work are:

- This paper presents the first cooperative live multicast protocol that supports simultaneous transmission from two sources, i.e., it allows interfering transmissions. Such protocol can extract performance benefits in a wide range of SNRs which is very relevant in dense small cell networks that are continuously being deployed today.
- Performance results for PSNR show video quality improvements in the order of 2-10 dB. Such performance gains are achieved due to the significant increase in the average effective video rates.

II. System Setup & Model

In our network model illustrated in Figure 1, there are two BSs with overlapping coverage. They stream a live video to $N$ multicast users given the wireless resources such as the frequency and time. Each BS receives the complete video stream as it would in typical orthogonal transmission scheme. In the figure, we illustrate only one non-relay user, namely S3 to avoid clogging the figure. In this work, we assume that two of the multicast users are willing to help as relays and are selected among the potential number of relays, $N_p$, based on the relay selection rule described in Section III-C. We do not investigate more advanced relay selection rules since the problem is out of the scope of this paper. However, we investigate different channel conditions between BS1-BS2 and S1-S2 as well as between S1-S2 and S3 to evaluate different relay/user channels in a multicast setting. Considering that S3 is a representative multicast user, changing the channel conditions between S1-S2 and S3 effectively mimics different multicast users. As in the conventional multicast designs, the sender can then select the user with the worst average channel conditions among all multicast users and send the video in a rate that ensures all the multicast users receive the live video.

The video is assumed to be packetized, however, our system model can also support segments such as the ones used in DASH (even though DASH specific details are out of the scope of this paper). We also assume the BSs do not have any type of communication through the backhaul network. With our protocol, the packets are transmitted using a certain Modulation and Coding Scheme (MCS) in two hops, each having their own time slot. In the first hop, the BSs transmit their packets simultaneously as illustrated in Figure 2. In the second hop, each relay attempts to decode the interfering signals. Upon correctly decoding each packet, the two relays transmit an ACK (consistently with WLAN and LTE-based link-layer protocols), allowing each BS to select a new unique packet\footnote{The selection of a unique packet can be done in different ways, for example each relay can be responsible for transmitting the odd/even packets, or a higher level coordination can be arranged.} for the next transmission round as illustrated in Figure 2. The relays independently apply a DSTC that is adaptive to the result of SIC decoding in order to improve the reliability of the transmission towards the remaining users.

A. Channel Model

Every node in our system model has a single omnidirectional antenna that can be used both for transmission and reception while all nodes have the same average power constraint. We denote the channel from the $s$-th base station to the $r$-th relay as $h_{s,r}$, and the channel from the $r$-th relay to destination as $h_{r,d}$. We assume that the fading coefficients are independent and complex Gaussian random variables with zero mean and unit variance, i.e., $h_{s,r} \sim \mathcal{CN}(0,1)$, $h_{r,d} \sim \mathcal{CN}(0,1)$. All the channels from base stations to relays and relays to destinations are considered to be block-fading...
In case $x_1$ is correctly decoded, it is then subtracted from the aggregate signal $y_i$. The implementation of the cancellation mechanism is executed at the level of PHY frames. The successful decoding of $x_1$ is verified with the use of an error correcting cyclic redundancy check (CRC) code. Thus, upon the successful decoding, and with $h_{BS1,i}$ available at the relay, we can completely remove/cancel a complete block from the aggregate received signal $y_i$ allowing the decoding of the second signal.

**B. DSTC for Interfering Signals (DSTCIF)**

Next, we briefly describe the main concept behind DSTCIF that employs an adaptive DSTC depending on the result of SIC. This protocol was first proposed and evaluated in [16] in the context of improving the spectral efficiency in multi-user relay networks.

The main idea behind DSTCIF is that after the first hop transmission, the relays may be able to either decode or not decode the signals received. Based on what they decode, they will transmit different signals in the second time slot. To denote these signals that the relays transmit, we use the notation $q_{r,i}$ where $r$ indicates the relay and $i$ denotes the transmitting BS. Note that the behavior of each relay is independent of the other participating relay. The relay pre-processing can be modeled compactly as follows:

$$q_{r,1}[1] = x_1 + a_{r,1}x_2 + b_{r,1}w_r$$

$$q_{r,2}[1] = x_2 + a_{r,2}x_1 + b_{r,2}w_r$$

The adopted signal notation covers every possible packet decoding outcome at the relay through the complex gain variables $a_{r,1}, b_{r,1}, a_{r,2}, b_{r,2}$. The main concept is that $q_{r,2}$, for example, will always contain an equalized version of symbol $x_2$ from the second BS plus whatever signal remains depending on the SIC results.

The resulting signals $q_{r,1}, q_{r,2}$ are used as input to the space-time coder. The matrix of transmitted symbols is

$$\mathbf{Z} = \sum_{r=1}^{2} \sum_{i=1}^{2} q_{r,i} \mathbf{A}_{r,i} + \mathbf{B}_{r,i} q_{r,i}^*.$$ 

In the above matrix the rows and the columns are indicated by the relay and the time slot, respectively. $q_{r,i}$ is the power scaling coefficient for signal $q_{r,i}$. This is essentially the DSTC codeword but in the general case it contains completely different signals and thus it is not in the well-known form of the Alamouti STC. Hence, if both information blocks are decoded then the system would operate like a typical STC encoder. However, when $q_{r,i}$ does not contain a decoded symbol, the application of the STC effectively operates on this composite symbol which makes it effectively an analog-type of code [18]. The same process takes place at both relays, while the space time decoder at the destination decodes the two signals. The final destination receives the STC codeword forwarded over the second slot, and applies a MMSE MIMO-SIC decoder to ensure optimality:

$$\mathbf{\hat{x}} = \text{HDD}\{ (\mathbf{H}^H\mathbf{\Sigma}_w^{-1}\mathbf{H} + \mathbf{I})^{-1}\mathbf{H}^H\mathbf{\Sigma}_w^{-1}\mathbf{\hat{y}}_d \},$$

Rayleigh. The channel coefficients are quasi-stationary, that is they remain constant for the complete duration of a block transmission for each base station/relay and relay/destination pair. Additive white Gaussian noise (AWGN) is assumed at the relays and the destinations. Each transmitted block consists of $L$ symbols.

**B. Channel State Information (CIS)**

Regarding the required CSI at a relay, only the knowledge of the channel from the base stations to that specific relay is needed in order to decode the interfering symbols and calculate the power scaling factor. No further channel knowledge is required. However, channel state information at the final receiver/destination (CSIR) is required and it can be obtained by sending training signals from the relays and the base stations.

**III. Protocol for Multicast Video Delivery from Multiple Base Stations**

With the proposed cooperative video multicast delivery protocol, the transmissions take place in two slots. In the first slot, the BSs transmit their corresponding packets simultaneously and each relay attempts to decode the interfering signals on its own (Section III-A). In the second slot, the relays cooperatively transmit their signals using a DSTC that is adaptive to the result of SIC (Section III-B). The end-to-end packet error rate is computed based on this two-slot transmission system and in order to handle these losses, application level FEC is applied (Section III-D). Finally, the effective video rate and the corresponding video quality is computed (Section III-E).

**A. Opportunistic Successive Interference Cancellation (OSIC)**

Let us assume that BS1 and BS2 communicate at a rate of $l$ and $m$ bits/symbol, respectively. Thus, the baseband model for the received interfering information blocks at relay $r$ is:

$$y_i(1) = h_{BS1,r}x_1 + h_{BS2,r}x_2 + w_i.$$  

(1)

In this expression $w_i$ is the AWGN sample at the relay. A relay attempts to decode the two symbols $x_1, x_2$ by employing ordered OSIC. That is, the symbol with the highest energy/bit is decoded first while the other symbol is treated as noise [14]. If there was no interference, the following condition must be true so that block $x_1$ from BS1 is decoded:

$$\log_2(1 + \frac{P_{BS1}|h_{BS1,r}|^2}{\sigma^2}) \geq l \Rightarrow \frac{P_{BS1}|h_{BS1,r}|^2}{\sigma^2(2^l - 1)} \geq 1$$

The fractional term in the RHS of the last derivation is essentially the normalized SNR/bit that is required for decoding $l$ bits/symbol [17]. We can get a similar expression for the data from BS2, and by assuming $\mathbb{E}[|x_1|^2] = \mathbb{E}[|x_2|^2] = 1$, we conclude the following condition must be true so that $x_1$ is decoded first:\

$$\frac{P_{BS1}|h_{BS1,r}|^2}{2^l - 1} > \frac{P_{BS2}|h_{BS2,r}|^2}{2^m - 1}$$  

(2)

\(^2\)It is possible that different rules are used for selecting the symbol to be decoded first or even a completely different IC scheme. Our central concept is to cancel the interference of the BS1 and extract the BS2 data block.
In the above, HDD stands for hard decision decoding. Thus, decoding at the final destination is conditioned on what is decoded at the relays. From the perspective of a communication system, this is a viable approach since the final receiver must have this knowledge in order to know how to equalize. Once the successful delivery of this packet is validated with a CRC code, the receivers increase the value of the measured throughput in case of a successful reception. This throughput is denoted as $\bar{R}(l,m)$ and it is a function of the MCSs used at the two BSs.

C. Relay Selection

In a multicast scenario, there may be many users, hence many potential relays. In order to select the best two relays for DSTCIF, we use a simple relay selection rule. In particular we consider that the nodes that receive the two strongest signals from both sources are selected as the two relays. More thorough investigation regarding relay selection in multi-source communications protocols can be found in [16].

D. Rate Estimation with FEC

While the PHY-layer communication scheme allows us to reach a throughput level indicated by $\bar{R}$, there might still be losses in the packet level. In this paper, we use an application-layer FEC encoder to handle the packet level losses. The basic idea of application-layer FEC is that redundant information is sent a priori by the source station, in order to be used by the receivers to correct errors/losses without contacting the source station. The advantage of using application-layer FEC for multicasting is that any parity packet can be used to correct independent packet losses among different nodes. This way, we can avoid the feedback implosion problem, which occurs when the source station is overwhelmed by feedback messages from the receivers in a large multicast system. However, such a scheme introduces overhead since extra parity packets are now transmitted by the source station. Furthermore, since the FEC is applied across packets, it also introduces additional delay which will be discussed in Section IV. Despite additional overhead and delay, considering the benefits for error recovery, such a scheme is widely used in a multicast environment.

With the proposed streaming system source packets are sent to the application-layer FEC Reed Solomon (RS) encoder. The RS encoder generates $I - J$ additional packets for $J$ input source packets, corresponding to a FEC rate of $\gamma = J/I$. FEC is applied across the source packets so that each generated transport packet contains parts of both the source payload and the parity bits. At the encoder, the source packets $J$ along with the parity packets $I - J$ are transmitted. Once the packets are transmitted, at the receiver, they pass through the application-layer RS decoder. The source packets can be recovered correctly at the RS decoder as long as the total number of received packets is greater than $J$. While evaluating the performance of the system, for given $J$ and average wireless packet error rate (PER) before FEC error recovery, we numerically determine $\gamma$ so that FEC decoding failure probability is below a threshold [13]. Note here that the selected RS code is the same for all packets. Due to the additional complexity of an unequal error protection (UEP) scheme, we do not include it here although performance benefits are expected.

E. Effective Video Rate

For calculating the effective video rate at the sender, we utilize the throughput estimate at the final destination $\bar{R}$. Therefore, if the video payload consists of $L_d$ bytes, and the combined protocol overheads is $L_h$ bytes, then the effective video rate is given by:

$$T = \frac{L_d}{L_d + L_h} \ast \gamma \ast \bar{R}$$

Note that in multicast, there are multiple users with different channel qualities, hence the effective rate for each user will be different. In conventional multicast, the sender arranges its transmission rate and the video rate so that the furthest away user received the video. In this paper, we consider a representative multicast user, S3, and compute the effective video rate based on that user. When there are multiple users, the sender can select the lowest effective video rate based on all multicast users as in the conventional multicast designs.

IV. Performance Evaluation

We implemented the proposed cooperative protocol and we evaluated its performance through Monte Carlo simulations. We consider the transmission of 5000 blocks with $L=1000$ bits each. We study three MCSs used by 802.11a [19] and in all the figures we provide their IDs as specified in the standard, namely MCS=2 (QPSK with coding rate 0.75), MCS=4 (QAM16 with coding rate 0.75), MCS=6 (QAM64 with coding rate 0.75). The channel is assumed to be block fading Rayleigh channel with average channel gains between all the nodes is equal to 1 unless otherwise specified. For each value of the transmit SNR, we tested different channel coding rates for all systems and we selected the optimal one to present in the figures. The system performance is calculated by considering the modulation scheme, channel coding, and $L$ whereas the video quality was calculated by taking into the application layer FEC and the R-D function of the layered video into account. The R-D function of the videos\(^3\) used are as in [13].

We first consider a basic setup where we have two BSs, a certain number of potential relays and a destination that is a representative multicast user. For the DSTCIF, the relay selection method always selects the two best relays among the potential relays. For the RDSTC case, relays are selected among the potential relays based on the successful packet reception at each packet under a total sum power constraints [13]. In other words, RDSTC scheme can use up to $N_p$ relays whereas DSTCIF always uses the best 2 relays and the total relay transmit power is fixed for both schemes.

\(^3\)Three different video clips (Soccer, Foreman, Bus; 352 288 ; 30 frames/s) are encoded with an H.264/SVC encoder using the JSVM software at a base rate of 110 kb/s. These videos possess a good variety of motion and texture characteristics.
We study symmetric links where the SNR between the BSs and relays are equal to the SNR between the relays and the destination. We vary the SNR and also the MCS and illustrate the BER and effective video rate results in Figure 3 and Figure 4 respectively. For a certain MCS, we observe that the proposed scheme outperforms R-DSTC scheme for every SNR value. We then consider the same setup for different videos. We assume the multicast session can use only the 10% of the total available effective video rate and we present the PSNR values for different videos in Figure 5. We observe quite different PSNR curves for different channel conditions due to different R-D curves of the videos, nevertheless, the DSTCIF outperforms the RDSTC for all different type of videos we considered for all the SNR values.

As discussed previously, FEC introduces an additional delay to the system. In a system that adds $I - J$ parity packets to each block of $J$ source packets, the receiver must wait for $I$ packets before FEC decoding. Therefore the maximum delay due to FEC decoding is the time needed to transmit $I$ packets. Since the proposed cooperative multicast system can support higher video rates, it also leads to a smaller delay compared to the orthogonal transmission.

We also investigate the effect of different number of potential relays ($N_p$) on the performance. In order to simplify the figure, for each SNR value, we choose the best MCS that provides the highest effective video rate. We then illustrate the effective video rate performance for different SNR values considering symmetric links in Figure 6. We observed that as the number of potential relays increase, the effective video rate values increase for both schemes. For the RDSTC, the effective video rate increases due to diversity even under a total sum power constraint. For the DSTCIF, the effective video rate increases since the system chooses the best two relays among the potential relays. For all different number of potential relays, the DSTCIF scheme outperforms the RDSTC.

We then evaluate the performance of the proposed scheme under different multicast performance metrics. We assume that the users are randomly uniformly distributed between 10dB and 40dB SNR range. We consider two different metrics: mean and min. The min value indicates the effective video rate at the edge of the multicast. The mean value indicates the average effective video rate of all the users. For both of the schemes considered, the best MCS has been choosen among all options and the results are illustrated in Figure 7. We observe that DSTCIF outperforms RDSTC over the all SNR range for both mean and min values. For higher SNR values the mean value is much higher than the min value since more users get high quality video.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a cooperative video multicast protocol for small cells with overlapping coverage. The small cells are allowed to have interfering transmissions in the first hop and transmit to the relays who follow the optimal strategy under interference conditions and apply successive interference cancellation (SIC) decoding. In the second hop, in order to increase the reliability of the system as well as improve the performance, the relays apply a Distributed Space Time Code (DSTC) that dynamically adapts itself based on the result of SIC. This cooperative transmission scheme is complemented with application layer FEC in order to handle the remaining packet level losses. We investigate the performance of such a scheme under different channel conditions and compare the performance with a scheme that applies orthogonal transmission in the first hop. Our results illustrate that the proposed protocol leads to superior performance compared to similar cooperative orthogonal transmission schemes.

In this paper, we consider the transmission of the same video from the two BSs to multicast users where BSs use the same power. A future direction is to investigate the performance when BSs can dynamically adapt their power to improve the performance. In this case, note that, the effective transmission rate of each BS will be different. However, due to interference, these rates also depend each other. Hence, we need a model to effectively estimate these rates jointly.

REFERENCES

Fig. 5: The video quality for different sequences ($N_p=8$).

Fig. 6: Effective video rate for different $N_p$.

Fig. 7: Average effective video rate and the effective video rate at the edge of the multicast with adaptive modulation and coding in broadband wireless data systems,” *IEEE/ACM Transactions on Networking (TON)*, vol. 20, no. 1, pp. 57–68, 2012.


