Optimized Network-coded Scalable Video Multicasting over eMBMS Networks

Andrea Tassi*, Ioannis Chatzigeorgiou*, Dejan Vukobratović‡ and Andrew L. Jones*

*School of Computing and Communications, Lancaster University, United Kingdom

‡Department of Power, Electronics and Communication Engineering, University of Novi Sad, Serbia

Abstract—Delivery of multicast video services over fourth generation (4G) networks such as 3GPP Long Term Evolution-Advanced (LTE-A) is gaining momentum. In this paper, we address the issue of efficiently multicasting layered video services by defining a novel resource allocation framework that aims to maximize the service coverage whilst keeping the radio resource footprint low. A key point in the proposed system mode is that the reliability of multicast video services is ensured by means of an Unequal Error Protection implementation of the Network Coding (UEP-NC) scheme. In addition, both the communication parameters and the UEP-NC scheme are jointly optimized by the proposed resource allocation framework. Numerical results show that the proposed allocation framework can significantly increase the service coverage when compared to a conventional Multi-rate Transmission (MrT) strategy.

Keywords—Unequal error protection, random linear network coding, multimedia communication, resource allocation.

I. INTRODUCTION

Over the last few years, the technological evolution of communication devices has fuelled a surge in demand for new multimedia services over fourth generation (4G) and next-generation cellular networks. To give an example, up to 67% of the global mobile Internet traffic will be represented by video content by 2018 [1]. Various standards for video coding and compression have been proposed to enable layered video streaming. Among them, the H.264 Scalable Video Coding (H.264/SVC) standard is gaining popularity [2].

The key idea behind H.264/SVC compression is the generation of a layered video stream, which consists of one *base layer* and multiple *enhancement layers*. The base layer ensures a basic reconstruction quality, which progressively improves with the number of recovered enhancement layers. Hence, a scalable video service can be transmitted to multiple network users and potentially be decoded by them at different quality levels, depending on the radio propagation conditions.

In this paper, we consider a 3GPP Long Term Evolution-Advanced (LTE-A) communication network multicasting H.264/SVC video services. Since its first release, the LTE-A standard manages Point-to-Multipoint (PtM) services by means of the evolved Multimedia Broadcast and Multicast Service (eMBMS) framework [3]. A PtM service can be delivered by using the Single-Cell (SC-) or the Single Frequency Network (SFN-) eMBMS modes. In the SC-eMBMS mode, the service is delivered by each base station independently of the others. On the other hand, the SFN-eMBMS mode allows multiple neighbouring base stations (forming the SFN)

to deliver the same PtM service in a synchronous fashion. Hence, base stations of the same SFN do not interfere with each other.

3GPP addresses the reliability issues of PtM communications by proposing the adoption of Application Level-Forward Error Correction (AL-FEC) schemes based on Raptor codes [4]. However, Magli *et al.* [5] noted that this kind of sparse codes, as well as LT codes, are usually designed to be applied over large source messages. As a consequence, the use of Raptor or traditional LT codes can lead to a non-negligible communication delay. As noted in [5], this issue can be mitigated by using Random Linear Network Coding (RLNC) strategies, applied over short source messages.

This paper draws inspiration from [6] and [7], which utilize the RLNC principle to protect the reliability of Point-to-Point (PtP) and PtM communications, respectively. Unlike [6] and [7], this paper takes into account the fact that the layers of a video stream have different importance levels and hence we adopt the Unequal Error Protection (UEP) implementation of RLNC [8], which we shall refer to as UEP-NC. UEP-NC allows the transmitter to adjust the error protection capability of the code according to the importance level of the transmitted video layer.

In addition to tackling reliability issues, the definition of efficient resource allocation models suitable for delivering PtM services is also a challenging task. Among candidate resource allocation models, the family of Multi-rate Transmission (MrT) strategies is the most fitting for exploiting the layered nature of a service [9]. MrT strategies attempt to split users into subgroups. The transmission parameters of each video layer are then optimized by considering the propagation conditions experienced in each user subgroup. In this way, all subgroups receive the same video stream but each subgroup is able to recover a different set of video layers and, thus, experience a different level of video quality [10], [11].

A MrT resource allocation model can be designed to achieve a variety of objectives. For instance, [10], [11] maximize the total service level experienced by all the users. On the other hand, [7] concentrates on minimizing the amount of radio resources needed to broadcast a video service. In contrast to [7], [10], [11], this paper interprets the number of recovered video layers as the *system profit*, while the amount of radio resources required to deliver the PtM service is considered the *system cost*. According to a fundamental economics principle, we propose a novel MrT-based allocation model, which we

call UEP Resource Allocation Model (UEP-RAM). Our model maximises the system profit-cost ratio and ensures that at least a desired fraction of users achieve a predetermined service level. A key aspect in the proposed UEP-RAM is that system and UEP-NC parameters are jointly optimized.

The rest of the paper is organized as follows. Section II presents the theoretical framework for the assessment of the service level achieved by each user. The integration of the UEP-NC principle into the LTE-A protocol stack is also presented. Section III describes the proposed UEP-RAM and provides a novel heuristic strategy to efficiently derive a good quality solution. Results are discussed in Section IV while the main findings of the paper are summarized in Section V.

II. SYSTEM MODEL

We consider an eMBMS network that delivers the same *target scalable video stream* to a multicast group of U users. If the eMBMS network consists of a single base station, the video service is transmitted according to the SC-eMBMS mode. On the other hand, if the eMBMS network is formed by two or more spatially contiguous base stations, they will form a SFN [3].

A. Unequal Error Protection Random Linear Network Coding

We denote by $\mathbf{x} = \{x_1, \ldots, x_K\}$ a *scalable* source message composed of K elements. As shown in Fig. 1, the elements of $\mathbf{x}$ can be grouped into L *layers*. Each layer signifies a different importance level and consists of a fixed number of elements. For example, the ℓ-th layer consists of k_ℓ elements. We assume that layers are arranged in descending order of importance, that is, the first layer – which comprises the first elements of the message – is the most important layer, while the L-th layer is the least important layer. We define the user Quality-of-Service (QoS) level as the number of *consecutive* layers that a user can recover, starting from the first layer.

In an effort to improve reliability, each source message is transmitted to users according to the RLNC principle [12]. A typical RLNC PtM scheme generates a stream of $N \geq K$ coded elements $\mathbf{y} = \{y_1, \ldots, y_N\}$ obtained by linearly combining elements of $\mathbf{x}$. Hence, coded element y_j is defined as $y_j = \sum_{i=1}^{K} g_{j,i} x_i$, where each *coding coefficient* $g_{j,i}$ is selected uniformly at random over a finite field $\mathrm{GF}(q)$ of size q. A user can recover the source message $\mathbf{x}$ as soon as it collects K linearly independent coded elements [12].

A key issue of conventional RLNC-based schemes is that a user is not able to recover all the K source elements if it has not collected at least K linearly independent coded elements. To address this issue, we adopt the UEP implementation of the RLNC [8], which we shall refer to as UEP-NC. In this case, instead of generating coded packets by linearly combining source packets over the whole message $\mathbf{x}$, the RLNC process takes place over a nested structure of L *expanding windows*. As depicted in Fig. 1, the ℓ-th expanding window $\mathbf{x}_{1:\ell}$ consists of all the source elements belonging to the first ℓ layers. In other words, $\mathbf{x}_{1:\ell}$ is formed by the first $K_\ell = \sum_{i=1}^{\ell} k_i$ elements of the source message $\mathbf{x}$.
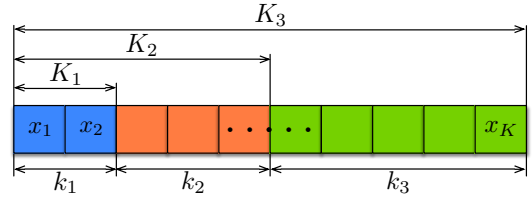

Fig. 1. Example of layered source message $\mathbf{x}$ comprised of $L = 3$ layers.

In this paper, we map coded elements associated with an expanding window onto Packet Data Units (PDUs). The PDU description and its role in the LTE-A protocol stack will be clarified in Section II-B. For the sake of simplicity, we assume that coded elements associated with different expanding windows are not mixed within the same PDU. The selection of the coding coefficients for the generation of each coded element is governed by a Random Number Generator (RNG). As explained in [6], the RNG seed associated with the first coded element in a PDU is transmitted along with standard LTE-A signalling information. The RNG seeds associated with the remaining coded elements in the PDU are then incrementally computed from that associated with the first coded element.

Let us focus on one of the N_ℓ PDUs associated with the ℓ-th expanding window $\mathbf{x}_{1:\ell}$. This PDU can hold n_ℓ coded elements[1], which are transmitted over a broadcast erasure channel using a Modulation and Coding Scheme (MCS) with index m_ℓ. The probability that user u will fail to recover the PDU is denoted by $p_{u,\ell}$. Note that the larger the value of m_ℓ is, the higher the modulation order or the lower the error-correcting capability of the corresponding MCS is. We can infer that the value of m_ℓ has an impact on both $p_{u,\ell}$ and n_ℓ; if m_ℓ increases, then n_ℓ and $p_{u,\ell}$ are also likely to increase.

Based on our adopted model, if N_ℓ PDUs associated with expanding window $\mathbf{x}_{1:\ell}$ are transmitted, $r_\ell \leq N_\ell$ of them will be successfully received by user u. We initially refer to the case where source elements of expanding window $\mathbf{x}_{1:\ell}$ can be retrieved from PDUs related to the first ℓ expanding windows. For convenience, we introduce sets $\mathbf{N}_{1:\ell} = \{N_1, \ldots, N_\ell\}$ and $\mathbf{r}_{1:\ell} = \{r_1, \ldots, r_\ell\}$ to enumerate all transmitted and received PDUs, respectively, which are associated with each of the first ℓ expanding windows. The probability that user u achieves the QoS level ℓ can be quantified by the probability $\mathrm{P}_u(\mathbf{N}_{1:\ell})$ of recovering the source elements of the ℓ-th expanding window, expressed as

$$\mathrm{P}_u(\mathbf{N}_{1:\ell}) = \sum_{r_1=0}^{N_1} \cdots \sum_{r_\ell=0}^{N_\ell} \overbrace{\prod_{i=1}^{\ell} \binom{N_i}{r_i} (1-p_{u,i})^{r_i} \, p_{u,i}^{N_i-r_i}}^{\mathrm{R}_u(\mathbf{r}_{1:\ell})} g(\mathbf{r}_{1:\ell}) \quad (1)$$

where $\mathrm{R}_u(\mathbf{r}_{1:\ell})$ represents the probability that user u successfully receives r_i out of N_i coded elements, for $i = 1, \ldots, \ell$. Function $g(\mathbf{r}_{1:\ell})$ is the probability of recovering the ℓ-th expanding window, given that user u has successfully collected the set of PDUs described by $\mathbf{r}_{1:\ell}$.

[1] As a consequence, the total number of transmitted coded elements associated with the ℓ-th expanding window is $N_\ell \, n_\ell$.

(a) $n_\ell = 2$ for $\ell = 1, \ldots, L$ (b) $n_\ell = 5$ for $\ell = 1, \ldots, L$
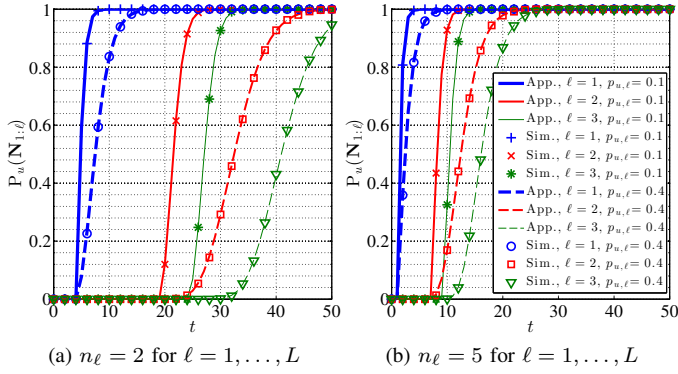
Fig. 2. Performance comparison between the approximated ("App.") and simulated ("Sim.") version of $P_u(\mathbf{N}_{1:\ell})$.

To the best of our knowledge, a closed form expression for $g(\mathbf{r}_{1:\ell})$ is not known. However, we know that if the field size q tends to infinity, the probability of receiving a coded element, which is linearly independent of previously received coded elements, approaches one. Following the rationale of [13] and considering large values of q, we set $g(\mathbf{r}_{1:\ell}) = 1$ if the number $r_\ell n_\ell$ of successfully received coded elements associated with the ℓ-th expanding window is equal to or greater than a threshold value $r_{\min,\ell}$; otherwise, we set $g(\mathbf{r}_{1:\ell}) = 0$. We can thus write

$$g(\mathbf{r}_{1:\ell}) \simeq I\left(r_\ell n_\ell \geq r_{\min,\ell}\right) \qquad (2)$$

where $I(s)$ is the indication function, i.e., $I(s) = 1$ if statement s is true, otherwise $I(s) = 0$. Using the same reasoning as in [13], we can express $r_{\min,\ell}$ as the following recursion

$$r_{\min,\ell} = K_\ell - K_{\ell-1} + \max\left(r_{\min,\ell-1} - r_\ell n_{\ell-1},\ 0\right) \qquad (3)$$

for an initial value of $r_{\min,1} = K_1$.

In this paper, we invoke approximation (2) for the calculation of (1). To validate the accuracy of the proposed approximation, we compare the probability values derived from (1) with those obtained by simulation. Fig. 2 shows both the approximated and simulated probability values of $P_u(\mathbf{N}_{1:\ell})$ in the case of a three-layer source message with $K_1 = 10$, $K_2 = 50$ and $K_3 = 100$, when the finite field has $q = 2^8$ elements. We assume that each PDU can hold the same number of coded elements ($n_\ell = 2$ or 5) and is affected by the same erasure probability ($p_{u,\ell} = 0.1$ or 0.4). Probabilities have been plotted as a function of t, where $N_\ell = t$, for $\ell = 1, 2, 3$. We observe that the performance gap between simulation and approximated calculations is smaller than $7 \cdot 10^{-3}$, which confirms that the proposed approximation is accurate for $q \geq 2^8$. In particular, in the rest of the paper, we will refer to $q = 2^8$.

B. Layered Video Multicasting over eMBMS Networks

Let us first consider a network operating in SC-eMBMS mode within which a single base station delivers a H.264/SVC video stream. The stream consists of L layers $\{v_1, \ldots, v_L\}$, where v_1 represents the base video layer, and $\{v_2, \ldots, v_L\}$ represent the $L-1$ enhancement video layers. We remark

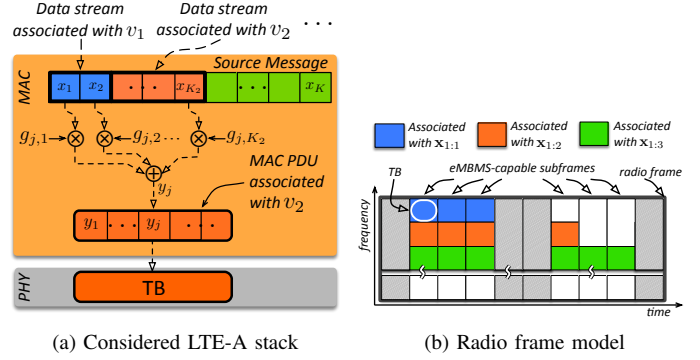
(a) Considered LTE-A stack   (b) Radio frame model

Fig. 3. Generation of a MAC PDU associated with $\mathbf{x}_{1:2}$ (left side) and TB allocation on an LTE-A radio frame for $L = 3$ (right side).

that the user QoS improves with the number of recoverable *consecutive video layers*, starting from v_1.

In our system model, each data stream associated with a video layer traverses the Packet Data Conversion Protocol (PDCP), the Radio Link Control (RLC) and the MAC layers of the LTE-A protocol stack of the base station. In this paper, similarly to the protocol stack presented in [6], we make use of a modified implementation of the standard MAC layer, which allows the transmission of the video stream by means of the UEP-NC principle as explained in Section II-A.

The output of an H.264 video encoder can be modelled as a stream of Group of Pictures (GoPs), each GoP has a duration of d_{GoP} seconds. Since the video decoding process takes place on a GoP-by-GoP basis [2], a GoP represents a layered source message. According to the layered model presented in Section II-A, the ℓ-th layer of a source message consists of k_ℓ elements. To fit our model, the data stream associated with the ℓ-th layer of a GoP is segmented by the MAC layer into k_ℓ source elements. If H is the number of bits comprising each source element and b_ℓ is the bit-rate of the ℓ-th layer of the GoP, the number of source elements forming the ℓ-th layer is $k_\ell = \lceil (b_\ell d_{\text{GoP}})/H \rceil$.

As shown in Fig. 3a, the modified MAC layer generates coded elements associated with the ℓ-th expanding window. All coded elements are then mapped onto one or more MAC PDUs, each of which can hold n_ℓ coded packets. The sequence of N_ℓ MAC PDUs is then forwarded to the physical layer. The LTE-A standard imposes that a MAC PDU is mapped onto exactly one physical layer Transport Block (TB). For this reason, we use the terms PDU and TB interchangeably.

Fig. 3b depicts the frequency-time structure of an LTE-A radio frame. Every radio frame comprises 10 subframes, each of which has a time duration of one Transmission Time Interval (TTI), which is equal to $d_{\text{TTI}} = 1\,\text{ms}$. Since at most 6 subframes can convey eMBMS data traffic [3], we use $f_{\text{eMBMS}} = 0.6$ to refer to the maximum fraction of eMBMS-capable subframes per radio frame. Notice that a TB is a fraction of the radio frame. It has a *fixed* time duration of one TTI, and consists of N_{RBP} Resource Block Pairs[2]

---

[2]A resource block pair can be modelled as a frequency-time structure spanning a bandwidth of $180\,\text{kHz}$ and a transmission time interval of 1 TTI.

(RBPs) [3], not shown in Fig. 3b. The number n_ℓ of coded elements that a TB can hold depends on the adopted MCS, characterized by index m_ℓ, the value of N_{RBP} and the element size H. Table I reports the values of the ratio n_ℓ/N_{RBP} for $H = 2$ KB and different MCS indexes. The reported values have been selected in order to fit the actual bit capacity of a TB [14].

Consider Fig. 3b, we assume that at most one TB associated with the same expanding window can be transmitted during the same subframe. The remaining part of the radio frame that has not been reserved for the delivery of the layered video service under consideration (marked in white) can be used by other LTE-A eMBMS services.

We now shift our focus from the SC-eMBMS to the SFN-eMBMS mode. In this case, base stations belonging to the same SFN are connected to the Multicell Coordination Entity (MCE), which is in charge of synchronizing the transmissions of all base stations [3]. The LTE-A standard imposes that all SFN base stations multicast the same video stream using the same physical signals. To this end, we assume that all RNGs, used by the SFN base stations during the network encoding process, are kept synchronized by the MCE. Consequently, all SFN base stations generate the same set of coded elements.

In both SC- and SFN-eMBMS modes, we assume that each user can provide Channel Quality Indicator (CQI) feedback on the downlink channel conditions. CQI feedback is sent to the only base station in SC-eMBMS or the base stations that is closest to the user in SFN-eMBMS. In the latter case, the reported CQI is forwarded to the MCE. In the CQI feedback, each user specifies the index of the MCS that ensures the maximum transmission rate and a TB error probability or, equivalently, a MAC PDU erasure probability not greater than $\hat{p} = 0.1$ [3].

III. SCALABLE VIDEO MULTICASTING OPTIMIZATION

In this section, we propose UEP-RAM to maximize the QoS level experienced by each user during the transmission of scalable video streams, while keeping the number of TB transmissions low. The constraint set of the proposed model ensures that a predetermined fraction of users shall achieve the desired QoS level with at least a given probability. In the case of SC-eMBMS, we assume that the proposed resource allocation strategy is performed at the MAC layer along with the scheduling operations defined by the standard. In the case of SFN-eMBMS, the MCE is in charge of scheduling services within the SFN. For this reason, we assume that UEP-RAM is also handled by the MCE.

We denote by $m^{(u)} \in [1, 15]$ the MCS index reported in the CQI feedback of a user $u \in [1, U]$. Upon receiving the $m^{(u)}$ values for any value of u, base stations employ the same MCS with index m_ℓ to broadcast TBs associated with the ℓ-th expanding window. If $m_\ell \leq m^{(u)}$ for a user u, the TB error probability $p_{u,\ell}$ experienced by that user will not be greater than $\hat{p}$. However, base stations cannot compute $\mathbf{P}_u(\mathbf{N}_{1:\ell})$ using (1) as they are not aware of the exact value of $p_{u,\ell}$. To alleviate

this problem, our proposed optimization strategy sets $p_{u,\ell} = \hat{p}$ if $m_\ell \leq m^{(u)}$, otherwise $p_{u,\ell} = 1$.

Let $\delta_{u,\ell}$ be an indication variable associated with user u and QoS level ℓ, so that $\delta_{u,\ell} = 1$ if the first ℓ service layers are recovered by user u with a probability of at least $\hat{Q}$; otherwise, $\delta_{u,\ell} = 0$. We understand that user u will recover the first ℓ service layers if it successfully recovers the ℓ-th expanding window or *any* of the expanding windows having index greater than ℓ. Thus, the indication variable $\delta_{u,\ell}$ can be expressed as

$$\delta_{u,\ell} = I\left(\bigvee_{i=\ell}^{L} \mathbf{P}_u(\mathbf{N}_{1:i}) \geq \hat{Q}\right) \qquad (4)$$

where $\vee$ denotes the logical OR operator.

We define the *profit* made by the system as the number of video layers that any user can recover with a given probability. Using (4), we can express the profit as $\sum_{u=1}^{U} \sum_{\ell=1}^{L} \delta_{u,\ell}$. On the other hand, we represent the *cost* incurred by the system as the number of required TBs for the delivery of a GoP, which is captured by $\sum_{\ell=1}^{L} N_\ell$. In order to optimize the profit while keeping the cost low, the proposed UEP-RAM aims at maximizing the *profit-cost ratio* as follows

$$(\text{UEP-RAM}) \quad \max_{\substack{m_1,\dots,m_L \\ N_1,\dots,N_L}} \sum_{u=1}^{U} \sum_{\ell=1}^{L} \delta_{u,\ell} \Big/ \sum_{\ell=1}^{L} N_\ell \qquad (5)$$

$$\text{subject to} \quad \sum_{u=1}^{U} \delta_{u,\ell} \geq U \hat{t}_\ell \qquad \ell = 1,\dots,L \qquad (6)$$

$$0 \leq N_\ell \leq \hat{N}_\ell \qquad \ell = 1,\dots,L. \qquad (7)$$

Objective (5) represents the system profit-cost ratio. Constraint (6) ensures that the fraction of users recovering the first ℓ video layers with a probability of at least $\hat{Q}$ shall not be smaller than $\hat{t}_\ell$. Constraint (7) imposes that the transmission of data associated with the ℓ-th expanding window shall not require more than $\hat{N}_\ell$ TBs. Of course, the transmission of a GoP should not exceed the GoP duration d_{GoP}. For this reason, the values of $\hat{N}_1,\dots,\hat{N}_L$ have to be such that $\max(\hat{N}_1,\dots,\hat{N}_L) \leq \lfloor f_{\text{eMBMS}} \cdot d_{\text{GoP}} / d_{\text{TTI}} \rfloor$.

Unfortunately, the UEP-RAM is a computationally complex optimization problem because of the coupling constraints among optimization variables introduced by (6). For this reason, we propose a novel heuristic strategy, summarized in Procedure 1, which can be used to find a good quality solution of the UEP-RAM in a finite number of steps.

Consider constraint (7) which allows N_ℓ to be equal to 0. In that case, coded elements associated with $\mathbf{x}_{1:\ell}$ will be neither generated nor transmitted. Let us assume that the first s expanding windows are not delivered, where $0 \leq s \leq L-1$. The desired service coverage can only be offered if (i) the first $(s + 1)$ service layers are recovered with a probability of at least $\hat{Q}$ by the user fraction $t'_{s+1} = \hat{t}_1$, and (ii) the remaining layers are recovered by the user fractions $t'_i = \hat{t}_i$, where $i = (s + 2),\dots,L$.

The rationale for the proposed heuristic can be summarized as follows. The value of s is initially set to $L-1$ (line 1 of

Procedure 1 Heuristic UEP-RAM

```
1:  s ← L − 1
2:  while s ≥ 0 do
3:  │   Nℓ ← 0 and mℓ ← 0, for any ℓ ∈ [1, . . . , L]
4:  │   t′ₛ₊₁ ← t̂₁, and t′ᵢ ← t̂ᵢ, for any i ∈ [s + 2, . . . , L]
5:  │   mℓ ← solution of S1-ℓ problem, for any ℓ ∈ [s + 1, . . . , L]
6:  │   for ℓ ← s + 1, . . . , L do
7:  │   │   if S2-ℓ can be solved then Nℓ ← solution of S2-ℓ
8:  │   if constraint (6) holds then
9:  │   │   m_INT,ℓ ← mℓ and N_INT,ℓ ← Nℓ, for any ℓ ∈ [1, . . . , L]
10: │   │   for ℓ ← L, . . . , s + 2 do
11: │   │   │   if Nℓ₋₁ > 0 and Nℓ > 0 then
12: │   │   │   │   mℓ ← mℓ₋₁, and Nℓ ← 0
13: │   │   │   │   if S2-ℓ can be solved then
14: │   │   │   │   │   Nℓ ← solution of S2-ℓ, and mℓ₋₁ ← 0
15: │   │   │   │   else if Nℓ > N_INT,ℓ then
16: │   │   │   │   │   mℓ ← m_INT,ℓ, and Nℓ ← N_INT,ℓ
17: │   │   if Σ^L_{ℓ=1} N_INT,ℓ < Σ^L_{ℓ=1} Nℓ then
18: │   │   │   return m_INT,1, . . . , m_INT,L and N_INT,1, . . . , N_INT,L
19: │   │   else
20: │   │   │   return m₁, . . . , m_L and N₁, . . . , N_L
21: │   else s ← s − 1
22: return no solution found.
```

Paramter		Value
Inter-Site-Distance (ISD)		500 m
System Bandwidth		20 MHz
Duplexing Mode		FDD
Carrier Frequency		2.0 GHz
Transmission Power		46 dBm per-sector
Base station and user Antenna Gains		see Table A.2.1.1-2 [15]
Pathloss and Penetration Loss		see Table A.2.1.1.5-1 [15]
Channel Model		ITU-T PedA [16]
H		2 KB
n_ℓ / N_{RBP}, for $m_\ell = 4, \ldots, 15$		2, 3, 5, 6, 8, 10, 12, 14, 17, 20, 66, 72
$\hat{Q}$		0.99
$\hat{N}_\ell$, for $\ell = 1, \ldots, L$, where $n_{\min} = \min_{m_\ell \in [4, \ldots, 15]} n_\ell$		$\lceil k_\ell / n_{\min} \rceil + \lceil \hat{p} \lceil k_\ell / n_{\min} \rceil \rceil$, $n_{\min} = 2 N_{RBP}$
Stream A [17] ($L = 3$)	$\{\rho_1, \ldots \rho_3\}$ [dB]	$\{27.9, 35.9, 45.8\}$
	$\{b_1, \ldots, b_3\}$ [Kbps]	$\{47.3, 326.1, 1396.7\}$
	$\{\hat{t}_1, \ldots, \hat{t}_3\}$	$\{0.99, 0.8, 0.6\}$
Stream B [18] ($L = 4$)	$\{\rho_1, \ldots \rho_4\}$ [dB]	$\{28.1, 33.4, 39.9, 46.4\}$
	$\{b_1, \ldots, b_4\}$ [Kbps]	$\{36.8, 79.4, 303.4, 835.9\}$
	$\{\hat{t}_1, \ldots, \hat{t}_4\}$	$\{0.99, 0.9, 0.75, 0.6\}$
Stream A, B	d_{GoP}	0.533 s

Procedure 1). The while-loop (lines 2-21) iteratively tries to find a feasible solution to the UEP-RAM. However, if constraints (6) and (7) are not met, then the value of s is decreased (line 21). In particular, for any $\ell \in [s + 1, L]$, the following problem is solved (line 5)

$$(\text{S1-}\ell) \quad \underset{m_\ell \in [1, \ldots, 15]}{\arg\max} \left\{ \sum_{u=1}^{U} I \left(m_\ell \leq m^{(u)} \right) \geq U\, t'_\ell \right\}. \quad (8)$$

The solution to (8) is the largest possible value of m_ℓ for which at least $U\, t'_\ell$ users experience a TB error probability equal to or smaller than $\hat{p}$.

Let $P(\mathbf{N}_{1:\ell})$ be the user-agnostic version of (1) representing the probability of recovering the ℓ-th expanding window when $\hat{p}$ is the erasure probability of all TB transmissions. The value of N_ℓ is obtained by iteratively solving, for $\ell = s + 1, \ldots, L$, the problem below[3] (lines 6-7)

$$(\text{S2-}\ell) \quad \underset{N_\ell \in [0, \ldots, \hat{N}_\ell]}{\arg\min} \left\{ P(\mathbf{N}_{1:\ell}) \geq \hat{Q} \;\wedge\; 0 \leq N_\ell \leq \hat{N}_\ell \right\}. \quad (9)$$

where $\wedge$ denotes the logical AND operator.

Both the MCS indexes and the number of TB transmissions associated to each expanding window obtained so far define the *intermediate solution* of the heuristic UEP-RAM $\{m_{\text{INT},1}, \ldots, m_{\text{INT},L}, N_{\text{INT},1}, \ldots, N_{\text{INT},L}\}$. If the solution of S2-$\ell$ does not meet constraint (6), the value of s is decreased and the while loop continues until all valid values of s are exhausted. If constraint (6) is fulfilled, a refinement procedure (lines 9-20) aiming at improving the quality of the intermediate solution is initiated. In particular, the MCS index m_ℓ for $\ell \in [s + 2, L]$ is potentially decreased by setting it to $m_{\ell-1}$, $N_{\ell-1}$ is set to 0 and N_ℓ is recomputed. We remark that if N_ℓ is set to zero, coded packets associated to $\mathbf{x}_{1:\ell}$ will not been generated and transmitted. If the refined solution requires a total number of TB transmissions which is smaller than that associated with the intermediate solution, the procedure

---

[3]Since S1-ℓ and S2-ℓ are characterized by small discrete feasible spaces, they are simple enough to permit the use of exhaustive search to find their solutions.

returns the refined solution. Otherwise, the procedure returns the intermediate solution.

IV. NUMERICAL RESULTS

The performance of the proposed resource allocation framework is investigated in this section. We consider a network composed of 19 base stations, each of which controls three hexagonal cell sectors. In the case of a SC-eMBMS network, 18 interfering base stations are organized in two concentric rings centred on the base station delivering the target scalable video stream. In the case of a SFN-eMBMS configuration, we consider a SFN formed by 4 base stations surrounded by the remaining 15 interfering base stations. Table I describes the main simulation parameters and target values for two H.264/SVC video streams, namely stream A and B [2].

To assess the quality of the UEP-RAM solutions generated by our proposed heuristic strategy, we shall compare them to resource allocation solutions obtained by directly solving the UEP-RAM using a genetic strategy [19], hereafter referred to as *direct solution*. Genetic approaches are known to provide a tight approximation of the optimum solution to the considered problem class [20]. Both the direct and the heuristic solutions shall be compared in terms of the profit-cost ratio $\tau = \sum_{u=1}^{U} \sum_{\ell=1}^{L} \delta_{u,\ell} / \sum_{\ell=1}^{L} N_\ell$ that they can achieve. For clarity, we remark that genetic strategies are characterized by a considerable computational complexity, which makes them unsuitable for use in practical network scenarios [19].

To assess the effectiveness of the proposed UEP-RAM, we compare it against the MrT resource allocation idea proposed in [10], [11], which relies on a standard LTE-A MAC layer and does not employ NC or AL-FEC error protection strategies. If ρ_ℓ denotes the highest Peak Signal-to-Noise Ratio (PSNR) achieved by user u after the recovery of the first ℓ layers [2], while $\dot{P}_{u,\ell}$ is the probability that user u will recover the first ℓ layers, the MrT strategy defines the *user performance level* as

$$\dot{\rho}_u = \max_{\ell=1, \ldots, L} \left\{ \rho_\ell \cdot \dot{P}_{u,\ell} \right\}. \quad (10)$$
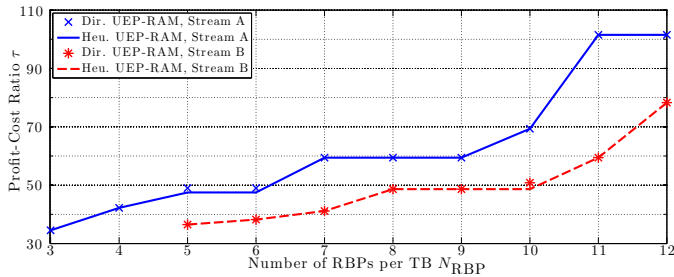
Fig. 4. Profit-cost ratio associated with the direct ("Dir.") and the heuristic ("Heu.") solution of the UEP-RAM model vs. the number of RBPs per-TB.

The considered MrT allocation method aims at maximizing the sum of all user performance levels by optimising the MCSs used for the delivery of each layer of the video stream. The MrT optimization problem can be summarized as

$$(\text{MrT}) \quad \max_{m_1,\dots,m_\ell} \sum_{u=1}^{U} \dot{\rho}_u \qquad (11)$$

$$\text{subject to} \quad m_{\ell-1} < m_\ell \qquad \ell = 2,\dots,L \qquad (12)$$

where constraint (12) allows the model to adopt a different MCS for each video layer and, thus, exploit user heterogeneity in terms of radio propagation conditions.

For the sake of clarity, note that we refer to the approximation of $p_{u,l}$, described in Section III only during the resource resource allocation operations. When assessing the user QoS, as presented in the following sections, we refer to the actual PDU error probability experienced by each user.

A. SC-eMBMS Delivery Mode

In order to better study the behaviour of the proposed UEP-RAM strategy, we adopt a user disposition pattern that ensures significant user heterogeneity. In the case of SC-eMBMS, users are placed on the radial line representing the symmetry axis of one of the cell sectors, which is controlled by the base station transmitting the target video service. We consider $U = 80$ users, which are spaced 2 m apart with the first user being placed 90 m from the base station transmitting the target video service.

Fig. 4 compares the profit-cost ratios associated with resource allocation solutions obtained by directly solving the UEP-RAM and by using the proposed heuristic strategy, as a function of N_{RBP} for both stream A and stream B. We observe that the performance gap between the direct and the heuristic solutions of UEP-RAM is negligible. Heuristic solutions deviate at most 4.3% from direct solutions in Fig. 4. Having established that the proposed heuristic strategy generates good quality solutions, we utilize it in the performance comparison of the MrT and UEP-RAM methods.

We first draw attention to the fact that $\hat{t}_\ell$, which is the fraction of users that shall receive the first ℓ video layers with probability of at least $\hat{Q} = 0.99$, can also be interpreted as a distance from the centre of the cell. For example, if we consider the value $\hat{t}_4 = 0.6$ and we recall the user disposition pattern that we have adopted, we understand that $0.6 \cdot 80 = 48$ of the users, who are located within a distance
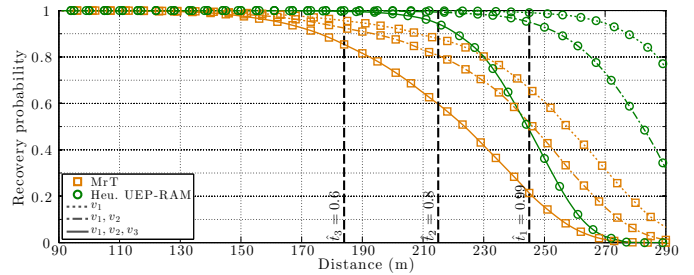
(a) Stream A
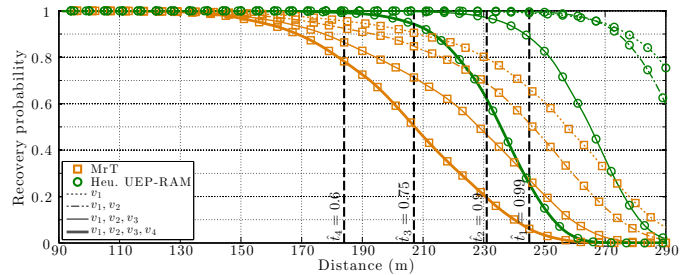
(b) Stream B

Fig. 5. Probability of recovering the first ℓ video layers (with $\ell = 1, \dots, L$) vs. the distance from centre of the cell.

of $90 + 47 \cdot 2 = 184$ m from the centre of the cell, shall recover the first four video layers with probability of at least $\hat{Q} = 0.99$.

Figs. 5a and 5b show the probability of recovering the first ℓ video layers as a function of the distance from the base station in the centre of the cell. Vertical dashed lines denote the target distance for each QoS level. We observe that the proposed UEP-RAM comfortably meets the service coverage constraints. In addition, UEP-RAM ensures a service coverage which is considerably greater than that of MrT. For instance, if UEP-RAM is applied on stream A (Fig. 5a), the coverage area will be 64% greater than that of MrT for base layer provisioning (i.e., layer v_1 only) and 41% greater for all-layer provisioning (i.e., layers v_1, v_2 and v_3). The same trend is observed for stream B (Fig. 5b).

B. SFN-eMBMS Delivery Mode

Similarly to the SC-eMBMS scenario, the user distribution in the SFN-eMBMS case is characterized by significant heterogeneity. We consider a network composed of $U = 1700$ users located at the vertices of a regular square grid with step 20 m. The user grid is placed on the playground in such a way that the area spanned by the SFN is uniformly covered, starting from inside the SFN.

In this section, we use the maximum PSNR experienced by a user as the performance metric. As in the previous section, the adopted user distribution allows us to convert the maximum PSNR at each user into contours of the playground. Consequently, a visual comparison of the coverage areas offered by MrT and UEP-RAM can be performed. We explained that the maximum PSNR for the MrT strategy, denoted by $\dot{\rho}_u$, can be obtained from (10). The formulation of the maximum PSNR for the proposed UEP-RAM can be easily adapted from (10)
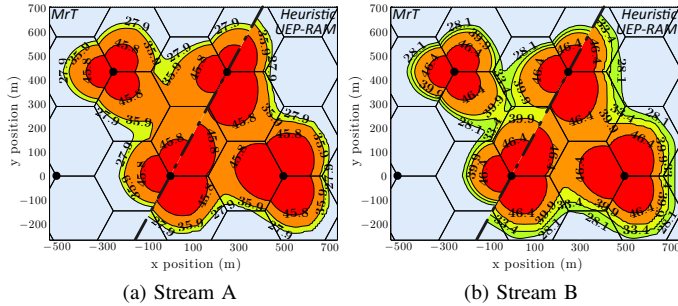
| | (a) Stream A | (b) Stream B |

Fig. 6. Maximum PSNR values (dB) associated with the MrT and heuristic UEP-RAM strategies (black circles represents the base stations).

as follows

$$\rho_u = \max_{\ell=1,\dots,L} \left\{ \rho_\ell \cdot \mathrm{P}_u(\mathbf{N}_{1:\ell}) \right\}. \qquad (13)$$

Fig. 6 shows the aforementioned performance metric associated with each point of the playground for the two resource allocation strategies and the two video streams when $N_{\mathrm{RBP}} = 5$ RBPs. Coloured regions depict areas where users can achieve a maximum PSNR value that is equal to or greater than that of ρ_ℓ shown in Table I. Table II reports the fraction of users in Figs. 6a and 6b that can recover the first ℓ video layers, for $\ell = 1, \dots, L$, with at least a target probability $\hat{Q} = 0.99$. We observe that the heuristic UEP-RAM ensures a service coverage which is considerably greater than that of MrT. For instance, the heuristic UEP-RAM resource allocation solutions allow 65% of users to recover all four video layers of stream B and all users to recover the base layer. In comparison, the MrT strategy successfully delivers all four video layers to about 33% of users, while the base layer is received by less than 40% of users.

V. Conclusions

In this paper we presented a viable method for the incorporation of the UEP-NC scheme into LTE-A networks as a means of improving the reliability of a H.264/SVC video multicast service. Furthermore, we proposed a novel resource allocation framework that aims to improve the service coverage with a reduced resource footprint. Inspired by a fundamental economics principle, the goal of the proposed optimization framework is achieved by maximizing the system profit-cost ratio. In spite of the natural complexity of the proposed modelling, we defined a novel heuristic strategy that was used to efficiently derive a good quality resource allocation solution of the considered problem. Numerical results showed that the proposed resource allocation framework ensures a service coverage which is up to 2.5-times greater than that of the considered conventional MrT strategy.

Acknowledgemt

TABLE II
Fraction of users recovering each set of video layers.

Recovered Layers	Stream A		Stream B	
	MrT	Heur. UEP-RAM	MrT	Heur. UEP-RAM
v_1	39.9%	100%	39.9%	100%
v_1, v_2	36.1%	81%	36.1%	100%
v_1, v_2, v_3	34.1%	74%	34.3%	87.4%
v_1, v_2, v_3, v_4	-	-	33.1%	65%

References

[1] Cisco Visual Networking Index, "Global mobile data traffic forecast update, 2013-2018," 2014.

[2] P. Seeling and M. Reisslein, "Video Transport Evaluation With H.264 Video Traces," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 1142–1165, Fourth 2012.

[3] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*. Elsevier Science, 2011.

[4] D. Lecompte and F. Gabin, "Evolved multimedia broadcast/multicast service (eMBMS) in LTE-advanced: overview and Rel-11 enhancements," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 68–74, 2012.

[5] E. Magli, M. Wang, P. Frossard, and A. Markopoulou, "Network Coding Meets Multimedia: A Review," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1195–1212, 2013.

[6] C. Khirallah, D. Vukobratović, and J. Thompson, "Performance Analysis and Energy Efficiency of Random Network Coding in LTE-Advanced," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4275–4285, 2012.

[7] A. Tassi, C. Khirallah, D. Vukobratović, F. Chiti, J. Thompson, and R. Fantacci, "Resource Allocation Strategies for Network-Coded Video Broadcasting Services over LTE-Advanced," *IEEE Trans. Veh. Technol.*, Jul. 2014.

[8] D. Vukobratović and V. Stanković, "Unequal Error Protection Random Linear Coding Strategies for Erasure Channels," *IEEE Trans. on Commun.*, vol. 60, no. 5, pp. 1243–1252, May 2012.

[9] R. Afolabi, A. Dadlani, and K. Kim, "Multicast Scheduling and Resource Allocation Algorithms for OFDMA-Based Systems: A Survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 240–254, 2013.

[10] D. Munaretto, D. Jurca, and J. Widmer, "Scalable video broadcast in cellular networks: impact on QoS and network resources," in *Proc. of IEEE ISCC 2010*, Riccione, Italy, IT, 2010.

[11] H. Zhang, Y. Zheng, M. Khojastepour, and S. Rangarajan, "Scalable video streaming over fading wireless channels," in *Proc. of WCNC 2009*, Budapest, HU, Apr. 2009.

[12] M. Médard and A. Sprintson, *Network Coding: Fundamentals and Applications*, ser. Academic Press. Elsevier, 2012.

[13] A. Tassi, I. Chatzigeorgiou, and D. Vukobratović, "Resource Allocation Frameworks for Network-coded Layered Multimedia Multicast Services," *IEEE J. Sel. Areas Commun.*, 2014.

[14] D. Vukobratović, C. Khirallah, V. Stanković, and J. Thompson, "Random Network Coding for Multimedia Delivery Services in LTE/LTE-Advanced," *IEEE Trans. Multimedia*, vol. 16, no. 1, pp. 277–282, Jan. 2014.

[15] 3GPP TR 36.814 v9.0.0 (Release 9), "Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects," 2010.

[16] S. Schwarz, J. Ikuno, M. Simko, M. Taranetz, Q. Wang, and M. Rupp, "Pushing the limits of LTE: A survey on research enhancing the standard," *IEEE Access*, vol. 1, pp. 51–62, 2013.

[17] "Video Traces 2 for Terminator." [Online]. Available: http://trace.eas.asu.edu/videotraces2/cgs/cif/Terminator1_G16B15_DQP15_2EL_48_33_18/

[18] "Video Traces 2 for Gandhi." [Online]. Available: http://trace.eas.asu.edu/videotraces2/cgs/cif/Gandhi_G16B15_DQP10_3EL_48_38_28_18/

[19] S. Sivanandam and S. Deepa, *Introduction to Genetic Algorithms*. Springer, 2007.

[20] K. Deep, K. Pratap Singh, M. Kansal, and C. Mohan, "A real coded genetic algorithm for solving integer and mixed integer optimization problems," *Applied Mathematics and Computation*, vol. 212, no. 2, pp. 505–518, 2009.