Analysis and optimisation of video transmission in LTE networks

Marco Martins Castanho, Luís M. Correia, and Ricardo Jorge Dinis

Abstract—This work focuses on the study of video transmission over Long Term Evolution (LTE), more specifically on the evolved Multimedia Broadcast Multicast Service (eMBMS). The service is assessed in terms of user satisfaction, analysing the video quality perceived by the user and the probability of the video playback to stop, both these metrics being based on the packet loss ratio. Several parametric scenarios are tested, where the influence of each parameter on the performance metrics is assessed. One of these parametric scenarios assesses the influence of eMBSFN being enabled or disabled based on its main properties. Additionally, two realistic scenarios are tested, based on real urban and rural networks. It is verified that the Quality of Service (QoS) provided to the user is always better when Multicast Broadcast Single Frequency Networks (MBSFN) is enabled. Base on the results from the urban scenario, all outdoor users are expected to be provided with at least an acceptable service. The coverage for indoor users was initially obtained as 25%, but an extrapolation based on the parametric results suggests that this value can raise to 91% under more advantageous conditions, such as lower base frequency and Modulation and Coding Scheme (MCS) index. Measurements performed in an indoor hotspot evaluate the impact of the MCS index in the performance of this service, and an analysis is performed on the balance between maximum throughput and coverage. As a low modulation enhances coverage, therefore serving more users, high modulation increases the maximum throughput, hence allowing for more simultaneous video streams. Results from simulations and measurements are compared.

Keywords—LTE, Video Transmission, eMBMS, MBSFN, MCS Index, Coverage.

I. INTRODUCTION

Mobile phones have become essential in our lives and made their own single stand. Once regarded as luxury, they are now the most indispensable device in our everyday life, replacing old tools and empowering new ways of communication through the emergence of smartphones and tablets. But no matter how advanced these devices are, they are nothing but uninteresting gadgets if they are not connected to a network. Therefore, the connection and the network itself have a major importance in mobile communications, with statistics showing that in Europe there are more mobile services subscriptions than inhabitants [1]. In fact, statistics show that the average mobile connection data rate grew 81% in 2013, with the number of connected mobile devices reaching approximately 7 billion, i.e., on average one for each human being in the world [2]. In 2012, mobile video traffic exceeded 50% of the global mobile data for the first time, with 51% of traffic by the end of the year. Also, it is expected that two-thirds of the global mobile data traffic will be video by 2017 [3]. In order to accommodate the increase in mobile data traffic organisations like the 3rd Generation Partnership Project (3GPP) have been developing standards for wireless communication of high-speed data for mobile phones, such as LTE.

The multicast standard for LTE, eMBMS, allows multimedia content to be sent once and received by many end users. This distribution mode can be a valuable alternative to unicast when a large number of users is interested in the same content, as an effective way for service providers to lower cost per bit. Also, data is sent over synchronised SFN where tightly synchronised, identical transmissions from multiple cells appear to the device as a transmission from a single large cell, since the several received signals are combined into a better resulting signal instead of being seen as interference.

The main scope of this work is to analyse the impact of several parameters in video transmission over LTE networks, from a user Quality of Experience (QoE) perspective, through scenarios implemented and simulated using the OPNET Modeler simulation tool [4]. Although this subject is not fresh, since LTE provides high bit rates that galvanises the usage of high demanding services, like video streaming, this work focuses mainly on the eMBMS service and its advantages over the primitive unicast transmission. The objective is to analyse video quality, as perceived by an end user, and to determine the conditions for an acceptable QoE to be provided.

This work is divided into five parts. Part II describes the performance parameters being analysed, Part III elaborates on the models used and/or developed for measuring the parameters and Part IV presents the analysis of the results obtained. Part V is accountable for the conclusions of the work made here.

II. PERFORMANCE PARAMETERS

Video quality evaluation can be classified into subjective and objective. Subjective video quality is concerned with how video is perceived by a viewer and designates his opinion on a particular video sequence. Subjective video quality tests are often expensive in terms of time and human resources. Objective video evaluation techniques are mathematical models that estimate results of subjective quality assessment, but are based on criteria and metrics that can be measured objectively and automatically evaluated by an algorithm.

The parameters that define quality in a video stream for IP video applications can be related to the video coding process or to the network through which the video is transmitted. The first ones include resolution, frame rate, compression. Higher resolution means that each individual picture frame is represented by more pixels, decreasing spatial granularity. The
frame rate is the number of frames shown per second. This is a measure of time granularity, and is usually represented in frames per second (fps). Together they strongly influence the video bitrate. Higher resolution and/or frame rate means better video quality, but it also means higher video bit rate, and hence a higher bandwidth needed to transmit the video. Compression techniques can be either lossless or lossy. As the name indicates, in the former no video quality is lost, although the compression ratios are usually not satisfactory for video streaming, while the latter strongly reduce the video bitrate and consequently the bandwidth required to transmit it, but also, as the name suggests, degrade video quality.

The main network related performance parameters that have an impact on the quality of IP video applications are packet loss and delay variation. Packet loss occurs when packets of data being transmitted across an IP network fail to reach their destination, or are corrupted on arrival. This can happen due to a large number of factors, such as network congestion, weak radio signals due to distance or multi-path fading, faulty networking hardware, or faulty network drivers. When reliable delivery is necessary, packet loss increases latency due to additional time needed for retransmission. When this is not the case and no retransmission is performed, packets experiencing high enough delays might be dropped, resulting in lower latency overall at the price of data loss. In terms of the impact this has on video, it can vary from a few disruptions in the image to total frame loss. Packet delay variation is the difference in latency between retransmitted and received packets. Even in the absence of packet loss, packet delay variation may cause the video to freeze until a frame with higher delay arrives to the decoder. The effects of packet delay variation in video streams can be removed by a properly sized play-out buffer at the receiver, which may only cause a detectable delay before the start of media playback, due to the time the buffer takes to be filled up to a minimum threshold.

III. SIMULATION AND MODELS DESCRIPTION

A. Video Quality Assessment Models

In order to evaluate the video quality based exclusively in the degraded video stream, only No-Reference (NR) models can be used. The first of these models used in this work is proposed in ITU-T Recommendation G.1070 [5]. The model incorporates network, application and terminal quality parameters of high importance for QoE planning and, opposite to previous recommendations, calculation of video quality is now done without the need of using a video signal, which is a key feature for this work as mentioned before. The input parameters used in the model are:

- Video codec specifications and codec implementation;
- Spatial resolution – theoretical spatial resolution employed in a codec;
- Video display size – the size of the user equipment (UE) display where video is being played;
- Key frame interval – the time interval in which the video is coded solely from intra-frame information;
- Video packet loss ratio, \( P_{pl} \) [9], referring to end-to-end IP packet loss ratio;
- Video frame rate, \( F_{fr} \) [fps], used in the encoder;
- Video bit rate, \( F_{br} \) [bps], at the encoder.

The video packet loss ratio parameter is obtained from OPNET. The remaining parameters are encoding parameters and are characteristic of the video sequence.

The output of this model is an estimation of subjective video quality, \( V_q \), in the form of an estimated Mean Opinion Score (MOS) – an arithmetic mean of the scores attributed by different individuals, which can range from 1 (worst) to 5 (best) [6]. The estimation for subjective video quality is calculated as:

\[
V_q = 1 + I_p I_c
\]

where \( I_p \) is a representation of the video quality degradation introduced by the packet losses in the transmission process, and can be expressed as:

\[
I_p = \exp \left( -\frac{P_{pl}}{D_{fr}} \right)
\]

where \( D_{fr} \) is the packet loss robustness factor and represents the degree of video quality robustness due to packet loss.

In (1), \( I_c \) is a representation for the basic video quality affected by the coding distortion under a combination of \( F_{fr} \) and \( B_r \), and can be expressed as:

\[
I_c = I_{opt} \exp \left( \frac{(\ln F_{fr} - \ln O_{fr})^2}{2 D_{fr}^2} \right)
\]

where \( D_{fr} \) is the degree of video quality robustness due to frame rate \( F_{fr} \) and \( O_{fr} \), an optimal frame rate that maximises the video quality at each video bit rate \( B_r \), and is expressed as:

\[
O_{fr} = v_1 + v_2 B_r
\]

If the video frame rate is the optimum that maximises video quality, \( F_r = O_{fr} \), then \( I_c = I_{opt} \), where \( I_{opt} \) represents the maximum video quality at a video bit rate \( B_r \), and is expressed as:

\[
I_{opt} = v_3 + \frac{v_4}{1 + (\frac{B_r}{v_5})^2}
\]

The degree of video quality robustness due to frame rate \( F_r \) is expressed as:

\[
D_{fr} = v_6 + v_7 B_r
\]

The packet loss robustness factor is expressed as:

\[
D_{pl} = v_{10} + v_{11} \exp \left( -\frac{F_{fr}}{v_{13}} \right) + v_{12} \exp \left( -\frac{B_r}{v_{14}} \right)
\]

Coefficients \( v_1, v_2, \ldots \) and \( v_4 \) are dependent on the first four input parameters listed above and can be found in [5].

In [7], authors propose a new metric representing the weighted percentage of slice loss, \( p_{w} \), which is expressed as:

\[
p_{w} = x_1 P_{pl1} + x_2 P_{pl2} + P_{plB}
\]

where \( P_{pl1} \) and \( P_{pl2} \) are the packet loss ratios in I, P and B frames, respectively, and \( x_1 \) and \( x_2 \) are two coefficients representing the average number of affected slices when there is an error in an I or in a P slice, respectively. Also, the authors have found that \( I_p \) can be correlated with \( p_{w} \), selecting the appropriate values for \( x_1 \) and \( x_2 \), as:

\[
I_p = \frac{1}{1 + k p_{w}}
\]

where \( k \) is a constant. The values of \( x_1, x_2 \) and \( k \) that minimise the Root Mean Square Error (RMSE) between the actual \( I_p \) values, derived from subjective tests, and the values obtained from (9) are provided in [7]. With these values, the obtained Pearson Correlation (PC) is 0.84 and the RMSE between the
actual and the derived $I_p$ values is 0.16, these values being much better than those obtained by using the G.1070 model.

Assessment of these models showed high correlation with each other. Consequently, only the model proposed in [7] is considered for results analysis.

B. Leaky Bucket Model

One of the parameters that affect QoS of a video broadcast service is the probability of occurrence of rebuffering events which occur when the decoder buffer is emptied and it provokes the video playback to stall. As this metric strongly depends on the size of the decoder buffer, the buffer size should be designed so that the decoder can decode the video bit stream without suffering from buffer overflow or underflow. For a proper choice of the buffer size the authors in [8] propose a hypothetical reference decoder (HRD) whose buffer size is determined using a leaky bucket model which, as the name suggests, is based on an analogy of a bucket that has a hole in the bottom through which any water it contains will leak away at a constant rate, until or unless it is empty. Water can be added intermittently, i.e., in bursts, but if too much is added at once, or it is added at too high an average rate, the water will exceed the capacity of the bucket, which will overflow. In this analogy the bucket represents the decoder buffer and water represents bits of information being added to and decoded out of the buffer. This analogy also suggests that the leaky bucket model is designed for Constant Bit Rate (CBR) channels, but it can also be used with Variable Bit Rate (VBR) channels.

The VBR case will be considered, since both the video trace file used in this work was coded with VBR and the radio channel is a VBR one. For these channels, having a sustainable peak bit rate $R'$ greater than the long-term average bit rate $R$ of the encoded video sequence, it is beneficial to characterise the encoded sequence using a leaky bucket with the higher leak $R''$ rate at the encoder, yet allowing the leaky bucket to underflow when the channel drains the bucket faster than the encoder can fill it. To illustrate, Figure 1 shows the encoding and decoding schedules using a leaky bucket with the higher leak rate. The later/lower bound on the encoding schedule is the schedule by which bits drain from the encoder leaky bucket and are transmitted or packetised. In the figure, flat spots in this transmission schedule indicate intervals in which the bucket is empty because there is nothing to transmit.

The earlier/upper bound on the encoding schedule represents the capacity constraint of the leaky bucket: an upward shift of the transmission schedule by $B'$ bits. Since the bucket drains at a rate $R'$ greater than $R$, the bucket can have a capacity $B'$ smaller than $B$, while still not overfilling, and can start with an initial state $F'$ smaller than $F$, where $B$ and $F$ are the buffer size and initial state, respectively, for the CBR case. The transmitted bits enter the decoder buffer after a constant transmission delay $\delta$. If, after the first bit enters the decoder buffer, the decoder delays at least:

$$D' = \frac{F'}{R'}$$  \hspace{1cm} (10)

seconds before decoding the first frame, then the decoding schedule is guaranteed not to underflow the decoder buffer. Furthermore, with delay $D'$, if the capacity of the decoder buffer is at least $B'$, then the decoding schedule is guaranteed not to overflow the decoder buffer. In either the CBR and VBR case a single leaky bucket is specifiable by three parameters $(R, B, F)$, where $R$ is the peak transmission bit rate (in bits per second) at which bits may leave the encoder buffer and enter the decoder buffer after a constant delay, $B$ is the capacity (in bits) of the encoder or decoder buffer and $F$ is the initial decoder buffer fullness (in bits) before the decoder can start removing bits from its buffer. $F$ and $R$ determine the initial or start-up delay $D$, according to (10).

C. Rebuffering Delay Models

In this section, two rebuffering delay models developed by the author of this work are described. For both of them the decoder buffer size is needed and is computed using the Leaky Bucket Model presented before. By applying it with the video trace file used for this work, the plot of leaky bucket parameters $(R, B)$, represented in Figure 2, was obtained.

The corresponding buffer size $B$ was obtained using linear interpolation between the 2 150 and 2 200 kbps points:

$$B = B_1 + \frac{B_2 - B_1}{R_2 - R_1} (R - R_1)$$  \hspace{1cm} (11)

where: $(R_1, B_1)$ is the point in the lucky bucket plot corresponding to 2 150 kbps and $(R_2, B_2)$ is the point in the lucky bucket plot corresponding to 2 200 kbps. The resulting buffer size is 2 183 kbits (approximately 273 kB), being the value used in the two rebuffering delay models presented following.

The first of the rebuffering delay models to be developed was based on the fact that a buffer with size $B$ bits, considered to be initially full, being leaked at an average rate of $R$ bits per second but not being filled, takes an average $D$ seconds to be completely emptied:

$$D = \frac{B}{R}$$  \hspace{1cm} (12)

Considering the buffer size obtained above and the average bit rate of the video sequence used in this work, 2 156 Mbps, the maximum rebuffering delay allowed for the tested video sequence is 1.44 s, which is rounded up to 2 s, because the sampling rate of the results statistics is set to 1 sample per second in OPNET. This means that, if a UE is 2 s consecutively without receiving any data, the buffer will be emptied and the video will stall. As a matter of fact, the UE does not need to stop receiving packets for 2 s to make the video to stall. From the video quality estimation models described before, it was observed that a MOS of 1, which is equivalent of no video, is achieved when the packet loss ratio is
above approximately 28%. This means that, if more than 28% of the information is lost, the decoder cannot decode the frames and no frame is displayed, hence, the model basic idea was reformulated: if a UE receives data with a packet loss ratio higher than 28% for 2 s consecutively, the video will stall. Applying the model to several UEs, it gives an estimation of the percentage of satisfied users, where a user is considered to be satisfied if the video never stalls during a simulation run. Another viewpoint is that, by applying the model to a single UE for several simulation runs, it gives the probability of that user being satisfied. This model is referred to as Max Delay model from this point on.

Since this is a very simplistic way of measuring the occurrence of rebuffering events, a more accurate model is proposed, trying to approximate what actually happens in a decoder buffer, therefore referred to as Realistic model. In the decoder buffer, bits are inserted at the rate at which they are received in the UE. After the buffer reaches a certain initial fullness \( F \) bits start being removed from the buffer at the rate at which they are decoded. As the packets are received, the buffer fullness is increased by the number of bits received. When the buffer reaches its initial fullness, defined here as 50% of the buffer capacity, bits start getting removed from the buffer at the rate defined in the video trace file, i.e., decreasing the buffer fullness. If the buffer fullness reaches 0, the video stalls and the decoder starts rebuffering. When removing bits from the buffer to be hypothetically decoded into a frame, the model ignores whether all the bits describing that frame where received or not at the UE, and the buffer fullness is decreased as if the bits were all available. This implies that even few packet losses (i.e., a number of bits equal to the buffer capacity being lost) during the entire simulation would cause the buffer to become empty, which is not realistic at all. As the Max Delay model, this one also required some adjustments. This was done by considering that, according to the packet loss ratio in each second, the UE either receives the total information it was supposed to, or it receives none. In other words, if the packet loss ratio in a second is low enough, it is considered that no losses occurred during that second and all the bits were received. On the other hand, if the packet loss ratio is high enough, it is considered that no packet was received and no bits are added to the buffer. For the reasons explained before, 28% was the value chosen for the packet loss ratio threshold. As in the Max Delay model, the purpose of this model is to estimate the percentage of satisfied users, i.e., users that never experiment rebuffering events, or the probability of a user being satisfied. Although this model is more accurate than the Delay model, it is still not very realistic, as rebuffering events might be provoked by events of total packet loss scattered over the whole simulation time and it is not supposed to happen.

Again, the assessment of these models showed high correlation with each other. Consequently, only the model Realistic model is considered for results analysis.

### D. Simulation Set Up

The setup for the base scenario, represented in Figure 3, is adapted from the default eMBMS scenarios of OPNET. As most modules remain unchanged for all scenarios, other will have their attributes and/or positions configured according to the scenario parameters.

Video is inputted in the simulator in the form of a video frame trace file, which is obtained from encoding of a video sequence. The simulator will generate random traffic based on the video frame sizes, i.e., with the same bit rate as the video, getting as close as possible to actually transmitting video. The video frame trace file used for this work was extracted from [9], being generated from the encoding of an excerpt of the film Terminator 2: Judgement Day. The parameters describing the trace file are summarised in Table I. Although the trace file was obtained from a 600 s video sequence, in most of the simulations only the first 100 s were used. This means that the values for the mean and peak frame bit rates in the simulations are not the ones represented in the table, but 1.516 Mbps for the mean bit rate and 2.188 Mbps for the peak bit rate.

![Figure 3. Basic simulation set up.](image)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Default values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encoder</td>
<td>H.264 FRExt</td>
</tr>
<tr>
<td>Encoding type</td>
<td>VBR</td>
</tr>
<tr>
<td>Frame Size</td>
<td>HD 1280×720 pixels</td>
</tr>
<tr>
<td>Frame Rate (fps)</td>
<td>30</td>
</tr>
<tr>
<td>GoP Size</td>
<td>12</td>
</tr>
<tr>
<td>Number of B frames</td>
<td>2</td>
</tr>
<tr>
<td>Quantiser</td>
<td>28</td>
</tr>
<tr>
<td>Mean Frame Bit Rate (Mbps)</td>
<td>2.214</td>
</tr>
<tr>
<td>Peak Frame Bit Rate (Mbps)</td>
<td>21.72</td>
</tr>
</tbody>
</table>

Each scenario was simulated 20 times, in order to assure statistical relevance. This number was obtained as the number of simulation runs from which both average and standard deviation values for the quality assessment models stop oscillating. After the simulation is completed, statistics of throughput, in packets per second, received at each UE, and throughput, in bits per second, received at the reference UE are collected and the quality assessment models are applied.

### IV. RESULTS ANALYSIS

#### A. Scenarios Description

The results analysis is done firstly by considering a simplistic reference scenario, and then 7 parametric scenarios are obtained by varying the one of the parameters of interest at a time, in order to assess their influence in the quality metrics previously described. Additionally, two realistic scenarios are studied, in urban and rural environments, both based on the real NOS network.

For the simplistic scenarios, the considered reference scenario consists of 2 eNBs in an urban scenario and 1 user between them. In this reference scenario, it is considered that the 2 BSs are separated by 600 m and that the UE is outside at 1/3 of the distance between the BSs, corresponding to a distance of 200 m from one of the BSs. The main parameters are summarised in Table II.

In the first scenario, the impact of increasing the distance between BSs is assessed. The distance between the 2 BSs is increased by 300 m in between simulations, starting at 300 m up to 3 000 m, with the UE always at 1/3 of the BS-BS distance from one of the BSs.

The second scenario has the objective of measuring the influence of the used frequency band, which is done by repeating the first scenario, but now for the other two frequency bands used by NOS: 800 and 1 800 MHz. Since the
choice of an adequate propagation model depends on the base frequency and the UMi model is only valid for base frequencies above 2 000 MHz, a different model should be used for these two bands. For base frequencies under 2 000 MHz the 3GPP Urban Microcell model is the most adequate. On the other hand, varying more than one parameter at once (base frequency and propagation model) may make it impossible to analyse the impact of one of those parameters alone, as wanted. Therefore, the first scenario presented (for the 2 600 MHz band) was repeated for this scenario, this time considering the 3GPP Urban Microcell as propagation model, in order to keep it coherent with the simulations for the other two bands. Knowing that the model is going to be applied with an important parameter out of its validity interval, some error is expected.

As stated in [11], the available bandwidth also depends on the frequency band, with only 10 MHz available for the 800 MHz band and 20 MHz available for the other two bands. For all bands (and in all scenarios), the entire available bandwidth is used. Also, for the 800 MHz band, additional 3 dB are considered for the BS maximum transmitted power, since it increases cell coverage and the use of this band aims for high coverage. Since path loss is expected to increase with the base frequency and the two bands assessed in this scenario have lower centre frequencies that the one from the first scenario, it is also expected that the BS ranges are higher in this scenario than they were in the previous one. Therefore, simulations are now made for BS-BS distances going up to 4 500 m. This means, following the same criteria as before for the UE position, that the UE-BS distance will now be ranged up to 1 500 m.

In the third scenario the objective is to assess how video quality is affected by UE mobility. For that purpose, two different tests were made: one where the UE is stopped in several positions; the other where the UE moves from one BS to the other at several different speeds. In the first case, the UE is set at a distance from the first BS ranging from 0 to 300 m, with a granularity of 50 m. For the second case, the values for the UE speed are 5, 20, 50, and 120 km/h. As UE speed is increased, the UE covers the distance between BSs quicker, which means that the simulation time needs to be adjusted to the UE speed, being the time it takes to cover the BS-BS distance. A consequence of this is that the number of samples collected, initially set to 1 sample per second, is now also varying. This would lead to an incoherence in the final results, with one curve being represented by 375 samples and another by just 18. In order to avoid this situation, the total number of samples in the effective simulation time was set to 50.

In the fourth scenario, the influence in video quality of a UE being indoor is measured. This scenario is simply repeating the stopped part of the previous one, but this time considering the UE to be indoor. In OPNET, this is made by selecting a different propagation model – the UMi Outdoor-to-Indoor. The model takes into account the angle between the LoS to the wall and a unit vector normal to the wall, and the distance between the UE and the wall. The first parameter is not configurable as it depends on the building and BS positions; the latter is configurable but the default value of 5 m was kept. With these parameters, the additional path loss ranges from 16.5 to 31.5 dB, depending on the angle. The results are then compared with the Outdoor case from the third scenario.

The fifth scenario has the objective of measuring the influence of varying the MCS index which, by definition, corresponds to varying modulation and coding. Simulations were made for the MCS index 10, 15, 19, 21 and 24, which were chosen in order to correspond to the values used in the measurements described further, performed prior to the scenarios definition. Contrary to the reference and the parametric scenarios described above, where the UE was set at 200 m away from the first BS, in this scenario that distance is changed to 300 m. The reason for this change is that, for MCS index 21, at 200 m distance (correspondent to the reference scenario), the video quality estimation is almost the maximum possible. As the quality is expected to increase when the MCS index is lowered and the MCS index 21 is the second highest tested value, for all the values under (at least) 21 no change in quality would be noticed. Therefore, one must deteriorate the channel conditions just enough so that the video quality estimation drops below the maximum, for all MCS indexes, without reaching the minimum, for any MCS index. This is achieved by setting the UE at a wider distance from the first BS, specifically at 300 m away.

In the sixth scenario, the influence in the quality estimation of considering the BSs from the first 3 interfering tiers is assessed. For that purpose, the reference scenario with only 2 BSs in compared with cases with 7, 19 and 37 BSs, corresponding to the first, second and third interfering tiers, respectively. The BSs are placed in a virtual hexagonal grid, with each BS at the centre of a hexagon (i.e., a cell). The hexagons radii are set to half the reference BS distance of 600 m, resulting that each BS is 600 m apart from each of its neighbours. Ten UEs are placed in a circular configuration around the centre BS, each one of them 300 away from that BS and not 200 m as in most other scenarios. The reason for this is same as presented for the fifth scenario; if the 200 m distance was kept, the video quality estimation would always be at its maximum value.

The seventh scenario is simply a repetition of the sixth, but this time with MBSFN disabled. The major impact of this change is that interference changes from constructive to destructive, which means that increasing the number of BSs in now expected to be prejudicial for the video quality.

The realistic urban scenario, as the name suggests, tries to simulate more accurately the delivery of an eMBMS service to several users in an urban scenario. The main aspect that gives realism to this scenario is that the positions of the eNBs were provided by NOS and are based on the operator’s real network in an urban area, where 3 interference tiers are considered.

Three simultaneous scenarios are considered: one where UEs are stopped indoor, one where UEs are stopped outdoor, and another where UEs are outdoor but with mobility. For the
stopped cases, UEs are set at 5 different distances from the centre eNB, ranging from 50 to 250 m, with granularity of 50 m. For the mobility case, the same speed values from scenario 3 are considered. Hence, 15 simultaneous cases are tested. For each case, 5 UEs are used resulting in a total of 75 UEs. Summing, in terms of UE parameters, this scenario is a mixture of scenarios 3 and 4, but with several BSs and several UEs for each case. This being an urban scenario, where capacity is usually an important constraint, the chosen frequency band is 2 600 MHz, since a 20 MHz bandwidth is available, enabling a higher network capacity. This means that the UMi propagation model is used, with the Outdoor to Indoor variant of the model used for the indoor UEs. The simulation time was reduced from the parametric scenarios to 40 s due to long simulation duration.

1) **Realistic Urban scenario**

The realistic rural scenario has the same objective as the realistic urban described above, but now for a rural environment. Again, the eNB coordinates were provided by NOS, based on the operator’s real network in an rural area. The most obvious difference from the urban scenario figure is the much wider BS-BS distances, resulting that, with the same number of BSs, the simulated area has now a width of approximately 90 km, compared to the approximately 3 km of the urban scenario. Knowing that, in rural environments, coverage is the main constraint, the 800 MHz band with the available 10 MHz bandwidth and the extra 3 dB of maximum BS transmission power are used. This means that the M.2135 Rural Macro (RMa) propagation model is used for the outdoor UEs. Since there is no outdoor-to-indoor model for rural environments defined in OPNET, the same model is considered, but with an extra 20 dB attenuation. This extra attenuation is introduced in the simulator by decreasing by 20 dB the UE antenna gain.

The same scenarios as before are tested, but now with wider distances between stopped UEs and the centre eNB. Stopped UEs are now placed at distances to the centre BS ranging from 400 to 2 000 m, with granularity of 400 m and the moving UEs move at the same speeds as before.

**B. Results Analysis**

Usually, a video service is regarded as perfect if it has a MOS of 5, but a service with a MOS of 4 is still regarded as having good quality and as being the least acceptable quality. This means that in what follows the conditions for obtaining a service with MOSs of 4 and 5, for each scenario, are evaluated.

The MOS results obtained for Scenario 1 are represented in Figure 4.5. Keeping in mind that MBSFN is enabled, it is expected that the perceived video quality decreases as BSs are more apart from each other, due to the reduction in the constructive interference. This effect is intensified by the fact that the UE-BS distance is also being increased.

From Figure 4, one can see that a MOS of approximately 5 is achievable for a BS-BS distance up to approximately 800 m, and a MOS of 4 is achievable for a BS-BS distance up to approximately 1 000 m. Again, for a UE at half the distance between the BSs, these results would be lower. Keeping in mind that these results are obtained for the 2 600 MHz band, which is mainly used to provide capacity in urban networks, and knowing that micro-cells in this type of environment have typical radii from 300 to 800 m, these results show that, under the conditions of this scenario, it is very likely that a user is provided with an excellent QoE.

The second scenario can be seen as an assessment of the eMBMS coverage for the 3 frequency bands, merged into Figure 5.

The results show that, the higher the base frequency is the lower the perceived quality, for a fixed BS BS distance. This can be interpreted from another angle: as the BS-BS distance increases, the higher the base frequency is the faster the service quality drops. The lower coverage for the 2 600 MHz results are explained by the use of a different propagation model.

For the 800 MHz band, the excellent quality is achieved for BS-BS distance up to approximately 1 000 m, and the minimum acceptable quality is achieved up to approximately 2 500 m. As this band is commonly the most used in rural areas, due to the low path loss, hence, the highest coverage, and knowing that typical rural cell radius are usually between 10 and 15 km, these results seem to be lower than expected. However, these results were obtained in an urban environment where path loss is much higher due to higher density and size of obstacles (buildings), and with a propagation model for urban micro cells. In an urban environment, the typical cell coverage in the 800 MHz band is approximately between 1 and 3 km, which makes the obtained results acceptable.

In the 1 800 MHz band, the results from Figure 5 show that an excellent QoE is provided within the first 500 m, and an acceptable service is provided to users up to 800 m apart from the reference BS. Knowing that the typical cell radius in a urban environments is between approximately 400 and 1 000 m, these results mean that it is expected that an acceptable service to be provided in most cases. If one considers the cell radii to be 600 m, there is an 83% coverage for excellent quality, and it is expected that 100% of the users have an acceptable service.

The results obtained from simulations of Scenario 3 are represented in Figure 6. As the user is stopped in several positions along the path from one BS to the other, almost no variation in the service quality is noted, as it is always excellent. As soon as mobility starts to be considered, some degradation in quality the half way between the BSs starts to be perceived, due to signal degradation due to a change in the signal power spectrum and fast fading caused by the Doppler shift.

For the 5 km/h case, the quality degradation in the half-way zone is still not enough to highly degrade the QoE, but it stops being excellent. As user speed is increased, the Doppler shift increases accordingly, raising the probability of the signal power to drop below the sensitivity level, hence leading to a decrease in QoE. This accentuated decrease in quality is...
observable in Figure 6 for almost all cases. For 20 km/h, the quality drops almost to the acceptable minimum, going below that level in higher speeds. Although the severity of the quality drop increases with user speed, its time duration obviously decreases. For the drop below the reference MOS 5, the duration may range from approximately 40 s for 5 km/h to just 2 s for 120 km/h; for the reference MOS 4, it does not last for more than 2 s for any case.

In Figure 7, the results for Scenario 4 are presented, with the goal of measuring the impact of Outdoor to Indoor transmission. The figure represents the results for a stopped UE (0 km/h) from the previous scenario, along with equivalent results but for an Indoor user.

The results show that an excellent QoE is only achieved in the proximity of 20 m to one of the BSs, and that an acceptable quality only within 60 m to a BS. This result seems somewhat sharp, meaning that one can only enjoy an excellent service at home if there is a BS on the top of the neighbour building, and assuming that the streets are less than 20 m wide. One possible reason for these severe results is that it is considered that the building is in NLoS with the BS, which is not always the case. In urban environments, BS are usually on top of buildings, which means that the near surrounding other buildings probably have LoS with the BS. Also, the use of the 1 800 MHz band should be considered instead of the 2 600 MHz one for these simulations, due to the impact on path loss. Another way to improve these results would be using a lower MCS index. Under poor channel conditions, a lower order modulation makes the signal more resilient to errors, hence resulting in a better QoE. The cost of this measure would be a lower throughput, meaning less TV channels available. Due to the additional wall penetration loss, results were expected to be lower than for the outdoor case, but probably not this lower. In conclusion, this result is not acceptable, as it suggests that eMBMS is not viable for indoors.

Making a rough estimation of how these results could be if the 1 800 MHz band was used instead, based on Figure 5, one can see that the 1 800 MHz band has a coverage approximately 100 m larger than the 2 600 MHz one. If that would be the case, the coverage for an acceptable service would increase to approximately to 120 m, which is still not good, but considerably better. If indoor coverage remains limited to 120 m, it means that only 40% of the cell (in terms of radius) is covered, which is still far below the objective of 90%. As for the gain in coverage by considering a lower MCS index, the only results collected that allow such analysis are the ones obtained for comparison with the measurements presented in Figure 8, which represent the variation of MOS as a user is moving away from an eNB, for each of the MCS indexes tested. This makes the estimation even rougher, knowing that these results were obtained for a moving outdoor user, in the 2 600 MHz band.

This figure shows that decreasing the MCS index from 24 to 15 leads to a gain in coverage for MOS 4 of approximately 80 m. With this one gets an estimated coverage of 200 m, which represents 60% of the cell radius, significantly better than previously, but still below the objective.

The results for Scenario 5 are presented in Figure 9, where the impact of the MCS index is assessed. With the exception of the first value presented in the figure, one can observe that the video quality decreases as the MCS index is increased. This is due to the fact that, as the MCS index is increased, the available throughput increases, but the transmission is less robust to errors. As the scenario is simulated with the same video sequence for all the cases, only the latter effect is evident. In the conditions of this scenario, the results show that it is expected that an excellent QoE is provided when the MCS is under index 19. Moreover, they also show that an acceptable video quality is expected to be achieved for all the MCS indexes tested. This counter-balance between available throughput and channel error robustness (which influences coverage), adjusted by the MCS index, allows the operators to tune the transmission according to the desired service.

For Scenario 6, which assesses the impact of increasing the number of BSs, the results are shown in Figure 10. Although the effect is very subtle, by zooming into the MOS scale, one can observe an almost linear improvement in the perceived quality as the number of BSs increases. The justification for these results comes from the properties of MBSFN, i.e., interference becomes constructive. This means that all signals arriving the UE from a BS belonging to the MBSFN cluster where the UE is located are combined into a better signal. Hence, the quality of the resulting signal is expected to be higher if more BSs are considered for the cluster, resulting in a higher quality video service.
In Figure 11, the results for Scenario 7 are presented, showing the same curve from Figure 10 along with a new one obtained under similar conditions, with the only exception of MBSFN being now disabled. As a consequence, each UE will be receiving unwanted signals from the surrounding BSs, in the same frequency in which the UE is receiving his wanted signal, probably a delayed version of the wanted signal, since all users in the network are considered to be using the same service. This makes it difficult for the UE to distinguish between his signal and the interfering signals it receives and consequently reduces the service quality. As more surrounding BSs are considered, more sources of interfering signals exist, hence poorer service quality. As a result of this, nearly the opposite of what happened in scenario 6 is expected and verified to happen now. As can be seen in the figure, higher slope from 2 to 7 BSs exists, but no zooming is needed now for it to be noticeable. The reasons for this is that the impact of the first interfering tier (7 BSs) is much higher that the remaining tiers. The impact of increasing the number of BSs in this scenario becomes worse knowing that no frequency re-use schemes are considered and that BSs are omnidirectional by default.

The results for the urban and rural scenarios are presented in Figure 12 and Figure 13, respectively. The reason for dividing the results in two parts for each scenario is that 3 different cases are assessed and in two of them the variable parameter is distance, whereas in the third it is speed. In each of the figures, each square represents the MOS of the 5 users in very similar channel conditions over the whole simulated time. For the UEs moving case, each of the 5 users moving at the same speed is at a different position, hence experiencing different channel conditions.

Regarding the urban scenario, the orange line on the left of Figure 12 shows that, for the users being closer to the centre BS the perceived quality is slightly higher than for the users being in a zone in between BSs, because they are always well covered by the same large number of BSs in the scenario. In fact, some of the users who are 250 m away from the centre are less than that distance away from a tier 1 BS. Thus, it would not be surprising if MOS would start increasing in the highest distance values simulated. As for the blue line in the same figure, it represents users in the same configuration as before, but now indoors. In this case, the perceived quality rapidly decreases for users who are at larger distances to the centre. This figure represents the same assessment as made for scenario 4. In fact, despite having different network configuration in terms of BSs positioning, Figure 12 (left) show high resemblance with Figure 7. Two subtle differences can be observed between the figures: first, in the realistic case, the acceptable quality level is achieved for a slightly larger distance; second, in the realistic case, MOS never exactly reaches 1. The explanations for both aspects might be the same: there are more BSs in the realistic case and the average inter-BS distance is marginally shorter than in the parametric scenario.

The same analysis of coverage estimation for the indoor case that was made for Scenario 4 can, and must, be made here. In this scenario, it would result in coverage increasing from approximately 70 to 250 m, corresponding to approximately 91% of cell radius coverage, since the average inter-BS distance for this scenario is approximately 550 m. With this correction, the result is already within the objective cell coverage of 90%. The results for the moving users in this scenario are represented by a grey line on the right of Figure 12. The fact that, even for low speeds, the MOS does not reach 5 is explained by the fact that the 5 users in each case move along parallel trajectories 250 m apart from each other. This means that each of these 5 users has different channel conditions, in some cases moving in zones away from any BS. This is compensated by the fact that other UEs always move close to one or more BSs. As a result, the perceived quality is always acceptable. This figure can be seen as different way of representing the same type of results as the ones obtained for Scenario 3. There, several lines represented the instantaneous video quality as a user moved. Now, each square represents the average quality of five users during the whole run. Despite of this, the results from both scenarios are coincident.

In rural environments coverage is the main concern, since there is not usually a number of users large enough to overflow network capacity. Additionally, path loss is lower than in urban environments, due to absence of tall buildings, which justifies that BSs are more separated in rural than in urban environments. Consequently, larger UE-BS distances were considered for this scenario, in comparison with the urban scenario. For this scenario, UE-BS distances up to 2 km for the indoor case and up to 5 km for the outdoor are considered, which, in the latter, correspond to half of the typical distance between BSs. Looking at the left side of Figure 13, the most obvious result is that the coverage for indoor UEs is immensely lower than for outdoor UEs. An acceptable service quality is achieved for up to approximately 600 m apart from a BS, in the case of indoor users, whereas for the outdoor ones, that the maximum distance increases to approximately 3 km. This means that the coverage for an acceptable video quality is approximately 5 times higher outdoors than it is indoor. In this scenario, both cases represent not acceptable results. If one considers the inter-BS distance to be 10 km, consequently the cell radius to be 5 km, the cell coverage for an acceptable service quality is merely 12% and 60% for the indoor and outdoor cases, respectively. In this scenario, one cannot make the same approximation made for the previous indoor scenarios, first, because the lowest frequency band is already being used, and second, because there are no results collected for the influence of MCS index in rural environments. Even so, it is expected that the decrease of the MCS index brings the cell coverage percentages to more acceptable values.

Comparing this figure with the equivalent for the urban scenario, the most obvious difference in the lines representing
the outdoor case is that, where in the rural scenario MOS decreases to 1, in the urban one there is no significant drop in quality. The seemingly most evident reason for this fact is that the distances tested are not the same in both scenarios. However, for each scenario, the maximum distance tested is approximately half of the typical distances between BSs for that scenario, which means that in that case, UEs are in the middle distance between BS. The explanation for this difference can be taken from the results in Figure 4 and Figure 5. The urban scenario uses the 2.600 MHz band, and BS-BS distance is, on average, approximately 500 m. From Figure 4, one can see that, with such distance, MOS is still far from dropping above 5, and in this scenario only 2 BSs were considered. The equivalent results for the 800 MHz band used in the rural scenario, represented in Figure 5, show that MOS never quite reaches 5 and drops below 4 when BSs are approximately 2 km apart, whereas BSs in the rural scenario are, on average, approximately 8 km apart.

The results for the case of UEs moving are presented on the right of Figure 13. The figure shows that MOS is almost constant, with a very subtle decrease as the UE speed is increased, showing similarity with its equivalent for the urban scenario presented in Figure 12. The fact that the rural line is almost constant at a lower value than the urban one is justified with the same arguments as for the differences for the outdoor stopped case. In the urban scenario, the moving UEs are positioned up to 500 m away from the centre BS (which means that they may be less than that value away from some tier 1 BS), hence their perceived quality is near perfect. In the rural scenario, these UEs are now positioned up to 2 km, which means that the quality of the video they receive is on the edge of not being acceptable.

C. Measurements

1) Procedure

Assessment of a real functioning eMBMS service was performed at NOS headquarters in Lisbon. This assessment consisted of measuring the quality of the received signal on the UE and correlating it with the subjective quality of the video being played, for 5 different MCS indexes. BS antennas located in the 1st to 4th floors of the 10-storey building and measurements were made at 60 different points for each MCS index, in the 2nd and above floors of the building, using an UE, connected to a laptop via the USB port. Real-time values of the Signal-to-Noise Ratio (SNR), Reference Received Power (RRP), Reference Signal Received Quality (RSRQ) and Received Signal Strength Indicator (RSSI) received at the UE were read from a data acquisition software tool installed on the laptop. The UE was playing a live TV news channel, from the few available, with the video coded in H.264/AVC and the audio coded in MPEG-2. At each point, the values mentioned above, shown on the software tool, and the floor at which they were taken were registered and a subjective evaluation of the video quality was made, based on a 5 grade MOS scale.

2) Results

The values read from the software tool and the subjective quality values were stored on a spreadsheet and later converted into graphs, such as showed in Figure 14, where, for each MCS index, the SNR and the subjective quality values of each measurement point are represented by black dots, as well as a logistic trend line that makes an approximation of the real dependency of MOS on SNR.

By analysing the black dots only, in any of the graphs, one can observe three distinct zones: a first zone where all the points have an MOS of 1, no matter how small SNR is, since this is the lowest score possible; a second zone where, as expected, the MOS increases approximately linearly with SNR; and a third zone where all the points have an MOS of 5, no matter how large SNR is, since this is the highest score possible.

More interesting than reading the figure above alone is to merge them and to analyse the differences between the figures for each MCS index, i.e., to analyse the impact of the changing of MCS index in the MOS vs. SNR curves. Consequently, the trend lines for the five cases were merged into a new one represented in Figure 15. Analysing the trend lines in the figure together, one can observe that they all have approximately the same slope in zone two, which means that the score is affected equally by variations (not absolute values) in SNR for all MCS indexes.

On average, an increase of approximately 3 dB in SNR results in the score increasing by 1. More interesting than that, one can also observe that the lower the MCS index is, the closer the correspondent curve is to the vertical axis, and vice-versa. This demonstrates that, as expected, in order to keep a constant score, as the MCS index is increased, the radio channel conditions (i.e., SNR) have to increase as well. Analysing this from another angle, for a constant SNR, an increase in the MCS index results in a degradation of the video quality and therefore in a lower score.

D. Comparison of Simulation and Measurement Results

As that the main analysis made for the measurements results focuses on the impact of the variation of the MCS index in the perceived quality, by means of analysing the required SNR for each quality score to be achieved, comparison with the simulations is only possible having equivalent results obtained from simulations. Among the scenarios described before there is in fact one scenario where the impact of the variation of the MCS index is assessed, thus new simulations are performed in order to obtain SNR statistics for that scenario. The results obtained from the new simulations were first represented in figures similar to Figure 14 and then merged together in Figure 16.

The major difference between the results lies in their nature, i.e., the measurements results are based on a statistical mensuration whereas the simulations results are based on deterministic models implemented in the simulator. As a consequence, under the same conditions the simulations results are always the same, whereas in the measurements it is
impossible to assure constant channel conditions. Another major difference lies in the reliability of the MOS results. As the measurements results are, as it suggests, truly subjective evaluations of the video quality made by an actual user, the simulations results are rough estimations employing simplistic and inaccurate estimation models.

![Figure 16. Logistic trend lines for the simulations results.](image)

The most obvious difference between the figures is in the interval between the lowest and the highest MCS index curves. In the measurements results, this interval is approximately 8 dB whereas in the simulations figure it is approximately 3 dB. This means that the impact of changing the MCS index is much higher in the measurements than it is in the simulations. Another difference between the figures is that the curves from the simulations results are shifted to the right in approximately 15 dB comparing with the measurements. This means that, in the simulations, better channel conditions are required for the same quality to be achieved, using the same MCS index.

V. CONCLUSIONS

Most of the results obtained in the parametric simulations are satisfactory. The results from the realistic scenarios are not so acceptable, since they were obtained under adverse conditions. Although the results from the measurements show a clear relationship between channel conditions, video quality and the influence of the MCS index, these measurements were performed in an indoor hotspot environment inside an office building, not being representative of the more common scenarios where BSs being outside and users both indoors and outdoors.

The major outline conclusion drawn by the results is that eMBMS improves the quality of a video streaming. The improvement in the perceived quality is clearly visible in the parametric scenario that addresses it. One of the objectives of this thesis from the point of view of MOS was to assess whether this service is viable and if it is worth the necessary investment. This is done by assessing whether it is possible to provide a good QoE to all customers in different conditions and which are the necessary conditions for that. Focusing primarily on the results from the realistic scenarios, one would be tempted to conclude that this service is not of interest for MOS since most indoor clients would not be satisfied, with only 25% and 12% of the cell radius being covered with an acceptable service for the urban and rural scenarios, respectively, under the simulated conditions. However these results were obtained in very unfavourable conditions, i.e., using high modulation and coding for both scenarios and using the frequency band with the highest path loss for the urban scenario. The results obtained from the parametric scenarios allow for a rough estimation of what these percentages of coverage could be under more favourable conditions. The percentages of satisfied users were then estimated to increase to 91% for the urban if more favourable conditions had been used. For the outdoor case, the results for the urban scenario show that all users can be provided an excellent service, whereas in the rural scenario an acceptable service can only be provided in approximately 60% of the cell radius. Again, this result is unacceptable for the operator, but using a smaller modulation order could improve the service coverage and consequently the percentage of satisfied clients. The impact of mobility provokes the perceived quality to slightly decrease from near-perfect to just acceptable in urban environments, which is guaranteed to users moving at velocities below 120 km/h. In rural environments, almost no impact is noticed, with the quality always being just above acceptable. Again, this could be improved by choosing a lower MCS index. Summarising these results, in the urban environment only indoor costumers can have complaints, even in the most favourable conditions. A possible solution could be the use of indoor hotspots. In the rural environment, all 3 cases tested have low coverage under the simulated conditions. As already mentioned, the necessary solution is to use lower order modulation by setting a lower MCS index. Another important conclusion is that the use of the eMBMS service in the 2600 MHz band is not advised.

Regarding future work, it would be ideal to simulate the transmission of real video sequences and to use more accurate quality estimation models. Also, the measurements were made only on an indoor hotspot scenario. Measurements for outdoor, outdoor-to-indoor and outdoor with mobility could improve this work and validate the simulations. Also, it would be interesting to consider additional users in the network using other LTE services. The assessment of the influence of this additional users in the QoS perceived by the eMBMS users would be an interesting aspect to be assessed in the future.

REFERENCES