VoIP Quality Monitoring in LTE Femtocells

Cristian Olariu
and Mícheál Ó Foghlú
Telecommunication Software and Systems Group
Waterford Institute of Technology,
Cork Road, Waterford, Ireland
Email: {colariu, mofoghlu}@tssg.org

Philip Perry
and Liam Murphy
School of Computer Science and Informatics
University College Dublin
Belfield 4, Dublin, Ireland
Email: {philip.perry, liam.murphy}@ucd.ie

Abstract—The increasing number of users demanding voice and data communication through cellular networks has driven the need for higher network throughput rates and lower latency. LTE femtocells address this pressing problem by offloading cellular service providers networks and increase both coverage and capacity for their users. Assuming a wired DSL backhaul for these femtocells, this paper shows simulations exploring a case where the DSLAM represents the main bottleneck when the cellular network operator and the DSL provider do not collaborate. This paper introduces the concept of Intermediary Mean Opinion Score which may be employed at femtocell gateways to isolate network problems and feed into customer experience management. We also propose and investigate a technique of mapping the human audio recency into the MOS calculation. Results are presented to illustrate the information that can be extracted from a lightweight monitor in the network.

Index Terms—VoIP, MOS, femtocells, LTE, network management solution.

I. INTRODUCTION

The use of small cells deployed in residential premises is often referred to as femtocell technology. This offers a way to improve cellular coverage for indoor users and also increases network capacity with minimal infrastructure costs. Such femtocells have been deployed for 3G cellular networks with future deployments expected to be based on *Long Term Evolution* (LTE) and *Long Term Evolution Advanced* (LTE+) standards. Other wireless access methods like WiMAX or CDMA2000 have been considered too but with fewer deployment initiatives.

Most commonly, femtocells are expected to use the subscriber's broadband connection to carry the traffic to and from the cellular network [1]. In 3GPP's Release 8, femtocells are denoted with *Home enhanced Node B* (HeNB). Figure 1 depicts a typical femtocell deployment in the LTE context. It can be seen that a femtocell can provide service to any equipment featuring a cellular interface (e.g. cellphone, laptop, PDA). All these types of equipment are generically called *User Equipment* (UE).

Femtocell's backhaul connectivity is provided by a Broadband Router which may be accessed by other network equipment like a video client (e.g. set-top box). The traffic originated in these equipments is in direct contention for Internet bandwidth with the femtocell originating traffic and we assume that the broadband router has priority classes enabled. A cloud symbolizes the network between the *Internet Service Provider* (ISP)'s network and the *Home enhanced Node B Gateway* (HeNBGw).

For the immediate future, the main cellular application will continue to be voice telephony which, from a backhaul perspective, will appear identical to a *Voice over Internet Protocol* (VoIP) call. As the migration towards femtocellular deployments will coincide with an increase in high bandwidth multimedia applications [2] which require the delivery of high quality video content over the broadband connections, the broadband access network will become a major bottleneck [3].

The anticipated increase in the number of base stations arising from femtocell deployment presents severe scalability problems for the management and control functions of cellular networks. Some of these problems will be solved by the increasing use of autonomic nodes and a convergence of management and control functions [4]. Novel monitoring techniques are a key requirement to implement these autonomic control and management loops.

From data traffic point of view, VoIP is regarded as a stream of relatively small packets that are time sensitive. Network delay, jitter, and packet loss rate are taken into account whilst measuring the overall or instantaneous objective quality of a conversation. These network metrics are combined in ITU-T's *Quality of Service* (QoS) metric, called Transmission Factor (denoted R). From this factor the commonly used *Mean Opinion Score* (MOS) is directly obtained. Although MOS is used to calculate the quality of a conversation on a end-to-end basis, we propose that this metric could be calculated at intermediary points so that the degradation of the call between the UE and that intermediary point could be estimated.

For femtocells we propose that one *Intermediary Mean Opinion Score* (IMOS) should be calculated between the UE and the HeNBGw. This IMOS will then play an important role in the admission control of calls through that HeNBGw. Figure 2 depicts our proposed implementation. The ISP's network and the HeNBGw are connected via the *Metropolitan Area Network* (MAN). The voice packets to and from the HeNB should be analyzed at the HeNBGw and a corresponding IMOS should be computed for both directions (i.e., up-link and down-link). The motivation behind this is that at this point, decisions could be taken about the current call or about possible calls handed over from the macrocell. Such a metric
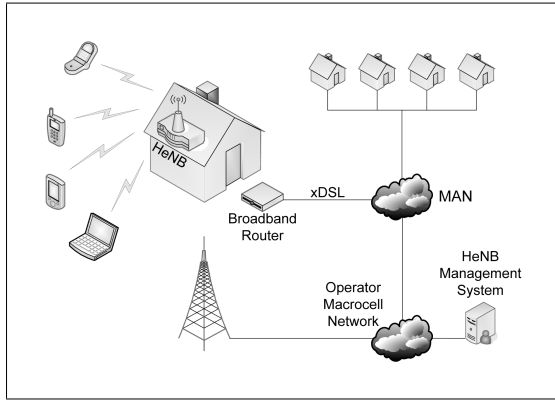
Figure 1. Typical deployment for Femtocells in LTE

has further uses for fault detection, where a problem in the network could be isolated to either the up-link or the down-link between the HeNB and the HeNBGW. Such a metric will also be useful for *Customer Experience Manager* (CEM) and *Service Quality Manager* (SQM), possibly in connection to Service Level Agreement (SLA) compliance and negotiation.

Network delay could be obtained through estimations based on related network metrics such as packet interarrival deviation or through direct calculation, i.e. difference between local time and the time obtained from packets' timestamps, in this later case the time of the two nodes must be synchronised. A femtocell must meet rigorous frequency and timing synchronisation requirements to comply with the air interface standard. The HeNBGws must also have their time synchronised, so it is assumed here that time synchronisation is implemented, and the packets' delays can be easily calculated using *Real-time Transport Protocol* (RTP) [5] timestamps correlated with UTC timestamps contained in corresponding *Real-time Transport Control Protocol* (RTCP) packets [5].

In this paper we explore the impact of video streaming on the quality of femtocell VoIP calls. In particular we examine the congestion at the *Digital Subscriber Line Access Multiplexer* (DSLAM) at the local exchange where femtocell calls must contend with video traffic for access to and from the Metropolitan Area Network (MAN). In the scenario investigated, it is assumed that the cellular service provider has no control over the DSLAM traffic and will not be able to manage traffic to increase the VoIP quality so all traffic is of equal priority at the DSLAM queues.

In this paper we present one implementation we have used to calculate the MOS score as an instantaneous measure and an algorithm to assess the quality as an overall measure. One feature of these two is the presence of the recency effect, introduced by the human perception about an audio stream of information.

The paper begins with a related work survey. A section describing the implementation and use of MOS follows, the results are then presented, conclusions are drawn and future work is outlined.

II. Related Work

The main idea behind femtocells is to increase users' link quality and carriers' system capacity. One major improvement in increasing cell capacity was obtained by reducing the cell size and transmit distance [1]. A deployed femtocell will cover an area with high quality access for mobile phones and the connection, or backhaul, from a femtocell into the operator's network can be wired or wireless. In order to maintain low infrastructure costs, residential femtocells could use the user's broadband connection as backhaul, which would be typically an xDSL connection.

Many researchers have focused their attention towards the possible problems raised by the radio aspects of a femtocell [6], [7] or the use of radio backhaul [8], [9]. To date the use of wired backhaul has been taken as a simple IP connection rather than analyzed in detail.

In terms of VoIP quality monitoring, there is a need to take into consideration the recency effect imposed by humans' perception and audio memory [10]. It is estimated that the human audio memory can recall from the last 30 seconds [11], thus at any point in a conversation a subject will make his/her *Quality of Experience* (QoE) estimation based only on the recent 30 seconds.

Another human characteristic metric was presented by [11] who showed by subjective experimental means that a typical human senses a transition from good to bad with a 5 seconds delay and 15 seconds from bad to good. This is an important hint in choosing the averaging window for the calculated R values.

The initial R calculation proposed by ITU-T comprises metrics that are specific to old analog telephony, such as crosstalk over wires [12]. To ease the use of this model, a simplified version of the model for packet switched networks is proposed in [13].

MOS calculation involves two networking metrics, packet delay and packet loss rate, each one defining a term is the MOS formula. There is no implementation defined as standard, and various implementations of the E-model have been proposed in [14], [13]. E-model implementation is not disclosed to the large public by proprietary implementations, hence using MOS as reference for qualifying VoIP calls without specifying which implementation is used, can lead to inconsistency.

The work presented by [13] is of particular interest, where the *Transmission Factor* (R)'s formula is adapted for different codecs. The method presented there uses curve fitting to map subjective results into a specific analytical formula. Both delay and equipment impairment factors' formula are obtained by plotting data from tables provided by ITU-T's subjective testing with specified voice codecs [12]. The best approximation function was obtained by empirically adjusting weighting parameters. The formula is generalized to comprise parameters for other codecs with various packet size and error mask distribution.

Packet delay is one of the most important metrics in a packet switched network, but measuring it can be prone to a certain

level of uncertainty. In particular, for VoIP quality estimation it is crucial to keep this uncertainty as low as possible. Two error factors affect this measure, i.e. local host time offset and local clock frequency instability, which can be adjusted by different time synchronisation algorithms. For large distances between users, NTP [15] could be employed with the level of clock offset uncertainty lying in the range of tens of miliseconds. Another algorithm, *Precision Time Protocol* (PTP) [16], relies on timing events at the Ethernet interface. It is designed for small distances (e.g. *Local Area Network*s (LANs)) and its level of uncertainty drops in range of microseconds.

Mapping these to VoIP, *Network Time Protocol* (NTP) fails because of the large interval of uncertainty while PTP fails in distance allowed between users and because is technology dependent.

LTE standards demand clock synchronisation to be within few Hz so some hardware or *Global Positioning System* (GPS) solution is required. Here we assume that such a solution has been implemented in our nodes and our simulation environment (i.e. *Qualnet*) distributes the same clock time to all simulation nodes.

VoIP requirement of precise time and frequency keeping is fulfilled by the more strict requirement for femtocells. For the case of femtocells, NTP and PTP are regarded as possible candidates only in conjunction with another clock synchronisation technique [17]. NTP/PTP could be employed if the wired backhaul of the femtocell is supporting time synchronisation. PTP supports the concept of boundary clocks, thus in key points of the backhaul, like Local Exchange or DSLAM, these could be implemented or are already supported by some service providers.

A similar concept of VoIP monitoring is presented in previous works [18], [19], [20] but with at least two major differences from our proposal. First, our approach uses time synchronisation and not just an estimation for calculating the delay; and second, all nodes involved in our approach are owned by the mobile operator, while the other nodes in the network (such as the DSLAM) are assumed not to implement our monitor. This makes our architecture different from a regular IP network.

III. METRIC CALCULATION

ITU-T has made various recommendations on assessing the quality of a VoIP conversation, for example, mouth-to-ear delay is one of the most problematic requirements, since it is recommended to be lower than 150ms [21]. An agreed estimation of core network delay is 100ms, thus giving an overall 50ms for other delay causing impairments such as *Media Access Control* (MAC) buffering in an *Internet Protocol* (IP) network.

In order to calculate the MOS, we must first calculate the Transmission Rating Factor (R):

$$R = R_0 - I_s - I_d - I_e + A \qquad (1)$$

In equation (1), R_0 represents the signal-to-noise ratio; I_s is a combination of all impairments which occur more or less simultaneously; I_d represents the impairments caused by delay; I_e represents the impairments caused by low bitrate codecs and A represents the advantage factor.

In order to obtain the MOS, a conversion formula is provided in [12]:

$$MOS = \begin{cases} 1 & \text{if } R < 0, \\ 1 + 0.035 \cdot R + 7 \cdot 10^{-6} \\ \cdot R(R-60)(100-R) & \text{if } 0 \le R \le 100, \\ 4.5 & \text{if } R > 100. \end{cases} \qquad (2)$$

According to [13], equation (1) can be reduced to

$$R = 93.4 - I_d - I_e \qquad (3)$$

The default values for I_d provided in G.107 were determined through subjective testing, where the voice packets were delayed in a range from 0 to 400 ms, resulting in an I_d factor lying in a range from 0 to 31. The resulting plot of these values is used to determine the best curve fit. For I_d, the resulting formula is:

$$I_d = 0.0024 \cdot d + 0.11(d - 177.3) \cdot H(d - 177.3) \qquad (4)$$

where $H(x)$ is the Heavyside function:

$$H(x) = \begin{cases} 0 & if \ x < 0, \ else \\ 1 & for \ x \ge 0 \end{cases} \qquad (5)$$

In (4), d represents the total end-to-end delay a packet may suffer and is

$$d = d_{codec} + d_{de-jitter \ buffer} + d_{network}$$

and 177.3 is a constant and represents the delay at which the plot shows a breaking point. It is a common practice in curve fitting to use the summation of multiple simple functions which are enabled at particular points by the Heavyside function.

Again, the E-model does not provide any analytical formula for I_e, thus it has to be determined as per voice codec, since I_e is relying on codec's capacity to recover voice fragments missing due to packet loss. After analyzing the behavior of G.711 and G.729a codecs, a generic formula is provided in [13]. The formula for I_e is:

$$I_e \sim \gamma_1 + \gamma_2 \cdot ln(1 + \gamma_3 \cdot e) \qquad (6)$$

In equation (6), e represents the total loss rate with a value between 0 and 1 and $\gamma_1, \gamma_2, \gamma_3$ are fitting parameters. Packet loss could be bursty or random and this leads to I_e's two different behaviours. Voice codecs may have a *Packet Loss Concealment* (PLC) implemented and it could be employed or not. Combining packet loss behavior (i.e. bursty or random) with the status of the PLC (i.e. enable or not), then four combinations are possible and all this must be taken into consideration when determining the specific formula for I_e for one codec.
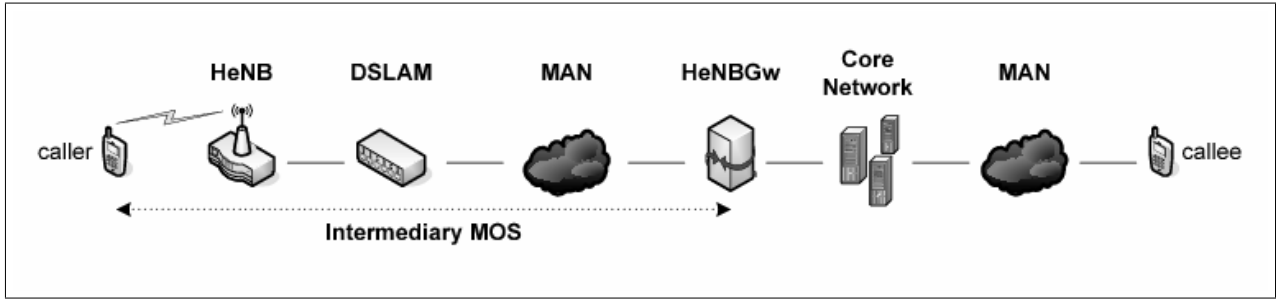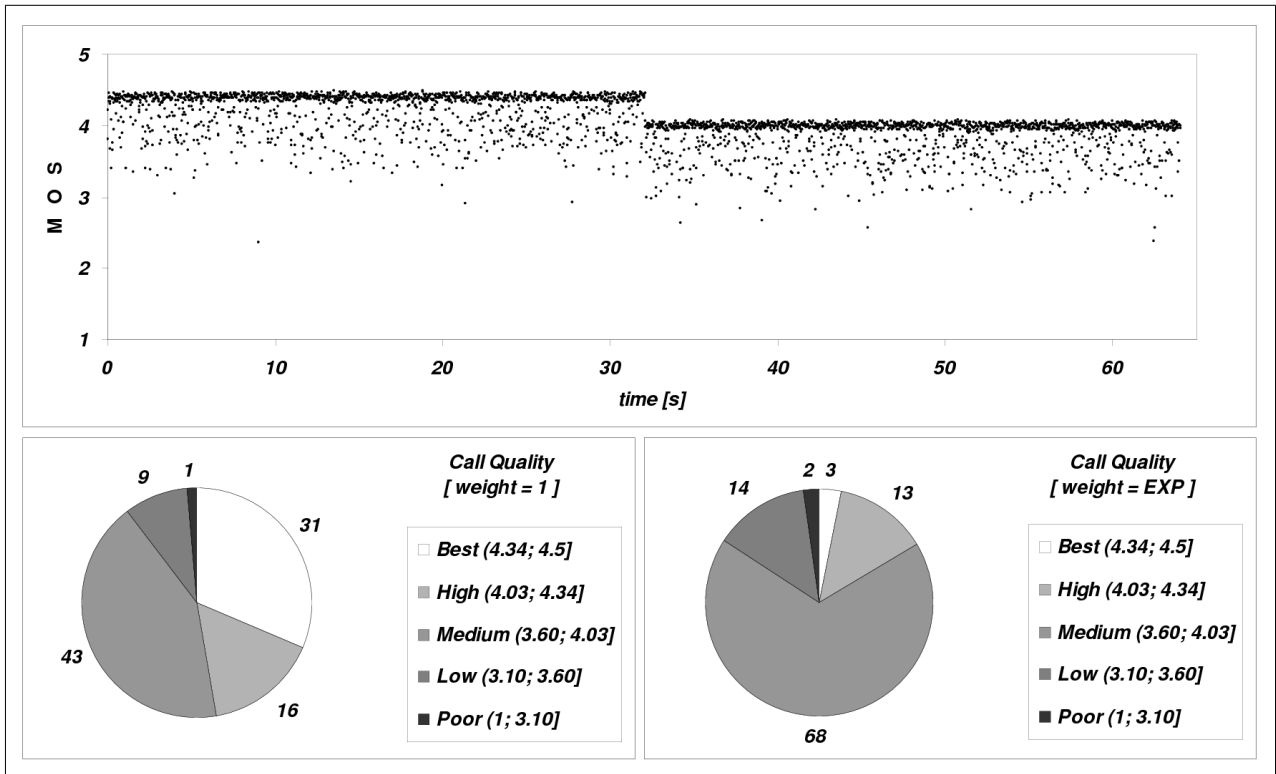
Figure 2.   Intermediary MOS concept

Figure 3.   Weighted MOS score mapped into corresponding Call Quality pie charts

IV. MOS BASED VoIP QUALITY ASSESSMENT

Based on the description presented in the previous subsection about how MOS is computed, in this part we discuss processing those instantaneous values into a set of measures that can be used to determine actions regarding a call (e.g., handover a call from a HeNB to the macrocell base station).

One approach is the instantaneous approach, which is employed during the call to asses the quality and another approach is the post-processing of the MOS values obtained.

For the instantaneous approach, an averaging sliding window is employed. We have chosen the length of the window to be 1 second. This is based on the consideration that human perception has inertia and on the experimental results presented in [11]. The window slides with each new arriving voice packet and the average for that window is stored in a table. Each value is compared against the *Quality rating* thresholds presented in the following table and a counter for the associated quality level is updated accordingly:

Quality rating	MOS	R
Best	4.34 ... 4.50	90 ... 100
High	4.03 ... 4.34	80 ... 90
Medium	3.60 ... 4.03	70 ... 80
Low	3.10 ... 3.60	60 ... 70
Poor	1.00 ... 3.10	$R < 50$

At any time during a conversation the quality level counter values can be used to show the percentage time that the users' satisfaction level reached a specific level. Based on the

$$w(t) = H(30 - t_f) \cdot \left(e^{4.605 \cdot \left(1 - \frac{t_f - t}{30} \right)} \right) + H(t_f - 30) \cdot \left(e^{4.605 \cdot \left(1 - H(t_f - 30 - t) - H(t - (t_f - 30)) \cdot \frac{t_f - t}{30} \right)} \right) \qquad (7)$$

agreement between the user and his/her Service Provider, a set of thresholds could be agreed upon to be monitored during a call. For example, it could be established that the total time of *Best* conversation quality should be above 95%. A pie chart of *Quality rating* percentages is obtained at conversation end and could be used to feed back the quality of the call to the user.

We mentioned the need to introduce the recency factor and the following will present different approaches in mapping human recency in a VoIP conversation. First we present in Figure 3 a typical MOS score that could be obtained through a regular conversation. This example is particularly chosen to cover all 5 *Quality rating* groups.

Based on [11] a human can recall the last 30 seconds form the audio memory. This means that in a VoIP conversation the MOS values corresponding to the last 30 seconds should have a bigger weight in the histogram than those before.

For simplicity, we propose a 100 scale for weighting factors. Equation (7) represents the formula we propose for implementing the recency factor into the Call Quality assessment.

In (7) t_f is the time corresponding to the last arrived packet and t is sweeping the interval between the conversation start time and t_f. It could be observed that again the Heavyside function in employed to enable parts of the equation based on the current point in time where the weight is computed (i.e. t).

It can be observed that the MOS score degrades at conversation midpoint. This degradation is captured by the exponential weighting factor which reports a lower score that the constant one. Hence, the usage of this technique closes the analytical method, based on pure network metrics, to the QoE reported by an end user.

V. SCENARIOS AND SIMULATIONS

A. Typical HeNB call scenarios

The different simulation scenarios are differentiated by the nature of the session's endpoints. In a session involving only two VoIP UE, the session could be seen as fully monitored. When the other end is using a circuit switched technology interacting through a gateway, it is half-monitored. For consistency, we name the UE in the HeNB the *caller* and the other side the *callee*.

In Figure 4 it can be observed that six different scenarios can be used to model a call between two pairs. The first is a HeNB-to-HeNB call where both HeNBs belong to the same Service Provider. It can be observed that the *callee's* backhaul is similar to *caller's*, and this scenario captures even the case when both HeNBs use the same DSLAM as backhaul. The call is fully monitored (i.e., the network between both UEs is all-IP network). In this case the HeNBGw will provide four MOS values: two characterizing the link between itself and the *caller* and two for the link between itself and the *callee*.

The second scenario involves a *callee* in the same Provider's network, attached to a macrocell tower. This scenario is similar to the previous one and a fully monitored call can provide information about four links in the network.

The third scenario depicts a situation where the *callee* belongs to the Service Provider's fixed telephone network which employs a circuit switched technology, which is interfaced to the packet switched network through a media gateway. We call this case half-monitored.

The other three scenarios are similar with the first three presented above with the difference that the *callees* in these scenarios belong to a different Service Provider's network.

For simplicity, in this paper we are mainly focused on the first scenario presented. This is a *All IP Network* (AIPN) case, in which each subnetwork of the communication chain transfers information as IP packets.

B. Simulation results

We present our results obtained to sustain our hypothesis that the DSLAM is the principal candidate as the network bottleneck. We simulated the scenario depicted in Figure 5. We established a call and then used network traffic generators to send traffic through the same non-prioritizing DSLAM serving all HeNBs. This extra traffic is increased until the MOS in any direction decreases under a certain pre-established limit and store the parameters used to generate the extra traffic.

We have used *Qualnet 4.5* [22] open source network simulator to obtain the results. *Qualnet* features the VoIP application in its available protocol stack. A change needed to be operated on the source code: by default VoIP is implemented as a half-duplex connection and that was fixed to implement a full-duplex VoIP session. The voice codec we used was the standard *G.711* to remove the complexity of rate adaptation that may have obscured the efficiency of the approach. The results are expected to be broadly similar for an adaptive codec such as AMR with users distributed throughout the coverage area. *Qualnet* implements Equation (2), (3), (4), and (6) for MOS calculation. Equation (6) with specific parameters for the G.711 codec turns into:

$$I_e = 30 \cdot ln(1 + 15 \cdot e) \cdot H(0.04 - e) + 19 \cdot ln(1 + 70 \cdot e) \cdot H(e - 0.04)$$

A batch simulation was run, in which the number of VoIP sessions increased from 1 to 60 and the background traffic was *Constant Bitrate* (CBR). The background traffic load is increased in steps until the MOS's *best* quality category percentage drops under 90 and the corresponding load is recorded. The results presented in Figure 6 clearly show that the amount of traffic that is sent through the same DSLAM decreases as the number of VoIP calls increase.

One solution is that the DSLAM should prioritize and even allocate more bandwidth to the voice traffic coming from and towards a HeNB. However, it cannot be assumed
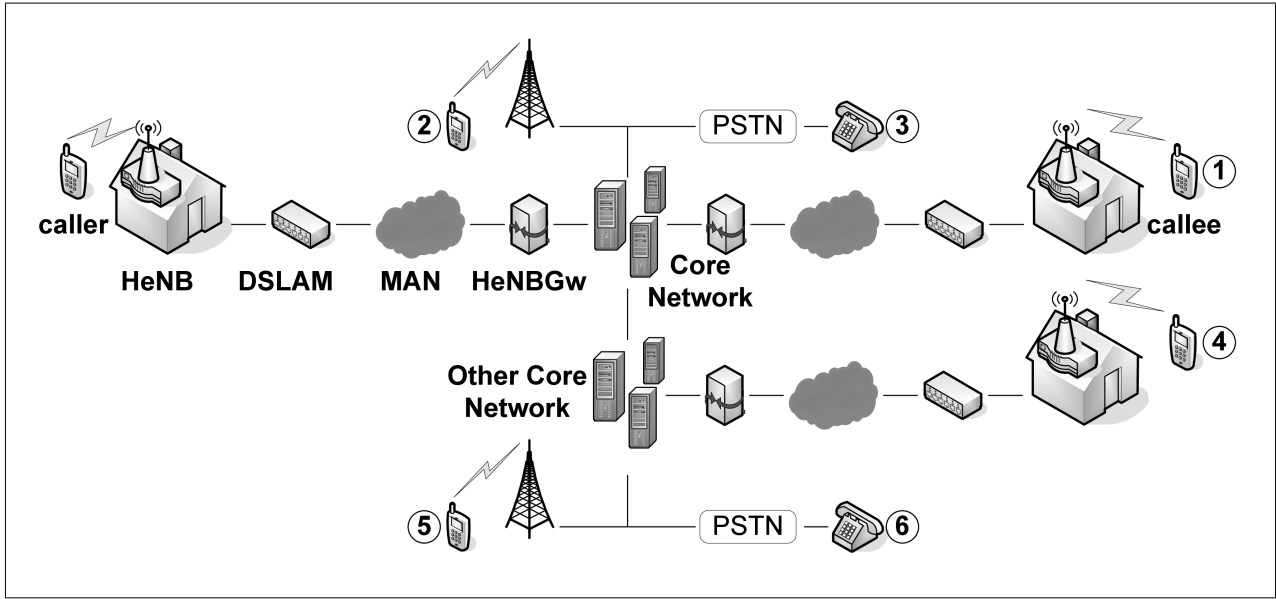
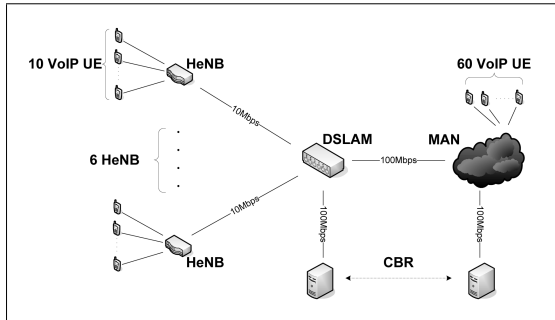Figure 4. Typical deployment for Femtocells in LTE

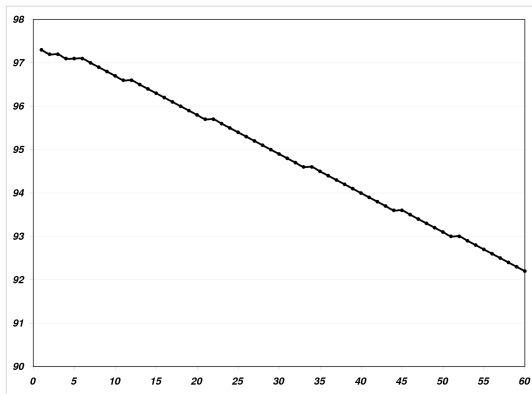Figure 5. Scenario used for pointing the DSLAM as a bottleneck

Figure 6. CBR traffic versus no. of VoIP conversations

that the Internet broadband connection over xDSL (*Digital Subscriber Line* (DSL)) is offered by the same body offering the femtocell, so we assume that users will be served by a DSLAM which is not prioritizing VoIP traffic from or towards the femtocell. In other words VoIP will compete with other high volume traffic such as video for medium access.

C. Intermediary MOS

Next, we show through simulations how IMOS works. Simulations have been run with a scenario matching the first case presented in Figure 4. The actual scenario, MOS pie charts, and loading traffic used are depicted in Figure 7. The HeNBGw of the *caller* has the capability of extracting the needed information from each voice packet to obtain the MOS at that place in the network.

Because a VoIP session is implemented as a full-duplex connection, there are two voice streams for one voice session. The *direction* of each stream plays an important role at the HeNBGw as much as at both ends. For example, the MOS score obtained by the *caller* is based on packets sent from the corresponding *callee*, so we say that this MOS characterizes the downlink and only other downlink traffic could degrade the QoS in this case.

It could be seen that the MOS obtained at the HeNBGw for packets arriving from the *callee* falls 100% in the *Best* quality category and the same packets arriving at the *caller* determine a lower quality. This leads to the conclusion that the link between the HeNB and its HeNBGw has caused the quality degradation. The reason for this degradation is a heavy load the DSLAM has to support from a user using the downlink to access some multimedia content.

Since we assume that no priority mechanism is implemented at the DSLAM, contention is causing impairments for the voice quality. The traffic load we used to cause quality
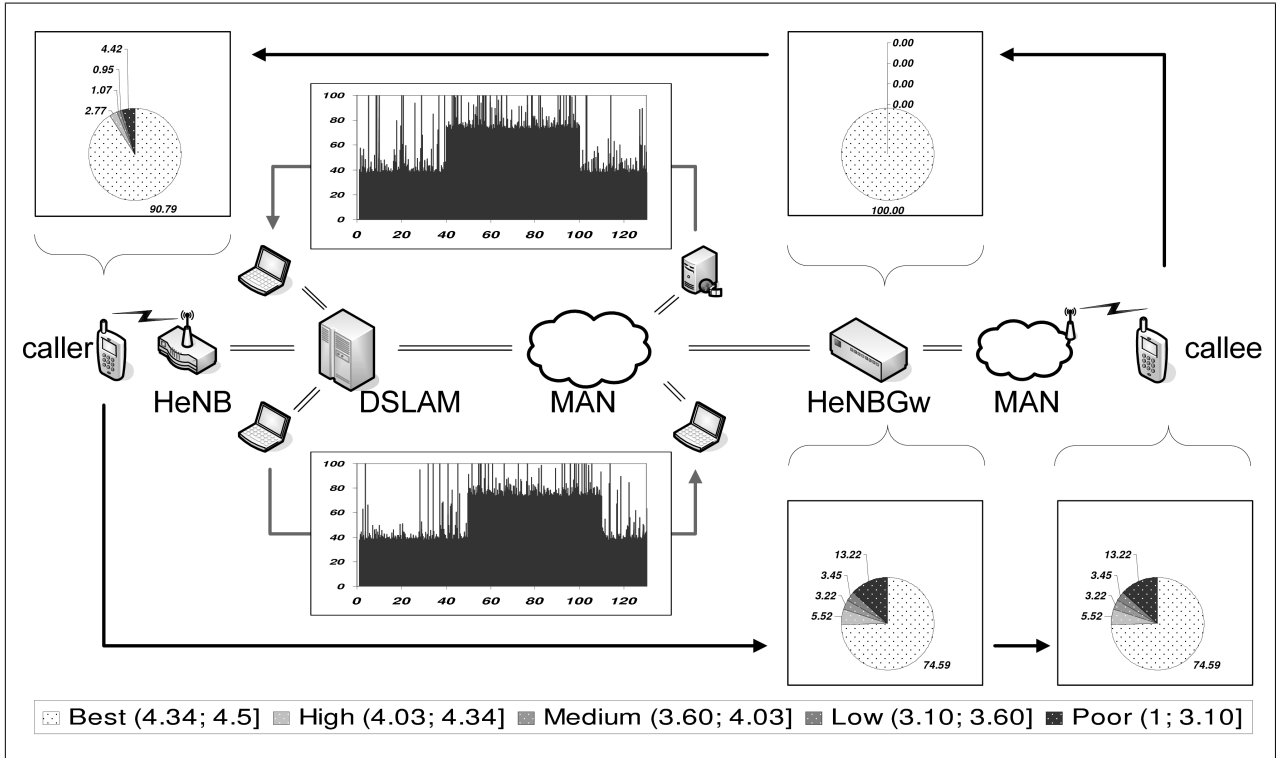
Figure 7.   Simulated scenario with results

degradation, starts at the 40th second and ends at the 100th second of the simulation.

The same situation happens in the opposite direction, where voice packets content with other multimedia carrying packets. The traffic load has almost the same pattern with the difference that it is 10 seconds delayed towards the end of the conversation. Considering the traffic peak randomness and the recency factor in our MOS calculation, this delaying explains the difference in the final MOS figures at the *caller* respectively at the *callee*. The reason the *callee* shows the same overall quality as the HeNBGw is that the network link residing between those, causes no important delay or loss.

Each one from the two voice streams passing the HeNBGw will result in a corresponding MOS score. Because the aggregation of MOS scores happens at the HeNBGw we find this a being the most appropriate place in the network to place the SQM. Within SQM, a CEM could be implemented.

VI. CONCLUSIONS AND FUTURE WORK

One important point highlighted in this paper emphasizes that monitoring *Mean Opinion Score* (MOS) can be achieved in femtocellular networks with high accuracy due to clock synchronisation of all key nodes in such scenario (i.e. *Home enhanced Node B* (HeNB) and *Home enhanced Node B Gateway* (HeNBGw)).

A new concept derived from MOS is presented here. This is the *Intermediary Mean Opinion Score* (IMOS) and implies

MOS calculation at intermediary nodes (e.g. HeNBGw). This concept isolates problems in the network by partitioning the link between the call ends in small quality monitored segments.

Humans' audio recency factor mapped into the MOS calculation is implemented by a weighting algorithm and compared with the non-weighting one. The results seam to be more accurate and future subjective tests are needed.

The IMOS measure can be used to drive *Customer Experience Manager* (CEM) applications. It may be used to make admission control decisions at the HeNB. This paper emphasizes the need to use the IMOS concept in taking handover decisions.

For future work, we aim to implement our IMOS concept on a real test-bed and conduct a series of *Perceptual Evaluation of Speech Quality* (PESQ) tests for validation purposes. For this, the key aspect of clock synchronisation inside important nodes needs to be considered, since it plays a crucial role in real-time multimedia networks.

REFERENCES

[1] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell Networks: a Survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, september 2008.

[2] G. Muntean, P. Perry, and L. Murphy, "Objective and subjective evaluation of QOAS video streaming over broadband networks," *IEEE Transactions on Network and Service Management*, vol. 2, no. 1, pp. 19–28, 2005.

[3] G. Muntean, L. Murphy, and P. Perry, "Quality-oriented adaptation scheme for video-on-demand," *Electronics Letters - IEE*, vol. 39, no. 23, pp. 1689–1689, 2003.

[4] B. Jennings, S. van der Meer, S. Balasubramaniam, D. Botvich, M. Foghlu, W. Donnelly, and J. Strassner, "Towards autonomic management of communications networks," *IEEE Communications Magazine*, vol. 45, no. 10, pp. 112 –121, oct. 2007.

[5] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," 2003.

[6] H. Claussen, L. Ho, and L. Samuel, "Self-optimization of Coverage for Femtocell Deployments," *Wireless Telecommunications Symposium*, pp. 278–285, april 2008.

[7] D. López-Pérez, A. Valcarce, G. de la Roche, and J. Zhang, "OFDMA Femtocells: A Roadmap on Interference Avoidance," *IEEE Communications Magazine*, vol. 47, no. 9, pp. 41–48, september 2009.

[8] R. Pabst, B. Walke, D. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. Fettweis, "Relay-based Deployment Concepts for Wireless and Mobile Broadband Radio," *IEEE Communications Magazine*, vol. 42, no. 9, pp. 80–89, sept. 2004.

[9] C. Hoymann, A. Rácz, N. Johansson, and J. Lundsjo, "A Self-backhauling Solution for LTE-Advanced," *WWRF 21-WG4 Stockholm, Sweden*, 2008.

[10] A. Clark, "Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality," *Proc. IP Telephony Workshop*, 2001.

[11] ITU Study Group 12, "Study of the relationship between instantaneous and overall subjective speech quality for time-varying quality speech sequences: influence of a recency effect," *Contribution D.139*, 2000.

[12] ITU-T G.107, "The E-Model, A Computational Model For Use In Transmission Planning," 2005.

[13] R. G. Cole and J. H. Rosenbluth, "Voice over IP performance monitoring," *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 2, p. 9, 2001.

[14] L. Carvalho, E. Mota, R. Aguiar, A. Lima, J. de Souza, and A. Barreto, "An E-Model Implementation for Speech Quality Evaluation in VoIP Systems," *10th IEEE Symposium on Computers and Communications*, pp. 933–938, 2005.

[15] D. Mills, J. Martin, J. Burbank, and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification," *Request for Comments*, no. 5905, Jun. 2010.

[16] J. Eidson, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," *IEEE Std 1588-2002*, pp. i –144, 2002.

[17] Femto Forum, "Femtocell Synchronization and Location," no. June, 2010.

[18] A. Conway, "A passive method for monitoring voice-over-IP call quality with ITU-T objective speech quality measurement methods," vol. 4, pp. 2583–2586, 2002.

[19] R. Birke, M. Mellia, M. Petracca, and D. Rossi, "Understanding VoIP from backbone measurements," pp. 2027–2035, 2007.

[20] S. Anderson, S. Niccolini, and D. Hogrefe, "SIPFIX: A scheme for distributed SIP monitoring," pp. 382–389, 2009.

[21] ITU-T G.114, "One-way Transmission Time," 2003.

[22] Scalable Network Technologies, "Qualnet 4.5," Website, http://www.scalable-networks.com/.