End-to-End Delay Performance for VoIP on LTE System in Access Network

Ng Liang Shen[1,*], Noraniah Abdul Aziz[2], and Tutut Herawan[3]

[1] Faculty of Computer System and Software Engineering
Universiti Malaysia Pahang Lebuhraya Tun Razak, 26300 Gambang, Malaysia
[2] TATI University of College 24000 Kemaman Terengganu, Malaysia
[3] Department of Mathematics Education Universitas Ahmad Dahlan
Yogyakarta 55166, Indonesia
ngliangshen@gmail.com, noraniah@tatiuc.edu.my, tutut@uad.ac.id

Abstract. The desire Quality of Service (QoS) of Voice over Internet Protocol (VOIP) is of growing importance for research and study Long Term Evolution (LTE) is the last step towards the 4th generation of cellular networks. This revolution is necessitated by the unceasing increase in demand for high speed connection on LTE networks particularly for under variable mobility speed for VoIP in the LTE. This paper mainly focuses on performance of VOIP and the impact of resource limitations in the performance of Access Networks particularly important in regions where Internet resources are limited and the cost of improving these resources is prohibitive. By determine rate communication quality, is determined by end to end delay on the communication path, delay variation, packet loss. These performance indicators can be measured and the contribution in the access network can be estimated using simulation tool OPNET Modeler in varying mobility speed of the node. The overall performance of VOIP thus greatly improved significantly by deploying OPNET Modeler.

Keywords: Quality of Service (QoS), Voice over Internet Protocol (VOIP), Internet, performance, Future, Network Access, analysis, delay.

1 Introduction

The startling growth of Internet technology, coupled with the relatively low deployment cost of IP networks, has pushed for an integrated “IP-based core” - a single network for data, video and voice access. However, the diverse service-requirements and novel traffic characteristics of the emerging Internet applications have posed many technical challenges that the Internet community must address in the near future, as the emerging multimedia applications begin to constitute an ever-increasing fraction of Internet traffic. High quality interactive voice and video applications can tolerate little delay variation and packet loss. Quality of Service (QoS) is a defined level of performance in a data communications system. As an

* Corresponding author.

K. Mustofa et al. (Eds.): ICT-EurAsia 2013, LNCS 7804, pp. 162–172, 2013.

example, to ensure that real time voice and video are delivered without irritating blips, a guaranteed bandwidth is required. The plain old telephone system (POTS) has delivered the highest quality of service for years, because there is a dedicated channel between parties. However, when data is broken into packets that travel through the same routers in the LAN or WAN with all other data, QoS mechanisms are the only way to guarantee quality by giving real time data priority over non-real time data. A large number of factors are involved in making a high-quality VoIP call. These factors include the speech codec, packetization delay, packet loss, delay (coding, transmission, propagation and queuing), delay variation, and the network architecture to provide QoS. Other factors involved in making a successful VoIP call include the call setup signalling protocol, call admission control, security concerns, and the ability to traverse Network Access Translation (NAT). Although VoIP involves the transmission of digitized voice in packets, the telephone itself may be analog or digital. The voice may be digitized and encoded either before or concurrently with packetization. In making a high quality VOIP call and the engineering trade-offs that must be made between delay and the efficient use of bandwidth [1,3]. VoIP becoming popular can be mainly attributed to the cost advantages to consumers over traditional telephone networks. The traditional business model for telephone services has been that most people pay a flat monthly fee for local telephone call service and a per-minute charge for long-distance calls.Introduction of Long Term Evolution (LTE), the 4th Generation (4G) network technology release 8 specifications are being finalized in 3GPP have developed and planning to globalize extensively compared to 3rdGeneration (3G) and 2nd Generation (2G) networks [1]. LTE determines goals peak data rate for Downlink (DL) 100 Mbps and Uplink (UL) data rate for 50Mbps, increased cell edge user throughput, improved spectral efficiency and scalable bandwidth 1.4 MHz to 20 MHz [2]. VoIP capacity of LTE has to show better performance as Circuit Switch voice of UMTS. LTE should be at least as good as the High Speed Packet Access (HSPA) evolution track also in voice traffic. It discusses the challenging issues that need to be faced by computer networks to transmit the VoIP applications. It gives the description idea about the VoIP over LTE and their functionality and design parameters of the LTE networks. After determining the problems it is necessary to identify the research questions that lead the research process to be in the scope [2].

a) How much the maximum throughput is support in the different bandwidth (e.g. 1.4 MHz, 3 MHz, 5 MHz, 10 MHz, 15 MHz and 20 MHz)?
b) What is the impact on the VoIP quality in terms of E2E delay when the network is congested with VoIP only?
c) To what extent do the performances of packet loss for interactive voice vary?

The 3GPP LTE is a new standard with comprehensive performance targets, therefore it is necessary to evaluate the performance and stability of this new system at an early stage to promote its smooth and cost-efficient introduction and deployment. The motivation behind the design models presented in this report is to discuss issues related to traffic behavior for VoIP alone as well as along with other traffic in the LTE network. E2E delay for VoIP is a matter of fact for performing real-time application efficiently over the Internet. Today, emergence of the real-time application demands more resources. The main motivation of our paper work is to ensure fast and reliable voice communication for huge number of users in wireless network. LTE evolved packet system (EPS) is the bearer of the QoS level of granularity. This system also establishes the packet flow between the user terminal (UE or MS) and the packet data network gateway (PDN-GW). The traffic running between a particular client application and the service can be wrecked into split service data flows (SDFs). Mapping the same bearer, SDFs receive common QoS activities (e.g., scheduling policy, queue management policy, rate shaping policy, and radio link control (RLC) configuration). A scalar value referred to as a QoS class identifier (QCI) with the help of bearer, specifies the class to which the bearer belongs. Set of packet forwarding treatments referred by QCI (e.g., weights scheduling, ad-mission thresholds, configuration of link layer protocol and queue management thresholds) preconfigured through the operator on behalf of each network element. The class-based technique applies in the LTE system to improve the scalability of the QoS framework. In the LTE framework, bearer management and control follows the network-initiated QoS control paradigm that initiated network establishment, modification, and deletion of the bearers. Two types of bearers in LTE [2,3,4]:

a. **Guaranteed bit rate (GBR):** Dedicated network resources correlated to a GBR value connected with the bearer and permanently allocated when a bearer becomes established or modified.
b. **Non-guaranteed bit rate (non-GBR):** In the LTE system, non-GBR bearer is as-signed as the default bearer, similar to the preliminary SF in WiMAX, used to establish the IP connectivity. A non-GBR bearer has enough knowledge about congestion-related packet loss. In the framework, additional bearer is assigned as a dedicated bearer which is GBR or non-GBR.

Future Research need to cover further analyses could be carried out using the measurement data captured. A synthetic workload model for VoIP could be further developed, based on the workload characteristics identified in the measurement study.

2 How Is the Performance for VoIP on LTE System in Access Network

In LTE, the core network operations are completely based on packet switching domain, example, all the network interfaces are dependent on IP protocols, and hence it is known as Evolved Packet Core (EPC). The essence of EPC is to keep the number of operating nodes and interfaces as minimum as possible. The EPC divides the network components into control-plane objects such as data/barer-plane entity (i.e. a Serving Gateway) and the Mobility Management Entity (MME). The MME is considered as a signaling entity and used to represent the control plane function of the EPC. Such control functions include, among others, location function, the subscribers' equipment paging, and the bearer establishment and the connections establishment, roaming management [4]. The fundamental architecture of LTE system is presented in Figure 1. All the network interfaces are based on internet protocols (IP). The LTE system comprised of the core network and radio access network which represent the IP connectivity layer of LTE system.

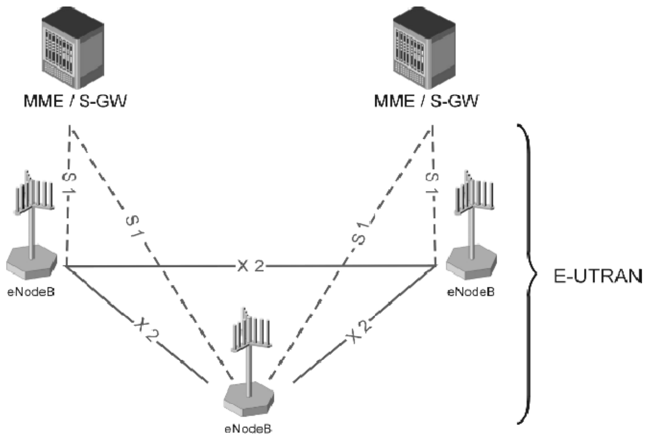

Fig. 1. Networks Architecture of LTE [16]

The Serving Gateway (S-GW) functions as switching as well as routing node to route and forward the data packets to and from the Base or Evolved-Universal Terrestrial Radio Access Network NodeB (eNB) [5]. Between the EPC and the external packet data network, a PDN-GW is often used as an inter-face point or an edge router. It is also possible that a UE has synchronized connectivity with more than one PDN GW. The responsibilities of the PDN-GW include establishment, maintenance, and deletion of GTP tunnels to S-GW or SGSN in the case of inter-RAT mobility scenarios. The PDN-GW routes the user plane packets by allocating the user's dynamic IP addresses. Apart from that, it provides functions for lawful interception, policy/QoS control, and charging [6].

3 LTE QoS Framework and Voice over IP (VoIP)

LTE evolved packet system (EPS) is the bearer of the QoS level of granularity. This system also establishes the packet flow between the user terminal (UE or MS) and the packet data network gateway (PDN-GW). The traffic running between a particular client application and the service can be wrecked into split service data flows (SDFs). Mapping the same bearer, SDFs receive common QoS activities (e.g., scheduling policy, queue management policy, rate shaping policy, and radio link control (RLC) configuration) [19, 20]. A scalar value referred to as a QoS class identifier (QCI) with the help of bearer, specifies the class to which the bearer belongs. Set of packet forwarding treatments referred by QCI (example, weights scheduling, admission thresholds, configuration of link layer protocol and queue management thresholds) preconfigured through the operator on behalf of each network element [7]. The class-based technique applies in the LTE system to improve the scalability of the QoS framework. [8].

3.1 Two Types of Bearers in LTE

There are two types of bearer in LTE which are GBR and non-GBR. In the framework, additional bearer is assigned as a dedicated bearer which is GBR or non-GBR. Real Transport Protocol (RTP) is the basic protocol in VoIP engineering, which is, not only for transporting media streams but also to initialize the media session in concord with SIP. It is also used for media stream supervision and intended to provide out-of-band control information for the RTP flow. In response to the media quality that supplies to the other members in the media session via separate UDP port, there are many additional functionalities of RTP. Audio and video synchronization and quality improvements through low compression instead of high compression are a few of them. VoIP transmissions are deployed through traditional routing. A typical VoIP structural design is shown on the Figure 2, though many "possible" modifications of this architecture are implemented in existing systems. In VoIP engineering, original voice signal is sampled and is encoded to a constant bit rate digital stream at the end of the sending process.In place of circuit-switched voice transmission and traditional dedicated lines, these packets flow over a general-purpose packet-switched digital to analog signal in the receiving end for it to be easily detected. The VoIP infrastructure can be visualized as three layers: end user equipment, network components, and a gateway to the traditional telephone network as stated in Figure 3.

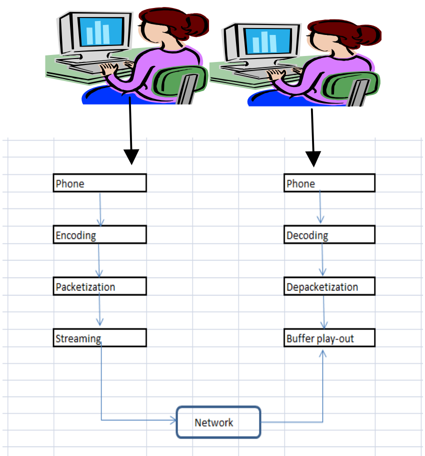

Fig. 2. VoIP Architecture of LTE

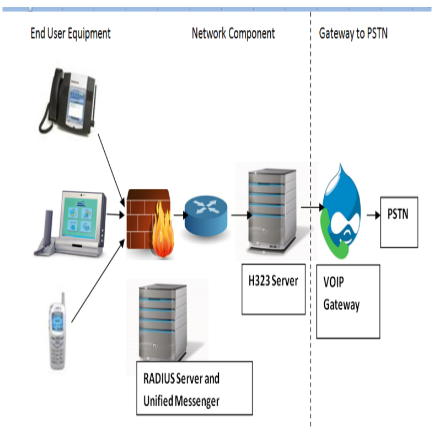

Fig. 3. VOIP. Network Archtecture

3.2 Characteristics VoIP Traffic Delay

AMR codec provides the VoIP traffic along with the Voice Activity Detector, Relieve Noise Generation and Discontinuous Transmission. Depending on the speed activity of the traffic, AMR provides a constant rate of small packets transmission. During the active period, one VoIP packet took at 20 ms intervals and 160 ms interval for one Silence Description (SID) packet during silent period. To improve the spectral efficiency of the VoIP traffic, UDP, IP and RTP headers in LTE are also compressed with Robust Header Compression (ROHC). According to [9], for voice signal, 250 ms is the maximum tolerable mouth-to-ear delay and around 100 ms delay for the Core Network and also less than 150 ms acceptable delay for Medium Access Control (MAC) buffering and Radio Link Control (RLC). Both end users are LTE users and assume less than 80 ms acceptable delay for buffering and scheduling. For 3 GPP performance evaluations 50 ms delay has been bound for variability in network end-to-end delays. End-to-end delay means the time required for a packet to be traversed from source to destination in the network and is measured in seconds. Generally, in VoIP network there are three types of delays occurring during the packet transverse. They are: sender delays when packets are transverse from source node, network delay and receiver delay. In one direction from sender to receiver for VoIP stream flow, end-to-end delay can be calculated by the equation. Let say in seconds-that a network needs to send a message with n bytes as $T(n) = \alpha + \beta n$ In mathematical terms, this is a line equation, where α is the *constant*, and β is the line *slope* or *gradient* [9,10].

4 Simulation Design and Methodology

In real world scenarios, performance evaluation of a well-designed network model and the model itself carries significant importance. To cope with the challenge, practice using simulator OPNET (Optimized Network Engineering Tool), as kernel source code is not open source. [11,12]. In order to determine the maximum number of calls that can be supported by an existing network, while limiting VoIP delay constraint utilize queuing analysis to approximate and determine the maximum number of calls that the existing network can support while maintaining a delay of less than 80ms.,the Principles of Jackson theorem for analysing queing [12]. Analysis by decomposition is summarized in first isolating the queuing network into subsystems, example, and single queuing node[13,14,15,16]. Next, analysing each subsystem separately, considering its own network surroundings of arrivals and departures. Then, finding the average delay for each individual queuing subsystem finally, aggregating all the delays of queuing subsystems to find the average total end-to-end network delay [17,18,19,20].

Figure 4 presents an algorithm computes the maximum number of calls considering VoIP delay constraint.

a. Initially, no calls are introduced only traffic in the network background traffic.
b. A new call is added, according to the call distribution.
c. For each network element, $\lambda = \lambda VoIP + \lambda bg$ is computed. Where, $\lambda VoIP$ is the total added new traffic from a single VoIP in pps, and λbg is the background traffic in pps). The value λbg is known for each element; however, $\lambda VoIP$ can get affected by introducing a new call depending on the call traffic flow, i.e. whether or not the new call flow passes through the network element.
d. For each network element, the average delay of a VoIP packet is computed.
e. The end-to-end delay is computed by summing up all the delays of step (d) encountered for each possible VoIP flow. This includes all external and internal flows, with internal flows consisting of intra-floor and inter-floor.

f. If the maximum network delay is less than 80 ms, then the maximum number of calls has not been reached. Therefore a new call can be added, and hence go to steps (b)-(f). If not, the maximum delay has been reached. Therefore the number of VoIP calls bounded by the delay is one less than the last call addition.

Input n : number of network of elements
$\lambda[1..n]$: background traffic for network element 1,2, ... n
Delay [1..n]: delay for network element 1,2 ... n
P: set of call − flow paths (p) where p is a subset of {1,2, ... n}
Output: MaxCalls: maximum number of calls
λVoip ← 100pps, or 180.8kbps;
VoIP MaxDelay ← 80; // network delay for VoIP call in ms
MaxDelay ← 0;
MaxCalls ← −1;
Delay [1..n] ← 0;
While MaxDelay < *VoIP MaxDelay Do*
MaxCalls ← MaxCall + 1
Generate a call according to call distribution and let P_C be its flow path
for each element i in Pc do
$\lambda i \leftarrow \lambda i +$, or λvoip
if i is a link then
Delay ← $(1 - \lambda 2\mu i)/(\mu i - \lambda i)$
Else
Delay ← $1/(\mu i - \lambda i)$
end if
end for
for each p in P where $p \cap P_C \neq \emptyset$ do
PathDelay (p) ← $\sum_{n=1}^{\infty}$ Delay where i is a network element in path p
if PathDelay (p) > *MaxDelay then*
MaxDelay ← PathDelay (p)
end if
end for

Fig. 4. Algorithms number of calls VoIP [8]

5 Result and Analysis

5.1 End-to-End (E2E) Delay Performance

For VoIP applications, the packet E2E delay should not exceed 150 ms to evaluate that the quality of the created VoIP calls are accepted. In this section, the packet E2E delays result for different scenarios are presented in various statistical plots. Scenario 1, correspond to the Baseline VoIP, while Scenario 2 correspond to the Congested with VoIP. In all the scenarios, the sources and destinations happen to be started at 100 seconds as the start time of the profile and application configuration has been set to 40 seconds and 60 seconds, respectively. the figures presented in following sections, X axis represents simulation time in seconds; Y axis represents in seconds. Figure 5 illustrates the comparable performance of the E2E delay under the different case scenarios of Baseline VoIP network. E2E delay is measured for the VoIP traffic flows between source node and destination node through Baseline network with various node speeds the blue line shows the E2E delay of the scenario where the node speed is fixed (0 mps). The blue line show the E2E delay of scenarios where the node speeds are 23.83 mps, whereas red line shows in Figure 6, the congested VoIP network level of about 60.84 mps. With the increasing simulation time, the both curves settle at Simulation time 2.5 sec.

Table 1. Statistics of E2E delay of Based-line Congested VoIP Network

Baseline Before Congestion	Congestion End to End Delay	Simulation time	Baseline Before Congestion	Congestion End to End Delay	Simulation time
23	75	3.26	48.82	49.18	1.01
23.83	74.17	3.11	49.66	48.34	0.97
24.67	73.33	2.97	50.49	47.51	0.94
25.50	72.50	2.84	51.32	46.68	0.91
26.33	71.67	2.72	52.16	45.85	0.88
27.17	70.84	2.61	52.99	45.01	0.85
28.00	70.00	2.50	53.82	44.18	0.82
28.83	69.17	2.40	54.65	43.35	0.79
29.66	68.34	2.30	55.49	42.51	0.77
30.50	67.50	2.21	56.32	41.68	0.74
31.33	66.67	2.13	57.15	40.85	0.71
32.16	65.84	2.05	57.99	40.01	0.69
33.00	65.00	1.97	58.82	39.18	0.67
33.83	64.17	1.90	59.65	38.35	0.64
33.83	64.17	1.90	60.49	37.52	0.62
34.66	63.34	1.83	61.32	36.68	0.60
35.50	62.51	1.76	62.15	35.85	0.58
36.33	61.67	1.70	62.98	35.02	0.56
37.16	60.84	1.64	63.82	34.18	0.54
37.99	60.01	1.58	64.65	33.35	0.52
38.83	59.17	1.52	65.48	32.52	0.50
39.66	58.34	1.47	66.32	31.68	0.48
40.49	57.51	1.42	67.15	30.85	0.46
41.33	56.67	1.37	67.98	30.02	0.44
42.16	55.84	1.32	68.82	29.19	0.42
42.99	55.01	1.28	69.65	28.35	0.41
43.83	54.18	1.24	70.48	27.52	0.39
44.66	53.34	1.19	71.31	26.69	0.37
45.49	52.51	1.15	72.15	25.85	0.36
46.32	51.68	1.12	72.98	25.02	0.34
47.16	50.84	1.08	73.81	24.19	0.33
47.99	50.01	1.04	74.65	23.35	0.31

Packet loss refers to the failure of one or more transmitted packets to reach their destination across a network. A VoIP user is satisfied if more than 98 % of its voice packets are delivered successfully. Packet loss is determined using the following equation.

$$\left(\text{Packet Loss} = \frac{\mathit{Send\ Packet} - \mathit{Received\ Packet}}{\mathit{Send\ Packet}} \times 100\right)$$

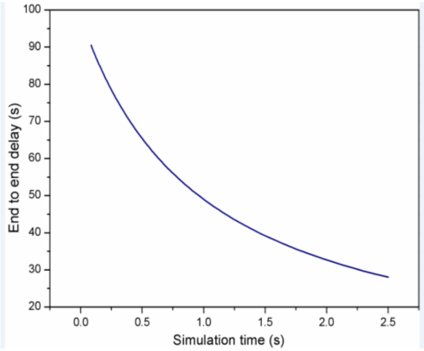

Fig. 5. End-to-End Delay of Baseline Network

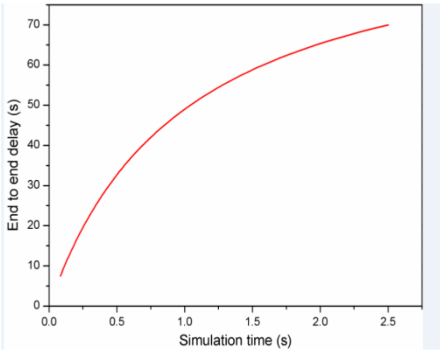

Fig. 6. End-to-End Delay of Congested VoIP Network

5.2 Packet Loss Performance for Baseline VoIP Network

In congested VoIP network. Can observe in figure, sent and received data are overlapped between 100 seconds to 140 seconds in all the simulation scenarios. After 140 seconds when the load is around 95%, the packet loss is found in all the cases. In the Congested VoIP and VoIP congested with FTP network scenario, on an average, packet loss in the VoIP Congested with FTP network for case 1 is about 74% higher than that of Congested VoIP network. In other both cases, average packet losses are 88%, 95% and 84% higher respectively. To wrap it, all of the voice traffic corresponding to each case in VoIP Congested with FTP experience higher packet loss than Congested VoIP network.

6 Conclusion

Comprehensive study, analysis and evaluation of the end-to-end delay performance evaluation for VoIP in the LTE network have been done. The evaluation is made by simulating in OPNET Modeler based on performance metrics such as end-to-end delay and throughput. Three network scenarios have been simulated: Baseline VoIP network scenario, congested VoIP network scenario and VoIP congested with FTP network scenario. It has been found that maximum throughput increased as the bandwidth incremented. Out of the three methods simulated for throughput measurement, the scenario with highest bandwidth (example 20 MHz) exhibited maximum throughput. After that, quality measurement of VoIP has been done with respect to E2E delay, both for a network congested exclusively with VoIP and VoIP with FTP. Four scenarios have been created for this evaluation, one case with stationary node and other three cases with mobile nodes (gradually increasing the node speed). The simulation results showed that when the node is not moving, E2E delay is slightly higher for network congested with VoIP only. In other cases, better E2E is obtained for this network due to the presence of moving node. Finally the rate of packet loss for the remains quite minimal for congested with VoIP network regardless of the node speed. Mean-while, for VoIP with FTP network, packet loss rate is also quite insignificant for fixed node case but the rate upshots as the node starts moving [19,20].

References

1. 3GPP, Release 8 V0.0.3.: Overview of 3GPP Release 8: Summery of all Release 8 Features (November 2008)
2. Puttonen, J., Puupponen, H.-H., Aho, K., Henttonen, T., Moisio, M.: Impact of Control Channel Limitations on the LTE VoIP Capacity. In: The Proceedings of 2010 Ninth International Conference on Networks, pp. 77–82 (2010)
3. Puttonen, J., Henttonen, T., Kolehmainen, N., Aschan, K., Moisio, M., Kela, P.: Voice-Over-IP Performance in UTRA Long Term Evolution Downlink. In: The Proceedings of IEEE Vehicular Technology Conference, pp. 2502–2506 (2008)
4. Khan, F.: LTE for 4G Mobile Broadband: Air Interface Technologies and Performance, 1st edn. University Press, Cambridge (2009)
5. Martín-Sacristán, D., Monserrat, J.F., Cabrejas-Peñuelas, J., Calabuig, D., Garrigas, S., Cardona, N.: On the Way towards Fourth-Generation Mobile: 3GPP LTE and LTE-Advanced. EURASIP Journal on Wireless Communications and Networking, 1–10 (2009)
6. Holma, H., Toskala, A.: LTE for UMTS - OFDMA and SC-FDMA Based Radio Access. John Wiley & Sons Ltd., Finland (2009)
7. Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestri-al Radio Access Network (E-UTRAN). 3GPP TS 36.300 version 8.11.0 (Janauary 2010)
8. Olorunda, O., Olorunda, A.: Bridging the Digital Divide - The Social and Cultural Impact of VoIP in Developing Countries: Nigeria as a Case Study. In: The Proceedings of PTC 2006 (2006)
9. One Way Transmission Time. ITU-T recommendation G.114 (May 2003)
10. M. J. Karam and F. A. Tobagi.: Analysis of delay and delay jitter of voice traffic in the Internet. Speedtrak Communications, San Mateo, CA 94401, USA, Computer Systems Laboratory, Stanford University, Vol. 40, pp. 711-726 (2002)
11. OPNET. Opnet modeler 16.0
12. `http://www.opnet.com/solutions/network_rd/modeler.html`
13. Salah, K., Alkhoraidly, A.: An OPNET-based simulation approach for deploying VoIP. International Journal of Network Management 16(3), 159–183 (2006)
14. Dantu, R., Fahmy, S., Schulzrinne, H., Cangussu, J.: Issues and challenges in securing VoIP. Computers & Security 28, 743–753 (2009)
15. Salah, K., Almashari, M.: An Analytical Tool to Assess Readiness of Existing Networks for Deploying IP Telephony. In: The Proceeding of 11th IEEE Symposium on Computers and Communications, ISCC 2006, pp. 301–305 (2006)
16. Masum, M.E., Babu, M.J.: End-to-End Delay Performance Evaluation for VoIP in the LTE network. Postgraduate Thesis, Blekinge Institute of Technology 23 (2011)

17. Gupta, J.D.: Performance issues for VOIP in Access Networks. Postgraduate Thesis, Department of Mathematics and Computing Faculty of Sciences, University of Southern Queensland, 87–96 (2005)
18. He, Q.A.: Analysing the Characteristics of VoIP Traffic. Postgraduate Thesis, Department of Computer ScienceUniversity of Saskatchewan, 62–77 (2007)
19. Muhleisen, M., Walke, B.: Evaluation and Improvement of VoIP Capacity for LTE. ComNets Research Group, pp. 1–7. Aachen University, Germany
20. Holma, H., Toskala, A.: LTE for UMTS: OF DMA and SC-FDM Based Radio Access, pp. 259–280. John Wiley & Sons (2009)