UNIVERSITY OF MÁLAGA
Departamento de
Lenguajes y Ciencia de la Computación

PhD Thesis

A methodology for obtaining More Realistic Cross-Layer QoS Measurements in mobile networks: A VoIP over LTE Use Case

Author:
Francisco Javier Rivas Tocado

Supervised by:
Pedro Merino Gómez
Almudena Díaz Zayas

Málaga - Spain
2015

Publicaciones y
Divulgación Científica

UNIVERSIDAD
DE MÁLAGA

AUTOR: Francisco Javier Rivas Tocado

http://orcid.org/0000-0003-2513-0496

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga

El Dr. Don Pedro Merino Gómez, Titular de Universidad del Área de Ingeniería Telemática de la Universidad de Málaga, y la Dr. Almudena Díaz Zayas

Certifican que D. Francisco Javier Rivas Tocado, Ingeniero de Telecomunicación, ha realizado en el Departamento de Lenguajes y Ciencias de la Computación de la Universidad de Málaga, bajo su dirección, el trabajo de investigación correspondiente a su Tesis Doctoral titulada:

A methodology for obtaining More Realistic Cross-Layer QoS Measurements in mobile networks: A VoIP over LTE Use Case

Revisado el presente trabajo, estiman que puede ser presentado al tribunal que ha de juzgarlo, y autorizan la presentación de esta Tesis Doctoral en la Universidad de Málaga. Así mismo certificamos que los trabajos enumerados en el apartado de Publicaciones y que justifican esta tesis no han sido utilizados en tesis anteriores.

Málaga, Octubre de 2015

Fdo.: Pedro Merino Gómez
Titular de Universidad del Área de Ingeniería Telemática.

Fdo.: Almudena Díaz Zayas
Doctora por la Universidad de Málaga.

Contents

List of Figures

List of Tables

Acknowledgements

To my loving wife and family, for their constant support and for all the time I have borrowed from them. To Pedro Merino, for his mentoring and guidance, and for being an early supporter of the research of IP communications over mobile phones. To my work and lab colleagues, that have inspired me through all these years and have made this journey as pleasant as it can be.

Resumen (in Spanish)

Los servicios de voz han sido durante mucho tiempo la primera fuente de ingresos para los operadores móviles. Incluso con el protagonismo creciente del tráfico de datos, los servicios de voz seguirán jugando un papel importante y no desaparecerán con la transición a redes basadas en el protocolo IP. Por otra parte, hace años que los principales actores en la industria móvil detectaron claramente que los usuarios no aceptarían una degradación en la calidad de los servicios de voz. Es por esto que resulta crítico garantizar la experiencia de usuario (QoE) en la transición a redes de nueva generación basadas en conmutación de paquetes.

El trabajo realizado durante esta tesis ha buscado analizar el comportamiento y las dependencias de los diferentes servicios de Voz sobre IP (VoIP), así como identificar configuraciones óptimas, mejoras potenciales y metodologías que permitan asegurar niveles de calidad aceptables al mismo tiempo que se trate de minimizar los costes.

I. Introducción

La caracterización del rendimiento del tráfico de datos en redes móviles desde el punto de vista de los usuarios finales es un proceso costoso que implica la monitorización y análisis de un amplio rango de protocolos y parámetros con complejas dependencias. Para abordar desde la raíz este problema, se requiere realizar medidas que relacionen y correlen el comportamiento de las diferentes capas. La metodología de caracterización propuesta en esta tesis proporciona la posibilidad de recoger información clave para la resolución de problemas en las comunicaciones IP, relaciolándola con efectos asociados a la propagación radio, como cambios de celda o pérdida de enlaces, o con carga de la red y limitaciones de recursos en zonas geográficas específicas.

Dicha metodología se sustenta en la utilización de herramientas nativas de monitorización y registro de información en smartphones, y la aplicación de cadenas de herramientas para la experimentación extensiva tanto en redes reales y como en entornos de prueba controlados. Con los resultados proporcionados por esta serie de herramientas, tanto operadores móviles y proveedores de servicio como desarrolladores móviles podrían ganar acceso a información sobre la experiencia real del usuario y sobre cómo mejorar la cobertura, optimizar los servicios y adaptar el funcionamiento de las aplicaciones y el uso de protocolos móviles basados en IP en este contexto.

Las principales contribuciones de las herramientas y métodos introducidos en esta tesis son los siguientes:

- Una herramienta de monitorización multicapa para smartphones Android, llamada TestelDroid, que permite la captura de indicadores clave de rendimiento desde el propio equipo de usuario. Asimismo proporciona la capacidad de generar tráfico de forma activa y de verificar el estado de alcanzabilidad del terminal, realizando pruebas de conectividad.

- Una metodología de post-procesado para correlar la información presente en las diferentes capas de las medidas realizadas. De igual forma, se proporciona la opción a los usuarios de acceder directamente a la información sobre el tráfico IP y las medidas radio y de aplicar metodologías propias para la obtención de métricas.

- Se ha realizado la aplicación de la metodología y de las herramientas usando como caso de uso el estudio y evaluación del rendimiento de las comunicaciones basadas en IP a bordo de trenes de alta velocidad.

- Se ha contribuido a la creación de un entorno de prueba realista y altamente configurable para la realización de experimentos avanzados sobre LTE.

- Se han detectado posibles sinergias en la utilización de instrumentación avanzada de I+D en el campo de las comunicaciones móviles, tanto para la enseñanza como para la investigación en un entorno universitario.

II. Estudio del tráfico de Internet sobre redes móviles

De acuerdo a Cisco [1], el tráfico de datos global creció un setenta por ciento en 2012, alcanzando cerca de doce veces el tamaño de todo el tráfico de Internet en el año 2000. Por otra parte, el tráfico móvil de datos excedió el 50 por ciento por primera vez, poniendo de manifiesto la importancia creciente del tráfico multimedia. Asimismo, según el Ericsson Mobility Report [2], en 2014 el número de subscripciones de smartphones superaba los 2600 millones, y las previsiones esperan que en 2020 se hayan doblado.

De forma un tanto inesperada, los operadores han visto recientemente cómo terceras partes han acaparado buena parte del mercado de voz proporcionando servicios sobre sus redes a un coste muy reducido o incluso nulo, pero lamentablemente esto ha repercutido directamente en la calidad de los servicios. En este contexto, los usuarios demandan a los operadores una conectividad global que mantenga altos niveles de calidad en cualquier lugar. Sin embargo, no resulta infrecuente encontrarse también en sus redes con problemas de rendimiento y conectividad que impactan de manera muy severa la experiencia de los usuarios móviles.

Para mantener niveles adecuados de calidad, los operadores, proveedores de contenido y los desarrolladores requieren herramientas apropiadas para monitorizar el rendimiento de los servicios de datos. Los operadores de red tienen acceso típicamente a diferentes niveles de información, obteniendo indicadores clave de rendimiento (KPI) sobre el funcionamiento de sus redes. Los KPI se basan esencialmente en el registro de contadores de rendimiento que son recogidos de los distintos elementos de red. Estos contadores de rendimiento son definidos por las especificaciones técnicas del 3GPP (TS 32.405, TS 32.450, etc) e implementados en elementos de red como los RNC, MSC, SGSN o e-NodeBs. Sin embargo, estos contadores pueden no reflejar la experiencia real de los usuarios debido a la no utilización de sofisticadas funciones de filtrado y correlación en las funciones estáticas de los elementos de red [3].

Los proveedores de contenido y los desarrolladores de aplicaciones tradicionalmente han tenido que probar sus aplicaciones en un entorno de usuario sin ningún conocimiento sobre la configuración de red que estaba siendo usada. Tradicionalmente, las herramientas de monitorización usadas para evaluar la experiencia de usuario, requerían la utilización de hardware dedicado, y estaban al alcance únicamente de los operadores de red por su coste prohibitivo. Sin embargo, recientemente ha sido posible usar los propios smartphones como dispositivos de monitorización con la ayuda de aplicaciones software.

III. Sobre la necesidad de un entorno de prueba realista para experimentación controlada

La mejora de la calidad de servicio (QoS) de forma sostenible es un objetivo clave para los operadores de red porque las tareas de gestión han ido adquiriendo una compejidad incremental. A pesar de los esfuerzos de los organismos de estandarización, todavía hay un salto significativo que cubrir para optimizar tanto QoS como QoE. A día de hoy, a pesar todavía son frecuentes las estimaciones de QoS y QoE derivadas de costodas campañas de prueba en campo, también conocidas como drive test. Además, se ha requerido con frecuencia de analistas experimentados para ajustar la configuración de las redes.

Por otra parte, la mayoría de las configuraciones de red y de servicios propuestas en la literatura, están derivadas de simulaciones [4] [5] [6] [7] [8] [9]. Como es ampliamente conocido, en el proceso de modelado de los sistemas de comunicaciones para ser simulados, muchos detalles pueden ser abstraídos o pasados por alto. Estas inexactitudes pueden causar que los resultados obtenidos puedan derivar en conclusiones inexactas o erróneas Por ejemplo, es muy frecuente encontrar que el consumo de recursos asociado al plano de control es directamente ignorado cuando se evalúan diferentes estrategias de planificación. Así pues, es imprescindible contrastar y complementar los resultados de las simulaciones con medidas obtenidas directamente en los terminales de los propios usuarios. Es por tanto necesario correlar dichas medidas con la información recogida por la propia red para reducir los costes y ajustar de forma más precisa la operación de las redes a la calidad percibida por los usuarios finales. Más aún, como han expresado las organizaciones y alianzas participadas por operadores de red, como NGMN [10], “la optimización de QoS requiere todavía de desarrollos “reales” para seguir estudiando la dirección hacia la que moverse”.

En esta tésis, se propone el uso de un entorno de prueba controlado [11] en que poder realizar experimentos sobre tecnología 4G LTE (Long Term Evolution) en un contexto realista. De esta forma se pretende permitir la correlación de las configuraciones radio con la calidad de servicio percibida a nivel de aplicación. La ejecución de campañas de medida exhaustivas usando este entorno permitirá la identificación de nuevos contadores de rendimiento, la relación de los mismos entre sí, y la puesta en práctica de casos de uso para la optimización de QoS y QoE en redes LTE.

El foco de esta tesis está depositado en los servicios de voz sobre IP (VoIP) en redes LTE, los cuales representan nuevos desafíos comparados con las tecnologías previas. En LTE, las llamadas de voz son proporcionadas mediante una red basada enteramente en IP (all-IP), en lugar de una red basada en conmutación de circuitos, lo que implica que la voz tiene que competir por ancho de banda con otros servicios proporcionados por la red. Resulta vital para los operadores garantizar al menos la misma calidad de experiencia usando VoIP que aquella disponible en los sistemas pre-LTE como 2G GSM (Global System for Mobile Communications) y 3G UMTS (Universal Mobile Telecommunications System).

De no ser así, los operadores perderán una importante oportunidad para diferenciarse en calidad respecto de los proveedores de servicio externos, y los usuarios tendrán menos motivos para no decantarse por opciones de menor coste.

El entorno de prueba ha sido concebido para permitir la validación del rendimiento de las configuraciones de red y de los problemas presentes en la literatura de investigación. Especialmente en el caso de los servicios de VoIP, hemos correlado las configuraciones de parámetros radio LTE de capa 1 (PHY) y capa 2 (MAC), con indicadores de rendimiento a nivel IP y de calidad de voz como el MOS (Mean Opinion Score) obtenido a partir de la aplicación del algoritmo PESQ (Perceptual Evaluation of Speech Quality).

En orden a capturar los parámetros de rendimiento IP y las medidas MOS, se han acometido dos desarrollos en el marco de esta tesis:

- Una herramienta software para dispositivos Android, que proporciona funciones avanzadas de monitorización, tales como el registro de parámetros radio y la captura de tráfico IP.

- Una cadena de post-procesado enfocada en los procesos de captura y análisis de indicadores clave para servicios de tiempo real como VoIP. La variación del retardo o jitter, la detección de paquetes perdidos, el espaciado entre paquetes y el algoritmo PESQ son algunas de las métricas utilizadas. De esta forma se permite una caracterización completa del comportamiento de los servicios sobre redes celulares.

La organización de esta tésis es la siguiente. El Capítulo 1 introduce la necesidad de obtener medidas de rendimiento que capturen la QoS y QoE como es percibida por los usuarios finales, así como los nuevos desafíos introducidos por la provisión de servicios de llamada de voz sobre LTE. Este capítulo proporciona de igual forma una breve introducción al servicio de VoIP sobre LTE y menciona algunas configuraciones de red reseñables propuestas en la literatura para optimizar la operativa de LTE. En el Capítulo 2 se realiza una revisión más detallada del estado del arte originario sobre servicios VoIP y su provisión sobre LTE.

El Capítulo 3 introduce TestedDroid, una herramienta de monitorización para dispositivos móviles, y la cadena de herramientas de post-procesado para el análisis de la información obtenida. El Capítulo 4 se enfoca en la validación de las herramientas y de la metodología introducida en el capítulo previo. Este capítulo proporciona los resultados obtenidos durante una serie de campañas de prueba del servicio de VoIP en redes móviles comerciales, tanto en escenarios estáticos como vehiculares. Un análisis del rendimiento de VoIP con especial énfasis en escenarios de alta velocidad se presenta en el Capítulo 5. El Capítulo 6 describe el entorno de prueba propuesto para la realización de experimentos avanzados en un entorno realista al tiempo que controlado. El capítulo 7 proporciona los resultados de las pruebas realizadas para demostrar y validar las funcionalidades del entorno de prueba propuesto en el capítulo 6. Los detalles sobre las configuraciones utilizadas para la realización de las medidas, los resultados obtenidos durante el análisis de rendimiento IP, y su correlación con los parámetros LTE se han incluido de igual manera. Por último, el capítulo 8 se presentan los trabajos futuros basados en un instrumento de nueva generación como el UXM Wireless Test Set de Keysight technologies, para su uso en múltiples campos de investigación.

IV. Usando los smartphones como herramientas de medida

Los smartphones se han ido convirtiendo en una de las mayores plataformas de ejecución de servicios de internet a medida que dispositivos más potentes se han hecho disponibles a un precio razonable. La verificación del rendimiento de dichos servicios ha ido por tanto moviéndose desde las antiguas configuraciones, en las que era necesario utilizar un modem externo para dar conectividad a un ordenador, hasta la situación actual en la que se utilizan los smartphones u otros dispositivos análogos de forma autocontenida. En línea con esta evolución, ha sido necesario un desarrollo paralelo de las herramientas de evaluación y medida, de manera que fuera posible el análisis de servicios y aplicaciones específicamente diseñados para dispositivos móviles. Es por tanto comprensible que las propias herramientas de medida se hayan desarrollado e integrado de forma nativa como aplicaciones en los propios terminales móviles.

Siguiendo el mencionado enfoque, en trabajos anteriores se desarrolló la herramienta SymPA [12] para terminales basados en el sistema operativo Symbian OS. Dicha herramienta fue aplicada en el estudio del servicio de video streaming en redes móviles [13]. Con TestelDroid, hemos pretendido incrementar las funciones de monitorización incorporando las capacidades proporcionadas por el sistema operativo Android, y de esta forma extender nuestro campo de investigación y experimentación a los dispositivos, aplicaciones y servicios basados en Android.

Las primeras herramientas diseñadas para funcionar en dispositivos móviles, como Qualipoc [94], estaban centradas en la observación de parámetros de accesibilidad de los servicios, tales como la disponibilidad, el ancho de banda proporcionado, la tasa de error, etc. Para tal fin, incluían funcionalidades activas en modo cliente para hacer uso de los servicios, desde iniciando conexiones FTP (File Transport Protocol) hasta llamadas de voz y mensajes de texto. Ese tipo de clientes weran usados para probar los servicios y verificar los parámetros generales. Sin embargo, la creciente complejidad de las redes móviles y de las aplicaciones ha requerido de soluciones que fueran un paso más allá hacia el análisis del rendimiento radio y de las aplicaciones. Mediante ese enfoque, se ha buscado identificar las fuentes reales de los problemas de comunicación y así ayudar a mejorar la experiencia de los usuarios.

Aunque la prueba de servicios mediante generación activa tiene utilidad innegable en algunos casos de uso, resulta claramente inviable escalar esa técnica para incluir clientes para cada nueva aplicación o servicio presente o futuro. Es por tanto que la monitorización pasiva representa una alternativa natural y escalable para el análisis del rendimiento de la comunicación en aplicaciones móviles. Es posible tener en cuenta múltiples planos mediante un enfoque pasivo, incluyendo tanto una perspectiva de rendimiento de comunicación, como de uso de memoria o consumo de batería entre otras.

En el marco de esta tésis se ha desarrollado y aplicado la herramienta TestelDroid en el análisis del tráfico VoIP en un entorno de comunicación móvil real. Esta herramienta proporciona múltiples y variadas funcionalidades para capturar tráfico, monitorizar la red móvil, geolocalizar las medidas mediante GPS y evaluar el consumo de potencia como muestra de sus capacidades. Los datos recogidos por TestelDroid son almacenados para su posterior post procesado, utilizando una cadena de herramientas que importa funcionalidad entre otras fuentes de wireshark o Google Earth.

Como se ha mencionado, TestelDroid ha sido aplicada para realizar un anális de la calidad de VoIP sobre redes móviles, en concreto utilizando tecnologías GPRS (General Packet Radio Service), UMTS(Universal Mobile Telecommunications System) y HSPA (High Speed Packet Access).

Adicionalmente, Testeldroid complementa las funciones de monitorización con generación activa de tráfico y pruebas de conectividad. La generación activa está orientada a facilitar las pruebas en conexiones TCP en un entorno de móvil a móvil, el cual no puede apoyarse en la presencia de servidores externos. Así pues, un dispositivo móvil se configura como servidor mientras que el otro dispositivo adopta el rol de cliente. Mediante el intercambio de un fichero de tamaño configurable, se genera un volumen de información suficiente para monitorizar la calidad del enlace. Las funciones de prueba de conectividad proporcionan dos modos de diagnóstico para facilitar la operativa de las pruebas, lo cual es particularmente necesario para detectar problemas relacionados con firewalls y NATs. El primero permite verificar la alcanzabilidad del otro extremo de comunicación mediante el intercambio configurable de información ICMP (Internet Control and Mesage Protocol), también conocida como ping. El segundo permite el establecimiento de conexiones TCP originadas desde el extremo móvil, dada una dirección IP y un puerto TCP destino.

Testeldroid recolecta la información que necesita utilizando interfaces de programación estándar (API) del sistema operativo Android, del núcleo de Linux y del entorno de ejecución de Android. El API de Android proporciona acceso a las funciones de red, tales como la consulta del identificador de celda, las celdas vecinas disponibles y la intensidad de señal, así como el acceso a datos sobre la batería y la localización GPS. El entorno de ejecución de Android se requiere para implementar las funciones activas de chequeo de conectividad TCP y de generación de tráfico. Por último, para implementar la captura de paquetes, el intercambio de información ICMP y para tener acceso al consumo de batería, es necesario disponer de acceso al propio núcleo de Linux, lo que implica disponer de privilegios de superusuario.

La información proporcionada por TestelDroid en estos tres niveles, se puede apreciar en mayor detalle en la figura 3.3.

Hemos realizado una caracterización del rendimiento de TestelDroid, monitorizando llamadas VoIP en un terminal móvil Nexus One durante 8 horas, en términos de utilización de CPI, uso de memoria y consumo de energía. La utilización media de CPU es menor del 1%. El uso de memoria es de 22 MBytes, lo cual es un valor habitual para un proceso Android. Finalmente, el consumo de energía es solamente un 0.2

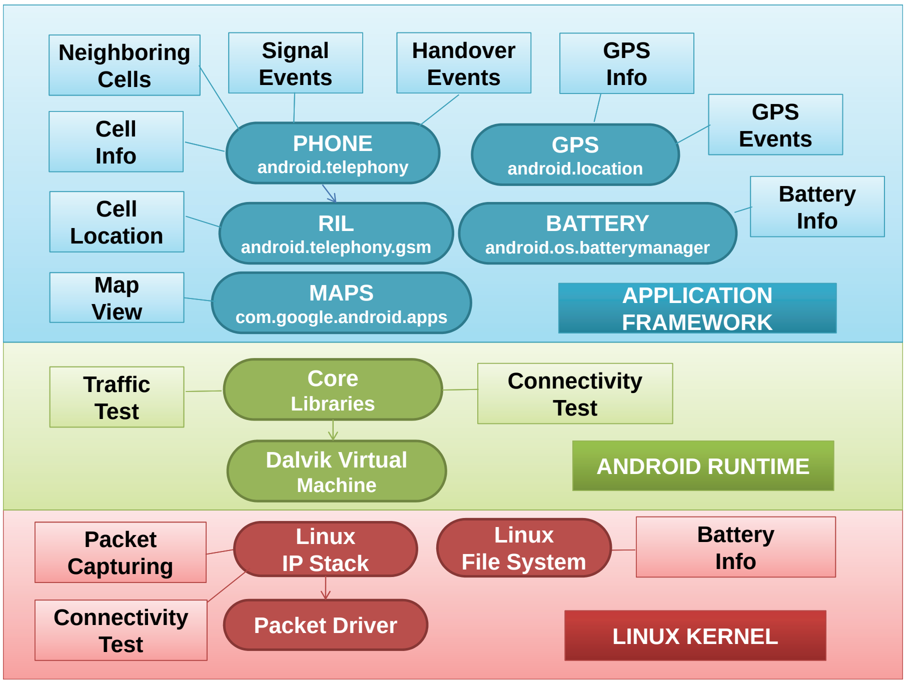

Figure 1: TestelDroid diagram

V. Metodología y cadena de herramientas de post-procesado

En el marco de esta tesis se ha desarrollado también una metodología de post-procesado automático para correlar las distintas fuentes de información proporcionadas por TestelDroid. La herramienta de procesado proporciona múltiples resultados, las cuales permiten una caracterización cruzada entre capas del rendimiento de las comunicaciones IP en terminales móviles.

En particular, para identificar la fuente de potenciales problemas de comunicación, se han desarrollado dos diferentes representaciones que relacionan las magnitudes bajo análisis: la representación temporal y la vista geográfica. En la primera, se presenta la evolución de la información de las distintas capas sobre un eje temporal. Esta vista es útil para estudiar con precisión la temporización de eventos específicos, como por ejemplo la duración de los handovers entre distintas tecnologías y su impacto en el tráfico a nivel IP (pérdidas de paquetes, interrupción de la conexión, etc).

Por otra parte la representación geográfica es de gran ayuda para situar las medidas en el contexto en que fueron obtenidas. Este es un aspecto clave para entender la raíz de problemas asociados a ubicaciones específicas. En un análisis generalista de otras magnitudes, estos problemas podrían pasar fácilmente desapercibidos. Sin embargo, ser consciente de la información de localización asociada a las medidas, puede permitir hacer un seguimiento directo de las degradaciones e interrupciones de servicio hacia zonas de cobertura especialmente limitada, por citar uno de los posibles usos de esta representación.

Más detalles pueden consultarse en [15].

Siendo VoIP el objeto de interés en este trabajo, durante los experimentos llevados a cabo para validar esta metodología se han usado señales de prueba de audio recomendadas por ITU-T P.501 [86] para medidas de telefonía.

Esta metodología se sustenta en el uso de Testeldroid y un conjunto completo de scripts de procesado que proporciona una representación tanto numérica como gráfica de las medidas obtenidas. El tráfico IP capturado es almacenado en formato libpcap. Dichos ficheros son posteriormente filtrados usando TShark, una utilidad incluida en la distribución del conocido analizador de protocolos Wireshark, para detectar los identificadores de SSRC (Synchronization Source) de los flujos RTP (Real-time Transport Protocol) contenidos en el tráfico IP capturado. Los identificadores SSRC se asignan usando un valor aleatorio buscando que sea globalmente único para una sesión RTP particular. Usando los SSRCs, es posible separar los paquetes RTP por flujos y analizarlos de forma independiente para obtener información sobre la evolución temporal de parámetros clave: retardo entre paquetes, variación del retardo, pérdida de paquetes y errores de secuencia. Esta información es entonces correlada con la información de nivel radio de cara a analizar el impacto sobre el tráfico IP de la propagación radio (ruido, desvanecimienos, interferencias) y de los procedimientos de movilidad como reselecciones de celda o handovers.

El algoritmo PESQ (Perceptual Evaluation of Speech Quality) [92] es un algoritmo objetivo que compara una señal de referencia con una señal degradada y proporciona una métrica de calidad llamada MOS (Mean Opinion Score). En nuestra metodología, el algoritmo PESQ ha sido aplicado a la señal degradada reconstruida en el extremo receptor para estimar de esta forma la pérdida de calidad en la transmisión de información extremo a extremo de VoIP. El sonido VoIP se recupera extrayendo el contenido encapsulado en los paquetes RTP capturados a nivel IP, de esta forma se consigue también independizar el resultado de las características específicas del terminal de la aplicación receptores. Por ejemplo, no se vería afectado por el uso de gestión adaptativa del buffer de compensación de variación de retardo ni por las implementaciones específicas de algoritmos de gestión de paquetes perdidos.

VI. Caracterización del rendimiento del tráfico en redes móviles

Este caso de uso se enfoca en la evaluación del rendimiento de VoIP sobre HSPA a nivel de aplicación, aunque durante los experimentos la tecnología de acceso o RAT (Radio Access Tecnology) era susceptible de cambiar a UMTS o GPRS, dependiendo de la cobertura y del estado de la red. Hay muchos trabajos en la literatura sobre este tópico, pero generalmente están basados en simulaciones y enfocados en las capacidades del interfaz aire, funcionalidades de MAC-hs como planificación rápida de paquetes [95], y en el impacto de parámetros RRC (Radio Resource Control) como los ciclos de recepción discontínua (DRX) [96]. Estos trabajos estudian exhaustivamente la planificación de paquetes y la asignación de recursos para la obtención de resultados relevantes desde el punto de vista de la gestión de red. Sin embargo, no parece suficientemente probado el impacto en el rendimiento a nivel de aplicación desde el punto de vista de usuarios individuales.

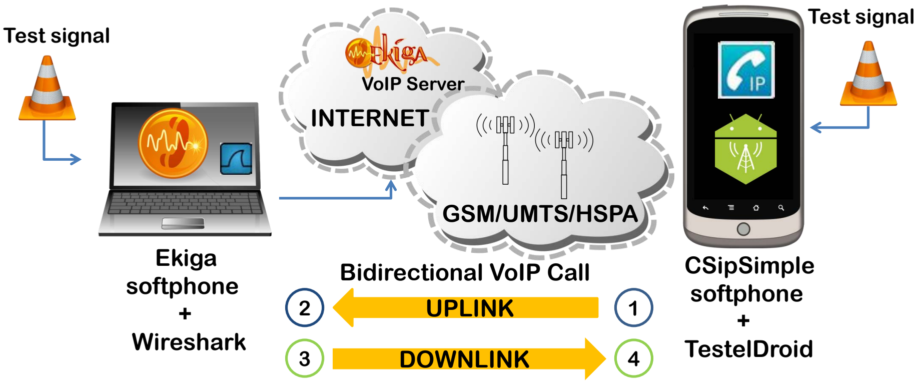

Figure 2: VoIP experimental scenario over live cellular networks

Parámetros como la tasa de pérdida de paqutes, errores de secuencia y variación del retardo no han sido considerados en los trabajos anteriormente referenciados. Aunque estos parámetros han sido objeto de estudio en [97] y [98], dichos estudios están orientados a probar la bondad de diferentes algoritmos de planificación mediante simulaciones.

Los estudios llevados a cabo en [99] y [100], muestran cierta similaridad con el trabajo aquí introducido, con la particularidad de que en este se proporciona una metodología abierta que permite el estudio del rendimiento de cualquier aplicación teniendo en cuenta las características de la propagación radio.

VI.I. Escenario de pruebas de campo

Con el propósito de verificar TestelDroid en un despliegue móvil, se ha usado la herramienta para verificar el rendimiento de VoIP tanto en escenarios estáticos como vehiculares sobre redes comerciales. Para ello hemos usado un proveedor de VoIP comercial, Ekiga.net, que proporciona también una aplicación software de telefonía para PC. La figura 4.1 muestra la estructura del experimento, donde dos clientes de VoIP son configurados para establecer comunicación bidireccional. Los datos recogidos son posteriormente proporcionados para obtener las métricas de rendimiento buscadas.

De cara a validar la utilización de TestelDroid, hemos comparado los valores obtenidos de variación media del retardo o jitter con los de trabajos anteriores. Promediando los resultados de las medidas estáticas, hemos obtenido un jitter de 6 ms para HSDPA y de 13 ms para HSUPA, que están en línea con los resultados obtenidos en [99] y [100]. En entorno vehicular, hemos obtenido un jitter medio de 23 ms, aunque los resultados muestran una dispersión alta causada por las pérdidas de paquetes introducidas por los cambios de celda y por el impacto de la propagación radio en movilidad. En todo caso, estos valores de jitter son también coherentes con los proporcionados en [99].

Por otra parte, hemos observado algunas degradaciones dramáticas en la calidad de la señal de voz recibida. Correlando la información temporal sobre la tecnología de acceso, ha sido posible detectar que la reducción de calidad ocurría en transiciones a GPRS. En el experimento se ha utilizado el códec G.711, de 80kbps de ancho de banda constante. Sin embargo en conexioens GPRS el ancho

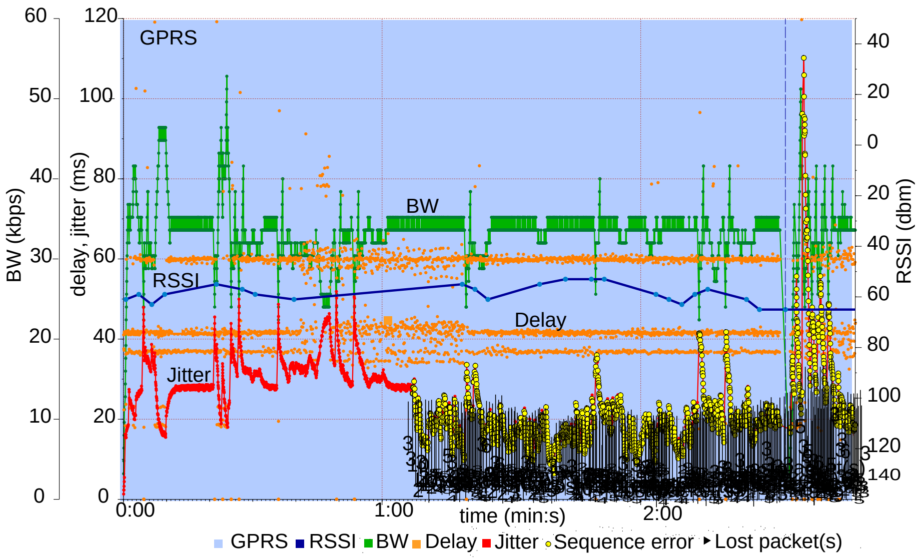

Figure 3: VoIP traffic received by CSipSimple during a GPRS connection (Downlink receiver side)

de banda disponible se reduce a valores tan reducidos como 35 kbps (ver Figura 4.5) al tiempo que los clientes de VoIP analizados no reaccionan a esta variación en la capacidad del canal de comunicación (ver Figura 4.4). Esto trae como consecuencia que las formas de onda de audio enviadas, por el cliente Ekiga desde el PC, no puedan ser recibidas en tiempo real y por tanto se aprecie un retardo continuamente incremental. La Figura 4.6 muestra el comportamiento descrito, en el que se muestra la señal original transmitida en la parte inferior, la señal recostruida de acuerdo a las marcas de tiempo RTP en el medio, y por último la señal tal como fue recibida en la parte superior. Puede observarse claramente cómo esta última está expandida, acumulando un incremento en el retardo de 500 ms en tan sólo 5 segundos.

Huelga decir que la comunicación efectiva es imposible en esas condiciones, lo que pone de manifiesto una limitación presente en la provisión de servicios de voz desde nivel de aplicación sin tener control de la configuración de red subyacente ni de los recursos asignados para garantizar una calidad de servicio adecuada. Casos similares representan por otra parte una oportunidad para que los operadores de red ofrezcan un servicio diferenciado en calidad.

VII. Estudio del rendimiento del tráfico de Internet en ferrocarriles de alta velocidad

El transporte en ferrocarril utiliza tecnologías específicas como GSM-R para el intercambio de señalización de tráfico ferroviario, sin embargo las redes móviles comerciales podrían también proporcionar valor añadido a los pasajeros y compañías de transporte, como servicios de televisión en circuito cerrado o telemetría. Más aún, los usuarios móviles demandan con mayor frecuencia una

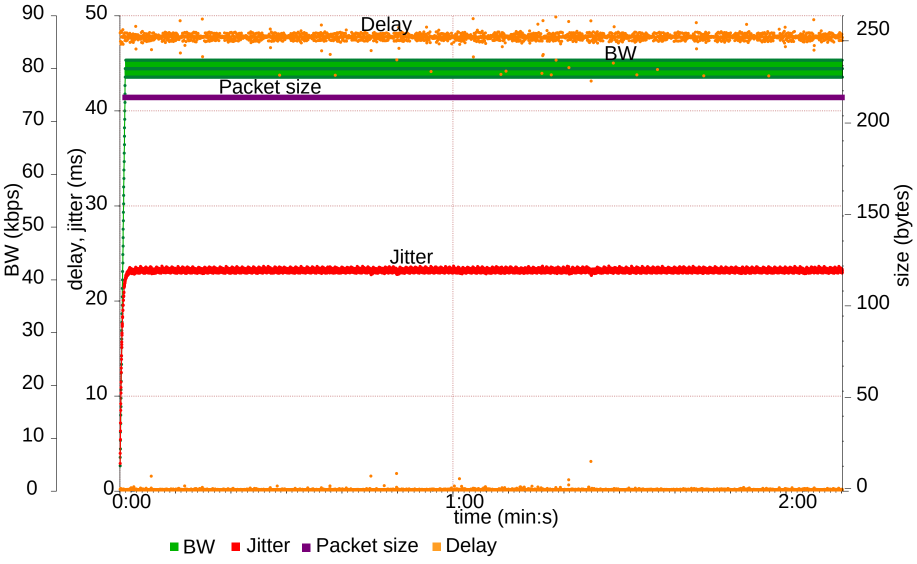

Figure 4: VoIP traffic transmitted by Ekiga softphone during a conversation with a mobile connected through a HSDPA connection (Downlink source side)

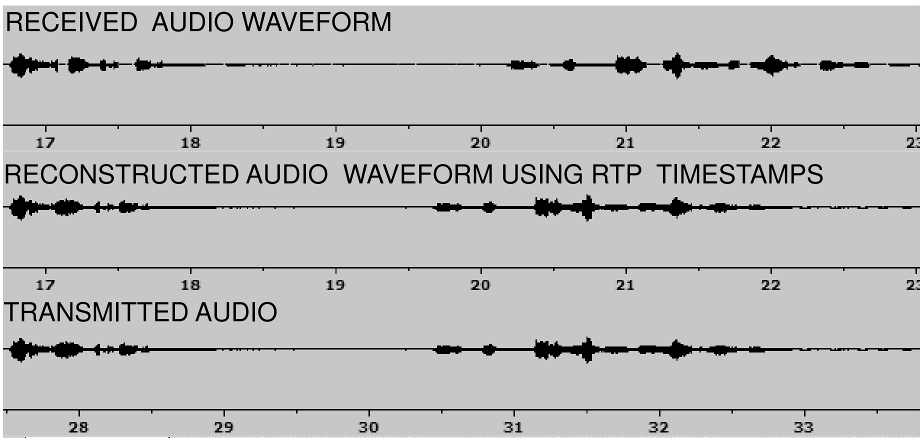

Figure 5: Transmitted and received waveforms comparison in HSDPA

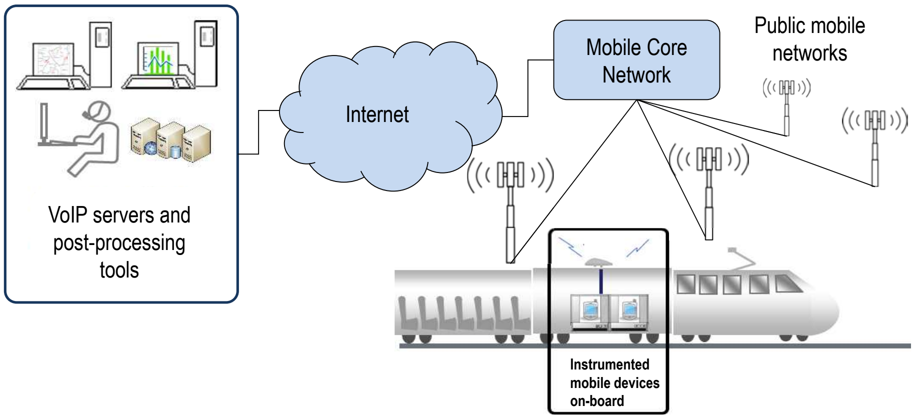

Figure 6: On-board monitoring with instrumented mobile devices

conectividad continua mientras viajan. Aunque las tecnologías desplegadas como 3G proporcionan cobertura en general, en lo respectivo a escenarios de alta movilidad como el ferrocarril, se requiere de un mayor análisis para garantizar una adecuada experiencia de usuario.

En el marco de esta tesis, se han aplicado las metodologías y herramientas desarrolladas para recoger y analizar información sobre el tráfico de servicios de Internet en redes móviles comerciales en un entorno ferroviario. En concreto, hemos realizado un estudio extensivo sobre una línea de alta velocidad en el sur de España. Las campañas de medida han abarcado cientos de sesiones VoIP, comparando diferentes operadores de red sobre más de 155 kilómetros.

Hemos llegado a la conclusión de que la metodología de caracterización puede ser útil para proporcionar información a los operadores de ferrocarril (por ejemplo para la migración de GSM-R a LTE), operadores de telecomunicación (para verificar y adaptar sus despliegues de red) y para desarrolladores de software móvil (para adaptar sus aplicaciones a un entorno de alta velocidad).

VII.I Sistema de monitorización embarcado

La Figura 7.1 describe configuración del entorno de medida utilizado en ferrocarriles de alta velocidad. Hemos usado la herramienta TestelDroid, desarrollada en la Universidad de Málaga, para capturar el tráfico IP recibido en los propios terminales así como información adicional de red usando teléfonos comerciales Android. Las campañas de medida se han realizado sobre una línea de ferrocarril de 155 km entre las ciudades de Málaga y Córdoba. Hemos utilizado diferentes terminales, durante múltiples trayectos en ambos sentidos para obtener información estadísticamente relevante. Para transportar el tráfico generado durante los experimentos, hemos utilizado el acceso a Internet comercial por defecto proporcionado por los dos operadores principales de telefonía móvil.

VII.II. Medidas de VoIP

De cara a facilitar la caracterización de red hemos utilizado una fuente de tráfico UDP/IP de tasa de bit constante. En concreto, para ese fin se han utilizado

Test	Parameter	Operator 1	Operator 2
VoIP	max jitter > 25 ms	87.37 %	71.46 %
	mean jitter > 25 ms	2.52 %	7.47 %
	packet loss > 1%	1.26 %	10.85 %
	max packet loss	63 %	65 %
	max interpacket delay	21 s	27 s
	2.5 < PESQ MOS < 3.5	4 .04 %	9.34 %
	PESQ MOS ≤ 2.5	6.56 %	9.09 %
FTP	SACK missing segments	0.03 %	3.25 %
	SACK duplicated packets	0.74 %	0.09 %
	Out of order	0.00 %	1.34 %
	mean transfer time	248.57 s	270.20 s

Table 1: A comparative summary of the results

sesiones de tráfico VoIP empleando el códec G.711. El tráfico de VoIP se ha generado desde un servidor Asterisk y se ha recibido en los terminales Android que ejecutan el software de monitorización TestelDroid en segundo plano.

Asimismo, en paralelo hemos realizado más de 500000 medidas del tiempo de ida y vuelta mediante desde el servidor utilizado a la red del operador mediante ICMP. El propósito de dichas medidas ha sido garantizar que las variaciones del retardo observadas se debieran única y exclusivamente a la operativa de la red móvil y a las condiciones de alta velocidad. La mayoría de los paquetes ICMP (97.5%) ha experimentado un retardo exactamente de 12 ms, con una desviación típica de 0.59 ms.

La Tabla 5.2 muestra un resumen comparativo de los resultados obtenidos.

Los resultados de las medidas revelan que las ráfagas de paquetes perdidos y la eventual presencia de grandes picos de retardo son los principales problemas detectados a nivel IP. Estos pueden deberse a los siguientes factores:

- Frecuentes cambios de celda, que habitualmente se han observado que no están libres de errores.
- Condiciones radio deficientes en los extremos de las celdas, particularmente cuando el RSSI disminuye por debajo de -100dBm, debido a una elevada distancia a las estaciones base.
- Asignación temporal de portadoras de datos de ancho de banda limitado, que no son capaces de sostener el volumen de tráfico necesario para un servicio de tasa constante.
- Expropiación de recursos por condiciones de carga de la red y competencia con llamadas de voz de mayor prioridad
- Interferencia de celdas vecinas, particularmente en las transiciones entre zonas rurales y urbanas como ciudades o pueblos.

VII.III. Medidas del rendimiento de TCP

Adicionalmente, hemos analizado el comportamiento de las conexiones TCP correspondientes al intercambio de ficheros mediante FTP. Un total de 40 descargas FTP se han llevado a cabo en el mismo escenario de alta movilidad ferroviaria. Se han utilizado ficheros de 26Mbytes y los tamaños por defecto para la ventana de recepción (65536 bytes). La duración medida de las transferencias oscila entre los 2 y los 7 minutos, con una desviación típica de 70 segundos, con el operador 1 comportándose un 10% mejor en promedio. En la Tabla 5.2 se muestran estadísticas adicionales de forma comparativa. Los terminales móviles usan reportes de confirmación selectiva (SACK) para indicar intervalos de números de secuencia que pueden haber sido perdidos o recibidos fuera de orden, pero también pueden usarse para indicar recepción de paquetes duplicados.

Podemos ver que el operador 1 consigue minimizar la presencia de reportes de paquetes perdidos pero por el contrario se manifiesta un mayor número de paquetes duplicados. Más del 1% de los paquetes del operador 2 son recibidos fuera de orden, y el 3% de los paquetes en sentido contrario contienen reportes SACK avisando de segmentos TCP no recibidos, lo cual es dos órdenes de magnitud superior respecto del operador 1.

Lamentablemente, hemos encontrado también algunos comportamientos no esperados que degradan la comunicación TCP significativamente. En concreto, hemos observado que en ocasiones a nivel TCP hay problemas para recuperar la dinámica de la conexión cuando se recibe la notificación mediante SACK de que se ha perdido un conjunto grande de paquetes. En esta situación, TCP puede entrar eventualmente en un estado en el que se alterna un paquete nuevo con uno retransmitido en intervalos de uno a dos segundos sin que se incremente la ventana de transmisión. Esto origina que tome unos 30 segundos retransmitir los paquetes perdidos reportados y que entonces TCP se recupere de este estado. Este fenómeno ha ocurrido en la práctica totalidad de las llamadas del operador dos en las que se ha observado un aumento significativo del tiempo de conexión. Correlando el análisis del tráfico TCP con información radio, hemos detectado que el problema de rendimiento relacionado con SACK ocurre justo después de handovers, poniendo de manifiesto que el operador 2 puede no estar reenviando los datos pendientes entre las celdas involucradas en el handover.

Asímismo, con la ayuda de información GPS sincronizada, hemos identificado tres principales tipologías de áreas donde ocurren problemas de conectividad. La primera es en las transiciones de entorno urbano a rural, donde se encuentran celdas de distinto perfil, la segunda es en zonas rurales entre pequeños pueblos donde se detecta un nivel de intensidad de señal (RSSI) reducido, y por último hay una zona muy problemática cuando el tren reduce de 250 a 200 km/h y transita a la celda que da servicio en el largo túnel que atraviesa los montes de Málaga.

VIII. Un entorno de prueba para experimentación realista

En el Capítulo 7 se describe un entorno de prueba para experimentación que ha sido compuesto para verificar el rendimiento de aplicaciones internet sobre LTE de forma realista pero controlada.

El entorno de prueba se ha particularizado en esta ocasión para el estudio del servicio de VoIP, ya que la migración de los usuarios de servicios antiguos, basados en conmutación de circuitos, hacia tecnologías como LTE, que trabajan únicamente en el dominio de paquetes, requerirá de una comprensión adecuada de la calidad de servicio percibida por los usuarios.

Los beneficios claves de la solución propuesta incluyen la posibilidad de ejercitar un amplio rango de configuraciones de parámetros desde el nivel de radiofrecuencia hasta el nivel IP. La optimización conjunta de estos parámetros permitirá encontrar configuraciones que satisfagan los requerimientos de calidad de servicio.

VIII.I. Relación con soluciones previas

Los enfoques aplicados típicamente en la literatura involucran el uso de simuladores de red como ns-2 u Opnet, o incluso Matlab cuando se trata de analizar parámetros de nivel físico. Con frecuencia no es posible comparar los resultados de las distintas contribuciones por varios motivos. En primer lugar porque se utilizan modelos en lugar de implementaciones y dispositivos reales. Y en segundo porque los detalles de bajo nivel y los modelos en sí no se publican para ser contrastados.

Aunque el uso de modelos sencillos permite la escalabilidad del análisis cuando se trata de sistemas con muchos elementos, a la hora de hacer un análisis detallado la utilización de modelos puede resultar excesivamente simplificada y por tanto poner en cuestión la exactitud de los resultados obtenidos.

Por otra parte, para la realización de pruebas de campo para medir la calidad de servicio, es una opción utilizar herramientas comerciales (y muy costosas) como QXDM de Qualcomm para monitorizar parámetros LTE de capas bajas. Sin embargo, los resultados obtenidos en pruebas de campo no son reproducibles debido a la naturaleza variable de las condiciones de propagación radio (propagación multicamino, interferencia, desvanecimientos…). Asimismo hay normalmente una ausencia total de control sobre la carga de la red, su configuración y las políticas de asignación de recursos.

IX. Entorno de referencia para experimentación

Para dar evitar las limitaciones anteriormente descritas, nuestro objetivo ha sido crear un entorno de referencia para pruebas en el que poder realizar medidas de calidad de servicio. Así pues, para obtener resultados fiables y comparables, la repetibilidad será un punto clave, ya que además muchos comportamientos en comunicaciones tienen cierta componente aleatoria y requieren una base muestral lo suficientemente amplia para obtener representatividad estadística. Por otra parte se requiere de precisión en el ajuste de parámetros radio como la potencia absoluta y la relación señal a ruido, ya que tienen una implicación directa en la variación de la tasa de paquetes perdidos y otros parámetros clave en la calidad de servicio.

Para permitir que los resultados puedan ser reproducidos y verificados por otros experimentadores, una decisión clave ha sido el uso de teléfonos móviles comerciales. Al mismo tiempo se ha buscado que el entorno de prueba permita un grado de configurabilidad y control muy elevado. El tiempo es otro recurso muy valioso a tener en cuenta, ya que la solución buscada debe permitir llevar a cabo un gran número experimentos. Por último, para la evaluación de la calidad de voz se ha buscado la utilización de métodos objetivos, ya que la evaluación subjetiva es un proceso muy costoso.

Como resultado, hemos propuesto una arquitectura de prueba basada en emuladores de red de tiempo real, aplicaciones VoIP abiertas y algoritmos estándar de estimación de calidad de voz. La Figura 6.1 muestra la estructura funcional del sistema. Las llamadas de voz se han realizado utilizando el cliente Yate de telefonía software SIP y el servidor de VoIP Asterisk. El cliente usa una conexión a internet establecida mediante un modem USB-B3730, equipo terminal LTE de Samsung.

Estos dispositivos se comportan de la misma forma en que lo harían en una red real, pero de forma que sus puertos de antena están conectados mediante cables a un emulador de eNodeB E2010 de AT4 wireless (ahora T2010 de Keysight Technologies).La estructuración en capas del entorno de experimentación puede ser observada también en la Figura 6.2.

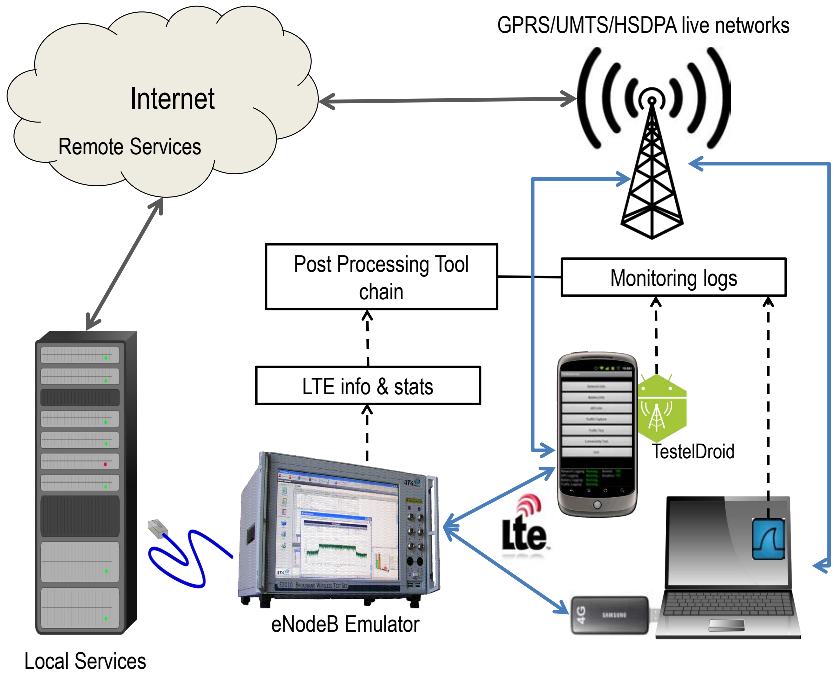

Figure 7: UMA testing facility setup

X. Obteniendo medidas de calidad de sevicio más realistas: Aplicación al estudio de VoIP sobre LTE

En el Capítulo 8 se proporcionan más detalles sobre la aplicación práctica del setup descrito en el capítulo anterior, demostrando su potencial de utilización en el estudio de VoIP sobre LTE al ser esta la línea argumental de la presente tesis. En dicho proceso, queda de manifiesto la extrema versatilidad del instrumento y

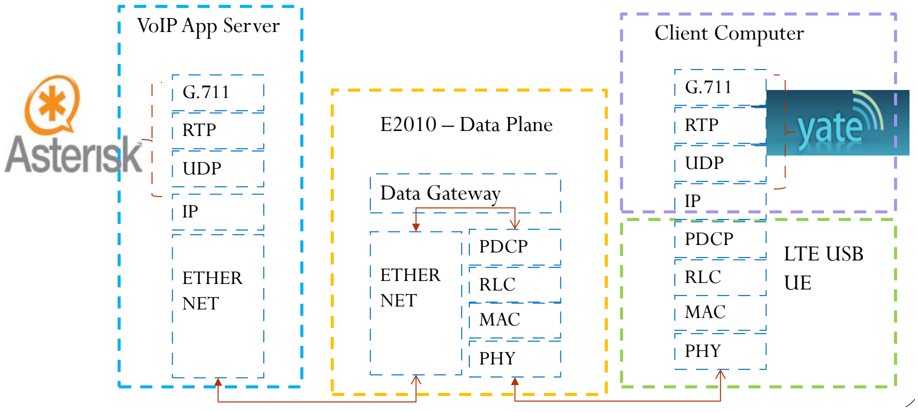

Figure 8: VoIP protocol stack

de la metodología usada, que podrá ser extendida al estudio de muchos otros campos y aplicaciones.

El T2010 es una plataforma genérica, con aplicación no solamente en la prueba y certificación de radiofrecuencia y señalización, sino también en la verificación de los diseños de fabricantes de terminales móviles. Las principales funciones que ofrecen son:

- Soporte físico para dos conectores de RF por terminal móvil
  - Control preciso de la potencia transmitida
  - Control de la frecuencia transmitida cubriendo las diferentes bandas comerciales
- Emulación conforme a la especificación de la señalización de capas RRC (Radio Resource Control) y NAS (Non-Access Stratum).
  - Configuración de parámetros de capa PDCP
  - Configuración de parámetros de capa RLC
  - Configuración de parámetros de capa PHY
- Control sobre el planificador
  - Tamaño de los bloques de transporte (TBS)
  - Esquema de modulación y codificación (MSC)
  - Asignaciones de recursos en tiempo y frecuencia
  - Control de la configuración HARQ (Hybrid Automatic-Repeat-Request)
- Funciones de medida de la calidad del receptor móvil
  - Medidas de tasa de error recibida a nivel físico y de enlace
  - Medidas estadísticas de éxito en retransmisiones
  - Análisis estadístico del reporte de calidad y rango del canal (CQI / RI)

- Emulación digital del canal radio
  - Configuración MIMO hasta 4x2
  - Utilización de perfiles estándar de desvanecimiento
  - Control sobre la correlación entre antenas transmisoras y receptoras
  - Generación de ruido blanco

Utilizando el entorno de propuesto, hemos realizado una campaña de pruebas con el objeto de verificar la consistencia de los resultados obtenidos. Como resultado hemos podido comprobar cualitativamente cómo, a pesar de la variabilidad estadística asociada a la emulación de canal, las medidas realizadas guardan una elevada correlación con las variables de entrada en el sentido esperado.

XI. Conclusiones y trabajos futuros

En el ámbito de esta tesis, hemos realizado una serie de contribuciones en dos campos principales. Por una parte hemos desarrollado una aplicación móvil para monitorización de comunicaciones de datos, y la hemos aplicado en el estudio del servicio de VoIP en redes comerciales, incluyendo casos de uso particularmente desafiantes como el de los ferrocarriles de alta velocidad. Por otra parte, hemos creado un entorno de referencia controlado para experimentación de servicios sobre tecnología 4G LTE, verificando asimismo su aplicación sobre VoIP.

En el futuro, seguiremos explotando el potencial de las herramientas de monitoriación desarrolladas. Extendiendo nuestras campañas de pruebas a la evaluación comparativa de las redes comerciales LTE y de tecnologías futuras a medida que vayan desplegándose. De particular interés para esta tesis hubiera sido la verificación del servicio comercial de VoLTE, ya que su despliegue había sido recientemente anunciado por Vodafone. Desafortunadamente, en el momento de la redacción de este documento, este servicio no está accesible a los usuarios en general por lo que ha sido imposible la realización de las pruebas correspondientes. También resultará de interés la exploración del rendimiento de servicios Over The Top (OTT), como los proporcionados por Skype, Whatapp y otros.

En el terreno de la experimentación en entorno controlado, también buscamos extender la experimentación a casos de uso adicionales y nuevos servicios. Asimismo, hemos analizado el potencial de aplicación de una nueva generación de emuladores de red. En concreto, el UXM wireless test set de Keysight Technologies, representa una evolución significativa sobre el equipo anteriormente usado en esta tesis, tanto en términos de disponibilidad de nuevas funcionalidades, como de mayor rendimiento. La utilización de dicho equipo debería permitirnos investigar escenarios adicionales, como redes heterogéneas con celdas de distinto tamaño, diferentes procedimientos de movilidad como handovers e incluso pruebas de máxima velocidad de transmisión hasta aprox. 600 Mbps.

Abstract

Voice has been the primary source of revenue for mobile operators during a long time, and even with data traffic being gaining focus it will not dissapear with the transition to ip based networks. Major players on the mobile industry has clearly stated that their customer base will not accept a decrease in the quality of their voice calls, thus it is critical to warrant the user experience in voice services deployed over new generation packet based networks.

Because of its importance, the work on this thesis aims to analyze the behaviour and dependencies of the VoIP communications over mobile networks. To that end, we have developed and applied in this field a mobile communications monitoring application that runs on Android smartphones, and a methodology to derive meaninful metrics and useful representations from the collected information.

To complement the measurements in live networks, as there is no control on the radio conditions and network configuration, we have composed a reference testbed for 4G LTE experimentation. This test environment will be the foundation for realistic and repeatable experiments with different network configurations, that could be controlled with high flexibility and accuracy because of the use of network emulators with digital embedded channel emulation.

We have validated the developed tools, methodologies and test environments in the measurement of VoIP performance, but in future works they could be directly applied to a wider range of target services and application scenarios.

Chapter 1
Introduction

1.1 Introduction

The characterization of traffic performance in cellular networks from the point of view of final users is a costly process which involves a large range of protocols and parameters with complex dependencies. To deal with this challenge cross-layer measurements and correlated analysis are required. The characterization methodology proposed in this thesis provides the possibility to collect key information for troubleshooting IP communications problems, correlating them with propagation issues, such as cell changes or link outages, and resource allocation problems at specific geographical locations. The methodology is based on a monitoring and recording tool for smartphones and a correlation tool chain which can be used in live networks and on an experimental testbed to carry out extensive experiment in a controlled and repeatable way. With the results obtained from the deployment of this methodology, mobile operators, services providers and mobile developers could gain access to real user experiences and specific users' data to improve radio coverage, adapt services and customize mobile applications and protocols based on IP to cope with mobility issues in cellular networks.

The main contributions of the tools and methods introduced in this thesis are the following:

- A cross-layer monitoring tool for Android smartphones, called TestelDroid, which enables capturing key performance counters at the user equipment. It also provides active traffic and connectivity tests.

- An automatic post-processing methodology for the cross-layer correlation of the collected measurements. Moreover, users can access traffic captures and radio measurements and use their own analysis methodologies to obtain new metrics.

- Validation use cases based on the performance evaluation of IP traffic on-board high-speed trains.

- A realistic and highly configurable testbed to carried out advance LTE experimentation.

1.2 Performance study of Internet traffic over mobile networks

According to Cisco [1], global mobile data traffic grew 70 percent in 2012, reaching nearly twelve times the size of the entire global Internet in 2000. In addition, mobile video traffic exceeded 50 percent for the first time, showing the increasing importance of multimedia traffic. By the end of 2013, the number of mobile-connected devices exceeded the number of people on earth. In this context, mobile subscribers demand high levels of quality everywhere. However, it is not unusual to find performance and connectivity problems that hugely impact on the user experience.

To maintain adequate levels of quality operators, content providers and developers need appropriate tools to monitor the performance of the services. Network operators usually have access to different levels of information, known as Key Performance Indicators (KPI), regarding the performance of their networks. KPIs are mainly based on counter values collected at network elements. These performance counters are defined by 3GPP technical specifications (TS 32.405, TS 32.450, etc) and implemented at network elements such as RNC, MSC, SGSN or e-NodeB. However, these counters usually do not reflect the suscriber's experience because highly sophisticated filtering and correlations functions are not implemented in the static functions of network elements [3].

Content providers and application developers traditionally have to test their apps in a user environment without any knowledge of network information. Typically, communication monitoring tools for evaluating the user experience required dedicated hardware and were only available to network operators because of their prohibitive cost. However, it has recently become possible to use current smartphones themselves as monitoring devices with the help of software applications.

1.3 The need of a realistic testbed for repeatable testing

The enhancement of QoS in a sustainable manner is a critical goal for network operators as management tasks are becoming increasingly complex. Although some initial efforts have been carried out by the standardization bodies there is still a significant gap to be covered in QoS and also in QoE optimization. Actually, current efforts towards improving QoS and QoE are typically based on estimations derived from costly drive test campaigns. Furthermore, involvement of human expertise is required to manually tune network configurations. On the other hand, most of the service and network configurations available in the literature are derived from simulations [4] [5] [6] [7] [8] [9]. As is widely known, in the process of modeling communication systems to simulate them, some details may be missed and thus misleading results may be derived. For example, it is very common to find that the consumption of control resources is ignored when evaluating different scheduling methods. However, control elements and data are both actually competing for time-frequency resources, and the allocation of control resources to multiplex users in PDCCH is not a trivial task. In this context, providing optimized network configurations based on measurements obtained directly from the subscribers' terminals and correlated with the information collected at the network will pave the way for a reduction of costs and more accurate tuning of network operation from the point of view of the QoS perceived by final users. Moreover, as stated by standards organizations (SDOs) or alliances with the participation of network operators such as NGMN [10], *the optimization of QoS still requires "real" developments to further study the direction in which to move forward*.

In this thesis we also propose the use of an experimental testbed [11] to carry out specific LTE (Long Term Evolution) experiences in a real context and to extract the correlation between LTE radio configurations and QoS parameters perceived at the application level. The execution of exhaustive measurements campaigns using this testbed will enable the identification of specific performance counters, correlations between them and use cases for QoS and QoE optimization in LTE networks.

The focus of the thesis is on VoIP calls over LTE, which pose new challenges over previous technologies. In LTE voice calls are now delivered through an all-IP network (VoIP) instead of a circuit switched one, which means that voice has to compete for bandwidth with other services provided in the network. It is vital to at least guarantee the same QoE for VoIP calls that was available in pre-LTE technologies such as GSM (Global System for Mobile Communications) and UMTS (Universal Mobile Telecommunications System). This will be required to avoid significant impact on customers, who will demand a good service in all-IP mobile networks. Due to situations like this, the testbed has been conceived to validate the performance of the network configurations and problems presented in the research literature. Specifically for VoIP service we have correlated layer 1 and layer 2 LTE radio parameters with IP performance parameters and MOS (Mean Opinion Score) measurements based on the PESQ (Perceptual Evaluation of Speech Quality) algorithm.

In order to collect IP performance parameters and MOS measurements two new developments have been undertaken as part of this thesis:

- A software tool for Android devices which provides advanced monitoring functionalities such as the logging of radio parameters and the capture of IP traffic.

- A post-processing tools chain focuses on the processes of capture and analysis of key performance indicators (packet losses, jitter, inter-packet delay, PESQ, etc.) in real time services, such a VoIP, which lead to the achievement of a complete characterization of their behavior over cellular networks.

The organization of the thesis is as follows. Chapter 1 introduces the necessity of obtaining performance measurements which capture the QoS and QoE as perceived by final users and the new challenges presented in the provision of voice call services in LTE. This chapter provides also a brief state of the art on VoIP over LTE and some noteworthy LTE configurations proposed in the literature to optimize its performance. In chapter 2, additional information on the VoIP state of the art are provided. Chapter 3 introduces TestelDroid, a monitoring tool for Android devices and the post-processing tool chain. Chapter 4 is focused on the validation of the tools and method introduced in the previous chapter. This chapter provides the results obtained during the testing of VoIP services

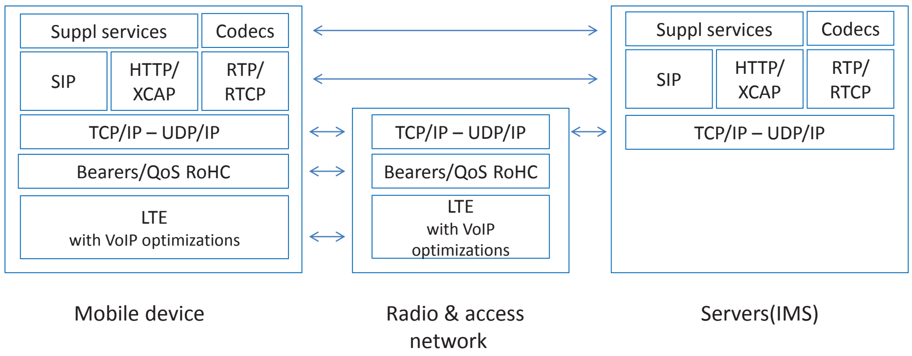

Figure 1.1: Scope of VoIP deployment over LTE by GSMA PRD IR.92 v6.0

in live networks, in static and vehicular scenarios. A special analysis of VoIP performance in high speed scenarios is described in Chapter 5. Chapter 6 presents the proposed testbed to carry out advanced and realistic experimentation in a controlled environment. Chapter 7 provides the results of the tests carried out to validate the functionalities of the testbed proposed in Chapter 6. The configuration under which the measurements were collected, the results obtained during the analysis of IP performance parameters and their correlation with LTE parameters are also included. Finally in Chapter 8 we present conclusions and future work.

1.4 Voice calls over LTE: a regular data service, the same quality as before

3GPP (3rd Generation Partnership Project) has standardized two solutions for the deployment of voice call service over LTE. The first is CSFB (circuit-switched fallback) [20], which implies a shift of the UE (User Equipment) access from LTE to 2G/3G during a voice call. The second is VoLTE [21] which on the contrary is based on IMS (IP Multimedia Subsystem) and does not require the use of legacy technologies. Another alternative available in the market is VoLGA (Voice over LTE via Generic Access)[22]. That specification has been developed by the VoLGA Forum, based on the existing 3GPP Generic Access Network (GAN) standard [23]. CSFB and VoLGA provide interim solutions for early LTE deployments while VoLTE offers a long term opportunity for mobile operators. VoLTE allows integrating voice and internet services and delivering new multimedia services in a permanent environment. This will enable the exploitation of the potential offered by mature LTE networks. In this context VoLTE has emerged as the preferred solution by carriers and the GSMA (GSM Association) is developing a specification for delivering integrated telephony services over LTE.

Specifically, the GSMA defines in [21] the minimum mandatory set of features that a wireless device and a network should implement to support a high quality IMS-based telephony service over LTE radio access. The scope provided by GSMA is shown in Figure 1.1.

Standardization bodies are confident about the necessity of introducing specific LTE configurations for the deployment of VoIP service. In [21] GSMA proposes a list of LTE configurations which we aim to verify and extend if possible with new optimizations in future works using the proposed testbed.

At the early stages of this thesis, as stated in the 3GPP initiative, the MultiService Forum had

already demonstrated successful VoLTE calls, and also MMTel (Multi Media Telephony) services [24]. During these tests, equipment from 19 manufacturers was used. The tests performed focused on validating the interoperability between the interfaces defined in the 3GPP Technical Specifications.

Network operators have also performed testings experiments for QoS (Quality of Service) measuring. E.g. different performance metrics such as latency or throughput are evaluated for the TeliaSonera network in [25]. However, our work aims to go a step beyond and not only measure the performance of interfaces and terminals using individual metrics but also correlate all these measurements with specific LTE parameters in order to identify optimum configurations.

Nowadays, VoLTE is been deployed in many contries, but despite having being publicly announced by some operators, it is not yet available to all the users.

1.4.1 Some considerations about the transport of VoIP over LTE

As proposed by GSMA in [21] SIP (Session Initiation Protocol) is the protocol used to register UEs in the IMS server. RTP (Real-time Transport Protocol) and UDP (User Datagram Protocol) are the protocols recommended to voice transport, and RTCP (RTP Control Protocol) to provide link aliveness information while the media are on hold.

The most restrictive performance indicators for interactive real-time services such as VoIP are the end-to-end delay and jitter. The maximum allowed one-way delay for voice service is 300 ms as stated in [26], with a recommended value lower than 150 ms. The low latency of LTE access (20-30 ms) reduce the end-to-end delay obtained in previous cellular technologies. However, LTE radio bearers cannot guarantee fixed delay. Instead, fast retransmissions are used to repair erroneous transmissions, and uplink and downlink transmissions are controlled by schedulers. Consequently, LTE transmissions introduce jitter, what implies UEs must implement efficient de-jitter buffers. The minimum performance requirements for jitter buffer management of voice media are described in [27].

The objective of radio resource management (RRM) procedures in LTE is to ensure an efficient use of the resources [28]. RRM algorithms at the eNodeB involve functionalities from Layer 1 to Layer 3. Admission control mechanisms, QoS management and semi-persistent scheduling are deployed at layer 3, while Hybrid Adaptive Repeat and reQuest (HARQ) management, dynamic scheduling and link adaptation are in layer 2 and CQI (Channel Quality Indicator) manager and power control in layer 1.

3GPP specifies RRM signaling but the actual RRM algorithms are not provided [28] in order not to constraint implementations and thus also allow operators to differentiate in this field. The combination of radio bearers that a UE must support for Voice over IMS profile are defined in [29] Annex B. Concretely, the voice traffic requires a guaranteed bit rate (GBR) bearer, as described in [30]. The network resources associated with the EPS (Evolved Packet System) bearer supporting GBR must be permanently allocated by admission control function in the eNodeB at bearer establishment. Reports from UE, including buffer status and measurements of UEs radio environment, must be required to enable the scheduling of the GBR as described in [31]. In uplink it is the UEs responsibility to comply with GBR requirements.

In the following subsections we will analyze separately those parameters and configurations which have been identified in the state of the art as optimized LTE solutions for VoIP.

1.4.1.1 Quality class indicator

The characteristics of the bearers are signalized with a QoS class identifier (QCI). The QCI is a pointer to a more detailed set of QoS attributes, including layer 2 packet delay budget, packet loss rate and scheduling priority. As defined in [32] QCI 1 is intended for conversational voice.

1.4.1.2 RLC mode configuration

As specified in [21] the unacknowledged mode (UM) should be configured at RLC (Radio Link Control) layer for EPS bearers with QCI 1 to reduce traffic and latency.

1.4.1.3 DRX mode

Support of LTE discontinuous reception (DRX) methods for both UE and network are mandatory to reduce power consumption on mobile devices. The idea behind DRX methods is that the terminal pauses the monitorization of control channels during some periods of time, allowing it to turn the radio off. DRX parameters can be tuned depending on RRC (Radio Resource Control) status or service. Decisions about when the radio should be activated again can be based on QoS indicators. The simulations conducted in [4] give, for VoIP applications, a potential saving of about 60 percent.

1.4.1.4 Compression

In order to optimize radio resources the UE and the network must support Robust Header Compression (RoHC) to minimize the size of IP packets during VoIP calls [21]. As we have already said the use of UM at RLC and reduced sequence number sizes also decrease overhead. The reduction of packet size will enable the improvement of coding efficiency which is especially important for uplink scenarios and to improve the quality of data connection in areas with poor coverage.

1.4.1.5 Semi-persistent scheduling

The mechanisms involved in packet scheduling at the eNode-B are three: dynamic packet scheduling, link adaptation and Hybrid Adaptive Repeat and Request management (H-ARQ). Semi-persistent scheduling significantly reduces control channel overhead for applications that require persistent radio resource allocations such as VoIP. There are many different ways in which semi-persistent scheduling can configure persistent allocations.

Simulated results obtained in [5] show that in uplink direction semi-persistent scheduling can support higher capacity than dynamic scheduling while at the same time guaranteeing VoIP QoS requirement, but with the cost of sacrificing some statistical multiplexing gains from HARQ. As dynamic scheduling is already needed for other services, they suggest using dynamic scheduling by default for VoIP. As an exception, semi-persistent scheduling should be applied for some VoIP users only in situations where the signaling load becomes too high.

1.4.1.6 Admission control

Admission control mechanisms determine whether a new EPS (Evolve Packet System) bearer request should be admitted by checking that QoS requirements of at least all the bearers with high priority are fulfilled. Specific rules and algorithms for admission control are not specified by 3GPP. An interesting technique is proposed by AT&T Labs Research in [6] based on an intelligent blocking algorithm (IBA) for the admission process of VoIP calls subject to individual customers' blocking objectives and which is only invoked when the total bandwidth in use is close to its engineering limit. The algorithm has also been validated with simulations.

1.4.1.7 Packet bundling in Downlink

Traffic generated by voice codecs can be very bursty due to the nature of the service, i.e. silences in the conversations, and the nature of the codecs, e.g. traffic from the AMR (Adaptive Multi-Rate) codec. The combination between VoIP traffic patterns with dynamic packet scheduling can lead to inefficient resource utilization that might be improved by the use of packet bundling in the downlink. Several simulation studies have been done for both UTRAN (Universal Terrestrial Radio Access Network) [7] and E-UTRAN (Evolved Universal Terrestrial Radio Access) [8] probing that dynamic packet bundling approaches improve VoIP services.

1.4.1.8 RLC segmentation and TTI bundling in Uplink

UEs have limited transmission power and at the edge of LTE cells there is a high probability of obtaining an error in the transmission of VoIP packets because the device is not able to gather enough energy during one TTI (Transmission Time Interval) (1ms) to send the packet. In the event of an unsuccessful transmission, HARQ retransmissions are required, which will imply the introduction of 8 ms delay per retransmissions. As consecuence, a very large number of retransmissions would involve an intolerable increase in the delay for a conversational voice service and reduction of the transmission efficiency.

The conventional approach used to reduce delays and improve the coverage at the cell edge is RLC segmentation. It consists of the segmentation of RLC SDUs (Service Data Unit) and their transmission in consecutive TTIs. However this is not an optimum solution from the point of view of the overhead introduced in the control signaling and the increase of the vulnerability of packet loss due to HARQ feedback errors. The idea behind TTI bundling is that a given transport block is transmitted a fixed number of consecutive TTIs without waiting for the HARQ feedback. The eNodeB sends the corresponding HARQ feedback only when it has received the whole bundle of transmissions. This approach reduces the amount of HARQ feedback significantly [9]. The L2 header overhead is also reduced because there is lower need for segmentation, as well as the signaling overhead required for uplink grants (i.e. resource allocations) because a single grant is required for each TTI bundle.

Chapter 2

Survey on VoIP over mobile networks

2.1 Background

There has been extensive research about VoIP, and a number of very good surveys have been conducted. As the purpose of the present work is to provide the reader with updated information focused on VoIP over mobile networks, the user is encouraged to access those for a more detailed introduction to general aspects of VoIP.

In [33], in addition to more general aspects of VoIP, there is a section dedicated to the perceived speech quality over emerging mobile and pervasive networks. Although the analysis is focused on bluetooth, wimax and mesh networks, there are interesting references to the Mobisense project. Some reference works are [34], that discusses about alternative transport and [35], that provides information about VoIP codecs in a WLAN environment.

2.2 Survey on Skype

In [36], a complete analysis of the Skype traffic is performed including active and extensive passive measurements, traffic identification, use pattern characterization and discussing NAT and firewall related issues. Different codecs are studied both for audio and video traffic using TCP and UDP traffic as transport protocols forcing different configuration options. The scenario where the measurements has been carried out is a campus LAN. The work also deals with the study of the adaptation capabilities of the Skype traffic to changing network conditions. No influence is observed from the use of TCP in the bandwidth, it is not clear if there should be a clear impact.

The paper provides also a traffic pattern characterization which includes useful statistical data on call durations for experiments. One interesting outcome is that about 80% of the traffic is actually originated by Skype control signalling. As the scenario is not related to a mobile environment no statement is made about the effects of such signalling amount on power consumption. However, as it is based on LAN, the study does not cover the impact of mobility management (handovers/cell reselections) or other impairments associated to a cellular technology.

In [37] it is introduced a methodology to measure the quality of the voice transmission using Skype. The work is based on the use of analog audio signal using microphone and earphone connections as the contents of the Skype traffic are not available. As no signal reconstruction is possible from traffic sniffing, the degradation of the analog audio signal is measured using the PESQ methodology. As no statistical information is provided, it cannot be concluded if the cuantitative results are representative, e.g. some throughput measurements in the uplink are higher than for the downlink but no qualitative explanation is provided.

In [38] the quality of the voice service using Skype is analyzed in both an UMTS environment and an emulated bottleneck LAN network setup to reproduce controlled changing conditions. The PESQ metric is used to evaluate the quality. The concept of Network Utility Function is introduced to analyze the dependency of the quality on the network conditions, using packet losses and delay parameters as input. As PESQ is based on the comparison of audio signals, the signals are generated with Winamp and feeded to Skype with no microphone to avoid the associated degradation.However, at the receiving side the audio is recorded in a different PC using a cable connected to the microphone input. Emulated channel rates varies between 16 Kbps and 384 Kbps, whereas live traffic is assigned 384Kbps in the downlink and 64kbps in the uplink, which were typical values for UMTS Rel99 but have been superseded by current mobile technologies. ADSL rates used in the experiments may be now also obsoletes. The measured quality is normalized to compensate the effect of the Skype source codec (by generating and receiving in the same machine) and the degradation associated to the recording microphone+cable connection.

In the buffer scenario, data are buffered during 500 ms and sent at the maximum LAN rate, this seems not to be a realistic operation mode. It is also mentioned that Skype provides a ”technical info” window that shows information about the packet loss ratio at the remote end. In UMTS connections, Skype seemed to choose a different codec based on local information. Nearly no packet loss was observed during the UMTS experiments (0.1%), probably that indicates that the connections were originated in a static scenario with no mobility nor handovers. The paper includes useful information on the cumulative distribution function for packet interarrival time and on its evolution over the time, at sender and receiver. Packet interarrival time is quantized at 1 ms steps at sender rate, the authors state that it could be originated by the PCMCIA driver.

In [39] is analyzed the performance of Skype on UMTS. It is stated that the network control is shifted to the network edges (with traffic control at application side) and that it impacts Quality of Experience. Similar work seems to be introduced in [40]. At the time of writing, the authors identified as future work items the analysis of video and the interaction with fixed telephony with SkypeIN and SkypeOUT. They indicate that ITU recommends a delay lower than 150 ms for VoIP call, as indicated in ITU-T Recommendation G.114 for one way transmission time [87]. It is referenced a model to derive the perceived quality from network input parameters as defined in ITU-T Recommendation G.107, The E-model [88]. A description of the model can be found in [41]. They used the samples available at Signalogic, Speech Codec- Wav Samples, `http://www.signalogic.com/index.pl?page=codec_samples`. They also relied on Winamp to reproduce the wav files and windump to capture the traffic. The authors also identify the codec being used by Skype observing detailed information provided by Skype itself.

The former Global IP Sound codecs (ISAC and ILBC) are no longer aplicable for Skype to Skype calls from version 3.2. For SkypeOUT/SkypeIN, the G.711 and G.729 codecs are applicable. Newer codecs such as SVOPC and the current SILK are not considered. SVOPC is reviewed in [36] and SILK in [42].

2.3 Surveys on VoIP

In [43] is is introduced a survey on VoIP. This paper introduces several methods used in the assessment of voice quality: Percertual Speech Quality Measure (PSQM), Perceptual Analysis Measurement System (PAMS), Perceptual Evaluation of Speech Quality (PESQ), E-Model and P.563. Also voice codecs use for the compression of audio signals and header compression techniques are described. The three main performance indicators that characterize the quality of voice communications over internet are analyzed in detail: delay, jitter and packet loss. One way delay increases the echo impairment effect and affects the quality as described in ITU G.114. The ETSI TYPHON project recommends a end-to-end dealy lower than 100 ms for toll quality. It is recommended to prioritize voice packets to prevent delays.

Human ear is highly intolerant of short-term audio gaps, therefore jitter should be kept to a minimum. Ideally it should be less than 30 ms however, 30 and 75 ms can be acceptable as well as per [44]. The mechanism used to compensate the jitter is a buffer. The jitter buffers may have a static size or variable according to the evolution of the jitter. There must be a tradeoff between the delay included by the play-out buffers and the packets being dropped.

Packet loss ratio is another key metric. More than 99 percent of packets are required for VoIP toll quality and 97 percent for business quality [45][44]. Reasons for lossing packets are buffer overflow in network nodes and routers, discarded at destination because of excess of delay and fading and other wireless propagation impairments (if non reliable operation is being used). For an acceptable voice quality VoIP service can tolerate bit error rates (BER) up to 10e-5 but then its quality is drastically reduced. A mapping of packet losses to speech quality is provided in [46] and [47]. In [47] they conclude that voice quality is acceptable below 2 percent of packet loss.

To compensate packet losses Forward Error Correction techniques are proposed. These techniques help to decrease packet losses but increase the delay and the required throughput. There are also other mechanisms based on the use of codecs that replicate packets and replace silences. [48] provides a detailed survey on packet loss recovery techniques in general.

2.3.1 Quality Evaluation

Providing the expected voice quality is a critical point in VoIP services, as typically voice has not been the primary design goal in data networks. To verify the quality of voice services it is required to use reference methods, which may be divided in subjective and objective methods. Subjective methods are based on averaging evaluations carried out by groups of users when listening references speech signals. This process is called Mean Opinion Score (MOS) [49] and is defined in an ITU recommendation. The outcome of the MOS is a value between 1 (worst quality) and 5 (best quality). When evaluating signal quality, it can be considered either absolute (ACR) or relative (RCR) category ratings. Relative ratings is based on comparing a degraded signal to a reference signal. Subjective tests involve a slow and costly process, and the results may be affected by impairments in the environment between different experiments and by the actual user base. As consequence it is difficult to achieve repeatability in these evaluations. Objective methods relay on analyzing network parameters such as one way delay, jitter, packet losses or signal parameters compared to a reference source As objective tests are suitable to be reproduced, they are often used to analyze the impact of variations in different input parameters. Different types of objective tests may be used either using reference speech patterns as signal source or obtaining information from live communications, thus categorized in intrusive or non-intrusive respectively. Non-intrusive methods allow for larger evaluations as they do not require specific test environments, but are generally more complex and their results are less reliable than intrusive methods. Intrusive tests usually use both male and female speech patterns as they usually provide different signal characteristics such as frequency an may be differently impacted by network transmission.

2.4 Recommendations for VoIP transmissions

In [50] it is recommended to consider the effect of IP communications using the E-Model, but it is focused on narrow band signals and does not fully consider the impairments originated by wireless and satellite technologies neither all the combined effects. [50] provides information about the subjective quality appreciated for different codecs with different percent of packet losses. Particularly, different packet loss concealment options of G.711 (Namely Annex A or B) show a different behaviour. The bursty packet losses scenario, with Annex B shows a deeper decrease in the subjective quality with losses higher than 2%. Note that both the type of losses and the concealment method vary.

A unique feature of the E-Model is its flexibility to deal with the impairments introduced by speech compression and packet loss via the Equipment Impairment Factor (Ie). To optimize the delay, the G-711 codec is recommended because it has the lowest Ie-value and allows additional delay for a defined quality. The recommendation is to use G.711 unless the link speed demands compression.

As per [51], to ensure interoperability, in line with 3GPP, the baseline profile mandates support of the adaptive multirate (AMR) narrowband speech codec. However, seen in terms of an MMTel business solution, the recommended speech codec for voice over LTE access is AMR wideband (high-definition voice). This codec is mandated in 3GPP networks and terminals that support 16KHz speech sampling. It is also stated that the frame loss rate is very low because of the MAC layer HARQ retransmissions (not present in 2G and 3G voice services) but introduce some delay variations (jitter) that need to be handled with appropriate de-jittering buffers. Even with the effect of jitter, the delay is on par or better that the obtained in 2G or 3G. The effect of delay on signal quality is low in LTE as the end-to-end delay is in the order of 20-30 ms.

Also, in LTE, in addition to header compression (RoHC), "short packet data convergence protocol (PDCP)/RLC sequence numbers"' are used to reduce message overhead and allow for transmission formats with higher redundancy, thus reducing the error ration and increasing the coverage. Unacknowledge RLC mode is used for voice sessions as a certain level of packet losses is acceptable (up to 1 percent is typically considered acceptable).

Chapter 3
Monitoring tools for Android devices

3.1 Smartphones as monitoring tools

Smartphones are becoming a major platform for the execution of Internet services as more powerful and less expensive devices are becoming available. The verification of the performance of such services has also moved from the traditional PC plus modem setups to smartphone built-in scenarios. Testing configurations based on computer based tools are not enough anymore for the verification of Internet services on mobile phones. They are giving the way to measurement tools specifically tailored for smartphones, enabling performance analysis of services and applications designed for mobile devices. In this transition, it becomes critical to develop powerful tools providing native measuring capabilities for mobile based data services. Thus, a wide range of monitoring functionalities are needed to allow the fusion of information from many different sources and view points, obtaining a complete profiling of mobile applications and services.

Following this approach the SymPA [12] tool was developed and successfully applied in the study of video streaming service in cellular networks [13]. With its successor TestelDroid we pretend to increase the monitoring functionalities with the features provided by Android and thus extend our research into Android services, applications and devices.

As commented before, the proliferation of mobile Internet services for smartphones makes it more necessary than ever to use measurement tools in mobile devices to monitor their connectivity performance. In this chapter we introduce TestelDroid, a software tool for Android devices which provides advanced monitoring functionalities for analyzing multimedia services and applications specifically designed for smartphones.

The potential beneficiaries of this tool are mobile operators, service providers and mobile developers, as TestelDroid offers the foundation to apply a unified and independent methodology to analyze performance parameters from the application level, taking into account underlying issues.

Mobile operators require scalable monitoring solutions because network complexity and mobile subscribers use patterns are increasingly challenging. Providing smartphones with advanced measurement capabilities paves the way for distributed and scalable monitoring systems as the dependency on dedicated expensive hardware is avoided. This scheme could complement and enhance network self optimization procedures extending the observation points to additional protocol layers and to the user device and applications perspective.

The first tools designed to operate in mobile devices, such as Qualipoc [94], were centered on the observation of service accessibility parameters such as availability, provided bandwidth, error rate, etc. To that end, they included active client features ranging from FTP (File Transport Protocol) to voice calls and message sending. These clients were used to test the services and verify general parameters. However, the rising complexity of mobile networks and applications requires test solutions to go a step further both in application and radio performance monitoring. By doing so, it will be possible to identify the actual sources of communication issues and to improve the user experience, as subscribers do not care whether communication problems are related to application, protocol or radio aspects, since they all result in experience degradation. Measuring solutions aiming to deal with many perspectives must therefore correlate information from all the communication layers.

Although active service probing is useful for some purposes, it is clearly unfeasible to include clients for every single application and service available, present or future. Passive monitoring represents a scalable alternative for general purpose analysis of mobile applications' performance. Different aspects may be studied jointly using the passive approach including communication performance, memory usage and battery consumption among others.

In this chapter we will introduce the usability of TestelDroid in the detection of VoIP traffic behaviors associated to cellular environments. In section II we will introduce the features provided by TestelDroid for traffic capture, cellular network monitoring, GPS measurement localization and power consumption. Some noteworthy implementation details are also included. Section III presents its applicability in the analysis of the performance of multimedia services in live cellular networks providing a qualitative analysis of VoIP quality over GPRS (General Packet Radio Service), UMTS (Universal Mobile Telecommunications System) and HSPA (High Speed Packet Access) technologies. Finally, section IV provides a summary of the main contributions of this work.

3.2 Related work on monitoring tools for smartphones in cellular network

Mobile phones have traditionally been used in telecommunications consulting to diagnose under performance behaviors in mobile network deployments. However, tools such as TEMS Pocket [52], NEMO Handy [53] or Qualipoc [54] are engineering tools which run on special smartphones. These tools provide very powerful features, being specifically designed for network operators. A wide variety of measurements and representations can be generated. They can also be integrated with non handset based solutions to create larger coordinated test environments. However, as evidenced in the associated documentation, they are restricted to being executed in specific device models. This is because they require modified ROM and kernel images to access the proprietary information provided by the manufacturer's chipset. As a general rule, the cost of these tools is high and typically reserved for network operators. These hardware limitations, and their high cost, pose a major drawback for a wider adoption by application developers and users other than network operators. In comparison with these tools TestelDroid is not able to access chipset data, however the information retrieved from the operating system is still very useful to carry out cross-layer analysis of the IP traffic vs radio performance.

Numerous applications are also available on the Android market, such as Netmonitor, GSM Signal Monitoring, Field Test-SignalSitemap, G-Mon, CellMapper and Network Signal Info, which provide monitoring capabilities and run on commercial smartphones. Such applications monitor radio parameters, store cell information in databases and generate maps with the location of the cells. However none of them provide the functionality of traffic capturing nor the advance post-processing methodology to correlate radio measurements with IP traffic issues. In contrast, TestelDroid requires root access to execute some of its functionalities.

Finally, research oriented applications, such as 3GTest [55], LiveLab [56] or PowerTutor [57], also incorporate sophisticated performance evaluation techniques. In [55] monitoring functionalities are limited to a fixed number of services and a set of well-defined metrics to evaluate network and service performance. The main contribution of our work compared to 3GTest is that our tool and method is independent of the service. All the IP traffic is captured so that any application can be tested using the capture to obtain the metrics corresponding with the service. Our methodology also provides post-processed results which make it possible to obtain conclusions about the service and the network under test; however, the metrics can be easily extended by the researchers because traffic captures and radio parameters logging files are accessible in a standard format. This also enables researchers to use their own tools to process the results. Furthermore, some of the measurements provided by 3GTest require the use of an external server.

In [56] the authors introduce an iPhone OS tool the logging functionalities of which are very similar to those provided by TestelDroid. However, the focus of the tool is the characterization of smartphone application usage (for example, to find which applications are most used and how often they are accessed). They also apply a user-collaborative approach to characterize networks, but only limited information is provided in this respect. Our work goes deeper on the characterization of low level traffic related issues and their correlation to radio metrics.

The PowerTutor tool [57] is oriented to measuring power consumption by monitoring major system components such as CPU, network interface, display, GPS and applications. This tool provides a very detailed power model taking into account all these components, while TestelDroid and the method introduced in this chapter focus on the monitoring of power consumption and the impact of cellular connections on battery life.

3.3 TestelDroid, on-device monitoring tool

TestelDroid [58] is a cross-layer monitoring tool which uses the engineering features provided by current commercial Android smartphones to profile the communication performance of mobile applications running on mobile devices. With TestelDroid smartphones become actual measurement probes that gather information closely related to final users. This proximity grants the methodology

Table 3.1: Features provided by TestelDroid

Network	Operator RAT (Radio Access Technology) CID (Cell Identification) LAC (Location Area Code) RSSI (Radio Signal Strength Indicator) PSC (Primary Scrambling Code)
Neighboring cells	PSC RSSI RSCP (Received Signal Code Power) RAT (Radio Access Technology)
Battery	Battery level Temperature (°C) Voltage (mV) Current (mA)
GPS	Longitude Latitude Altitude Speed
IP traffic	Pcap format Arrival timestamps Promiscuous mode
Connectivity test	Ping Open ports
Active traffic test	Server-Client mobile-to-mobile Transfer of auto-generated file Bit rate monitoring Average transfer speed

the potential to characterize the performance of the mobile applications from the point of view of subscribers. Moreover, user application level analysis, which includes KPI such as application specific download speed, upload speed, latency, jitter as defined in [59] and packet loss, requires access to performance counters which are only available at the user equipment.

Four different sets of parameters are obtained: IP traffic, GPS coordinates, battery consumption, and network related information. You can find a full description of the monitored parameters in Table 3.1.

All the collected data can be logged using highly analyzable plain text files (except for traffic capture, stored on *pcap* format). Logging is implemented as an Android service, thus it can be running on background while performing other actions on the phone. The parameters to be logged (network, neighbor cells, battery, GPS, traffic) are configurable under the preferences menu, as shown in Figure 3.1. The tool has been tested on Nexus One (HTC), Nexus S (Samsung) and Galaxy S (Samsung) devices with Android versions ranging from 2.2 (Froyo) to 2.3.4 (Gingerbread). During the development of the tool we noticed that PSC (for the current cell), current consumption and neighboring cell information was missing on Samsung phones.

The information retrieved by TestelDroid can be logged and exported for further analysis with our post-processing tools and other tools, such as Wireshark, or integrated with other applications such as Google Earth, as shown in Figure 3.2.

Other features provided by TestelDroid are active traffic and connectivity tests. Active traffic test functionality is oriented to testing a mobile to mobile TCP connection. One mobile is configured as the server and the other side as

Figure 3.1: TestelDroid screenshots, from left to right: Preferences configuration, GPS measurement localization and Network information

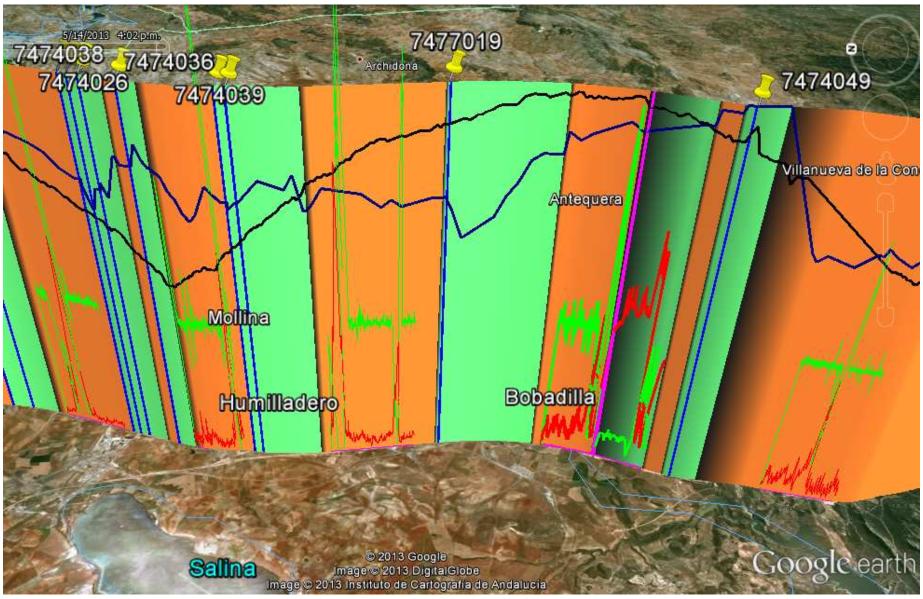

Figure 3.2: Geographical representation of the measurements collected by TestelDroid

the client. An auto generated file with the size specified by the user is sent while speed connection is monitored and the mean throughput is calculated. Connectivity test functionalities provide two diagnostic modes. The first one is based on the traditional ping functionality. The user can configure the number of pings, packet size and timeouts. The second mode enables checking the establishment of mobile originated TCP connections, given an IP address and a TCP port, which is useful for detecting reachability problems.

3.3.1 Some implementations details

TestelDroid retrieves information through the standard APIs (Application Programming Interfaces) provided by the Android OS, Android runtime and Linux kernel. The Android API provides access to network features such as cell identifier, available neighbors and signal strength, battery data and GPS location. The Android runtime is required to implement TCP connectivity and traffic tests. Finally, to implement packet capturing, ping and battery consumption, it is necessary to access the Linux kernel itself, which implies root access.

The information provided by TestelDroid in these three different levels is shown in more detail in the Figure 3.3. At the application framework level (blue area), Android API provides information such as current cell related information, available neighbors or GPS location. An event oriented approach is used to notify of changes such as signal strength variations, handovers, or GPS location updates. Battery data is also provided at this level except for current consumption, which will have to be retrieved from the Linux file system. Google Maps API is used as well to have a visual representation based on GPS data. Android runtime level (green area) provides part of the connectivity test (check whether a socket can be opened at a specified host/port) and traffic test (transfer of a file on a client/server model) features. Finally, some tasks are accomplished directly at the Linux kernel itself (red area), such as packet capturing, part of the connectivity test (ping) and current consumption information for the battery.

We have profiled TestelDroid performance, monitoring VoIP calls on a Nexus One mobile device over an 8 hours period, in terms of CPU utilization, memory usage and energy consumption. Mean CPU utilization is less than 1%. Memory usage is 22 MB, which is a normal value for an Android process. Finally, the energy consumption is only the 0.2% of the total power during a measurement session. These results prove that TestelDroid presents a good performance for background monitoring during long periods.

3.4 On-desktop post-processing tools and methodology

We have developed an automatic post-processing methodology to correlate the data collected by TestelDroid. The tool provides multiple outputs which in turn provide a cross-layer characterization of the performance of IP communications in cellular networks.

In particular, two different representations correlating the magnitudes under analysis have been developed to identify the source of potential communication issues: time based and geographic views. In the former, the evolution over time of cross-layer parameters is presented. This view is useful to accurately study the

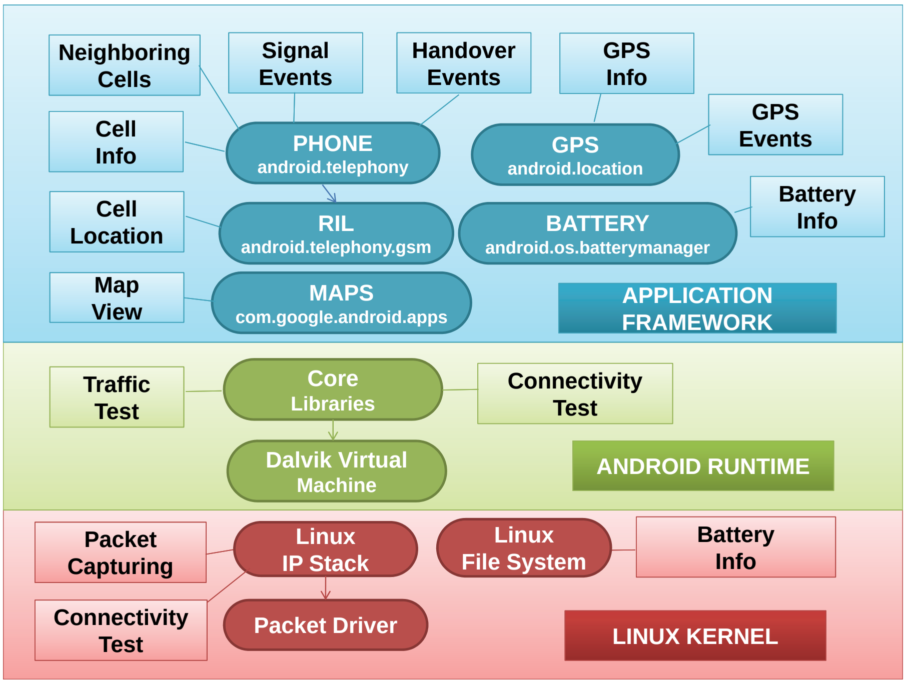

Figure 3.3: TestelDroid diagram

timing of specific events, such as handover duration and its impact at IP level (packet losses, connection interruption, etc). We then see that time representations provide lots of meaningful information for network characterization. In addition, the location based representation is of great help to put measurements in their geographical context. This is a key factor in understanding the root of communication issues. By being aware of the measurement locations, it is straightforward to then trace service interruptions back to critical coverage areas, to name just one of its many potential applications. More details can be found in [15].

VoIP is the service under test in the proposed methodology. Audio test signals recommended in ITU-T P.501 [86] for telephonometry measurements have been used during the experiments carried out to validate this methodology.

This methodology introduces the use of TestelDroid and also a complete set of processing scripts which will provide a numerical and graphical representation of the measurements collected.Data files collected by TestelDroid are the starting point of the analysis methodology depicted in Figure 3.4. The IP traffic is captured and stored in libpcap format. Traffic files are filtered using TShark, an utility included in the distribution of the protocol analyzer Wireshark, to detect the SSRC (Synchronization source) identifiers of the RTP (Real-time transport protocol) flows contained in the IP traffic capture. SSRC identifier is a randomly chosen value meant to be globally unique within a particular RTP session. Using the SSRCs the RTP packets are separated in individuals files containing isolated RTP flows. Each flow is independently analyzed to extract information about the temporal evolution of key performance parameters: inter-packet delay,

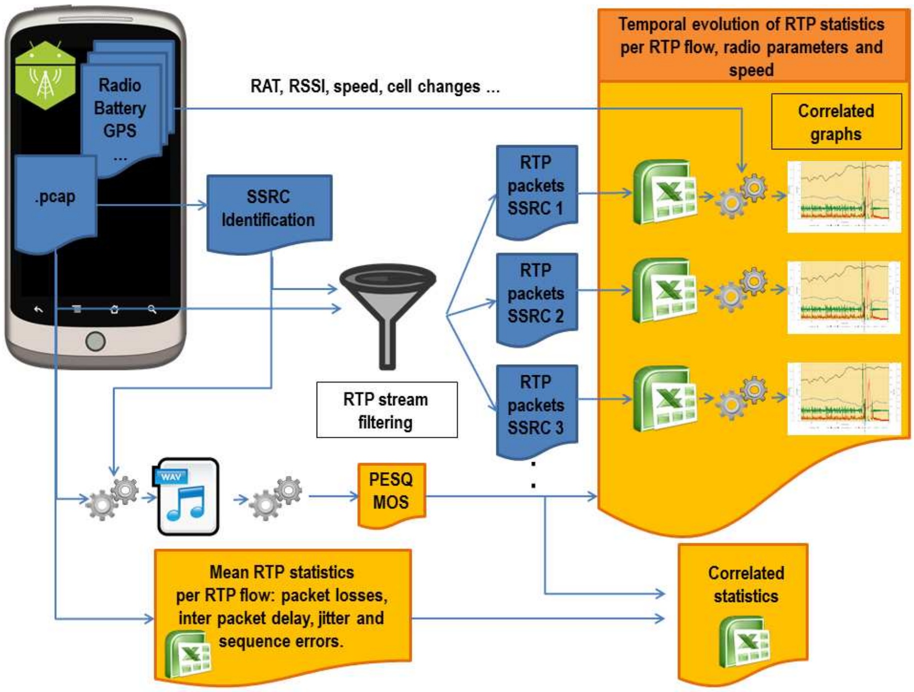

Figure 3.4: The post-processing methodology developed as part of this thesis includes the inspections of radio access parameters, QoS parameters and QoE values

jitter, packet losses and sequence errors. This information is correlated with the information capture at radio level in order to analyze the impact of mobility issues, such as fading, cell reselections and handovers, over IP traffic.

PESQ (Perceptual Evaluation of Speech) [92] is an objective algorithm which compares a reference signal to a degraded signal and provides a MOS (Mean Opinion Score) value. In our methodology the PESQ algorithm is applied to the degraded signal reconstructed on the receiver side to obtain the end-to-end quality of VoIP service. The VoIP speech is recovered extracting the payload from RTP packets captured at IP level, in order to avoid the side effects on the PESQ algorithm introduced by the techniques applied at application level, such as packet loss concealment and adaptive jitter buffer [93]. The MOS values obtained are also correlated with the rest of parameters calculated: mean jitter, packet losses, inter-packet delay, mean RSSI, etc.

3.4.1 RTP flow processing chain

In order to process the RTP (Real-time Transport Protocol) flows associated with each session, the sequence below has been followed:

1. Identify the list of RTP flows. Each flow has associated a randomly chosen 32 bit synchronization source (SSRC) field intended to provide a globally unique identifier during an RTP session.

2. Generation of IP statistics summary

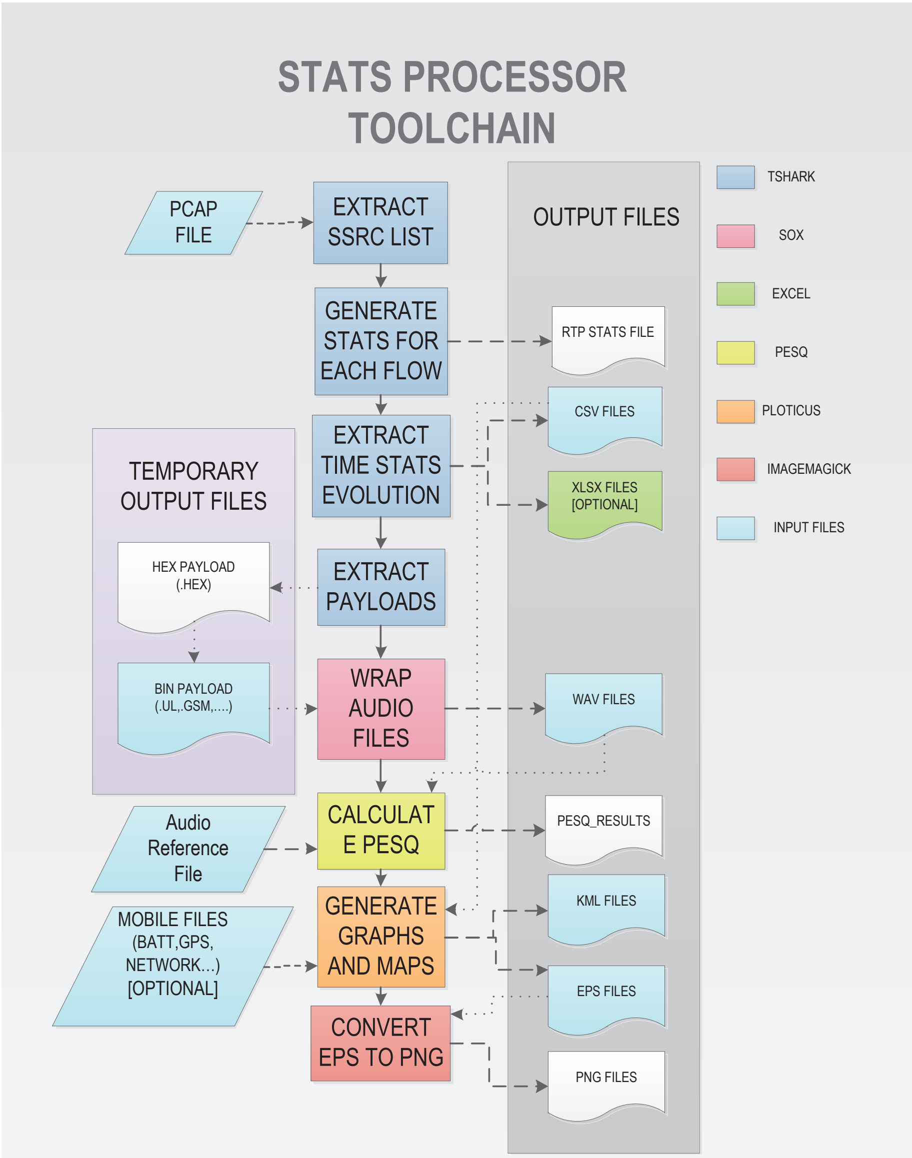

Figure 3.5: Data processing chain

  - One file is generated containing average results for all the flows in the captured pcap file.

3. Extraction of the temporal evolution of the flows. For each flow two files are generated

   - A .csv file containing the changes related to every IP statistic.
   - An equivalent .xlsx file including also a graphical representation.

4. The payload transported by each flow is extracted and converted to the appropriate binary format.

5. The binary contents are then wrapped in a .wav file.

   - The resulting .wav files thus contain the audio reconstructed from the IP traffic received at the mobile device.

6. Voice call quality is estimated using the PESQ algorithm for each flow, using as input the generated wav files and the source reference waveform.

   - As a result PESQ MOS results are generated for each flow

7. Vectorial graphics are generated using the files containing the temporal evolution of the IP traffic and radio parameters.

   - The generated .eps files contain both network information and IP parameters calculated during processing.

8. The resulting graphs are converted to .PNG format. Although the quality is slightly decreased, this format allows a faster visualization

9. In addition, advanced tridimensional representations are created using GPS positioning if available for the measurement sessions. This will be further explained later in Section 5.3.3.

Figure 3.5 contains a diagram of the complete processing chain.

3.5 Conclusion

We presented TestelDroid, a software tool which runs on Android phones. It offers a high configurability of the monitoring functionalities provided. The next chapter provides multiple measurements in different scenarios, carried out to characterize the performance of VoIP in HSPA and inter-RAT with legacy technologies. A correlation of all the monitored parameters is provided to prove the potential of the tool in troubleshooting tasks.

TestelDroid will be available through our web site: http:// www.lcc.uma.es/∼pedro/mobile and in the Android market.

The fundamental novelty of the testing methodology introduced in this thesis is the ability to correlate key performance indicators obtained at the Android application level with radio access parameters and speech quality. Mobile operators, developers, content providers and researchers from the academia can take advantage of the methodology presented in this chapter to evaluate the performance of cellular networks, mobile applications and new mobile protocols.

Chapter 4

Live testing for Characterization of Traffic Performance in Cellular Networks

4.1 Case study: VoIP over HSPA and inter-RAT mobility

This case study focuses on the evaluation of VoIP performance over HSPA at the application level, although during the experiments the RAT was eventually changed to UMTS or GPRS, depending on the coverage. There is a great deal of related work in the literature which covers this topic, but usually it is based on simulations and focused on air interface capabilities, MAC-hs functionalities such as fast packet scheduling, link adaptation and HARQ (Hybrid Automatic Repeat Request) [95], and RRC (Radio Resource Control) performance such as discontinuous reception (DRX) cycles [96]. These works exhaustively study packet scheduling and resource allocation, and obtain relevant results from the point of view of network management. However the performance of the results obtained at the application level and from the point of view of individual users has not been proved. Parameters such as IP bit rate, inter-packet delay, jitter, IP packet losses and sequence errors have not been evaluated in referenced works. Although these parameters have been the subject of study in [97] and [98], such studies are oriented to prove the performance of scheduling algorithms but by means of simulations.

Studies carried out in [99] and [100] show a close similarity with the work introduced in this chapter, with the singularity that we provide an open methodology which enables studying the performance of any application taking into account radio propagation particularities. They perform test campaigns on live HSDPA (High Speed Downlink Packet Access) networks and study jitter, packet loss and end-to-end delay. Both works prove their validity to obtain relevant statistics from live testing in cellular networks, but it is worth noting, at this point, that a remarkable contribution of our thesis is the possibility of evaluating the impact of radio network management procedures over mobile

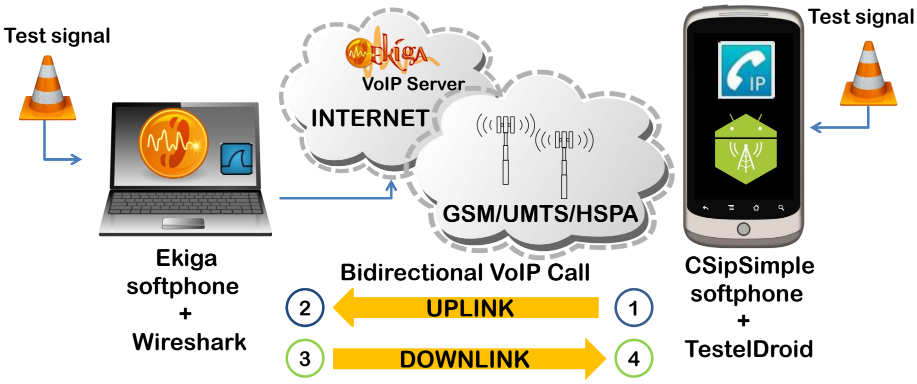

Figure 4.1: VoIP experimental scenario over live cellular networks

application and protocols performance by the simple inspection of the results provided by TestelDroid. In the following sections we will prove the soundness and the complementary functionalities offered by our tool for supporting field measurements on Android mobile phones.

4.1.1 Field trials setup

In order to verify TestelDroid in a mobile deployment, the tool was used to measure VoIP performance over live cellular networks. Both static and vehicular scenarios were covered. For the test we used a commercial VoIP service provider, Ekiga.net, which also offers a PC softphone client application. Figure 4.1 shows the experiment setup, where two VoIP clients were configured to establish a bidirectional communication. Data collected were post-processed and most relevant results are depicted in figures in the following sections.

TestelDroid was deployed on a Nexus One mobile phone running in parallel with CSipSimple, a VoIP client for Android. The other side of the connection was hosted by a PC running the Ekiga softphone. Both the CSipSimple and the softphone were configured to use Ekiga VoIP service. The Ekiga softphone auto answering functionality was enabled to allow automated VoIP call tests. The PC was placed in the campus network of the University of Malaga and thus connected to Internet using high speed data access. All the IP traffic exchanged during the experiment was captured using Wireshark in the computer end and TestelDroid on the mobile side. Four measurement points allow them to measure downlink and uplink traffic at both ends, as indicated in Figure 4.1. To provide a reference source for the experiment, a standard audio sample was injected in both devices. The sample was extracted from the artificial voices described in the recommendation ITU-T P.501, which specifies test signals which are applicable for several purposes in telephonometry. Using reference sample signals allows the analysis of the impact of impairments observed in IP-based networks, and will enable us to provide objective speech quality metrics in future experiments.

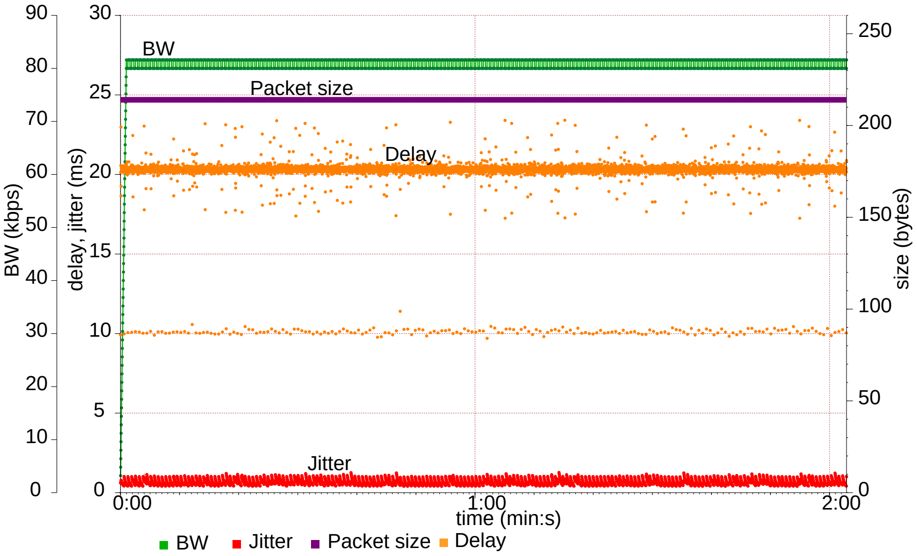

Figure 4.2: VoIP traffic transmitted by Ekiga softphone during a conversation with a mobile connected through a HSDPA connection (Downlink source side)

4.1.2 Static characterization of VoIP traffic

In a first step we characterized the traffic generated by both VoIP clients in a static scenario, before running the measurement campaigns in the vehicular scenario. Traffic inspection allowed us to determine the codec negotiated for both sides of the conversation, which in this case was G.711. We have observed that there are no changes in the codec used during the call once it is established. In the remaining experiments we have forced the codec to stay the same to ensure the results are suitable for comparison.

The size of the packets sent by the Ekiga client installed in the PC is 214 bytes, with a time period between audio packets of 20 ms with a small dispersion, and a fraction of the packets at 10 ms intervals. We can appreciate this in Figure 4.2, where inter-packet delay is depicted with orange points. This dispersion introduced a small jitter at the source (red line). Transmission rate is represented by the green line, which has an average value of 80 kbps, as we could expect because it is close to the rate obtained with 214 bytes packets at 20 ms. This is one of the standardized bit rates for G.711 codec. During the experiments the traffic bandwidth is calculated by accumulating the length of the packets received during the last second.

Figure 4.3 shows the parameters which define the VoIP stream transmitted by CSipSimple in HSDPA. TestelDroid facilitated the analysis of RTP (Real-time Transport Protocol) traffic emitted, providing access to useful information not normally available on mobile devices. In this analysis the inter-packet delay was observed to be restricted to a discrete number of values. These values differ from those observed at the PC softphone: 0, 15 and 20 and 35 ms, and present higher dispersion than previously had been observed. This dispersion introduced a jitter of 5 ms seconds at the source that must be taken into account in the analysis of the traffic on the receiver side.

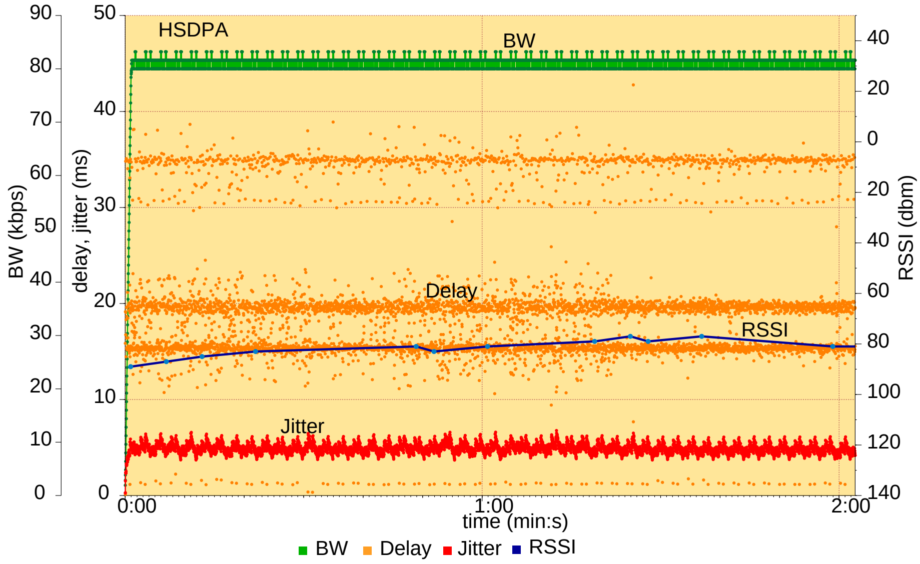

Figure 4.3: VoIP traffic generated by CSipSimple softphone during a HSDPA connection (Uplink source side)

The packet size and the transmission bit rate average are 214 bytes and 80 kbps respectively, which are equivalent values to those from the Ekiga client. Figure 4.3 also depicts RSSI (blue line) and RAT in use (orange background color indicates HSDPA). Adding information not only from the IP level, but also from the radio connection will enable us to explain some traffic patterns. As we can see in the following figures, radio parameters have a great impact on traffic parameters and without this information users would miss the actual source of the behavior observed.

In order to validate the results obtained from the information collected by TestelDroid we have compared our mean jitter results with values obtained in previous work. Averaging the results obtained during the static measurements we have obtained a mean jitter of 6 ms for HSDPA and 13 for HSUPA (High Speed Uplink Packet Access), which are in the same range than the results obtained in [99] and [100]. In the vehicular scenario we obtained a mean jitter of 23 ms but results show a high dispersion due to packet loss introduced by cell changes and fading. This mean jitter is also coherent with those provided in [99].

When the available bandwidth is reduced in GPRS, we have been able to observe that the signal quality is dramatically impacted because the codec requires higher transmission rates. Thus, in GPRS connections we have seen that the bandwidth may be reduced to 35 kbps (see Figure 4.5) whereas the analyzed VoIP clients do not react to this variation in the communication channel capacity, as per Figure 4.4. In consequence, we can observe in Figure 4.6 that the encoded waveform sent by the Ekiga client at the PC cannot be received at real time and it is expanded because of the increasing delays. The figure contains three waveforms: the original transmitted signal (bottom), the reconstructed waveform using timestamps to compensate the transmission delays (middle) and the actual received signal (top) which can be clearly identified as expanded

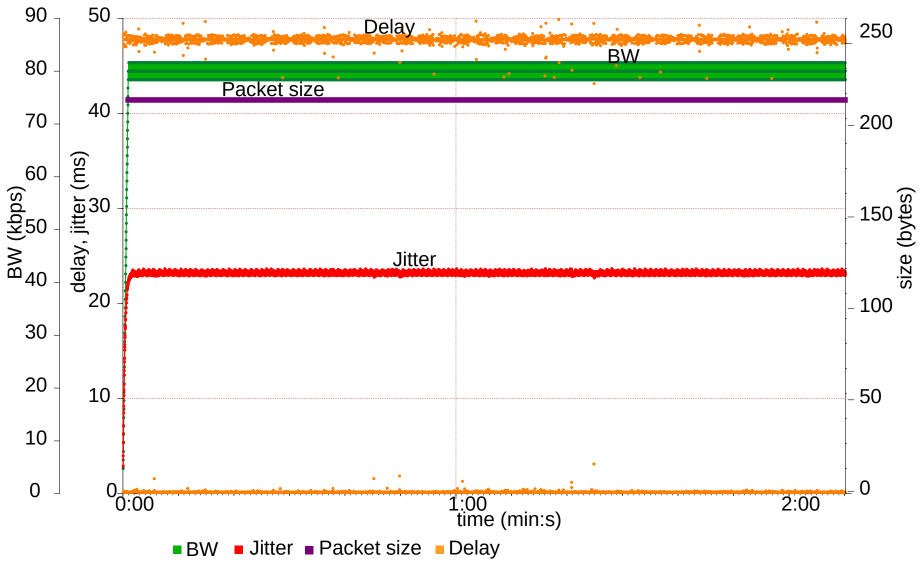

Figure 4.4: VoIP traffic transmitted by Ekiga softphone during a conversation with a mobile connected through a GPRS connection (Downlink source side)

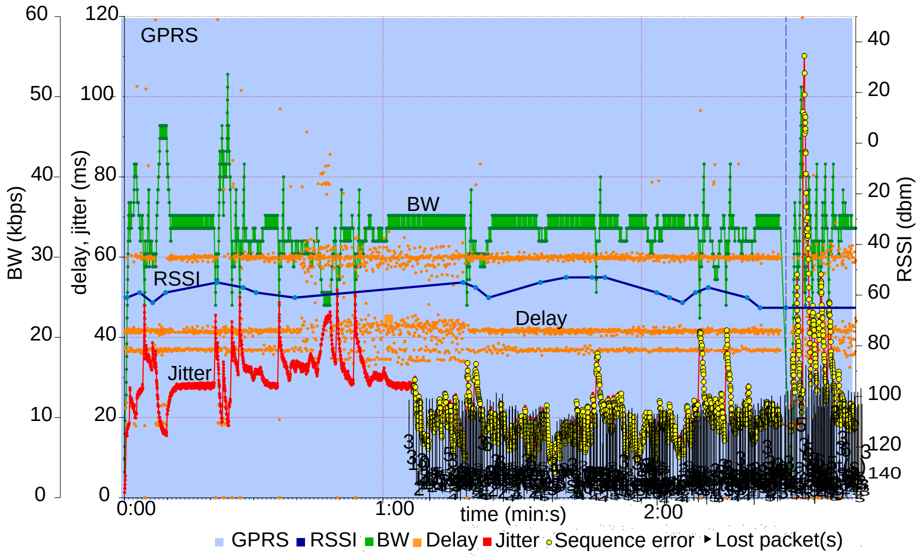

Figure 4.5: VoIP traffic received by CSipSimple during a GPRS connection (Downlink receiver side)

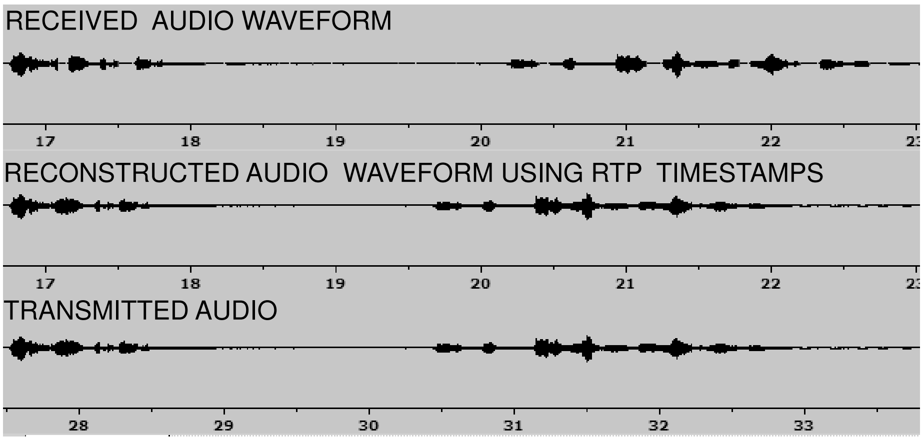

Figure 4.6: Transmitted and received waveforms comparison in HSDPA

because it accumulates more than 500 ms delay increments in just 5 seconds. However, because of the difference in transmission and reception rates (80 kbps and 35 kbps respectively), the network can not afford to buffer the downlink traffic indefinitely. After 80 seconds of reception in the downlink, 40 seconds of delay is accumulated which indicates that the buffering has reached its limit (more than 3 Mbits is being performed by the network) and lots of packets are dropped (marked with black arrows in Figure 4.5) interleaved with received packets leading to a drop rate of nearly 50% on average. If no buffering had been performed by the network, the drop rate would have increased to 56%=(80-35)/80, but the huge delay would have been avoided. It must be noted that as VoIP is a very interactive service, large delays are not acceptable. Although no additional figures are included, in the analysis of the uplink scenario associated to the downlink session in Figure 4.5, a similar behavior has been observed but with sequences of bursts of more than 100 lost packets.

It clearly shows that applications must negotiate quality of service parameters tailored to their needs and that the network must commercially provide connection profiles adapted to real time services.

4.1.3 Vehicular characterization of VoIP traffic

One of the observations we can make on the basis of the vehicular results is the impact of HSDPA and UMTS over inter-packet delay. The behavior observed during the tests is depicted in Figure 4.7. Inter-packet delay presents a clear difference in its behavior if we compare UMTS intervals (green background) and HSDPA intervals (orange backgroud).

During a vehicular test where data connection changed several times from UMTS to HSDPA we can appreciate how values acquired show a different behavior. In the case of HSDPA six values can be discerned ranging from 0 to 50 ms in steps of 10 ms. A lower inter-packet delay granularity was expected in HSDPA, as its TTI (Transmission Interval Time) duration is only 2 ms, but a similar result was also obtained in [100] during jitter calculation. When compared to the results in [100] the instantaneous jitter is not directly comparable because we calculate it as an average of the delay (following the formula in RTP RFC)

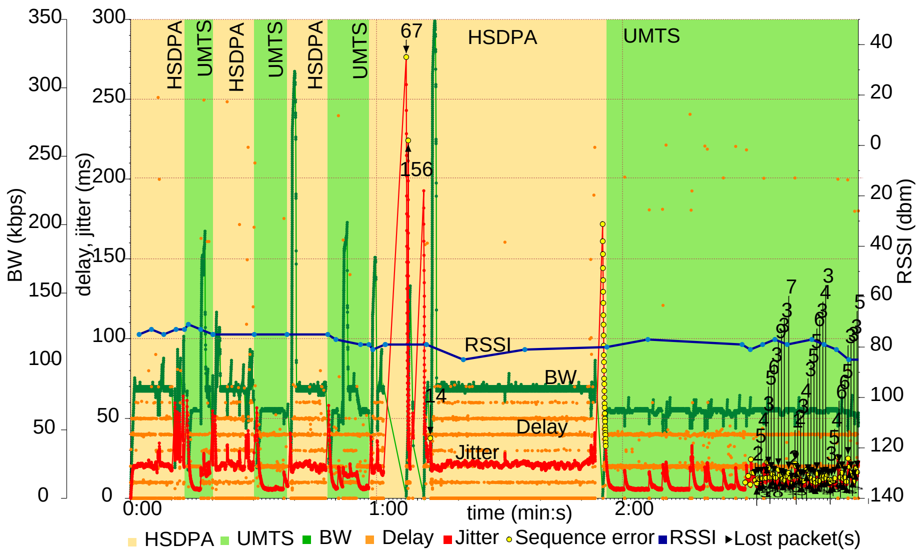

Figure 4.7: Radio access technology impact over inter-packet delay (Downlink receiver side)

which results in a low pass filter of the signal that hides the high frequency fluctuations. In our study case we can appreciate the same behavior for the inter-packet delay. That paper pointed out that it was related to the Iub flow control and the management of data buffers used between nodeB and RNC.

Regarding the packet losses and the end of session shown in Figure 4.7, we cannot conclude that UMTS is not suitable for VoIP because it could reach speeds of up to 384 kbps. However, by inspecting the information provided by TestelDroid we have seen that only a 64 kbps resource allocation has been dedicated to the connection by the network. In consequence a similar issue as the one introduced in subsection III.B appears when network buffer overflows. Real time interaction is also heavily impacted upon because of the high buffering delays, that reach up to 11 seconds and finally result in packet losses.

4.2 Conclusion

This chapter has introduced tools and methods to measure the performance of data communications in cellular networks. The methodology has focused on the cross-layer monitoring of the performance counters available at the user equipment. These counters are beyond the scope of probes deployed by the operators in their own cellular networks and enable measuring the quality of service as perceived by final users. To that end, we have developed TestelDroid, a monitoring tool for Android commercial smartphones. In addition, cross-layer methodology has been developed to automatically correlate the information collected.

After the analysis of the results from the measurement campaigns, we can conclude that to improve user experience it is necessary to analyze not only the behavior of services and applications or radio technologies separately but their combination as well, since their interaction may originate performance degradations. The use of tools to correlate IP traffic with key radio parameters could help to verify the quality levels in new technologies such as LTE, where a voice call service is provided as an all-IP service. Tools and methods introduced in this chapter have been used in the project Tecrail, which aims to evaluate the viability of the migration from GSM-R to LTE in railway environments (see http://tecrail.lcc.uma.es).

We also observed an interesting dependence between inter-packet delay behavior and radio access technology in use.

Chapter 5

Performance study of Internet traffic on high speed railways

5.1 Introduction

Railway transportation uses specific technologies such as GSM-R for traffic signaling purposes, but commercial networks could also provide added value services to passengers and transport companies such as CCTV or telemetry. Furthermore mobile users increasingly want to be always connected even when they are traveling. Although current deployed technologies such as 3G provide wide coverage in general, when it comes to high mobility scenarios such as the railway, further analysis is required to ensure appropriate user experience. In this chapter we propose a methodology to collect and to organize traffic information related to the behavior of Internet services over commercial mobile networks (UMTS and HSDPA). We apply our method to conduct an extensive study on a high speed train line in the south of Spain. The measurement campaigns comprised hundreds of VoIP sessions, comparing different network providers over more than 155 kilometers. We conclude that the characterization method is useful to provide information for railway operators (for instance, to migrate from GSM-R to LTE), telecom operators (to fit their network deployments) and for mobile software developers (to adapt their applications to the high speed environment).

GSM-R is the railway extension of the GSM telephony standard designed specifically to satisfy railway radio communication requirements such as group calls, call priority assignment or broadcast calls. Specific features called Advanced Speech Call Items (ASCI) were standardized and introduced inside the GSM documentation to cope with these new requirements. GSM-R is used for the transport of ETCS (European Train Control System) signaling for train control applications (data) and train radio (voice). Because GPRS is not part of the GSM-R specifications, control and signaling information are carried via Circuit Switch Data (CSD) as the bearer service. GSM-R has a dedicated frequency band (876-880/921-925 MHz) of only 4 MHz, which means that at present it is only possible to have 19 channels. Due to this limited capacity the GSM-R Industry Group started working on the adoption of new technologies such as

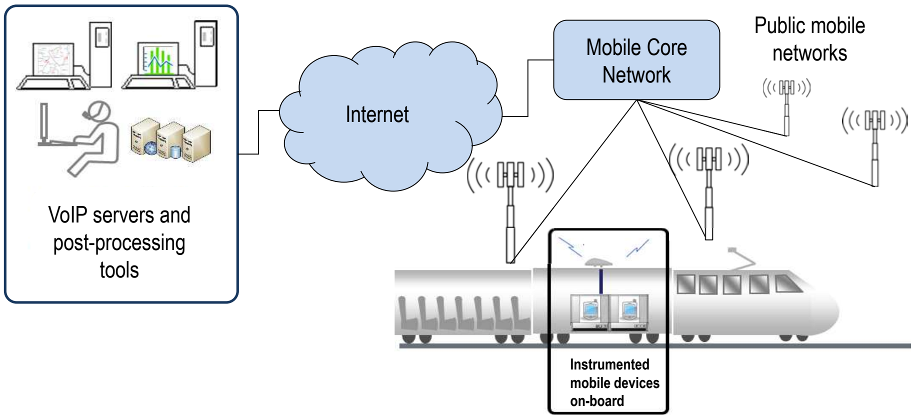

Figure 5.1: On-board monitoring with instrumented mobile devices

GPRS and more recently LTE/SAE for a Future Railway Mobile Radio System (FRMRS). Preliminary studies indicate that LTE will offer approximately 20 times the performance of past 2G solutions, and can even co-exist with GSM-R over similar frequency bands.

The limitations of GSM-R have already been recognized because of the need to transmit increasing amounts of data. Nevertheless, railway industry commitment to support and develop GSM-R until 2025 and the interest of rail owners in recovering their investments in GSM-R equipment make the migration towards new communication technologies difficult. In the mean time, real-time services such as access to huge amounts of on-board data telemetry or video-surveillance for train monitoring are emerging in the railway operations domain [60][61]. However it is not possible to provide them using current GSM-R deployments. Also providing their railway passengers with Wi-Fi or other short range wireless systems, as well as mobile phone-based services represents a unique opportunity for rail operators [62][63] to enhance their service portfolio. Potential services could use passengers' own devices for on-line entertainment or stream content to flat screens available on-board. In this scenario the link between the passenger devices and the base station is provided by other on-board wireless deployments, such as Wi-Fi, WiMax or even LTE, and the backhaul network is based on cellular technologies (UMTS, HSDPA, LTE, etc.).

During the transition towards a new communication infrastructure for railways, all these services could be offered by public mobile networks based on more advanced technologies. In fact, UMTS, HSDPA and LTE provide higher bit rates and lower latencies compared to GSM-R. However the deployment in these networks of real-time services in high speed scenarios still has to cope with demanding challenges such as rapidly time-varying radio channels, frequent handovers, Doppler shifts and multi-path reception. All these factors increase the probability of service interruptions, that are particularly harmful for services demanding low latencies and continuous connectivity. In this context the measurement of communications performance on-board high speed trains is a need to guarantee passengers satisfaction, safety and security when used for railway signalling.

In this sense there is a lack of work oriented to characterize Internet traffic in railway environments based on real experimentation. Previous studies like Sivchenko et al.[64] and Aguado et al. [65] were based on simulations, but not on real measurements. Real experimentation is complex and time-consuming but it is always needed to detect real problems and particularities of each deployment, due to, for example, the orography. Other studies [66] are based on field trials focus on propagation measurements, without taking into account IP traffic issues or the quality of service provided to the applications level, which is one of the main objectives of this work.

The tools and methods introduced in previous chapters have been used to collect key information for troubleshooting IP communications problems correlating them with propagation issues, such as cell changes or link outages, and resource allocation problems at specific geographical locations, in busy areas (stations, etc.). With the results obtained from field trials carried out for this work, mobile operators, rail operator's infrastructure (tracks owners) and train operators have access to real and specific data to improve radio performance and refine cell planning near to railway tracks.

5.2 On route IP traffic inspection

This section describes the monitoring experiments carried out on high speed railways and the results obtained.

5.2.1 On-board monitoring set-up

We have used the TestelDroid tool [58], developed at the University of Málaga, to capture the IP traffic as well as additional network information using commercial Android phones. A general overview of the field test set-up used during the measurement campaigns is provided in Figure 7.1. Analyzing the collected information we are able to derive different quality metrics. The measurement campaigns have been scheduled on a 155 km railway line, between the cities of Málaga and Córdoba in the south of Spain. We have used different mobile devices during multiple trips to obtain statistically relevant results. To transport the generated traffic, the default mobile internet access from two major network operators have been used during the measurements. In order to ease network characterization, we have selected a constant rate UDP traffic source. Concretely, series of VoIP sessions with G.711 codec have been used for that purpose. The VoIP traffic is generated in an Asterisk server and received by the Android phones running the TestelDroid monitoring software. Using the formerly described setup, in this approach we have focused on analyzing the downlink communication performance in railways scenarios. The reference values used to evaluate VoIP performance are shown in Table 5.1.

5.2.2 Description of monitored parameters

During the tests, the monitoring software probes different sources of information and generates independent traces for later processing of each magnitude under analysis. In the collection process all the stored data are appropriately time-stamped with precision better than 1 ms. These time marks help the processing tools to easily correlate and merge the different information layers. 4 different sets

Mean Packet Delay	$\leq$ 150ms
Packet loss rate	$\leq$1%
Mean Jitter	$\leq$ 25 ms
Mean Opinion Score	$\geq$3.5

Table 5.1: VoIP metrics reference values

of parameters are obtained, IP traffic, GPS coordinates, Battery consumption, and network related information. Most of the collected parameters are stored in human readable formats, with the exception of the IP traffic that is saved in a pcap file.

The Network information layer contains information about the mobile network where the phone is connected. The radio parameters monitored are as follows:

- Cell identifier
- Location Area Code (LAC). A location area typically comprises several cells.
- Received Signal Strength Indicator (RSSI). This is a measurement of the signal power received at the phone's antenna connectors.
- Primary Scrambling Code (PSC). This is a WCDMA physical layer parameter that allows differentiation of one cell from its neighbors.

Network information may be very valuable for network operators, as it will allow them to trace the issues detected back to their actual deployment and configuration of base stations.

The Location related information is obtained using internal GPS receivers in the mobile phones used for the tests. The following GPS parameters are maintained in addition to timestamp:

- Latitude
- Longitude
- Speed

GPS coordinates are useful for locating measurements within their context, and the speed is very important in a high speed train, where 300 km/h may be exceeded.

Battery consumption is continuously recorded because it is a very valuable resource in mobile devices. Even in normal environments, some studies [67] have identified energy consumption as a key factor in the user's perceived experience, particularly when low battery levels are involved. Although high speed trains reduce travel times, battery drain is even more critical in a railway environment where distances of hundreds of kilometers are involved and higher transmission powers are required because of the larger cell sizes.

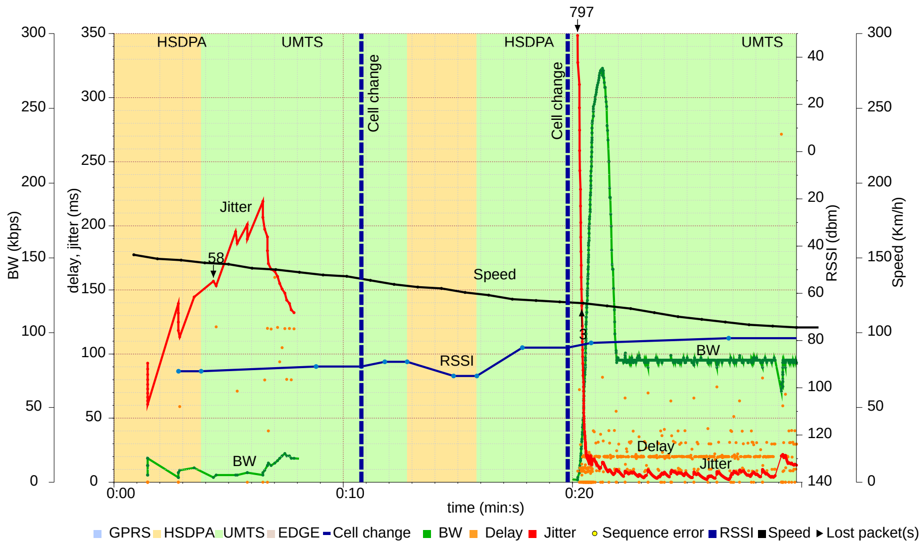

Figure 5.2: Time based representation of RTP flow 0x2e90aaf3

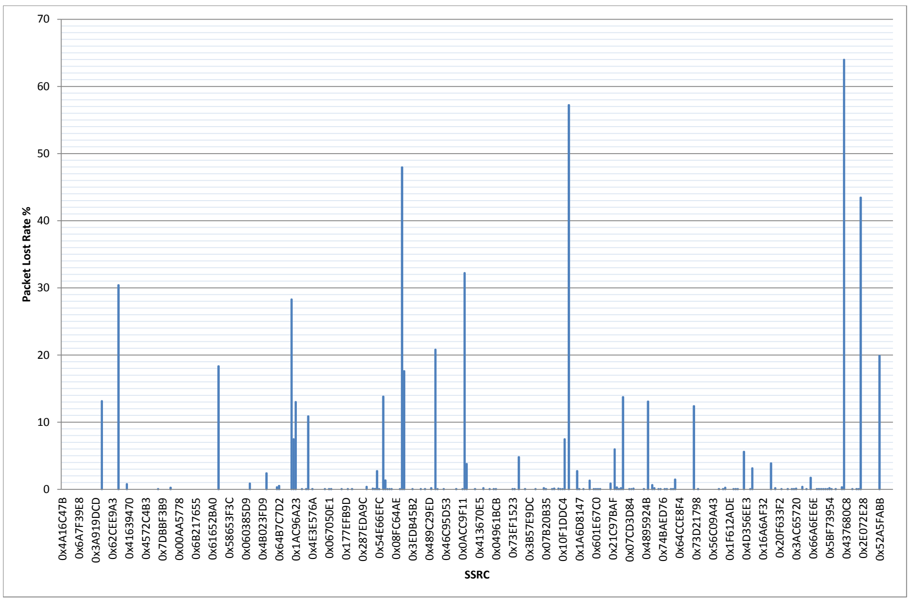

Figure 5.3: Packet loss rate Operator 1

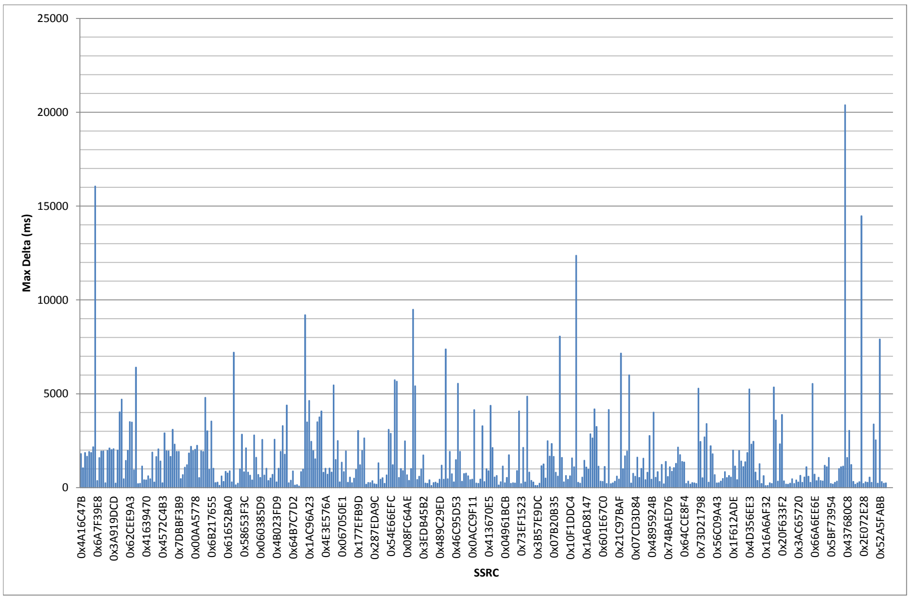

Figure 5.4: Max delta Operator 1

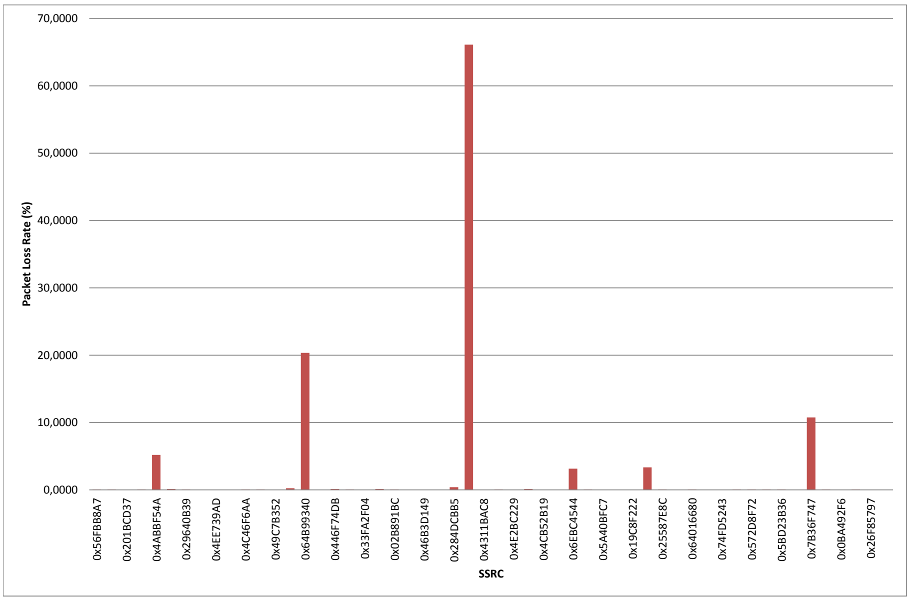

Figure 5.5: Packet loss rate Operator 2

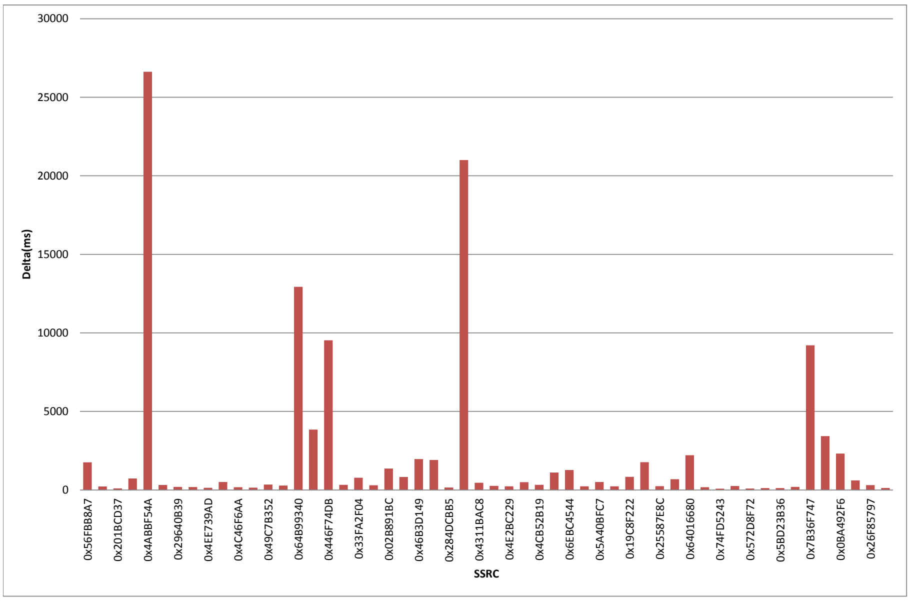

Figure 5.6: Max delta Operator 2

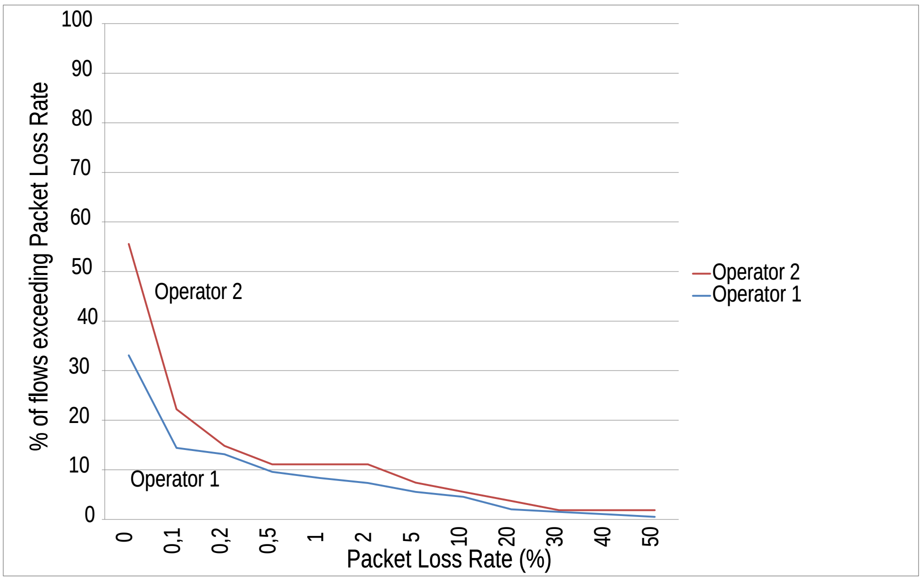

Figure 5.7: Packet loss comparative cumulative percent

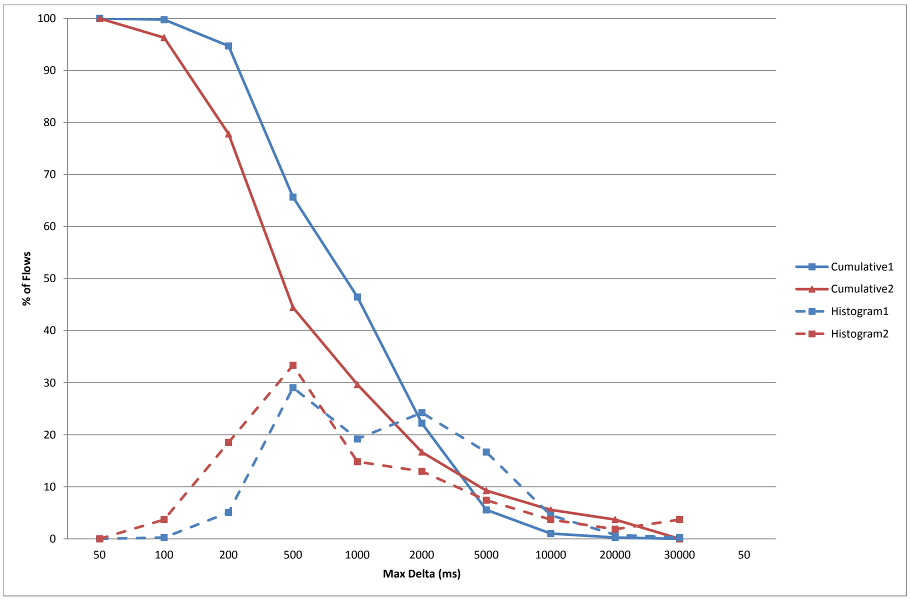

Figure 5.8: Max delta cumulative and histogram

5.3 VoIP measurement results

Each VoIP session is associated to one RTP flow, and for each flow we generate average statistics and charts for different key magnitudes. A Packet Loss chart contains the average rate of packet losses per flow. The packet loss rate has large influence on the user experience, having a large weight in quality models such as the E-Model. The Max Delta chart shows the maximum delay between two consecutive packets within a flow. As the source generates one packet every 20 ms, it is a clear metric of connection interruptions. The Mean Jitter chart contains the calculated jitter for every flow. The jitter has been derived according to [59], which involves an exponential filtering of the instantaneous changes of packet delay. The jitter provides an estimation of how variable the communication conditions are for a session. The PESQ MOS measurements provide an estimation of the perceived quality for each flow. In Figures 5.35.45.55.65.75.8 we provide detailed results on packet loss rates and inter-packet delays.

Two graphs show the rate of packet losses for every measured flow of operator 1 (Figure 5.3) and operator 2 (Figure 5.5). The graphs for the two operators have different densities because we have conducted 400 experiments with operator 1 and 55 with operator 2. Although most flows present a very low rate of lost packets, a non negligible set of sessions have exceeded the 1% target threshold. VoIP streams with losses exceeding that level are typically considered of poor quality. In addition, the packet loss rate reaches unacceptable levels for some flows, with more than 60 percent of the transmitted packets being dropped for specific cases in both operators. In Figure 5.7 we represent the cumulative distribution of flows exceeding given packet loss rates. Approximately, one in ten flows exceeds the selected target. It is also worth mentioning that at least one packet is lost in more than 30-50% of the sessions depending on the operator.

Furthermore, for both operators, results exceeding 10% of losses are in the order of 5% of the flows. In a comparative analysis, we can conclude that operator 1 consistently provides lower loss rates, but the differences are reduced in the range of bad performing flows. It is interesting to compare the obtained values with those defined in [68] as higher target loss rates that can be negotiated between the phone and the network. When establishing a data context, the maximum selectable Service Data Unit (SDU) error rate is 1% for conversational traffic. However, the results obtained clearly surpass the limit of 0.1% defined for background and interactive traffic classes, which are the profiles typically offered by network operators for most users. Regarding the time distribution of the packet losses, we have seen that in most cases the lost packets are grouped in bursts. We provide more details on this topic in Section 5.3.2.

Inter-packet delay statistics are represented on the Figures 5.45.65.8. In Figure 5.4 and Figure 5.6 the flow statistics are presented separately for operators 1 and 2 respectively. The maximum gap (delta) between consecutive packets will help us to estimate the duration of handovers and link outages at the IP level. The duration of handovers in high-speed trains has been studied in [69] from a simulation point of view. The results given there point to a mean duration of 389 ms with an standard deviation of 23 ms. However, in our measurements we find that the median session would contain a maximum gap of nearly 500ms or 1 second depending on the operator and a much wider variability between sessions. In Figure 5.8 it is possible to see that in general operator 2 has shorter interruptions than operator 1. However, in the range of flows with 10% higher gaps operator 2 shows a worse performance, with some interruptions even lasting more than 20 seconds.

5.3.1 Identification of problematic sessions

As shown in the previous section, in some sets of measurements the performance observed has been unexpectedly low. We have used the visualization tools introduced in Chapter 2 to help identify the sources of communication issues.

Different conflictive sessions are used to illustrate the features and benefits of the methodology and tools used in this work. These are scenarios not meeting the formerly mentioned quality targets for packet losses, jitter and MOS. Each VoIP session is associated with one RTP flow and we have analyzed the temporal evolution of its IP statistics inside each flow. Furthermore, we have related them to received signal strength, cell changes, train speed, geographical coordinates and other parameters as previously stated.

To identify the source of potential communication issues, we have developed different representations using the magnitudes under analysis. In fact, two main perspectives have been selected for that purpose, time based and geographic views. In the former, the evolution over time of a selected set of parameters is presented. This view is of use to accurately study the timing of specific events, such as handover duration. We will see next that time representations provide lots of meaningful information for network characterization. In addition, the geographical representation is of great help to put measurements into their real context. This is a key factor in understanding the root of communication issues. By being aware of the measurement locations, it is straightforward to trace service interruptions back to critical coverage areas to name just one of its many potential applications.

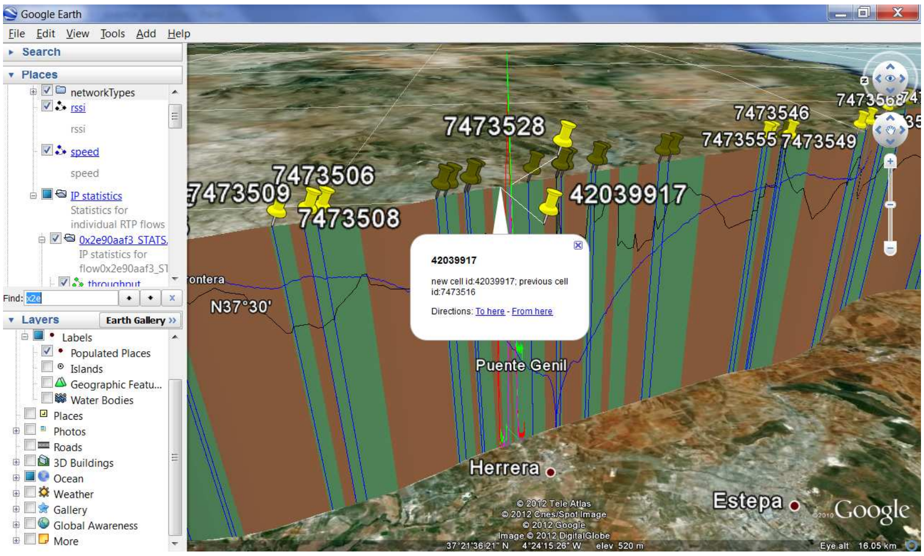

Figure 5.9: Map corresponding to RTP flow 0x2e90aaf3 (Location: Puente Genil)

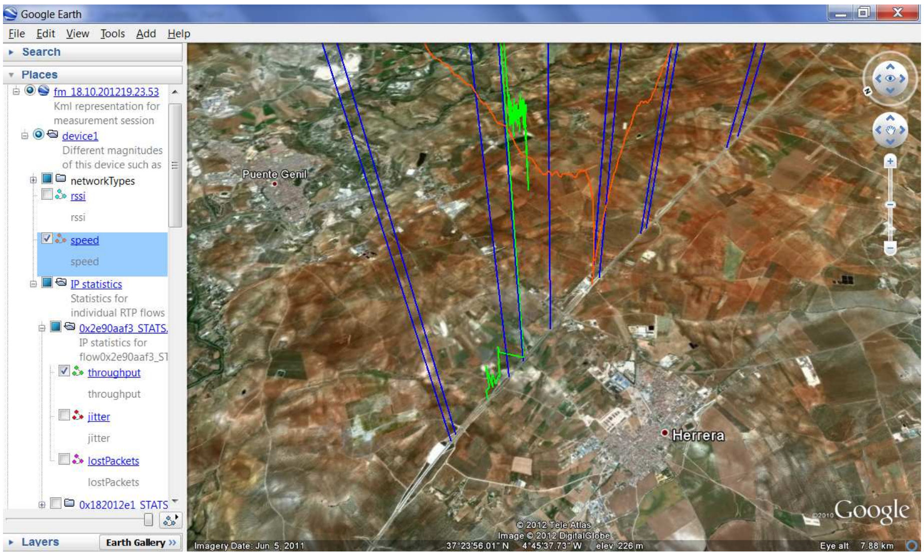

Figure 5.10: Zoom on map corresponding to RTP flow 0x2e90aaf3 (Location: Puente Genil)

5.3.2 Time based representation

We will use Figure 5.2, obtained in step 7 of the RTP processing chain, as an example of session diagnostics to explain the time based representation. In this measurement the train is decelerating, passing from 150 km/h to 100 km/h during the session. The technology in use is shown as the background color, with orange for HSDPA and green for UMTS in this case. As cell changes are also illustrated by adding vertical blue dashed lines, it is easy to infer that the large interruption in the call session has been originated by an apparently faulty handover between two cells. It must be noted that the IP connectivity is not actually recovered until a second cell change has occurred. As a consequence, the communication has been lost for more than 12 seconds. Orange dots mark the delay between consecutively received packets, and to the right of the figure they clearly highlight the 20 ms pattern of our VoIP session. However it can be seen that, even in the final favorable conditions, the effect of the mobile environment increases the inter-packet delay to 80 ms and even up to 250ms at the very end of the session. In the adverse network conditions at the beginning of the session, a very limited number of packets is received with frequent gaps of longer than 1 second. The green line represents the throughput averaged in the last second, that is very poor at the beginning and not applicable during the gap. After the last cell change, a peak in the bandwidth can be appreciated which is not related to the source traffic. Remember that the voice codec was intentionally chosen to keep a constant rate. This peak is clearly caused by the network configuration, that uses buffers for incoming undelivered packets and forwards them to the phone in the new cell increasing the throughput to above 250 kbps until the buffered data are flushed and the normal rate of 80kbps is recovered. The reader may recall that the throughput peak is not wide enough (nearly 2 seconds) to compensate the poor beginning of the session (comprising a total of 20 seconds). Simply with a quick graphical inspection, the buffering capabilities offered by the network can be estimated as approximately 500 kbps. As a consequence of the limited buffer size, most of the initial packets have been lost. Bursts of lost packets are inferred in our methodology from sequence numbers and are also plotted with black arrows indicating the size of the bursts. A total of more than 800 packets have failed to reach the phone, thus missing more than 16 seconds of traffic. For some services such as voice it would not make sense to deliver buffered packets after such a delay, because preserving the end to end delay of the fresh traffic would increase the perceived quality. On the contrary, non-real time traffic like TCP based services may be severely affected by packet losses related to buffer overflows. Providing service specific bearer capabilities would be required to maximize perceived service quality.

5.3.3 Geographical representation

The tridimensional representation generated in step 8 of the RTP processing chain is based on Keyhole Markup Language (KML) descriptors, which takes advantage of the integration of KML with Google Earth. Thus, it is possible to dynamically modify the information under analysis. Simply by selecting different layers of information to be displayed, the magnitudes of interest are shown or hidden depending on user preference. This flexibility makes it particularly suitable troubleshooting communication issues related to the physical network deployment, e.g. because of the impact of the terrain, natural obstacles or nearby densely populated areas. The results are grouped using specific measurement magnitudes in a particular session as the lower level positioning entity, scaling up to RTP flow entities at an intermediate level and complete measurement campaigns during full trips as overall containers. Using this approach, it is possible to filter information at will on a per flow basis. Using context information such as the length of the measurement path, appropriate scaling factors are applied for convenient visualization. Each positioning entity is assigned specific location information and a recommended perspective for visualization based on parameters such as the train's speed and direction. This information allows the geographic viewers to quickly set an appropriate initial visualization perspective. The visualization can later be customized by varying the zoom level and many other parameters.

The geographic view of the former scenario is shown in Figure 5.9. Only the information belonging to the RTP flow of interest has been selected for visualization. In addition to the information layers provided by the measurement and processing tools, external information sources may also contribute to analysis. The process to follow when inspecting an RTP flow would be simple as we now explain. The user would only need to type a fragment of the flow number of interest using the find function and the flow will be automatically selected in the places side bar. By double clicking on it, the view will be directly focused above it. Cell changes are marked with yellow marks, also including the cell identifier. As cells changes may appear close together, additional information is added to each cell mark and can be accessed by simply clicking on the marks. Both source and target cells for each handover or cell reselection are indicated in this way. Selecting the cell changes in our example, it became clear that all the cells involved are different and thus the phone has not bounced back to the original cell after the initial handover.

In a similar way as the time based representation, the geographic representation contains network related information and IP traffic statistics. The default appearance of measurement results is the same for both representations but each element may be customized as desired during runtime. For example, the background color associated with the technology in use may be changed, made partially transparent or totally hidden as shown in Figure 5.10. In that view, the train speed has been assigned an orange color and only cell changes and IP throughput have also been enabled. By selecting a customized perspective, it can be clearly appreciated that the communication issues appear when the train is passing between two towns close together, before reaching a railway station. The algorithms for cell selection are mainly driven by received signal strength according to some parameters configurable by the network operator. Thus it would not be unlikely that a similar treatment were applied to train passengers and to nearby town residents. As a result a short range cell deployed in one town could have been assigned and probably interferes with cells from the neighboring town cells which may affect communication. In future work we plan to also represent the location of known cells to help diagnose communication issues. In that way, network operators could import sets of cells to the tool chain from their cell databases and use this additional information layer for enhanced diagnosis.

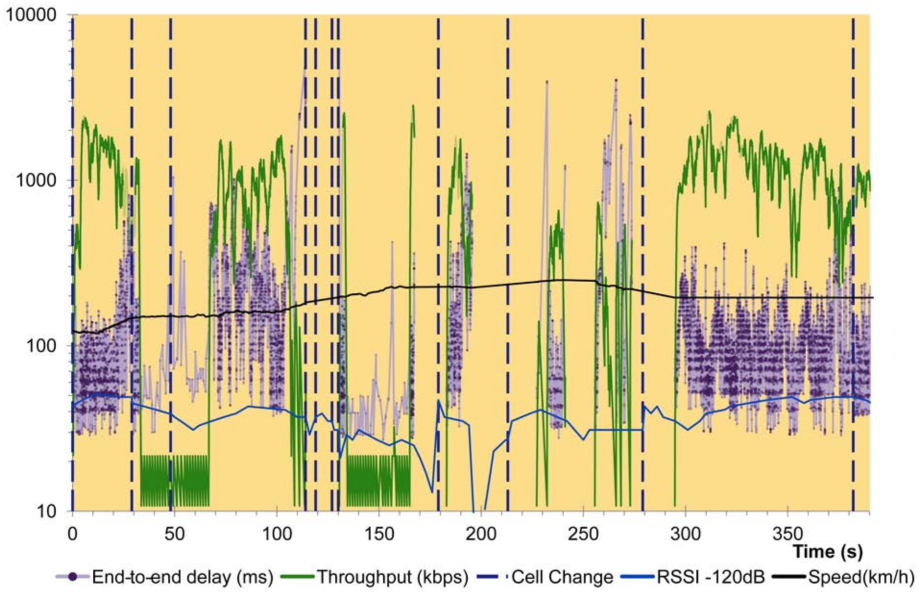

Figure 5.11: FTP throughput and end-to-end delay

5.4 FTP performance measurements

We have also analyzed the behavior of TCP connections corresponding to file transfers. A total number of 40 FTP downloads were conducted on-board a high speed train. 26 Mbytes files and default settings for the receive window (65536) have been used. The transfer lengths measured range from 2 to 7 minutes, with a typical deviation of 70 seconds, with operator 1 performing 10% better in average. Table 5.2 compares additional statistics. The mobile device uses selective acknowledgement (SACK) reports to indicate ranges of sequence numbers that can be lost or received out of order, but they can also be used to indicate duplicate reception of packets. We can see that operator 1 minimizes missing packet reports but delivers a higher number of duplicated packets. More than 1% of the packets from operator 2 are received out of sequence and 3% of the packets in the reverse path contain SACK reports indicating missing TCP segments, which is two orders of magnitude higher than in operator 1.

In Figure 5.11 the end-to-end transmission delay is concentrated in steps of 10ms, the radio frame length, because of the low delay variations outside the air interface.

5.4.1 Identification of TCP communication issues

We have identified significant performance degradation when TCP recovers from large sets of lost packets being notified using SACK. In this scenario, TCP eventually enters a state where a new packet alternates with a retransmitted packet in 1 or 2 second intervals without increasing the transmission window. It takes 30 seconds on average to retransmit the reported lost packets and recover from this state. This issue appears in almost every operator 2 call, increasing

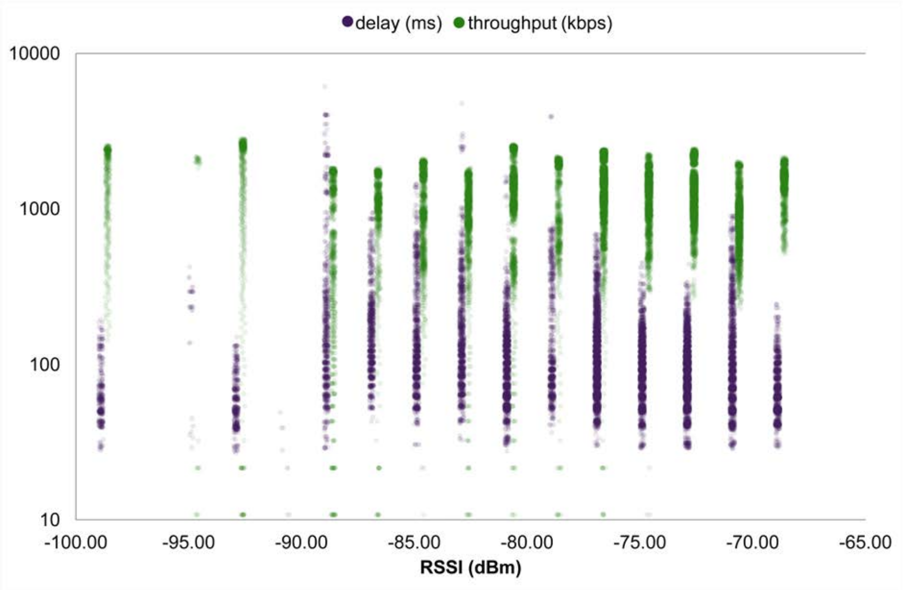

Figure 5.12: Correlation of Delay and throughput with RSSI for a FTP session

the transmission time. A call containing two occurrences of this issue around seconds 50 and 150 is shown in Figure 5.11. It can be observed that despite transmission delays below 100 ms the throughput does not increase until all the pending packets have been received.

We have also observed that RLC level retransmission delays also have some impact on TCP performance, as throughput decreases even after buffered PDUs are delivered when failed PDUs are successfully retransmitted. The presence of RLC retransmissions can be detected because very abrupt bursts are received at IP level. These sets of data are not buffered in the network but in the actual receiving phone waiting for pending RLC data in order to keep the in sequence delivery. In future experiments we will analyze if a larger receive window improves performance in the presence of RLC retransmissions.

Correlating the TCP traffic analysis with radio information, we have detected that the SACK related performance issue appears right after handovers, indicating that operator 2 is not forwarding pending data between cells.

With the help of synchronized GPS information we have also identified the areas where connectivity gaps appear. The first one is associated to four consecutive cell changes when the train has just left the urban area of the city of Malaga. The next gaps appear when a rural area between small villages is crossed with low RSSI, as the connection stalls when RSSI goes below -100 dBm. Above that value, we have not found dependencies between RSSI and throughput, as can be seen in Figure 5.12. The last gap occurs while the train decelerates from 250 to 200 km/h and enters the cell that serves the long tunnel under the mountains of Malaga.

Test	Parameter	Operator 1	Operator 2
VoIP	max jitter > 25 ms	87.37 %	71.46 %
	mean jitter > 25 ms	2.52 %	7.47 %
	packet loss > 1%	1.26 %	10.85 %
	max packet loss	63 %	65 %
	max interpacket delay	21 s	27 s
	2.5 < PESQ MOS < 3.5	4 .04 %	9.34 %
	PESQ MOS $\leq$ 2.5	6.56 %	9.09 %
FTP	SACK missing segments	0.03 %	3.25 %
	SACK duplicated packets	0.74 %	0.09 %
	Out of order	0.00 %	1.34 %
	mean transfer time	248.57 s	270.20 s

Table 5.2: A comparative summary of the results

5.5 Measurement conclusions

We conducted 500,000 ping tests from the server to the operator's network to verify that delay variations are negligible outside the air interface. Most packets (97.5%) are delayed exactly 12 ms, with a typical deviation of 0.59 ms.

The measurement results reveal that bursts of packet losses and delay spikes are the main issues detected on high speed scenarios. They can be caused by the following factors:

- Frequent changes of cells, which usually are not error free.
- Poor radio conditions at cell edges, particularly when RSSI decreases below -100dBm, because of the distance to the base stations.
- Allocation of low bandwidth bearers that are unable to sustain the amount of traffic required by a fixed rate service.
- Resource preemption because of higher priority voice calls.
- Interference from neighbor cells, specially in transitions between rural urban areas such as cities or small villages.

In the presence of RLC acknowledge mode, the impact of these factors could be probably minimized with increased redundancy in resource allocations, specially in retransmissions, if possible according to the network load. The results obtained have proved the features provided by TestelDroid enables a cross-layer analysis of radio and IP traffic performance and leverage location to be carried out to better diagnose mobility-related performance issues.

Chapter 6

A realistic experimentation testbed

6.1 Introduction

This chapter describes an experimental test setup composed to verify the communication performance of internet applications over LTE (Long Term Evolution) networks.

The test environment has been particularized to study the VoIP service, as migrating customers of legacy circuit-switched voice services to the data centered LTE technology will require appropriate understanding of the perceived QoS (Quality of Service).

The key benefits of the proposed solution are described, such as the ability to verify a wide range of network parameter configurations from RF (Radio Frequency) to IP (Internet Protocol) levels. Cross tuning these parameters will allow obtaining optimum configurations that satisfy the QoS requirements of these services.

Initially, the drawbacks of typical QoS evaluation environments are highlighted and the design goals of the testbed are introduced. The proposed solution is then described in detail, particularly the powerful features provided by Keysight Technologies' solutions used for LTE eNodeB and Radio Channel emulation, as well as the open applications used for VoIP communication. The voice test vectors and the processing methods for QoS estimation are then presented. Finally the key benefits of the proposed scheme are summarized.

LTE has proven to be the primary choice for network operators to provide high rate data services with an evolving path towards real 4G speeds, which will be provided by LTE Advanced through a set of 3GPP specification releases. As LTE is an All-IP data centered technology, there is no support for circuit switched voice services. However, LTE is expected to provide voice services over IP and IMS (IP Multimedia Subsystem), as intended by the VoLTE (Voice over LTE) initiative.

As key network operators such as Telia Sonera stated at the Mobile World Congress in Barcelona 2011, their customers are used to a high voice quality, and any migration path to LTE should not compromise the provided quality of service. Voice is still the core business for many operators, and obtaining accurate knowledge on the trade-offs involved in resource management and QoS over LTE will be critical for their success.

With the aim of generating reference results, we suggest a novel approach that will hopefully help key market players in making decisions.

6.2 Previous solutions

Typical test approaches involve the use of well known network simulators such as ns-2 or Opnet, or even Matlab when aiming to analyze physical layer parameters. Quite often, the results from different contributions are not directly comparable because in the end those techniques are based on models and not on actual implementations and devices. Although they are useful for initially addressing a research topic, the intrinsic nature of models may lead to excessive simplification.

In the past, field testing was a frequent practice to obtain information related to the perceived quality of service, and some commercial (and expensive) tools such as Qualcomm QXDM are available to monitor low level LTE parameters. However, the results obtained are not easy to reproduce because of variable radio conditions due to phenomena like fading, multipath propagation and interference. In addition, typically there is limited or no information about the network configuration, which cannot be modified and load conditions is limited or non-existent.

6.3 Reference test environment

To overcome the limitations of the former approaches, our goal is to obtain a reference test environment for QoS measurements. To obtain reliable and comparable results, repeatability will be a key point because many communication effects are random in nature and require repeating experiments a number of times to be statistically meaningful. Accuracy is also required because e.g. some of the most QoS impacting parameters, such as the loss rate, are related to the signal to noise ratio.

As we intend to identify relations between network parameters and perceived QoS, a high degree of configuration will be required. Furthermore, the behavior of the actual user devices should be analyzed, instead of working with models or assumptions.

Time is a very valuable resource, and as the suggested solution should encompass a large number of experiments, automation capabilities will be required to speed up the process. Subjective quality estimation is an example of a specially costly process, as it would involve the participation of a large base of listeners to evaluate the received voice quality. Hopefully, objective methods should provide a valid reference for speech quality estimation.

In order to meet the target requirements, we suggest a test architecture based on high-end network emulation, open VoIP applications and standard voice quality estimation methods.

Voice calls are originated between the Yate softphone SIP client, and the Asterisk VoIP server. The client uses an internet connection established by a GT-B3730 USB modem, a commercial LTE device from SamsungTM. An LTE network is established by an T2010 eNodeB emulator from Keysight.

A custom tool chain has been integrated to allow automatic processing of received traffic parameters and voice quality.

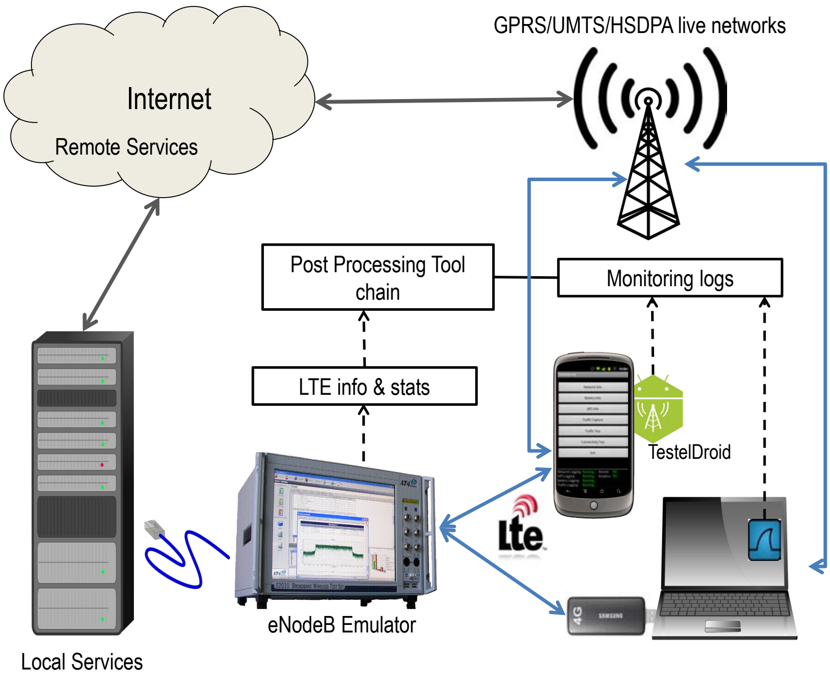

Figure 6.1: UMA testing facility setup

The Asterisk VoIP server is used in the setup as it is a well known open source implementation. A useful feature for test automation is the auto answering capability, which enables the playback of a stored speech when a client application requests the establishment of a voice call towards the server.

The T2010 operates as a data gateway during the experiments, providing access to local or external servers forwarding LTE IP traffic to an Ethernet interface. Only the data plane is represented in the diagram, but control signaling is also required, e.g. RRC (Radio Resource Control) signaling is required to establish LTE connections and SIP (Session Initiation Protocol) is involved in VoIP session establishment.

The G.711 codec has been selected for the initial tests, following the recommendations in [50] because it provides better quality under equivalent delay conditions. It must be noted that other codecs will be analyzed in future works, especially AMR (Adaptive MultiRate) since it is used by 3GPP (3G Partnership Project) [21]. In this configuration, a USB LTE user equipment is used and thus the VoIP client runs on a laptop. However, when commercial LTE smartphones become available, the VoIP clients will be hosted in the actual LTE phones.

The T2010, is a generic platform used not only in conformance RF and signaling testing but also for design verification. The main features provided are the following:

- RF (Tx and Rx) with support for 2 antenna ports.
- RRC and NAS signaling
- Resource scheduling

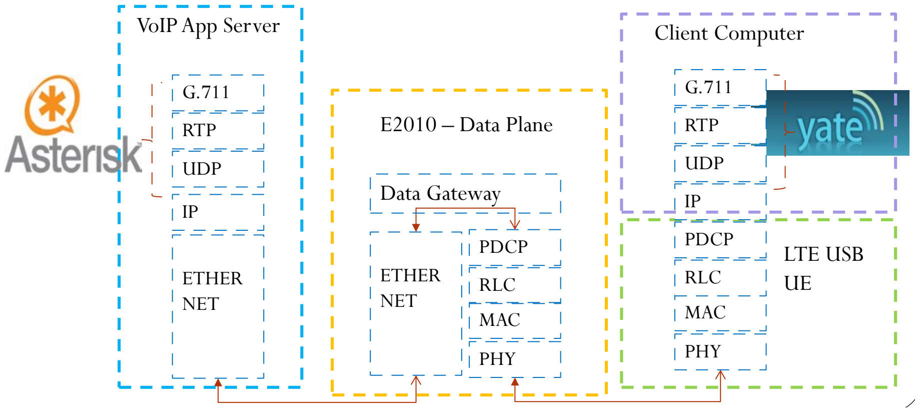

Figure 6.2: VoIP protocol stack

  - Transport Block Size,
  - Modulation and Coding Scheme
  - Frequency and Time allocations
  - HARQ configuration

- On board measurement capabilities

- Channel Emulation
  - MIMO
  - Fading
  - AWGN

- RF parameters
  - RSSI, channel bandwidth, ...

In addition to LTE signaling and RF connection features, it also integrates channel emulation and digital generation of impairments such as AWGN, which is a critical feature to achieve high accuracy when setting SNR conditions. Standard multipath fading profiles defined by 3gpp are supported to emulate reference propagation conditions.

MIMO is a key feature in LTE, as it is one of the foundations of the technology's high rates and spectral efficiency. The T2010 provides up to 4x2 integrated MIMO features, thus increasing the range of test possibilities with interesting network configurations.

The Mobile Test Application is a control and measurement application oriented to design verification using the T2010 platform. It exposes a set of key network parameters, and an intuitive interface that isolates the user from the internals of LTE signaling and ensures a consistent state according to the configuration.

The Test Application also provides measurements oriented to RX performance [78] such as PDCCH and PDSCH demodulation error rate, and collects quality information reported by the user equipment such as channel quality indicator

Figure 6.3: LTE eNodeB emulator T2010

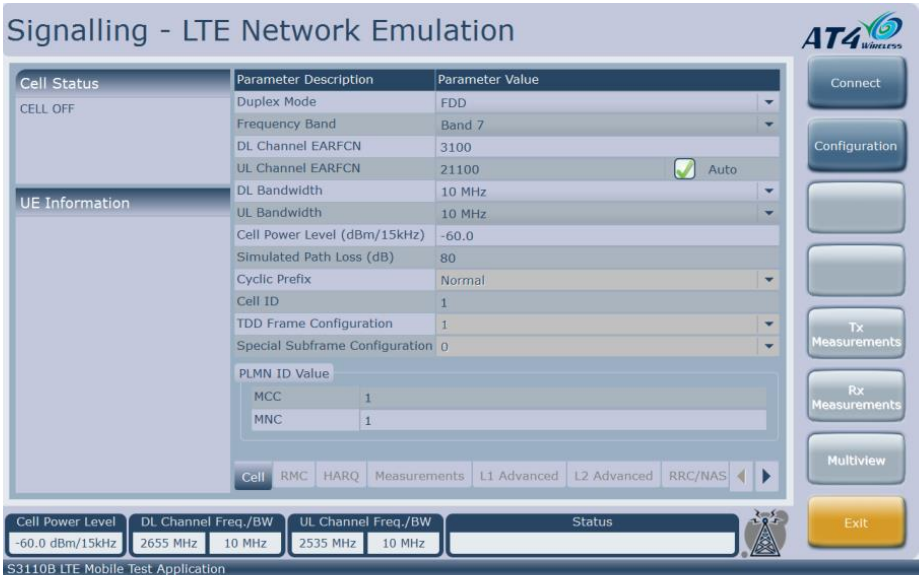

Figure 6.4: Mobile Test Application

(CQI) , reference signal received power (RSRP) and reference signal received quality (RSRQ) defined in [83].

For the purpose of test time optimization, a remote command interface allows automation.

To verify the operation of the testbed, we have carried out some preliminary measurements. These tests have been performed for proof of concept purposes.

Some of the configuration parameters are detailed below:

- MIMO configuration: 2x2
- Channel Bandwidth: 10MHz
- Reference Signal Power: -60 dBm/15kHz
- Noise Power: test specific (*)

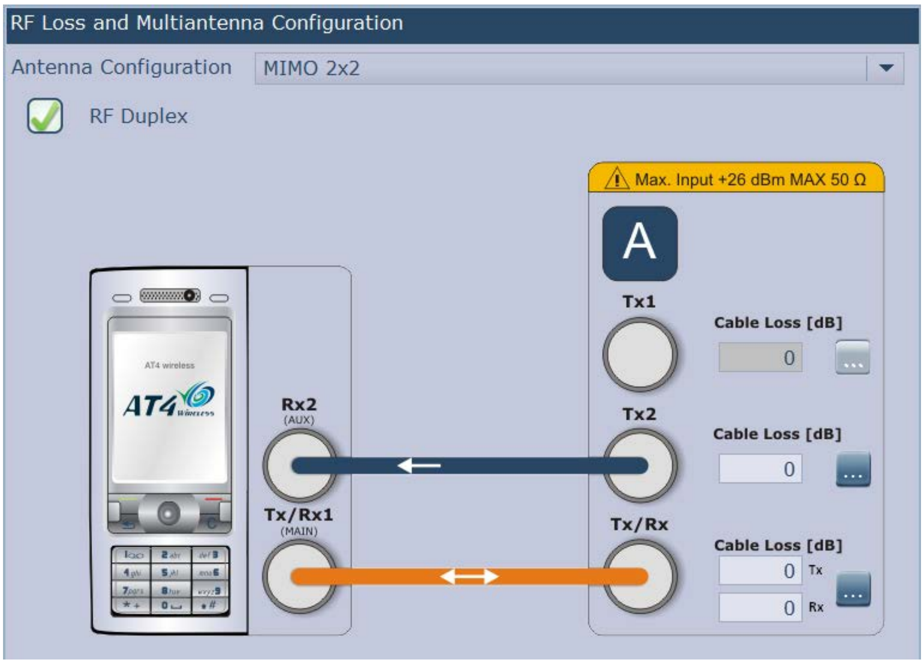

Figure 6.5: Mobile Test Application, RF Loss and Multiantenna Configuration

- Max HARQ retransmissions: 3
- Peak PDSCH bandwidth: 120 Kbps (**)
- Resource allocation: periodic, 5ms
- Modulation: 16QAM
- MCS (Mod.and coding index): 13
- PHICH duration: Normal
- PHICH resources: 1/6
- Number of PDCCH Symbols: 1
- Specific Aggregation level: 2
- Fading profile: ETU70

(*) Different noise powers have been modified to obtain different signal to noise ratios and thus obtain different rates of packet losses.

(**) Although the required nominal IP throughput is 80 Kbps, the scheduling configuration allows peaks of up to 120Kbps to accommodate PDSCH retransmissions.

6.4 Network QoS statistics

The proposed tool chain allows inspection of the IP traffic both at source and destination to obtain typical communication statistics such as packet loss rate, throughput, inter packet arrival, end to end delay and jitter.

Combined with the information provided about the LTE configuration and the reception stats, we can obtain useful cross layer information. As an example, the figure shows the relation between the PDCCH (Physical Downlink Control Channel) and PDSCH (Physical Downlink Shared Channel) BLER (Block Error Rate) and the IP packet losses. It can be appreciated that the relation is not strictly linear because of the randomness associated to AWGN and fading impairments, resulting in bursty behaviors. This relation is also governed by the operation of HARQ, as the incremental redundancy increases the probability of correct reception with each retransmission. An RLC SDU containing an IP packet is lost only after the maximum number of unsuccessful retransmissions has been reached.

This is one example that illustrates the need for advanced LTE network configuration tuning to achieve the QoS levels demanded by specific service requirements such as VoIP. The proposed solution will help us generate knowledge regarding the effect of the different network parameters on the QoS delivered at IP level.

In our previous works [12][58] we clearly identified the need for appropriate network monitoring tools and the advantages of merging information from multiple communication layers.

During the experiments, reference audio samples suggested by ITU-T P.501 [86] were used, as recommended by [90]. These audio samples are specially suited for speech quality evaluation on telephone networks.

To evaluate the suitability of the voice quality analysis, we have used the PESQ (Perceptual Evaluation of Speech Quality) method defined in ITU-T recommendation P.862 [77]

PESQ analysis the relative degradation between the original and the received voice signals. In consequence, both waveforms have to be provided as input for the algorithm. The source signal is directly fed to the server, but for comparison purposes we obtain it from the generated IP traffic to isolate it from server encoding effects, whereas the received voice is reconstructed from the IP traffic recorded at the destination end.

The PESQ method generates an objective score that can be mapped to MOS (Mean Opinion Score) [91] and represents an estimation of the received transmission. That mapping is defined in P.862.1 [89]. It must be noted that PESQ results range from -0.5 to 4.5 while MOS varies from 1 (poor) to 5 (excellent). We have performed some additional tests to validate the estimated quality for different packet loss rates, finding that a fair quality (MOS ≥ 3) is estimated for losses up to 2%. Although it is often suggested that in order to optimize voice quality losses should be kept below 1% [50], other studies estimate the acceptable loss rate at 3%[43] or 4% [25].

Although [50] suggest using G.711 to minimize the delay, because of the scarcity of radio resources and for other reasons such as to minimize the need for transcoding, other codecs like AMR or Skype proprietary codecs will be addressed in future works. AMR codecs have been chosen for the VoLTE initiative as part

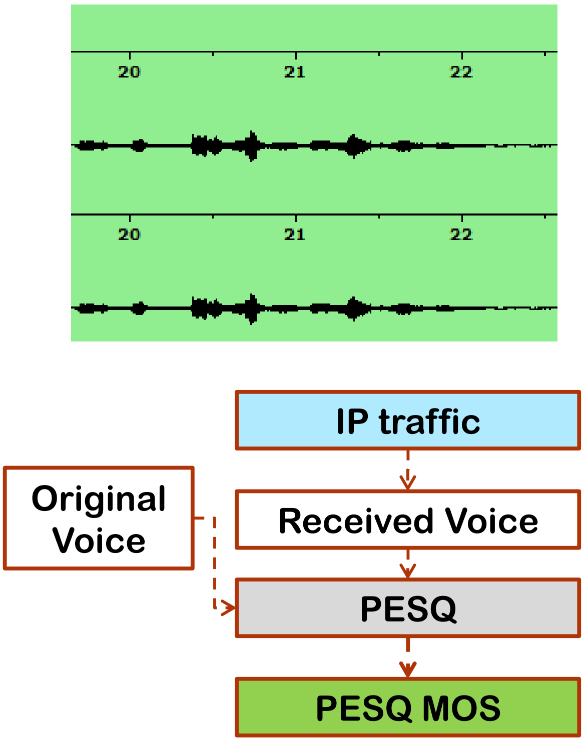

Figure 6.6: Voice quality analysis

of the Voice over IMS profile. Skype codecs are used by a large base of users and appropriate analysis of its performance over LTE networks is also required. The effect of the delay has not been considered in these initial tests, but we will also consider it in our ongoing work because the end-to-end delay has a critical impact on voice quality. In [87], the authors indicate that end-to-end delays lower than 150ms are acceptable by most users, and that delays higher than 400 ms are clearly unacceptable. However, the E-model [88] assumes that the impairments associated to the delay are additive and all factors should be studied jointly, e.g. a lower delay will provide a better perceived quality with the same rate of packet loss. The impact of network configuration on the delay will be addressed in subsequent works. E.g. A high number of HARQ retransmissions will reduce the probability of packet loss at the expense of higher peak delay and jitter.

As stated in [50] the effect of packet losses will also depend on how random or bursty the losses are. In future experiments, the length and frequency of loss bursts will be examined for different propagation conditions and configurations of network parameters.

6.5 Conclusions

In summary, we have proposed a reference scenario for the analysis of the QoS-related tradeoffs for IP services over LTE.

Since real devices are used, the actual quality of service is directly measured without requiring models and simulations. To obtain repeatable results, our proposal relies on maintaining controlled test environments due to the accuracy and configurability of the T2010 LTE eNodeB emulator.

Some preliminary results have been provided to illustrate the possibilities of the proposed approach, highlighting the ability to correlate information from different layers.

Chapter 7

Obtaining more realistic cross-layer QoS measurements: A VoIP over LTE use case

7.1 Introduction

Finally in this thesis we will provide further details on the tests performed to verify the real-time experimentation testbed introduced in the previous chapter, which enables more realistic analysis of quality of service (QoS) in LTE networks. This testbed is envisioned for the improvement of QoS and quality of experience (QoE) through the experimentation with real devices, services and radio configurations. Radio configurations suggested in the literature typically arise from simulations, the testbed provides a real and controlled where such configurations can be validated. The added value of this testbed goes a long way not only in the provision of more realistic results but also in the provision of QoS and QoE cross-layer measurements through the correlation of information collected at different layers: from service and IP levels to radio and protocol parameters. Analyzing the inter layer dependencies will allow us to identify optimal settings for the radio access network and service parameters. This information can be used to suggest new cross-layer optimizations to further improve quality of experience of mobile subscribers. As a use case we examine VoIP service over LTE, which is currently an open issue.

7.2 Testbed configuration for VoIP testing

We have composed an experimental testbed [11] with the aim of providing a realistic test scenario where previous and new radio configurations could be deployed. Additionally it is possible to analyze their interactions and to verify cross-layer performance of Internet applications and services over LTE. Moreover the testbed can be used to reproduce, in a controlled environment, behaviors captured in field test campaigns [11].

The testbed includes a LTE test base station from Keysight Technologies

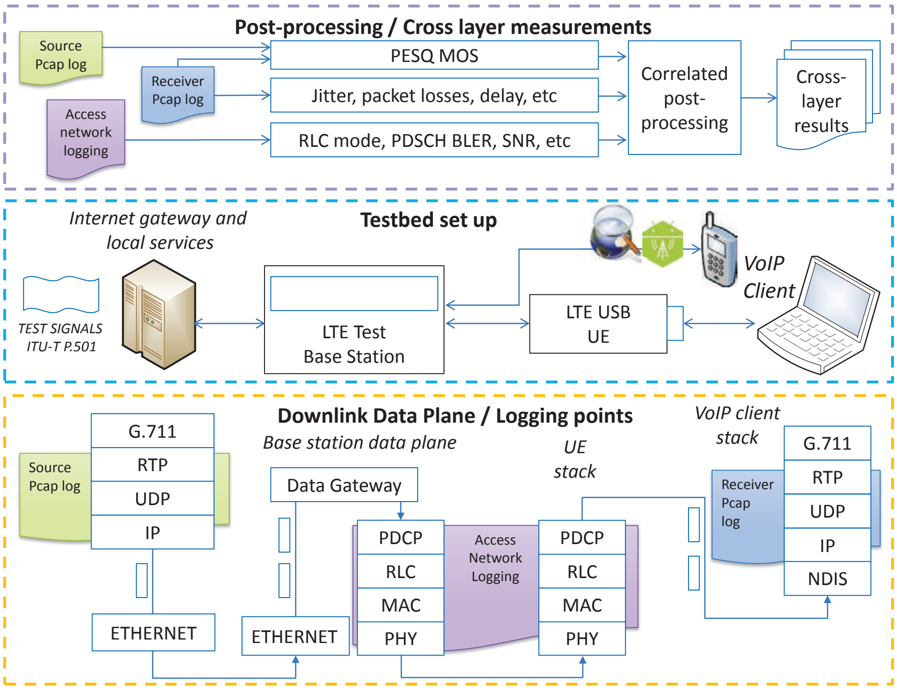

Figure 7.1: Testbed configuration

(formerly from AT4 wireless) which provides high performance protocol and radio capabilities behaving as an actual LTE RAN (Radio Access Network), as shown in Figure 7.1. It also includes features such as emulation of channel propagation that allows modeling fading and additive white gaussian noise impairments, in addition to a high degree of configurability of the LTE stack and logging functions. The LTE test base station supports the connection of real LTE terminals and the transport of IP traffic generated by commercial applications installed on them. The NAS (Non Access Stratum) signalling exchange is provided by a core network emulation. Although the effect of core network transportation is also important for QoS, it has not been analyzed in the present work because of the focus on RAN and will be addressed in the future. The mobile terminals also incorporate advanced monitoring software [12][58]. Finally, the testbed includes post-processing tools which enable the testing and identification of IP connectivity issues and LTE mismatches through the correlation of logs collected at different points as shown in Figure 7.1.

During the experiment, VoIP calls are initiated by a commercial VoIP client running in a laptop. The laptop uses the Samsung GT-B3730 USB LTE modem connected, via a radio frequency (RF) wire, to the T2010 eNodeB emulator from Keysight Technologies. The emulator is connected to the Internet and to a local Asterisk server via a proprietary data gateway. The core network is not presented in this version of the testbed. In this work we focus on the study of radio access interface performance from the point of view of QoS and QoE perceived at the UE (User Equipment). In order to automate the establishment of the calls the Asterisk server is configured to provide a callback service, so that this service

Table 7.1: Resource scheduling and radio frequency configuration in the LTE test base station

Parameter	Configuration
MIMO configuration	2x2
Channel Bandwidth	10MHz
Reference Signal Power	-60 dBm/15kHz
Noise Power	-67 to -73dBm/15KHz
Max HARQ retransmissions	3
RLC Transmission mode	UM
RLC sequence number size	5
PDCP discard policy	no discard
PDCP sequence number size	7
Peak PDSCH bandwidth	120 Kbps
Resource allocation	periodic, 5ms
Modulation	16QAM
MCS (Mod.& coding index)	13
PHICH duration	Normal
PHICH resources	1/6
Number of PDCCH Symbols	1
Specific Aggregation level	2
Fading profile	EPA5

reproduces a 30 second recording each time a call is received in a pre-configured VoIP extension. Records have been extracted from audio samples provided in the ITU-T (International Telegraph Union Telecommunication Standardization Sector) recommendation P.501 to speech quality evaluation on telephone networks. The codec used during the transmission is the G.711 because it is well known and its constant bit rate eases the initial analysis of the impact in throughput and similar metrics. Other codecs such as AMR (codec recommended for VoLTE profile in [21]) or SILK (Skype) will be addressed in future experiments.

Wireshark is used to capture the IP traffic on both sides. The emulator also provides low level EUTRAN traces that are valuable for detailed examination of behaviors of interest. The post-processing of the results collected is carried out using a tool developed in our research group which, among other things, obtains delay, packet losses, jitter and MOS values of the VoIP calls. PESQ (Perceptual Evaluation of Speech Quality) algorithm defined in [77] is used to calculate MOS. PESQ analyzes relative degradation between the original and the received voice signals. In order to apply the algorithm both waveforms have to be provided as input. The source signal is directly fed to the server, but for comparison purposes we obtain it from the generated IP traffic to isolate it from server encoding effects, whereas the received voice is reconstructed from the IP traffic recorded by Wireshark at the destination end.

Table 7.1 contains an example configuration deployed in the LTE test base station.

Different fading and noise propagation conditions have been applied. Multi-path fading conditions are typically experienced in mobile environments as a

Table 7.2: Extended Pedestrian A channel model

Excess tap delay (ns)	Relative power (dB)
0	0.0
30	-1.0
70	-2.0
90	-3.0
110	-8.0
190	-17.2
410	-20.8

result of the user mobility. A typical environment with low delay spread is represented with the EPA5 profile as defined in [78]. EPA stands for Extended Pedestrian A channel model, which contains 7 channel taps with an average delay spread of 45ns and a maximum tap delay of 410ns. The EPA5 profile has an associated maximum Doppler frequency of 5Hz and the associated tap delay and relative power is shown in Table 7.2.

As demonstrated in the table, the Signal to Noise Ratio (SNR) has been swept in a range of 7 to 13 dB to analyze the results under moderate packet loss conditions. As the VoIP service has real time requirements, the RLC layer is configured to operate in unacknowledged mode (UM), that does not re-transmit. However, although the RLC UM does not retransmit unconfirmed data, the Hybrid Automatic Repeat Request (HARQ) at MAC (Medium Access Control) level provides convenient fast retransmission with incremental redundancy. Thus, even in the presence of a moderate Physical Downlink Shared Channel (PDSCH) BLock Error Rate (BLER), a higher layer Protocol Data Unit (PDU) will only be lost if the maximum number of HARQ retransmissions is reached at MAC level.

7.3 Analysis of IP performance

The aim of this chapter is to introduce the importance of cross-layer results in the context of the VoIP service. Using the testbed described in the previous chapter a campaign of experiments have been carried out to obtain the results referred to. Also cross-layer correlations between LTE mechanism and IP performance are explained.

7.3.1 IP parameters

As we have stated in previous sections, some variable bit rate codecs are expected to be used in VoIP over LTE, however in this approach we have chosen the G.711 codec to compare because it is a standard and widely studied codec with constant bit-rate (CBR). G.711 codec has a 64 kbps voice bandwidth. The constant sampled rate of 20 ms and the fixed 160 bytes of the payload plus 40 bytes of IP/UDP/RTP header produce a flow with a bandwidth of 80 kbps at the IP level. At the radio access a peak PDSCH bit rate of 120kbps have been scheduled to provide enough throughput headroom. We have analyzed the IP

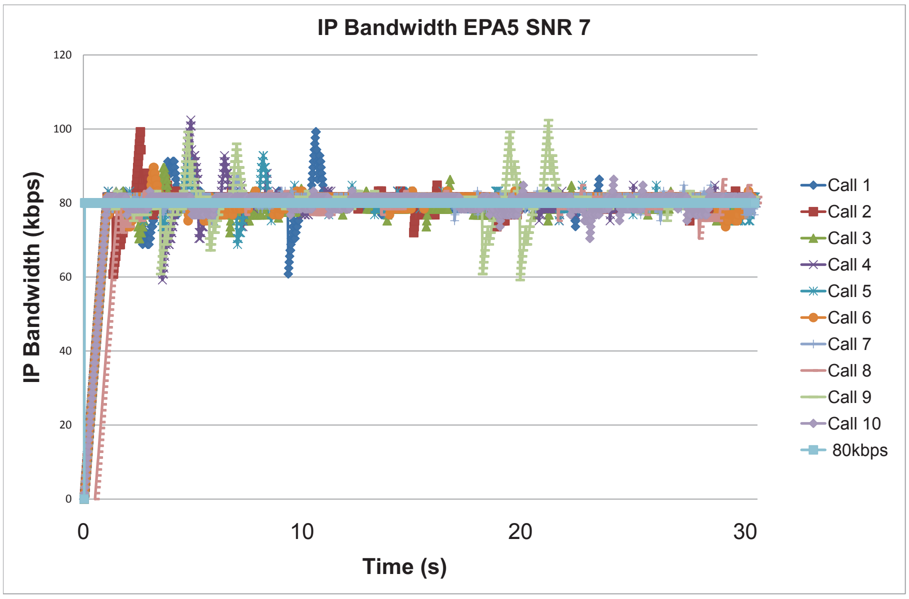

Figure 7.2: Instant IP bandwidth measurements during 10 VoIP calls in a EPA5 LTE scenario with a 7 dB of SNR

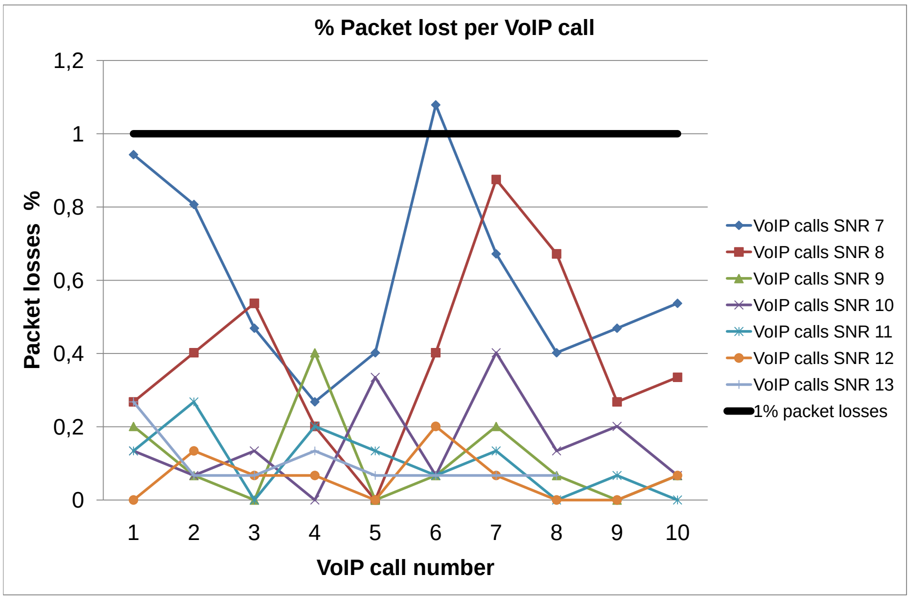

Figure 7.3: Packet losses per VoIP call for different levels of SNR in a EPA5 scenario

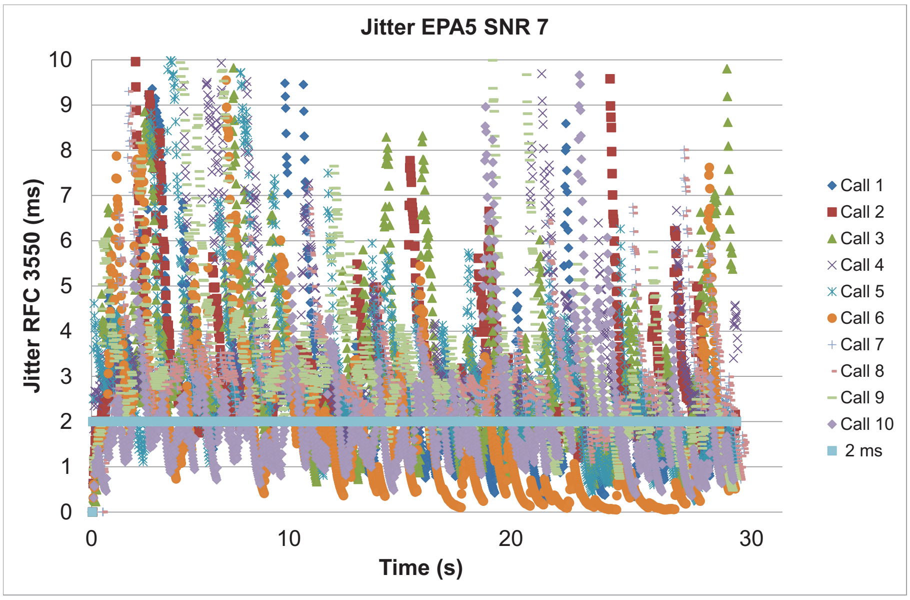

Figure 7.4: Instant jitter measurements during 10 VoIP calls in a EPA5 LTE scenario with a 7 dB of SNR

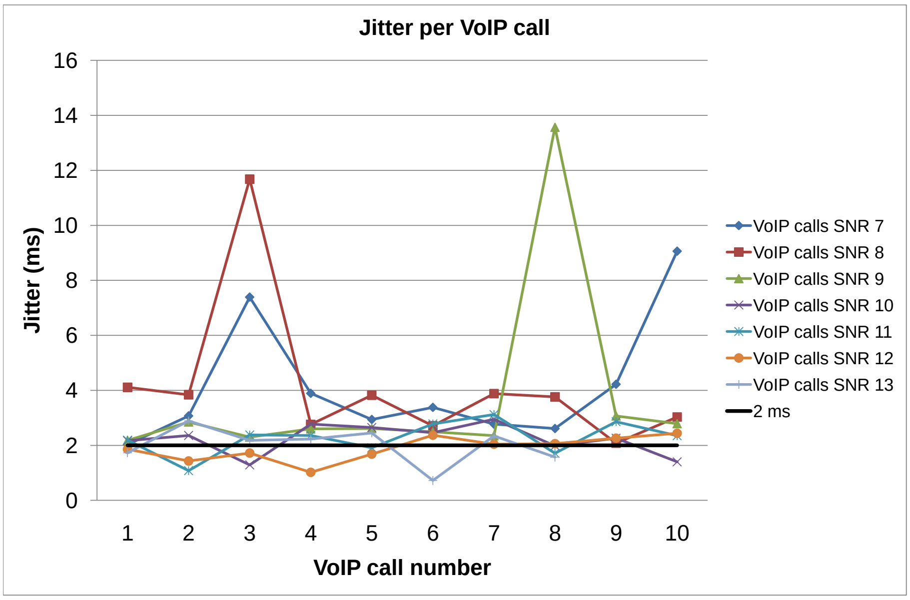

Figure 7.5: Mean Jitter per VoIP call for different levels of SNR

bandwidth of the flow received by the mobile device during 10 consecutive VoIP calls in the worst-case scenario, that is the scenario with 7 dB of SNR. The call length is 30 seconds. Results are depicted in Figure 7.2. The instantaneous evolution of the IP bandwidth for the 10 calls is compared with the nominal 80 kbps constant bit rate generated by the source. The IP bandwidth fluctuations obtained at the destination are caused by lost and delayed packets, which cause instantaneous decrements of the received bit rate. Successful retransmissions generate bandwidth peaks one second after the decrement, because the calculation is made averaging the received packets during the last second. Although we have used the LTE same fading profile and nominal signal to noise ratio in all the calls, different instantaneous results have been obtained because of the random nature of the fading and noise generators.

A packet loss rate close to 0% and a jitter lower than 2 ms can provide good quality VoIP calls, even comparable with a PSTN (Public Switched Telephone Network) call. However most codecs used in the VoIP service are not tolerant of higher packet losses. For the "standard" G.711 codec or the G.729 codec, a 1% packet loss rate significantly degrades a call [79]. In Figure 7.3 we depict packet losses obtained during sessions where a different SNR were configured. We can see that packet losses are higher than 1% only for VoIP call with the lower configured SNR, 7 dB.

The interarrival jitter is calculated as defined in [59] using the IP traces captured at the mobile subscriber terminal. Each RTP packet contains a timestamp which reflects the sampling instant of the first octet in the RTP data packet. The instantaneous variation of the delay is obtained by comparing the elapsed time between two received packets with the difference between their timestamp. The jitter is then derived applying a filter to the instantaneous delay variation. In Figure 7.4 we observe the temporal evolution of instant jitter during VoIP calls conducted in the scenario configured with 7 dB of SNR, while Figure 7.5 shows the mean jitter obtained in all the scenarios. This is a traditional analysis based on only IP parameters, which is very useful to characterize the performance of the service under study. However it is not enough for the adaptation and optimization of the service to the underlying transport technology. This can be better appreciated by observing Figure 7.5. Concretely it can be seen that despite the delay variations introduced by HARQ retransmissions the average jitter is kept in only a few ms, although eventually the instantaneous delay may vary in the order of tens of ms. However to obtain a better comprehension of VoIP performance over LTE it is necessary to monitor low level parameters and correlate them with IP parameters. In the following section we will analyze the correlations between parameters monitored at different LTE layers.

7.4 Cross-layer measurement analysis

In this section a further analysis of the results obtained in previous experiments is provided. Specifically we will present a correlation of the SNR configured in the experiments with different IP and RF measured parameters. In addition, we will also depict the mapping of voice quality measurements. Known functions (linear, polynomial, exponential and logarithmic functions) have been applied to obtain the correlation between the parameters. To retrieve the degree of correlation the coefficient of determination R^2 have been calculated. R^2 ranges from 0 (indicating the absence of a systematic correlation) to 1 (indicating a perfect correlation).

7.4.1 HARQ and Packet Losses

Figure 7.6 shows the effect of the SNR on the packet loss rate. We have represented the mean value, as well as the minimum and the maximum packet loss rate to illustrate the maximum variability for a given SNR value. It must be noted that although in some points the slope seems to change, these effects may appear because of the randomness of the propagation conditions, these magnitudes require large statistical analysis and in a limited set of experiments small deviations may appear in the results.

We have also compared the PDSCH BLER with the IP packet loss rate. The PDSCH BLER represents the ratio of correctly acknowledged transport blocks to the total number of transmitted transport blocks. As the SNR decreases, the effect of the noise makes the PDSCH BLER increase. In Figure 7.7 we correlate the PDSCH BLER with the packet loss rate. In absence of HARQ the PDSCH BLER should match the packet loss rate, but HARQ reduces the rate of packet losses at the cost of additional use of PDSCH resources to allocate retransmissions. As we are operating in relatively ideal conditions, the relation between the BLER and the effective bandwidth reduction is approximately linear as a single retransmission will succeed typically. In worse SNR conditions the ratio of lost packets rate to PDSCH BLER would be even further reduced at the expense of a more noticeable impact on the bandwidth.

Other useful magnitudes are also related with SNR and parameters. The CQI is a magnitude reported by a mobile device, with a configurable periodicity, that provides an estimation of the instantaneous quality of the channel. The larger the reported CQI, the higher coding rate (lower redundancy) that can be used for transmission. We have verified that the reported average CQI decreases consistently as the noise increases and in future work we will provide detailed results of the mapping of PDSCH BLER to reported CQI for different cell configurations and propagation conditions. Furthermore, we will also analyze different CQI adaptive schedulers, that will react to the instantaneous received CQI to provide appropriate resource allocations.

In general low R^2 values have been obtained (0.7 or less) with the equations used, which exposes the complexity of the relationships between the cross-layer parameters analyzed. In future work we will apply objective-driven simulations [80] to obtain more accurate patterns between cross-layer parameters under study.

7.4.2 Voice quality measurements

In this section we provide the voice quality results associated to the former experiments. VoIP voice quality has been calculated using the objective PESQ algorithm standardized by ITU-T. The PESQ algorithm outcome provides a quality metric mapped to the MOS. The algorithm requires injecting a known speech signal into the system under test, and the degraded output signal is compared with the original (reference) input. Values between 4 and 4.5 represent toll quality, which is the typical quality offered by the PSTN, while values below

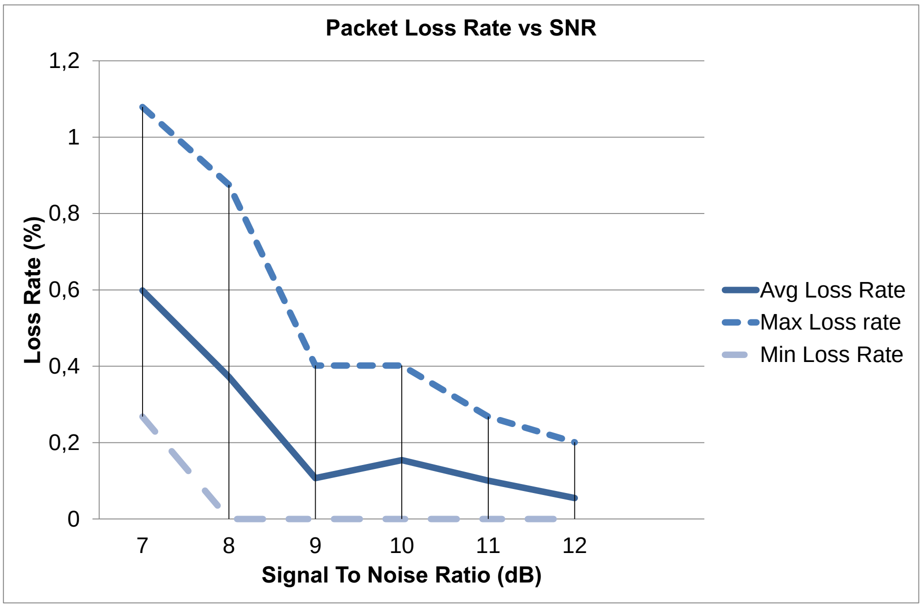

Figure 7.6: Packet Loss vs SNR

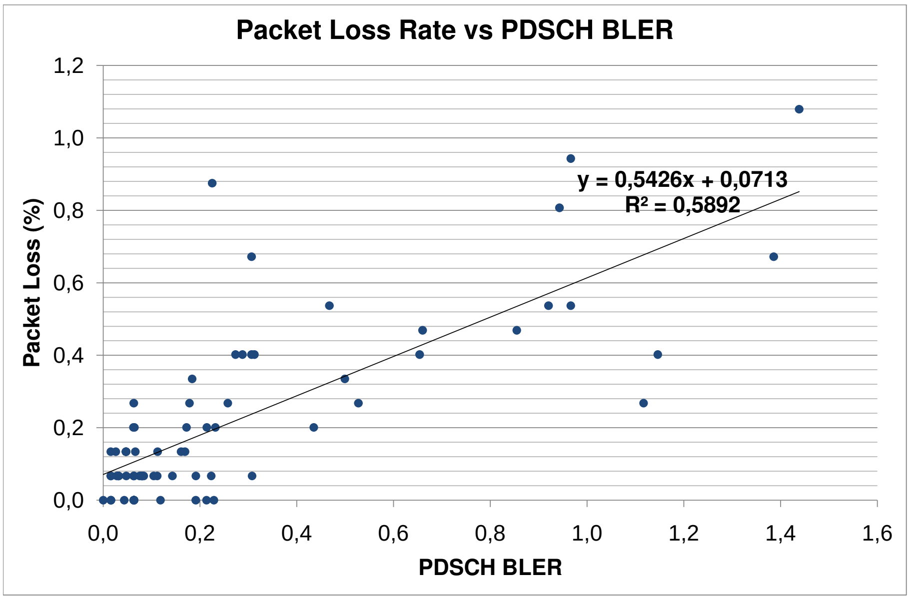

Figure 7.7: Packet Loss vs PDSCH BLER

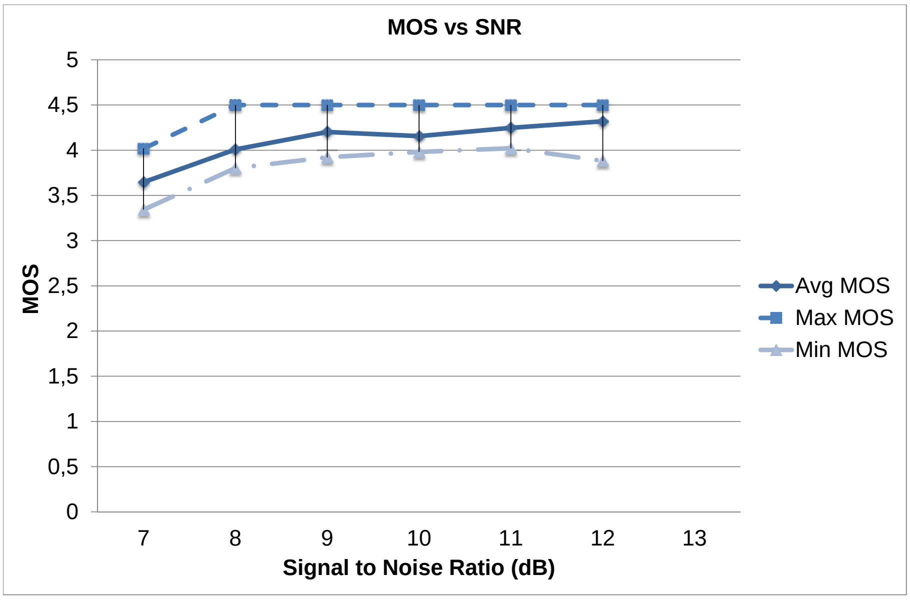

Figure 7.8: MOS vs SNR

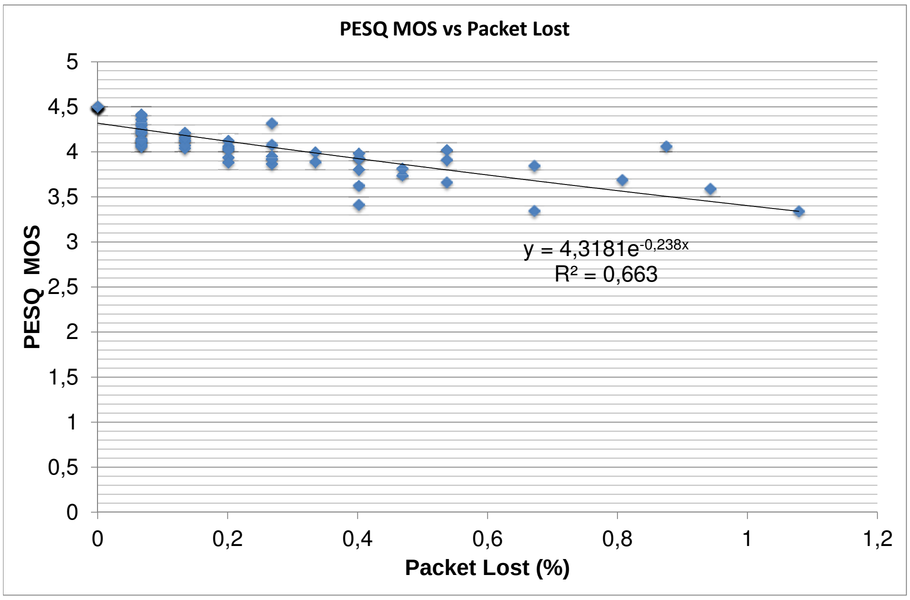

Figure 7.9: PESQ MOS vs Packt Lost

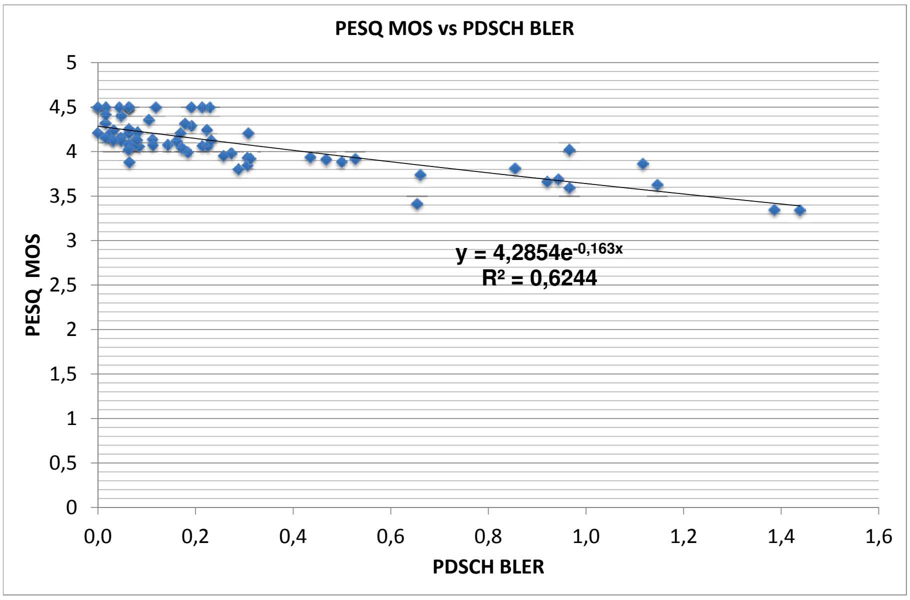

Figure 7.10: PESQ MOS vs PDSCH BLER

3,5 are often considered by some users. In Figure 7.8, we can see that an SNR of 8dB or higher produces a an average MOS higher than 4, whereas an SNR of 7dB results in an MOS of below 3,5 for some calls.

Although the mapping of MOS to packet losses has been analyzed in the literature, it is also represented in Figure 7.9 to verify that a linear estimation can be derived. Particularly, it can be verified that for packet losses close to 0% the maximum quality is reported by PESQ, whose maximum output is 4.5. It must be noted that the PESQ algorithm does not consider factors such as the end to end delay that affect the subjective quality and are considered in other methods such as the E-model.

In Figure 7.10 the mapping of MOS to the PDSCH BLER is represented. Although it will depend on the cell configuration, in future work we will consider using delay aware quality algorithms such as E-model to review the relation between MOS and PDSCH BLER, as the delay introduced by HARQ retransmissions could have also an impact on quality depending on the configuration and the end to end delay budget.

Chapter 8

Conclusions and Future Works

Within the scope of this thesis, we have made a number of contributions in two main fields. On one hand we have developed and used mobile applications to monitor VoIP communications performance in commercial mobile networks, including particularly challenging use cases as high speed railways to provide valuable insights. On the other hand we have created a controlled reference test environements for experimentation over 4G LTE, verifying also its application on the evaluation of VoIP services.

In the future, we will continue to further exploit the potential of the capabilities for network monitoring. Extending our network assessment campaigns to the comparative evaluation of 4G LTE and future technologies as they appear. Of particular interest will be the verification of VoLTE, as its deployment has recently been announced by Vodafone. Unfortunately, at the time of writing this service is not publicly available to all the users. We will also explore the comparative performance of voice services from Over The Top (OTT) providers such as Skype, Whatsapp and others.

Taking advantage of the flexibility of our tools and methodologies, we will extend also the our evaluation scope to other applications and services that relay on IP based communications

On the evolution of testbeds, we have recently analyzed the potential application of a new generation of network emulators for improved performance and extended functionality to cover a wider range of research scenarios. In the next section it can be found an initial assesment of capabilities and potential research application of the UXM wireless test set from Keysight Technologies, that represents and evolution from the test equipment we currently use.

8.1 UXM assessment for future experiments

Although sometimes it is not widely present in outside of the manufacturing and test industry, mobile phones pass through a demanding test process before receiving the approval to be used in mobile networks. For that purpose, the standardization organizations like 3GPP (3rd Generation Partnership Project) define extensive conformance specifications. Advanced test platforms such as Keysight's UXM [72] Wireless Test Set implement those tests behaving as real base stations from the point of the mobile phones being tested. In addition,

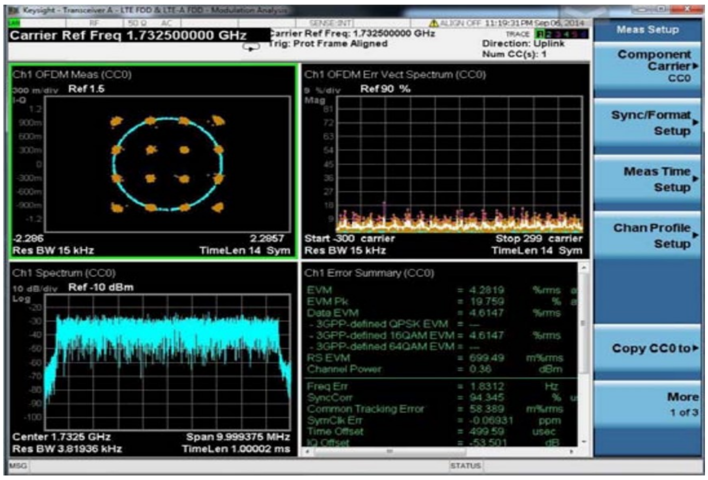

Figure 8.1: XApps Measurement tools

mobile device manufacturers and other users of UXM also benefit from a great variety of features aimed to provide a powerful and flexible R&D test environment.

Leveraging the experience gained using the E2010, we have identified many potential gains in adopting a new generation instrument like UXM to address a wider range of research areas with improved performance.

The UXM is an extremely versatile instrument, being able to emulate multiple base stations with potentially different radio access technologies, including 4G LTE and 3G WCDMA. Additionally it can operate simultaneously also as a radio channel emulator, noise and arbitrary waveform generator to generate impairments and as a signal.

One of the key UXM advantages is its ease of use of the E7530A and E7630A LTE/LTE-A Test and Lab Applications [74] that run inside the instrument itself. When being used locally, the windows based GUI can be also accessed through its large touch screen. When remotely, it can be operated via SCPI commands and also through a Remote Desktop access. Next we will briefly introduce some of the UXM key features and their potential use cases for research purposes.

8.1.1 Transmitter RF and Baseband Signal Analysis

Among its many capabilities UXM provides native support for Keysight's XApps measurement tools that are the de-facto standard for signal analysis. A wide range of measurements can be performed such as Error Vector Magnitude (EVM) measurements, spectrum analysis, equalizer flatness, IQ constellations, measurement of dynamic power control and more.

8.1.2 Receiver performance analysis

The receivers of mobile devices are typically evaluated in terms of the probability of receiving data in presence of radio impairments with an error ratio below defined thresholds.

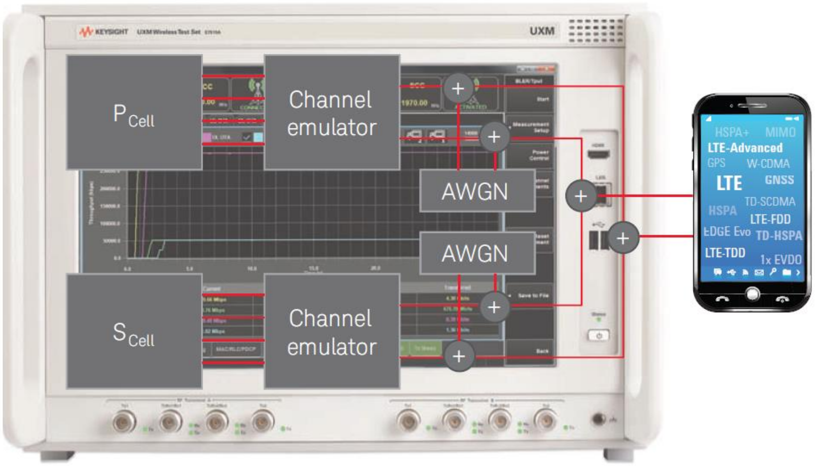

Figure 8.2: Internal Radio Channel Emulation

To replicate the effect of radio propagation conditions, the UXM features an embedded digital channel emulator that provides greater simplicity and improves accuracy by avoiding the uncertainty contributions typically involved in external RF interconnections.

In addition to the emulation of fading scenarios, other impairments including CW (Continuous Waveform), AWGN (Additive White Gaussian Noise) and arbitrary waveforms.

8.1.3 Mobility procedures

Understanding the internals of mobile network signaling protocols exchanges may be a complex task if not harnessing to real examples. Multiple cells can be generated using UXM, including different Radio Access Tecnologies (RAT), and its interfaces for operation and logging are intuitive and easy to use.

In addition to controlling RF parameters such as frequency bands, channel bandwidths, it is straight to perform and monitor many mobility procedures such as e.g. network attach, detach or cell reselections. Another notable application scenario could be the evaluation of different types of handovers (intra-frequency, inter-band, inter-band, inter-RAT) upon realistic controlled radio impairments with fading and noise.

8.1.4 End to end IP multimedia performance analysis

Smartphones are no longer expected to be used only for voice calls. Instead, IP based multimedia connections have become prevalent nowadays. Another key characteristic of current use cases is the wide variety of new different user profiles and traffic patterns.

In this scenario it is important to ensure that mobile devices can transmit data up to their maximum published capabilities. UXM high performing capabilities

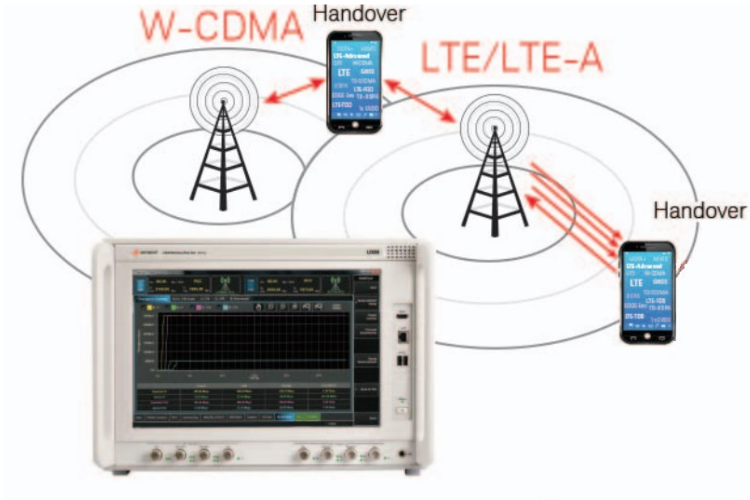

Figure 8.3: UXM driven mobility procedures

Figure 8.4: E2E IP and MAC level throughput graphs

allows testing throughputs as high as those of state of the art LTE category 11 devices (587.5 Mbps).

It is also possible to can gain insight on E2E traffic dynamics using realtime monitoring graphs. They illustrate the evolution of the IP throughput and the related MAC level throughput statistics.

This type of representation is particularly useful to detect traffic artifacts that could be related to complex cross layer interactions impact on application level communication performance.

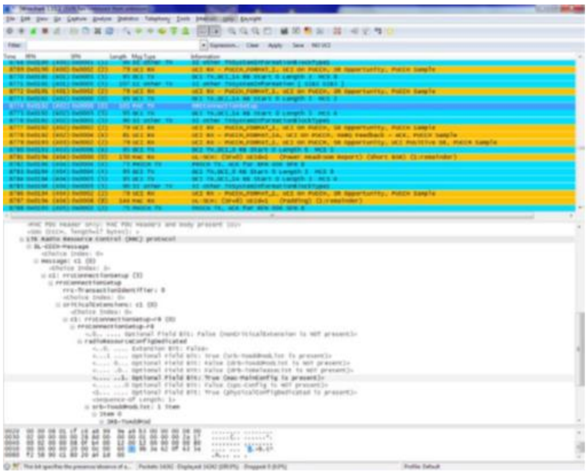

Figure 8.5: Protocol logging and analysis software

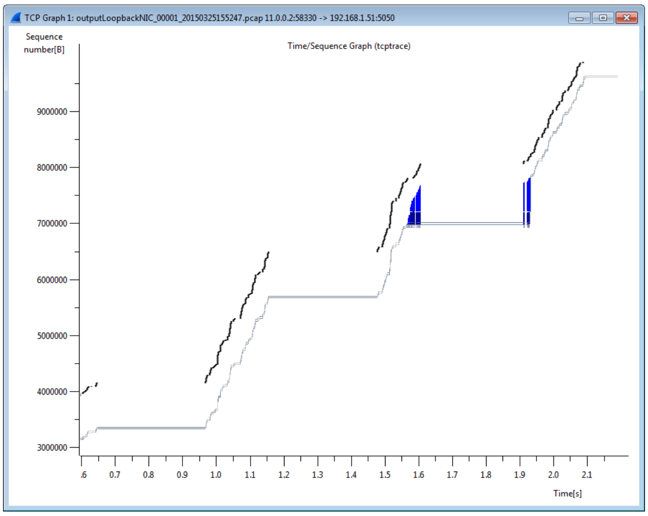

Figure 8.6: Integration with Wireshark tools for data stream analysis

8.1.5 Protocol logging and analysis

For an in depth traffic analysis, UXM can deploy a powerful protocol and logging analysis software [75], the E7515A-L01. This software allows UXM users to control the generation and filtering of logs from a wireshark integrated interface.

Complex traffic patterns can be analyzed in combination with UXM ability to emulate realistic radio propagation conditions. Of particular interest can be the use of wireshark integrated flow analysis functions to study the dyamic sevolution of TCP flows.

It is not only possible analyze upper layer signaling, but also the underlying

Figure 8.7: IMS-SIP and VoLTE tools

control information and MAC PDUs. Even a more sophisticated analysis can be obtained from the correlation of the user plane performance user plance data with the control plane.

8.1.6 VoLTE and multimedia testing

UXM provides multiple radio access features required by the VoLTE profile [21], including SPS (Semi-Persistent Scheduling), multiple DRBs (Dedicated Radio Bearers), DRX (Discontinuous Reception) and more.

As an all IP network, LTE relies on IMS (IP Multimedia Subsystem) to register IMS capable UEs and to establish multimedia sessions.

UXM also integrates an IMS-SIP server [76] that can also be combined with an IMS-SIP software client to test VoLTE and other multimedia functional scenarios.

8.1.7 Heterogenous networking

Heterogeneous networking is and has been for years a research area of tremendous activity. These scenarios typically combine multiple cells with different powers, and thus it is required to coordinate in time their transmission patterns to preserve quality particularly in the cell borders using enhanded Inter-Cell Interference Coordination (eICIC). For that purpose, mobile devices can be instructed to measure signal quality and power at different sets of time intervals. A special low power transmission scheme named ABS (Almost Blank Subframes) has also been added to reduce the interference caused to neighbor cells both on the data reception and channel quality estimation. With a correct alignment of the measurement periods to the coordinated ABS transmission patterns, it can be identified the two MCS (Modulation and Coding Schemes) that will be required both when high interference from neighbor cells is present or when their power is reduced or even fully removed in ABS.

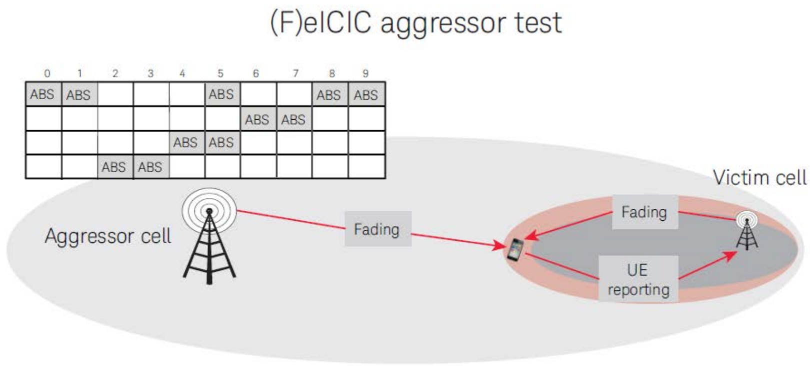

Figure 8.8: Heterogeneous networking with eICIC

The 3GPP test specifications define a set of reference patterns and cell configurations, but UXM applications provides a highly flexible interface to explore the performance with complex cell aggression scenarios. It is worth noting that multiple configurations are possible in a cell for ABS to achieve power reduction, where typically at least CRS (Cell specific Reference Signals) would be present for the convenience of the UEs camped in it. Depending on the physical cell identities of the aggressing and victim cells, their CRS could be mapped in different frequency resources. If not taken into consideration, this effect could lead to optimistic MCS estimations because the aggressing CRS in an ABS might not overlap with the victim CRS which are used for the channel estimation but it will definitely overlap with the victim frequency resources for data.

8.1.8 Additional Applications

There are many other potential applications, but the ability to transmit in unlicensed spectrum and the integration with energy measurement equipment are of particular interest.

Appendix A

VoIP measurements under different fading profiles

Signals transmitted in mobile radio channels suffer from different propagation related effects such as "fading".

The multipath propagation conditions consist of several parts:

- A delay profile in the form of a "tapped delay-line", characterized by a number of taps at fixed positions on a sampling grid. The profile can be further characterized by the root mean square (r.m.s.) delay spread and the maximum delay spanned by the taps.
- A Doppler spectrum, characterized by a spectrum shape and a maximum Doppler frequency that is determined from the mobile speed.
- A set of correlation matrices defining the correlation between the UE and BS antennas in case of multi-antenna systems.

Channel models are defined by combining a delay profile with a Doppler spectrum, with the addition of correlation properties in case of a multi-antenna scenario.

	Number of paths	Delay spread (r.m.s)	Maximum delay
Extended Pedestrian A (EPA)	7	45 ns	410 μs
Extended Vehicular A (EVA)	9	357 ns	2.51 μs
Extended Typical Urban (ETU)	9	991 ns	5μs

Table 1: Summary of delay profiles for LTE channel models [3GPP TS 36.803]

Path Number	Extended Pedestrian A (EPA)		Extended Vehicular A (EVA)		Extended Typical Urban (ETU)	
	Delay (ns)	Power (dB)	Delay (ns)	Power (dB)	Delay (ns)	Power (dB)
1	0	0	0	0	0	-1
2	30	-1	30	-1.5	50	-1
3	70	-2	150	-1.4	120	-1
4	90	-3	310	-3.6	200	0
5	110	-8	370	-0.6	230	0
6	190	-17.2	710	-9.1	500	0
7	410	-20.8	1090	-7	1600	-3
8			1730	-12	2300	-5
9			2510	-16.9	5000	-7

Table 2: Tapped delay line models [3GPP TS 36.803]

Maximum doppler frequency	Corresponding UE speed
5 Hz	3.1 km/h
70 Hz	43 km/h
300 Hz	185 km/h

Table 3: Uplink doppler frequencies and corresponding UE speeds (Operating Band 3: 1710 - 1785 MHz)

Maximum doppler frequency	Corresponding UE speed
5 Hz	2.9 km/h
70 Hz	41 km/h
300 Hz	176 km/h

Table 4: Downlink doppler frequencies and corresponding UE speeds (Operating Band 3: 1805 - 1880 MHz)

Delay profiles

The delay profiles are selected to be representative of low, medium and high delay spread environments. The profiles for low and medium delay spread are based on the ITU Pedestrian A and Vehicular A channel models respectively, originally defined for the ITU-R evaluation of IMT-2000 [102]. The high delay spread model is based on the Typical Urban model used for GSM [81] and in some of the evaluation work for LTE. The resulting model parameters are summarized in Table 1 and Table 2. The models are defined on a [10 ns] sampling grid. They can be adapted to any desired sampling grid used in a simulation or test setup using the procedure defined to align sampling grids shown in Annex B of TR 25.943 [82].

Delay spread	Doppler frequency	Model	Comment
Low	Low	EPA 5Hz	Low delays spread model representing small cell and indoor cases.
Medium	Low	EVA 5Hz	
Medium	Medium	EVA 70Hz	
High	Medium	ETU 70Hz	Represents high delay spread environments, with a delay span of the same order as the cyclic prefix
High	High	ETU 300Hz	

Table 5: LTE Radio Channel Models defined by the 3GPP

Doppler frequency

A set of three Doppler frequencies spanning the requirement range as high, middle and low Doppler frequencies is selected in TR 36.803 [84]:

- Common high speed scenarios for moderately high mobile speeds. It is stated in TR 25.913 [85] that high performance should be maintained up to mobile speeds of 120 km/h. The corresponding maximum Doppler frequency for fc =2690 km/h is fD =299 Hz. Based on this, the high Doppler frequency is selected as 300 Hz.

- TR 25.913 also state that "The E-UTRAN shall support mobility across the cellular network and should be optimized for low mobile speed from 0 to 15 km/h." For testing purposes, too low mobile speeds are not attractive, since testing times may be very long. The lowest Doppler frequency in UTRA propagation conditions is 5.4 Hz, corresponding to between 2.3 and 7 km/h in the existing frequency bands. Based on this, the low Doppler frequency is selected as 5 Hz.

- An intermediate Doppler frequency can be set at the "logarithmic" average of the 5 and 900 Hz, being 67 Hz. Based on this, the medium Doppler frequency is selected as 70 Hz.

- The LTE requirements for mobility in TR 25.913 state that "Mobility across the cellular network shall be maintained at speeds from 120 km/h to 350 km/h (or even up to 500 km/h depending on the frequency band". This special case is called High speed train scenario and it is not contemplated in this work.

The UE speed that the Doppler frequencies will correspond to will vary between the Operating bands. Table 3 and 4 shown the corresponding UE speeds for the carrier frequencies at the centre of each uplink and downlink for band 3. Informative values for the rest of band can be consulted in [84].

In accordance with these parameters several multi-path models for cellular systems have been specified for low, medium and high delay spread environments as shown in Table 5.

ETU 70

Duplex Mode	FDD
Frequency band	Band 3
DL Channel EARFCN	1575
DL Bandwith	15 MHz
Cell Power Level	-90.0
UL Channel EARFCN	19575
UL Bandwith	15 MHz
Cell ID	2
Cyclic Prefix	Normal
Channel Model	ETU70
AWGN Generated SNR (dBm15kHz)	-95 dBm15kHz

Table 6: Configuration summary: ETU 70;Cell power -90 dBm/kHz; Noise power -95 dBm/kH; 6 HARQ retransmissions

DL Relative ACK Number9(%)	96.77.94
DL Absolute ACK Number	86457
DL Relative NACK Number (%)	3.23
DL Absolute NACK Number	2883
DL Relative statDTX Number (%)	0.00
DL Absolute statDTX Number	0
DL PDSCH Relative BLER (%)	3.23
DL PDSCH Maximum BLER (%)	6.64
DL PDCCH Relative BLER (%)	0.00
DL PDCCH Maximum BLER (%)	0.00
DL Average Throughput (Kbps)	80.81
DL Minimum Throughput (Kbps)	0.00
DL Maximum Throughput (Kbps)	1838.40
DL Throughput Limit (Kbps)	4944.00
Minimum CQI	5
Maximum CQI	12
Average CQI	6
Medium CQI	6
CQI Index Range to Check	From 8 to 15
CQI Reports in Range (%)	28.69
RSRP (dBm)	-92
RSRQ (dB)	-10
Frames for BLER Meas	0
Num. CQI Reports for Meas	0

Table 7: RX measurement summary: ETU 70; Cell power -80 dBm/kHz; Noise power -90 dBm/kHz; 6 HARQ retransmissions

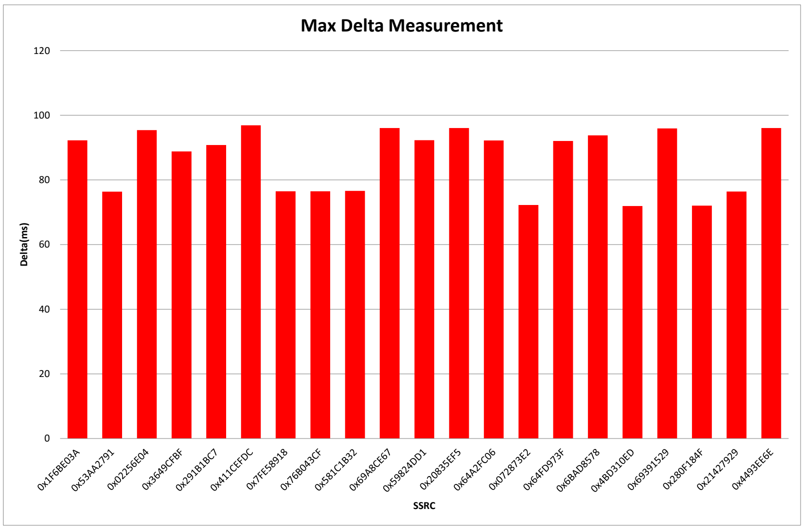

Figure 9: IP measurement: Delta (ETU 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions)

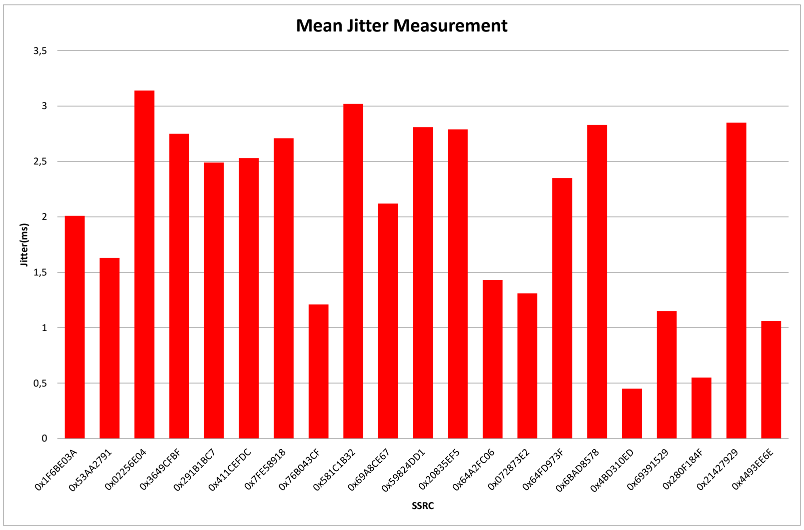

Figure 10: IP measurement: Jitter (ETU 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions)

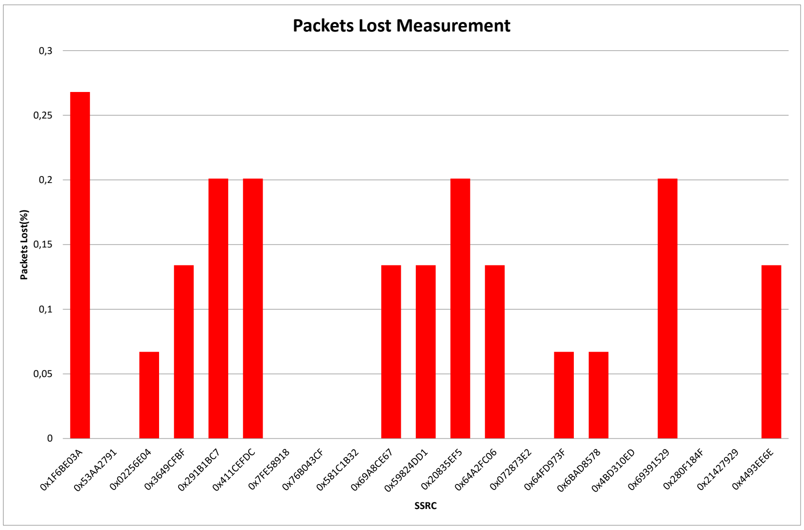

Figure 11: IP measurement: Packet lost (ETU 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions)

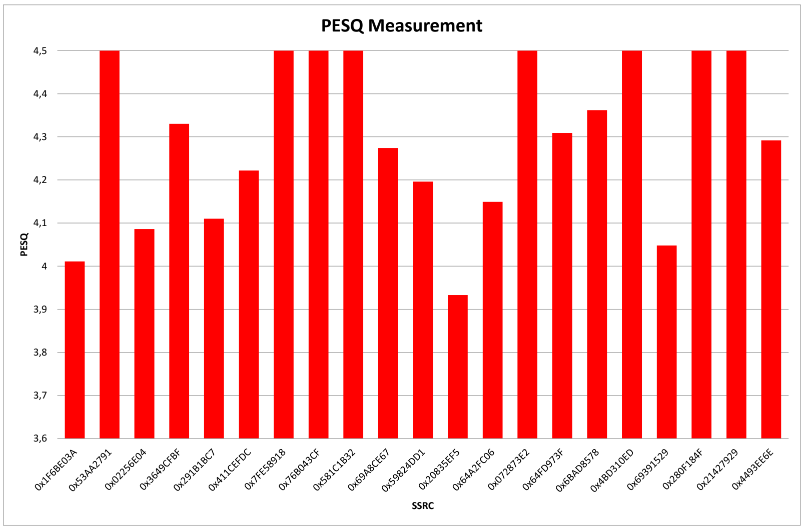

Figure 12: Voice quality: PESQ (ETU 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions)

EVA 70

Duplex Mode	FDD
Frequency band	Band 3
DL Channel EARFCN	1575
DL Bandwith	15 MHz
Cell Power Level	-90.0
UL Channel EARFCN	19575
UL Bandwith	15 MHz
Cell ID	2
Cyclic Prefix	Normal
Channel Model	EVA70
AWGN Generated SNR (dBm15kHz)	-95 dBm15kHz

Table 8: Configuration summary: EVA 70;Cell power -90 dBm/kHz; Noise power -95 dBm/kH;6 HARQ retransmissions

DL Relative ACK Number (%)	96.75
DL Absolute ACK Number	90214
DL Relative NACK Number (%)	3.25
DL Absolute NACK Number	3026
DL Relative statDTX Number (%)	0.00
DL Absolute statDTX Number	0
DL PDSCH Relative BLER (%)	3.25
DL PDSCH Maximum BLER (%)	75.00
DL PDCCH Relative BLER (%)	0.00
DL PDCCH Maximum BLER (%)	0.00
DL Average Throughput (Kbps)	142.40
DL Minimum Throughput (Kbps)	0.00
DL Maximum Throughput (Kbps)	1288.32
DL Throughput Limit (Kbps)	4944.00
Minimum CQI	5
Maximum CQI	9
Average CQI	6
Medium CQI	6
CQI Index Range to Check	From 8 to 15
CQI Reports in Range (%)	21.19
RSRP (dBm)	-92
RSRQ (dB)	-9.5
Frames for BLER Meas	0
Num. CQI Reports for Meas	0

Table 9: RX measurement summary: EVA 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions

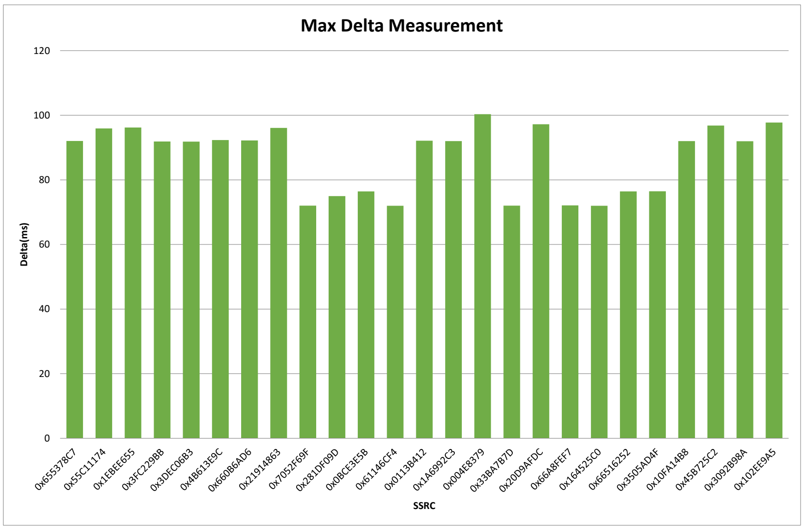

Figure 13: IP measurement: Delta (EVA 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions)

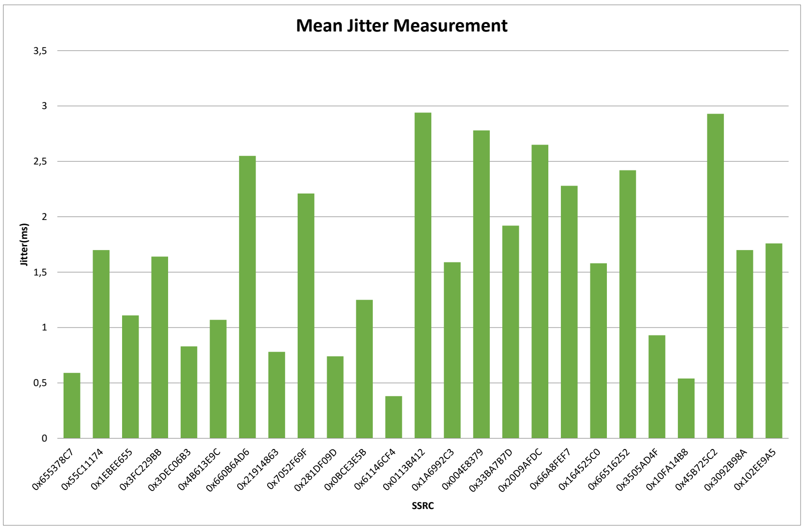

Figure 14: IP measurement: Jitter (EVA 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions)

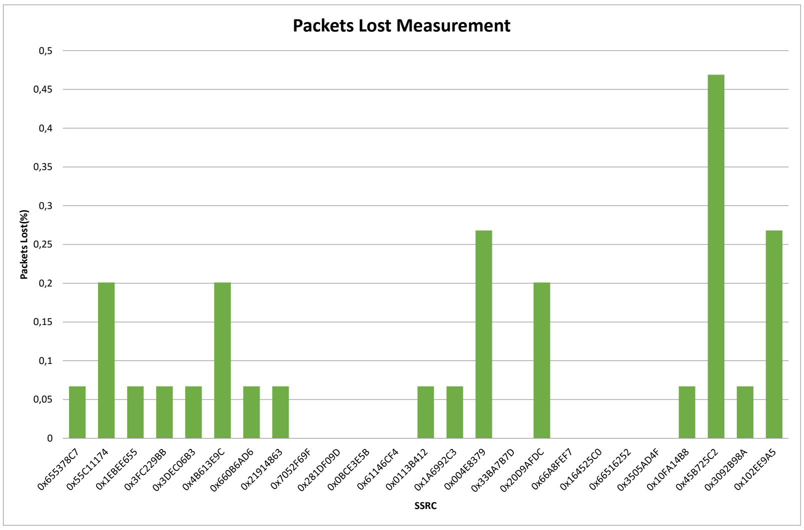

Figure 15: IP measurement: Packet lost (EVA 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions)

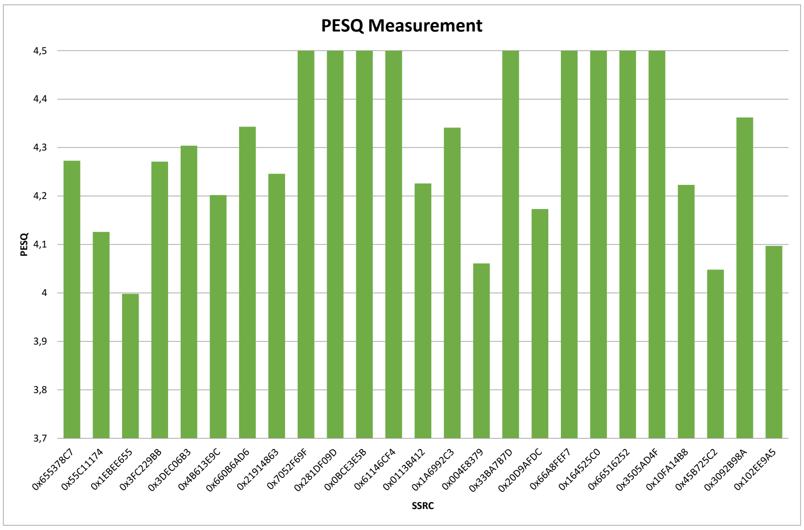

Figure 16: Voice quality: PESQ (EVA 70; Cell power -90 dBm/kHz; Noise power -95 dBm/kHz;6 HARQ retransmissions)

EVA 5

Duplex Mode	FDD
Frequency band	Band 3
DL Channel EARFCN	1575
DL Bandwith	15 MHz
Cell Power Level	-85.0
UL Channel EARFCN	19575
UL Bandwith	15 MHz
Cell ID	2
Cyclic Prefix	Normal
Channel Model	EVA5
AWGN Generated SNR (dBm15kHz)	-95 dBm15kHz

Table 10: Configuration summary: EVA 5;Cell power -85 dBm/kHz; Noise power -95 dBm/kH; 3 HARQ retransmissions

DL Relative ACK Number (%)	83.86
DL Absolute ACK Number	2384699
DL Relative NACK Number (%)	16.14
DL Absolute NACK Number	458809
DL Relative statDTX Number (%)	0.00
DL Absolute statDTX Number	0
DL PDSCH Relative BLER (%)	16.14
DL PDSCH Maximum BLER (%)	59.68
DL PDCCH Relative BLER (%)	0.00
DL PDCCH Maximum BLER (%)	0.00
DL Average Throughput (Kbps)	3900.08
DL Minimum Throughput (Kbps)	395.52
DL Maximum Throughput (Kbps)	4696.80
DL Throughput Limit (Kbps)	4944.00
Minimum CQI	6
Maximum CQI	14
Average CQI	8
Medium CQI	8
CQI Index Range to Check	From 8 to 15
CQI Reports in Range (%)	87.62
RSRP (dBm)	-88
RSRQ (dB)	-9.5
Channel Model	EPA5
AWGN Generated SNR (dBm/15kHz)	-95 dBm/15kHz
Frames for BLER Meas	0
Num. CQI Reports for Meas	0

Table 11: RX measurement summary: EVA 5; Cell power -85 dBm/kHz; Noise power -95 dBm/kHz; 3 HARQ retransmissions

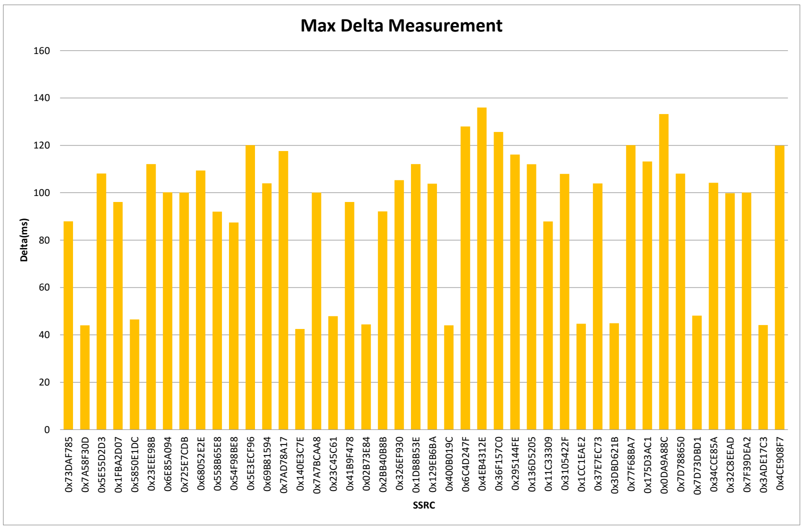

Figure 17: IP measurement: Delta (EVA 5; Cell power -85 dBm/kHz; Noise power -95 dBm/kHz; 3 HARQ retransmissions)

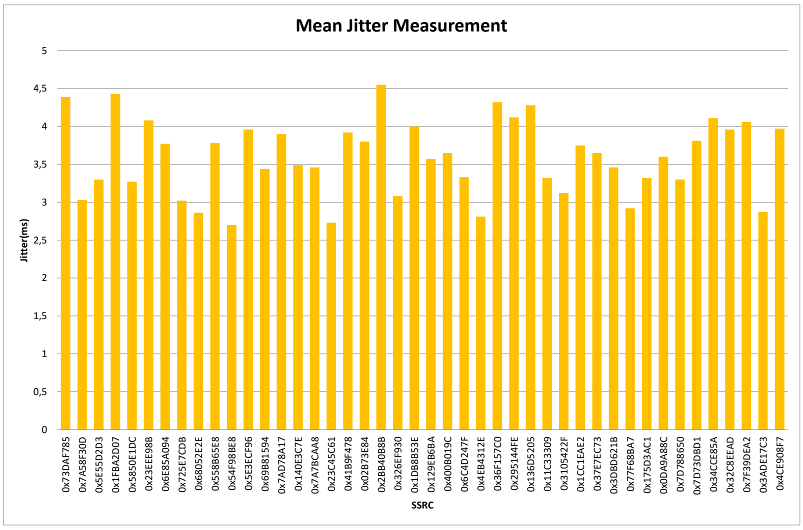

Figure 18: IP measurement: Jitter (EVA 5; Cell power -85 dBm/kHz; Noise power -95 dBm/kHz; 3 HARQ retransmissions)

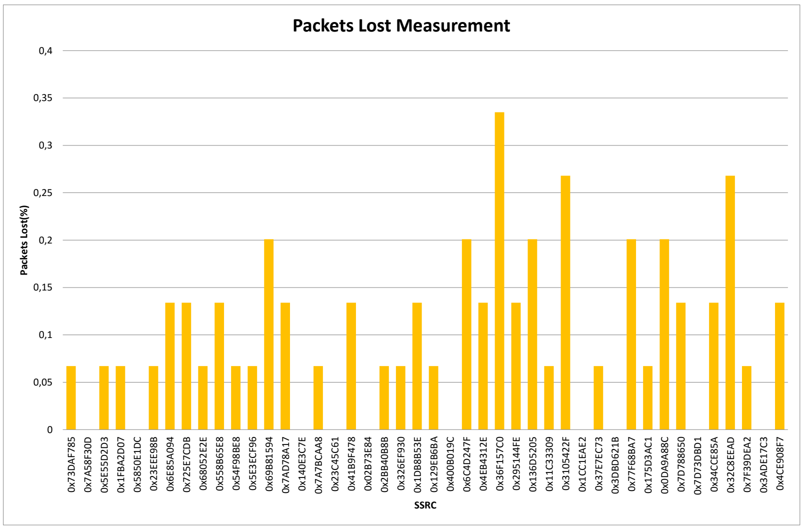

Figure 19: IP measurement: Packet lost (EVA 5; Cell power -85 dBm/kHz; Noise power -95 dBm/kHz; 3 HARQ retransmissions)

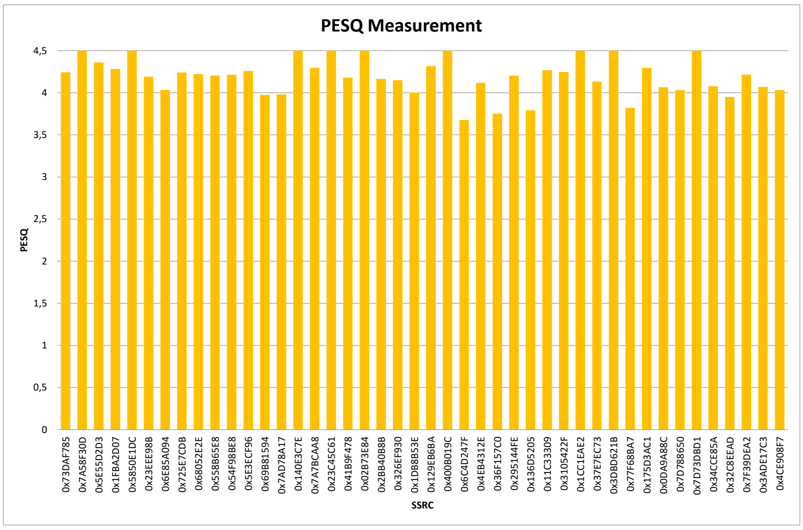

Figure 20: Voice quality: PESQ (EVA 5; Cell power -85 dBm/kHz; Noise power -95 dBm/kHz; 3 HARQ retransmissions)

EPA 5

Duplex Mode	FDD
Frequency band	Band 3
DL Channel EARFCN	1575
DL Bandwith	15 MHz
Cell Power Level	-90.0
UL Channel EARFCN	19575
UL Bandwith	15 MHz
Cell ID	2
Cyclic Prefix	Normal
Channel Model	EPA5
AWGN Generated SNR (dBm15kHz)	-95 dBm15kHz

Table 12: Configuration summary: EPA 5; Cell Power: -85 dBm/15kHz Noise Power: -95dBm/15kHz; 3 HARQ retransmissions

DL Relative ACK Number (%)	80.36
DL Absolute ACK Number	1839173
DL Relative NACK Number (%)	19.64
DL Absolute NACK Number	449369
DL Relative statDTX Number (%)	0.00
DL Absolute statDTX Number	0
DL PDSCH Relative BLER (%)	19.64
DL PDSCH Maximum BLER (%)	26.48
DL PDCCH Relative BLER (%)	0.00
DL PDCCH Maximum BLER (%)	0.00
DL Average Throughput (Kbps)	3737.31
DL Minimum Throughput (Kbps)	0.00
DL Maximum Throughput (Kbps)	4696.80
DL Throughput Limit (Kbps)	4944.00
Minimum CQI	5
Maximum CQI	14
Average CQI	9
Medium CQI	9
CQI Index Range to Check	From 8 to 15
CQI Reports in Range (%)	87.62
RSRP (dBm)	-88
RSRQ (dB)	-9.5
Frames for BLER Meas	0
Num. CQI Reports for Meas	0

Table 13: RX measurement summary: EPA 5; Cell Power: -90 dBm/15kHz Noise Power: -95dBm/15kHz; 3 HARQ retransmissions

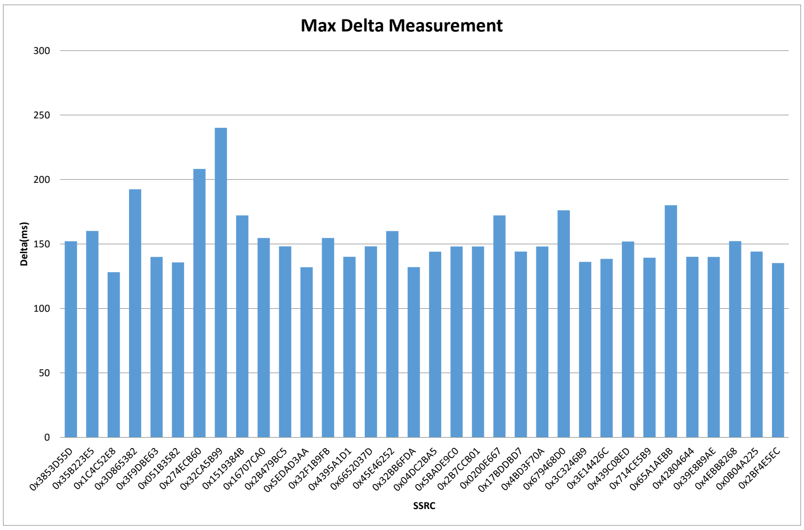

Figure 21: IP measurement: Delta (EPA 5; Cell Power: -85 dBm/15kHz Noise Power: -95dBm/15kHz; 3 HARQ restransmissions)

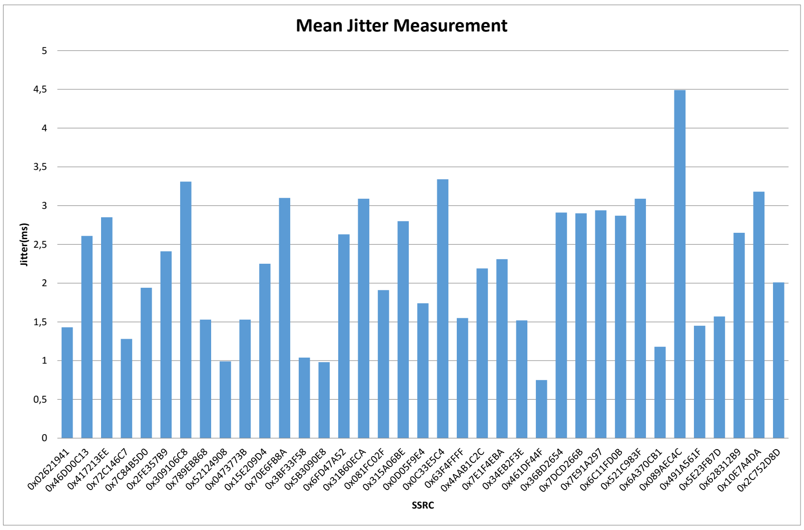

Figure 22: IP measurement: Jitter (EPA 5; Cell Power: -85 dBm/15kHz Noise Power: -95dBm/15kHz; 3 HARQ restransmissions)

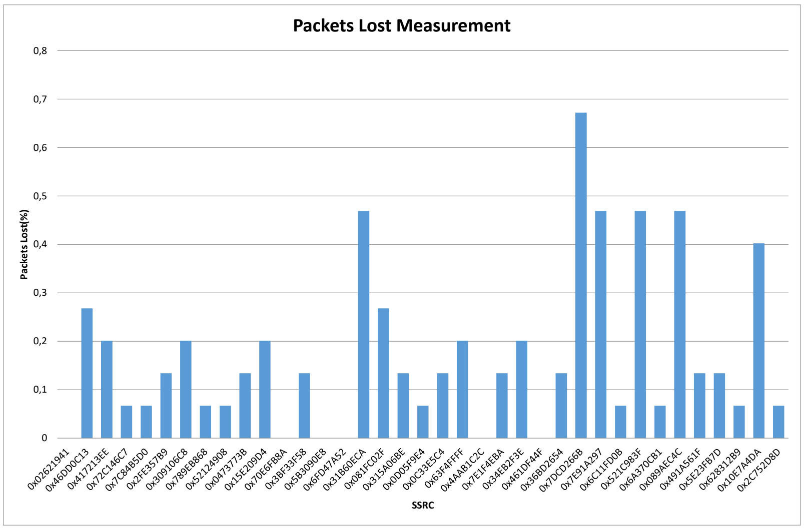

Figure 23: IP measurement: Packet lost (EPA 5; Cell Power: -85 dBm/15kHz Noise Power: -95dBm/15kHz;3 HARQ restransmissions)

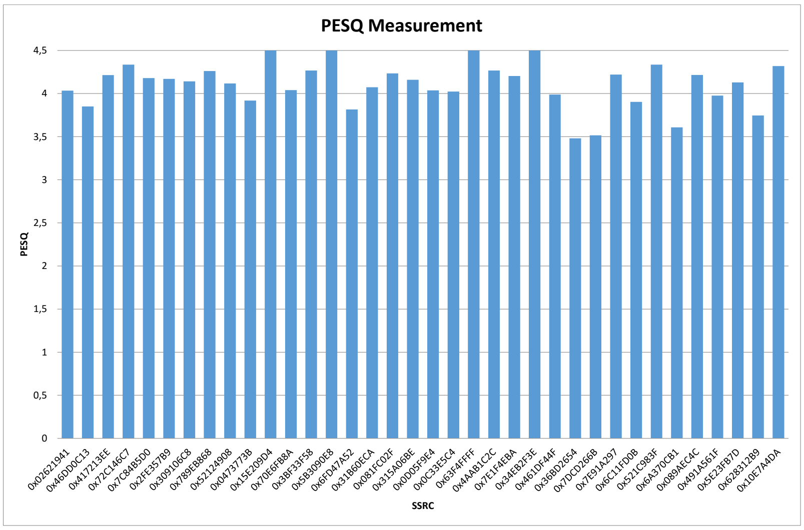

Figure 24: Voice quality: PESQ (EPA 5; Cell Power: -90 dBm/15kHz Noise Power: -95dBm/15kHz)

Publications

The work presented in this thesis has been internationally published in different conferences and journals.

International journals

1. Francisco Javier Rivas Tocado, Almudena Díaz Zayas, Pedro Merino Gómez, "Characterizing Traffic Performance in Cellular Networks," Internet Computing, IEEE , vol.18, no.1, 2014

   (a) For 2014, the journal IEEE INTERNET COMPUTING has an Impact Factor of 1.713 in the category COMPUTER SCIENCE, SOFTWARE ENGINEERING

      i. Total Journals 104

      ii. Category Journal Rank 16

      iii. Category Quartile Q1

2. Francisco Javier Rivas Tocado, Almudena Díaz Zayas, Pedro Merino Gómez, "Obtaining more realistic cross-layer QoS measurements: A VoIP over LTE use case", Journal of Computer Networks and Communications, Hindawi, vol. 2013, 2013 (Journal of Computer Networks and Communications currently has an acceptance rate of 15%.)

Book chapters

2. Francisco Javier Rivas Tocado, Almudena Díaz Zayas, Pedro Merino Gómez,"Testing LTE configurations and applications", Lecture Notes of the institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer, vol. 44, 2012

3. Francisco Javier Rivas Tocado, Almudena Díaz Zayas, Pedro Merino Gómez, "UMA Testing Facility", Lecture Notes of the institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer, vol 44, 2012

International conferences

1. Francisco Javier Rivas Tocado, "Performance study of internet traffic on high speed railways", 14th International Symposium and Workshops on A World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2013

2. Andrés Álvarez Muñiz, Almudena Díaz Zayas, Pedro Merino Gómez, Francisco Javier Rivas Tocado, "Field measurements of mobile services with Android smartphones", Consumer Communications and Networking Conference (CCNC), IEEE, 2012

3. Andrés Álvarez Muñiz, Almudena Díaz Zayas, Pedro Merino Gómez, Francisco Javier Rivas Tocado, "Mobile Application profiling with TestelDroid", Consumer Communications and Networking Conference (CCNC), IEEE, 2012 (BEST DEMO AWARD)

4. Almudena Díaz Zayas, Pedro Merino Gómez, Francisco Javier Rivas Tocado, "Test environment for QoS testing of VoIP over LTE", Network Operations and Management Symposium (NOMS), IEEE, 2012

Bibliography

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012—2017 ," tech. rep., 2013.

[2] Ericsson, "Mobility report", tech. rep., Jun. 2005.

[3] R. Kreher, *Performance Measurement: A Practical Guide to KPIs for the UTRAN Environment.* Wiley, 2006.

[4] C. Bontu and E. Illidge, "DRX mechanism for power saving in LTE," *IEEE Communications Magazine*, vol. 47, no. 6, pp. 48–55, 2009.

[5] D. Jiang, H. Wang, E. Malkamaki, and E. Tuomaala, "Principle and Performance of Semi-Persistent Scheduling for VoIP in LTE System," in *Proc. Int. Conf. Wireless Communications, Networking and Mobile Computing WiCom 2007*, pp. 2861–2864, 2007.

[6] X. Mang, Y. Levy, C. Johnson, and D. Hoeflin, "Admission control for VoIP calls with heterogeneous codecs," in *Proc. 18th Annual Wireless and Optical Communications Conf. WOCC 2009*, pp. 1–4, 2009.

[7] O. Fresan, T. Chen, K. Ranta-aho, and T. Ristaniemi, "Dynamic packet bundling for VoIP transmission over Rel'7 HSUPA with 10ms tti length," in *Proc. 4th Int. Symp. Wireless Communication Systems ISWCS 2007*, pp. 508–512, 2007.

[8] J. Puttonen, T. Henttonen, N. Kolehmainen, K. Aschan, M. Moisio, and P. Kela, "Voice-Over-IP Performance in UTRA Long Term Evolution Downlink," in *Proc. IEEE Vehicular Technology Conf. VTC Spring 2008*, pp. 2502–2506, 2008.

[9] R. Susitaival and M. Meyer, "LTE Coverage Improvement by TTI Bundling," in *Proc. IEEE 69th Vehicular Technology Conf. VTC Spring 2009*, pp. 1–5, 2009.

[10] "NGMN Top OPE Recommendations Version 1.0.21," tech. rep., NGMN Alliance, 2010.

[11] A. Diaz, P. Merino, and F. J. Rivas, "Test environment for QoS testing of VoIP over LTE," in *Proc. IEEE Network Operations and Management Symp. (NOMS)*, pp. 780–794, 2012.

[12] A. Diaz, P. Merino, and F. J. Rivas, "Mobile Application Profiling for Connected Mobile Devices," *IEEE Pervasive Computing*, vol. 9, pp. 54–61, 2010.

[13] A. Diaz, P.Merino, and F. Rivas, "QoS Analysis of Video Streaming Service in Live Cellular Networks," *Computer Communications, Elsevier, (to be published)*.

[14] "SwissQual - The independent network quality company."

[15] F. J. Rivas, A. Diaz, and P. Merino, "Performance study of internet traffic on high speed railways," in *Proc. The 14 th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WOWMOM)*, 2013.

[16] B. Wang, K. Pedersen, T. Kolding, and P. Mogensen, "Performance of voip on hsdpa," in *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, vol. 4, pp. 2335–2339 Vol. 4, May 2005.

[17] K. Aho, I. Repo, T. Nihtila, and T. Ristaniemi, "Analysis of voip over hsdpa performance with discontinuous reception cycles," in *Information Technology: New Generations, 2009. ITNG '09. Sixth International Conference on*, pp. 1190–1194, April 2009.

[18] R. Shreevastav, C. McGoldrick, and M. Huggard, "Delivering improved qos and cell throughput in umts based hsdpa networks," in *World of Wireless, Mobile and Multimedia Networks Workshops, 2009. WoWMoM 2009. IEEE International Symposium on a*, pp. 1–9, June 2009.

[19] M. Jurvansuu, J. Prokkola, M. Hanski, and P. Perala, "Hsdpa performance in live networks," in *Communications, 2007. ICC '07. IEEE International Conference on*, pp. 467–471, June 2007.

[20] 3GPP, "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Circuit Switched (CS) fallback in Evolved Packet System (EPS); Stage 2," "TS" 23.272.

[21] "Permanent Reference Document (PRD) IR.92 IP Multimedia Subsystem (IMS) Profile for Voice and SMS."

[22] V. Forum, "Voice over LTE via Generic Access; Requirements Specification; Phase1," TS 1.4.0, 2010.

[23] ETSI, "Digital cellular telecommunications system (Phase 2+); Generic Access Network (GAN); Stage 2 ," TS 43.318, 2012.

[24] M. Forum, "MSF VoLTE Interoperability Event 2011. Multivendor testing in global LTE and IMS Networks," whitepaper, 2011.

[25] "LTE 'Real World' Performance Study. Broadband and Voice over LTE (VoLTE) Quality Analysis: TeliaSonera," tech. rep., Epitiro, 2011.

[26] 3GPP, "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Services and service capabilities," TS 22.105.

[27] 3GPP, "Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction," TS 26.114.

[28] A. T. Harri Holma, *LTE for UMTS: Evolution to LTE-Advanced.* Wiley, 2011.

[29] 3GPP, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification," TS 36.331.

[30] 3GPP, "LTE; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access," TS 23.401.

[31] 3GPP, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2," TS 36.300.

[32] 3GPP, "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture," TS 23.203.

[33] S. Jelassi, G. Rubino, H. Melvin, H. Youssef, and G. Pujolle, "Quality of experience of voip service: A survey of assessment approaches and open issues," *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, pp. 1 –23, 2012.

[34] Scheets, G.; Parperis, M.; Singh, R., "Voice over the internet: A tutorial discussing problems and solutions associated with alternative transport," in Communications Surveys and Tutorials, IEEE , vol.6, no.2, pp.1-10, Second Quarter 2004

[35] C. a. O. A. Sfairopoulou, B. Bellalta, "A Comparative Survey of Adaptive Codec Solutions for VoIP overMultirateWLANs: A Capacity versus Quality Performance Trade-Off," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, pp. 2050–2090, 2011.

[36] D. Bonfiglio, M. Mellia, M. Meo, and D. Rossi, "Detailed Analysis of Skype Traffic," *IEEE Transactions on Multimedia*, vol. 11, pp. 117–127, Jan. 2009.

[37] M. Cardenete-suriol, J. Mangues-bafalluy, Álvaro Masó, M. Gorricho, M. Cardenete-suriol, J. Mangues-bafalluy, Álvaro Masó, and M. Gorricho, "Characterization and comparison of Skype behavior in wired and wireless network scenarios," in *Global Telecommunications Conference, 2007. GLOBECOM 07. IEEE* , pp. 2167–2172, 26-30 Nov. 2007.

[38] K. Tutschku, T. Hoßfeld, A. Binzenhoefer, and M. Fiedler, "Measurement and analysis of skype voip traffic in 3g umts systems," in *Proceedings of the 4th International Workshop on Internet Performance, Simulation, Monitoring and Measurement (IPS-MoMe 2006)*, February 2006.

[39] A. B. Tobias Hoßfeld *, "Analysis of Skype VoIP traffic in UMTS: End-to-end QoS and QoE measurements," *Computer Networks*, vol. 52, pp. 650–666, Dec. 2007.

[40] K.-T. Chen, C.-Y. Huang, P. Huang, and C.-L. Lei, "Quantifying skype user satisfaction," in *Proceedings of the 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, SIGCOMM '06, (New York, NY, USA), pp. 399–410, ACM, 2006.

[41] L. Carvalho, E. Mota, R. Aguiar, A. Lima, and J. De Souza, "An e-model implementation for speech quality evaluation in voip systems," in *Computers and Communications, 2005. ISCC 2005. Proceedings. 10th IEEE Symposium on*, pp. 933–938, June 2005.

[42] D. Schlosser, M. Jarschel, V. Burger, and R. Pries, "Monitoring the user perceived quality of silk-based voice calls," in *Telecommunication Networks and Applications Conference (ATNAC), 2010 Australasian*, pp. 78–83, Oct 2010.

[43] F.-N. P. Stylianos Karapantazis, "VoIP: A comprehensive survey on a promising technology," *Computer Networks*, vol. 53, pp. 2050–2090, Jan. 2009.

[44] "Dtr/tiphon-05001, telecommunications and internet protocol harmonization over networks (tiphon); general aspects of quality of service (qos), tr 101 329," TR 101329, ETSI, Oct, 1998.

[45] D. Houck and G. Meempat, "Call admission control and load balancing for voice over {IP}," *Performance Evaluation*, vol. 47, no. 4, pp. 243 – 253, 2002. Internet Performance and Control of Network Systems.

[46] R. Reynolds and A. Rix, "Quality voip - an engineering challenge," *BT Technology Journal*, vol. 19, no. 2, pp. 23–32, 2001. cited By 31.

[47] T. Uhl, "Quality of service in voip communication," *AEU - International Journal of Electronics and Communications*, vol. 58, no. 3, pp. 178–182, 2004. cited By 9.

[48] C. Perkins, O. Hodson, and V. Hardman, "A survey of packet loss recovery techniques for streaming audio," *Network, IEEE*, vol. 12, pp. 40–48, Sept 1998.

[49] "Recommendation p.800, methods for subjective determination of transmission quality," tr, ITU-T, Aug, 1996.

[50] TSB-116-A: Voice Quality Recommendations for IP Telephony, TIA, March 2006

[51] P. S. Lennart Norell, Eric Parsons, "Telephony services over LTE end-to-end," *ERICSSON review*, vol. 1, pp. 34–38, 2010.

[52] Ascom, "Tems pocket - a complete measurement smartphone system in your hand." `{http://www.ascom.com/}`, Apr. 2013.

[53] Anite, "Nemo Handy - Network Measurement and Troubleshooting on the go." `http://www.anite.com/businesses/network-testing/products/nemo-handy`, Jan. 2013.

[54] SwissQual, "Qualipoc android brochure." `http://android.swissqual.com/docs/android-brochure.pdf`, 2012.

[55] J. Huang, Q. Xu, B. Tiwana, Z. M. Mao, M. Zhang, and P. Bahl, "Anatomizing application performance differences on smartphones," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*, MobiSys '10, (New York, NY, USA), pp. 165–178, ACM, 2010.

[56] C. Shepard, A. Rahmati, C. Tossell, L. Zhong, and P. Kortum, "Livelab: measuring wireless networks and smartphone users in the field," *SIGMETRICS Perform. Eval. Rev.*, vol. 38, pp. 15–20, Jan. 2011.

[57] L. Zhang, B. Tiwana, Z. Qian, Z. Wang, R. P. Dick, Z. M. Mao, and L. Yang, "Accurate online power estimation and automatic battery behavior based power model generation for smartphones," in *Proceedings of the eighth IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis*, CODES/ISSS '10, (New York, NY, USA), pp. 105–114, ACM, 2010.

[58] A. Alvarez, A. Diaz, P. Merino, and F. J. Rivas, "Field measurements of mobile services with android smartphones," in *Proc. IEEE Consumer Communications and Networking Conf. (CCNC)*, pp. 105–109, 2012.

[59] H. Sshulzrinne, S. Carner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applicacionts," rfc 3550, IETF, July 2003.

[60] M. Aguado, O. Onandi, P. Agustin, M. Higuero, and E. Jacob Taquet, "Wimax on rails," *Vehicular Technology Magazine, IEEE*, vol. 3, pp. 47 –56, sept. 2008.

[61] M. Edwards, I. Donelson, J., W. Zavis, A. Prabhakaran, D. Brabb, and A. Jackson, "Improving freight rail safety with on-board monitoring and control systems," in *Rail Conference, 2005. Proceedings of the 2005 ASME/IEEE Joint*, pp. 117 – 122, march 2005.

[62] B. Lannoo, J. Van Ooteghem, D. Pareit, T. Van Leeuwen, D. Colle, I. Moerman, and P. Demeester, "Business model for broadband internet on the train," *JOURNAL OF THE INSTITUTE OF TELECOMMUNICATIONS PROFESSIONALS*, vol. 1, no. 1, pp. 19–27, 2007.

[63] K.-D. Lin and J.-F. Chang, "Communications and entertainment onboard a high-speed public transport system," *Wireless Communication Magazine*, vol. 9, pp. 84–89, Feb. 2002.

[64] D. Sivchenko, G. Xu, B.and Zimmermann, and S. Hischke, "Internet traffic performance in high speed trains," in *HET-NETs '04: Second Int'l Working Conf.: Performance Modelling and Evaluation of Heterogeneous Networks,(ITST),2009 9th International Conference on*, pp. 26–31, July 2004.

[65] M. Aguado, E. Jacob, M. Berbineau, J. Astorga, and N. Toledo, "Simulation framework for performance evaluation of broadband communication architectures for next generation railway communication services," in *Intelligent Transport Systems Telecommunications,(ITST),2009 9th International Conference on*, pp. 453 –457, oct. 2009.

[66] I. Gaspard and G. Zimmermann, "Investigations for broadband internet within high-speed trains," in *Advances in Radio Science*, vol. 3, pp. 247–252, May 2005.

[67] S. Ickin, K. Wac, M. Fiedler, L. Janowski, J.-H. Hong, and A. Dey, "Factors influencing quality of experience of commonly used mobile applications," *Communications Magazine, IEEE*, vol. 50, pp. 48 –56, april 2012.

[68] 3GPP, *3GPP Technical Specificacion Group Services and System Aspects; QoS Concept and Architecture (Release 4)*, 2002.

[69] S. Lin, W. Jiang, and Z. Zhong, "Analysis and requirements design of transmission interference of high-speed railway communication," in *Service Operations, Logistics, and Informatics (SOLI), 2011 IEEE International Conference on*, pp. 408 –412, july 2011.

[70] A. Nassar, M. Mohammed, E. A., and K. Elleithy, "Virtual wireless and mobile communication laboratory," *Network, IEEE*, vol. 2, no. 1, pp. 19–24, 2012.

[71] A. Kara, E. Aydin, R. Oktem, and N. Cagiltay, "A remote laboratory for training in radio communications: Errl," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, pp. 1–5, Sept 2007.

[72] "Keysight technologies."

[73] C. A. Garcia-Perez, A. M. Recio-Perez, A. Diaz-Zayas, and P. Merino-Gomez, "PerformLTE: a testbed for LTE testing in the future internet," in *Wired/Wireless Internet Communications, WWIC 2015, LNCS 9071*, pp. 313–326, 2015.

[74] Keysight, *Keysight Technologies E7530A and E7630A LTE/LTE-A Test and Lab Applications. Technical Overview.* Keysight, 2015.

[75] Keysight, *Keysight E7515A-L01 Protocol Logging and Analysis Software.User's Guide.* Keysight, 2014.

[76] Keysight, *Keysight E6966B IMS-SIP Network Emulator. Technical Overview.* Keysight, 2015.

[77] ITU-T, "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," Recommendation P.862.

[78] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) conformance specification; Radio transmission and reception; Part 1: conformance testing," TS 36.521-1.

[79] O. Heckmann, *The Competitive Internet Service Provider: Network Architecture, Interconnection, Traffic Engineering and Network Design.* Wiley, 2007.

[80] A. Diaz, P. Merino, and A. Salmeron, “Obtaining Models for Realistic Mobile Network Simulations using Real Traces,” *Communications Letters, IEEE*, vol. 15, pp. 782 –784, july 2011.

[81] 3GPP TS 45.005 V7.8.0 (2006-11), “3rd Generation Partnership Project; Technical Specification Group GSM/EDGE Radio Access Network; Radio transmission and reception (Release 7)”.

[82] 3GPP TR 25.943 V6.0.0 (2004-12), “Deployment aspects (Release 6)”.

[83] 3GPP TS 36.214, ”Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer - Measurements

[84] 3GPP TR 36.803 v1.1.0 (2008-04), ”User equipment (UE) radio transmission and reception (Release 8)”.

[85] 3GPP TR 25.913: “Requirements for Evolved UTRA and UTRAN”.

[86] ITU-T Rec. P.501 Test Signals for use in Telephonometry

[87] ITU-T Rec. G.114: One-way transmission time

[88] ITU-T Rec. G.107: The E-model: a computational model for use in transmission planning

[89] ITU-T Rec. P.862.1: Mapping function for transforming P.862 raw result scores to MOS-LQO

[90] ITU-T Rec. P.862.3: Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2

[91] ITU-T Rec. P.800: Methods for subjective determination of transmission quality

[92] ITU-T Rec. P.862: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs

[93] Zizhi Qiao; Lingfen Sun; Ifeachor, E.; , ”Case study of PESQ performance in live wireless mobile VoIP environment,” Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on , vol., no., pp.1-6, 15-18 Sept. 2008

[94] Qualipoc handheld, Web site: `http://www.swissqual.com/index.php/systems/qualipoc_handheld.html`

[95] Wang B.; Pedersen, K.I.; Kolding, T.E.; Mogensen, P.E.; ”Performance of VoIP on HSDPA”, IEEE 61st Vehicular Technology Conference (VTC-Spring), 2005

[96] Aho, K.; Repo, I.; Nihtila, T.; Ristaniemi, T.; ”Analysis of VoIP over HSDPA Performance with Discontinuous Reception Cycles”, Sixth International Conference on Information Technology: New Generations (ITNG), 2009

[97] Weimin X.; Ghosh, A.; Schaeffer, D.; Downing, L.; "Voice over IP (VOIP) over cellular: HRPD-A and HSDPA/HSUPA," IEEE 62nd Vehicular Technology Conference (VTC-Fall), 2005

[98] Shreevastav, R.; McGoldrick, C.; Huggard, M.; "Delivering improved QoS and cell throughput in UMTS based HSDPA networks," IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops (WoWMoM), 2009

[99] Arjona, A.; Westphal, C.; Yla-Jaaski, A.; Kristensson, M.; , "Towards High Quality VoIP in 3G Networks - An Empirical Study," 4th Advanced International Conference on Telecommunications (AICT), 2008

[100] Jurvansuu, M.; Prokkola, J.; Hanski, M.; Perala, P.; , "HSDPA Performance in Live Networks," IEEE International Conference on Communications (ICC), 2007

[101] Lunden, P.; Aijanen, J.; Aho, K.; Ristaniemi, T.; "Performance of VoIP over HSDPA in Mobility Scenarios", IEEE Vehicular Technology Conference (VTC-Spring), 2008

[102] Recommendation ITU-R M.1225, "Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000", ITU-R, 1997.