



UNIVERSIDAD DE CARABOBO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE
TELECOMUNICACIONES
DEPARTAMENTO DE SEÑALES Y SISTEMAS



**DESARROLLO DE UN CLASIFICADOR DE LA CALIDAD DE LA VOZ
PARA EL ESTANDAR UMTS MEDIANTE EL USO DE INTELIGENCIA
ARTIFICIAL**

MALPICA JOSE
BALDAN JUAN

Bárbula, 26 de noviembre del 2016



UNIVERSIDAD DE CARABOBO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE
TELECOMUNICACIONES
DEPARTAMENTO DE SEÑALES Y SISTEMAS



**DESARROLLO DE UN CLASIFICADOR DE LA CALIDAD DE LA VOZ
PARA EL ESTANDAR UMTS MEDIANTE EL USO DE INTELIGENCIA
ARTIFICIAL**

TRABAJO ESPECIAL DE GRADO PRESENTADO ANTE LA ILUSTRE UNIVERSIDAD DE
CARABOBO PARA OPTAR AL TÍTULO DE INGENIERO DE TELECOMUNICACIONES

MALPICA JOSE
BALDAN JUAN

Bárbula, 26 de noviembre del 2016



UNIVERSIDAD DE CARABOBO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE
TELECOMUNICACIONES
DEPARTAMENTO DE SEÑALES Y SISTEMAS



CERTIFICADO DE APROBACIÓN

Los abajo firmantes miembros del jurado asignado para evaluar el trabajo especial de grado titulado «DESARROLLO DE UN CLASIFICADOR DE LA CALIDAD DE LA VOZ PARA EL ESTANDAR UMTS MEDIANTE EL USO DE INTELIGENCIA ARTIFICIAL», realizado por los bachilleres MALPICA JOSE, cédula de identidad 22.224.945, BALDAN JUAN, cédula de identidad 20.728.186, hacemos constar que hemos revisado y aprobado dicho trabajo.

Firma

Prof. OSMAN AHMAD

TUTOR

Firma

Prof. HERRERA BENJAMIN

JURADO

Firma

Prof. HERNANDEZ ELIMAR

JURADO

Bárbula, 26 de noviembre del 2016

Dedicatoria

A Dios principalmente que me dio la vida y la fortaleza para alcanzar esta meta.

A mi madre quien es mi motor principal y siempre me ha brindado su apoyo incondicional.

A mi padre por sus consejos que de alguna u otra forma me motivaron para seguir luchando.

A mis familiares que siempre han creído en mi y me han tendido la mano en todo momento.

MALPICA JOSE

A mi madre Haydee, a mi padre Ramón, a mi hermano Guillermo y a todas esas personas que me apoyaron a lo largo de esta etapa de mi vida.

BALDAN JUAN

Agradecimientos

A Dios primordialmente porque gracias a el he finalizado esta etapa de mi vida.

A mis familiares por tanto cariño y afecto brindado.

A mi compañero Juan Baldan por su comprensión, tolerancia y el gran esfuerzo dedicado en este Trabajo de Grado llevado a cabo.

Al profesor Ahmad Osman por su valiosa colaboración y constante dedicación al guiarnos para la culminación de este proyecto.

A mis compañeros de estudios que más que compañeros los considero como unos hermanos por todas las experiencias vividas en el transcurso de estos años de vida universitaria.

MALPICA JOSE

Primeramente quiero agradecer a Dios, por haber podido culminar este Trabajo de Grado y junto con el, esta importante etapa de mi vida.

A mi madre, mi padre y mi hermano, quienes siempre me brindaron el inmenso apoyo y la gran comprensión, que sirvieron como soporte e inspiración para alcanzar exitosamente esta meta en mi vida.

A todos mis familiares que de una u otra forma me apoyaron a lo largo de estos años de carrera universitaria.

A todos esos amigos y compañeros de estudio que tuve la dicha de conocer a lo largo de este gratificante camino, por su apoyo, su amistad y los buenos momentos vividos con todos y cada uno de ellos.

A mi compañero de tesis y amigo José Yoel Malpica por todo el esfuerzo y energía dedicados a la culminación de este proyecto.

A nuestro tutor el profesor Ahmad Osman por brindarnos su apoyo y guiarnos durante toda la realizacion de este Trabajo de Grado.

Y finalmente a todos aquellos profesores que a lo largo de estos años se esforzaron por brindar sus conocimientos para prepararnos como profesionales.

A todos... GRACIAS!!

BALDAN JUAN

Índice general

Índice de Figuras	IX
Índice de Tablas	XI
Acrónimos	XIII
Resumen	XV
I. Introducción	1
1.1. MOTIVACIÓN	1
1.2. OBJETIVOS	2
1.2.1. Objetivo General	2
1.2.2. Objetivos Específicos	2
1.3. ALCANCES	3
II. Marco Conceptual	5
2.1. Estándar UMTS y su Arquitectura	5
2.1.1. La voz en los sistemas UMTS	7
2.1.1.1. Interfaz Aire del estándar UMTS	9
2.1.1.2. Técnica de Acceso WCDMA	11
2.1.1.3. Funcionamiento de la Tecnología WCDMA	12
2.2. Inteligencia Artificial	14
2.2.1. Redes Neuronales Artificiales y su estructura	14
2.2.2. Deep Learning y las redes neuronales profundas	15
2.2.2.1. El Perceptrón Multicapa	16
2.2.3. Entrenamiento Supervisado	18
2.2.3.1. Red Feed-Forward y el Algoritmo Backpropagation	19
2.2.3.2. Levenberg-Mardquardt	20
2.2.3.3. Función de Activación	20
III. Procedimientos de la investigación	21
3.1. Investigación y documentación.	21
3.2. Simulación de la Interfaz Aire del Estándar UMTS.	21

3.3. Validación del esquema de simulación	22
3.4. Creación del banco de muestras	22
3.4.1. Construcción del banco de muestras	23
3.4.2. Diseño experimental	24
3.5. Entrenamiento de la red MLP.	24
3.5.1. MLP de una capa oculta.	25
3.5.2. MLP de dos capas ocultas.	25
3.6. Verificación de resultados.	26
IV. Análisis, Interpretación y Presentación de los Resultados	27
4.1. Simulador de la interfaz aire del estándar UMTS.	27
4.2. Clasificador de voces de forma objetiva.	30
V. Conclusiones y recomendaciones	37
5.1. Conclusiones	37
5.2. Recomendaciones	38
Referencias Bibliográficas	41
Anexos	
A. Manual de construcción de un simulador de la interfaz aire para el estándar UMTS	
B. Manual de construcción de una red neuronal profunda del tipo MLP para la clasificación de voces	

Índice de figuras

2.1. Arquitectura de una Red UMTS(Fuente[7])	7
2.2. Diagrama de bloques básico de la Interfaz Aire del estándar UMTS (Fuente [17])	10
2.3. Proceso de Spreading y Despreading en WCDMA (Fuente:[17])	14
2.4. Representacion del Perceptron Multicapa	17
4.1. Simulador de la interfaz aire para el estándar UMTS	28
4.2. Gráfica de la BER vs SNR(dB)	29
4.3. Señal antes y despues del Spreading	30
4.4. Clasificación de una muestra de audio	31
4.5. Estructura del MLP	32
4.6. Clasificador final	33
1.1. Matriz Hadamard	
1.2. Estructura de árbol de los códiigos OVVSF	
2.1. Clasificador final	

Indice de tablas

3.1. Factores que distorsionan la voz en el canal RF	23
3.2. MOS para las muestras de audio procesadas	24
4.1. Representación del banco de muestras	31

Acrónimos

MOS	Mean Opinion Score
3GPP	3rd-Generation Partnership Project
ETSI	European Telecommunications Standards Institute
UMTS	Universal Mobile Telecommunications Systems
WCDMA	Wideband Code Division Multiple Access
ITU-T	International Telecommunication Union - Telecommunications
QPSK	Quadrature Phase Shift Keying
QOS	Quality Of Service
IMT-2000	International Mobile Telecommunications - 2000
UTRAN	UMTS Terrestrial Radio Access Network
UU	User UTRAN
UE	User Equipment
CN	Core Network
RNS	Radio Network Subsystem
RNC	Radio Network Controller
PLMN	Public Land Mobile Network
MSE	Mean Squared Error

DESARROLLO DE UN CLASIFICADOR DE LA CALIDAD DE LA VOZ PARA EL ESTANDAR UMTS MEDIANTE EL USO DE INTELIGENCIA ARTIFICIAL

por

MALPICA JOSE y BALDAN JUAN

Presentado en el Departamento de Señales y Sistemas
de la Escuela de Ingeniería en Telecomunicaciones
el 26 de noviembre del 2016 para optar al Título de
Ingeniero de Telecomunicaciones

RESUMEN

El desarrollo de un clasificador de calidad de la voz para el estándar UMTS mediante Inteligencia Artificial, basa su funcionamiento en el uso de una Nota Media de Opinión o MOS (del inglés Mean Opinion Score) que trabaje de forma objetiva gracias al uso de técnicas de Deep Learning con Redes Neuronales. Esto tiene como finalidad el estudio de la voz en el Canal de Transporte Dedicado, específicamente en la Interfaz Aire del Estándar de la Telefonía Móvil UMTS, para de esta manera poder calificarla de una manera objetiva, es decir dicha calificación no dependa de opiniones subjetivas de sujetos de pruebas, las cuales se pueden ver influenciadas por distintos factores.

Dicho clasificador consistirá de una Red Neuronal previamente entrenada con una base de datos creada a partir de distintas muestras de voces procesadas a través de un esquema de transmisión y recepción del estándar realizado en la herramienta de simulación Matlab-Simulink. Dicho esquema ha sido creado siguiendo de manera rigurosa las especificaciones dadas por la 3GPP, con la finalidad de poder simular el estándar con la mayor exactitud posible.

La cantidad de muestras procesadas y que previamente serán evaluadas para la creación del MOS, tendrá un número considerable para poder realizar un entrenamiento eficiente de la red neuronal que supondrá el uso eficiente de la inteligencia artificial.

Palabras Claves: clasificador, inteligencia, artificial, objetiva, neuronales, interfaz, subjetiva, estándar, telefonía

Tutor: OSMAN AHMAD

Profesor del Departamento de Señales y Sistemas

Escuela de Telecomunicaciones. Facultad de Ingeniería adscrito al Laboratorio X

Capítulo I

Introducción

1.1. MOTIVACIÓN

En la actualidad las empresas de telefonía móvil poseen problemas en lo que respecta a la calidad de la voz durante las diversas conversaciones entre sus usuarios, dichos problemas son ocasionados por múltiples factores que afectan el canal de comunicación, tales como el ruido de fondo (ruido blanco gaussiano), el eco, el desvanecimiento por multitrayectoria, entre otros. [1] [2] [3]

Dichas empresas realizan pruebas con la finalidad de determinar la calidad de voz sobre sus sistemas de telefonía móvil, pero debido a que dichas pruebas son realizadas tradicionalmente siguiendo la recomendación ITU-T P.800, la cual supone un método subjetivo, los resultados no pueden ser del todo confiables y medibles ya que dependen de la elaboración de una Nota Media de Opinión o MOS que está dado por las percepciones individuales de cada uno de los sujetos que participan en las pruebas.[4] [1] [2] [3]

Anteriormente se han desarrollado proyectos que guardan relación con lo antes expuesto, un ejemplo de ello son los estudios realizados en la Facultad de Ingeniería de la Universidad Militar de Granada-Colombia, los cuales están basados en la aproximación de la calidad de voz y cobertura en una red GSM de emergencia.

Donde hacen énfasis en medir los parámetros probabilísticos: GoS (Grade of Service), DCR (Dropped Call Rate) y MOS (Mean Opinion Score) cuya propuesta de telecomunicaciones tiene como objetivo la atención de desastres en situaciones de emergencia haciendo uso de software libre para interconectar llamadas. [5]

Por otro lado en la Universidad de Tsukuba-Japón se realizaron estudios acerca de la evaluación de la calidad subjetiva y objetiva para reducir el ruido al hablar en sistemas de telefonía móvil. Dicho estudio describe los resultados experimentales obtenidos en la evaluación una vez reducido el ruido ocasionado por los múltiples factores que afectan las comunicaciones. Se explica además que dicha evaluación de la reducción del ruido se realizó mediante el uso de un MOS subjetivo correspondiente al método especificado en ITU-T Rec. P.835. [6]

Es por todo lo antes expuesto se presenta este trabajo especial de grado, que tiene como finalidad el desarrollo un clasificador de voces que permita caracterizar la calidad de la voz en función de las perturbaciones presentes en los canales de telefonía móvil, el cual trabajaje de una forma objetiva mediante el uso de inteligencia artificial, utilizando técnicas de Deep Learning, específicamente mediante el uso del Perceptrón Multicapa o MLP por sus siglas en inglés, para de esta manera poder establecer la calidad de la voz sobre estos canales.

1.2. OBJETIVOS

1.2.1. Objetivo General

Desarrollar un clasificador de la calidad de la voz para el estándar UMTS mediante el uso de inteligencia artificial.

1.2.2. Objetivos Específicos

- Analizar los factores que afectan la calidad de la voz en función del canal de comunicaciones del estándar UMTS.

- Construir un esquema de simulación de la interfaz aire del estándar UMTS en función del canal, para la generación de un banco de datos con los cuales alimentar la red neuronal.
- Estudiar el funcionamiento de Deep Learning como inteligencia artificial para la creación de un clasificador de voces.
- Entrenar una red neuronal utilizando técnicas de Deep Learning, para clasificar la calidad de la voz, usando los datos obtenidos de las simulaciones.

1.3. ALCANCES

El trabajo especial de grado desarrollado implementa un clasificador de la calidad de la voz para el estándar UMTS, dicho clasificador trabaja de una manera objetiva mediante el uso de Deep Learning, específicamente mediante la implementación de un Perceptrón Multicapa (MLP) el cual ha sido entrenado utilizando como datos los factores que afectan las comunicaciones en el canal RF y que además podemos medir, esto debido a que en nuestro país las compañías de telefonía móvil realizan el procesamiento de la voz sobre canales que trabajan con dichos estándares (GSM, UMTS).

Capítulo II

Marco Conceptual

2.1. Estándar UMTS y su Arquitectura

El estándar UMTS o Sistema de Telecomunicaciones Móviles Universal es la propuesta de ETSI para la tercera generación (3G) de telefonía celular, sucesor de GSM, miembro de la familia global IMT-2000 del sistema de comunicaciones móviles de tercera generación de la ITU. Proporciona servicios que se adaptan a las necesidades y preferencias de los usuarios. La principal ventaja de UMTS sobre la segunda generación móvil (2G), es la capacidad de soportar altas velocidades de transmisión de datos y tener una gama de servicios mucho más amplia. Esta capacidad sumada al soporte del protocolo de Internet (IP), se combinan poderosamente para prestar servicios multimedia interactivos y nuevas aplicaciones de banda ancha, tales como video telefonía, video conferencia y la transmisión de audio y video en tiempo real.[7] [8]

La arquitectura del estándar UMTS se estructura de la siguiente forma: la UTRAN es su principal red de acceso, la cual tiene como función mantener, controlar y gestionar la conexión entre el UE o User Equipment (Equipo de Usuario) y el CN o Core Network (Núcleo de Red). Como se puede observar en la figura 2.1, la red UTRAN está conformada por varios RNS o Radio Network Subsystem (Subsistema de Red de Radio), los cuales a su vez, poseen un RNC o Radio Network Controller

(Controlador de Red Radioeléctrica) y varios Nodos B. El Nodo B se encarga de proveer la cobertura de la señal para el UE a través de la interfaz Uu también conocida como interfaz aire. El RNC se comunica con el Nodo B a través de la interfaz IuB, y tiene como tarea controlar y gestionar la conexión establecida por el Nodo B.

Las principales interfaces que la conforman son las siguientes: [9] [7] [1]

- Interfaz Iu: se encuentra entre un RNS y el CN. Sus objetivos principales son: brindar soporte a la interconexión del RNC con cualquier Punto de Acceso del Núcleo de la Red que se encuentre dentro de una misma PLMN (Red Móvil Terrestre Pública), y permitir el inter-funcionamiento con cualquier Núcleo de Red GSM. [10]
- Interfaz Iur: se encuentra entre dos RNS. Su principal tarea es la de permitir el intercambio de información entre dos RNC, inclusive si estos son de distintos fabricantes. [11]
- Interfaz Iub: se encuentra entre el Nodo B y su RNC. Se encarga de mantener la interconexión entre el RNC y sus Nodos B, siendo capaz de ofrecer compatibilidad aun si estos son de distintos fabricantes, además de permitir el manejo de los recursos radioeléctricos entre ambos, como por ejemplo agregar y eliminar celdas controladas por un Nodo B, para apoyar la conexión dedicada entre el UE y el RNC. [12]
- Interfaz Uu: se encuentra entre el UE y el Nodo B. Es la interfaz aérea del estándar UMTS, su funcionamiento se basa en la técnica de acceso múltiple WCDMA y el Espectro Ensanchado por Secuencia Directa (DSSS). [7]

En la figura 2.1 se observa la Arquitectura de una Red UMTS con sus distintos componentes e interfaces.

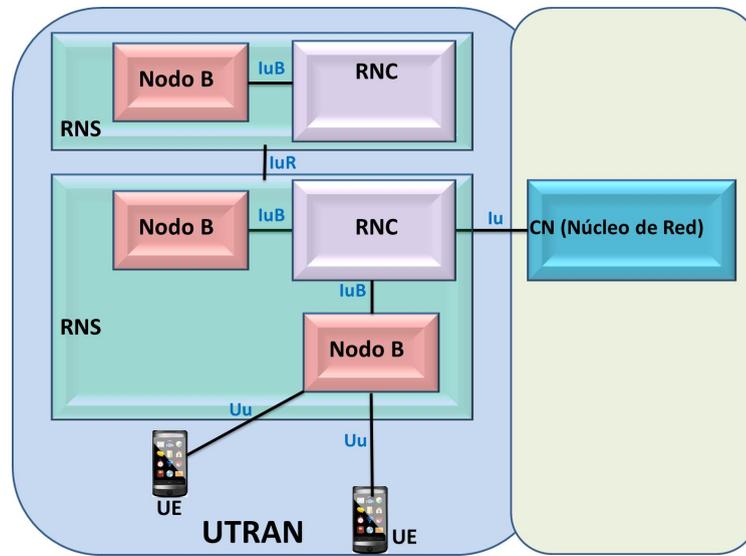


Figura 2.1: Arquitectura de una Red UMTS(Fuente[7])

2.1.1. La voz en los sistemas UMTS

Siempre que el sistema esté funcionando de manera correcta no deberían existir interferencias y alteraciones en la voz durante una conversación entre usuarios, sin embargo esta es una condición muy ideal que generalmente no se cumple. En la interfaz aire también conocida como interfaz Uu del estándar UMTS, se puede degradar la calidad de la voz durante una llamada, por ejemplo el proceso de Handover, encargado de transferir el servicio de una estación base a otra, cuando la calidad del enlace es insuficiente, produce un bajo rendimiento de la interfaz que contribuye en gran medida a la degradación de la calidad de la voz. Otros factores que afectan la calidad de la voz en UMTS son: las interferencias por ruido, el eco, el desvanecimiento por multitrayectoria, interferencias por otros usuarios o co-canal y las interferencias por otras celdas o canal adyacente. En este trabajo de grado nos enfocamos en el estudio y simulación de los tres primeros. [3]

- Las interferencias por ruido: se define como ruido a cualquier señal indeseable en un sistema de telecomunicaciones. El ruido es un fenómeno natural,

inevitable y generalmente incontrolable, en otras palabras, siempre estará presente en cualquier sistema de comunicaciones y contribuirá, en mayor o menor medida, al deterioro de la señal en su llegada al receptor. El ruido blanco gaussiano aditivo AWGN suele ser la forma típica de modelar el ruido. El factor que determina la calidad de la señal se le conoce como relación señal a ruido SNR y se mide en dB. Cuanto más baja sea esta, más ruidosa se vuelve la señal. En los sistemas UMTS la SNR tiende a ser mas baja a medida que el móvil se aleja del Nodo B, valores por debajo de los -10dB provocan distorsión en la señal recibida al punto de volverse imperceptible. [13][3]

- El desvanecimiento por multitrayectoria y la velocidad del móvil (desplazamiento Doppler): en los sistemas de telefonía la señal se propaga desde el transmisor hasta el receptor a través de múltiples caminos gracias a los obstáculos y reflectores existentes en el canal RF. La propagación por múltiples trayectos está compuesta por la presencia de una componente con línea de vista y otras sin ésta conocidas como componentes difusas; entre el Nodo B y el móvil. Cuando dicho móvil se encuentra en movimiento su velocidad provoca un cambio en la frecuencia de la señal transmitida a lo largo de cada trayecto de la misma, este fenómeno se conoce como desplazamiento Doppler el cual es capaz de distorsionar la señal a medida que aumenta, y está regido por la siguiente ecuación:

$$f_d = \left(\frac{v}{c}\right)f_c$$

Donde v corresponde a la velocidad con la que se mueve el dispositivo, c es la velocidad de la luz, f_c es la frecuencia de la portadora, y f_d el máximo desplazamiento Doppler en Hz, así que por ejemplo para los sistemas de telefonía móvil si f_c es igual a 1.8GHz se podrían exponer tres casos:

1. Un peatón que camina a 1m/s, proporciona una f_d igual a 6Hz.
2. Un carro que viaja a 20m/s (72Km/h), proporciona una f_d igual a 120Hz
3. Un tren que viaja a 42m/s (151Km/h), proporciona una f_d igual 250Hz[14]

- El eco: es el fenómeno en el cual parte de la señal que se transmite regresa al emisor de forma retrasada y distorsionada. Todas las conversaciones tele-

fónicas sufren la presencia de eco. El eco depende del retardo involucrado. Si el retardo es corto, el eco es imperceptible, pero si el retardo es grande, el eco torna difícil la comunicación. Estadísticamente se puede suponer el umbral para el retado de 25mseg, es decir que si el retado supera esto el eco es perceptible por el oído humano, retardos de entre los 120mseg y 250mseg pueden suponer grandes problemas en una conversación. Los retardos mas grandes que se generan en un sistema telefónico se deben a largas distancias físicas, por ejemplo en una transmisión por satélite este puede ser superior a 520mseg.[15] [3]

Durante una llamada todos estos factores contribuyen al aumento de la Tasa de Error de Bit o BER (Bit Error Rate). Cuando la BER se incrementa el decodificador de voz recibirá cada vez menos información acerca de la voz codificada lo que tendrá como consecuencia la degradación de la calidad de la misma. [3]

Estos tres factores han sido estudiados y simulados, para su posterior uso en el llenado del banco de muestras para el entrenamiento de la red.

2.1.1.1. Interfaz Aire del estándar UMTS

La primera parte de este trabajo de grado, se enfoca en la construcción y simulación de la Interfaz Aire del estándar UMTS, esto debido a que es en el canal RF donde se originan la mayoría de las interferencias relacionadas a la degradación en la calidad de la voz, un ejemplo de ello, como se explico anteriormente, son las interferencias debidas al desvanecimiento por multitrayectoria y desplazamiento Doppler, así como los problemas por ruido y eco que afectan en gran medida al canal radioeléctrico. [16][14] [15] [3]

La Interfaz Aire o Interfaz Uu del estándar UMTS consta de las siguientes etapas, para la transmisión primeramente se realiza la Codificación, luego se produce el Spreading (Ensanchamiento) y el Scrambling debidos a la técnica de acceso WCDMA usada por el estándar, después de esto se lleva a cabo la Modulación para que de esta manera se lleve a cabo la transmisión por el canal RF; para la recepción

el proceso viene a ser el inverso, siguiendo el orden: Demodulación, De-Scrambling, De-Spreading y Decodificación. En la figura 2.2 se puede observar el diagrama de bloques básico de la Interfaz. [7] [17] [18] [1]

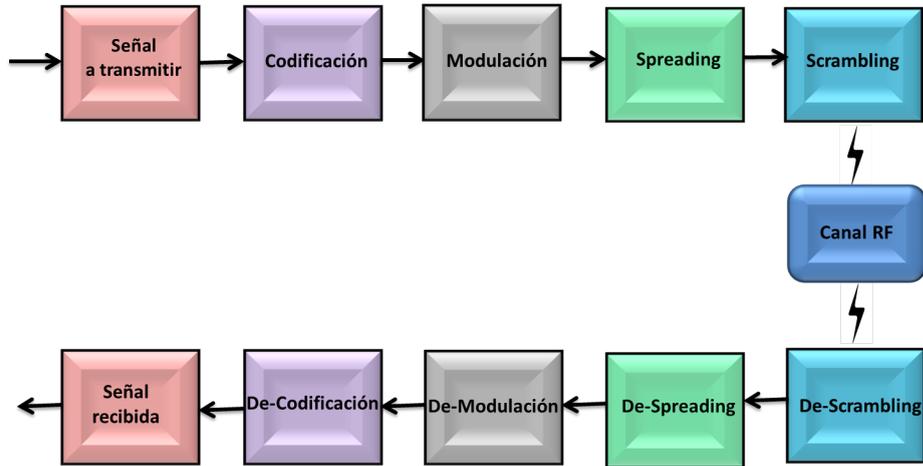


Figura 2.2: Diagrama de bloques básico de la Interfaz Aire del estándar UMTS (Fuente [17])

Una breve explicación de cada etapa se da a continuación:

- **Modulación y Demodulación QPSK:**

QPSK (Quadrature Phase Shift Keying) es una versión de la modulación de frecuencia donde la fase de la onda portadora se modula para codificar bits de información digital en cada cambio de fase. Con cuatro fases, QPSK es capaz de codificar dos bits por símbolo, los bits son asignados a cada símbolo mediante el código Gray, de manera tal que entre dos símbolos adyacentes, los símbolos solo se diferencian en 1 bit, con lo que se logra minimizar la tasa de error de bits. [17] [7] [1]

- **Codificador Convolutivo y Decodificador Viterbi**

La codificación convolutiva junto con la decodificación de Viterbi es una de las principales técnicas FEC (Forward Error Correction) empleadas en la codificación de canales que son corrompidos principalmente por ruido aditivo

gaussiano (AWGN). En el codificador convolucional la secuencia de bits codificada depende de los bits previos, debido a que el mismo realiza una codificación continua. Por su parte el algoritmo de decodificación de Viterbi basa su funcionamiento en el llamado diagrama de Trellis. El cual intenta reconstruir las transiciones que se produjeron al codificar una secuencia de entrada dada.[17] [1][7]

- Spreading y De-Spreading:

La información a transmitir es multiplicada por un código que ensancha la señal aumentando su ancho de banda. Este proceso se explica detalladamente mas adelante.[17][7] [19] [20]

- Scrambling y De-Scrambling:

La señal es multiplicada por un código pseudoaleatorio utilizado para diferenciar las señales de otras. El proceso se explica detalladamente mas adelante. [17][7] [20] [19]

El diagrama de bloques explicado, ha sido utilizado para la construcción detallada del esquema de simulación de la interfaz aérea en UMTS, y de esta manera hacer las simulaciones necesarias en este trabajo de grado.

2.1.1.2. Técnica de Acceso WCDMA

WCDMA es una tecnología derivada del CDMA tradicional utilizado en las redes IS-9540. Las principales diferencias entre ambas tecnologías, es que WCDMA utiliza una señalización y un canal de control diferente, así como un mayor ancho de banda (5 MHz) para su funcionamiento, el cual le provee la capacidad para transmitir datos a velocidades de hasta 2 Mbps. [7] [1]

Esta tecnología emplea una técnica de ensanchamiento, es decir, la señal de datos es ensanchada para que ocupe todo el ancho de banda asignado para la transmisión. Este ensanchamiento se realiza con un código de ensanchamiento específico

para cada usuario, con el cual se establece la diferencia entre cada equipo conectado a la red. Este procedimiento es llamado DS-WCDMA Direct Sequence WCDMA (WCDMA de Secuencia Directa). [7] [17] [18] [1]

La principal ventaja que ofrece WCDMA a sus usuarios, es que gracias a la técnica de espectro ensanchado, proporciona una alta seguridad y protección contra posibles interceptaciones en sus llamadas. [1] [7]

WCDMA posee dos modos fundamentales de transmisión, los cuales son: FDD y TDD. En el desarrollo de WCDMA se hizo énfasis en el modo FDD ya que es el más difundido. [7] [17] [18] [1]

2.1.1.3. Funcionamiento de la Tecnología WCDMA

- Spreading y Despreading: para la tecnología WCDMA el funcionamiento es el siguiente: la información a transmitir es multiplicada por un código, el resultado de esta multiplicación produce una señal de mayor ancho de banda, específicamente de 3.84MHz, que representa el ancho de banda asignado para la transmisión en modo FDD-WCDMA, a este procedimiento se le denomina «Spreading». El receptor capta la señal ensanchada y la sincroniza con el mismo código que se utilizó para la transmisión, el resultado será: la información transmitida más algunos armónicos de alta frecuencia que no forman parte de la información y que pueden ser filtrados con facilidad, a este procedimiento se le denomina «Despreading». El Spreading y el Despreading son realizados tanto por el Nodo B como por el teléfono móvil, debido a que la información transita en dos sentidos, uno desde el teléfono hacia el Nodo B (Uplink) y otro desde el Nodo B al teléfono (Downlink). Los códigos utilizados para el proceso de Spreading son los denominados códigos ortogonales (OVSF codes), en el Anexo A se explica esto con detalles.[17] [19] [7] [18] [21] [20] [1]
- Código de Scrambling y De-Scrambling: es utilizado en la segunda parte del proceso de ensanchamiento, su finalidad es diferenciar la información de cada usuario en el trayecto radioeléctrico, en el proceso de Scrambling este código es asignado por la red al usuario antes de una transmisión de tal forma

que ambos lo conocen y lo utilizan para la separación de la información en el proceso de De-Scrambling. Usualmente se utilizan códigos de ruido pseudo-aleatorio (PN codes).[17] [19] [7] [18] [21] [20] [1]

- Chips: Un concepto utilizado en WCDMA es el «chip», que corresponde a los bits utilizados en el código de ensanchamiento. La velocidad del código de ensanchamiento no se expresa en bits/ segundos (b/s), sino en chips/s (chips por segundo), de tal forma que la velocidad del código de ensanchamiento es de 3.84 Mchips/s, que es lo necesario para que la señal de datos se ensanche a los 3.84MHz. La velocidad de códigos es conocida como SCR System Chips Rate (Tasa de Chips del Sistema).[17] [19] [1]
- Factor de ensanchamiento: el factor de ensanchamiento es una cifra que describe el número de chips por cada símbolo utilizado para el ensanchamiento de la señal, se representa de la siguiente forma: [17] [19] [7] [18] [21] [20] [1]

$$\text{Factor de Ensanchamiento} = \frac{\text{Tasa de chips del sistema}}{\text{Símbolos generados por los datos}}$$

- Símbolos: Los símbolos dependen de la modulación. Un símbolo es un elemento de transmisión como resultado de una modulación. Las modulaciones utilizadas en WCDMA son QPSK, para el enlace descendente, en el cual se utilizan cuatro símbolos para la transmisión, donde se transmiten dos bits por cada símbolo. Para el enlace ascendente se utiliza Dual QPSK el cual utiliza dos símbolos para la transmisión, donde se transmite un bit por cada símbolo. En la figura 2.3 se puede observar el proceso de Spreading y Despreading de la señal. [17] [19] [7] [18] [21] [20] [1]

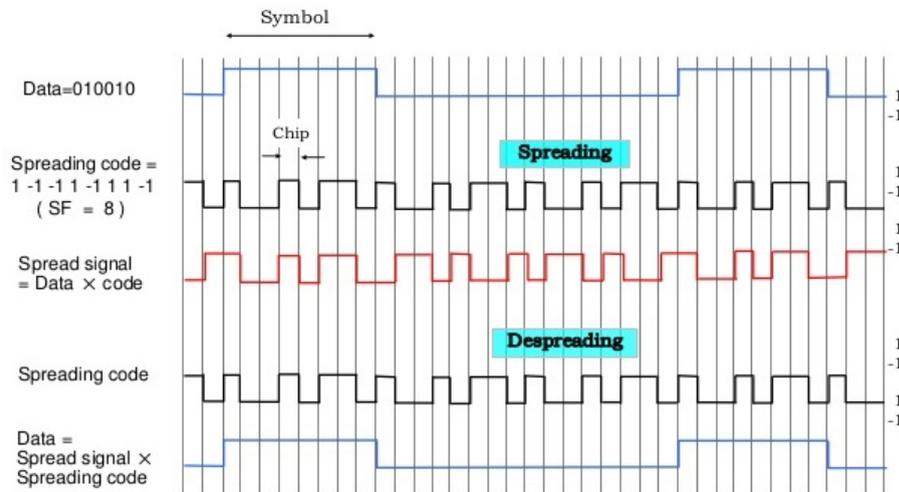


Figura 2.3: Proceso de Spreading y Despreading en WCDMA (Fuente:[17])

2.2. Inteligencia Artificial

La Inteligencia Artificial es la parte de la ciencia que se ocupa del diseño de sistemas de computación inteligentes, es decir sistemas que exhiben las características que asociamos con la inteligencia en el comportamiento humano. La IA se encarga de desarrollar sistemas que realicen tareas que, cuando las realizan las personas, las asociamos con la inteligencia. [22]

2.2.1. Redes Neuronales Artificiales y su estructura

Las redes neuronales artificiales (RNA), son una aproximación a la inteligencia artificial fundamentada en el cerebro de los mamíferos, más explícitamente en las neuronas biológicas. Las mismas suponen una implementación sencilla de un comportamiento local observado en nuestros cerebros. El cerebro está compuesto de neuronas, las cuales son elementos individuales de procesamiento. La información viaja entre las neuronas, y basado en la estructura y ganancia de los conectores neuronales, la red se comporta de forma diferente.

Una red neuronal, es un sistema de procesadores paralelos conectados entre sí en forma de grafo dirigido. Esquemáticamente cada elemento de procesamiento (neuronas) de la red se representa como un nodo. Estas conexiones establecen una estructura jerárquica que tratando de emular la fisiología del cerebro busca nuevos modelos de procesamiento para solucionar problemas concretos del mundo real.[23] [24] [25]

Las redes neuronales artificiales están formadas por una gran cantidad de neuronas, estas no suelen denominarse neuronas artificiales sino nodos o unidades de salida. Un nodo o neurona cuenta con una cantidad variable de entradas que provienen del exterior (X_1, X_2, \dots, X_m). [24]

Por otro lado se dispone de una sola salida (Y_j) que transmitirá la información al exterior o hacia otras neuronas. Cada Y_j o señal de salida tiene asociada una magnitud llamada peso, este se calculará en función de las entradas, por lo cual cada una de ellas es afectada por un determinado peso ($W_{j0} \dots W_{jq} + m$). Los pesos corresponden a la intensidad de los enlaces sinápticos entre neuronas y varían libremente en función del tiempo y en cada una de las neuronas que forman parte de la red. [24]

El proceso de aprendizaje consiste en hallar los pesos que codifican los conocimientos. Una regla de aprendizaje hace variar el valor de los pesos de una red hasta que estos adoptan un valor constante, cuando esto ocurre se dice que la red ya "ha aprendido". Al conectar varias neuronas de un determinado modo, se consigue una red.[24]

2.2.2. Deep Learning y las redes neuronales profundas

Un tipo de Inteligencia Artificial muy común en la actualidad es Deep Learning, el cual es un conjunto de técnicas y procedimientos algorítmicos basados en Machine Learning para lograr que una máquina aprenda de la misma forma que lo hace un ser humano. Siendo más precisos, hablamos de una familia de algoritmos cuyo propósito es simular el comportamiento que lleva a cabo nuestro cerebro para reconocer imágenes, palabras o sonidos.

Aunque existen varias maneras de implementar Deep Learning, una de las más comunes es utilizar redes neuronales. Lo importante en el desarrollo de estas es su útil comportamiento al aprender, reconocer y aplicar relaciones entre objetos y tramas de objetos propios del mundo real. En este sentido, son utilizadas como una herramienta para resolver problemas difíciles.[23] [24] [25]

Las Redes Neuronales Profundas o Deep Neural Networks, se refiere a una familia de algoritmos que simulan el proceso que lleva a cabo el cerebro cuando el mismo trata de reconocer o clasificar rostros, imágenes, voces o palabras. Son algoritmos que funcionan en base a un proceso por capas. El aprendizaje profundo simula el funcionamiento básico del cerebro, que se realiza a través de las neuronas. Existen dos tipos de redes neuronales profundas muy utilizadas en la actualidad, estas son las CNN (Redes Neuronales Convolucionales) y el MLP (Perceptrón Multicapa), la segunda es la que se utiliza en este Trabajo de Grado y se explica a continuación. [25][23]

2.2.2.1. El Perceptrón Multicapa

El perceptrón multicapa (MLP) es una técnica de Deep Learning que consta de una red neuronal de tipo feedforward (conexión hacia adelante), la cual contiene dos o más capas ocultas de neuronas, que utiliza funciones de activación no lineales tales como la función sigmoïdal o la función tangente hiperbólica. Una red de tipo MLP se compone de varias capas de neuronas. En la figura 2.4 se puede observar la representación del MLP.

El MLP de la figura 2.4 aparece totalmente interconectado, es decir podemos observar que cada neurona de una capa tiene una conexión a cada neurona de la capa anterior y de la siguiente capa. Hay casos en que existe interconexión parcial lo que se traduce en que cada neurona se conecta solamente con un subconjunto de las neuronas de la capa anterior o de la siguiente, aunque esto último no es muy habitual. También es posible la conexión entre neuronas de una misma capa en las denominadas redes competitivas, o incluso la realimentación de la señal de salida a las entradas, en las redes llamadas recurrentes[27] [24][28]

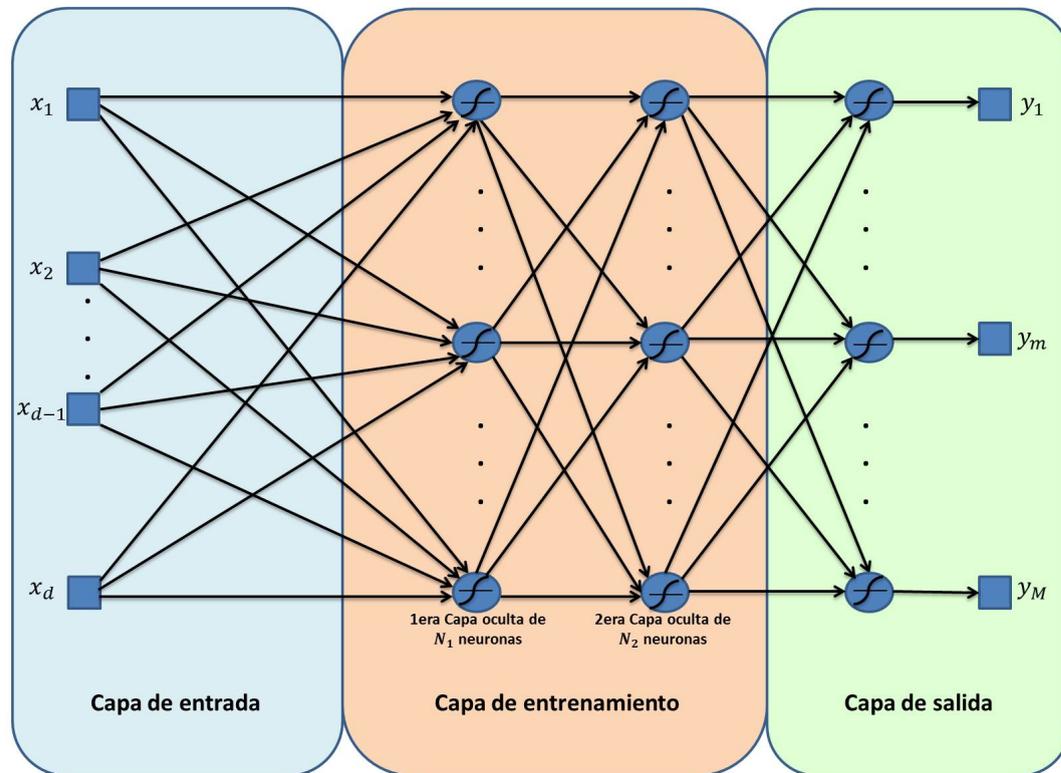


Figura 2.4: Representación del Perceptrón Multicapa

- La capa de entrada de un MLP realiza la función de permitir la entrada de muestras a la red neuronal, por lo que no es una capa de “neuronas” en el sentido estricto de la palabra. No contiene pesos ni umbrales, ni se realiza en ella ninguna operación matemática, sino que simplemente “lee” los datos que se presentan a su entrada y los distribuye a la primera capa oculta de neuronas. Visto de otra forma, se puede también considerar que la capa de entrada de un MLP contiene nodos cuya activación representa valores de las variables del dominio del problema en el cual la red es aplicada.
- Por su parte, la capa de salida representa las decisiones que toma el MLP. Existirán al menos tantas neuronas de salida como variables de salida tenga el problema a considerar. A menudo la función de activación de las neuronas de salida es lineal, de manera que sólo realiza la suma ponderada de las señales procedentes de la última capa oculta. También se pueden usar funciones

sigmoidales según el problema a resolver.

- En lo que respecta a las capas ocultas estas representan relaciones entre las variables de entrada que no pueden interpretarse a simple vista. Representan variables que no pertenecen al dominio del problema, sino que codifican relaciones entre ellas a más alto nivel. Es en esta etapa donde se ejecutan los algoritmos avanzados tales como el Gradiente Conjugado y el Levenberg-Mardquardt. El número de capas ocultas es un parámetro a optimizar cuando se diseña un MLP. El Teorema de Aproximación Universal especifica que “Una red MLP de una capa oculta es suficiente para realizar el mapeo entrada-salida que proporciona una realización aproximada de cualquier función continua en un intervalo dado”. En la práctica se observa que en muchas ocasiones, aun siendo suficiente el uso de una capa oculta, al utilizar dos capas ocultas, el rendimiento de la red MLP mejora considerablemente. Hecho que se ha podido constatar durante la elaboración de este trabajo, en donde inicialmente se realizaron pruebas con una sola capa oculta, y aunque los resultados no fueron del todo desfavorables, una vez implementada una segunda capa oculta se logra optimizar notablemente el entrenamiento de la red MLP.[27] [24] [28]

Se podría pensar que mientras más capas ocultas se utilicen, mejor serán los resultados a obtener, sin embargo, esto depende de la cantidad de muestras a procesar por el MLP, en la mayoría de los casos utilizar mas de dos capas ocultas conducirá a un entrenamiento más lento, que a la larga puede no ser beneficioso para cumplir los objetivos deseados.[27] [24]

2.2.3. Entrenamiento Supervisado

Este es el tipo más popular de entrenamiento y el mismo que ha sido utilizado en el desarrollo de este trabajo de grado. En este modo de entrenamiento, se puede considerar intuitivamente que existe un “profesor” que “enseña” a la red, proporcionando información para llevarla a emular la función deseada. Supongamos que

un conjunto de muestras de entrada se aplica a la entrada de la red, entonces la respuesta a la salida se compara con la respuesta deseada que proporciona el “profesor”, este informa a la red, ya sea el resultado correcto o incorrecto, para actualizar los pesos en consecuencia. Para este tipo de aprendizaje hace falta conocer la salida deseada de las muestras de entrenamiento.[27] [29]

2.2.3.1. Red Feed-Forward y el Algoritmo Backpropagation

La red Feed-Forward es el modelo más utilizado en lo que se refiere al Perceptrón Multicapa, se caracteriza por su forma de trabajar, en la que las neuronas de cada capa reciben información de las neuronas de la capa precedente y la transmiten a la capa siguiente, además de que las neuronas no están interconectadas dentro de cada capa de la red.

El algoritmo Backpropagation por su parte, es un método de entrenamiento de redes multicapa con aprendizaje supervisado. Su funcionamiento consiste en el aprendizaje de un conjunto predefinido de pares entradas-salidas, empleando un ciclo de propagación-adaptación de dos fases: una vez que se ha aplicado un patrón a la entrada de la red como estímulo, éste se propaga desde la primera capa a través de las capas intermedias u ocultas de la red, hasta generar una salida. La señal de salida se compara con la salida deseada y se obtiene un error para cada una de las salidas. El error se propaga hacia atrás, partiendo de la capa de salida, pasa por las capas ocultas hasta llegar a la capa de entrada. Las neuronas de las capas ocultas y de entrada, solo reciben una fracción del total del error, basándose aproximadamente en la contribución relativa que haya aportado cada neurona a la salida original. Este proceso se repite capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su contribución relativa al error total. Para reducir el error en la red, se minimiza el error con respecto a los pesos de esta intentando llevarlo a cero. El software de simulación se encarga de llevar a cabo esta minimización con métodos como el de Levenberg-Marquardt, utilizado en este trabajo de grado y que explicaremos mas adelante.

La unidad procesadora básica de la red Feedforward, que usa la Backpropagation como método de entrenamiento la podemos observar en la figura 2.4. [27] [29]

2.2.3.2. Levenberg-Marquardt

El algoritmo de Levenberg-Marquardt (LM) es otro algoritmo de optimización no lineal basado en el uso de las derivadas de segundo orden. Como la Backpropagation, computa los cambios de peso, sin embargo dicho método es más restrictivo. Los requerimientos de memoria para el algoritmo LM aumentan en función del cuadrado del número de pesos, por lo tanto se limita a redes relativamente pequeñas. El algoritmo LM es una combinación entre los métodos de Backpropagation y el método de Newton. El método supone que la función que se modela es lineal y que el error mínimo se puede encontrar en un solo paso. Se calcula la variación del peso para hacerlo en un único paso. Testea la red con estos nuevos pesos para determinar donde es más pequeño el error. Una variación en el error sólo es aceptada si esta mejora el error. Cuando el error decrece, la variación del peso es aceptada y la hipótesis lineal se ve reforzada por la disminución de un parámetro de control. Cuando el error aumenta, la variación del peso es rechazada y, como en Backpropagation, se sigue el gradiente para aumentar el control y así restar importancia a la suposición lineal. [27] [29]

2.2.3.3. Función de Activación

En la realidad se sabe que las neuronas envían señales a otras neuronas a través de una señal eléctrica a lo largo del axón. Esto se modela a través del uso de una función de transferencia o activación. Cuando usamos Backpropagation, esta función de transferencia es comúnmente una sigmoide(logística o tangente hiperbólica) o una función lineal.

Capítulo III

Procedimientos de la investigación

3.1. Investigación y documentación.

Se recopiló toda la documentación necesaria para la realización del proyecto en general. Empezando por la revisión del estándar UMTS y el estudio de los parámetros de funcionamiento del mismo; la selección del software más adecuado para la implementación del esquema para la simulación del estándar; la forma en que se realizaría la clasificación de la calidad de la voz y su desarrollo mediante deep learning utilizando el Perceptron Multicapa MLP.

3.2. Simulación de la Interfaz Aire del Estándar UMTS.

En esta etapa se llevó a cabo la parte práctica del proyecto referente a la simulación del estándar UMTS tomando en cuenta los parámetros y consideraciones teóricas que fueron analizadas previamente. Para la ejecución de este proceso de diseño y desarrollo se realizó de forma planificada lo siguiente:

- Montaje del esquema UMTS en simulink siguiendo de forma exhaustiva las recomendaciones de la 3GPP.

- Prueba de transmisión y recepción de audio con las etapas de modulación-demodulación y codificación-decodificación.
- Prueba del funcionamiento de la técnica de acceso WCDMA, simulando diversos usuarios en la red, junto con pruebas de transmisión y recepción de audio.

En dicho esquema de simulación se implementa la interfaz aire del estándar con una etapa Transmisora, el Canal de Comunicaciones RF y una etapa Receptora.

La descripción detallada de la construcción del esquema, así como el funcionamiento de cada bloque utilizado en el mismo se encuentra en el Anexo A de este Trabajo de Grado.

3.3. Validación del esquema de simulación

La validación de dicho esquema se realizó mediante múltiples estudios. Para la verificación de la técnica de acceso WCDMA, se simuló más de un usuario en el esquema, con la finalidad de comprobar el funcionamiento de la técnica del Espectro Ensanchado (Spreading/Scrambling), además de la verificación en las gráficas de los bits enviados, procesados a través del proceso de Spreading/Scrambling-Despreading/Descrambling y recibidos correctamente.

Igualmente se construyeron las gráficas de BER en función de los 3 factores estudiados referente las perturbaciones del canal RF, es decir BER vs SNR(dB), BER vs Eco(ms) y BER vs Fd(Hz). Y de esta forma comprobar el aumento de la Tasa de Error de Bits (BER), a medida que se incrementan cada uno de estos factores en las simulaciones, produciendo el deterioro en los archivos voz tal y como se esperaba.

3.4. Creación del banco de muestras

Esta fase fue una de las más importantes para este trabajo de grado, la misma consistió en la elaboración de un banco de muestras, a partir de la planificación de

un diseño experimental. En la Tabla 3.1 se muestran los factores que distorsionan la calidad de la voz en el canal RF, tales como el ruido aditivo gaussiano AWGN, expresado mediante la SNR(dB); el Eco, medido en retardo(ms) y el desplazamiento Doppler f_d , medido en Hertz(Hz), todos estudiados anteriormente en el Capitulo 2. Así como el rango de valores entre los cuales se ha variado cada uno de estos factores, un total de 10 variaciones para cada caso.

El procedimiento seguido fue, simular los múltiples casos posibles de combinaciones aleatorias entre estos factores afectando una muestra de audio a la vez, por ejemplo: una muestra de audio con un nivel de SNR de -10dB, eco con 50ms de retardo y desplazamiento Doppler de 80Hz.

Parámetro/Niveles	1	2	3	4	5	6	7	8	9	10
SNR(dB)	-16	-14	-10	-8	-4	-2	0	2	4	6
Eco(ms)	0	10	30	50	70	90	150	180	210	300
f_d (Hz)	0	4	10	20	40	60	80	100	120	250

Tabla 3.1: Factores que distorsionan la voz en el canal RF

3.4.1. Construcción del banco de muestras

Para la construcción de dicho banco de muestras se elaboró un cuestionario, para simular el método de clasificación de la Nota Media de Opinión o MOS por sus siglas en inglés. El mismo fue enviado a una población de aproximadamente 200 personas, entre alumnos y profesores de la Escuela de Telecomunicaciones de la Universidad de Carabobo, junto con los grupos de muestras de audio procesadas y afectadas por las distorsiones del canal, haciendo uso del esquema de simulación. Con la intención de obtener los resultados de la clasificación por parte de los participantes, de la forma en que se muestra en la tabla 3.2.

Muestra de audio procesada	MOS
Audio perceptible en su totalidad sin ningún tipo de distorsión	5
Audio con muy poco ruido y/o eco, perceptible casi al 100 %	4
Audio con ruido y/o eco, pero perceptible	3
Audio con ruido y/o eco, apenas perceptible	2
Audio con mucho ruido, prácticamente imperceptible	1

Tabla 3.2: MOS para las muestras de audio procesadas

3.4.2. Diseño experimental

Para establecer la cantidad de muestras de audio a procesar a través del esquema, se utilizó el Método Factorial General, con la Ecuación de la Combinatoria, descrita a continuación:

$$\text{Nro. Combinaciones} = \frac{n!}{(n - r)!}$$

Donde "n" viene a ser la cantidad de valores distintos que toma cada parámetro, 10 para cada caso como se mencionó anteriormente, y "r" la cantidad de parámetros a combinar, un total de 3, representado por las distorsiones del canal que se simulan en el esquema. De manera tal que el número de combinaciones sería:

$$\frac{10!}{(10 - 3)!} = 720$$

Es decir un total de 720 casos posibles. En cuanto a la partición óptima del conjunto de datos, no existe ninguna regla establecida para efectuarla. Para este estudio se usaron un total de 660 muestras como datos de entrenamiento, y las 60 restantes como datos de prueba para el clasificador.

3.5. Entrenamiento de la red MLP.

El diseño para la arquitectura del MLP se llevó a cabo siguiendo las reglas establecidas para la construcción de este tipo de red neuronal profunda. A continuación

se da una explicación detallada.

3.5.1. MLP de una capa oculta.

Inicialmente se realizaron pruebas con una red MLP sencilla de una sola capa oculta de neuronas. El entrenamiento de la red se optimiza mediante el criterio que por defecto utiliza el software de simulación, el cual que consiste en la minimización del MSE (Error Cuadrático Medio). Según este, cuanto menor sea el MSE mejor se adaptan los datos. El número de nodos o neuronas en la capa oculta varía desde 2 hasta 30 mediante la adición sistemática de dos nodos a la red, es decir:

$$N = (2, 4, 6, \dots, 30),$$

Donde N(número de neuronas en la capa oculta), se varía hasta conseguir los mejores resultados con los datos de prueba. La configuración general de la red tiene la siguiente forma: estructura Feed-forward con algoritmo Back-Propagation, optimización mediante Levenberg-Marquardt y función de activación tangente sigmooidal. Si bien los resultados para este caso no son del todo desfavorables, se sabe que es posible mejorar esto agregando otra capa oculta a la red.

3.5.2. MLP de dos capas ocultas.

La segunda etapa de pruebas para el entrenamiento consistió en agregar otra capa oculta a la red, obteniéndose de esta forma el Perceptron Multicapa de dos capas ocultas, la arquitectura más básica de Deep Learning, ya que por definición una red MLP con más de una capa oculta se considera "profunda". En este modelo consideraremos N_1 y N_2 como el número de neuronas de las capas ocultas 1 y 2 respectivamente, luego N_1 es un número par que se hace variar desde 10 a 20 mediante la adición sistemática de dos neuronas, mientras que N_2 es un número par que, para cada valor de N_1 , se hace variar desde 2 hasta como máximo N_1 , siempre que no se exceda el número de 32 neuronas en total, es decir:

$$N_1 = (10, 12, 14, 16, 18, 20)$$

$$N_2 = (2, 4, 6, \dots, N_1 - 2, N_1), /N_2 \leq N_1; N_1 + N_2 \leq 32$$

Al igual que para el caso de una capa oculta, la configuración general de la red es la misma: estructura Feed-forward con algoritmo Back-Propagation, optimización mediante Levenberg-Mardquardt y función de activación tangente sigmooidal. Si bien el software provee otras configuraciones, esta fue la configuración que proporcionó un mejor rendimiento a lo largo de todas las pruebas.

3.6. Verificación de resultados.

Concluido el entrenamiento de la red MLP, el siguiente paso fue el análisis de los resultados y la verificación de los mismos de acuerdo a lo esperado. En esta fase se estudiaron tanto los resultados referentes al esquema de simulación de la interfaz aire del estándar UMTS, como los resultados arrojados por el clasificador en relación a los datos de prueba seleccionados para este fin.

Capítulo IV

Análisis, Interpretación y Presentación de los Resultados

4.1. Simulador de la interfaz aire del estándar UMTS.

A continuación se muestran los resultados referentes a la construcción del simulador de la interfaz aire para el estándar UMTS. El esquema final consta de una etapa transmisora, el canal RF y una etapa receptora, igualmente se aprecian los bloques que permiten colocar y variar los parámetros que afectan la calidad de la voz, los cuales fueron expuestos minuciosamente en los capítulos anteriores. Una explicación detallada de la construcción de este esquema de simulación, así como la configuración usada en cada uno de los bloques que en él se encuentran, se puede encontrar en el Anexo A de este Trabajo de Grado. En la figura 4.1 se puede observar el esquema final.

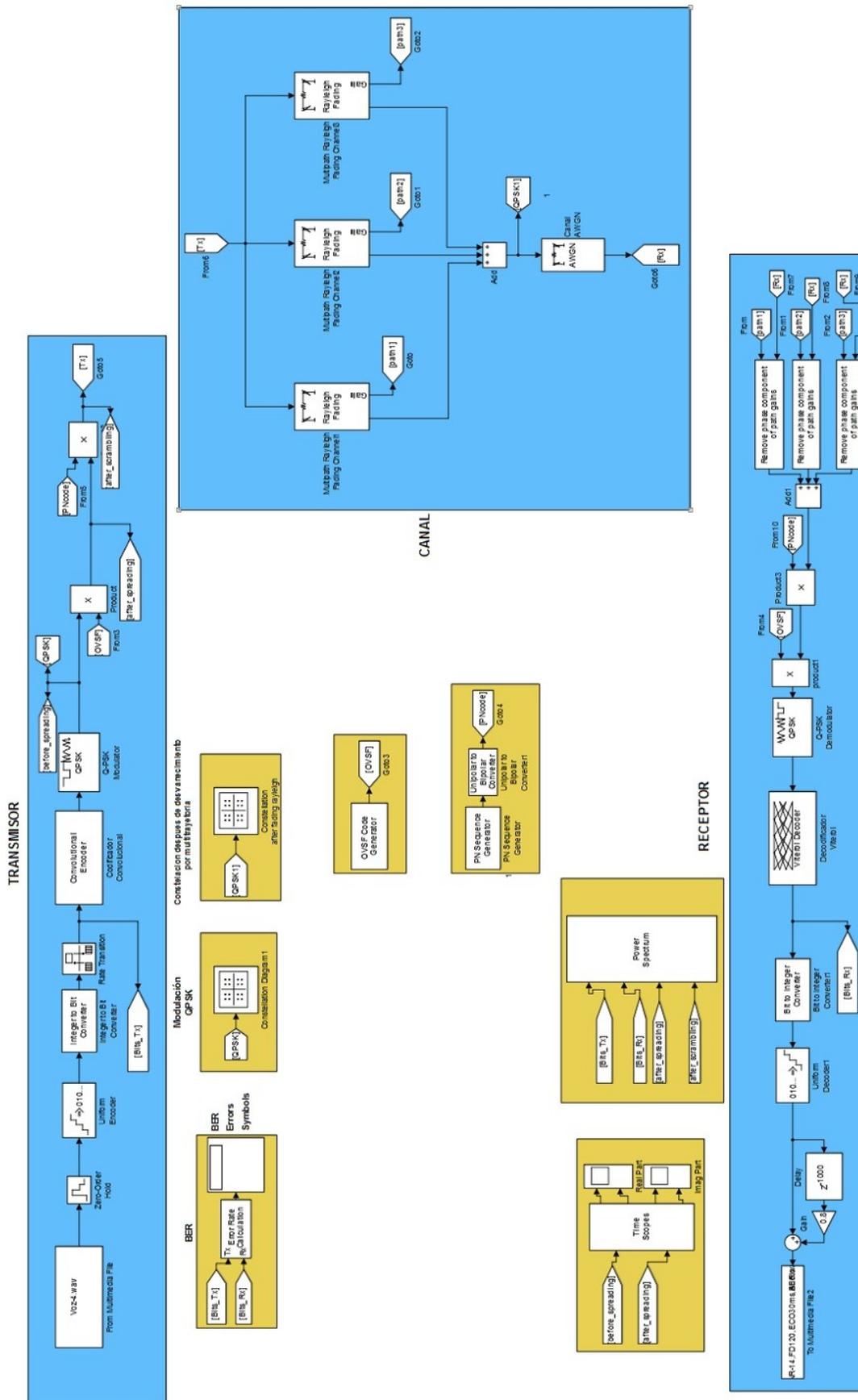


Figura 4.1: Simulador de la interfaz aire para el estándar UMTS

Como ya se dijo anteriormente, en el análisis de este esquema se estudiaron las relaciones entre la BER y los parámetros que afectan la calidad de la voz en el canal RF, es decir BER vs SNR(dB), BER vs F_d (Hz) y BER vs $E_{c/o}$ (ms). Con la finalidad de observar el aumento de la misma a medida que se van variando los niveles de las perturbaciones del canal, y de esta forma analizar cuales son sus consecuencias respecto al audio procesado a través del esquema de simulación. En la figura 4.2 se puede observar la grafica de BER vs SNR.

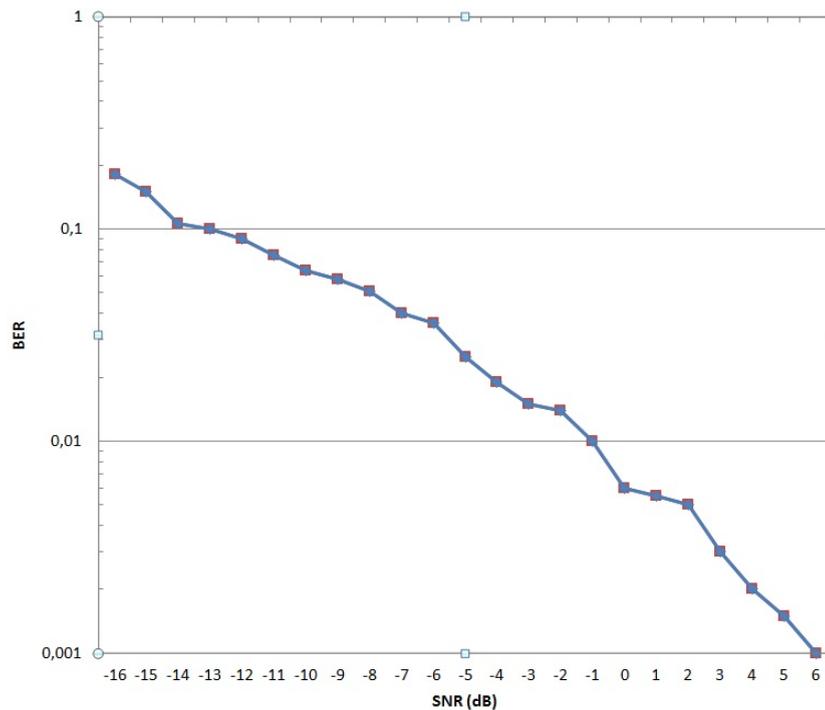


Figura 4.2: Gráfica de la BER vs SNR(dB)

Se puede apreciar que para valores cada vez más pequeños de SNR, se obtienen valores para la BER cada vez más altos, lo cual es un resultado esperado por previos estudios. Valores negativos de SNR son comunes en las técnicas de acceso que trabajan con el ensanchamiento del espectro, pueden ser causados por la introducción del ruido pseudoaleatorio, así como por las largas distancias de separación entre el UE y el Nodo B.

Igualmente es necesario el análisis referente al funcionamiento de la técnica de acceso con la cual trabaja la interfaz aérea del estándar, como ya se dijo anterior-

mente la técnica del espectro ensanchado produce el aumento en la velocidad de bits al multiplicar la información por el código de ensanchamiento, convirtiéndolos en chips, y el proceso inverso, es decir el desensanchamiento permite la recuperación de la información al multiplicar la señal nuevamente por el código de ensanchamiento, el proceso de ensanchamiento se puede apreciar en la figura 4.3.

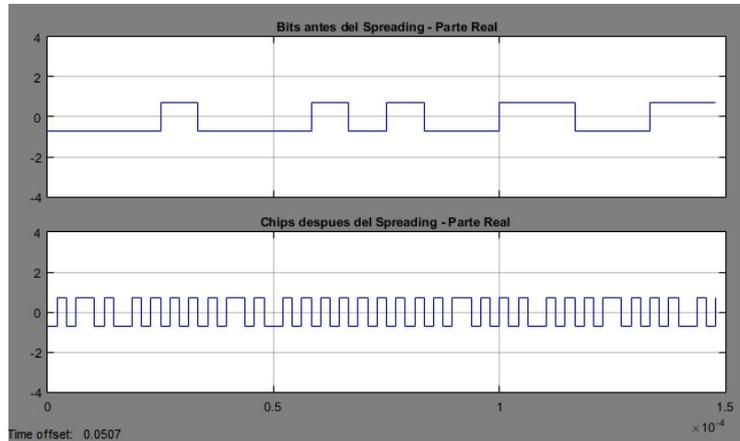


Figura 4.3: Señal antes y despues del Spreading

4.2. Clasificador de voces de forma objetiva.

EL clasificador de la voz es el resultado final de este Trabajo de Grado. Una vez finalizada la construcción de esquema de simulación para el procesamiento de las voces, fue necesario llevar a cabo dos importantes tareas, la construcción del banco de muestras o datos y el entrenamiento de la red MLP con dichos datos.

El procedimiento usado para la construcción del banco ha sido explicado con detalle en el Capítulo 3, en esta sección procederemos a mostrar el resultado del cuestionario enviado a la población del estudio. La figura 4.4 muestra un ejemplo de clasificación de una muestra de audio enviada a un grupo de personas.

AUDIO 6 (7 respuestas)

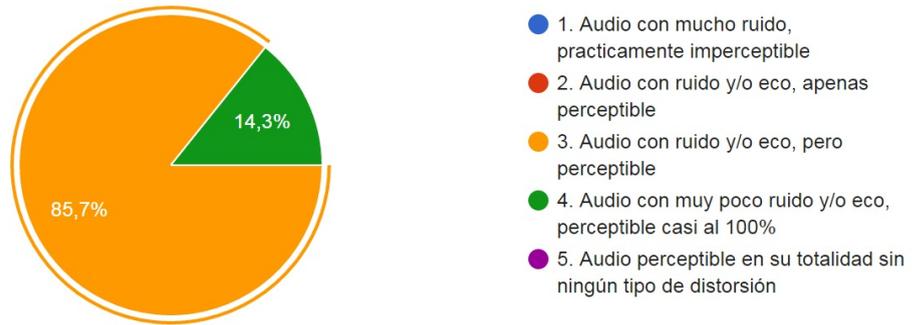


Figura 4.4: Clasificación de una muestra de audio

El valor que se le asigna a la muestra de audio es aquel que ha sido seleccionado por la mayor cantidad de personas.

Mediante el uso de este cuestionario se obtuvieron las respuestas para la construcción del banco de muestras, 660 para el entrenamiento y 60 para las pruebas, para un total de 720 muestras. Una representación de esto se observa en la tabla 4.1.

Muestra	SNR(dB)	Fd(Hz)	Eco(ms)	Clasificación
1	0	0	10	5
2	-4	60	50	2
3	-14	60	70	1
4	4	120	30	4
5	6	20	10	5
6	-8	40	0	3
7	-10	40	50	2
8	2	80	70	3
8	2	80	70	3
9	-16	0	10	1
10	0	20	300	3
11	0	4	50	5
12	-4	40	150	3
13	2	80	10	4
14	6	250	30	4
15	-14	80	10	1

Tabla 4.1: Representación del banco de muestras

El clasificador final ha sido construido haciendo uso de una red neuronal profunda del tipo Perceptrón Multicapa, con las siguientes características: arquitectura Feed-Forward con algoritmo de Backpropagation; 2 capas ocultas, 14 neuronas en la primera capa y 6 neuronas en la segunda, siguiendo las especificaciones explicadas en el Capítulo 3; optimización del entrenamiento mediante el algoritmo de Levenberg-Mardquardt y función de activación Tangente Sigmoidal. Cabe destacar que se llegó a esta configuración final luego de diversas pruebas. Un esquema de la estructura del MLP se muestra en la figura 4.5.

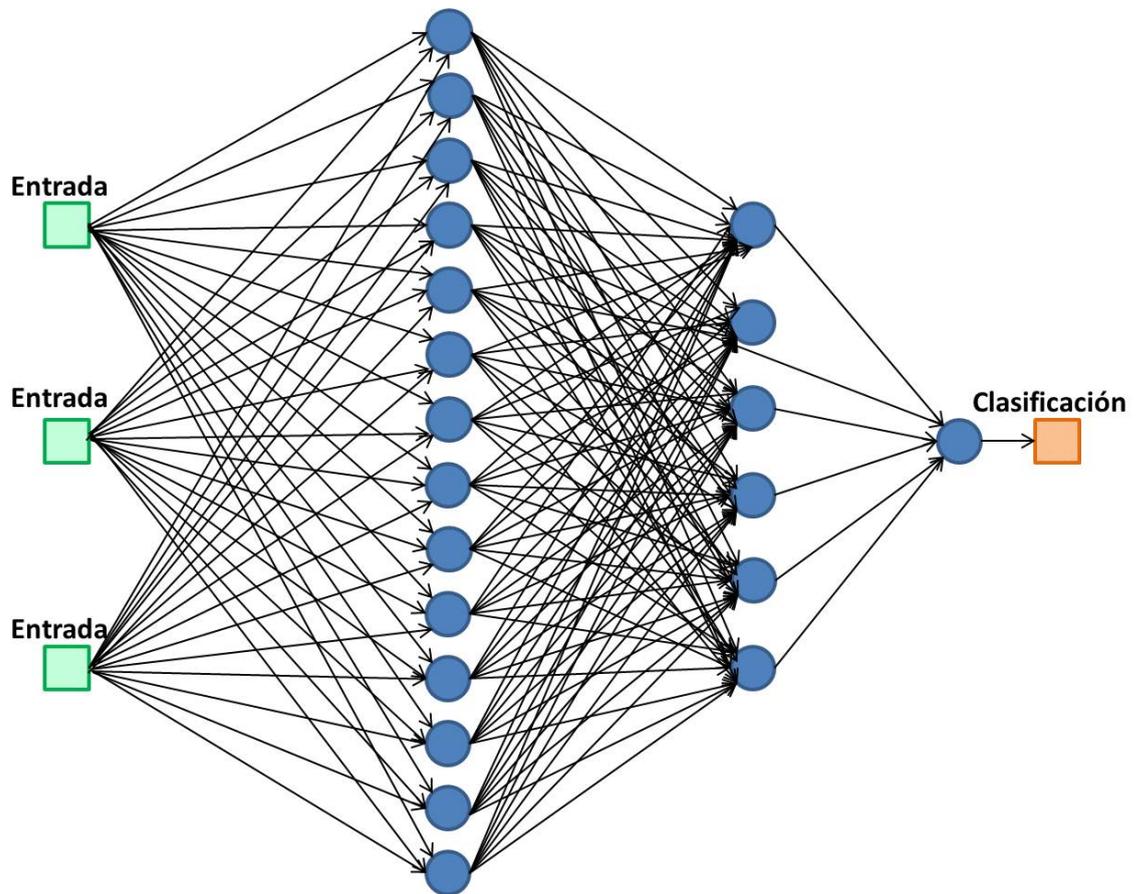


Figura 4.5: Estructura del MLP

La estructura de la red muestra una capa de entrada, la capa de entrenamiento con 2 capas ocultas y la capa salida para la clasificación.

Finalmente se tiene la interfaz gráfica del clasificador en el software de simulación.

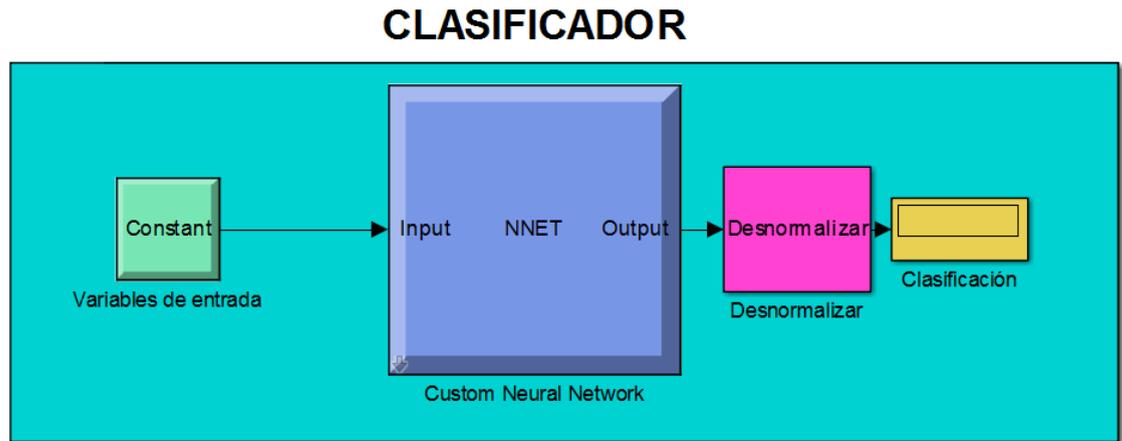


Figura 4.6: Clasificador final

Los parámetros a evaluar para definir la calidad de la voz se colocan manualmente en el bloque de entrada y el resultado de la clasificación se obtiene a la salida en el display.

La validación del funcionamiento del clasificador se ha hecho mediante el procesamiento de los datos de prueba a través del mismo, la tabla 4.2 muestra los resultados esperados de acuerdo a los datos recolectados y los arrojados por el clasificador.

Muestra	Esperado	Clasificador	Acertado
1	3	3	Si
2	2	2	Si
3	5	5	Si
4	2	2	Si
5	4	4	Si
6	4	4	Si
7	1	1	Si
8	1	1	Si

Muestra	Esperado	Clasificador	Acertado
9	4	4	Si
10	1	1	Si
11	2	2	Si
12	5	5	Si
13	4	5	No
14	2	2	Si
15	3	3	Si
16	1	1	Si
17	3	3	Si
18	4	4	Si
19	1	1	Si
20	1	1	Si
21	3	3	Si
22	2	2	Si
23	2	2	Si
24	4	5	No
25	4	4	Si
26	4	4	Si
27	5	5	Si
28	1	1	Si
29	4	4	Si
30	1	1	Si
31	4	4	Si
32	2	2	Si
33	2	2	Si
34	2	2	Si
35	4	4	Si
36	4	4	Si
37	5	5	Si

Muestra	Esperado	Clasificador	Acertado
38	1	1	Si
39	1	1	Si
40	4	4	Si
41	2	2	Si
42	2	2	Si
43	2	2	Si
44	4	4	Si
45	4	4	Si
46	2	2	Si
47	1	1	Si
48	1	1	Si
49	4	4	Si
50	1	1	Si
51	4	4	Si
52	5	5	Si
53	1	1	Si
54	1	1	Si
55	4	4	Si
56	4	4	Si
57	4	4	Si
58	4	4	Si
59	5	4	No
60	2	2	Si

De las 60 muestras procesadas a través del clasificador, se obtuvieron 3 errores (muestras 13, 24 y 59), lo que supone un 5 % del total de los datos de prueba, es decir que el clasificador acertó, según lo esperado en un 95 %, lo cual se puede considerar bastante bueno, considerando que es muy difícil obtener un 100 % de efectividad en este tipo de estudios.

Capítulo V

Conclusiones y recomendaciones

5.1. Conclusiones

El uso de la inteligencia artificial para clasificar la calidad de la voz en las llamadas, resultó ser una metodología bastante práctica. El método desarrollado permite calcular la experiencia del usuario con respecto a los factores que afectan la calidad de la voz en el canal de RF.

- La técnica de inteligencia artificial usada fue deep learning por sus grandes avances en la actualidad y su completa variedad de herramientas. Específicamente se usó la técnica más básica, el Perceptron Multicapa o MLP, esto debido a que la exigencia para obtener los resultados no era lo suficientemente considerable para el uso de herramientas de aprendizajes más avanzadas. Hay que destacar que se pudo haber implementado redes neuronales sencillas para la elaboración de este proyecto pero los resultados obtenidos no eran los mejores en comparación a lo anterior.
- La creación del esquema WCDMA en el software se llevó a cabo simulando solamente la interfaz aire, que es donde justamente la señal de voz se ve afectada mayormente por los diversos parámetros presentes en el canal. Los

canales de transporte pertenecientes a WCDMA no se agregaron al esquema porque se encargan del transporte de datos exclusivamente.

- El estándar UMTS-WCDMA utiliza el codec de audio .amr el cual no se pudo utilizar ya que fue una limitante del software usado, debido a que el mismo no reconoce ese tipo de formato de audio. Por lo que se trabajó con formatos .mp3 y .wav.
- Hay múltiples factores en el canal que pueden afectar la señal de voz del usuario. Se seleccionaron un número de parámetros determinados que son los más relevantes y de mayor influencia: ruido blanco gaussiano; desvanecimiento por multitrayectoria y efecto doppler y el eco. Esto se debe a que desde un principio se estableció un número finito de muestras de audio respectivamente afectadas por una cantidad de parámetros donde sus valores varían de forma discreta en un rango establecido.
- Al definir el parámetro desvanecimiento por multitrayectoria se consideró dejar fijo solo tres trayectos que tomaría la señal, si se considera variar el número de trayectorias se torna más complejo el estudio aumentando el número de posibles combinaciones y por ende mayor cantidad de muestras de audios a crear.
- Al esquema WCDMA creado se le agregaron dos usuarios para realizar pruebas de validación de la teoría de códigos ortogonales OVSA. Como resultado se obtuvo una transmisión sin interferencia entre usuarios, y en sus respectivos receptores la señal recuperada correctamente.

5.2. Recomendaciones

- Estudiar la posibilidad de simular otras perturbaciones del canal RF, tales como las interferencias co-canal y canal adyacente, nombradas en el Capítulo 2, así como también ampliar el rango de valores para realizar un análisis más completo de la calidad de la voz en los sistemas de comunicaciones móviles.

-
- Estudiar las diferencias al simular el enlace de subida (uplink) de la interfaz aire del estándar, ya que solo bastaría con hacer modificaciones sencillas al esquema construido, tales como la modulación.
 - Realizar estudios basados en otros tipos de redes neuronales profundas, tales como las Redes Neuronales Convolucionales (CNN), utilizando los datos obtenidos a partir de las voces procesadas, y de esta manera comparar los resultados respecto a los presentados en este Trabajo de Grado.
 - Seguir realizando estudios de deep learning en diversos software debido a las constantes librerías agregadas y avances obtenidos por ser un tipo de inteligencia artificial que está en constante desarrollo.
 - Aplicar ésta metodología de medidas objetivas sobre el estándar LTE, considerando que las operadoras de telefonía móvil en Venezuela buscan implementar en un futuro este tipo de red a nivel nacional.

Referencias Bibliográficas

- [1] Moe Rahnema. *UMTS Network planning, optimization and inter-operation with GSM*, Copyright© 2008 by John Wiley & Sons. Pgs: 63-69. Inc, 2008. URL: https://www.researchgate.net/publication/285260360_UMTS_Network_Planning_Optimization_and_Inter-Operation_with_GSM.
- [2] Ojemeni Uzoma. «Correlation between Uplink Noise, Uplink Load and Call Drop Rate in a WCDMA Network. Foundation of Computer Science». En: *International Journal of Computer Applications* (2013). URL: <http://research.ijcaonline.org/volume66/number5/pxc3885881.pdf>.
- [3] Ericsson. *User Description, Speech Quality Factors*. Ericsson. 2011. URL: <http://documents.tips/documents/speech-quality-factors.html>.
- [4] ITU-T. *ITU-T P.800.1. SERIES P: TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS. Methods for objective and subjective assessment of quality*. ITU-T. 2003. URL: <https://www.itu.int/rec/T-REC-P.800-199608-I/es>.
- [5] Elkin Gabriel Muskus Elvis Eduardo Gaona García Miguel Antonio Ávila Angulo. *APROXIMACIÓN DE LA CALIDAD DE VOZ Y COBERTURA EN UNA RED GSM DE EMERGENCIA*. Universidad Miitar Nueva Granada, 2014. URL: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0124-81702014000200002.
- [6] Nobuhiko Kitawaki; y Takeshi Yamada. *Subjective and Objective Quality Assessment for Noise Reduced Speech*. University of Tsukuba, 2007. URL: <https://portal.etsi.org/stq/workshop2007presentations/kitawaki.pdf>.

- [7] José A. Díaz P; Felicito A. Escobar; Felix A. Caramutti. *Estudio y planeación para la migración de una red de telefonía móvil de segunda generación (GSM/GPRS), a una red de tercera generación (UMTS/WCDMA) en la jurisdicción de el Salvador*. Tesis de grado. Abril 2009. URL: <http://biblioweb.ufg.edu.sv/virtual/index.cgi?idregistro=10638&tipologia=1&vista=C&busqueda=estudio+y+planeacion+para+la+migracion+de>.
- [8] Carlos Bonilla Tomás González Bolívar Ortiz. *Redes UMTS*. Universidad Tecnológica de Panamá, 2000. URL: <https://ehumir.files.wordpress.com/2013/04/articulo-redes-umts.pdf>.
- [9] TS GPP. *3GPP TS 25.201 V5. 0.0 (Dec. 2001), "3rd Generation Partnership Project. Physical layer - General description*. 2001. URL: http://www.etsi.org/deliver/etsi_ts/125200_125299/125201/10.00.00_60/ts_125201v100000p.pdf.
- [10] ETSI. *ETSI TS 125 410 V4.1.0 (2001-06) Universal Mobile Telecommunications System (UMTS); UTRAN Iu Interface: General Aspects and Principles*. (3GPP TS 25.410 version 4.1.0 Release 4). ETSI. URL: http://www.etsi.org/deliver/etsi_ts/125400_125499/125410/04.01.00_60/ts_125410v040100p.pdf.
- [11] ETSI. *ETSI TS 125 420 V10.0.0 (2011-04) Universal Mobile Telecommunications System (UMTS); UTRAN Iur interface general aspects and principles*. (3GPP TS 25.420 version 10.0.0 Release 10). ETSI. URL: http://www.etsi.org/deliver/etsi_ts/125400_125499/125420/10.00.00_60/ts_125420v100000p.pdf.
- [12] ETSI. *ETSI TS 125 430 V3.4.0 (2000-12) Universal Mobile Telecommunications System (UMTS); UTRAN Iub Interface: General Aspects and Principles*. (3GPP TS 25.430 version 3.4.0 Release 1999). ETSI. URL: http://www.etsi.org/deliver/etsi_ts/125400_125499/125430/03.04.00_60/ts_125430v030400p.pdf.
- [13] Dr. Amin Babiker A/Nabi Mustafa Shahd Alfadhil Al obeid Omer. *UMTS vs. GSM QOS Parameters Comparison (SNR, BER)*. Al-Neelein University, Faculty

- of engineering, 2014. URL: <http://textlab.io/doc/1101124/umts-vs-gsm-qos-parameters-comparison--snr--ber->.
- [14] Firas Mohammed Ali Al-Raie. *Simulation of Multipath Fading Effects in Mobile Radio Systems*. The Polytechnic Higher Institute of Yefren, Yefren, Libya. URL: http://www.microwavejournal.com/ext/resources/BGDownload/1/6/Multipath_Fading.pdf?1475747862.
- [15] Albana Nogueira. *CANCELADORES DE ECO, Aplicación de filtros adaptivos*. URL: http://iie.fing.edu.uy/ense/asign/tes/materiales/monografias/CanceladoresEco_Albana.pdf.
- [16] ITU-T. *Rec. UIT-R P.372-7UIT-R P.372-7, Ruido radioeléctrico*. ITU-T. 2001. URL: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.372-7-200102-S!MSW-S.doc.
- [17] ZTE Corporation. *ZTE Manual for WCDMA Network Planning & Optimization Training*. First Edition. ZTE University, China. 2009. URL: <https://es.scribd.com/doc/131949841/ZTE-WCDMA-Network-Planning-and-Optimization-Training>.
- [18] TS GPP. *3GPP TS 25.213 V5. 0.0 (1999), "3rd Generation Partnership Project. Spreading and modulation (FDD)*. 2000. URL: ftp://www.3gpp.org/tsg_ran/TSG_RAN/TSGR_05/Docs/Pdfs/rp-99477.pdf.
- [19] Alan Fakhru. *Simulation of Third Generation CDMA Systems. Tesis Doctoral*. Virginia Polytechnic Institute & State University, Diciembre 1999. URL: https://theses.lib.vt.edu/theses/available/etd-122199-153028/unrestricted/fakhru_thesis.pdf.
- [20] TS GPP. *3GPP TS 25.223 V2.1.0 (1999), "3rd Generation Partnership Project. Spreading and modulation (TDD)*. 1999. URL: ftp://www.3gpp.org/tsg_ran/TSG_RAN/TSGR_04/Docs/Pdfs/rp-99328.pdf.
- [21] I Ahmed. *Ahmed, I. Scrambling Code Generation for WCDMA on the StarCore™ SC140/SC1400 Cores. Application Note, AN2254, Rev, 1. Freescale Semiconductor*. Freescale Semiconductor. 2004. URL: http://www.datasheetlib.com/datasheet/850662/an2254_motorola-semiconductor.html.

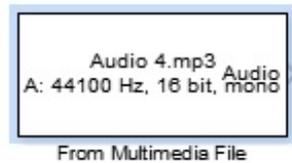
- [22] Ana Casali. *¿Qué es la Inteligencia Artificial?* Universidad Nacional de Rosario - Argentina, Junio 2007. URL: https://www.researchgate.net/profile/Ana_Casali/publication/268275299_Qu_es_la_Inteligencia_Artificial/links/551aa8450cf251c35b4e6fd0.pdf.
- [23] Jorge Aguilar J. *Introducción a las redes neuronales artificiales*. Flacso Andes Biblioteca Digital de Vanguardia, Junio 2008. URL: http://www.flacsoandes.edu.ec/comunicacion/aaa/imagenes/publicaciones/pub_25.pdf.
- [24] Sotolongo G; Gusmán Maria V. *Aplicaciones de las redes neuronales. El caso de la Bibliometría*. Centro de Investigación-Producción de Vacunas y Sueros. La Habana, Cuba, 2001. URL: <http://www.dynamics.unam.edu/DinamicaNoLineal/Articulos/MineriaDatos/Articulo03.pdf>.
- [25] Bryan García Navarro. *Implementación de Técnicas de Deep Learning*. Tesis de Grado. 2015. URL: <http://riull.ull.es/xmlui/bitstream/handle/915/1409/Implementacion%20de%20Tecnicas%20de%20Deep%20Learning.pdf?sequence=1>.
- [26] Hinton G; Deng L; Yu D; Dahl G. *Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups*. *Signal Processing Magazine, IEEE*, 29(6), 82-97. (2012). URL: <http://research.google.com/pubs/archive/38131.pdf>.
- [27] Jorge D. Martínez Pérez Rafael Gadea Gironés. *Redes neuronales y preprocesado de variables para modelos y sensores en bioingeniería*. Tesis Doctoral. Univesidad Politecnica de Valencia, 2012. URL: <https://riunet.upv.es/bitstream/handle/10251/16702/tesisUPV3874.pdf?sequence=1>.
- [28] Moritz August. «An Analysis of Deep Learning Architectures inspired by Neuroscience». En: *Department for Informatics, Technische Universitaet Muenchen* (2010). URL: http://www5.in.tum.de/persons/august/guided_research.pdf.
- [29] NORMA PELAEZ CHAVEZ. *APRENDIZAJE NO SUPERVISADO Y EL ALGORITMO WAKE-SLEEP EN REDES NEURONALES*. UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA, MEXICO, 2012. URL: http://jupiter.utm.mx/~tesis_dig/11612.pdf.

Anexo A

**Manual de construcción de un
simulador de la interfaz aire para
el estándar UMTS**

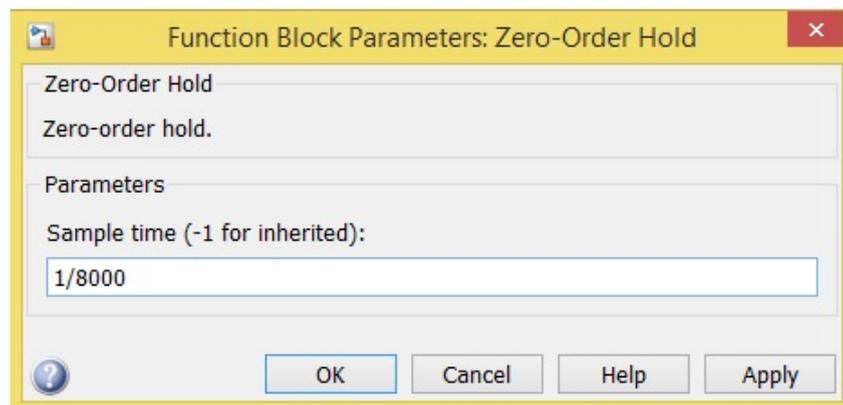
La elaboración del esquema WCDMA se fue desarrollando por etapas con el propósito de partir desde lo más básico hasta alcanzar lo más complejo. La etapa inicial comprende los siguientes bloques:

- From Multimedia File:



Por medio de este bloque se carga el archivo de audio (en cualquiera de los siguientes formatos .wav, .wma, .avi, .aif, .mp3, .flac, .ogg). El formato por defecto es .mp3.

- Zero-order hold:



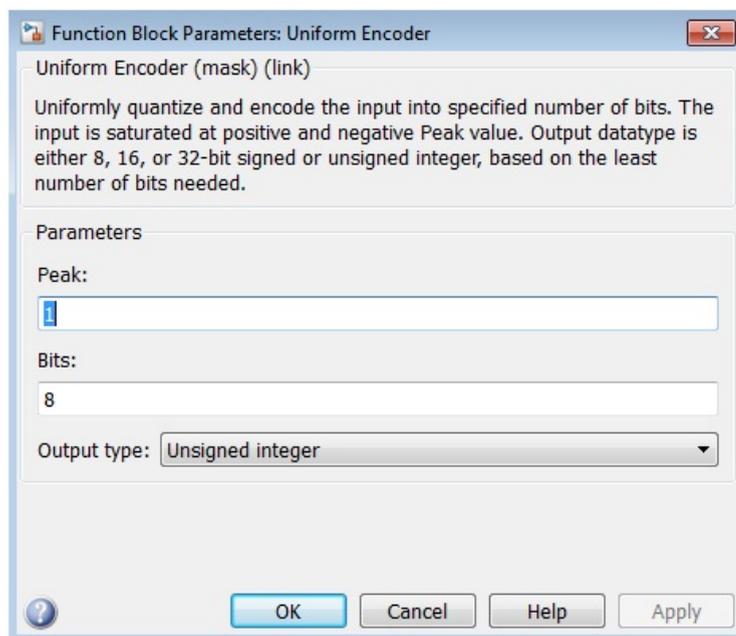
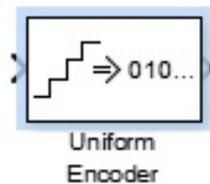
También conocido como muestreador. El tiempo de muestreo viene dado por

$$T_s = \frac{1}{2f},$$

y en telefonía el ancho de banda de la voz es 4KHz. Por ende el parámetro **Sample Time** se fijó en 1/8000. El proceso de muestreo consiste en tomar

muestras de la señal vocal a intervalos regulares. Estos intervalos deben ser tales que cumplan con el teorema de muestreo que establece: “La mínima frecuencia a la que puede ser muestreada una señal y luego reconstruida sin perder información es el doble de la frecuencia máxima de dicha señal”.

- Uniform Encoder

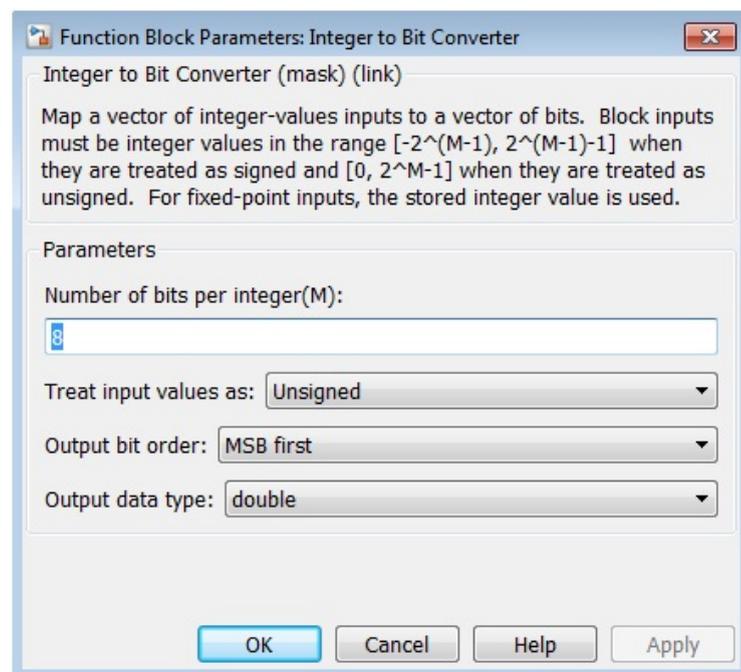


Denominado también cuantificador, es el segundo proceso dentro de la digitalización de una señal precedido por el muestro y seguido por la codificación. Luego de muestrear la señal analógica con la cual se ha discretizado la señal en tiempo, a continuación se aplicará lo mismo en amplitud. Por lo tanto se encarga de discretizar un rango continuo de amplitudes por aproximación o truncamiento de valores. Donde estos valores o niveles vienen dados por 2^{Bits} , asumiendo $\text{Bits} = 8$ se obtiene un total de 256 niveles. Los niveles de

cuantificación se eligen en función de la aplicación y del receptor, en este caso la señal es de voz, por lo que basta usar 256 niveles para que sea comprensible.

El parámetro **Peak** se fija en uno (1) para asignarlo como referencia, es decir, este sería el valor de amplitud más grande que pasaría por el codificador uniforme. En **Output type** se seleccionó **Unsigned integer** el cual establece que los valores tratados son enteros positivos.

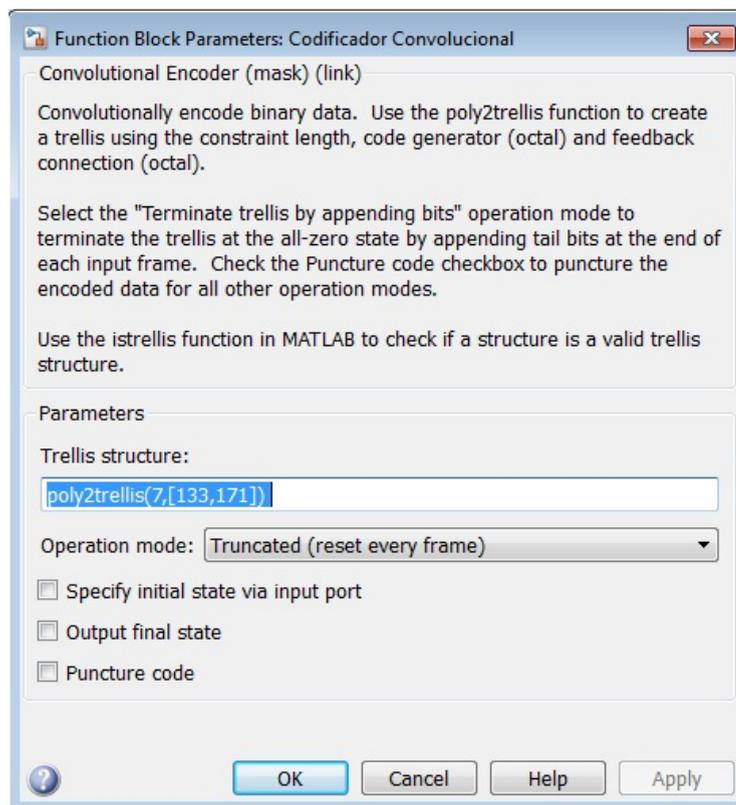
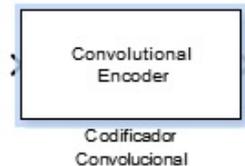
- Integer to Bit Converter:



Se encarga de convertir la señal de 256 niveles en pulsos comprendidos entre cero (0) y uno (1), es decir, a partir de este bloque la señal pasa a ser representada de forma binaria. Cada entero será representado por 8 bits. El parámetro **Treat input values as** se establece como **unsigned** e indica que los valores

tratados a la entrada son enteros positivos. **Output bit order** se define como **MSB first** y consiste en que los bits a la salida van en el orden desde más significativo a menos significativo.

- Convolutional Encoder:

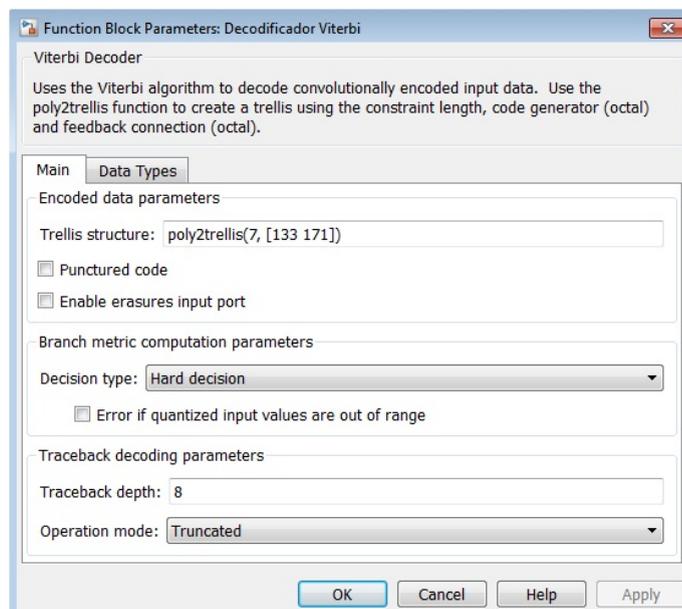


Codifica una secuencia de vectores de entrada binarios para producir una secuencia de salida de vectores binarios. La manera de definir el codificador convolutcional es usando la función **poly2trellis**, especificando la longitud de restricción y los polinomios generadores de códigos que por defecto se implementa de la siguiente manera **poly2trell(7,[133,171])**. El primer argumento de

la función indica la longitud de restricción de la codificación convolucional y los demás argumentos representan los polinomios generadores de códigos. La tasa de codificación se puede apreciar que es de $\frac{1}{2}$ por las secuencias generadoras establecidas. Cabe resaltar que la función considera estos valores suministrados de forma octal.

En la opción **Operation mode** se selecciona **Truncated**, esto indica la forma en cómo serán manejados los datos a la entrada del bloque. Se recomienda usar el modo de operación seleccionado porque el consiste en tratar cada entrada de forma independiente debido a que el bloque está expuesto a recibir entradas que longitud variable durante la simulación.

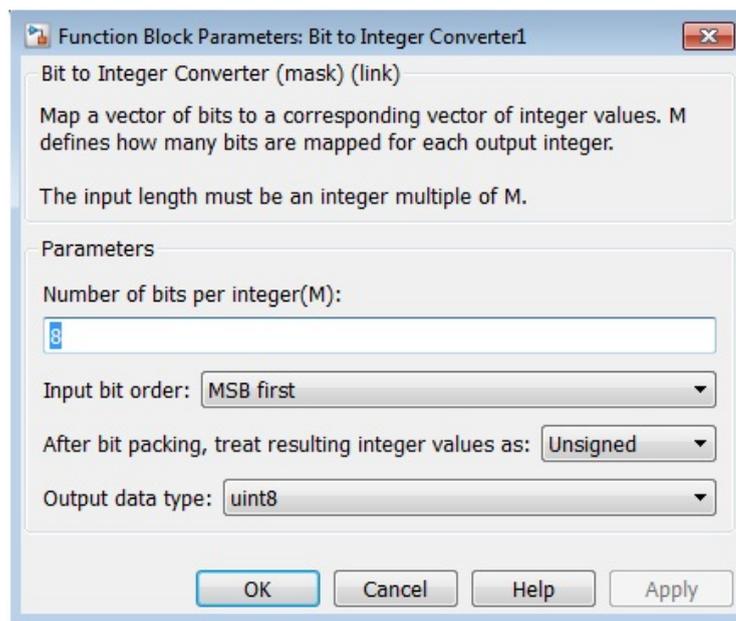
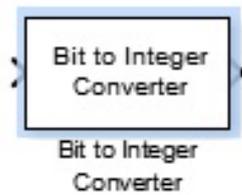
- Viterbi Decoder:



Decodifica los datos codificados por convolución utilizando el algoritmo de viterbi. Este bloque decodifica símbolos de entrada procesando varios a la vez

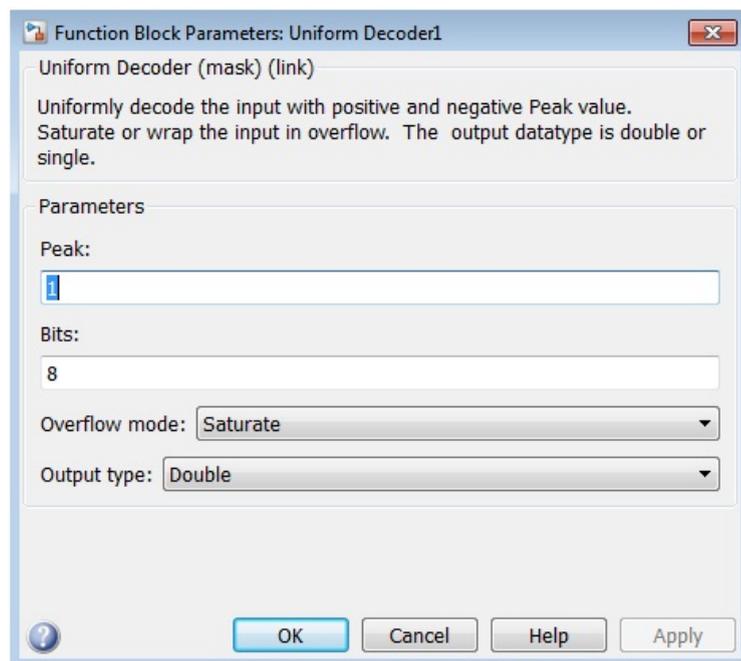
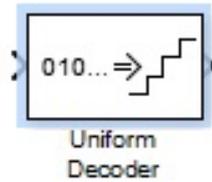
obteniendo un rendimiento más rápido para producir símbolos de salida binarios. Se usa la misma función con sus respectivos argumentos que se aplicó en la codificación. En **Decision type** se coloca la opción **Hard decision** el cual indica que los posibles valores a la entrada del decodificador es 0 ó 1 binarios adicionalmente implementa distancia de Hamming como cálculo de métrica.

- Bit to Integer Converter:



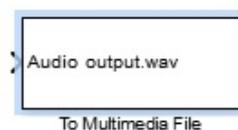
Realiza la operación inversa del bloque integer to bit converter.

- Uniform Decoder:



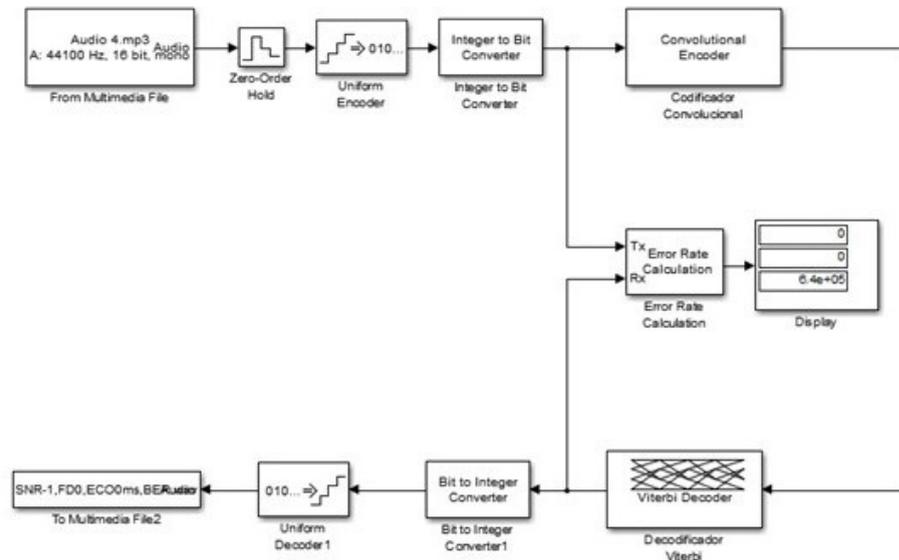
El bloque decodificador uniforme realiza la operación inversa del codificador uniforme.

- To Multimedia File:



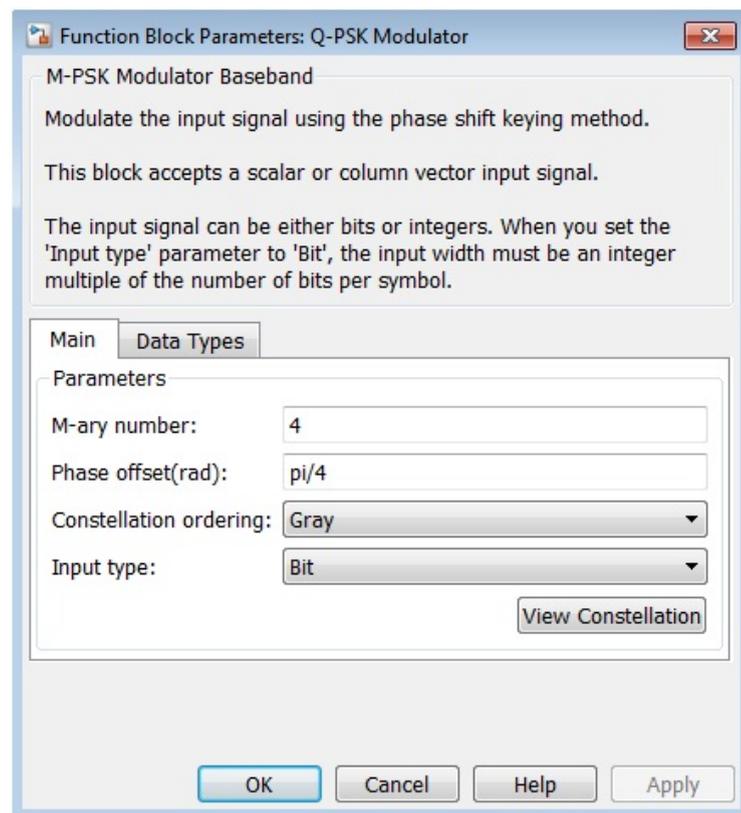
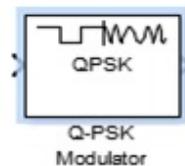
Este bloque permite asignar el nombre, tipo de formato y ubicación del archivo de audio recibido

El diseño completo de la etapa inicial representa una transmisión y recepción sencilla donde el audio pasa por un proceso de codificación convolucional a una tasa de $\frac{1}{2}$ y su respectiva decodificación mediante viterbi.



La segunda etapa del esquema se basó en incorporar el proceso de modulación y demodulación Q-PSK:

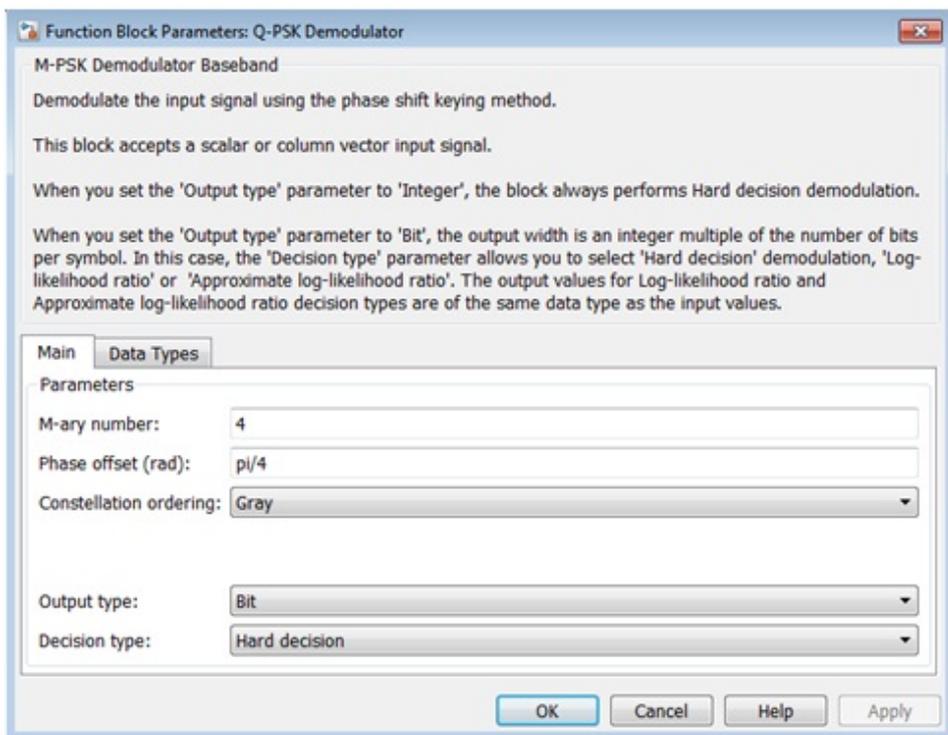
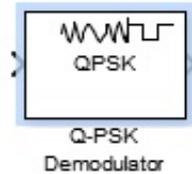
- QPSK Modulator:



La implementación de este bloque se debe a que el estándar UMTS en down-link se caracteriza por aplicar la modulación QPSK. Los parámetros a configurar son **M-ary number** el cual define el orden de la modulación y equivale al número de puntos en la constelación de la señal. El parámetro **phase offset** se refiere al desplazamiento de fase en radianes y en **Constellation ordering** se ajustó el parámetro a **Gray** debido a que en la asignación de bits a cada símbolo prevalece que entre dos símbolos adyacentes, estos solo se diferencian

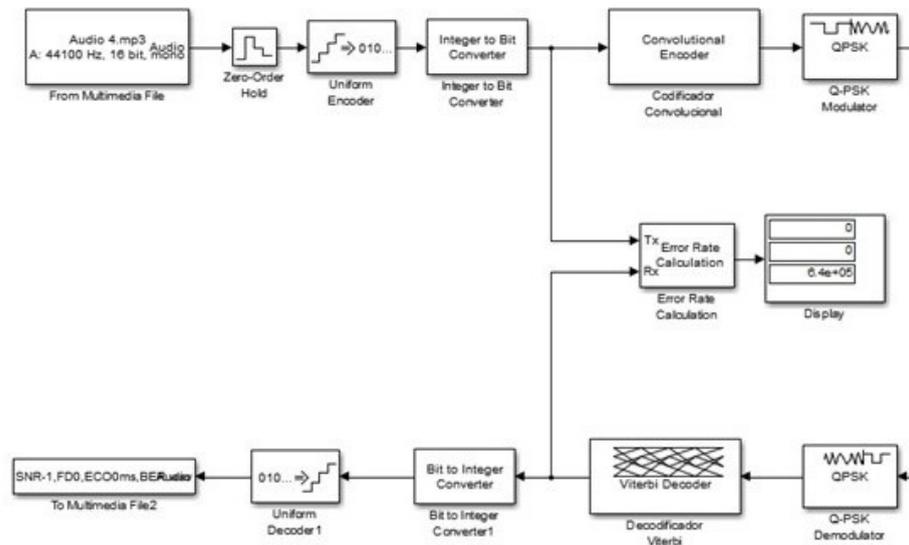
en un bit, con lo que se logra minimizar la tasa de bits erróneos.

- QPSK Demodulator:



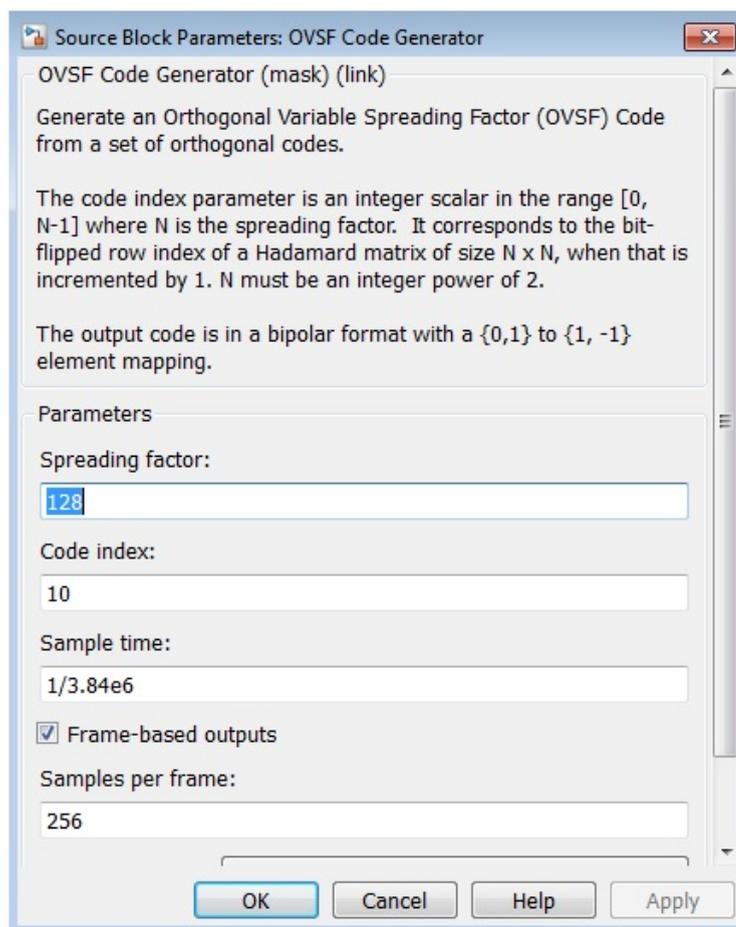
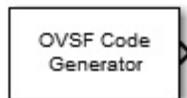
Se ajustan los mismos parámetros que en el bloque anterior.

Al agregar los bloques anteriormente mencionados y configurados se visualiza el esquema de la siguiente manera:



En la tercera etapa del montaje se llevó a cabo el ensanchamiento (Spreading) de la señal a transmitir y está constituida por los siguientes bloques:

- OVFSF Code Generator:



En WCDMA downlink el **Spreading Factor** (Factor de ensanchamiento también conocido por sus siglas SF) está comprendido en el siguiente rango de 4-512, valores obtenidos a partir de potencia base 2. El siguiente parámetro **Code index** (índice del código) es una valor entero desde 0 a SF-1 y específica

el código a usar de esa columna del árbol de códigos OVSF. Más adelante se explicará de forma más detallada lo antes descrito.

En **Sample time** por especificaciones del estándar se establece a una tasa de **1/(3.84Mchips)**, la palabra chip es un término usado en WCDMA que corresponde a los bits utilizados en el código de ensanchamiento. Y Por último el **samples per frame** indica el número de muestras por trama a la salida del bloque. Todos los valores ajustados a estos parámetros son tomados de referencias de las recomendaciones de la 3GPP.

Las siglas OVSF significan factor de ensanchamiento de variables ortogonales y fueron implementados por primera vez para los sistemas de comunicaciones 3G. Se utilizan principalmente para preservar la ortogonalidad entre los diferentes usuarios en un sistema de comunicación. Estos códigos se derivan de un árbol de código OVSF, y a cada usuario de le asigna un código diferente para separar las conexiones de los usuarios dentro de una misma celda. Un árbol de código OVSF es un árbol binario que refleja la construcción de matrices Hadamard.

$$H_1 = [1],$$
$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

y

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix} = H_2 \otimes H_{2^{k-1}},$$

Figura 1.1: Matriz Hadamard

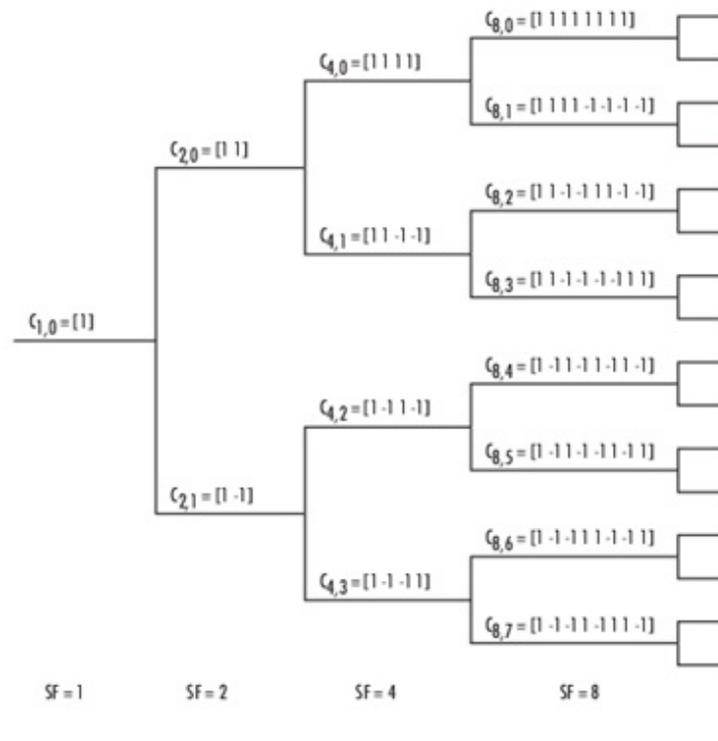


Figura 1.2: Estructura de árbol de los códigos OVFSF

Hay que tomar en consideración que dos códigos son ortogonales sí y solo sí ningún código se encuentra en la ruta desde el otro código a la raíz. A partir de la figura anterior se puede apreciar que los códigos $C_{4,0}$ y $C_{4,1}$ no son ortogonales a diferencia de $C_{4,0}$ y $C_{4,2}$ que sí lo son. Con respecto al índice del código en una estructura de árbol depende del SF seleccionado, por ejemplo si se establece $SF=4$ (ver figura 2) hay disponible cuatro posibles códigos en toda esa columna para dicho factor de ensanchamiento. Por lo tanto de acuerdo al índice de código seleccionado que en este caso puede tomar un valor entero entre 0 y $SF-1 = 3$, se determinaría el código OVFSF a implementar para el Spreading.

En el receptor para poder recuperar la información, la señal recibida tiene que multiplicarse con el mismo código que se usó para la transmisión de la misma. El principio de De-Spreading se observa a continuación:

El UE1 utiliza el c_1 para el De-spreading

$$(UE1 * c1 + UE2 * c2 + UE3 * c3) * c1 = UE1 * (c1 * c1) + UE2 * (c2 * c1) + UE3 * (c3 * c1)$$

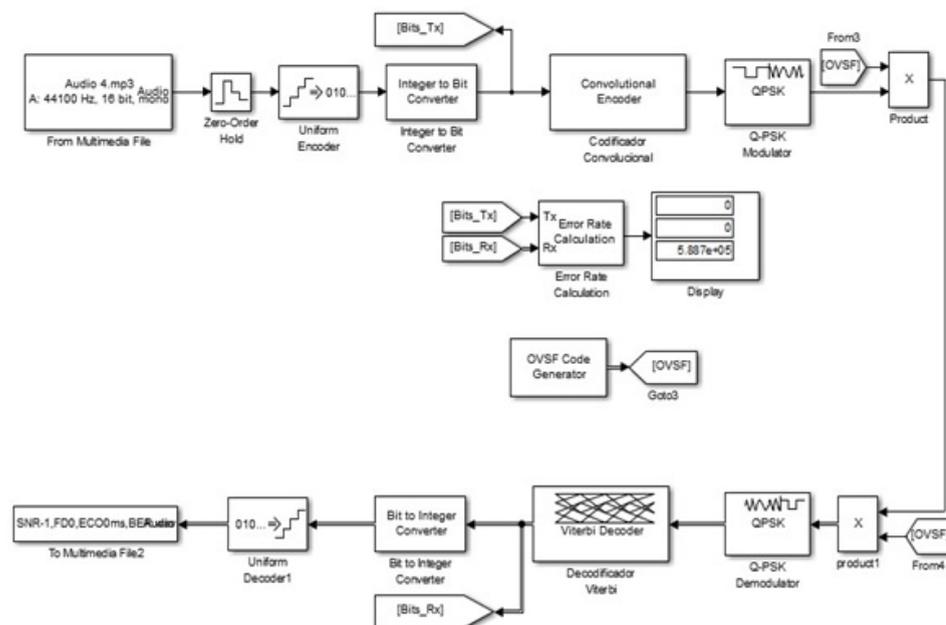
$$(UE1 * c1 + UE2 * c2 + UE3 * c3) * c1 = UE1 * 1 + UE * 0 + UE3 * 0$$

$$(UE1 * c1 + UE2 * c2 + UE3 * c3) * c1 = UE1$$

Se puede detallar que el código $c1$ al realizar el producto con señales que fueron ensanchadas con un código distinto no las recupera, a diferencia de efectuar esa multiplicación con señales que usaron el mismo código $c1$ al momento de ser transmitidas como lo es el caso del equipo de usuario 1 (UE1). De la misma forma, UE2 utiliza $c2$ para el proceso de De-spreading, y UE3 a $c3$, para de esta forma obtener sus propias señales

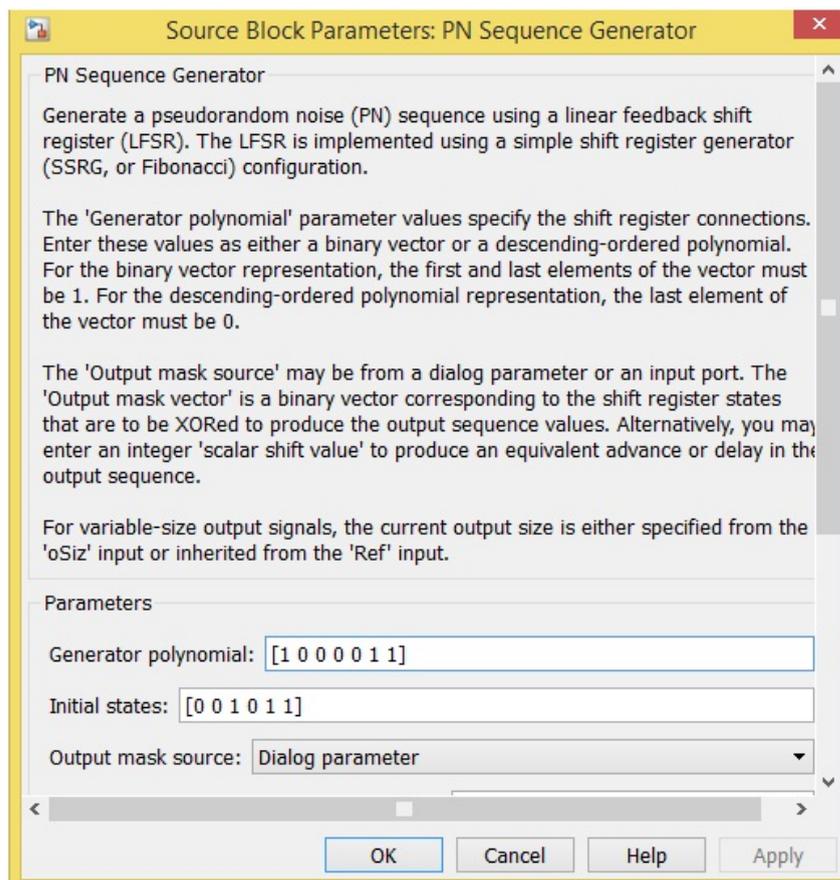
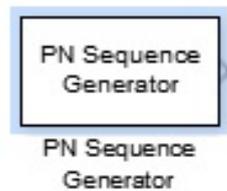
Los códigos ortogonales en el downlink dentro de cada BS (Base Station o también conocido como Nodo B) son administrados por el RNC (Controlador de la Red Radio) en la red.

La siguiente ilustración constituye el esquema WCDMA-Downlink con los códigos OVSF que son los encargados de aumentar el ancho de banda a 3.84MHz como se explicó previamente.



Posteriormente se agrego al esquema el código de scrambling el cual es un procedimiento adicional en el transmisor para separar las distintas celdas. Este código es utilizado sobre el ensanchamiento de tal forma que no cambia el ancho de banda y cabe destacar que con scrambling, no importa si los mismos códigos de ensanchamiento son utilizados por diferentes transmisores.

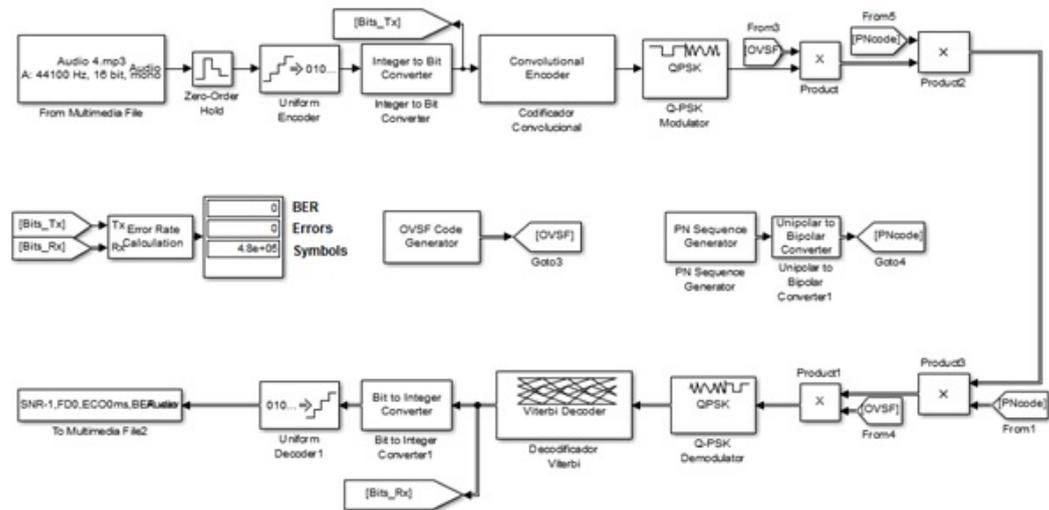
- PN Sequence Generator



El parámetro **Generator polynomial** como su nombre lo indica es un polinomio generador primitivo (de manera que sus coeficientes tengan uno como su

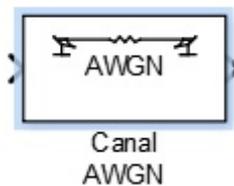
máximo común divisor). El vector ([1 0 0 0 1 1]) usado en la configuración del parámetro antes descrito representa el siguiente polinomio $z^6 + z + 1$ y el parámetro **Initial States** se refiere al valor inicial de los registros satisfaciendo los siguientes criterios que todos los elementos de ese vector deben ser números binarios y su longitud debe ser igual al grado del polinomio generador.

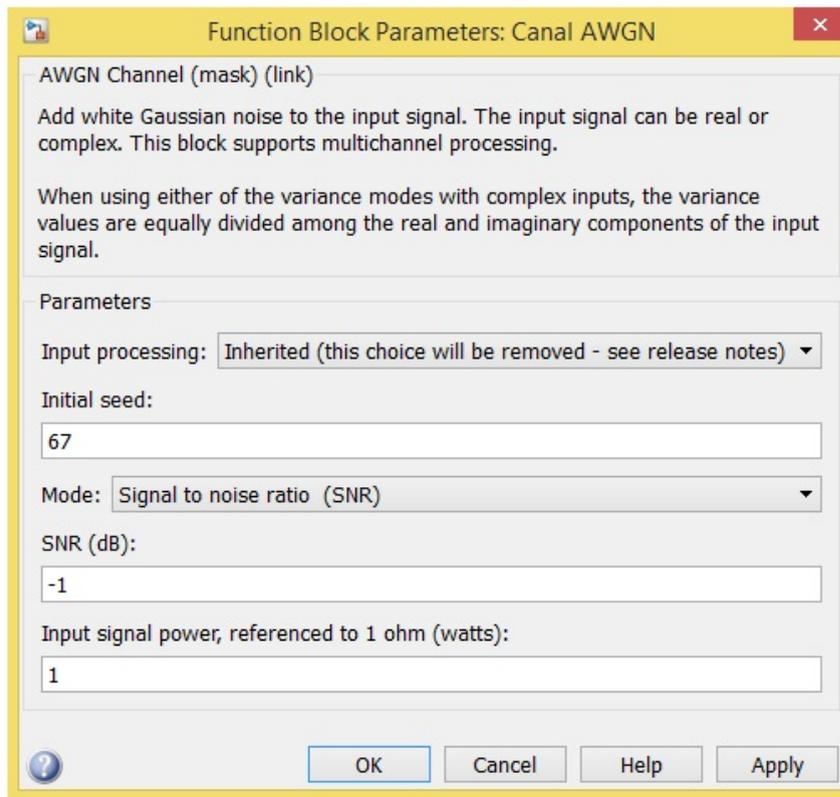
En Scrambling se emplean secuencias unicas pseudo-aleatorias o de pseudo-ruido (PN). Un código PN es una secuencia binaria que parece aleatoria, pero que puede reproducirse de forma deterministica si el receptor lo necesita. Es decir, esta secuencia genera una serie de números binarios que se repiten sin seguir un patrón determinado ya que de lo contrario no podría recuperarse la información en el receptor si fuera totalmente aleatoria. En la imagen se observa el bloque PN Sequence Generator agregado al esquema.



La última etapa en la elaboración del esquema consiste en agregar el canal y los diversos factores que lo afectan:

- Canal AWGN

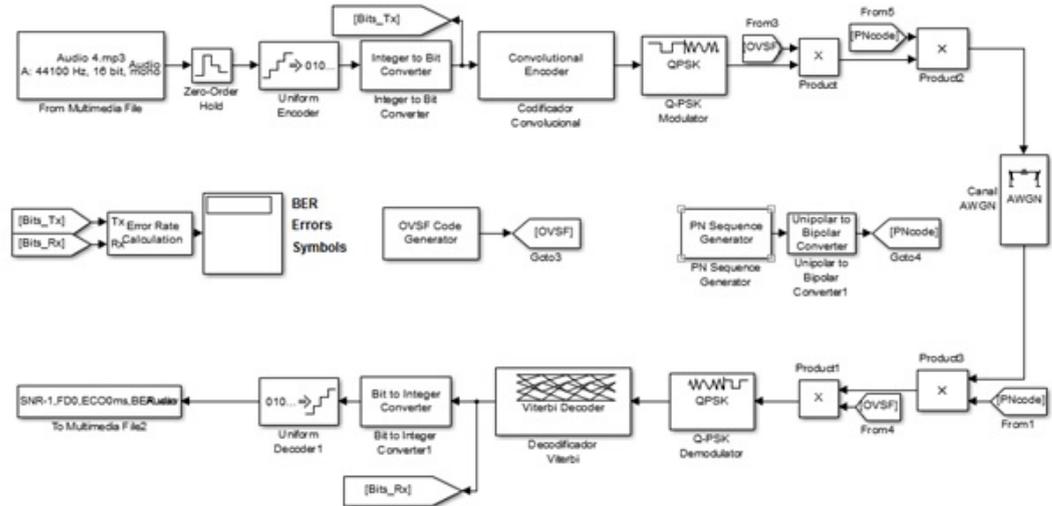




Este bloque a parte de cumplir la función de canal agrega ruido gaussiano blanco a la señal que pasa a través de él. Es por eso las siglas AWGN (Ruido Aditivo Blanco Gaussiano).

El parámetro **Initial seed** se considera como la semilla para generar ruido gaussiano a partir de una fuente aleatoria. Mayormente se deja establecido el valor asignado por defecto. En **Mode** se especifica la varianza del ruido generado por ese bloque para este caso se seleccionó la relación de potencia de señal a potencia de ruido (SNR), en decibelios. A éste parámetro posteriormente se le asignara un rango de valores para un análisis más detallado. El último parámetro hace referencia a la potencia media cuadrática de las muestras de entrada en watts y se deja establecido ese valor predeterminado.

La siguiente imagen muestra el esquema final WCDMA con el canal RF.



Anexo B

Manual de construcción de una red neuronal profunda del tipo MLP para la clasificación de voces

La construcción detallada del clasificador se explica a continuación. Una vez finalizado el esquema de simulación de la interfaz aire del estándar UMTS, el siguiente paso es procesar las muestras de audio a través del esquema, realizando las múltiples combinaciones posibles en los niveles de las distorsiones del canal RF, en nuestro caso hasta completar las 720 muestras establecidas por el diseño experimental. Estas muestras son divididas en grupos y son enviadas junto con un formulario, a una población de unas 200 personas que se encargaran de dar la clasificación de cada una de estas, para de esta forma construir el banco de muestras finales, el cual se distribuirá de la siguiente forma: 660 muestras para el entrenamiento de la red, y 60 para probar el funcionamiento una vez finalizado el entrenamiento. En la figura se muestra un ejemplo del cuestionario utilizado.

Clasificación de voces

En el siguiente cuestionario se les ha suministrado un total de 10 muestras de audio para su calificación del 1 al 5, de acuerdo a la calidad con la que se escuchan dichas muestras.

*Obligatorio

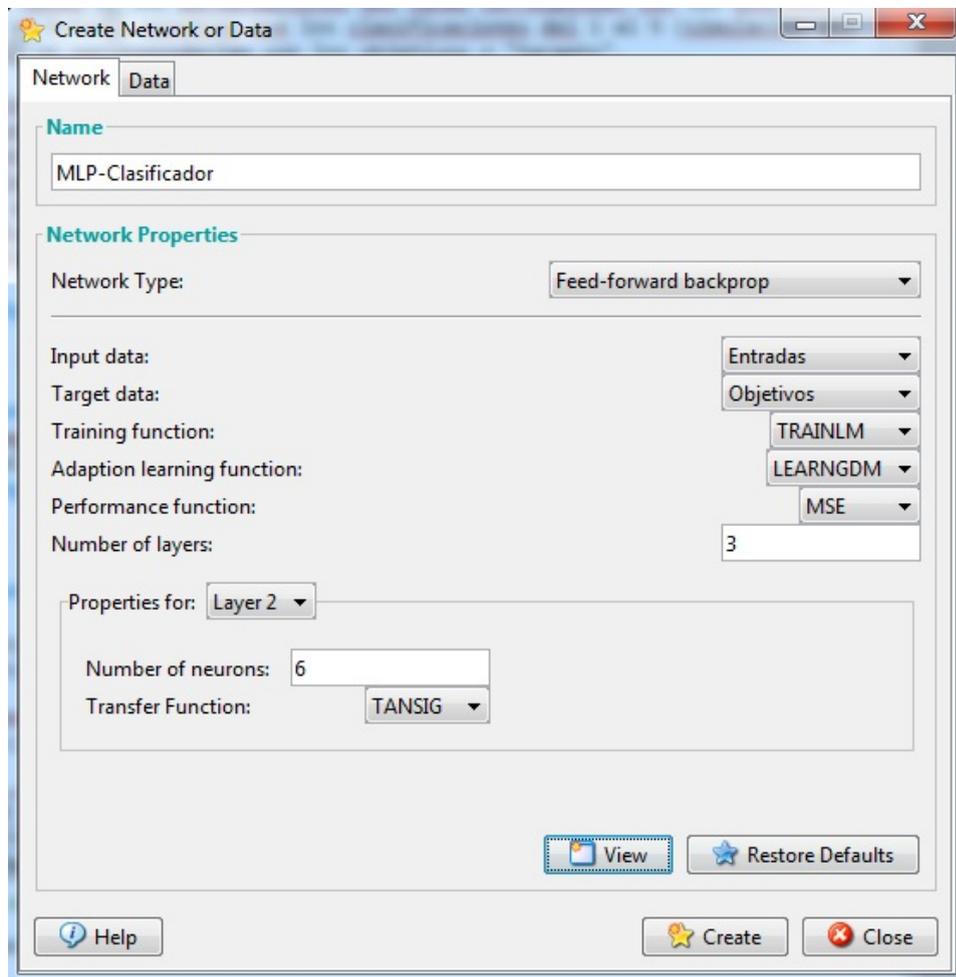
AUDIO 1

- 1. Audio con mucho ruido, practicamente imperceptible
- 2. Audio con ruido y/o eco, apenas perceptible
- 3. Audio con ruido y/o eco, pero perceptible
- 4. Audio con muy poco ruido y/o eco, perceptible casi al 100%
- 5. Audio perceptible en su totalidad sin ningún tipo de distorsión

Una vez que se obtuvieron todos los datos, se procedió a importarlos para bajarlos en el software de simulación.

Los datos de las perturbaciones del canal corresponden con las entradas (3 en total), mientras que los clasificaciones del 1 al 5 (simulación del MOS) se corresponderían con los objetivos o "targets".

Respecto a la configuración general de la red, al tratarse de un MLP, la misma será una red de Aprendizaje Supervisado (LEARNGDM), su estructura Feed-Forward con Backpropagation, la cantidad de capas ocultas se ha fijado en 2, con 14 y 6 neuronas para cada capa oculta respectivamente, el algoritmo de optimización Levenberg Mardquardt (TRAINLM), criterio de reducción del MSE y la función de activación la Tangente Sigmoidal (TANSIG). Dicha configuración se muestra en la siguiente imagen.



Luego de esto es necesario llevar a cabo el entrenamiento, la herramienta de redes neuronales se encargará de este proceso.

The screenshot shows the 'Neural Network Training (nntool)' window. At the top, the 'Neural Network' section displays a diagram of a feedforward neural network with 3 input nodes, two hidden layers (14 and 6 nodes), and 1 output node. Each layer includes weight (W) and bias (b) inputs, a summation node (+), and an activation function block.

The 'Algorithms' section lists the following settings:

- Data Division: Random (dividerand)
- Training: Levenberg-Marquardt (trainlm)
- Performance: Mean Squared Error (mse)
- Calculations: MATLAB

The 'Progress' section shows a table of training metrics:

Metric	Current Value	Target Value
Epoch	0	1000
Time	0:00:19	
Performance	0.00909	0.00299
Gradient	0.0105	1.00e-07
Mu	0.00100	1.00e+10
Validation Checks	0	268

The 'Plots' section includes buttons for 'Performance' (plotperform), 'Training State' (plottrainstate), and 'Regression' (plotregression). A 'Plot Interval' slider is set to 1 epoch.

At the bottom, a status bar shows 'Training neural network...' with a progress indicator. Two buttons are visible: 'Stop Training' and 'Cancel'.

Finalmente se debe utilizar el comando **gemsim(network,1)**, donde el parámetro **network** es el nombre de la red entrenada, mediante el cual podremos llevar nuestra red, una vez exportada, a un modelo que se pueda simular y probar en tiempo real.

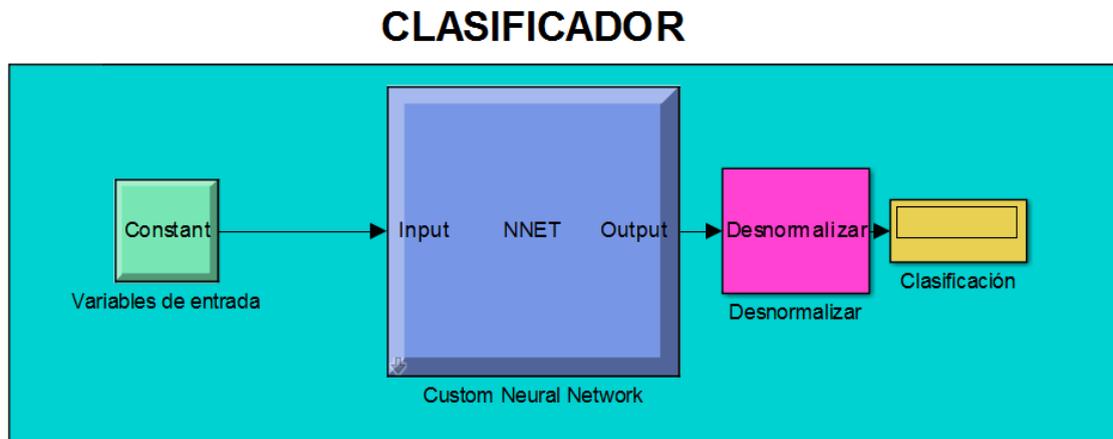


Figura 2.1: Clasificador final

Es importante resaltar que las respuestas dadas por el clasificador se encuentran normalizadas entre 0 y 1 debido a la función de activación que usa este, es por ello que es necesario agregar un bloque para llevar las respuestas a los valores reales de clasificación entre 1 y 5, para esto solo basta con multiplicar las salidas por 5.