



UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA MECÁNICA ELÉCTRICA

SISTEMA INTEGRADO DE MENSAJERÍA

CARLOS GABRIEL GÓMEZ VILLAGRÁN

ASESORADO: ING. ENRIQUE EDMUNDO RUIZ CARBALLO

GUATEMALA, ABRIL DE 2003

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

SISTEMA INTEGRADO DE MENSAJERÍA

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA POR

CARLOS GABRIEL GÓMEZ VILLAGRÁN

AL CONFERÍRSELE EL TÍTULO DE

INGENIERO ELECTRÓNICO

GUATEMALA, ABRIL DE 2003

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Sydney Alexander Samuels Milson
VOCAL I	Ing. Murphy Olympo Paiz Recinos
VOCAL II	Lic. Amahán Sánchez Alvarez
VOCAL III	Ing. Julio David Galicia Celada
VOCAL IV	Ing. Kenneth Issur Estrada Ruiz
VOCAL V	Br. Elisa Yazminda Vides Leiva
SECRETARIO	Ing. Pedro Antonio Aguilar Polanco

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Herbert René Miranda Barrios
EXAMINADOR	Ing. Edwin Alberto Solares Martínez
EXAMINADOR	Ing. Enrique Edmundo Ruiz Carballo
EXAMINADOR	Ing. Julio Cesar Solares Peñate
SECRETARIA	Inga. Gilda Marina Castellanos de Illescas

HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

SISTEMA INTEGRADO DE MENSAJERÍA

Tema que me fuera asignado por la Dirección de la Escuela de Ingeniería Mecánica Eléctrica con fecha 24 de febrero de 1998; No. Ref. EIME.053.98.

Carlos Gabriel Gómez Villagrán

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	xi
LISTA DE SÍMBOLOS	xv
GLOSARIO	xvii
RESUMEN	xxvii
INTRODUCCIÓN	xxix
1. SISTEMA BÁSICO DE RADIOLOCALIZACIÓN <i>PAGING</i>	1
1.1. Sistema básico	1
1.1.1. Codificador	2
1.1.2. Estación base	3
1.1.3. Receptores	3
1.2. Sistemas de radiolocalización	4
1.2.1. Sistema manual de radiolocalización	5
1.2.2. Sistemas automáticos de radiolocalización	6
1.2.2.1. Terminales de radiolocalización	6
1.2.2.2. Flujo de mensajes	6
1.2.2.2.1. Tono, voz y numérico	7
1.2.2.2.2. Alfanumérico	7
1.3. Cobertura de un sistema de radiolocalización	8
1.3.1. Factores que afectan la cobertura	9
1.3.1.1. Potencia del transmisor	10
1.3.1.2. Sensibilidad del receptor	10
1.3.1.3. Ganancia de antena	10
1.3.1.4. Altura de antena	11
1.3.1.5. Frecuencia de transmisión	11

1.3.1.6. Pérdidas en la trayectoria	12
1.3.1.7. Desvanecimiento	12
1.3.2. Transmisión simultánea	13
1.3.2.1. Ventajas operacionales de la transmisión simultánea	13
1.3.2.2. Requerimientos operacionales	14
1.3.3. Red de radiolocalización	15
1.4. Protocolos de señalización en radiolocalización (Formatos de codificación)	17
1.4.1. Formato análogo o digital	18
1.4.1.1. Formato de codificación por tonos	18
1.4.1.2. Formato de codificación binaria	19
1.4.1.3. <i>POCSAG (Post Office Code Standardization Advisory Group)</i>	20
1.4.1.4. <i>FLEXtm</i> Código de radiolocalización de alta velocidad	21
1.5. Soporte físico típico de la radiolocalización	22
1.5.1. Receptores o radiolocalizadores	22
1.5.1.1. Características comunes de los radiolocalizadores	23
1.5.2. Terminales de radiolocalización	27
1.5.2.1. Características del MPS2000 interruptor modular de radiolocalización	27
1.5.2.2. Características <i>Unipage</i>	28
1.5.2.3. Controladores	29
1.5.2.3.1. <i>C-net</i>	29
1.5.2.3.2. Núcleos transmisores	30
1.5.3. Estaciones base	30
2. Reconocimiento de voz	31
2.1. Reconocimiento de voz	33
2.1.1. Tipos de reconocimiento de voz	33
2.1.1.1. Palabra aislada y/o expresión simple	33

2.1.2.2.3.2.1.2.	Desarrollo de	
	banco de filtros	_68
2.1.2.2.3.2.1.2.1.	Respuesta de	
	impulso infinito	
	(<i>IIR</i>)	_____68
2.1.2.2.3.2.1.2.2.	Respuesta de	
	impulso finito	
	(<i>FIR</i>)	_____69
2.1.2.2.3.2.1.3.	Desarrollo de un	
	banco de filtros	
	uniforme basado	
	en la transformada	
	corta de <i>Fourier</i>	
	usando <i>FFT</i>	____74
2.1.2.2.3.2.1.4.	Desarrollo de un	
	banco de filtros <i>FIR</i>	
	no uniforme	____75
2.1.2.2.3.2.1.5.	Banco de filtros no	
	uniformes basados	
	en <i>FFT</i>	_____79
2.1.2.2.3.2.1.6.	Resumen de	
	consideraciones para	
	banco de filtros de	
	reconocimiento de	
	voz	_____79
2.1.2.2.3.2.2.	Análisis de predicción lineal para	
	el reconocimiento de voz (<i>LPC</i>)	81

2.1.2.2.3.7.2.2.1.3.	
Articulaciones	
sonoras y	
sordas	___ 100
2.1.2.2.3.7.2.2.2.	
Fonemas	
vocálicos	___ 101
2.1.2.2.3.7.3.	Módulo de análisis fonológico 103
2.1.2.2.3.7.4.	Módulo de análisis morfológico 103
2.1.2.2.3.7.5.	Módulo de análisis sintáctico _ 104
2.1.2.2.3.7.6.	Módulo de análisis semántico _ 104
2.1.2.2.3.7.7.	Módulo de análisis pragmático 105
2.1.2.2.3.8.	Procesos discretos de <i>Markov</i> _____ 107
2.1.2.2.3.8.1.	Características de un <i>HMM</i> ___ 109
2.1.2.2.3.8.2.	Reconocedor basado en
<i>HMM</i> 's	_____ 110
2.1.2.2.3.9.	Codificación de voz _____ 111
2.1.2.2.3.9.1.	<i>Vocoder</i> de canal _____ 114
2.1.2.2.3.9.2.	<i>Vocoder</i> de formantes _____ 115
2.1.2.2.3.9.3.	<i>Vocoder</i> articulatorio _____ 115
2.1.2.2.3.9.4.	<i>Vocoder</i> homomórfico _____ 115
2.1.2.2.3.9.5.	<i>Vocoder LPC</i> _____ 115
3. SISTEMA IRIDIUM	_____ 117
3.1. Constelaciones de satélites de órbita baja (<i>LEO</i>)	_____ 117
3.1.1. Satélites geoestacionarios (<i>GEO</i>)	_____ 117
3.1.2. Satélites de órbita media (<i>MEO</i>)	_____ 117
3.1.3. Satélites de órbita baja (<i>LEO</i>)	_____ 118
3.1.3.1. Ventajas	_____ 118

3.2.4.1. Voz	146
3.2.4.2. Radiolocalización	147
3.2.4.3. Fax	147
3.2.5. Comunicación pasó a paso del sistema	148
3.2.6. Mercado del sistema <i>Iridium</i>	149

4. RED SISTEMA INTEGRADO DE MENSAJERÍA

4.1. Planificación y dimensionado de red	151
4.1.1. Relación tráfico por usuario	152
4.1.1.1. Hora activa o pico	153
4.1.1.2. Tiempo medio de conversación	154
4.1.1.3. Número promedio de llamadas por usuario en la hora pico	156
4.1.2. Intensidad de tráfico telefónico	157
4.1.3. Grado de servicio	158
4.2. Diseño de red sistema integrado de mensajería	161
4.2.1. Acceso de llamadas	162
4.2.1.1. Fuentes de entrada	162
4.2.1.2. Conmutación virtual	164
4.2.1.2.1. Controlador de conmutación virtual (VCS)	169
4.2.1.2.2. Puerta de enlace o <i>gateway</i> de medios (MG)	171
4.2.1.2.3. Punto de control de servicio (SCP)	171
4.2.1.2.4. Nodo de servicios	172
4.2.1.2.5. H.323	172
4.2.1.3. Centro de llamadas de telefonía de paquetes (<i>PTCC Packet Telephony Call Center</i>)	177
4.2.2. Unidad de control	180
4.2.2.1. Arquitectura red <i>Lan</i>	181
4.2.2.2. Servidores de aplicaciones especiales	183
4.2.2.2.1. Servidor de base de datos y reconocimiento de voz	183

4.2.2.2.1.1. Base de datos _____	184
4.2.2.2.1.2. Reconocimiento de voz _____	191
4.2.2.2.1.3. Característica del <i>hardware</i> del servidor ____	193
4.2.2.2.2. Servidor de correo electrónico _____	194
4.2.2.2.3. Servidor de <i>Internet</i> _____	194
4.2.2.3. Terminal de radiolocalización _____	195
4.2.2.4. Protocolos del Sistema Integrado de Mensajería _____	196
4.2.2.4.1. Protocolo de la <i>PSTN</i> al sistema de radiolocalización __	197
4.2.2.4.2. Protocolos de <i>Internet</i> y correo electrónico _____	197
4.2.2.4.3. Protocolos sobre la infraestructura del sistema de radiolocalización _____	198
4.2.2.4.4. Protocolos entre la infraestructura del sistema de radiolocalización y los radiolocalizadores _____	199
4.2.2.5. Sistema de control de radiolocalización _____	204
4.2.2.6. Puertas de enlace (<i>gateway</i> 's) _____	205
4.2.2.6.1. Puertas de enlace (<i>gateway</i>) de <i>Internet</i> _____	205
4.2.2.6.2. Puerta de enlace (<i>gateway</i>) satelital _____	206
4.2.2.7. Medios de transmisión _____	206
4.2.3. Transmisión _____	207
4.2.3.1. Núcleos _____	209
4.2.3.2. Antenas _____	210
4.2.4. Recepción _____	211
4.2.4.1. Radiolocalizadores _____	212
4.2.4.1.1. Radiolocalizadores numéricos _____	212
4.2.4.1.2. Radiolocalizadores alfanuméricos _____	213
4.2.4.1.3. Radiolocalizadores <i>Iridium</i> _____	215
4.3. Evaluación “Sistema Integrado de Mensajería” _____	218
4.3.1. Estudio de preinversión _____	218
4.3.1.1. Idea _____	220

4.3.1.2. Perfil y prefactibilidad _____	220
4.3.1.3. Factibilidad _____	221
4.3.1.3.1. Estudio de mercado _____	223
4.3.1.3.2. Estudio técnico _____	227
4.3.1.3.3. Estudio económico _____	230
4.3.1.3.3.1. Período de recuperación de la inversión ____	236
4.3.1.3.3.2. Valor actual neto o valor presente _____	237
4.3.1.3.3.3. Tasa interna de retorno (TIR) _____	241
CONCLUSIONES _____	243
RECOMENDACIONES _____	245
REFERENCIAS BIBLIOGRÁFICAS _____	247
BIBLIOGRAFÍA _____	249
A. FORMATOS DE 2 Y 5/6 TONOS _____	253
B. <i>HANDOVER</i> _____	257

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Sistema básico de radiolocalización <i>paging</i> _____	4
2.	Diagrama de bloques de un receptor o radiolocalizador _____	22
3.	Sistema de reconocimiento de voz por aproximación acústica fonética _____	42
4.	Reconocimiento de voz por aproximación por medio del reconocimiento de la muestra _____	43
5.	Capacidad de corrección de palabras con sintaxis _____	44
6.	Procesador <i>bottom-up</i> _____	45
7.	Procesador <i>top down</i> _____	46
8.	Señales sonoras _____	47
9.	Señales no sonoras _____	47
10.	Modelo del tracto vocal _____	49
11.	Formas de onda de la voz _____	50
12.	Forma general de un filtro de primer orden _____	53
13.	Ventana de <i>Hamming</i> _____	54
14.	Aproximación por medio del reconocimiento de la muestra _____	55
15.	Aproximación acústica fonética _____	55
16.	Modelo de análisis de banco de filtros _____	57
17.	Modelo de análisis de <i>LPC</i> _____	57
18.	Proceso de filtrado y muestreo de una señal _____	58
19.	Modelo completo de análisis de banco de filtros _____	59
20.	Proceso de filtrado de una señal _____	60
21.	Formas de onda y espectros típicos en el análisis de banco de filtros _____	63
22.	Serie de filtros pasabanda de Q canales _____	65

23. Variación del ancho de banda con respecto a la frecuencia _____	67
24. Modelo general filtro <i>IIR</i> _____	69
25. Modelo general filtro <i>FIR</i> _____	69
26. Ventana de <i>Hamming</i> función coseno _____	73
27. Función ventana rectangular y de <i>Hamming</i> _____	74
28. Forma general de un banco de filtros <i>FIR</i> no uniforme _____	76
29. Anchos de banda de banco de filtros no uniformes de 3 y 7 bandas _____	77
30. Alineamiento temporal dinámico _____	86
31. Restricción local _____	87
32. Restricción global _____	88
33. Mezcla de <i>gaussianas</i> _____	94
34. Organización jerárquica de los módulos acústicos y lingüísticos _____	105
35. Estructura para un flujo inverso _____	106
36. Esquema de funcionamiento del proceso de clasificación basado en <i>HMM</i> _____	108
37. Modelo de codificación _____	111
38. Modelo de producción de voz _____	114
39. <i>Vocoder LPC</i> _____	116
40. Órbitas de la tierra _____	120
41. Sistema de satélites _____	122
42. Haces de cada satélite <i>Iridium</i> _____	123
43. Mapa de cobertura del sistema <i>Iridium</i> _____	124
44. Estructura de los satélites <i>Iridium</i> _____	125
45. Estructura de una puerta de enlace ó <i>gateway</i> _____	128
46. Asignación de frecuencias _____	131
47. <i>BPSK</i> _____	135
48. <i>QPSK</i> _____	136
49. Generación de estados de fase en <i>QPSK</i> _____	138
50. Transmisión <i>PCM</i> _____	140
51. Recepción <i>PCM</i> _____	141

52. <i>FDMA</i> _____	142
53. <i>TDMA</i> _____	143
54. Servicios del sistema <i>Iridium</i> _____	145
55. Comportamiento <i>poissoniano</i> del tiempo de duración de una llamada _____	155
56. Muestra de llamadas en la hora pico durante una semana _____	156
57. Diseño en bloques red Sistema Integrado de Mensajería _____	162
58. Conmutación de circuitos frente a conmutación de paquetes _____	165
59. Arquitectura de telefonía de paquetes abierta _____	167
60. Modelo de referencia <i>OSI</i> sobre <i>IP</i> _____	168
61. Arquitectura de redes de voz por paquetes _____	170
62. Modelo de interconexión de la red telefónica a la red <i>IP</i> _____	176
63. Centro de llamadas de telefonía de paquetes _____	179
64. Infraestructura común para los agentes de llamadas _____	179
65. Interrelación de elementos principales de la red _____	184
66. Módulo eventos del sistema _____	184
67. Módulo mantenimiento de radiolocalizadores _____	185
68. Módulo de mensajes _____	186
69. Módulo de operadores _____	187
70. Módulo sistema _____	187
71. Módulo cuenta abonados _____	188
72. Características en el proceso de reconocimiento de voz _____	192
73. Modelo de comunicación para el reconocimiento de voz _____	193
74. Patrón de radiación _____	211
75. Diagrama general “Sistema Integrado de Mensajería” _____	217
76. Esquema de administración de recursos _____	219
77. Implementación física _____	222
78. Organización recursos humanos _____	229

TABLAS

I.	Clasificación de las consonantes de acuerdo a su articulación _____	101
II.	Clasificación de los fonemas vocálicos _____	102
III.	Comparación <i>ISL</i> entre <i>Leonet</i> e <i>Iridium</i> _____	133
IV.	Componentes H.323 estándar _____	173
V.	Pila de protocolos en <i>VoIP</i> _____	175
VI.	Características generales de la red <i>Lan Ethernet</i> _____	181
VII.	Características topología <i>10BaseT</i> _____	182
VIII.	Límites máximos de datos del protocolo <i>FlexTM</i> _____	201
IX.	Comparación de características de los tipos de protocolo _____	202
X.	Matriz de comparación de los protocolos más recientes _____	203
XI.	Frecuencias de operación sistema de radiolocalización en Guatemala _____	208
XII.	Especificaciones eléctricas _____	210
XIII.	Características del radiolocalizador alfanumérico _____	214
XIV.	Costos de equipos de la red _____	228
XV.	Balance inicial _____	231
XVI.	Estimación de ventas _____	231
XVII.	Depreciaciones y diferidos _____	232
XVIII.	Financiamiento _____	233
XIX.	Programa de inversión _____	234
XX.	Beneficios anuales _____	235
XXI.	Flujo neto de efectivo _____	236
XXII.	Cálculo del VAN _____	238
XXIII.	Beneficios anuales con operadores _____	239
XXIV.	Flujo neto de efectivo con operadores _____	240
XXV.	Cálculo del VAN con operadores _____	240
XXVI.	Cálculo TIR _____	242

LISTA DE SÍMBOLOS

\otimes	Convolución circular
Σ	Sumatoria
α	Factor de crecimiento logarítmico
$*$	Convolución discreta
$\bar{\varepsilon}$	Relación tráfico por usuario
Δ	Intervalo

GLOSARIO

Alófono	Variante en la pronunciación de un sonido con respecto a su pronunciación ideal.
Anillos <i>Van Allen</i>	Son dos toroides cuyo eje es el geomagnético y están poblados de partículas que dañan el satélite.
Articulación	Posición adoptada por los órganos de la cavidad bucal en el momento de producirse un sonido.
Asignación dinámica de canales	<i>DCA</i> , técnica en la cual las portadoras están semipermanentemente asignadas a las celdas del satélite según las previsiones del tráfico.
Atenuación	Reducción en potencia de una onda electromagnética, debido al aumento en la distancia que recorre desde la fuente.
Banda K	Frecuencia comprendida entre los 18 y los 30 GHz., utilizada por el sistema <i>Iridium</i> para la comunicación entre satélites, y entre satélites y puertas de enlace (<i>Gateway</i>).
Banda L	Frecuencia comprendida entre 1 y 2 GHz., utilizada por el sistema <i>Iridium</i> para la comunicación entre el satélite y un teléfono o radiolocalizador <i>Iridium</i> .

<i>Bit</i>	Digito binario utilizado en un sistema de numeración binario. Este puede ser 0 ó 1.
<i>Bps</i>	<i>Bits por second</i> , velocidad de transporte de <i>bits</i> por segundo.
<i>Cap code</i>	Formato de identificación de un radiolocalizador.
Coarticulación	Variación de las características de un sonido en función de la articulación de los sonidos que lo preceden y lo siguen.
Cobertura	Región dentro de la cual un receptor recibe la transmisión de las señales de radiolocalización.
Codificación de voz	Representación eficiente en formato digital de la señal de voz para su almacenamiento y/o transmisión.
Correo electrónico	Aplicación en la cual pueden ser transmitidos mensajes de texto, sobre varios tipos de redes, usando varios tipos de protocolos.
<i>dB</i>	Decibeles, décima parte de un <i>bel</i> o unidad para la medición de la intensidad de niveles de potencia de señal.
Desvanecimiento	Fenómeno en el cual los niveles de señal varían en cortas distancias debido a la propagación por multitrayectorias.

<i>DTW</i>	<i>Dyanamic Time Warping</i> , técnica conocida como alineamiento temporal dinámica, la cual compara la variación en el tiempo de dos palabras.
<i>E1</i>	Esquema de transmisión digital de área extendida, predominantemente utilizado en Europa, en el cual se puede transportar información a velocidades de 2.048 Mbps.
Enlace cruzado	Radioenlace entre dos satélites, también conocido como conexión entre satélites.
Enlace de alimentación	Radioenlace que conecta la estación terrestre de una puerta de enlace (<i>Gateway</i>) con el satélite para transmitir datos de mensajes.
Estaciones base	Equipos de transmisión que reciben la señal de la unidad de control, para luego transmitirla a una frecuencia de radio asignada.
<i>FDMA</i>	<i>Frequency División Múltiple Access</i> (Acceso múltiple por división de frecuencia), asigna a los usuarios distintas frecuencias en las que pueden comunicarse.
<i>FFT</i>	Transformada rápida de <i>Fourier</i>
<i>FLEXTM</i>	Formato de radiolocalización de alta velocidad, el cual puede mover datos a una tasa de transferencia de 6400 bps.

Fonema	Unidad estructural más pequeña de sonido que se usa para distinguir una pronunciación.
Fono	Realización acústica de un fonema.
Fonología	Estudia la estructura o función de los sonidos del lenguaje.
Formantes	Propiedad física correspondiente a las frecuencias de resonancia del tracto vocal. Se definen perceptualmente.
Ganancia de antena	Medida de la direccionalidad de la señal si esta es recibida o enviada a través del antena.
<i>GEO</i>	<i>Geostationary Earth Orbit</i> (Órbita terrestre geostacionaria), mantiene una posición constante relativa a la superficie de la tierra.
Gramática	Conjunto de reglas que limita el número de combinaciones permitidas de las palabras del vocabulario.
<i>Handover</i>	Llamada en progreso de un abonado, que cruza la frontera de la celda adyacente, asignándole un nuevo canal a la llamada en la nueva celda, para evitar la terminación de la misma.
Haz de antena	Patrón de energía de radiofrecuencia emitido por un antena. Los satélites <i>Iridium</i> transmiten hacen radioeléctricos muy concentrados.

<i>HMM</i>	Modelos Ocultos de <i>Markov</i> , se distinguen tres clases generales: modelos continuos, modelos discretos y modelos semi-continuos.
<i>Internet</i>	Red global de interconexión de redes, donde cientos de miles de redes alrededor del mundo tienden a compartir recursos.
<i>IP</i>	Protocolo TCP/IP para envío de paquetes.
<i>Iridium</i>	Constelación de satélites de órbita baja.
<i>LAN</i>	Equipos conectados en una red geográficamente próxima, por ejemplo en el mismo edificio.
<i>LEO</i>	<i>Low Earth Órbita</i> (Órbita terrestre baja), altitud orbital inferior a 1,600 Km. sobre la superficie terrestre.
<i>LPC</i>	Análisis de predicción lineal, técnica que parametriza la señal en un número pequeño de patrones con los cuales es posible reconstruirla adecuadamente.
<i>Mel-cepstrum</i>	Parámetros que representan la envolvente espectral de la voz, utilizando un espaciamiento o resolución en frecuencia semejante a la que utiliza el oído.
<i>MEO</i>	<i>Médium Earth Orbit</i> (Órbita terrestre media), franja comprendida entre los 3,200 y los 16,000 Km. sobre la superficie terrestre.

Morfología	Estudia las reglas de formación de las palabras a partir de los morfemas elementales.
Parametización de la voz	Extracción de rasgos característicos de la voz.
POCSAG	<i>Post Office Code Standardization Advisory Group</i> . Formato de radiolocalización digital sincrónico que permite la transmisión de la información en una estructura de conjunto.
Pragmático	Estudia el contexto del significado.
Preénfasis	Proceso en el cual se suaviza el espectro de una señal, para reducir inestabilidades de cálculo de precisión finita.
Preprocesamiento	Proceso por el cual se extrae la información acústica de la señal de voz emitida.
Propagación	Radiación de ondas electromagnéticas.
Prosodia	Relación entre los diferentes sonidos del mensaje hablado que refleja tanto elementos lingüísticos (modalidades, énfasis, etc), como elementos no lingüísticos (características del hablante, estado de ánimo, etc.).
Protocolo	Juego de reglas diseñado para facilitar de manera eficiente y creíble la transferencia de información entre dos dispositivos.

Puerta de enlace <i>(Gateway) Iridium</i>	Punto de interconexión terrestre que conecta la red de satélites <i>Iridium</i> con la red telefónica pública conmutada.
Puerta de enlace <i>(Gateway)</i>	Se refiere a un dispositivo de propósito especial que realiza la conversión protocolar de información entre uno y otro nodo.
<i>QPSK</i>	<i>Quadrature Phase Shift Keying</i> , Modulación por cambio de fase en cuadratura.
Radio Frecuencia	Parte del espectro electromagnético entre las regiones de audio e infrarrojas (10 KHz. a 10,000,000 MHz.).
Radiolocalizador	Receptor de FM sintonizados a la frecuencia de RF de la estación base. Pequeño y ligero que funciona con pilas, recibe los mensajes alfanuméricos / numéricos que pueden leerse en una pantalla, alertando al abonado en formas acústicas, luminosas, vibratorias.
Radiolocalización	Servicio de comunicación de una vía que ofrece mensajería, en el cual los abonados reciben un mensaje alfanumérico / numérico en un receptor (radiolocalizador) designado.
Reconocimiento de voz	Conversión de una señal acústica de voz a texto.

Reparto fijo de**canales**

FCA, técnica en la cual un conjunto de canales está permanentemente asignado a cada celda.

Router

Dispositivo de capa de red que usa uno o más medios para determinar la trayectoria óptima a lo largo de la red.

Semántica

Estudia como se encadenan las palabras para dar significado a una frase.

SIM

Sistema Integrado de Mensajería.

Sintáctico

Estudia la combinación de las palabras para construir frases de forma correcta.

TAP

Telocator Alphanumeric Protocol. Protocolo de comunicación para enviar mensajes del usuario a la terminal de radiolocalización.

TDMA

Time División Múltiple Access (Acceso múltiple por división de tiempo), permite a los usuarios utilizar un solo canal de comunicaciones en tiempo compartido.

TNPP

Telocator Network Paging Protocol. Protocolo que hace posible la interconexión entre terminales de radiolocalización de diferentes fabricantes.

<i>Transponder</i>	Dispositivo que recibe las portadoras del enlace ascendente, las amplifica, cambia sus frecuencias y las retransmite a la Tierra en nuevas frecuencias.
Transmisión	
Simultanea	Envío de señal de radiolocalización desde múltiples transmisores al mismo tiempo, consiguiendo una gran área de cobertura.
Unidad de control	Interfaz inteligente, recibe, codifica los mensajes y avisos de acuerdo con un protocolo determinado, los sitúa en cola y los envía a los transmisores para su difusión.
Ventana <i>Hamming</i>	Ventana que no tiene tanta precisión frecuencial, pero su objetivo es suavizar la señal en los bordes.
Vocabulario	Número de palabras diferentes que debe reconocer el sistema de reconocimiento de voz.
<i>VoIP</i>	Voz transmitida sobre el protocolo <i>internet</i> , permite a los usuarios establecer llamadas de voz, fax, datos, etc., sobre conexiones <i>IP</i> .

RESUMEN

Este trabajo, consiste en el diseño de un Sistema Integrado de Mensajería (SIM) adecuado para tener cobertura nacional y global. El tema, versa sobre la adaptación a radiolocalización de las siguientes tecnologías: Reconocimiento de voz, sistema *Iridium* y voz sobre *IP*, que al integrarse proporcionan servicios de mensajería virtual. Se aportan suficientes conocimientos conjuntos en dichas tecnologías como para poder afrontar adecuadamente el diseño del SIM.

El resultado deseado fue el de un diseño híbrido del SIM, iniciándole con el debido dimensionamiento de la red, continuando con el desarrollo de cada parte que contiene el sistema, en cada una de las partes, se describe su función y sus requerimientos técnicos para su debido funcionamiento. Por ejemplo, en la parte de la base de datos se desarrollo la operabilidad de la misma y los campos que debe llevar como mínimo. Para finalizar la parte técnica, se desarrolla un diagrama general en bloques del sistema que provee el punto de partida para evaluar los servicios prestados y observar un sistema automático vanguardista en el presente y en un futuro inmediato.

Y finalmente se presenta la evaluación económica, la cual se enfoca solo en el estudio de preinversión (por ser este un sistema que no se implementará), en este estudio se profundiza en la factibilidad técnica económica, distinguiendo tres etapas del mismo, plan de mercado, estudio técnico (diseño del sistema) y el análisis económico por métodos que consideran el efecto que tiene el tiempo en el dinero y que determinan la conclusión más significativa, que el sistema en estudio es económicamente factible.

INTRODUCCIÓN

Se presencia en estos días, una progresiva proliferación por un aprovechamiento y optimación de los recursos de telecomunicaciones, por ello, resulta cada vez más atractivo desde un punto de vista de ingeniería y también comercial, pensar en una combinación de las tecnologías para que en un sistema se conjunten varios servicios que se le pueden prestar a un abonado, que debido a su movilidad, este siempre comunicado. Es por esta razón, que el Sistema Integrado de Mensajería propone una unión de servicios de comunicación, siendo la piedra angular del mismo, un sistema de radiolocalización conocido ya por todos como una forma de comunicación ágil, al cual se le incorporan sistemas innovadores como reconocimiento de voz (interfaz hombre máquina), voz sobre *IP* y el sistema *Iridium*, que juntos forman la solución híbrida para prestar servicios de radiolocalización automática con cobertura local y global, servicios de voz sobre *IP*, datos y otros que se pueden agregar beneficiándose de la versatilidad del sistema.

Ya que es fundamental el conocimiento de las partes del sistema, se inicia con la radiolocalización, la cual, se desglosa en aspectos relevantes que se requieren para su funcionamiento; primero, se describe lo básico, luego los sistemas de radiolocalización manuales y automáticos, cobertura, protocolos, y otros temas afines.

Reivindicando algo tan simple como, que la voz, es y seguirá siendo el interfaz más natural de comunicación para las personas, se incorpora a la red de radiolocalización un equipo de reconocimiento de voz, el cual hace posible la comunicación del hombre con las máquinas utilizando la voz. Esta comunicación no es más que el proceso de conversión de un mensaje hablado en texto que permita al usuario enviar su mensaje, con satisfacción y sin recurrir a operadoras. El capítulo dos,

contempla la teoría necesaria en la que se basan los sistemas de reconocimiento de voz, describiendo las disciplinas aplicadas en los mismos, tipos de reconocimiento de voz, y las formas utilizadas en el análisis del reconocimiento, siendo una de ellas más compleja que la otra, ya que trata de reconocer la voz desde un análisis espectral (utilizando técnicas completas como *LPC*, *MFCC*), pasando por un reconocimiento de fonemas (utilizando *DTW*, *HMM*) para determinar las características acústicas, y luego determinar características gramaticales. También se describen temas afines al análisis de reconocimiento como cuantización, alineamiento temporal dinámico (*DTW*), modelos de *Markov* y codificación.

La gran demanda de un sistema móvil global, la diversidad de normas celulares terrestres, las grandes áreas continentales y, sobre todo, marítimas con cobertura limitada o sin ningún tipo de infraestructura de telecomunicaciones y los servicios aeronáuticos sin aplicaciones globales son los principales argumentos para que naciera el sistema *Iridium* y diera respuesta a estas necesidades. Consiste en una constelación de satélites en órbita baja (*LEO*) que ofrece servicios de telefonía, datos y radiolocalización que además hace interfaz con las redes terrestres existentes. La incorporación de *Iridium* al sistema integrado de mensajería se debe a que ofrece ventajas significativas respecto a las redes terrestres, proporcionando una cobertura global, un rápido despliegue de la red, pretendiendo con ello que los abonados reciban mensajes de los satélites desde cualquier punto terrestre, marítimo o incluso en el aire (principalmente partes de Guatemala sin cobertura), teniendo con ello un sistema que integra “un mundo, un radiolocalizador y una factura”. El capítulo titulado sistema *Iridium* habla y desglosa los temas principales: Constelación de satélites de órbita baja (*LEO*) y funcionamiento del sistema.

El capítulo red Sistema Integrado de Mensajería, no es más que la unificación de los sistemas hablados en los capítulos del uno al tres, más la incorporación de voz sobre *IP*. Empieza con un tema importantísimo en el diseño de redes, que es la planificación y

dimensionado de red; continuando con el diseño del sistema, el cual se desglosa en bloques para ir describiendo cada bloque y cada componente que lo conforma (por ejemplo: servidores, terminales de radiolocalización, transmisión, recepción y otros) hasta llegar a un diseño general del Sistema Integrado de Mensajería y sus servicios. Se da también, una breve introducción a voz sobre *IP*, su arquitectura, funcionamiento del intercomunicador de redes de telefonía tradicional a voz sobre *IP*, características del estándar H.323 y protocolos utilizados en el proceso de una llamada. Y para finalizar con una evaluación económica del sistema, dicha evaluación solo comprende el estudio de preinversión que determina la bondad del proyecto analizada por métodos cuantificados.

Finalmente, se comentan la conclusiones, recomendaciones y se dan dos apéndices con temas interesantes, el primero, formatos de 2 y 5/6 tonos y el segundo, *handover* de la constelación *Iridium*

1. SISTEMA BÁSICO DE RADIOLOCALIZACIÓN

PAGING

El sistema de radiolocalización es probablemente el más simple de los servicios móviles vía radio que existen, es un sistema de comunicación unidireccional, personal y selectivo de alerta, sin mensaje o con un mensaje definido, compuesto por caracteres numéricos y alfanuméricos.

En este capítulo, se describen los fundamentos teóricos del sistema de radiolocalización, incluyendo su clasificación en sistemas manual y automático, características de factores determinantes en la cobertura del sistema, protocolos de señalización, estructura y características de los receptores y para finalizar, características de terminales de radiolocalización, controladores y estaciones base.

1.1. Sistema básico

Los componentes clave, incluyen una fuente de entrada, una red telefónica alambrada ya existente, un equipo de control de codificación y transmisión, el receptor o el radiolocalizador mismo.

Empezando con la fuente de entrada, un mensaje puede originarse en un teléfono, computadora con módem, un dispositivo de entrada de mensajes de oficina o a través de un operador que toma el mensaje por teléfono y lo envía en su nombre. Ya originado, el mensaje se envía a través de la Red Pública Telefónica (*PSTN*) hacia una terminal de radiolocalización para su codificación y transmisión a través del sistema. Típicamente, el codificador acepta el mensaje entrante, verifica la validez del número del

radiolocalizador y busca en el directorio o base de datos la dirección del abonado, conocida como *cap code*, convirtiendo la dirección y el mensaje al protocolo apropiado de señalización de radiolocalización. El *cap code* es una dirección única la cual, lo habilita para encontrar el radiolocalizador que se quiere alcanzar en cualquier lugar dentro del sistema. La señal codificada de radiolocalización se envía a los transmisores (estaciones base), mediante los sistemas de control de transmisión de radiolocalización, y es radiado al área de cobertura en la frecuencia especificada.

1.1.1. Codificador

Los dispositivos de codificación convierten números de radiolocalizador a códigos que pueden transmitirse. Hay dos formas en las cuales los codificadores aceptan números de radiolocalizador: manual o automáticamente.

En la codificación manual, un operador del sistema ingresa los números y mensajes mediante un teclado conectado al codificador. Mientras que en un codificador automático, alguien marca a una terminal automática y usa el teclado telefónico para ingresar los números del radiolocalizador.

El codificador genera el código del radiolocalizador correspondiente al número marcado y lo envía a la estación base para su radio transmisión.

1.1.2. Estación base

Los equipos de transmisión reciben la señal de la unidad de control, extraen la orden de disparo y transmiten los códigos en una frecuencia de radio asignada. La mayoría de las estaciones usadas están específicamente diseñadas para radiolocalización, y algunas diseñadas para voz bidireccional también se pueden utilizar.

1.1.3. Receptores

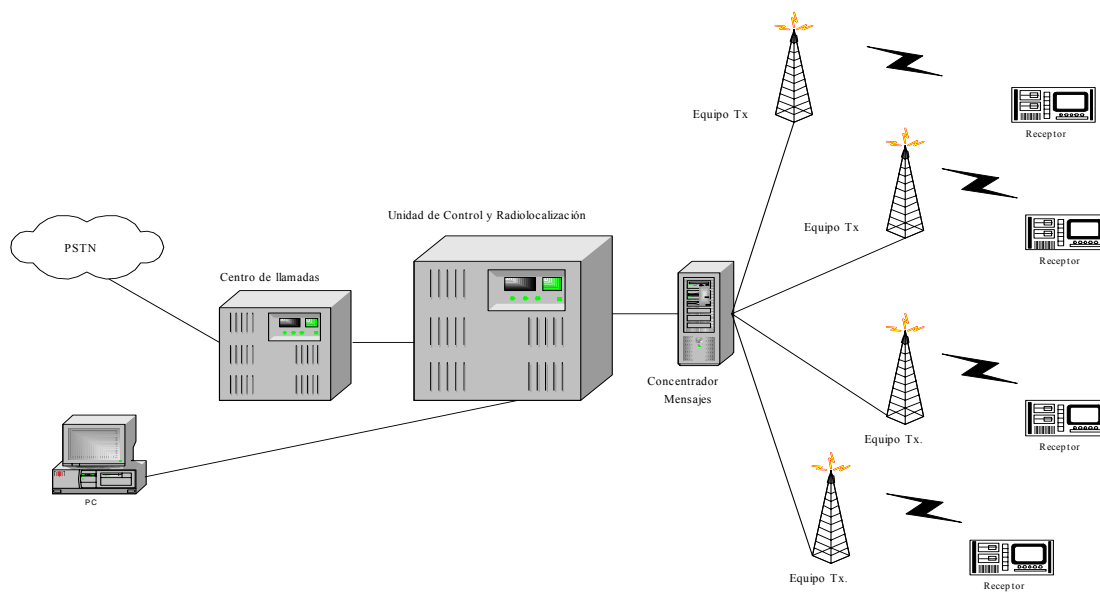
Estos son esencialmente receptores de FM sintonizados a la frecuencia de RF de la estación base del sistema. Una unidad decodificadora incorporada en cada radiolocalizador podrá reconocer su código y rechazará otros códigos. A los radiolocalizadores pueden asignárseles intencionalmente los mismos códigos para la radiolocalización de grupo. Hay también, radiolocalizadores que pueden asignárseles códigos de mensaje múltiple, típicamente a un máximo de cuatro, individualmente, de grupo y otras funciones de radiolocalización.

La función básica del radiolocalizador es alertar. Al recibir su código, el receptor pone una alarma que puede ser audible (tono), visual (indicador de pestañeo) o silenciosa (vibrador). Otras funciones de mensajería también pueden realizarse y estas incluyen voz y/o despliegue (numérico / alfanumérico) del mensaje en la pantalla. Algunos radiolocalizadores pueden realizar las tres funciones de mensajería.

1.2. Sistemas de radiolocalización

Los sistemas de radiolocalización pueden clasificarse en dos tipos: manuales y automáticos. A continuación se muestra en la figura 1, un diagrama en el cual se basan estos sistemas.

Figura 1. Sistema básico de radiolocalización *paging*



1.2.1. Sistema manual de radiolocalización

El remitente del mensaje usa un teléfono normal para llamar a un operador de radiolocalización y le dice, a que número o para quien es el mensaje. El operador ingresa el número en el codificador, el cual es como el de los teléfonos de 12 teclas. El codificador lo convierte en un código del radiolocalizador y el formato de salida codificado se envía por RF. Por ejemplo, un formato de radiolocalización de 2 tonos, la información codificada es de dos tonos secuenciales teniendo ciertas frecuencias y duraciones.

El operador monitoréa la frecuencia de radiolocalización por el tráfico. Si el canal está libre, el operador oprime el botón de transmisión para enviar el código del radiolocalizador y entonces pasa al siguiente mensaje.

La unidad de mando del transmisor energiza el transmisor de la estación base y envía el código del radiolocalizador para que se transmita. Todos los abonados sintonizados a esa frecuencia y dentro de la cobertura de la base reciben la señal de RF enviada por el transmisor. En alguno de los abonados, el receptor decodifica correctamente su código y activa la señal de alerta, lo que hace que éste tome alguna acción.

1.2.2. Sistemas automáticos de radiolocalización

1.2.2.1. Terminales de radiolocalización

Una terminal de radiolocalización, es un dispositivo codificador automático que contesta una línea de teléfono y permite a los usuarios marcar un número del radiolocalizador y un mensaje numérico o audible mediante un teléfono común. Algunas terminales, también aceptan mensajes textuales mediante conexiones con dispositivos como las computadoras personales. También las terminales de radiolocalización, incluyen normalmente una unidad de control para la estación base de manera que la misma pueda automáticamente transmitirlos.

Las características y funciones asociadas de una terminal de radiolocalización automática se describen a continuación:

- Conexión telefónica: respuesta automática y corte.
- Sugerencias (voz o tonos): le indica al usuario cuando proporcionar el número y el mensaje.
- Base de Datos: tiene los números, códigos, tipos de mensajes esperados..., etc., en un archivo de computadora.
- Codificador: genera códigos análogos o binarios transmisibles.
- Unidad de Control del transmisor: monitorea canales, transmite claves y envía señales al transmisor.

1.2.2.2. Flujos de mensajes

El flujo de un mensaje en un sistema automático se presenta a continuación: primero, para tono, voz y numérico y luego el alfanumérico.

1.2.2.2.1. Tono, voz y numérico

- El usuario marca el número.
- El usuario marca el mensaje numérico con su teclado o dice el mensaje; entonces indica el fin del mensaje.
- Verificación del número marcado; la aceptación de la llamada es enrutada al conmutador.
- Para llamadas por tonos, un sonido especial se envía para indicar al usuario que el mensaje ha sido aceptado.
- Para mensajes numéricos o de voz, un patrón específico de tonos o de voz inducen al usuario a enviar su mensaje o número telefónico.
- Un mensaje específico de aceptación por tono o voz se envía al usuario;
- Una señal de fin de llamada se envía a la telefónica.
- Las llamadas se enrutan a una terminal de radiolocalización.
- Se establece un canal entre el usuario y la terminal de radiolocalización.
- Desconecta el canal de voz.

1.2.2.2.2. Alfanumérico

- El usuario ingresa el número telefónico, el número del radiolocalizador y el mensaje alfanumérico en un dispositivo A/N.
- El dispositivo marca el número telefónico.
- El dispositivo señala.
- El dispositivo envía el número del radiolocalizador y mensaje de A/N.
- El mensaje se divide en segmentos, si es demasiado largo.
- El dispositivo envía una secuencia de desconexión.
- El dispositivo desconecta la línea telefónica.
- La terminal de radiolocalización responde.

- La terminal alerta al dispositivo para señalización (por clave) más adelante.
- Responde con una señal de aceptación a la señalización, envía secuencia de adelante cuando está lista, cada segmento es probado para asegurar su correcta recepción. De acuerdo al resultado de la prueba se da la apropiada respuesta, por ejemplo, un reenvío debería ser requerido si hay error de transmisión. Cuando el último segmento se recibe, se envía una señal de mensaje aceptado.
- La terminal de radiolocalización responde y desconecta la línea telefónica.
- Encamina la llamada a la terminal de radiolocalización.
- Coloca un canal de voz directo entre el usuario y la terminal de radiolocalización.
- Desconecta el canal.

1.3. Cobertura de un sistema de radiolocalización

La cobertura de un sistema de radiolocalización se refiere a la región dentro de la cual un receptor de radiolocalización puede recibir confiablemente la transmisión de las señales de radiolocalización. Los diferentes factores que determinan la cobertura de un sistema de radiolocalización se discutirán en las siguientes páginas.

Para los sistemas privados locales, por ejemplo aquellos que se usan en hospitales, hoteles y fábricas, la cobertura esta limitada y generalmente usan un solo sistema con transmisor compacto de bajo costo, poca potencia (hasta 10W).

La cobertura para sistemas de radiolocalización públicos puede variar significativamente en extensión. Para una ciudad pequeña o país pequeño, un transmisor debería ser suficiente, pero cuando el área a cubrirse es grande podría ser necesario tener varios sitios de transmisión. Cuando se usa más de un transmisor con el propósito de

mejorar la cobertura, la técnica de transmisión simultánea desde todos los transmisores necesita utilizarse.

También, una red de radiolocalizadores se examinará, este es un arreglo mediante el cual se encadena o conecta con otros sistemas de manera que el suscriptor fuera del rango de su sistema propietario aún puede ser localizado. Esto es posible transfiriendo los requerimientos desde su sistema a través de la red al sistema que cubre su nueva localización. Las redes habilitan localización a lo ancho de la nación de grandes países como los Estados Unidos y cuando se encadenan con satélites en la red se pueden realizar localizaciones internacionales o globales.

1.3.1. Factores que afectan la cobertura

La cobertura de un transmisor de radiolocalización depende de los siguientes factores:

- a) Potencia del transmisor
- b) Sensibilidad del receptor
- c) Ganancia de antena
- d) Altura de antena
- e) Frecuencia de transmisión
- f) Pérdida de trayectoria
- g) Desvanecimiento.

1.3.1.1. Potencia del transmisor

Aumentar la potencia de salida del transmisor de RF es obviamente una manera de aumentar el área de cobertura pero frecuentemente un método más efectivo para mejorarla es aumentar la altura de la antena de la estación. En la mayoría de las situaciones, un análisis muestra que duplicar la potencia del transmisor da como resultado que la potencia en el punto de recepción se incrementa solamente 1.4 veces, mientras que al duplicar el alto de antena, duplica la señal en el campo.

1.3.1.2. Sensibilidad del receptor

La sensibilidad del receptor esta definida como la potencia requerida en el campo para que un receptor responda a un porcentaje específico de llamadas transmitidas. Un radiolocalizador con una gran sensibilidad de recepción es deseable y esto significa que una pequeña señal es suficiente para responder correctamente al borde del área de cobertura. Sin embargo, usar radiolocalizadores de gran sensibilidad, resulta en una amplia cobertura comparados con otros que se usan de menos sensibilidad.

1.3.1.3. Ganancia de antena

La ganancia de antena es una medida de la direccionabilidad de la señal si esta es recibida o enviada a través de antena. Por ejemplo, una antena con ganancia unitaria nunca magnificará o atenuará la señal. Este es un medio mediante el cual la cobertura de un sistema de radiolocalizadores puede ser mejorada. El ingeniero de diseño tiene que sopesar el uso de antena con ganancia, versus usar una antena de ganancia unitaria con un transmisor de gran potencia. La direccionabilidad de las antenas viene a ser el precio de un ancho de banda estrecha.

Una antena omnidireccional irradia una señal con igual ganancia en todas las direcciones mientras que una antena direccional provee una ganancia relativamente grande en una dirección en particular. La elección del tipo de antena apropiado y la alineación y el ajuste dan un patrón consistente con la cobertura deseada. Algunas veces el patrón de cobertura también se ajusta con antenas direccionales, para minimizar la interferencia con otros sistemas de comunicación que usan una frecuencia similar o cercana a la del sistema de radiolocalización.

1.3.1.4. Altura de antena

Aumentar la altura de antena es una alternativa atractiva para aumentar la potencia de transmisión buscando mejor cobertura. Aumentar la altura de antena usualmente significa una larga línea de señal con sus pérdidas asociadas, y un sistema complejo de balance es necesario. Típicamente, los transmisores se colocan en lo alto de un edificio o colina para tener altura adicional, y consecuentemente cobertura. Si no hay edificios disponibles en el lugar deseado, se construye una torre con cabina en la base y antena montada lo más alto posible.

1.3.1.5. Frecuencia de transmisión

Frecuencias de transmisión mas ampliamente usadas para sistemas de localización son la banda baja de VHF (40 MHz) y la banda alta (150 MHz), UHF (450 MHz) y 900 MHz. En algunos países, 280 MHz también están disponibles. Generalmente, la penetración de la señal transmitida a través de edificios, es mejor a altas frecuencias. Los árboles y follaje tienden a atenuar más estas altas frecuencias. Las pérdidas en la trayectoria debidas a la dispersión de la onda y a su propagación a través de la atmósfera, es la misma a todas las frecuencias entre antenas que tienen una apertura constante con la frecuencia.

1.3.1.6. Pérdidas en la trayectoria

Aquí se refiere a la atenuación de la señal y a su propagación desde la antena de transmisión hasta el radiolocalizador. Un componente de estas pérdidas proviene de la dispersión de la onda y su propagación a través del aire. Cuando la trayectoria es larga y la altura de antena es baja de modo que la transmisión es cercana al suelo, la presencia de tierra modifica la generación y propagación de las ondas de radio, de manera que la intensidad del receptor es normalmente menor de lo que podría esperarse para una propagación de espacio libre.

Otras contribuciones a estas pérdidas incluyen obstrucción por la tierra misma, conocido como pérdidas de difracción dado a que la señal se arrastra en la superficie de la tierra y por colinas árboles y edificios.

1.3.1.7. Desvanecimiento

El desvanecimiento es un fenómeno mediante el cual los niveles de señal varían en cortas distancias debido a la propagación por multitr trayectorias. Adicionalmente, las condiciones atmosféricas varían, la trayectoria de transmisión se altera y esto podría incrementar o decrementar la efectiva transparencia de la trayectoria.

El severo desvanecimiento por multitrayectoria usualmente aumenta conforme la frecuencia se incrementa, y esto puede predecirse usando métodos estadísticos. Un margen de tolerancia se agrega con frecuencia mediante el cálculo de cobertura para cuidarse del desvanecimiento.

Para comprobar si un lugar está dentro del área de cobertura de una estación de radiolocalización, se calcula la intensidad del campo de recepción en ese lugar. Esto se hace iniciando con la potencia de transmisión, agregando pérdidas y ganancias a la antena y restándole todas las pérdidas mencionadas más las pérdidas en la línea de transmisión, conectores y filtros. En un punto en el margen del área de cobertura, debe tener un valor mayor de intensidad de señal que la sensibilidad del receptor, para tener algún margen de desvanecimiento.

1.3.2. Transmisión Simultánea

Es un método creíble de éxito en un gran área de cobertura. Esto involucra el envío de la señal de localización desde múltiples transmisores precisamente al mismo tiempo. Esta técnica tiene la ventaja obvia de un área de cobertura que proviene de la combinación de la cobertura de transmisores individuales. Hay otras ventajas:

1.3.2.1. Ventajas operacionales de la transmisión simultánea

- a) Métodos de alternativos de éxito en un gran área de cobertura incluyen transmisores consecutivos, donde los transmisores son secuencialmente operados y se operan múltiples transmisores en múltiples frecuencias. Los transmisores consecutivos acomodan pocos mensajes por canal, a causa del gran tiempo tomado para comunicar a una unidad del radiolocalizador. Los transmisores de múltiples frecuencias son espectralmente ineficientes.

- b) Una configuración especial de sectores la cual podría diseñarse en un sistema de múltiple transmisión permitiría programar transmisiones por sector. Cada sector puede tener varios transmisores. El control de sectores permitiría direccionar solamente el grupo requerido de transmisores para un servicio en particular. Esto permite también variar las tarifas a cobrar a los abonados.
- c) Las pérdidas por obstrucción debidas a colinas, árboles, y edificios, resultan considerablemente reducidas por las configuraciones de transmisores múltiples. Por ejemplo, si un abonado esta oculto por una colina con respecto a un transmisor, hay una alta probabilidad de que un segundo transmisor tenga una trayectoria libre.

1.3.2.2. Requerimientos operacionales

Una mensaje es iniciado por un usuario usando un codificador y/o una terminal. El controlador podría integrarlo dentro de una terminal de radiolocalización o bien podrían ser unidades separadas. Las señales de radiolocalización se envían a un sistema de distribución que pueden ser por línea física o radio enlace.

En transmisión simultánea, los siguientes parámetros deben mantenerse dentro de tolerancias firmes:

1. Las frecuencias relativas de las portadoras relativas de las transmisiones de radiolocalización.
2. Las fases de las frecuencias relativas de audio emanadas de los transmisores de radiolocalización.
3. Los niveles de audio relativos de los transmisores de radiolocalización.

Los transmisores incorporan osciladores sumamente estables para asegurarse que cualquier diferencia en la frecuencia de la portadora, se mantenga por debajo de los límites especificados para una apropiada operación en transmisión simultánea. A intervalos regulares se resintonizan los osciladores, por ejemplo las frecuencias de portadora se ajustan a límites aceptables.

Las diferencias en distancias desde el radiolocalizador hacia los diferentes transmisores provoca que las señales de audio lleguen en diferentes momentos, por ejemplo llegan fuera de fase una respecto de la otra. Los ecualizadores de audio, que son elementos de tiempo de retraso variable, se utilizan en las emisoras para ocuparse de este indeseable desfasamiento. Ajustes precisos en los niveles de audio se ejecutan en los transmisores para optimar el sistema de transmisión simultánea.

1.3.3. Red de radiolocalización

Este es un arreglo en el cual los abonados pueden ser localizados aun cuando vayan de una ciudad a otra, de un estado a otro o fuera del país. En este arreglo, los radiolocalizadores se conectan para formar una red de los mismos. Cada control de radiolocalizadores tiene su propia área de cobertura, pero cuando un abonado se desplaza fuera del área, el mensaje se pasa a través de la red que cubre el área donde se encuentra el abonado. Esto podría requerir que el abonado mantenga el sistema actualizado con aviso de sus movimientos y localizaciones.

Para comunicaciones entre redes de radiolocalizadores, se han desarrollado varios protocolos de comunicación. Un protocolo de comunicación, se define como un juego de reglas diseñado para facilitar de manera eficiente y creíble, la transferencia de información entre dos estaciones. Por ejemplo, un protocolo puede especificar que la información puede dividirse en paquetes de cierto tamaño y que el envío de la información de la dirección de destino debe preceder a los datos. Y si hay errores en un paquete recibido, la estación de destino puede pedir su reenvío desde el original.

Los protocolos propietarios son en su mayoría incompatibles con otros, de ahí que las redes de radiolocalizadores estén usualmente usando radiolocalizadores de un único proveedor.

A fin de permitir la interconexión de equipos provenientes de diferentes fabricantes en una misma red, deberán usarse puertas de enlace (*gateways*). Hay esencialmente dispositivos interconectables los cuales efectúan conversiones de protocolo es decir que actúan como traductores de idiomas, lo que permite que redes de equipo incompatibles hablen y trabajen juntos. La mejor solución es crear un protocolo estándar de la industria, tal como *X.25* y *TNPP*.

El *TNPP* son siglas de *Telocator Network Paging Protocol* es uno de los protocolos mas ampliamente aceptados y empleados para la creación de redes constituidas por equipos provenientes de diferentes fabricantes. También puede usarse en el caso de redes que usan equipo semejante. Es un protocolo de comunicaciones digitales punto a punto, que asegura confiablemente el envío de información de una terminal a otra. El enrutamiento de radiolocalización a través de la red no esta cubierto en las especificaciones del *TNPP*, es solamente un protocolo punto a punto. Sin embargo, en una gran red de radiolocalizadores, el enrutamiento debe efectuarse eficientemente y, en la mayoría de tales redes, los algoritmos de enrutamiento son siempre implementados para trasladar mensajes entre uno y más nodos.

El *TNPP* no está restringido a encadenar radiolocalizadores a través de línea física. Los requerimientos de radiolocalización pueden eventualmente ser radiados a satélite para terminales en lugares distantes. Esto puede extender la cobertura por sobre las fronteras internacionales en un sistema de red satelital, el protocolo asegura confiabilidad en el envío de datos mediante la transmisión de los mismos muchas veces desde la fuente, al mismo tiempo asegura que la información duplicada no se recibe en el destino. Esta técnica de múltiple retransmisión se requiere porque la comunicación a larga distancia involucra evitar el envío de señales desde el destino a la fuente, confirmando que los datos han sido correctamente recibidos. El *TNPP* a proveído facilidades tales como correo electrónico y monitoreo de estado remoto.

1.4. Protocolos de señalización en radiolocalización (Formatos de codificación)

En el sistema, la terminal de radiolocalización, después de aceptar un mensaje entrante y variarlo, codificará la dirección y el mensaje en el apropiado protocolo de señalización. El protocolo de señalización permite que radiolocalizadores individuales sean identificados exclusivamente y alertados para proveerlos con adicional mensaje de voz o despliegue si lo hay.

Se usan varios protocolos de radiolocalización para diferentes tipos de servicios de radiolocalizadores tales como tono solamente, tono y voz, etc. La mayoría de las redes están o pueden soportar muchos formatos diferentes en una sola frecuencia. Muchos formatos son específicamente facturados y son propietarios, pero hay protocolos de dominio público tales como *POCSAG*, esto permite a diferentes fabricantes producir radiolocalizadores compatibles.

1.4.1. Formato análogo o digital

Todos los formatos caen dentro de una de las dos técnicas básicas de codificación: formato por codificación por tonos y formato de binario codificado. El primer formato usa una técnica de transmisión análoga, mientras el segundo usa una técnica de transmisión digital para transferir información a los radiolocalizadores. En transmisión digital, la información a transmitirse se digitaliza primero. La transmisión digital tiene muchas ventajas sobre la transmisión análoga, algo muy importante es su mejor desempeño en presencia de ruido.

A continuación, se describen diferentes formatos análogos y digitales y se comparan.

1.4.1.1. Formato de codificación por tonos

En este formato, una secuencia de tonos de audio se usa para identificar cada radiolocalizador. Por ejemplo, podemos usar 3 entona f_1 , f_2 y f_3 , combinando dos a la vez para representar la identificación de seis radiolocalizadores así: f_1f_2 , f_1f_3 , f_2f_1 , f_2f_3 , f_3f_1 y f_3f_2 .

Detalles de los formatos de los radiolocalizadores de 2 tonos y 5/6 tonos puede encontrarse en el Apéndice A. Comparado con formatos de codificación binaria, los formatos de codificación por tonos son lentos y sólo pueden soportar unos pocos abonados. Además los sistemas en este formato producen un gran número de llamadas falsas al no existir ninguna redundancia. Sin embargo, en la actualidad muchos sistemas aún soportan este formato debido a la existencia de viejos radiolocalizadores.

1.4.1.2. Formato de codificación binaria

En los formatos de codificación binaria, la dirección y la información a enviarse se codifica mediante una terminal de radiolocalización en forma binaria y se transmiten en palabras de n *bits*. Las palabras en código binario se envían mediante la portadora del transmisor; cada palabra incluye un número fijo de *bits* de redundancia, generalmente de acuerdo con un código cíclico, lo cual permite una decodificación simple y un grado alto de eficiencia en la detección de errores. La portadora esta hecha para variarse dentro de dos frecuencias discretas de la misma manera que los códigos varían entre unos y ceros. Por ejemplo, con una portadora de 154.6 Mhz un uno puede representarse por 154.604 MHz y un cero por 154.596 MHz. Esta modulación digital se llama claveado por desplazamiento de frecuencia (*FSK*) y la velocidad típica de transmisión es 512 o 1200 bps. Cada transmisión comienza con un preámbulo de 576 *bits*, alternando 0 y 1, que permite la sincronización a *bits* entre transmisor y receptor.

Al preámbulo le sigue un conjunto de palabras de 32 *bits*, cada una de ellas con 21 *bits* de información y 11 *bits* de redundancia. Las palabras se transmiten en tramas continuas, cada trama comienza con una palabra de sincronización y 16 de información. El primer *bit* de cada palabra de información es una señal que indica si la palabra contiene información de dirección o de mensaje.

Los primeros 18 *bits* de una palabra de dirección son la identidad del receptor y cada receptor puede tener 4 direcciones. Cada conjunto se divide en 8 tramas de 2 palabras cada una. El ordinal de la trama ya forma parte de la dirección. Esto permite multiplicar por 8 la capacidad de direccionamiento del código que alcanza unos 2 millones de identidades y 8 millones de direcciones, reduciendo el número de llamadas falsas en razón de 8 a 1.

Los formatos de código binario proveen las siguientes mejoras y características sobre los formatos análogos previos:

- Mayor número de códigos y múltiple capacidad de direccionamiento.
- Expansión del tamaño de usuarios y flexibilidad.
- Capacidad de multifunción: timbre, tono, voz y mensajes de datos.
- Eficiente uso de las baterías.
- Señalización rápida.
- Excelente capacidad de corrección de error y probabilidad de recepción.

1.4.1.3. *POCSAG (Post Office Code Standardization Advisory Group)*

Este es el formato de radiolocalización digital del dominio público adoptado por muchos fabricantes alrededor del mundo, con el fin de crear un estándar, incluye transmisión de avisos y mensajes numéricos y alfanuméricos. Su empleo ha hecho disminuir los precios de los equipos y ha supuesto una mejora de las prestaciones y de las características, al trabajar distintos fabricantes en una misma tecnología, alcanzando sistemas más fiables. Tiene las siguientes características.

- a. Capacidad del Código: *POCSAG* puede acomodar dos millones de usuarios, cada uno capaz de soportar hasta cuatro direcciones. (para las funciones sólo tono, tono y audio, el despliegue numérico, etc.).
- b. Velocidad: *POCSAG* opera a velocidades de 2400 bits/seg. A esta tasa de transferencia, para enviar un simple timbre requiere solamente de 13 milisegundos. Esto es cerca de 100 veces más rápido que un protocolo de 2 tonos.

- c. Detección y corrección de error: la máxima ventaja de la velocidad es la capacidad de detección y corrección de error incorporado dentro de los formatos de codificación digital. Esto permite errores en la transmisión que pueden ser detectados y corregidos.

1.4.1.4. *FLEX*TM Código de radiolocalización de alta velocidad

Con la explosión de la tecnología inalámbrica y el dramático crecimiento en la industria de radiolocalizadores en muchos mercados asiáticos, existen redes que vienen a ser más grandes. En suma, el espectro de radiofrecuencia no está completamente disponible debido a las demandas de otras aplicaciones inalámbricas. En respuesta a este problema “**Motorola**” ha desarrollado el nuevo *FLEX*TM el cual mejora la eficiencia y penetración del sistema, permitiendo nuevas y poderosas características incorporadas en los radiolocalizadores así como en las redes. Las nuevas características beneficiarán a los operadores de los sistemas y a los abonados.

Uno de los éxitos del *FLEX*TM fue proveer un desempeño mejorado sobre *POCSAG*. En relación al *POCSAG* provee más de cuatro veces el número de abonados y canales de radiofrecuencia disponibles así como mucha protección contra desvanecimiento de señales. La combinación de una corrección mejorada con una superior protección contra el desvanecimiento mejora la probabilidad de recibir un mensaje intacto, especialmente largos mensajes alfanuméricos.

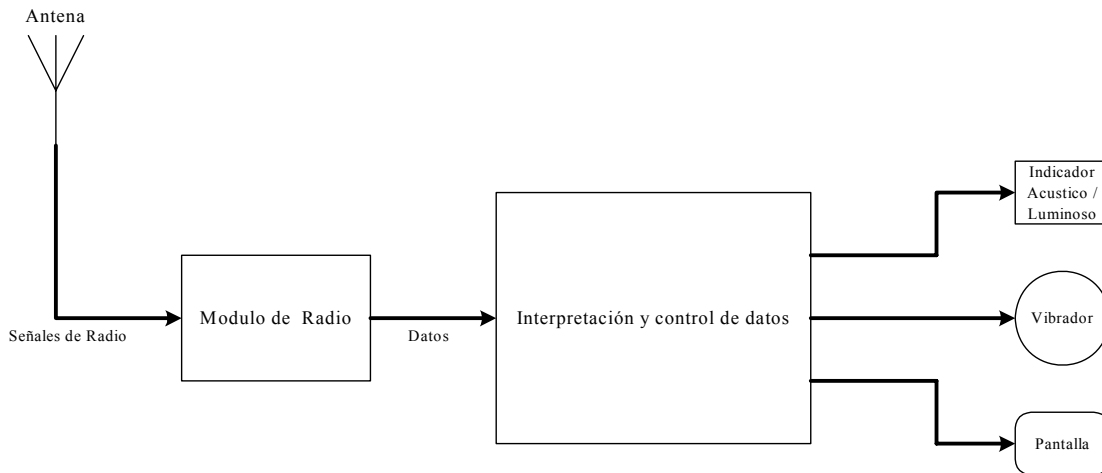
1.5. Soporte físico típico de la radiolocalización

El soporte físico de un sistema de radiolocalización se compone de receptores o radiolocalizadores, terminales de radiolocalización y estaciones base.

1.5.1. Receptores o radiolocalizadores

Los radiolocalizadores no han tenido un cambio dramático en término de sus funciones en la última década, sin embargo sus factores de forma son ahora mucho más atractivos que de los que se usaron antes. Robustos y mucho más pequeños en tamaño, ahora usan tecnología avanzada, en la figura 2 se muestra un esquema de bloques del funcionamiento de un radiolocalizador. Están controlados por un microprocesador y energizados por medio de una simple batería descartable AA o AAA.

Figura 2. Diagrama de bloques de un receptor o radiolocalizador



Fuente: Telecomunicaciones Móviles, Serie Mundo Electrónico, Pagina 34.

Basándose en la figura 2, un radiolocalizador moderno típico esta constituido de los siguientes módulos:

- a) Receptor: recibe y demodula las señales de localización.
- b) Decodificador: decodifica información binaria.
- c) Pantalla: presenta los mensajes y otras informaciones.
- d) Control: permite a los usuarios la colocación de parámetros y el acceso a la información.
- e) Batería: fuente de energía.

1.5.1.1. Características comunes de los radiolocalizadores

Rastreando desde los modestos radiolocalizadores de tono hasta los sofisticados receptores multifuncionales de datos y de información integrados con una computadora portátil, los radiolocalizadores modernos vienen con algunas de las muy comunes características explicadas en las siguientes páginas.

- a. Economizador de batería

Para conservar la energía de la batería, se han incorporado técnicas de ahorro dentro del formato de radiolocalización y el radiolocalizador. Dichas técnicas ponen al radiolocalizador en el modo de poco consumo por pequeños intervalos y de esa manera reduce la frecuencia del cambio de batería.

b. Llamada de grupo

El radiolocalizador esta provisto de dos códigos. Uno de estos para llamadas individuales y el otro para llamadas de grupo. Las llamadas de grupo permiten a un seleccionado grupo de individuos dentro del sistema que sean localizados simultáneamente por ejemplo: los bomberos y rescatistas.

c. Corrección y detección de error

La mayoría de los protocolos digitales de señalización, están equipados con códigos de detección y corrección de errores para asegurar una confiable recepción de los datos transmitidos. Esto involucra el envío de información adicional "con los datos" el cual habilita al receptor para verificar la presencia de errores en la transmisión y efectuar la corrección si es necesaria.

d. Alerta silente

Algunos localizadores tienen un vibrador para avisar la recepción de llamadas. Muy útil es en situaciones o lugares donde el tono de alerta no es aceptable, reuniones, servicios religiosos, etc.

e. Detección de mensajes duplicados

Si un mensaje nuevo es idéntico a un mensaje almacenado este no se memorizará. Aparecerá un indicador de duplicado para alertar que el mensaje ha sido recibido más de una vez. Esto previene el uso innecesario de memoria por mensajes duplicados.

f. Retención de memoria (Memoria no volátil)

Los mensajes y parámetros del radiolocalizador permanecen en memoria aún cuando este se apague o se le cambie la batería.

g. Contador de mensajes no leídos

Cada vez que un mensaje se recibe, se actualiza la pantalla para reflejar el número de mensajes no leídos en memoria.

h. Mensajes protegidos

Un cierto número de mensajes puede protegerse dentro de la memoria. Esto previene que tales sean sobrescritos por mensajes nuevos o se borren por error.

i. Mensaje congelado

El abonado puede detener un mensaje en la pantalla manteniendo presionado el botón *read*. Esto permite mantener el mensaje mientras se marca el número telefónico o se copia.

j. Presentador iluminado

Presionar un botón ilumina la pantalla. Esto permite leer en condiciones de poca luz.

k. Murmullo de alerta

Cuando un radiolocalizador esta puesto en modo vibrador, se emite un tono corto a la recepción de un mensaje. Esto es en lugar del timbrado habitual que puede molestar.

l. Indicador de memoria llena

Cuando el número de mensajes o caracteres almacenados alcanza su capacidad, el radiolocalizador presenta *Memorie Full* en la pantalla. Esto le da la oportunidad al abonado de ver sus mensajes y evitar así que algunos importantes sean sacados de la memoria por nuevos mensajes.

m. Programación

El *EEPROM* del radiolocalizador puede reprogramarse mediante una PC y un programa de protocolo. Este método puede usarse para configurar al radiolocalizador sus códigos y opciones remotamente sin que el usuario tenga ningún inconveniente. El radiolocalizador no tiene que ir necesariamente a un centro de servicio para ser configurado.

n. Indicador de fuente

Cuando selecciona, un indicador de fuente aparecerá al final de cada mensaje de datos. La fuente dirá que número telefónico se utilizó para enviar el mensaje y proveerá información adicional al abonado.

o. Bloqueo automático de timbre

El radiolocalizador bloqueará automáticamente el timbre después de un tiempo previamente asignado. Esto minimiza consumo de batería y el inconveniente de tener que apagarlo manualmente.

p. Servicios de información

Los códigos de los radiolocalizadores pueden programarse con servicios de información, por ejemplo resultados deportivos, bolsa de valores, estado del tiempo, etc. dicha información se recibe y se almacenan en sus propias localidades de memoria. Esto permite al usuario acceder a información específica en cualquier momento y siempre tendrá la información mas actualizada.

1.5.2. Terminales de radiolocalización

1.5.2.1. Características del MPS2000 interruptor modular de radiolocalización

- Diseño modular para fácil y efectiva expansión.
- Gran capacidad de procesamiento y más de un 1,000,000 de usuarios.
- Arquitectura flexible, procesador de entrada y salida que pueden encontrarse local o remotamente mientras se use un procesador central como único punto de facturación y mantenimiento.
- Manejo eficiente del tráfico: soporta un gran volumen de llamadas.
- Capacidad de red: provee un número de opciones de red incluyendo *TNNP*, *EUROMESSAGING*, *X.25*, y *Follow-me*.

- Anuncios de voz: tiene disponibilidad para mensajes de voz de alta calidad de norma y personalizados.
- Redundancia: Cualquiera de todos los módulos principales del sistema son configurables como redundantes, proveyendo una verdadera redundancia a múltiples niveles.

1.5.2.2. Características *Unipage*

La arquitectura del chasis modular permite un fácil y económico sistema de expansión, a continuación algunos puntos que lo caracterizan.

- Capacidad de mas de 1,000,000 de abonados y 1,900 troncales.
- En las troncales de entrada acepta señalización por pulso, *DTMF* y *MF*.
- Formatos de localización *POCSAG* (“502”, “1200”, “2400”), *GSC*, 2 tonos *MVS*, *RDS* y *FLEXTM*.
- Mensajes alfabéticos, numéricos, por tonos y voz.
- Soporta troncales análogas y digitales.
- Mensajes de voz: un sistema de mensajes versátil de voz y números puede proveer desde 6 hasta 600 Hrs. de almacenamiento de voz. También posee características adicionales como mensajes desechados, notificaciones, recordatorio de mensajes no contestados, saludos y anuncio de cambio de número.
- Transcripción alfabética: convierte todos los mensajes de voz a los mensajes alfabéticos a través de un operador para su posterior transmisión.
- Capacidad de red: ofrece toda la posibilidad de la interconexión por red.
- Mantenimiento flexible y fácil: de diagnóstico local o remoto y capacidad estadística, mantenimiento en línea y una opción de sistema de redundancia.

1.5.2.3. CONTROLADORES

1.5.2.3.1. *C-Net*

- Controlador avanzado de redes RF.
- Enlaces de la terminal a la estación de base.
- *C-net* puede conectar su estación base por medio de línea física, radio enlace, microonda o satélite.
- *C-net* reduce grandemente la necesidad de viajar a las estaciones base y aun permite actualizar su software sin tener que hacer el viaje.
- Soporta multiplexación de múltiples sistemas digitales localizadores montados en un enlace simple.
- Soporta señal de localización de más de 6,400 *bps.* y formatos de alta velocidad como el *FLEXTM*.
- Soporta corrección de error saliente para todos los datos de localización y en todos los formatos
- Soporta autoalineamiento de modulación en retraso.
- Soporta regeneración de señales en los puntos de transmisión.
- Provee un sistema comprensible y equipo de diagnóstico.
- Actualiza simultáneamente los programas de la estación base vía el enlace del sistema.

1.5.2.3.2. Núcleos transmisores

- Control por microprocesador ofrece la máxima flexibilidad para cambios en el sistema y parámetros de operación.
- Un sistema de procesamiento digital de la señal soporta cuatro niveles de operación y códigos de alta velocidad.
- Un diseño térmico óptimo y disipadores térmicos confiables para prolongar la vida de los componentes e incrementar la confiabilidad de la estación.
- Fácil mantenimiento debido a su diseño completamente modular, panel frontal de acceso y módulos de componentes tipo gaveta.
- Trabaja con diferentes distribución de medios: línea física, radio enlaces, microondas.
- Una amplia variedad de salidas de RF disponibles desde 20 hasta 350 watos.
- Disponible en diferentes bandas de frecuencia: VHF, 280 Mhz, 900 Mhz y UHF.
- Acomoda todos los formatos de localización: voz, 2 tonos, 5/6 tonos, *POCSAG*, *GSC*, y *FLEXTM* que pueden usarse solos o en combinación.

1.5.3. ESTACIONES BASE

Las estaciones base pueden ser local o remotamente controladas y monitoreadas. Un controlador remoto centralmente localizado, ofrece muchas ventajas operacionales. Esto es especialmente cierto en el caso de transmisión simultanea donde la ecualización del sistema de parámetros importantes en la estación base pueden realizarse automáticamente desde cualquiera de las estaciones.

2. RECONOCIMIENTO DE VOZ

La investigación en el área del tratamiento de la voz, ha permitido desarrollar sistemas de reconocimiento de voz que el hecho de comunicarse con un procesador ya no es una utopía, y mediante una serie de herramientas, se pueden crear programas en los que la interacción del usuario y la máquina sea lo más cercana posible a la relación común que se tiene con las demás personas.

La meta de este capítulo es dar una serie de principios de reconocimiento de voz, que permitan adecuarlo a un sistema integrado de mensajería ya que en teoría, el reconocimiento de voz nunca escucha conversaciones entre personas, y cuando se da un comando de parte del usuario, el ordenador nunca comete un error. Si los ordenadores tuvieran un reconocimiento de voz perfecto, todos estarían hablando con ellos en forma natural. Pasarán décadas antes de que esta tecnología sea perfecta. A pesar de este detalle, el reconocimiento de voz es muy útil en un sistema integrado de mensajería, ya que únicamente permite introducir nuevas interfaces del usuario para un mejor manejo del entorno informático, y así en combinación con otras tecnologías existentes poder ampliar las funcionalidades de este sistema y obtener un rendimiento máximo del mismo.

Uno de los aspectos más difíciles del reconocimiento de voz es su naturaleza interdisciplinaria, a continuación, se consideran las disciplinas aplicadas en el reconocimiento de voz:

1. Procesamiento de señal: es el proceso de extraer información pertinente de la señal de voz en una manera robusta y eficiente. Incluido en el procesamiento de señal, el análisis espectral usado para caracterizar las propiedades variantes en el tiempo de la voz, así como varios tipos de preprocesamiento y postprocesamiento de señal de la misma, para hacer la voz más robusta al ambiente magnetofónico y tener una señal de voz perfeccionada.
2. Acústica: la ciencia que comprende la relación entre la señal física de la voz y los mecanismos fisiológicos (mecanismo del tracto vocal humano) que producen la voz y la perciben (mecanismo del oído humano).
3. Modelo del reconocimiento de voz: juego de algoritmos usados para agrupar datos y crear prototipos en base a esos datos, y comparar un par de modelos en base a las características medidas de los patrones.
4. Teoría de la comunicación y la información: procedimientos para estimar parámetros de modelos estadísticos; los métodos para detectar la presencia de modelos particulares de voz, juego de algoritmos de codificación y decodificación (algoritmos de pila, decodificación *Viterbi*) usados para buscar un camino mejor en el reconocimiento del habla continua.
5. Lingüística: comprende la relación entre sonidos (fonología), palabras de un idioma (sintaxis), significado de las palabras habladas (semántica), y el sentido derivado del significado (pragmático). Incluido dentro de esta disciplina la metodología de la gramática y equivalencia de un lenguaje.
6. Fisiología: comprende los mecanismos de orden alto del sistema nervioso humano que cuenta para la producción y la percepción de la voz.

7. Ciencia de la computación: comprende el estudio de algoritmos eficientes para llevar a cabo, en *software* y/o *hardware*, los varios métodos usados en un sistema práctico de reconocimiento de voz.
8. Psicología: comprende los factores que permiten usar una tecnología para seres humanos en tareas prácticas.

2.1. Reconocimiento de voz

El reconocimiento de voz es la habilidad de un ordenador para entender y comprender palabras y frases habladas con el propósito de ejecutar una acción concreta con los datos obtenidos; en otras palabras, convertir una señal acústica de voz a texto. Gracias a las técnicas de reconocimiento de voz, este proporciona ventajas significativas a los usuarios y hace que las aplicaciones sean más fáciles de utilizar. En este caso, como es una aplicación de telefonía y como está se beneficia en gran medida del reconocimiento de voz, porque el único medio de comunicarse con el usuario es a través del teléfono. El reconocimiento de voz, es un gran sustituto de los menús realizados por los tonos de los botones del teléfono digital, no solo por su naturalidad y flexibilidad, sino porque muchos usuarios no tienen teléfonos que permitan la funcionalidad anteriormente comentada.

2.1.1. Tipos de Reconocimiento de voz

El reconocimiento de voz, se complica un poco más a la hora de clasificarlo, a continuación, se menciona la clasificación de las diversas técnicas de reconocimiento de voz. a) palabra aislada (*isolated word*) y/o expresión simple (*single utterance*), b) dependencia del hablante (*speaker dependent*), c) discurso continuo y discreto (*continuous and discret speech*) y d) independiente del hablante (*speaker independent*). A todas estas técnicas se les debe sumar una característica muy importante en el

reconocimiento de voz la cual es el tamaño del vocabulario, este puede incluir un vocabulario pequeño que permite a los usuarios dar órdenes simples a sus ordenadores; o grande que permite al usuario dictar un documento, este aumenta los requisitos de memoria y procesador del ordenador.

2.1.1.1. Palabra aislada y/o expresión simple

Se refiere al tipo de órdenes que se dan al sistema reconocedor. Este tipo de sistema reconocedor requiere una pausa entre dos órdenes, para que él comprenda cuando finalizan dichas órdenes. También se le conoce como sistema de expresión simple porque las órdenes que se emplean son cortas ráfagas de voz, tales como palabras o frases; “Salir”, “Hazlo” y “Consulta mi correo” son ejemplos de palabras aisladas o expresiones simples.

2.1.1.2. Dependencia del hablante

En este sistema se requiere que cada usuario utilice aproximadamente 30 minutos para entrenar al sistema con su propia voz; así, sólo reconoce de forma fidedigna las órdenes dichas por la persona en la que se especializa. La ventaja de un sistema dependiente del hablante es que, en último caso, sólo el usuario autorizado tiene el control del sistema. Por tanto, esta característica podría usarse para construir un dispositivo de seguridad basado en la voz de ese usuario en particular. Si la voz no coincide con la que la computadora reconoce, la computadora ignorará lo que se le diga.

2.1.1.3. Discurso continuo y discreto

Si el reconocimiento de voz es continuo, los usuarios pueden hablar de forma natural al ordenador. A medida que el usuario habla, el sistema reconocedor selecciona algunas palabras y frases para determinar qué se le está diciendo. Si es discreto, los usuarios necesitan hacer una pausa entre palabras. Estos tipos de reconocedores permiten las aplicaciones de dictado. Obviamente, se prefiere el reconocimiento continuo en lugar del discreto, con la salvedad de que el continuo requiere más tiempo de proceso para el computador.

2.1.1.4. Independiente del hablante

El reconocimiento de voz con independencia del hablante, trabaja directamente sin necesidad de entrenamiento alguno; de forma que el reconocedor entiende a cualquiera que hable. Este es el tipo de reconocimiento más difícil, porque hay una miríada de diferencias entre los hablantes. Por ejemplo, el acento de una persona del oriente del país es diferente al acento de uno, del centro y occidente.

Aunque cualquier combinación de las cuatro técnicas es posible, dos de ellas son las más utilizadas actualmente:

2.1.1.5. Comando y control

Este tipo de reconocimiento de voz es continuo, utiliza un vocabulario pequeño y un hablante independiente. Esto significa que los usuarios pueden utilizar varios centenares de órdenes diferentes o frases. Si un usuario dice una orden que no está en la lista, el sistema de reconocimiento de voz devolverá la frase “Comando no reconocido”, o pensará que oyó un comando sonoro similar. Dado que los usuarios de “Comando y

Control” sólo pueden decir frases específicas, éstas deben ser visibles en pantalla para que el usuario sepa qué decir en cualquier momento.

2.1.1.6. Dictado discreto

Utiliza un vocabulario discreto, grande y dependiente del hablante. Se utiliza para dictar en los procesadores de texto, mandar correo electrónico, o dictar comandos naturales del lenguaje que se utilice. Aunque los usuarios puedan decir lo que deseen, deben dejar pausas entre las palabras y realizar el dictado de forma antinatural, es decir que lo que se hable no sonará de igual forma que cuando lo hacemos con otras personas, debido a las pausas efectuadas. El dictado discreto requiere como mínimo un procesador Pentium 60 MHz. con 8 Mbytes de RAM.

2.1.2. Funcionamiento de la tecnología del reconocimiento de VOZ

La voz puede ser tratada digitalmente de varias formas, actualmente se utilizan dos formas:

1. Conversión Análoga – Digital
2. Análisis espectral y transformadas de *Fourier*

2.1.2.1. Conversión análoga – digital

Esta técnica es la forma más sencilla de utilizar y la de menor coste computacional, a continuación se describe el proceso de reconocimiento de voz por medio de la conversión analógica – digital.

La mayoría de formas de reconocimiento de voz, convierten datos de audio entrantes en fonemas específicos que se traducen en texto (un fonema es la unidad estructural más pequeña de sonido que puede usarse para distinguir una pronunciación de otro en un idioma hablado). Esta técnica tiene como base un programa de software residente que queda en memoria una vez instalado, en primer lugar el *software* debe de tener la capacidad para determinar cuando empieza y termina una orden, esto lo logra transformando la señal de voz que entra por el micrófono en tensiones. Cuando la señal de voz excede un cierto valor, en ese momento la orden habrá comenzado y cuando, la señal caiga por debajo de ese nivel, la orden habrá finalizado. La transformación de la señal de voz a tensiones requiere que se utilice un convertidor analógico – digital (ADC), las tensiones son convertidas en números binarios o un formato puramente digital, generalmente en el rango de -127 a 128 para un convertidor de 8 bits, y de -32768 a 32767 para un convertidor de 16 bits. Cada uno de los valores representa la posición de la onda de sonido en un instante determinado.

Esta cadena de *bits* resultante es almacenada en algún lugar de la memoria (*buffer*), después de que el *software* determina que se ha terminado de pronunciar la orden, generalmente captando una pausa significativa en la voz, pasará el contenido del *buffer* por una serie de filtros para eliminar el ruido de fondo que puede interferir en el proceso de reconocimiento de patrones.

Después del proceso de filtrado, el *software* realiza el algoritmo de reconocimiento de patrones. La cadena de *bytes* grabada por el *software* se compara con cada uno de los patrones disponibles, que son cadenas de *bits* que el sistema ha grabado previamente cuando se entrena para reconocer una serie de órdenes habladas. Al realizarse cada comparación, se calcula una puntuación basada en cuanto se parece la palabra al patrón. Cuando se ha calculado la puntuación para cada patrón, el *software* escoge aquella o aquellas con una puntuación mayor. Si la puntuación está dentro de unos márgenes de tolerancia determinados, el sistema ejecutará una macro asociada a ese patrón. La

ejecución de la macro sólo ocurre si un solo patrón es seleccionado por el *software*. Si selecciona más de uno, el sistema supone que no puede asociar la palabra pronunciada con uno de los patrones y simplemente ignora la orden.

A continuación se muestra un algoritmo general para este tipo de técnica de reconocimiento de voz.

Inicio;

 Recibir señal de voz

 Transformar señal de voz a formato digital

 Grabar formato digital de voz

 Filtrar formato digital de voz

 Inicio;

 Compara la voz con las pregrabadas

 Puntuar cada voz

 Fin;

 Escoger la voz regrabada con la puntuación mayor

 Ejecutar macro

Fin.

2.1.2.2. Análisis espectral y transformadas de *Fourier*

Esta técnica se utiliza para desarrollar sistemas de reconocimiento de voz más complejos que involucren una gramática continua e independiente del hablante, que define un juego de palabras y frases que pueden reconocerse. Una gramática puede usar reglas para predecir palabras, probablemente para seguir la palabra hablada, o puede definir el contexto que identifique el asunto de dictado y el estilo esperado de un idioma.

Típicamente el reconocimiento comienza con el muestreo digital del lenguaje. El siguiente paso es el proceso de la señal acústica. Las técnicas más frecuentes incluyen el análisis espectral; por ejemplo, análisis *LPC* (*Linear Predictive Coding*, predicción lineal codificada), *MFCC* (*Mel Frequency Cepstral Coefficients*), modelado de la cóclea, y otras.

El siguiente paso es el reconocimiento de fonemas, grupos de fonemas y palabras. Este paso puede ejecutarse de varias formas, mediante *DTW* (*Dynamic Time Warping*), *HMM* (*Hidden Markov Modelling*), *NNs* (*Neural Networks*), sistemas expertos y combinación de técnicas. Actualmente los sistemas basados en *HMM*'s son los más frecuentemente utilizados, y la aproximación más exitosa.

A continuación se describen los métodos aproximados de reconocimiento de voz y los procesos anteriormente mencionados que lleva producir un sistema de reconocimiento de voz, donde se pueda utilizar un dictado continuo y que sea independiente del hablante.

2.1.2.2.1. Aproximaciones del reconocimiento de voz

A continuación se presentan tres métodos aproximados para llevar a cabo el reconocimiento de voz:

1. Aproximación Acústica fonética.
2. Aproximación por medio del reconocimiento de la muestra.
3. Aproximación por medio de la inteligencia artificial.

2.1.2.2.1.1. Aproximación acústica fonética

El primer paso en este proceso es el análisis del sistema de voz el cual proporciona una apropiada representación espectral de las características de la señal de voz variantes en el tiempo. Las técnicas más comunes del análisis espectral son el banco de filtros y la predicción lineal codificada (*LPC*) los cuales se describirán más adelante, estas técnicas proporcionan descripciones espectrales de la voz en el tiempo.

El próximo paso en el proceso es el de detección de la característica de la voz. La idea aquí es convertir las medidas espectrales a un juego de características grandes que describen las propiedades acústicas de las diferentes unidades fonéticas. Entre las características propuestas para el reconocimiento esta la presencia o ausencia de la resonancia nasal, la presencia o ausencia de la excitación cuando se habla, las frecuencias de las tres primeras resonancias, la forma aperiódica o periódica, y la presencia de frecuencias altas y bajas. Estas características propuestas son binarias y continuas inherentemente. Este proceso normalmente consiste en un juego de detectores que operan en paralelo y tienen la lógica apropiada para tomar la decisión acerca de la presencia o ausencia de un valor o una característica.

Los algoritmos usados para estos detectores son bastantes sofisticados que hacen más fácil el procesamiento de señal.

El tercer paso en este proceso es la segmentación y calificación con que el sistema intenta encontrar regiones estables para luego calificar la región segmentada que otorga las características dentro de esa región para luego igualar las unidades acústicas fonéticas. Esta es la base del reconocedor acústico fonético y es la más difícil de llevar a cabo.

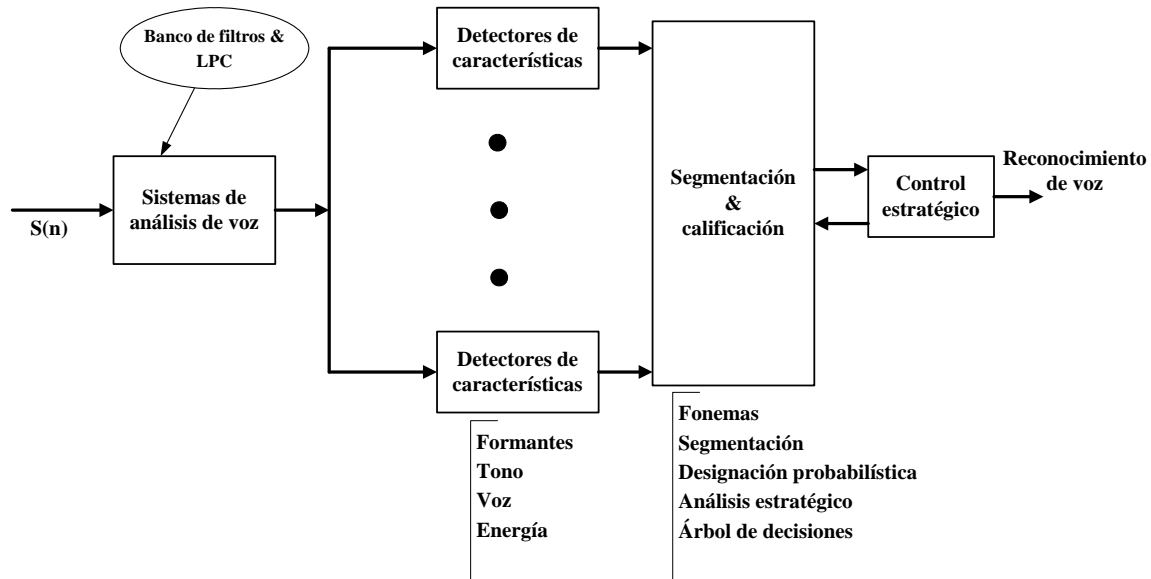
La figura 3 (página 42) muestra un diagrama en bloques de un sistema de reconocimiento de voz por aproximación acústica fonética.

2.1.2.2.1.2. Aproximación por medio del reconocimiento de la muestra

Este modelo se muestra en la figura 4 (página 42) y se basa en cuatro pasos que son los siguientes:

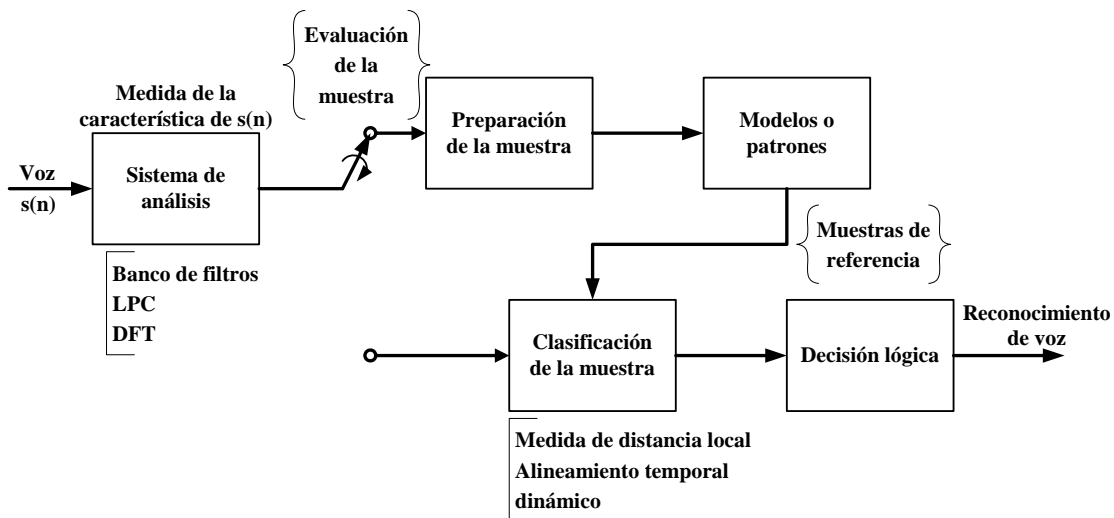
1. La medida de la característica, en la cual una sucesión de medidas se le hacen a la señal de entrada para definir el modelo de la muestra. Para las señales de voz las medidas características son normalmente las salidas de algún tipo de técnica de análisis espectral, como un analizador de un banco de filtros, una predicción lineal codificada (LPC), o un análisis discreto de la transformada de *Fourier (DFT)*.
2. Trazo de la muestra, en la cual se usan una o mas muestras de prueba que tengan los mismos sonidos de la voz, esto se hace para crear una muestra con esas características. La muestra resultante llamada muestra de referencia puede ser un modelo que caracteriza las estadísticas de las características de la muestra de referencia.
3. Clasificación de la muestra, en la cual las muestras de prueba se comparan con la muestra de referencia de cada clase y mide una similitud entre las muestras de prueba y las muestras de referencia.
4. Decisión lógica, aquí se decide que modelo de referencia tiene similitud con las muestras de prueba desconocidas.

Figura 3. Sistema de reconocimiento de voz por aproximación acústica fonética



Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 45.

Figura 4. Reconocimiento de voz por aproximación por medio del reconocimiento de la muestra



Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 51.

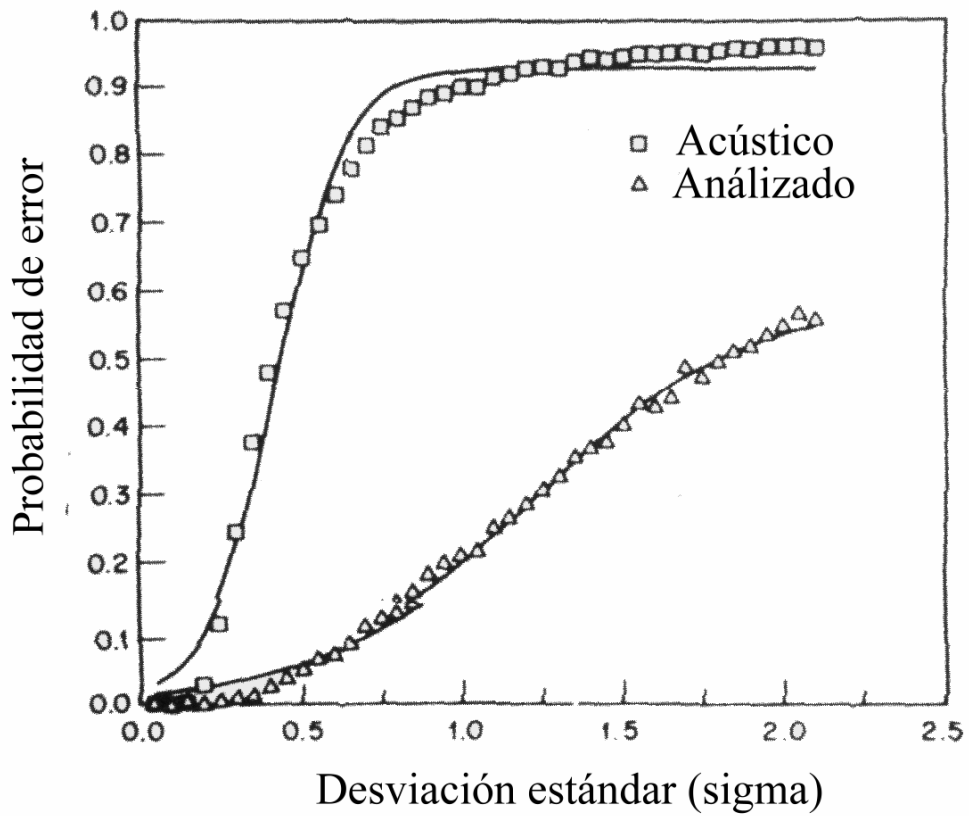
2.1.2.2.1.3. Aproximación por medio de la inteligencia artificial

La idea básica de usar la inteligencia artificial en el reconocimiento de voz es recopilar una variedad de fuentes del habla. Por ejemplo fuentes acústicas, léxicas, sintácticas, semánticas y pragmáticas.

La figura 5 (página 44) muestra la capacidad de corrección de palabras con sintaxis en el reconocimiento de voz. Cuando el parámetro de desviación se pone más grande, la probabilidad de error aumenta para ambos casos; sin embargo, sin la sintaxis la probabilidad de error llega rápidamente a 1.00, pero con sintaxis aumenta gradualmente con aumentos en el ruido.

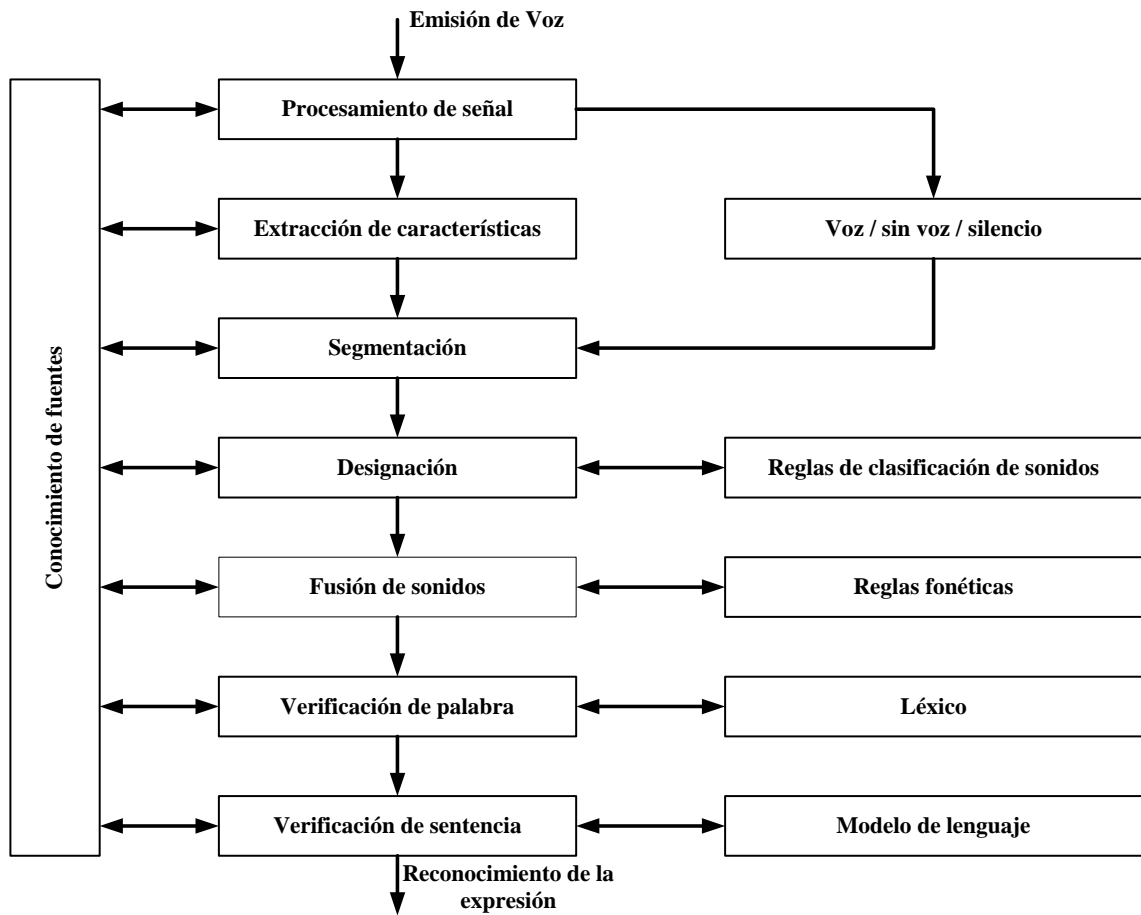
Hay varias maneras de integrar fuentes del habla en un reconocedor de voz. Quizás la aproximación más normal es un procesador *bottom – up* que se muestra en la figura 6 (página 45) en que los procesos de niveles bajos (ej. Descubrimiento de las características, decodificación fonética) precede a los procesos de niveles altos (decodificación léxica, modelo del idioma) de una manera secuencial para comprimir cada fase del proceso tan pequeño como sea posible. Otra alternativa es el llamado procesador *top down* en el que el modelo del lenguaje genera una hipótesis de palabras que se comparan contra la señal de voz sintácticamente y semánticamente las frases significativas se construyen en base a las palabras recopiladas, este modelo se muestra en la figura 7 (página 46).

Figura 5. Capacidad de corrección de palabras con sintaxis

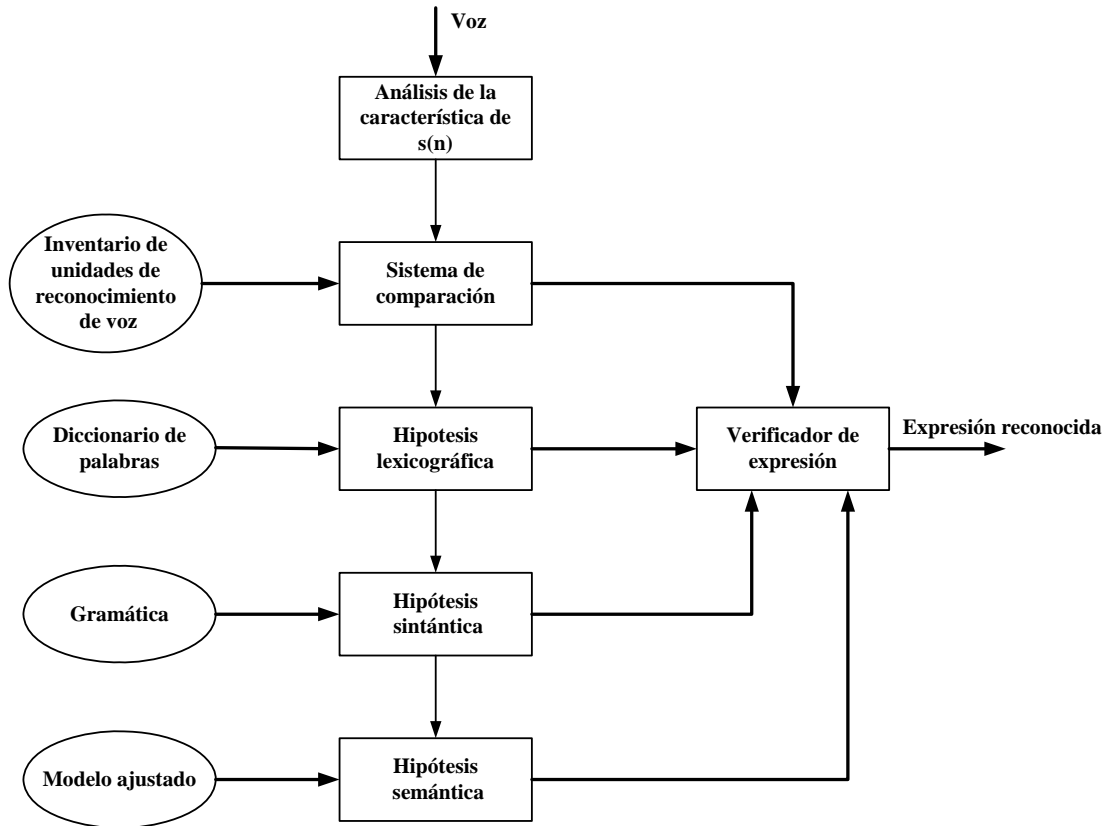


Fuente: *Lawrence Rabiner & Biing-Hwang Juang, Fundamentals of speech recognition, Pág. 54.*

Figura 6. Procesador *bottom – up*



Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 55.

Figura 7. Procesador *top – down*

Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 55.

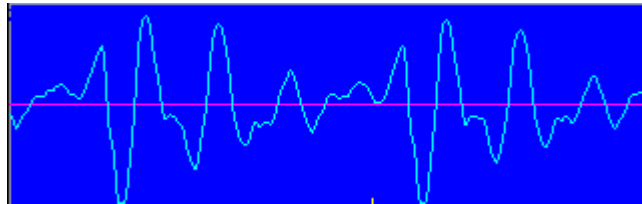
2.1.2.2.2. Tracto vocal y señal de voz

A continuación se describe la naturaleza del tracto vocal dando como resultado una señal de voz que puede ser caracterizada por un juego de propiedades espectrales y temporales que dependen de los rasgos acústicos fonéticos de los sonidos y son manifestados en la forma de onda de la voz.

El análisis de la señal de voz se lleva cabo mediante un modelo que describe el proceso del habla clasificando las señales en dos tipos:

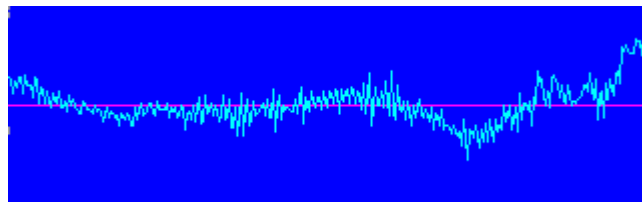
- a. Las señales sonoras que se caracterizan por tener alta energía y contenido frecuencial en el rango de los 300 Hz a 4000 Hz las cuales se generan por intermedio de las cuerdas vocales y además presentan cierta periodicidad como se muestra en la figura 8.

Figura 8. Señales sonoras



- b. Las señales no sonoras también conocidas como fricativas se caracterizan por tener baja energía y componente frecuencial uniforme presentando aleatoriedad en forma de ruido blanco se muestra en la figura 9.

Figura 9. Señales no sonoras



2.1.2.2.2.1. Modelo del tracto vocal

Este modelo se caracteriza en dividir la señal de voz en dos partes, anteriormente mencionadas las cuales se modelan como dos fuentes de excitación las que alimentan al sistema acústico que conforma el tracto vocal.

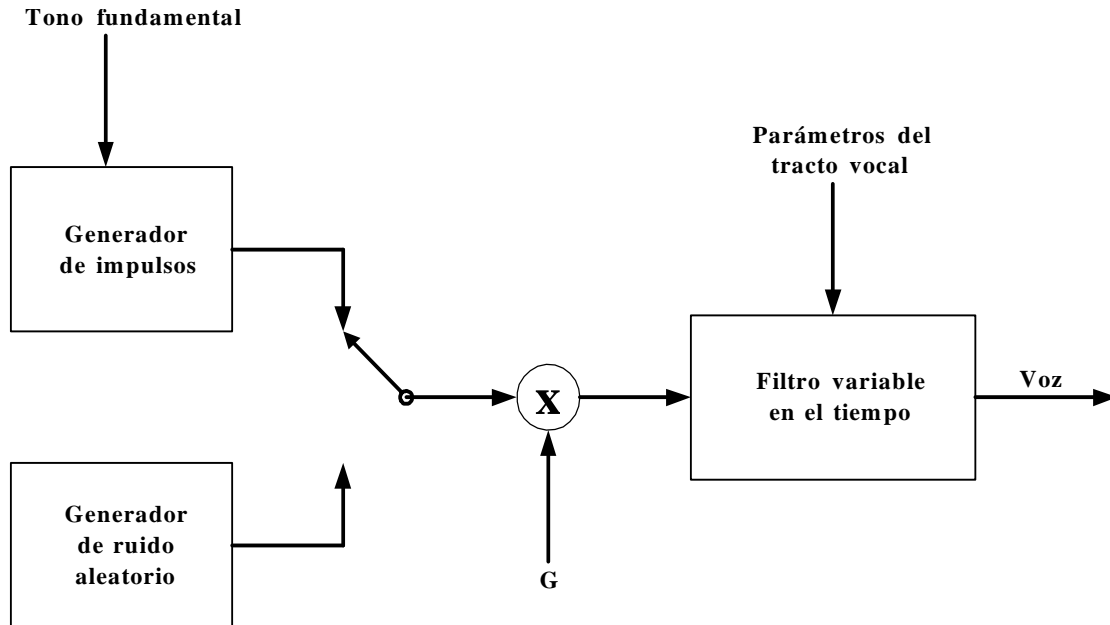
El tracto vocal modelado se manifiesta como un filtro variable en el tiempo cuyos parámetros varían en el tiempo en función de la acción consciente que se realiza al pronunciar una palabra.

El filtro variable en el tiempo tiene dos posibles señales de entrada que dependerán del tipo de señal, sonora o no sonora. Para señales sonoras la excitación será un tren de impulsos de frecuencia controlada, mientras que para las señales no sonoras la excitación será ruido aleatorio.

La combinación de estas señales simula el funcionamiento de la glotis. El espectro de frecuencias de la señal de voz puede obtenerse a partir del producto del espectro de la excitación por la respuesta en frecuencia del filtro.

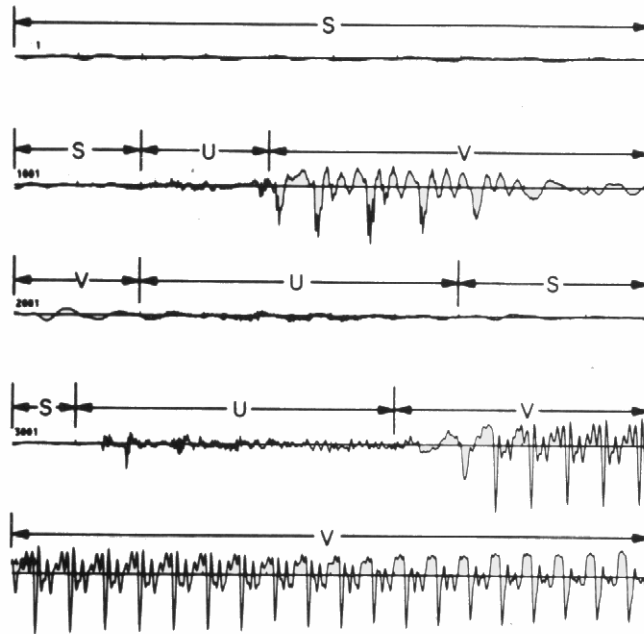
El tracto vocal manifiesta un número muy grande de resonancias, sin embargo, se consideran solo las tres o cuatro primeras que toman el nombre de formantes y cubren un rango de frecuencias entre 100 y 3500 Hz. Esto debido a que las resonancias de alta frecuencia son atenuadas por la característica frecuencial del tracto que tiende a actuar como un filtro pasabajo con una caída de aproximadamente -12 dB por octava.

Este modelo es una simplificación del proceso de voz. Los sonidos fricativos no se filtran por el tracto con la misma extensión en que lo hacen las señales sonoras por lo que el modelo no es muy preciso para este tipo de señales. Además, el modelo supone que las dos señales pueden separarse sin considerar ninguna interacción entre ellas, lo cual no es cierto ya que la vibración de cuerdas vocales es afectada por las ondas de presión dentro del tracto. Sin embargo estas consideraciones pueden ser ignoradas resultando el modelo lo suficientemente adecuado, el cual se muestra en la figura 10.

Figura 10. Modelo del tracto vocal

2.1.2.2.2. Voz representada en el tiempo y en el dominio de la frecuencia

La señal de voz en un espacio de tiempo es una señal variante, en el sentido que cuando se muestra en un período suficientemente corto (entre 5 y 100 msec.), sus características son bastantes estacionarias; sin embargo en los períodos largos de tiempo (en el orden de 1/5 segundos o más) las características señaladas cambian para reflejar los diferentes sonidos de voz hablados. Una ilustración de este efecto se da en la figura 11, que muestra la forma de onda que corresponde a los sonidos iniciales de una frase en el tiempo.

Figura 11. Formas de onda de la voz

Fuente: *Lawrence Rabiner & Biing-Hwang Juang, Fundamentals of speech recognition, Pág. 18.*

Analizando la figura 11, se tiene que cada línea de la forma de onda corresponde a 100 mseg de señal (1/10 seg.), y que cada parte entera abarca aproximadamente 0.5 seg.

La naturaleza variante de la señal en el espacio de tiempo puede empezar a ser vista en los primeros 100 mseg de la forma de onda (primera línea) que corresponde a el silencio de fondo que por consiguiente es baja en amplitud; en los próximos 100 mseg de la forma de onda (segunda línea), se empieza a mostrar pequeños aumentos en la amplitud que van marcando aumentos significativos para que la forma de onda tenga un cambio grande y empiece a tener una periodicidad.

Hay varias formas de clasificar los eventos de la voz, quizás la más simple está en vía de la producción de la voz como se muestra en la figura 11. Se pueden tener aquí tres estados que son:

1. Silencio (S): donde ningún sonido se produce,
2. Sin voz (U): donde los cordones vocales no están vibrando y la forma de onda es aperiódica, y
3. Voz (V): donde los cordones vocales se tensan y por consiguiente vibran periódicamente, entonces tenemos una forma de onda casi periódica.

Debe quedar claro que la segmentación de la forma de onda en varias partes de las expresiones de voz no es tan exacta, ya que a menudo es difícil distinguir un sonido débil de uno fuerte.

2.1.2.2.3. Procesamiento de señal y métodos de análisis del reconocimiento de voz

En un sistema de reconocimiento de voz, a su nivel más elemental, comprende una colección de algoritmos deducidos de una variedad grande de disciplinas como el modelo estadístico, teoría de la comunicación, procesamiento de señal, matemática combinatorial y lingüística, entre otros. Quizás el denominador común de todos los sistemas de reconocimiento de voz es el procesamiento de señal, que es el que convierte la forma de onda de la voz a algún tipo de representación paramétrica para el análisis extenso del mismo.

Una amplia gama de posibilidades existe para representar la señal de voz paraméricamente. Estos incluyen la energía de tiempo corto, ratas de cruce por cero, ratas de niveles de paso, y otros parámetros relacionados. Probablemente la representación paramétrica más importante de la voz es el espectro de tiempo corto, los métodos del análisis espectral son considerados por consiguiente como el centro del procesamiento de señal en un sistema de reconocimiento de voz. En este apartado se describen dos métodos dominantes del análisis espectral: el análisis espectral en el modelo del banco de filtros y el análisis espectral en el modelo de la predicción lineal

codificada (LPC). También se describe la técnica llamada cuantización vectorial que es un procedimiento que codifica la representación espectral continua, el conjunto de formas típicas espectrales codificadas hacen un índice codificado (*codebook*), conformando los niveles de cuantización, esta técnica de cuantización vectorial puede aplicarse a cualquier representación espectral y puede ser incluida en los modelos de banco de filtros y *LPC*.

2.1.2.2.3.1. Preprocesamiento y procesamiento de señal

Antes de seleccionar el proceso para llevar a cabo el análisis de procesamiento de la señal de voz se hace necesario realizar un preprocesamiento y procesamiento de la señal de voz. El procesamiento es igual al descrito anteriormente como conversión analógica - digital. En cambio el preprocesamiento se realiza a través de técnicas que permitan extraer la información acústica directamente de la señal de voz emitida. Dichas técnicas son el preénfasis y la aplicación de una ventana de *Hamming*.

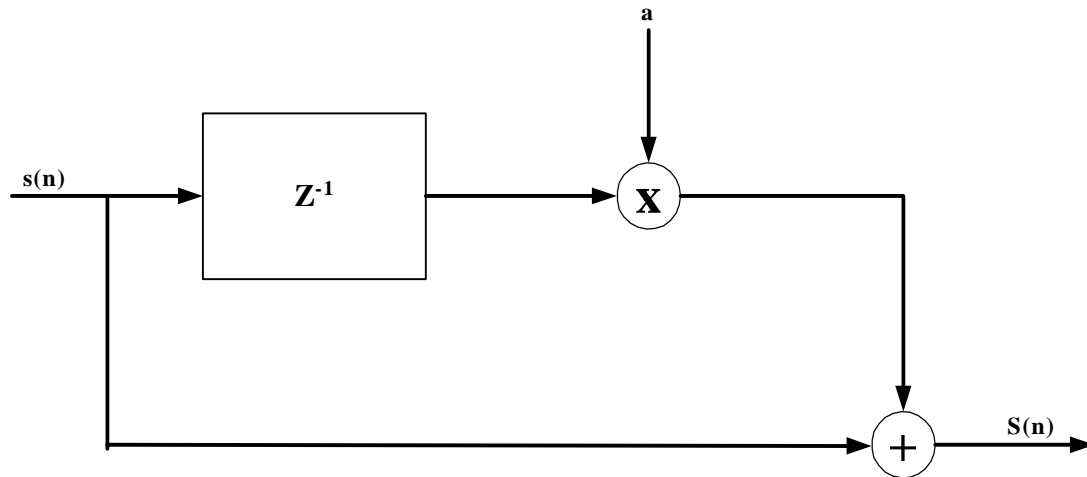
La etapa de preénfasis se realiza con el propósito de suavizar el espectro y reducir las inestabilidades de cálculo asociadas con las operaciones aritméticas de precisión finita. Además se usa para compensar la caída de -6 dB que experimenta la señal al pasar a través del tracto vocal. Se usa un filtro digital de primer orden como el que se muestra en la figura 12 y cuya función de transferencia es:

$$H(z) = az^{-1} \quad (2.1)$$

donde $a = 0.95$

La ecuación diferencial es:

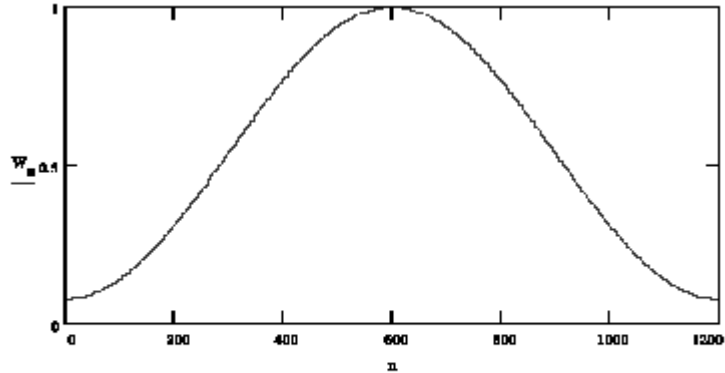
$$S(n) = s(n) - a * s(n-1) \quad (2.2)$$

Figura 12. Forma general de un filtro digital de primer orden

En la etapa siguiente, la señal preacentuada se toma cada 10mseg., por espacio de 20 mseg., y se la somete a una ventana de *Hamming* con el objeto de suavizar la señal en los bordes de dicha ventana, la cual se muestra en la figura 13.

Esta es la ventana que generalmente se usa para el análisis de señales de voz, y se define como:

$$W[nT] = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N}\right) \quad 0 < n < N \quad (2.3)$$

Figura 13. Ventana de *Hamming*

$$W[nT] = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N}\right), \quad 0 < n < N$$

2.1.2.2.3.2. Modelos del análisis espectral

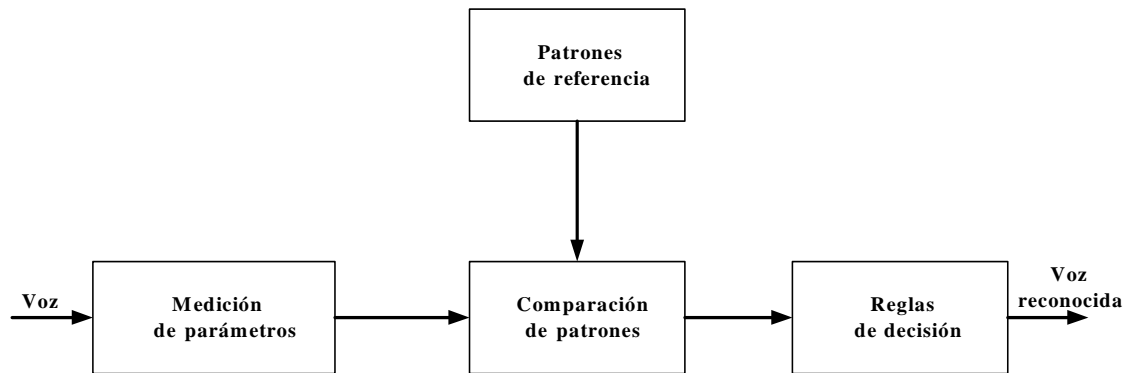
La figura 14 detalla en bloques, el reconocimiento de la muestra y la figura 15 muestra el modelo acústico fonético el cual se utiliza como ejemplo para describir el reconocimiento de voz, dan una idea en bloques de los pasos a seguir para el proceso de reconocimiento de voz y que a continuación se mencionan:

1. Medición del parámetro (en la cual la muestra de prueba se crea).
2. Comparación de la muestra.
3. Decisión.

La función de la medida de los parámetros esta representada por eventos acústicos relevantes de la señal de voz en términos simples y eficientes de los parámetros de voz. Aunque la decisión de cual de los parámetros utilizar debe considerarse por otras causas (ejemplo, eficiencia en el *software*, tipo de aplicación, memoria disponible), la forma en que la representación debe elegirse debe basarse estrictamente en las consideraciones de procesamiento de señal. En el modelo del reconocimiento acústico fonético, el primer paso en el proceso es esencialmente idéntico al utilizado en el reconocimiento de la

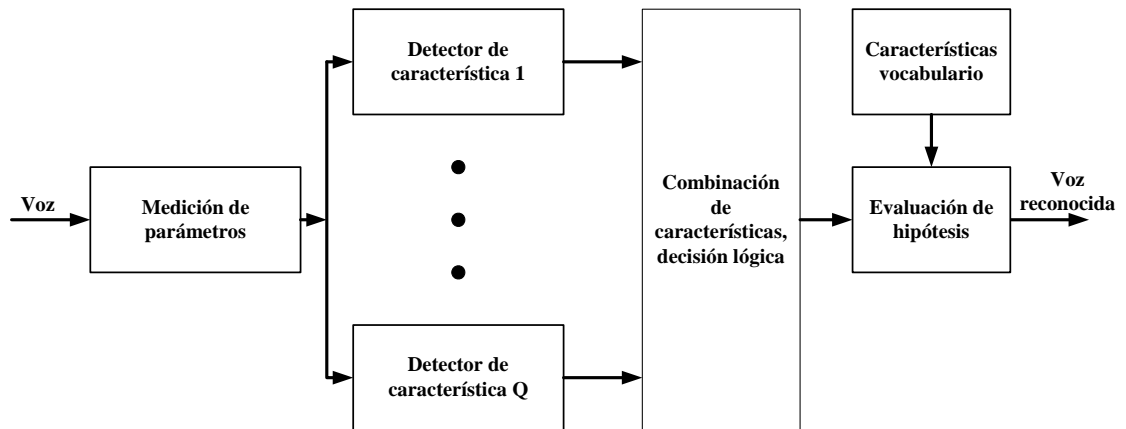
muestra, llamado medición de parámetros, aunque el siguiente paso es totalmente diferente.

Figura 14. Aproximación por medio del reconocimiento de la muestra



Fuente: *Lawrence Rabiner & Biing-Hwang Juang, Fundamentals of speech recognition, Pág. 71.*

Figura 15. Aproximación acústica fonética



Fuente: *Lawrence Rabiner & Biing-Hwang Juang, Fundamentals of speech recognition, Pág. 71.*

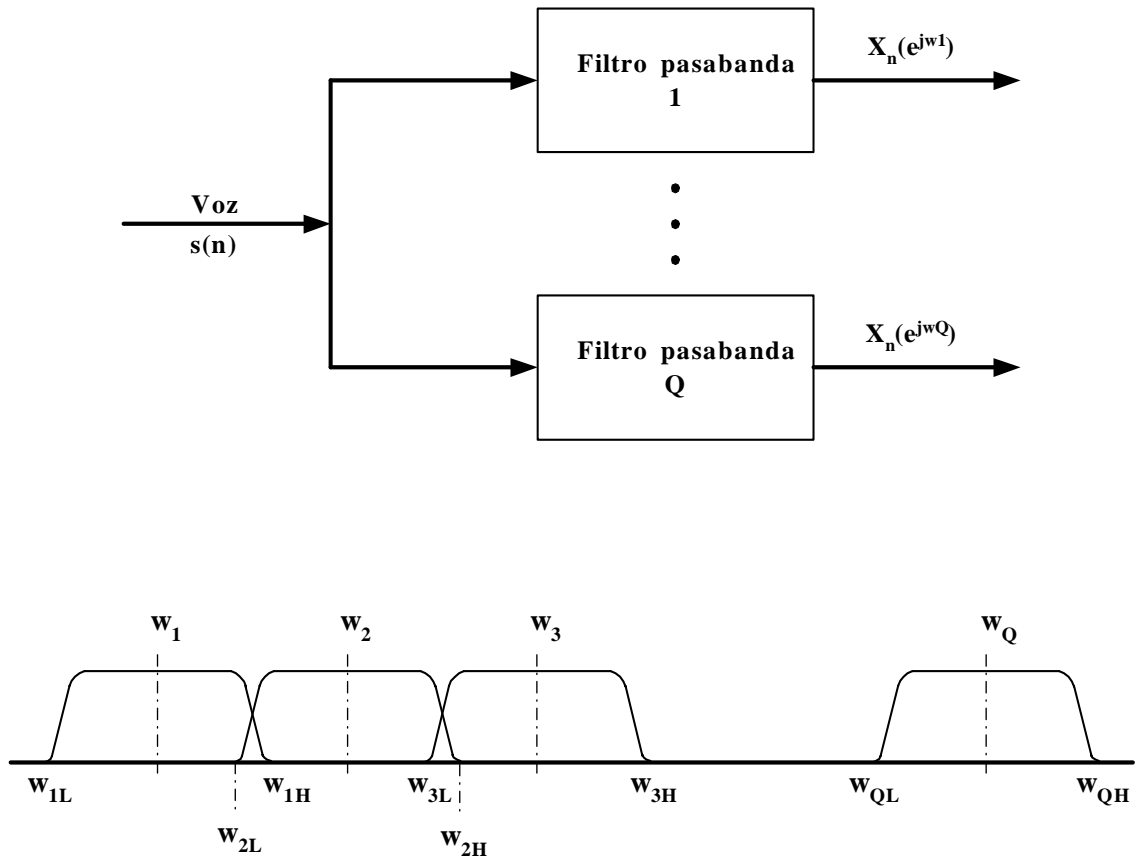
Como se mencionó previamente, los procesos de señal comúnmente seleccionados desde el principio y final del reconocimiento de voz son un modelo de banco de filtros y un modelo *LPC* (*Linear Predictive Coding*). La estructura completa del modelo del banco de filtros se muestra en la figura 16. La señal de voz, $S(n)$, es pasada a través de un banco de filtros pasabanda el cual cubre el rango de frecuencia del rango que nos interesa en la señal (ejemplo 100 – 3000 Hz para señales telefónicas de calidad, 100 – 8000 Hz para señales de banda ancha). Los filtros individuales pueden y generalmente se solapan en frecuencia, como se muestra en la figura 16. La salida del filtro pasabanda i^{th} , $X_n(e^{j\omega_i})$ (donde ω_i es la frecuencia normalizada $2\pi f_i/F_s$, y F_s frecuencia de muestreo) es una representación espectral en tiempo corto de la señal $S(n)$, en el tiempo n , visto a través del filtro pasabanda i^{th} con el centro de frecuencia en ω_i .

El modelo aproximado LPC se muestra en la figura 17, el análisis espectral se realiza sobre bloques de voz (fragmentos de voz) con un modelo limitado en polos completos. Esto significa que el resultado de la representación espectral $X_n(e^{j\omega})$ está limitada hacer de la forma $(\sigma/A(e^{j\omega}))$, donde $A(e^{j\omega})$ es un polinomio de orden p^{th} con una transformada en z .

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_p z^{-p} \quad (2.4)$$

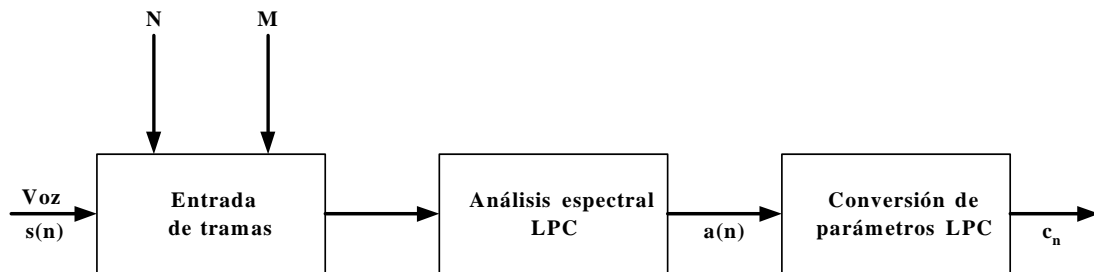
El orden p , es llamado análisis de orden *LPC*. Así la salida del bloque del análisis espectral LPC es un vector de coeficientes (parámetros de *LPC*) eso especifica (paramétricamente) el espectro de un modelo de todos los polos en el espectro señalado sobre un período de tiempo en el que el marco de muestras de la voz es acumulado.

Figura 16. Modelo de análisis de banco de filtros



Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 72.

Figura 17. Modelo de análisis de LPC



Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 72.

2.1.2.2.3.2.1. Procesador de banco de filtros

Como se sabe, para convertir una señal analógica a forma digital primero hay que limitarle el ancho de banda y después muestrearla, como se muestra en la figura 18.

Figura 18. Proceso de filtrado y muestreo de una señal



Antes de muestrear una señal es necesario filtrarla. Teóricamente la máxima frecuencia que puede representar es la mitad de la frecuencia de muestreo. En la práctica se utiliza una frecuencia de muestreo más alta para poder utilizar filtros no ideales.

La señal pasa a representarse en múltiplos del período de muestreo, T , como $s(nT)$, o lo que es lo mismo, s_n .

El sonido telefónico se muestrea de 8kHz. a 16kHz, y se considera suficiente para el reconocimiento y síntesis del lenguaje. El estándar de audio es una frecuencia de muestreo de 44kHz (*Compact Disc, CD*) o 48kHz (*Digital Audio Tape, DAT*) para representar frecuencias de hasta 20kHz.

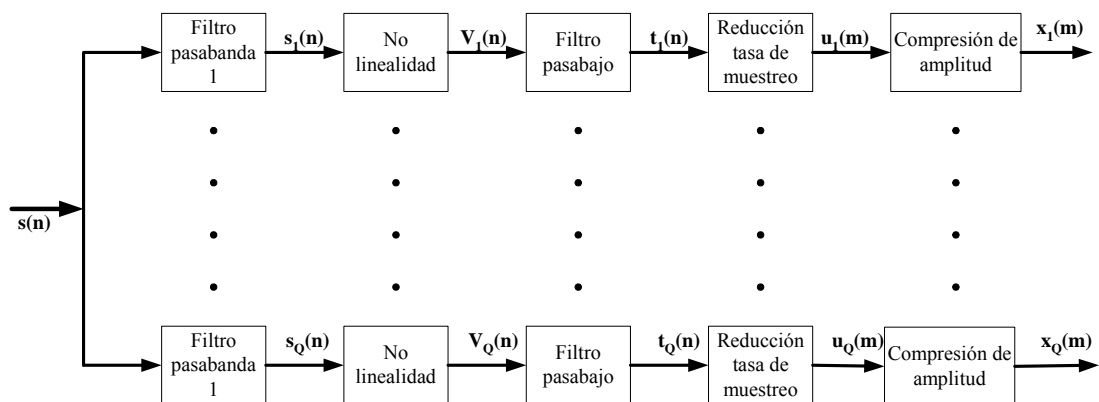
Un diagrama de bloques de la estructura canónica de un banco de filtros es mostrado en la figura 19. Por ejemplo la señal de voz $S(n)$, es pasada a través de un banco de Q filtros pasabanda, obteniendo las señales mostradas en las ecuaciones 2.5 y 2.5b, estos filtros cubren todos los rangos de frecuencia de modo que no se pierda información. Un ejemplo del filtrado se muestra en la figura 20.

$$S_i(n) = s(n) * h_i(n), \quad 1 \leq i \leq Q \quad (2.5a)$$

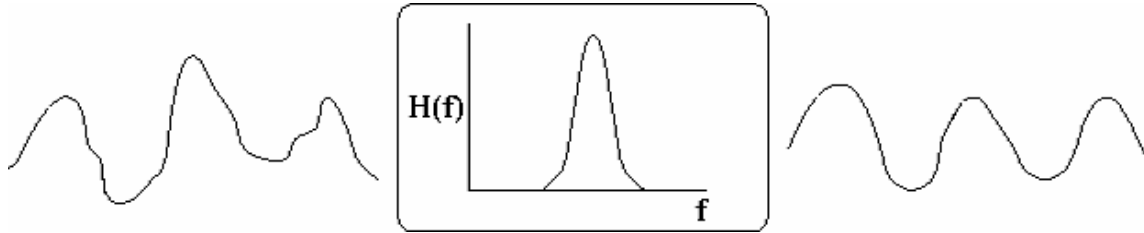
$$S_i(n) = \sum_{m=0}^{M_i-1} h_i(m)s(n-m) \quad (2.5b)$$

Donde se asume que la respuesta al impulso del filtro pasabanda i^{th} es $h_i(m)$ con una duración de M_i muestras; por lo tanto, se utiliza la convolución para representar la operación del filtrado para dar una expresión clara de $S_i(n)$, de la señal de voz filtrada en el pasabanda.

Figura 19. Modelo completo de análisis de banco de filtros



Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 74.

Figura 20. Proceso de filtrado de una señal

El propósito de analizar el banco de filtros es brindar una medida de la energía de la señal de voz en una banda de frecuencia dada, por lo que no es tan importante la fase de la señal, cada señal del pasabandas $S_i(n)$ es pasada a través de un dispositivo no lineal, semejante a un rectificador de onda completa o de media onda. El cambio no lineal del espectro de señal del pasabanda de la banda de frecuencia baja puede crear imágenes con frecuencia alta. Un filtro pasabajo se utiliza para eliminar imágenes de frecuencia alta, dando un juego de señales, $u_i(n)$, $1 < i < Q$, representando una estimación de la energía de la señal de voz en cada una de las bandas de frecuencia Q .

Para entender los efectos no lineales y del filtro pasabajo, se asume que la salida de un filtro pasabanda es una señal senoidal pura con frecuencia w_i .

$$S_i(n) = \alpha_i \text{sen}(w_i n) \quad (2.6)$$

Esta asunción es válida para la voz en el caso en que los sonidos de voz sean firmes cuando el ancho de banda del filtro es suficientemente estrecho que solo una simple armónica de voz pasa a través del filtro pasabanda. Si se utiliza un rectificador de onda completa, se tiene,

$$\begin{aligned} f(S_i(n)) &= s_i(n) && \text{para } s_i(n) \geq 0 \\ &= -s_i(n) && \text{para } s_i(n) \leq 0 \end{aligned} \quad (2.7)$$

entonces se puede representar la salida no lineal como

$$v_i(n) = f(s_i(n)) = s_i(n) \cdot w(n) \quad (2.8)$$

donde

$$w_i(n) = \begin{cases} +1 & \text{si } s_i(n) \geq 0 \\ -1 & \text{si } s_i(n) \leq 0 \end{cases} \quad (2.9)$$

esto se ilustra en la figura 21a y 21c.

Entonces la salida no lineal puede ser vista como una modulación en el tiempo, como se muestra en la ecuación (2.8), y en el dominio de la frecuencia se puede obtener el resultado siguiente:

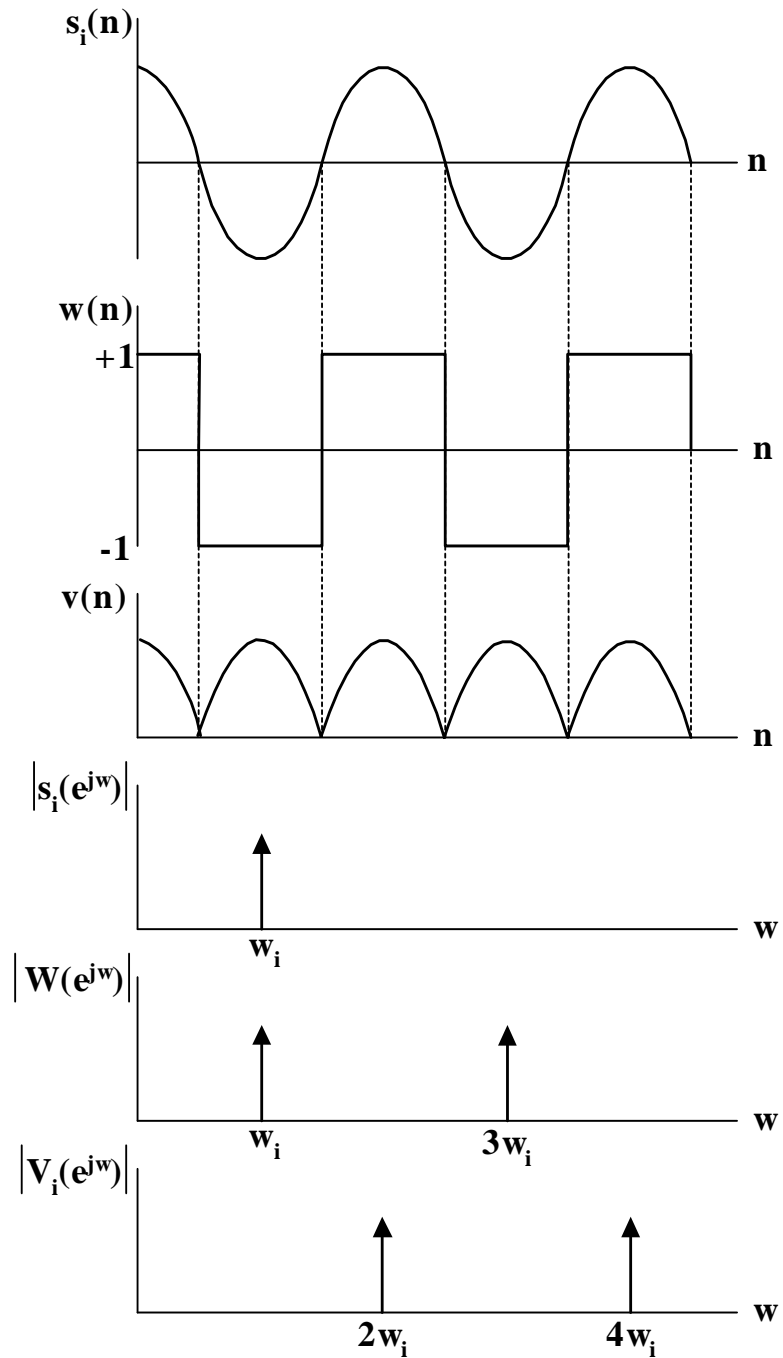
$$v_i(e^{j\omega}) = S_i(e^{j\omega}) \otimes W(e^{j\omega}) \quad (2.10)$$

donde $V_i(e^{j\omega})$, $S_i(e^{j\omega})$ y $W(e^{j\omega})$ son las transformadas de *Fourier* de las señales $V_i(n)$, $S_i(n)$ y $w(n)$ respectivamente, y \otimes es una convolución circular. El espectro $S_i(e^{j\omega})$ es un solo impulso con $\omega_0 = \omega_i$, mientras que el espectro $W(e^{j\omega})$ es un juego de impulsos de frecuencias armónicas extrañas $\omega_q = \omega_i q$, $q = 1, 3, \dots, q_{\max}$. Por lo tanto el espectro de $V_i(e^{j\omega})$ es un impulso con $\omega = 0$ y un juego de impulsos con amplitud mas pequeña con $\omega_q = \omega_i q$, $q = 2, 4, 6, \dots$, como se muestra en la figura 21d y 21f. El efecto del filtro pasabajo es retener la componente de DC de $V_i(e^{j\omega})$ y la salida del filtro tiene componentes de alta frecuencia a causa de la no linealidad.

El análisis anterior, solo es estrictamente correcto para una senoidal pura, es un buen modelo para voces casi periódicas de sonidos del habla, tan largos como el ancho del filtro pasabanda que no puede poseer dos o más señales armónicas fuertes. Porque la variación en el tiempo de la señal de voz, (ejemplo: es casi periódica), el espectro de la señal pasabaja no es un impulso de DC puro. Pero mientras la información de la señal este contenida en una banda baja de frecuencia alrededor de DC. El papel final del filtro pasabaja es eliminar los picos espectrales no deseados.

El ancho de banda de la señal, $V_i(n)$, esta relacionado con la rápida velocidad del movimiento de las armónicas del habla en una banda estrecha, y ésta generalmente se conoce en el orden de 20 a 30 Hz. Por lo tanto los dos últimos bloques del modelo de banco de filtros de la figura 19 son un ejemplo de la reducción de la velocidad de la señal filtrada con un filtro pasabajo, $t_i(n)$, esta es remplazada por una velocidad del orden de 40 a 60 Hz., y la señal de rango dinámico es comprimida utilizando una compresión de amplitud (ejemplo codificación logarítmica o codificación ley μ).

Figura 21. Formas de onda y espectros típicos en el análisis de banco de filtros



Fuente: *Lawrence Rabiner & Biing-Hwang Juang, Fundamentals of speech recognition, Pág. 75.*

2.1.2.2.3.2.1.1. Tipos de banco de filtros utilizados para el reconocimiento de voz

El tipo de banco de filtros más utilizado para el reconocimiento de voz es el uniforme para el cual la frecuencia central, f_i , del filtro pasabandas es definida como

$$f_i = \frac{F_s}{N} i, \quad 1 \leq i \leq Q \quad (2.11)$$

Donde F_s es una muestra de la velocidad de la voz y N es el número de espacios uniformes del filtro que requiere para ampliar un rango de frecuencia de la voz.

El número actual de filtros utilizados en el banco de filtros es Q , que satisfacen la relación

$$Q \leq N/2 \quad (2.12)$$

Con igualdad cuando el rango de frecuencia completa de la señal de voz es usada en el análisis

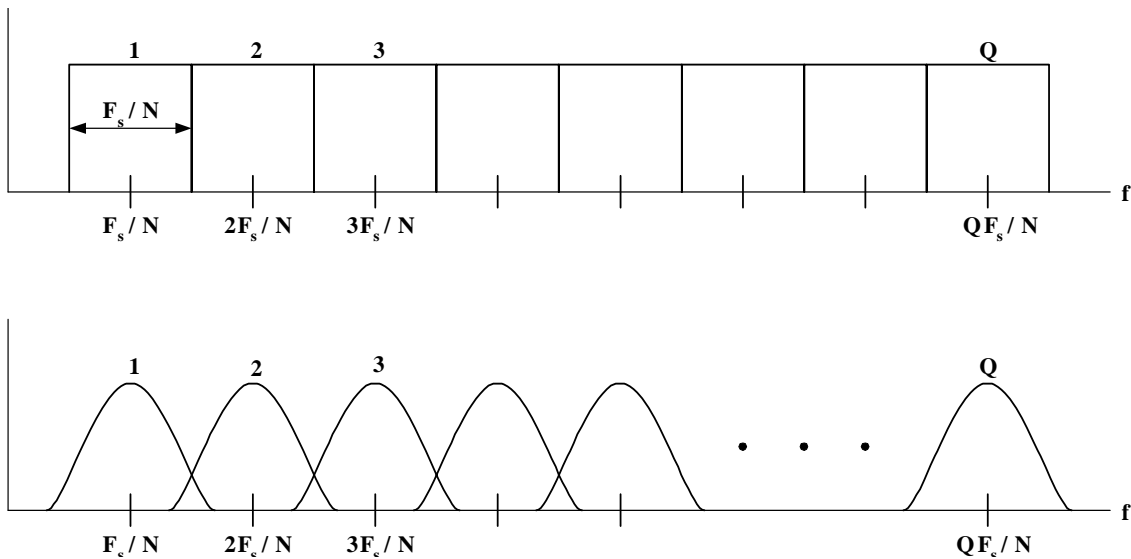
$$Q = N/2 \quad (2.13)$$

El ancho de banda, b_i , del filtro i^{th} , generalmente satisface la propiedad

$$b_i \geq \frac{F_s}{N} \quad (2.14)$$

Con igualdad significa que no hay frecuencia sobrepuesta entre canales de filtros adyacentes y sin igualdad significa la superposición en los canales de filtros adyacentes. (Si $b_i < F_s/N$ entonces las porciones seguras del espectro de voz se perderían en el análisis y el resultado del espectro de voz no sería considerado muy significativo). La figura 22a muestra un juego de Q filtros ideales, no sobrepuestos, filtro pasabanda de rango de cobertura de $F_s/N(1/2)$ a $(F_s/N)(Q+1/2)$. Al mismo tiempo, la figura 22b muestra un juego más realista de Q filtros sobrepuestos cubriendo el mismo rango aproximadamente.

Figura 22. Serie de filtros pasabanda de Q canales



Fuente: Lawrence Rabiner & Bñng-Hwang Juang, *Fundamentals of speech recognition*, Pág. 77.

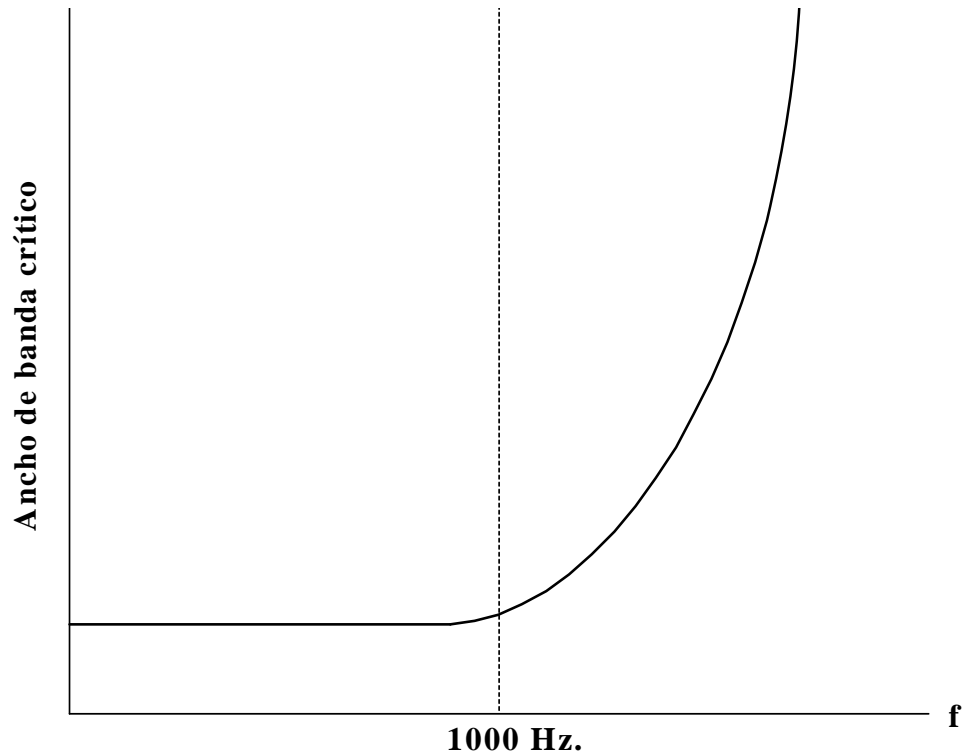
La alternativa para uniformar el banco de filtros es diseñar un banco de filtros no uniforme diseñado acorde a algún criterio para filtros individuales debiendo espaciar la frecuencia. Un criterio comúnmente utilizado es espaciar el filtro uniformemente a lo largo de la escala de frecuencia logarítmica (una frecuencia logarítmica se obtiene para justificar la percepción auditiva humana). Además para un juego de Q filtros pasabanda con frecuencia central f_i y ancho de banda b_i , $1 < i < Q$, se tendría

$$\begin{aligned}
 b_1 &= C \\
 b_i &= \alpha b_{i-1}, \quad 2 \leq i \leq Q \\
 f_i &= f_1 + \sum_{j=1}^{i-1} b_j + \frac{(b_i - b_1)}{2}
 \end{aligned} \tag{2.15}$$

donde C y f_1 son anchos de banda arbitrarios en frecuencias centrales del primer filtro y α es el factor de crecimiento logarítmico.

Los valores de α más comúnmente utilizados son $\alpha = 2$ el cual da un 1/8 del espacio de banda del filtro adyacente, y $\alpha = 4/3$ el cual da 1/3 del espacio del octavo filtro.

Un criterio para diseñar un banco de filtros no uniforme es utilizar directamente una escala de banda crítica. El espacio del filtro a lo largo de la banda crítica está basado sobre un estudio de percepción y está proyectado a escoger la banda que da contribución igual a la articulación de la voz. La forma general de la banda crítica está en la figura 23. La escala es cerrada hasta las frecuencias lineales debajo de 1000 Hz (el ancho de banda es esencialmente constante como una función de f), y es cercana a las frecuencias logarítmicas a partir de 1000 Hz (el ancho de banda es esencialmente exponencial como una función de f). Muchas variantes de la escala de banda crítica han sido utilizadas, incluyendo la escala *mel* y la escala *bark*.

Figura 23. Variación del ancho de banda con respecto a la frecuencia

Fuente: *Lawrence Rabiner & Biing-Hwang Juang, Fundamentals of speech recognition, Pág. 79.*

La mayor parte del diseño de uniformidad o no uniformidad está basado en la escala de bandas críticas. Ya que estas escalas de frecuencias no lineales aproximan la sensibilidad del oído humano y evitan el problema de relación entre frecuencia y tiempo, a continuación se mencionan varias de ellas:

- Q constante: Q es la relación del ancho de banda del filtro sobre la frecuencia central, por lo que implica una forma exponencial.
- Ancho de banda rectangular equivalente (ERB): se miden los anchos de banda de los filtros auditores.
- *Bark*: derivada de los experimentos de percepción.

- *Mel*: la solución de los ingenieros que tiene una forma analítica sencilla: $m=1125 \log(0.0016f + 1)$, donde la f es la frecuencia en Hz y m es la frecuencia resultante.

2.1.2.2.3.2.1.2. Desarrollo de banco de filtros

Un banco de filtros puede ser desarrollado de muchas maneras dependiendo del método utilizado para diseñar el filtro individual, los métodos de diseño para filtros digitales se reducen a dos clases:

1. Respuesta de impulso infinito (*IIR*)
2. Respuesta de impulso finito (*FIR*)

2.1.2.2.3.2.1.2.1. Respuesta de impulso infinito (*IIR*)

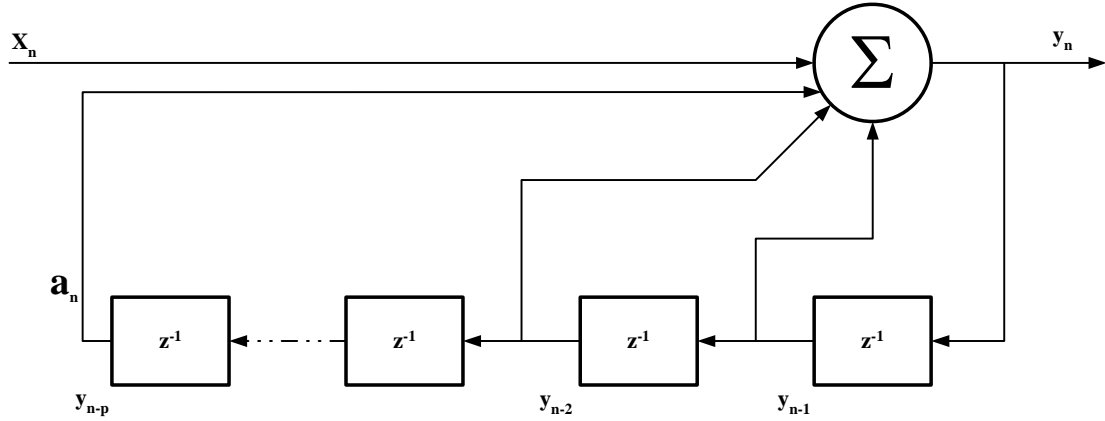
También llamados comúnmente filtros recursivos, su implementación es la más eficiente para realizar filtros individuales pasabanda tienen estructura paralela, pero no pueden diseñarse para que tengan una fase lineal exacta.

Un filtro de respuesta impulso infinito produce una salida, $y(n)$, que es la suma ponderada de la entrada actual y las salidas pasadas, esto se muestra en la figura 24

La forma general de un filtro *IIR* viene dada por:

$$y_n = \sum_{i=1}^p a_i y_{n-i} + \sum_{j=0}^q b_j x_{n-j} \quad (2.16)$$

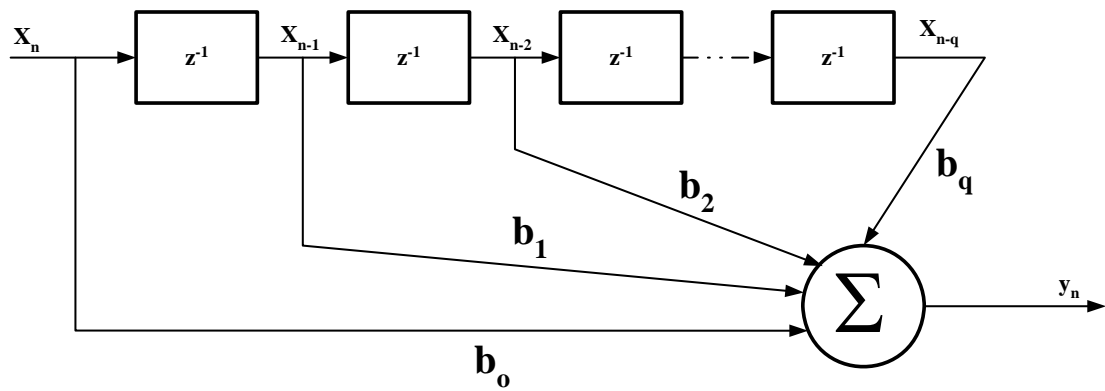
Figura 24. Modelo general de un filtro IIR



2.1.2.2.3.2.1.2.2. Respuesta de impulso finito (FIR)

Un filtro de respuesta de impulso finito (*Finite Impulse Response, FIR*) produce una salida, $y(n)$, que es la suma ponderada de la entrada actual y las entradas pasadas, $x(n)$. Esto se muestra en la figura 25 con Z^{-1} representando un retraso unitario.

Figura 25. Modelo general de un filtro FIR



La forma general de un filtro *FIR* viene dada por:

$$\begin{aligned} y_n &= b_0 x_n + b_1 x_{n-1} + b_2 x_{n-2} + \dots + b_q x_{n-q} \\ &= \sum_{j=0}^q b_j x_{n-j} \end{aligned} \quad (2.17)$$

Esta es la forma más directa y simple de desarrollar filtros individuales, son computacionalmente caros de implementar pero no necesitan introducir distorsiones de fase, por lo que son útiles en el reconocimiento de voz de alta calidad.

Como ejemplo para un banco de filtros, si denotamos la respuesta al impulso por el canal i^{th} como $h_i(n)$, $0 < n < L-1$, entonces la salida del canal i^{th} , $X_i(n)$, puede ser expresada como una convolución discreta finita de una señal de entrada, $s(n)$, con una respuesta al impulso, $h_i(n)$.

$$x_i(n) = s(n) * h_i(n) \quad (2.18a)$$

$$= \sum_{m=0}^{L-1} h_i(m) s(n-m) \quad (2.18b)$$

El cálculo de la ecuación anterior es repetida para cada canal i , donde $i = 1, 2, \dots, Q$.

Otra alternativa es considerar un filtro pasabajo donde se considera que cada filtro tenga una respuesta al impulso; cuya respuesta deseada es:

$$h_i(n) = w(n) e^{jw_i n} \quad (2.19)$$

Desarrollando las ecuaciones se tiene

$$\begin{aligned}
 x_i(n) &= \sum_m w(m) e^{jw_i m} s(n-m) \\
 &= \sum_m s(m) w(n-m) e^{jw_i(n-m)} \\
 &= e^{jw_i n} \sum_m s(m) w(n-m) e^{-jw_i m} \\
 &= e^{jw_i n} S_n(e^{jw_i})
 \end{aligned} \tag{2.20}$$

donde $S_n(e^{jw_i})$ es la transformada de *Fourier* en tiempo corto de $s(n)$ con frecuencia $w_i = 2\pi f_i$. La transformada corta de *Fourier* para $s(m)$ es definida como:

$$S_n(e^{jw_i}) = \sum_m s(m) w(n-m) e^{-jw_i m} \tag{2.21}$$

si se evalúa para $n = n_o$, se tiene la transformada convencional de la señal de ventana

$$S_{n_o}(e^{jw_i}) = FT[s(m)w(n_o - m)]_{w=w_i} \tag{2.22}$$

Con estas ecuaciones se puede establecer lo siguiente para L que es la duración de la muestra:

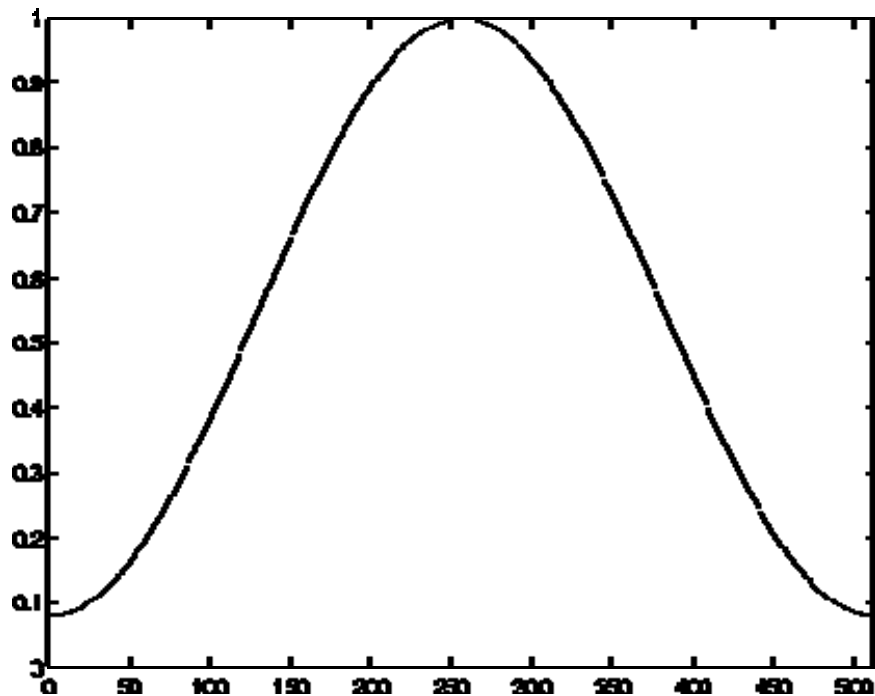
1. Si L es grande en relación a la señal periódica, entonces $S_n(e^{jw_i})$ brinda una buena resolución de la frecuencia. Esto significa que se pueden resolver los picos armónicos pero solo se puede verlo de una manera aproximada sobre el espectro de la sección de voz dentro de la ventana.

2. Si L es pequeña en relación a la señal periódica, entonces $S_n(e^{j\omega})$ brinda una resolución de frecuencia pobre (no hay picos armónicos involucrados), pero una buena estimación del grosor de la figura espectral se obtiene.

La ecuación 2.20 indica que se necesitan un número infinito de coeficientes, en realidad los filtros ideales son imposibles y los más cercanos son difíciles de diseñar, por lo que hay algunas soluciones que a continuación se describen:

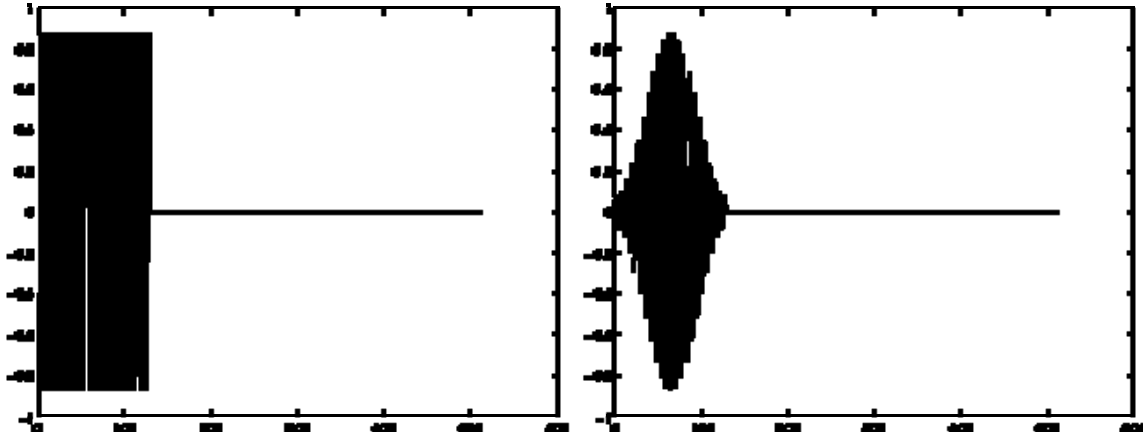
- Truncar la señal: es simple y efectivo si los ciclos son cortos.
- Ventana: en el proceso de la voz, se asume que la señal es estacionaria en intervalos de tiempo lo suficientemente cortos y se realizan transformadas de *Fourier* en cada intervalo, por lo que se procede a multiplicar la señal por una función ventana de *Hamming* cuyo valor fuera de un determinado rango casi cero; esta minimiza la energía en los extremos evitando discontinuidades, es por eso que no se utiliza una ventana rectangular. Una forma general de la ventana de *Hamming* se da en la ecuación (2.23) y es una función coseno que se muestra en la figura 26.

$$w = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) & 0 \leq n < N \\ 0 & \text{otro caso} \end{cases} \quad (2.23)$$

Figura 26. Ventana de *Hamming* de una función coseno

Como se puede observar una ventana de *Hamming* no tiene tanta precisión frecuencial, pero provoca efectos mucho menores que una ventana rectangular. A continuación en las figuras 2.27a y 2.27b se muestra una señal senoidal, con una función ventana rectangular la primera y una función ventana de *Hamming* la segunda.

Figura 27. Función ventana rectangular y de *Hamming*



- Usar un sistema de diseño de filtros más complejo, por ejemplo el algoritmo de intercambio *Parks-McClelland-Remez*, que diseña filtros *FIR* de fase cero óptimos con frecuencias arbitrarias.

2.1.2.2.3.2.1.3. Desarrollo de un banco de filtros uniforme basado en la transformada corta de *Fourier* usando *FFT*

Este procedimiento permite mejorar la eficiencia a nivel *software* de un banco de filtros uniforme con salida igual a la ecuación 2.20. Interesa entonces una frecuencia uniformemente espaciada que puede ser:

$$f_i = i \left(\frac{F_s}{N} \right), \quad i = 0, 1, \dots, N-1 \quad (2.24)$$

Entonces se puede volver a escribir la ecuación 2.20 como

$$x_i(n) = e^{j\left(\frac{2\pi}{N}\right)in} \sum_m s(m)w(n-m)e^{-j\left(\frac{2\pi}{N}\right)im} \quad (2.25)$$

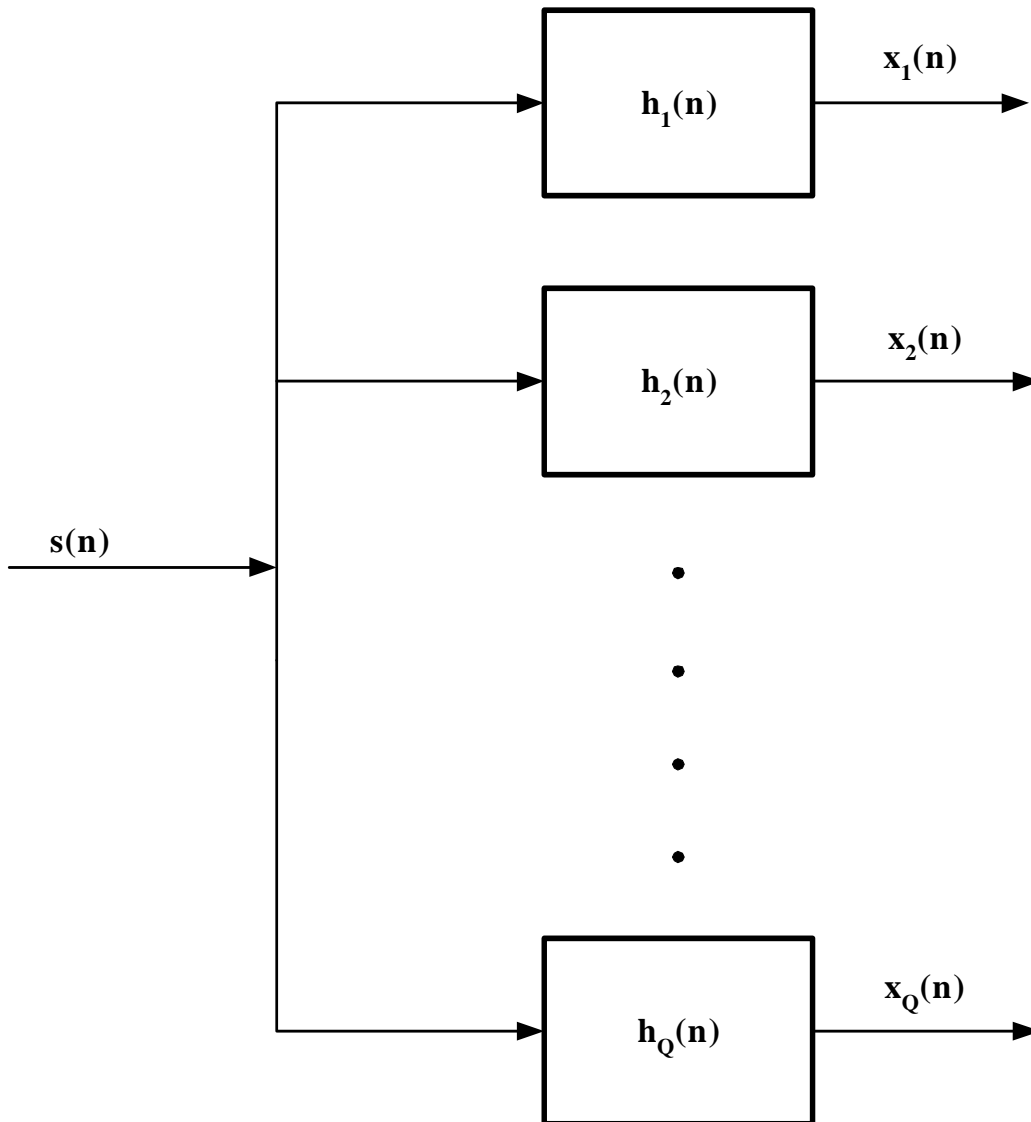
Si se parte sobre m muestras de tamaño N y se asume que $m = Nr + k$ donde $0 \leq k \leq N-1$ y $-\infty < r < \infty$; se tiene una señal de ventana $s_n(m) = s(m)w(n-m)$ donde $m = n - L + 1, \dots, n$, tiene L muestras de duración y se puede escribir la ecuación 2.25 como:

$$x_i(n) = e^{j\left(\frac{2\pi}{N}\right)in} \left[\sum_{k=0}^{N-1} u_n(k) e^{-j\left(\frac{2\pi}{N}\right)ik} \right] \quad (2.26)$$

donde $u_n(k) = \sum_r s_n(Nr + k)$, $0 \leq k \leq N-1$; $X_i(n)$ es una modulación DFT de N puntos de la secuencia de $U_n(k)$.

2.1.2.2.3.2.1.4. Desarrollo del banco de filtros *FIR* no uniforme

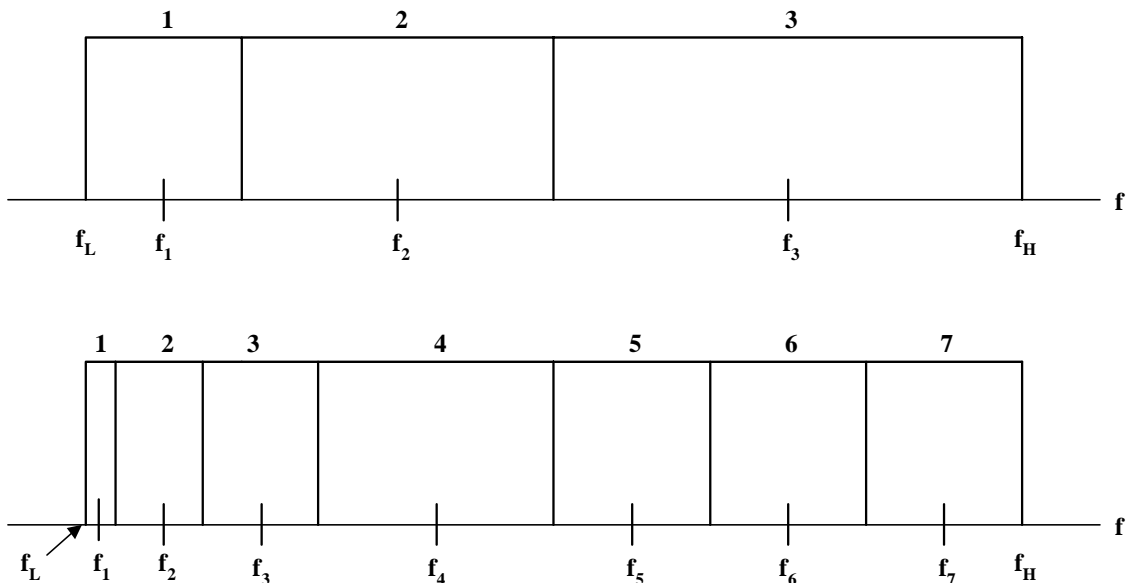
La forma mas general de un banco de filtros *FIR* no uniforme se muestra en la figura 28, donde el filtro pasabanda k^{th} responde al impulso $h_k(n)$ el cual representa un filtro con frecuencia central w_k y un ancho de banda Δw_k .

Figura 28. Forma general de un banco de filtros *FIR* no uniforme

Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 89.

En esta forma tan general cada filtro pasabanda es desarrollado vía una convolución directa cuya estructura *FFT* no eficiente puede ser utilizada. En este caso cada filtro pasabanda es diseñado según el método de ventana utilizando la misma ventana pasabaja, se puede mostrar que la frecuencia compuesta responde al canal *Q* del banco de filtros y es independientemente del número y distribución de los filtros individuales. Así que un banco de filtros con tres filtros tiene exactamente la misma frecuencia compuesta que la de uno de 7 filtros, esto se muestra en la figura 29.

Figura 29. Anchos de banda de bancos de filtros no uniformes de 3 y 7 bandas



Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 90.

Para mostrar esto se denota la respuesta del impulso del filtro pasabanda k^{th} como

$$h_k(n) = w(n)\tilde{h}_k(n) \tag{2.27}$$

Donde $w(n)$ es la ventana FIR, y $\tilde{h}_k(n)$ es la respuesta ideal al impulso del filtro pasabanda diseñado. La respuesta en frecuencia del filtro pasabanda k^{th} , $H_k(e^{jw})$, puede ser escrita como:

$$H_k(e^{jw}) = W(e^{jw}) \otimes \overline{H}_k(e^{jw}) \quad (2.28)$$

Así que la respuesta en frecuencia del banco de filtros compuesto, $H(e^{jw})$ puede ser escrita como:

$$H(e^{jw}) = \sum_{k=1}^Q H_k(e^{jw}) = \sum_{k=1}^Q W(e^{jw}) \otimes \overline{H}_k(e^{jw}) \quad (2.29)$$

Si se intercambia la sumatoria y la convolución se obtiene:

$$H(e^{jw}) = W(e^{jw}) \otimes \sum_{k=1}^Q \overline{H}_k(e^{jw}) \quad (2.30)$$

La sumatoria de la ecuación (2.30), es la sumatoria ideal de la respuesta en frecuencia de cada filtro, se ve que es independiente del número y distribución de los filtros individuales y se puede escribir como:

$$\hat{H}(e^{jw}) = \sum_{k=1}^Q \overline{H}_k(e^{jw}) = \begin{cases} 1, & w_{\min} \leq w \leq w_{\max} \\ 0, & \text{otro caso} \end{cases} \quad (2.31)$$

Donde w_{\min} es la frecuencia baja en el banco de filtros y w_{\max} es la frecuencia más alta. Entonces la ecuación (2.30) puede expresarse como:

$$H(e^{jw}) = W(e^{jw}) \otimes \hat{H}(e^{jw}) \quad (2.32)$$

La cual es el resultado deseado, independientemente del número de filtros ideales, Q , y de su distribución en frecuencia.

2.1.2.2.3.2.1.5. Banco de filtros no uniformes basados en *FFT*

Una manera posible de aprovecharse de la estructura *FFT* para llevar a cabo banco de filtros uniformes es diseñar un banco de filtros grande uniforme (ej., $N = 128$ o 256 canales) y entonces crear uno no uniforme combinando dos o más canales uniformes, se puede escribir esto como:

$$X'_k = X_k + X_{k+1} = \sum_{n=0}^{N-1} \left[x(n) 2e^{-j\frac{\pi n}{N}} \cos\left(\frac{\pi n}{N}\right) \right] e^{-j\frac{2\pi}{N}nk} \quad (2.33)$$

2.1.2.2.3.2.1.6. Resumen de las consideraciones para banco de filtros de reconocimiento de VOZ

Anteriormente se discutieron varios métodos para llevar a cabo banco de filtros para reconocimiento de voz. En este apartado se resumen las consideraciones que se deben tomar para escoger el número y tipos de filtros usados en las estructuras discutidas anteriormente.

La primera consideración para cualquier banco del filtro es el tipo de filtro digital a usar. Las opciones son *IIR* (recursivo) y *FIR* (no recursivo). Los diseños de *IIR* tienen la ventaja de que su desarrollo es simple, eficiente. La desventaja grande de filtros de *IIR* es que su respuesta es no lineal; por lo tanto, para minimizar esta desventaja en el diseño se toman las características de magnitud ideales y las características de magnitud

no ideales para tener una aproximación. Por otro lado, los filtros del *FIR* pueden lograr fase lineal sin considerar características de magnitud ideales. Para aplicaciones del reconocimiento de voz, se ha mostrado cómo una estructura de *FTT* puede aplicarse a menudo para aliviar la ineficacia del *software* en los bancos de filtro *FIR*, las estructuras de banco de filtro digitales más prácticas usan el *FIR* (normalmente en una realización de *FFT*).

Una vez decidido el tipo de filtro a utilizar, la próxima consideración es el número de filtros a ser usados en el banco de filtros. Para los bancos de filtros uniformes, el número de filtros, Q , no puede ser demasiado pequeño o la eficiencia del banco de filtros para filtrar el espectro de voz se daña grandemente; así los valores de Q son aproximadamente menos de 8. Semejantemente, el valor de Q no puede ser demasiado grande (a menos que hallan filtros solapados), porque los anchos de banda del filtro serían demasiados estrechos para algunos hablantes (ej., mujeres con alto diapason o niños), y habría una probabilidad alta de error ya que ciertos pedazos de voz tendrían baja energía (es decir, ningún prominente armónico caería dentro de la ventana). Así, en la práctica tienden a tener valores de $Q \leq 32$. Aunque en los bancos de filtros uniformes se pueden utilizar, muchos sistemas prácticos usan los no uniformes para reducir el *software* y para caracterizar el espectro de voz de una manera más consistente con la percepción humana.

Una consideración final para banco de filtros es la opción no lineal y filtro pasabajo usados para mejorar el rendimiento de la salida de cada canal. Típicamente para la no lineal se ha usado un rectificador de onda completa, rectificador de media onda, o un recortador. El espectro resultante es débil, sensible a la no linealidad. El filtro pasabajo usado en la práctica varía de un simple integrador a un filtro pasabajo *IIR* de buena calidad (típicamente un filtro de *Bessel*).

2.1.2.2.3.2.2. Análisis de predicción lineal para el reconocimiento de voz (LPC)

Una de las técnicas más usadas en el procesamiento de señales de voz es el análisis de predicción lineal. Esta técnica ha probado ser muy eficiente debido a la posibilidad de parametrizar la señal en un número pequeño de patrones con los cuales es posible reconstruirla adecuadamente. Los parámetros obtenidos mediante este método se caracterizan por variar lentamente durante las ventanas de análisis de tiempo.

Mediante esta técnica, se puede representar a la señal de voz mediante parámetros que varían en el tiempo y que están relacionados con la función de transferencia del tracto vocal y las características de la fuente sonora.

Otra ventaja es que no requiere demasiado tiempo de procesamiento, lo que es importante a la hora del desarrollo.

El modelo matemático expuesto establece que el tracto vocal puede modelarse mediante un filtro digital siendo los parámetros los que determinan la función de transferencia. El problema consiste en, dado un segmento de palabra, extraerle sus parámetros que en este caso vienen a ser los coeficientes del filtro.

El análisis de predicción lineal permite aproximar una señal a partir de señales pasadas. En este caso se trata de predecir señales de voz mediante un filtro FIR (filtro de respuesta infinita), cuya función de transferencia se deduce a partir de:

$$s(n) = \sum_{k=1}^p a_k s(n-k) + Gu(n) \quad (2.34)$$

Como se puede observar la señal de voz se representa por medio de señales anteriores y $u(n)$ es la entrada del filtro, el cual es un tren de impulsos periódicos o una fuente de ruido aleatorio.

El tren de impulsos producirá señales sonoras mientras que la fuente de ruido aleatorio producirá señales no sonoras a la salida del filtro. De esta manera el filtro viene a representar un modelo del tracto vocal.

La función de transferencia del filtro se obtiene sacando la transformada z a la relación anterior entonces se tiene:

$$H(z) = \frac{S(z)}{GU(z)} = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} = \frac{1}{A(z)} \quad (2.35)$$

Donde G es la ganancia del filtro y dependerá de la naturaleza de la señal. Dada la señal $s(n)$, el problema consiste en determinar los coeficientes de predicción a_k y la ganancia G . Ya que estos coeficientes se usarán como parámetros del reconocimiento de voz. Su determinación se realiza minimizando el error que se comete cuando se intenta realizar la aproximación de la señal.

Sea $\tilde{s}(n)$ la señal estimada de un número de muestras de la señal original, entonces:

$$\tilde{s}(n) = \sum_{k=1}^p a_k s(n-k) \quad (2.36)$$

El error entre la señal original y la señal estimada es:

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^p a_k s(n-k) \quad (2.37)$$

Mediante el método de los mínimos cuadrados, los coeficientes de predicción se calculan minimizando el error cuadrático medio con respecto a cada uno de los coeficientes. Entonces se tiene:

$$E_n = \sum_m e_n^2(m) \quad 2.38$$

$$E_n = \sum_m \left[s_n(m) - \sum_{k=1}^p a_k s_n(m-k) \right]^2 \quad (2.39)$$

Para resolver la ecuación anterior, para los coeficientes predictores, se realiza la diferenciación E_n con respecto a cada a_k ($\{\partial E_n / \partial a_k = 0\}$ $k = 1, 2, \dots, p$), entonces se tiene:

$$\sum_m s_n(m-i) s_n(m) = \sum_{k=1}^p \hat{a}_k \sum_m s_n(m-i) s_n(m-k) \quad (2.40)$$

Como la ecuación anterior, tiene términos de la covarianza de $s_n(m)$, esta puede escribirse como:

$$\phi_n(i,0) = \sum_{k=1}^p \hat{a}_k \phi_n(i,k) \quad (2.41)$$

Entonces el mínimo error cuadrático medio puede expresarse como:

$$\hat{E}_n = \sum_m s_n^2(m) - \sum_{k=1}^p \hat{a}_k \sum_n s_n(m) s_n(m-k) \quad (2.42)$$

Examinando el error cuadrático se puede notar que consiste en términos fijos y términos que dependen de los coeficientes predictores expresados por la covarianza y siendo esta función de (i,k) se reduce a una autocorrelación simple, por lo que se procede a realizar un análisis de autocorrelación; este proporciona una medida de la correlación de la señal con una copia desfasada en el tiempo de si misma. De aquí se obtienen los p coeficientes de autocorrelación los cuales deben tener valores entre 10 y 15. La función de autocorrelación es simétrica en $r_n(-k) = r_n(k)$, con lo cual da una matriz de p x p tipo *Toeplitz* (simétrica con todos los elementos diagonales iguales), la solución de esta matriz de autocorrelación es más eficiente por medio del algoritmo de *Durbin*.

Teniendo los coeficientes a_k del filtro se dispone de la forma con que la cavidad vocal se comporta y que junto con la señal de excitación se obtiene el sonido emitido en ese momento.

2.1.2.2.3.3. Comparación de la muestra

La siguiente etapa de análisis viene a ser la que se encarga de realizar la comparación de patrones acústicos; el éxito de este tipo de sistemas dependerá de la aplicación de una técnica conocida como alineamiento temporal dinámico (*Dynamic Time Warping*) la cual compara la variación en el tiempo de dos palabras.

El problema se presenta cuando se pronuncia una palabra la cual no siempre se realiza a la misma velocidad, produciendo distorsiones temporales afectando sus componentes acústicos. Estas variaciones no son proporcionales a la velocidad de locución y podrán variar de locutor a locutor; es por esto que se hace necesario un procedimiento que permita comparar dos palabras, sin considerar distorsiones temporales; todo esto lo hacen los algoritmos de programación dinámica.

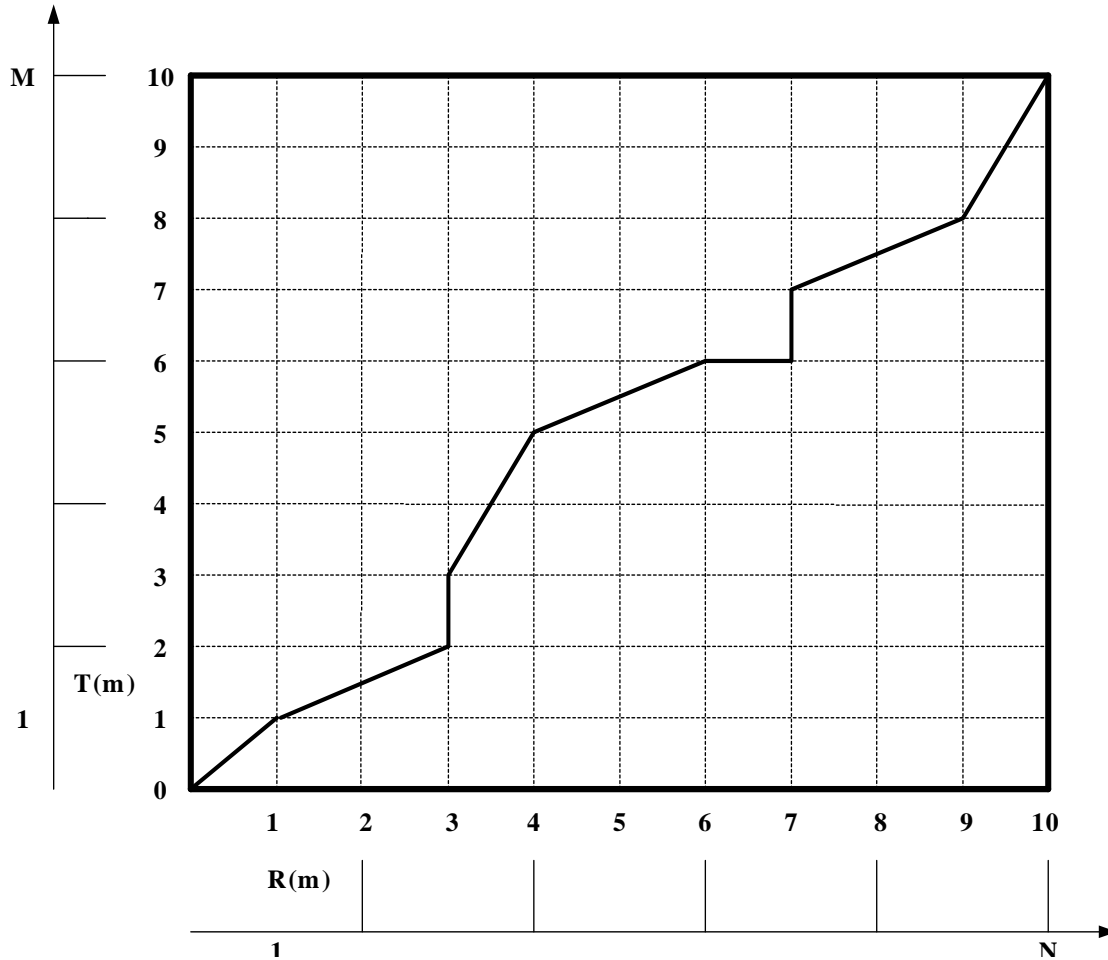
Estos algoritmos al comparar dos palabras proporcionan una medida de disimilitud o medida de distancia que es aprovechada en el reconocimiento de voz.

Sea $R(n)$ donde $n = 1, 2, \dots, N$ la secuencia de ventanas que representan el patrón de la palabra de referencia donde n es en general un vector de parámetros multidimensionales que describe las características de la n -ésima ventana de una palabra. Los componentes del vector en este caso vendrían a ser los coeficientes de predicción lineal, pudiendo ser algún otro tipo de parámetro.

Sea $T(m)$, donde $m = 1, 2, \dots, M$ la secuencia que representa el patrón de la palabra que se desea reconocer. Se necesita hallar la función de alineamiento $m = w(n)$, la cual describirá el camino óptimo. No se puede garantizar que este camino buscado sea una relación funcional simple por lo que en tales casos se hace necesario definir una tercera variable común a los ejes n y m tal que $n = i(k)$, $k = 1, 2, \dots, K$; $m = j(k)$, $k = 1, 2, \dots, K$ donde K es la longitud del eje temporal común.

En la figura 30 se muestran $R(n)$ y $T(m)$ distribuidos sobre un par de ejes coordenados donde cada abscisa n corresponde a un vector de catorce coeficientes los cuales representan el espectro de una ventana de 30 ms. de la palabra $R(n)$. Cada ordenada m corresponderá a un vector de 30 ms. de la palabra $T(m)$.

Figura 30. Alineamiento temporal dinámico



Fuente: Lawrence Rabiner & Biing-Hwang Juang, *Fundamentals of speech recognition*, Pág. 203.

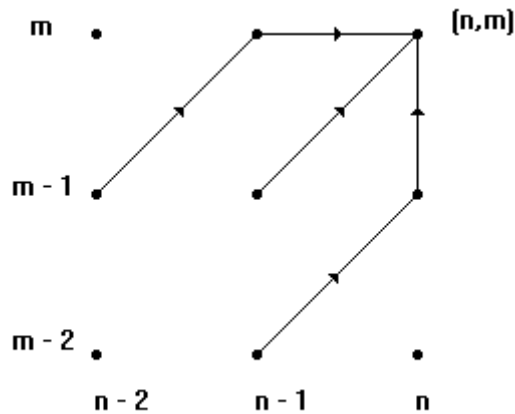
Para que el algoritmo tenga la mayor eficiencia posible, es decir para garantizar un camino óptimo se hace necesario que se cumplan la siguientes consideraciones.

2.1.2.2.3.3.1. Restricción local

Para determinar el camino óptimo debe especificarse una restricción local que garantice evitar excesivas compresiones o extensiones de las escalas de tiempo. Una primera restricción viene a ser: $i(k+1)^3 \leq i(k) \leq j(k+1)^3$ $j(k)$. Seguidamente se procede a restringir el espacio local del camino óptimo en la vecindad del punto (n,m) como

aparece en la figura 31, donde se muestra que los caminos válidos solo pueden provenir de los puntos $(n-1, m-1)$, $(n-1, m-2)$ o $(n-2, m-1)$.

Figura 31. Restricción local

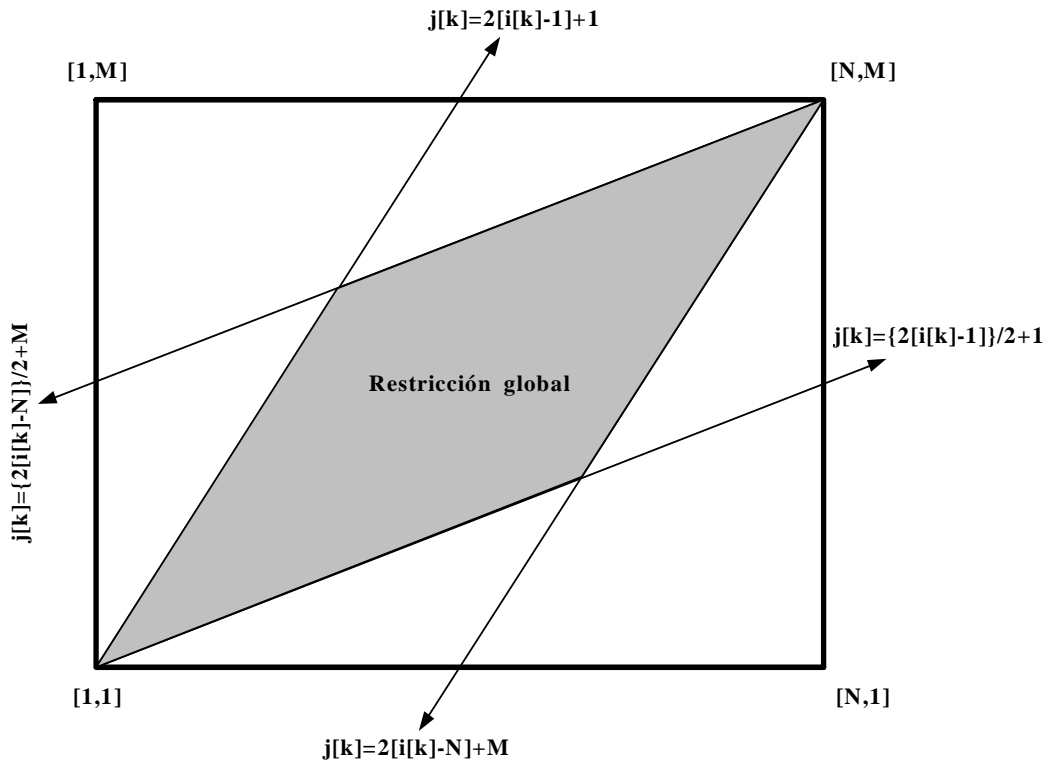


Fuente: *Lawrence Rabiner & Biing-Hwang Juang, Fundamentals of speech recognition, Pág. 210.*

2.1.2.2.3.3.2. Restricción global

Una consideración adicional debe tenerse en cuenta para consolidar el algoritmo, es la que toma en cuenta la existencia de regiones del plano (n,m) por las cuales se puede asegurar que no pasará el camino óptimo. Esto se debe a las restricciones locales explicadas en el punto anterior, y se puede aprovechar para evitar el procesamiento involucrado en regiones innecesarias. En la figura 32 se muestra la restricción global donde para el caso de las producciones usadas, la región sombreada será el único lugar donde se hallará el camino óptimo. De esta forma el procesamiento se concentrará en esta región, optimizando el algoritmo.

Figura 32. Restricción global



Fuente: *Lawrence Rabiner & Biing-Hwang Juang, Fundamentals of speech recognition, Pág. 215.*

2.1.2.2.3.3.3. Medida de disimilitud

Para la obtención del camino óptimo se necesita disponer de una distancia que defina la disimilitud entre dos palabras que a la vez provenga de la función de alineamiento. La forma genérica de dicha función de distancia es:

$$D(i(k), j(k)) = \frac{\sum (i(k), j(k)) \tilde{W}(k)}{N(\tilde{W})} \tag{2.43}$$

Donde $D(i(k),j(k))$ es una función, cuyo cálculo proporciona la distancia total a lo largo del camino óptimo, $d(i(k),j(k))$ es la distancia local entre las ventanas $i(k)$ de la palabra referencia y $j(k)$ de la palabra a reconocer, $W(k)$ es una función de ponderación para k , y $N(W)$ es un factor de normalización que depende de W . El camino corresponderá a aquel que minimice la función de distancia total representada por:

$$\hat{D} = \underset{(K, i(k), j(k))}{Min} \left\{ \sum_{n=1}^N D[i(k), j(k)] \right\} \quad (2.44)$$

La distancia que da los mejores resultados es aquella que se expresa como:

$$d_{cep} = \sum_{n=1}^p (W(n)(c_t(n) - c_r(n)))^2 \quad (2.45)$$

Se conoce como la distancia *cepstral* y está en función de los coeficientes *cepstrales* c_t y c_r de la palabra a reconocer y de la palabra referencia deducidos a partir de:

$$\begin{aligned} c(1) &= -a \\ c(n) &= -a_n - \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) a_k c(n-k), \quad 1 < n \leq p \end{aligned} \quad (2.46)$$

Como se puede observar es una distancia *euclídea* y su uso está muy difundido en procesamiento de voz debido a su simplicidad y porque resulta ser una aproximación de la distancia entre dos espectros logarítmicos representados por sus coeficientes *cepstrales*. Es independiente de como se efectúa el algoritmo *DTW*.

2.1.2.2.3.4. Cuantización vectorial

Una parte importante en cualquier tipo de procesamiento de voz viene a ser la optimización de los algoritmos en cuanto a velocidad y almacenamiento. La técnica que a continuación se expone permitirá un ahorro en memoria lo que a la vez permite que los algoritmos se ejecuten a mayor velocidad ya que no tendrán que hacer uso de dispositivos externos de memoria.

Las técnicas de parametrización de la señal vocal se realizan tomando una secuencia de ventanas de tiempo, cada una de las cuales es representada por un número p de parámetros. Cada ventana de tiempo se puede representar por un vector de p dimensiones.

Cuando se almacenan los parámetros, lo que generalmente se realiza es cuantificar cada parámetro separadamente usando un número determinado de *bits*. Esto se conoce como cuantificación escalar y no es la forma más económica de almacenamiento, ya que implica la ocurrencia uniforme de las ventanas de la señal de voz en el espacio vectorial.

En aplicaciones de codificación y reconocimiento resulta mas conveniente y económico el empleo de una técnica conocida como cuantización o cuantificación vectorial.

La idea principal de la cuantificación vectorial es hacer una partición del espacio vectorial en sectores, cada uno de los cuales será representado por un solo vector que puede ser el centroide. El conjunto de centroides viene a ser el libro índice (*codebook*) que conforman los niveles de cuantificación y a cada uno se le asigna una dirección o etiqueta. Como la parametrización se realiza por ventanas de tiempo pequeños representados por vectores de 14 dimensiones, a cada ventana se le asignará un vector.

Para efectuar la cuantificación de un vector de entrada, se hace asignándole la dirección del vector del libro índice más cercano, evaluado mediante una medida de disimilitud que puede ser la distancia *cepstral euclídea* o cualquier otra como la distancia de *Itakura*.

Un aspecto muy importante de cualquier sistema de cuantificación vectorial es la obtención del libro índice, el espacio vectorial debe ser dividido en sectores los cuales se hallan partiendo de vectores de entrenamiento. Dichos vectores deben representar fielmente el espacio de interés.

El libro índice se obtiene empleando un algoritmo conocido como *LBG*, cuyo nombre se deriva de los creadores *Yoseph Linde, Andrés Buzo y Robert Gray*.

El algoritmo *LBG* debe partir de un libro índice inicial \hat{A}_0 con el cual se compara cada vector del espacio a cuantizar con cada componente del libro índice. Se compone una partición de distorsión mínima $P(\hat{A}_m)$ clasificando cada vector mediante la distancia mínima con los vectores del libro índice. La suma de las distancias se compara con el umbral de distorsión, si resulta mayor se vuelve a calcular una nueva partición hasta que la de distancia total sea inferior al umbral. La recursión se muestra a continuación:

- Sea \hat{A}_0 el libro índice, $x_j : j=0, \dots, n-1$ los vectores de entrenamiento, $m = 0$ y $D-1 = \infty$.
- Dado $\hat{A}_m = \{y_i : i=1, \dots, N\}$ $N =$ niveles de cuantificación, encontrar la partición de distorsión mínima $P(\hat{A}_m) = \{S_i ; i=1, \dots, N\}$. S_i es el i -ésimo sector. x_j pertenecerá a S_i si $d(x_j, y_i) \leq d(x_j, y_l)$ para todo l . Se calcula la distorsión media:

$$D_m = D(\hat{A}_m, P(\hat{A}_m)) = \frac{1}{n} \sum_{j=0}^{n-1} \min_{y \in A_m} d(x_j, y) \quad (2.47)$$

- Si se cumple $\frac{D_{m-1} - D_m}{D_m} \leq \varepsilon$, entonces \hat{A}_m será el libro índice final, donde ε es el umbral de distorsión.
- Si no se cumple la relación anterior entonces se toman los centroides de los sectores hallados como componentes de un libro índice nuevo, al cual se le asigna \hat{A} . Se incrementa m en uno y se vuelve al segundo punto hasta que el error sea menor al umbral de distorsión ε .

Parte del algoritmo comprende la obtención del libro índice inicial, el cual se calcula partiendo del espacio vectorial hallando su centroide. Este centroide será el vector inicial y a partir del cual se obtendrán dos nuevos vectores $y_i + D$ e $y_i - D$ y se procede a calcular el libro índice con el algoritmo LBG para 2 vectores. A continuación se vuelve a fragmentar los centroides en dos y se emplea el algoritmo LBG para 4 vectores. Se procede así sucesivamente hasta el valor N que se requiera. Como se podrá apreciar N debe ser una potencia de dos y dependerá del grado de distorsión, la elección del valor de N . Su valor generalmente puede ser 64 o 128.

2.1.2.2.3.5. Análisis de formantes

Los formantes se definen perceptualmente y son una propiedad física correspondientes a la frecuencias de resonancia del tracto vocal, son muy útiles ya que los dos primeros formantes prácticamente identifican una vocal.

Para obtenerlos se pueden utilizar los métodos: *DFT (Discrete Fourier Transform)* y *LP (Linear Predictive)*.

2.1.2.2.3.5.1. Obtención de picos en el espectro suave

Se procede a examinar los picos en el espectro suave, utilizando las técnicas *cepstrales* para obtener la envoltura (Análisis de *Fourier* / Filtrado Homomórfico). Se toman N picos mayores de menor frecuencia.

2.1.2.2.3.5.2. Obtención de picos en el espectro LP

Utilizando las técnicas *cepstrales* para obtener la envoltura (Análisis predictivo lineal / espectro LP). Se toman N picos mayores de menor frecuencia o también se pueden escoger puntos donde la segunda derivada tenga valor más negativo.

El polinomio LP puede factorizarse y usando fracciones parciales la función de transferencia puede expresarse como:

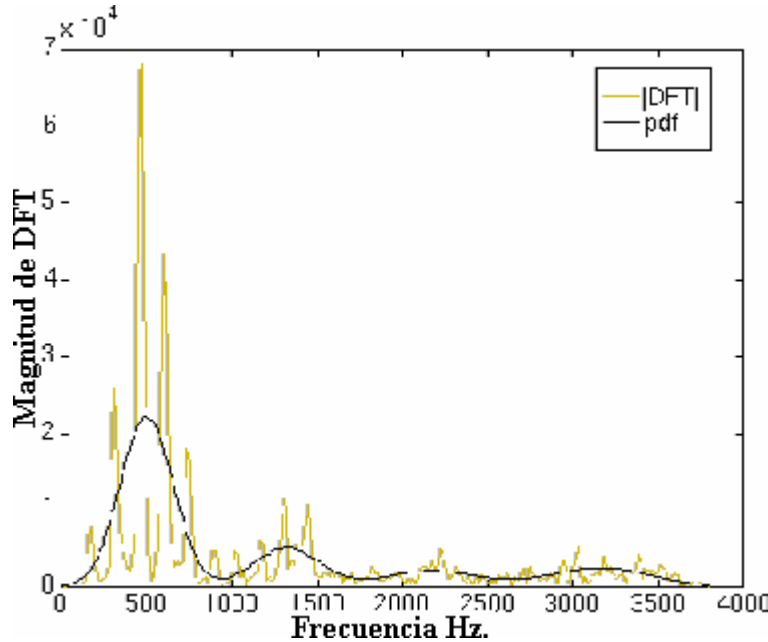
$$\frac{1}{1 + \sum_{k=1}^p a_k z^{-k}} = \frac{1}{\prod_{k=1}^p (1 + z_k z^{-1})} = \sum_{k=1}^p \frac{c_k}{(1 + z_k z^{-1})} \quad (2.48)$$

Este resultado da raíces en pares complejos conjugados (ya que el polinomio es real en a_i) y el ángulo define la frecuencia del formante. Como ocurre con el espectro LP, se pueden tomar todos los picos o aquellos cercanos al círculo unitario.

El comportamiento del formante es aproximadamente *gausiano*, por lo que se puede tratar su densidad del espectro de potencia como una función de densidad de probabilidad. Dicha función puede ser modelada como una mezcla de *gaussianas*, las cuales se muestran en la figura 33, estas encajan en un espectro de magnitud *DFT*.

$$b_j(\underline{x}) = \sum_{m=1}^M c_{jm} \frac{1}{\sqrt{2\pi}\sigma_{jm}} e^{-\frac{1}{2}\left[\frac{\mu - \mu_{jm}}{\sigma_{jm}}\right]^2} \quad (2.49)$$

Figura 33. Mezcla de gaussianas



Fuente: Lawrence Rabiner & Shafer, R., *Digital Processing of Signals*.

2.1.2.2.3.6. Análisis de Voz

Al analizar la voz es necesario conocer el tono y grado de voz.

2.1.2.2.3.6.1. Tono de voz

El tono se define perceptualmente y la correlación acústica es la frecuencia fundamental, por ejemplo en el caso de la voz telefónica esta a veces no aparece, pero incluso en voz con ancho de banda limitado el tono se mantiene claramente.

Las discontinuidades asociadas al uso de ventanas se pueden reducir si el análisis se lleva a cabo sincronizadamente respecto con el tono; en la mayor parte de los casos se aplica a las técnicas basadas en *LP*, ya que es más fácil trabajar con ventanas de longitud arbitraria.

Este análisis se puede realizar por los siguientes métodos:

- Puntos de cruce por cero.
- Picos en la función de autocorrelación.
- Pico en la función de autocorrelación del residuo *LP*.
- Función de la diferencia media de magnitud.
- Pico en el *cepstrum*.

Todos estos métodos producen aproximaciones dudosas, pero el tono posee continuidad y restricciones en la tasa de cambios; los errores más frecuentes son multiplicar o dividir el tono por dos.

2.1.2.2.3.6.2. Grado de voz

Es necesario determinar la existencia de voz, para eso se tiene que definir un método, el cual nos dice si esta ocurriendo una entrada de voz o no. Para esta determinación se usan las siguientes características:

- Tasa de cortes con los ejes.
- r_1 / r_0 , o la fracción de la potencia a bajas frecuencias.
- grado de fuerza del pico de tono.

Esto es difícil, provocando una enorme diferencia en la calidad subjetiva de un codificador *LP*; en la práctica se combinan varios métodos utilizando técnicas estadísticas de reconocimiento de patrones. El método más frecuente es del tipo binario, el cual indica si hay voz o no la hay; en ocasiones un valor real puede resultar útil esto se hace mezclando excitaciones de voz o de no voz con las potencias adecuadas, aplicando un filtro pasa alto a las excitaciones de no voz y un filtro pasa bajo a las excitaciones de voz, el grado de voz se define como el punto límite entre unas y otras.

2.1.2.2.3.7. Características acústicas y lingüísticas

Cuando el sistema debe reconocer frases o habla continua, es necesario acudir a otras fuentes de conocimiento además de las puramente matemáticas. Estas son por lo general reglas de tipo acústico y lingüístico, como se va a ver a continuación.

Antes de entrar en detalle sobre los niveles acústicos y lingüísticos en que se puede dividir un sistema de reconocimiento de voz, es necesario definir vocabulario y gramática.

- El Vocabulario: se conoce como el número de palabras diferentes que debe reconocer el sistema. Mientras mayor es el número de palabras más difícil es el sistema, por dos motivos: primero porque al aumentar el número de palabras es más fácil que aparezcan palabras parecidas entre si, y segundo porque el tiempo de tratamiento aumenta al aumentar el número de palabras a comparar.
- La gramática: es el conjunto de reglas que limita el número de combinaciones permitidas de las palabras del vocabulario. Ayuda a mejorar la tasa de reconocimiento al eliminar ambigüedades y puede ayudar a disminuir la necesidad de cálculo, al limitar el número de palabras en una determinada fase de reconocimiento (perplejidad de la gramática).

A continuación se describen los distintos niveles o módulos básicos en que se podría subdividir el sistema de reconocimiento basado en características acústicas y lingüísticas.

2.1.2.2.3.7.1. Módulo de procesado acústico

En este módulo se extraen, a partir de la forma de onda de la señal de voz, un conjunto de parámetros representativos, que pueden ser desde valores de la frecuencia fundamental, energía, densidad de cruces por cero y posición de los formantes, hasta otros que aporten información útil para comprender el sentido de la frase.

2.1.2.2.3.7.2. Módulo de análisis fonético

A partir de los parámetros obtenidos en el módulo anterior, la representación fonética más probable correspondiente a la señal de voz es el fonema, el cual se define como una unidad del habla y un conjunto de fonemas determina los sonidos con los cuales se pueden construir palabras en un lenguaje determinado. Por otro lado, un alófono es una de las diferentes pronunciaciones para un fonema en particular. El fonema es una abstracción y realmente no puede ser pronunciado excepto en términos de uno de sus alófonos.^[1] Tomando en cuenta lo anterior, se puede decir, que en cualquier lenguaje el número de fonemas es obviamente más pequeño que el número de fonos o alófonos. En el caso del español puede constar de 24 fonemas.

Se tiene la idea ortográfica de que es muy fácil separar la palabra “mesa” en cuatro sonidos: /m/e/s/a/, pero físicamente es un continuo de principio a final. Como consecuencia, en la vocal /e/ hay elementos de la /m/ y de la /s/. Además, por tratar de hablar de prisa no se llegan a efectuar los movimientos completos, y se dejan a medias. Por esas y otras razones, el reconocimiento automático de fonemas no es tan fácil como se quisiera.

2.1.2.2.3.7.2.1. Realización fonética

Los sonidos se producen cuando el aire espirado por los pulmones llega hasta la laringe, donde se encuentran las cuerdas vocales. Estas cuerdas vocales son dos músculos gemelos, elásticos, que vibran cuando el aire espirado pasa por ellas. Es entonces cuando se produce el sonido que llamamos voz. La voz pasa a la cavidad bucal o la cavidad nasal, donde los órganos, principalmente los de la boca, configuran y matizan los diversos sonidos en el habla.

Los fonemas se realizan por medio de sonidos. Los encargados de producir esos sonidos, de realizarlos fonéticamente son los órganos de fonación.

2.1.2.2.3.7.2.2. Clasificación de los fonemas

Los sonidos están divididos en dos grupos principales: consonantes, si la corriente de aire es detenida u obstruida y vocales, si el aire sale libremente.

2.1.2.2.3.7.2.2.1. Fonemas consonánticos

Se producen sonidos consonánticos cuando el aire al salir encuentra un obstáculo, ya sea cerrándole totalmente el paso, o dejándole una estrechez por donde pasa con fricción. Para clasificar a las consonantes se tendrá en cuenta el concepto de articulación. Se entiende por articulación de un sonido la posición adoptada por los órganos de la cavidad bucal en el momento de producirse un sonido.

En la producción y clasificación de las consonantes hay que tener en cuenta los siguientes factores:

- Lugar de articulación.
- Manera de articulación.
- Articulaciones sonoras y sordas.

2.1.2.2.3.7.2.2.1.1. Lugar de articulación

Para el idioma Español hablado en Guatemala hay nueve lugares de articulación para las consonantes:

1. Bilabial: el labio inferior contra o cerca del labio superior. Ejemplo: [m] en más.
2. Labiodental: el labio inferior contra o cerca del borde de los dientes frontales superiores. Ejemplo: [f] en fuente.
3. Dental: posición de la lengua contra el borde o atrás de los dientes superiores frontales. Ejemplo: [d] en dar.
4. Alveolar: posición de la lengua contra o cerca de la zona alveolar. Ejemplo: [l] en luna.
5. Palatal: posición de la lengua contra o cerca del paladar duro. Ejemplo: [c] en chico.
6. Velar: dorso de la lengua contra o cerca del velo. Ejemplo: [k] en calor.
7. Bilabio-velar: el labio inferior cerca del labio superior y al mismo tiempo el dorso de la lengua cerca del velo. Ejemplo: [w] en hueso.
8. Uvular: dorso de la lengua contra la úvula. Ejemplo: [rr] en carro.
9. Glotal: movimiento de cuerdas vocales. Ejemplo: [h] en jardín.

2.1.2.2.3.7.2.2.1.2. Manera de articulación

Independientemente de cuál sea la zona o punto de articulación, los órganos adoptan una postura o una posición al producir los fonemas. A esto se le llama modo de articulación. En el español hay seis maneras de articulación en las consonantes.

1. Oclusivas: en algún punto de la articulación del sonido la corriente de aire queda detenida y después es liberada con una pequeña explosión. Ejemplo: [p] en pasar.
2. Fricativas: la corriente de aire, sin ser detenida, es forzada a través del tracto vocal existiendo un cierre parcial, provocando que el aire salga con turbulencia.
3. Africativas (Oclusiva + Fricativa): la corriente de aire es detenida como en una oclusiva; pero en lugar de ser liberada abruptamente, es liberada con fricción como en un fricativo.
4. Nasaes: el velo es bajado y la corriente de aire pasa a través de la cavidad nasal con gran resonancia. Ejemplo: [m] en más.
5. Semivocales: la cavidad oral es cerrada a la mitad, pero la corriente de aire escapa por ambos lados del lugar de articulación. Ejemplo: [l] en luna.
6. Vibrantes: la posición de la lengua, bajo tensión, conecta el alveolar una vez que la corriente de aire pasa a través del tracto vocal. Ejemplo: [r] en pero y [rr] en perro.

2.1.2.2.3.7.2.2.1.3. Articulaciones sonoras y sordas

Todo fonema, sea cual fuese su punto o manera de articulación, puede producirse con vibraciones de las cuerdas vocales o sin ellas. En la tabla I, se muestra la clasificación de las consonantes de acuerdo a su articulación.

Tabla I. Clasificación de las consonantes de acuerdo a su articulación

Fonemas sonoros	Fonemas sordos
s, j, f, ch, p, t, k	b, d, g, m, n, ñ, l, r, rr, y

2.1.2.2.3.7.2.2.2. Fonemas vocálicos

Con las vocales, la posición de la lengua y la forma en que se posiciona la boca determinan su timbre. Debido a que el aire no es bloqueado o detenido, no hay contacto entre los articuladores superiores e inferiores. En su lugar, la posición de la lengua es significativa debido a que cambia el tamaño y forma de la cavidad oral.

Acústicamente, cuando se pronuncia una vocal hay unas bandas de frecuencia que se amplifican más según donde se ponga la lengua; una A, por ejemplo, se articula de forma más estrecha en la zona de la faringe y más ancha en la zona de la boca, lo cual lleva a la amplificación de un tramo de frecuencia que está alrededor de los 700Hz, y otro en los 1300Hz., que se conocen como formantes.

Hay tres posiciones verticales de la lengua, las cuales determinan el grado de abertura de la cavidad oral.

- Alto, cuando la lengua está cerca del paladar;
- Medio, cuando la lengua se encuentra a la mitad de la boca; y
- Bajo cuando la lengua está cerca del velo.

También hay tres posiciones de la lengua de frente hacia atrás.

- Anterior, cuando la parte alta de la lengua está muy cerca del alveolar;
- Central, cuando la parte más alta de la lengua está en el centro de la cavidad oral; y
- Posterior, donde la parte más alta de la lengua, el dorso, está muy cerca del velo.

Por esto se puede clasificar a las vocales en un cuadro bidimensional. En el español hay cinco posiciones principales, y se muestran en la tabla II.

Tabla II. Clasificación de los fonemas vocálicos

	Anterior	Central	Posterior
Alto	i		U
Medio	e		O
Bajo		A	

Las vocales son de mayor duración que las consonantes y están bien definidas espectralmente. Por esto las vocales usualmente son fáciles de reconocer y contribuyen significativamente al proceso de reconocimiento de voz. De acuerdo a la manera en que el tracto vocal se configure se determinan las frecuencias de resonancia del mismo (las formantes) y por ello el sonido se produce. Las vocales se identifican por sus formantes, las cuales son muy fuertes durante todo el fonema. Cada vocal tiene su propia configuración de formantes, esto es muy útil en el reconocimiento de voz, ya que las frecuencias al ser medidas pueden señalar en determinado momento de que tipo de fonema se está hablando.

2.1.2.2.3.7.3. Módulo de análisis fonológico

La fonología estudia la estructura o función de los sonidos dentro del lenguaje. Permite la adaptación de los datos obtenidos en los niveles anteriores a una determinada lengua. Es necesario definir cuales son las unidades fonológicas que van a ser reconocidas en el sistema ya que pueden ser alófonos, fonemas, difonemas, silabas o palabras.

Las reglas fonológicas aportan información de cómo varia la pronunciación de los fonemas, dependiendo del contexto. Para la realización de estas reglas, y un ajuste correcto de los parámetros, es necesario tener en cuenta la prosodia de la frase. Los valores de los parámetros obtenidos en el análisis acústico fonético ayudarán a determinar las silabas tónicas o átonas, si la frase es enunciativa o interrogativa, etc.

2.1.2.2.3.7.4. Módulo de análisis morfológico

Es importante conocer, para cada lenguaje, las reglas de formación de las palabras a partir de los morfemas elementales. Esta es una de las facetas que estudia la morfología. Por ejemplo, hay combinaciones de sonidos o de letras que están permitidas en unos lenguajes y en otros no, por lo que es necesario conocer esas reglas de formación específicas. También hay reglas de formación de palabras a base de utilizar prefijos o sufijos. La disponibilidad de estas reglas, o incluso de un diccionario, ayudará a la determinación de palabras dentro de la cadena de unidades fonéticas que ha salido del módulo acústico fonético.

Las reglas morfológicas ayudan también a la categorización gramatical de las palabras, lo que puede ser usado por otros módulos.

2.1.2.2.3.7.5. Módulo de análisis sintáctico

La sintaxis estudia como combinar las palabras para construir frases de forma correcta en un determinado lenguaje. En cada idioma, existe una serie de reglas de concatenación de palabras, constituyendo la gramática del lenguaje.

Este módulo ayuda en decidir una secuencia lógica de palabras, y en caso de dudas entre los módulos anteriores, elegirá aquella que sintácticamente sea correcta.

Si un sistema debe reconocer una frase como “Los caballos corren por el campo”, ha podido tener dudas si “caballos” va en singular o plural, dado que la terminación de la palabra es difícil de reconocer por el sistema y quizás el locutor no la ha dicho bien. Sin embargo, si ha reconocido con bastante seguridad el artículo previo “Los”, estará totalmente seguro que la palabra siguiente es “caballos”.

2.1.2.2.3.7.6. Módulo de análisis semántico

El conocimiento semántico está relacionado con como se encadenan las palabras para dar significado a una frase. Toma como partida el significado individual de la palabras, para deducir si una frase determinada tiene o no significado.

En este módulo y los siguientes es donde empiezan los problemas de reconocimiento, ya que no se dispone aún de una forma eficiente de introducir este conocimiento en las máquinas.

Piénsese que hay muchas frases o palabras que tienen significado en un contexto y no lo tienen en otro, o lo tienen pero diferente entonces ¿Cómo puede distinguir una máquina una opción de la otra?

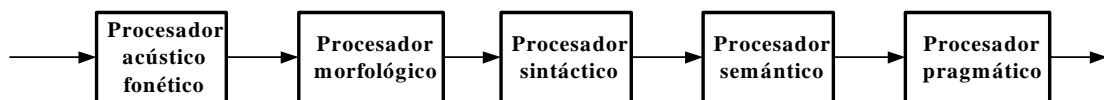
2.1.2.2.3.7.7. Módulo de análisis pragmático

El pragmatismo está relacionado con el contexto donde se están desarrollando las ideas. Está muy relacionado con el módulo de análisis semántico.

Puede darse el caso, en que existan frases, con un contenido sintáctica y semánticamente correctos, que tengan un contenido pragmático incorrecto; y sintácticamente mal formadas, que tengan un contenido pragmático correcto. Esto es necesario tenerlo en cuenta, en reconocedores de voz, ya que sucede más veces en el lenguaje hablado que en el escrito.

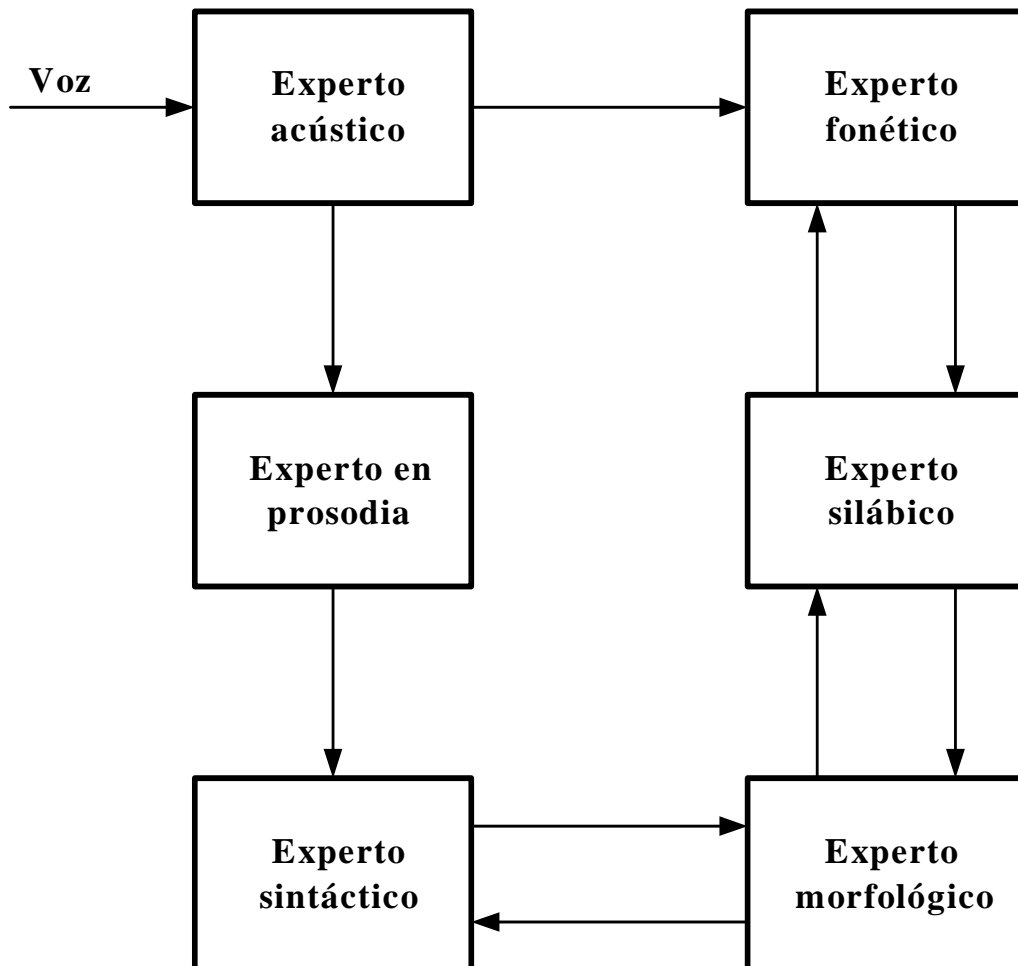
Luego de haber revisado los diferentes módulos que integran el reconocimiento acústico fonético, se pueden organizar en forma jerárquica bloques concatenados mostrados en la figura 34; El procesador acústico fonético analizando la forma de onda produce varias secuencias de fonemas, cada una de ellas correspondiente con un grado de probabilidad determinado a la transcripción fonética de la señal de entrada al sistema. El procesador morfológico genera una red con las palabras más probables, y esa red pasa al procesador sintáctico, que la depura y recorta, dejando solo las secuencias de palabras gramaticalmente correctas. El procesado semántico sigue limpiando esa red, eliminando las frases sin sentido. Por último, y en el supuesto caso de que quede más de un candidato, será el procesador pragmático quien tome la última decisión.

Figura 34. Organización jerárquica de los módulos acústicos y lingüísticos



Este sistema de organización no tiene ningún tipo de retroalimentación que pueda aumentar la eficiencia del sistema, se puede pensar en aprovechar mejor el sistema teniendo un tipo de flujo inverso entre algunos módulos. La figura 35 muestra una estructura de interconexión que refleja esta idea.

Figura 35. Estructura de interconexión para un flujo inverso



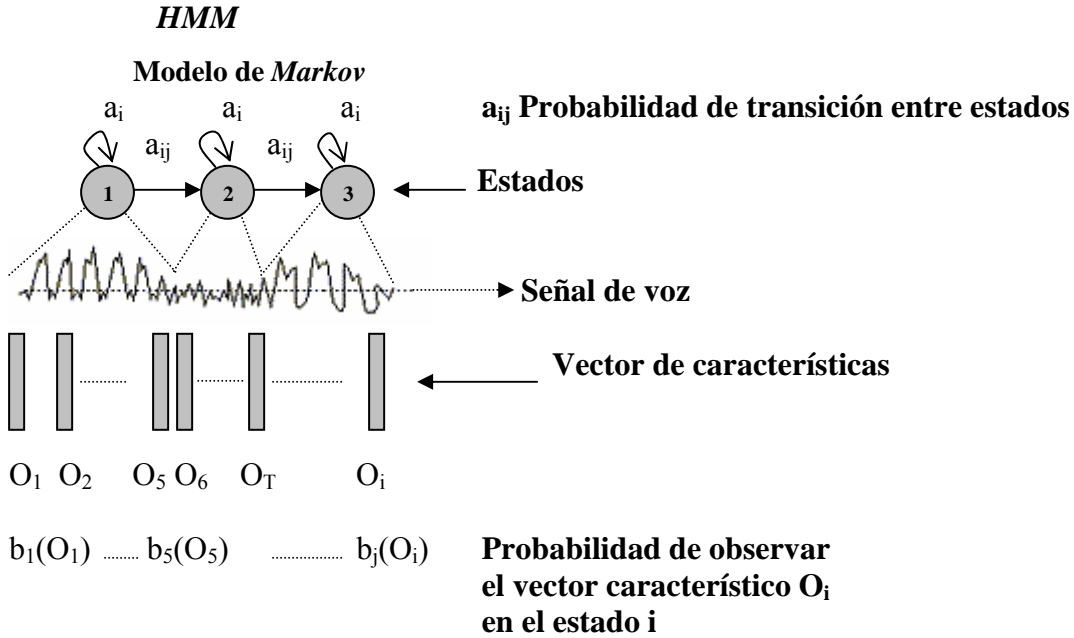
2.1.2.2.3.8. Procesos discretos de *Markov*

En los sistemas de reconocimiento de voz, es necesario adoptar un modelo estadístico para cada una de las palabras del vocabulario y este modelo lo compensa de buena forma los modelos ocultos de *Markov* (*HMM Hidden Markov Models*). Ya que requieren menos memoria física y ofrecen un mejor tiempo de respuesta, y como contrapartida tienen una fase de entrenamiento mucho más lenta y costosa.

Sin pretender entrar en detalle y en forma simplificada, se puede representar un HMM como un autómata de estados finitos en que el siguiente estado depende únicamente del estado actual, y asociado a cada transición entre estados se produce un vector de características. Se puede decir que un modelo de *Markov* lleva asociado dos procesos: uno oculto (no observable directamente) correspondiente a las transiciones entre estados, y otro observable (y directamente relacionado con el primero), cuyas realizaciones son los vectores de características que se producen desde cada estado y que forman la plantilla a reconocer.

En los *HMM* en reconocimiento de voz, se representa cada palabra del vocabulario con un modelo generativo (se calcula en la fase de entrenamiento) y posteriormente, se calcula la probabilidad de que la palabra haya sido producida por cada uno de los modelos de la base de datos del reconocedor. Para ello, se asume que durante la pronunciación de una palabra, el modelo realiza una segmentación de la misma asignando un tramo de la unidad a cada estado, como se ilustra en la figura 36.

Figura 36. Esquema de funcionamiento del proceso de clasificación basado en



Fuente: X. Huang, Y. ARIKI and M. Jack, *Hidden Markov Models for Speech Recognition*.

Adicionalmente, a cada vector de características asociado a un estado determinado, el modelo de *Markov* le asigna una probabilidad; la probabilidad total obtenida para todos los vectores de características será la puntuación que se utilice para realizar el proceso de clasificación. Así vista la generación de la palabra, las características espectrales de cada fragmento de señal dependen del estado activo en cada instante, y la evolución del espectro de la señal durante la pronunciación de una palabra depende de la ley de transición entre estados.

En cuanto a la generación de puntos de la plantilla, en estos modelos se asume que el primer vector de características se produce desde el primer estado, y el último se emite desde el último estado. Recuérdese que la secuencia de estados es la parte oculta del modelo: se conocen los vectores de características, pero no desde qué estado se han producido.

2.1.2.2.3.8.1. Características de un *HMM*

Un modelo M viene determinado por siguientes parámetros:

- N , número de estados del modelo.
- Matriz de transiciones, de dimensión $(N \times N)$. Define la estructura del modelo: cada uno de sus elementos, a_{ij} , define la probabilidad de pasar del estado i al estado j . Normalmente A será bidiagonal o triagonal, significando que desde cada estado se pueden producir dos o tres tipos distintos de transición.
- Conjunto de funciones de densidad de probabilidad que modelan estadísticamente las observaciones producidas desde cada estado. Habrá pues tantas funciones de densidad de probabilidades como estados.
- P vector de dimensión N . Cada uno de sus elementos, P_i indica la probabilidad de encontrarse inicialmente en el estado i . Para modelos de izquierda a derecha, $P_1 = 1$ y $P_j = 0$ para los demás estados.

Como en el caso *DTW*, la señal de voz viene representada por una plantilla o secuencia de vectores de características $O = \{O_1, O_2, \dots, O_T\}$, donde cada O_j es un conjunto de parámetros (coeficientes *LPC*, *Cepstrum*, ...) que caracteriza la señal de voz en una ventana de tiempo centrada en $t = i$, y T es el número de total de puntos de la plantilla. Los modelos *HMM* basados en este tipo de observaciones se llaman *HMM* continuos, y serán un conjunto de funciones de densidad de probabilidad continuas. Si, para simplificar las cosas, se hace pasar esa secuencia de observaciones $O = \{O_1, O_2, \dots, O_T\}$ a través de un cuantificador vectorial (en que cada vector de parámetros O_i es codificado como un número entero), la señal de voz quedará representada por una secuencia de centroides del cuantificador. Los *HMM's* que trabajan sobre este tipo de datos se conocen como *HMM* discretos y son una matriz con tantas filas como estados tenga el modelo y tantas columnas como centroides tenga el codificador vectorial, en

que cada elemento b_{jk} es la probabilidad de que, estando en el estado i , se produzca el centroide k .

2.1.2.2.3.8.2. Reconocedor de palabras basado en *HMM's*

Una vez definido lo que es un modelo de *Markov*, se describe a continuación como se aplica a un problema real: el de reconocimiento de palabras (la metodología a usar sería la misma si se utilizasen otras unidades acústicas: fonemas, demisilabas, frases cortas, etc.).

El reconocedor dispondrá de un modelo por cada palabra del vocabulario de reconocimiento, y la estructura de esos modelos se define en la fase de diseño: el número de estados (N) se elige "a priori" según la complejidad que se pueda permitir y la calidad deseada. Valores típicos de N son entre 5 y 15 estados. Lo mismo ocurre con el tipo de transiciones: la matriz A tendrá sólo ciertas componentes distintas de cero, y su número es un parámetro de diseño. El tipo de funciones estadísticas que se utilizarán para modelar las probabilidades de observación de los puntos de la plantilla desde cada estado, también se fija antes de entrar en la fase de entrenamiento de los modelos. Suelen ser *gaussianas* multivariadas, combinaciones lineales de *gaussianas* multivariadas, funciones *gamma*, etc.

Una vez fija la estructura de los modelos se lanza la fase de entrenamiento^[2], con el fin de calcular los valores óptimos de todos los parámetros que se han mencionado. Para ello, se usa un cierto número de repeticiones de cada palabra del vocabulario, que depende del tipo de reconocedor que se quiera construir (dependiente o independiente del locutor), de las prestaciones esperadas del sistema y del tipo de unidades que formen el vocabulario. Se puede decir que ese número de repeticiones varía entre 4 o 5 y unos cuantos centenares, lo que da idea del volumen de datos y de cálculos necesarios. Del

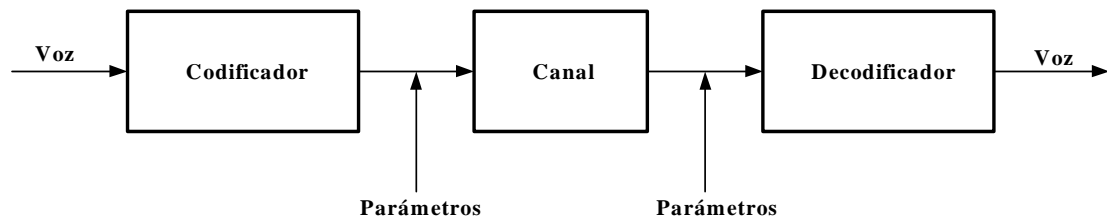
análisis de todas esas repeticiones saldrá el conjunto de parámetros que define cada modelo de *Markov*, y que formara la base de datos del reconocedor.

2.1.2.2.3.9. Codificación de Voz

La codificación es un campo básico dentro de las tecnologías del habla, desde las técnicas más sencillas que permiten el almacenamiento y acceso de los ordenadores a la señal de voz, hasta las más complejas que no sólo permiten un importante ahorro de recursos sino que también ofrecen representaciones y modelos útiles a todas las tecnologías del habla.

En la figura 37 puede verse el esquema básico de un sistema de codificación-decodificación.

Figura 37. Modelo de codificación



La señal de voz es una señal limitada en banda, aproximadamente entre 20 Hz y 20 KHz. Sin embargo, la mayor parte de la energía se concentra por debajo de 2 KHz, y se asegura casi toda la inteligibilidad con un ancho de banda entre 300 y 3400 Hz (salvo algunas consonantes fricativas, con rasgos distintivos hasta 667 KHz).

El margen dinámico de la voz es muy amplio, tanto entre locutores (unos 20 dB) como para un mismo locutor (hasta 40 dB de diferencia entre zonas sonoras y sordas). Su distribución de probabilidad se puede aproximar, para segmentos lo bastante largos, mediante una función *laplaciana* o *gamma*.

Al ser una señal limitada en banda, se puede muestrear y recuperar sin error si se muestrea por encima de la frecuencia de *Nyquist* siendo:

- f_s = frecuencia de muestreo
- W = ancho de banda

Valores típicos de frecuencia de muestreo son 8 KHz para aplicaciones telefónicas, y 44 KHz para codificación de muy alta calidad (*Compact Disc*).

Puesto que las muestras obtenidas tienen valores continuos dentro del margen dinámico, su entropía es infinita, y es imposible codificarlas y recuperarlas sin error. Sin embargo, la señal de voz es muy redundante, y el receptor final es muy robusto (el oído y el cerebro humano), y se pueden aplicar estrategias de codificación que permiten almacenar y transmitir la voz con un régimen binario moderado y recuperarla con una degradación aceptable de la calidad.

$$R = Mf_s \quad (2.50)$$

Siendo R = régimen binario

M = número medio de *bits* para codificar un símbolo del alfabeto

f_s = frecuencia de muestreo

Ya se ha visto que para poder recuperar la forma de onda, es necesario muestrear al menos a la frecuencia de *Nyquist*. Para ofrecer una señal con 4 KHz de ancho de banda (calidad telefónica), es necesario muestrear al menos a 8 KHz. Para reducir el régimen binario, sólo se puede actuar sobre el número de *bits* con el que se codifica cada muestra (número de estados permitidos). El estándar de calidad telefónica es de 12 *bits* (4096 estados) de codificación lineal, que se pueden reducir a 8 con técnicas de compresión / expansión (*PCM* a 64 *kbits*). Otras técnicas explotan la gran redundancia de la señal de voz para reducir a 4 el número de bits por muestra, mediante adaptación del tamaño del escalón de cuantificación y predicción del valor de las muestras a partir de los valores anteriores (*ADPCM* a 32 *Kbits*).

Sin embargo, si se renuncia a recuperar la forma de onda, se puede burlar el límite de la frecuencia de *Nyquist* en la transmisión, y reducir más el régimen binario.

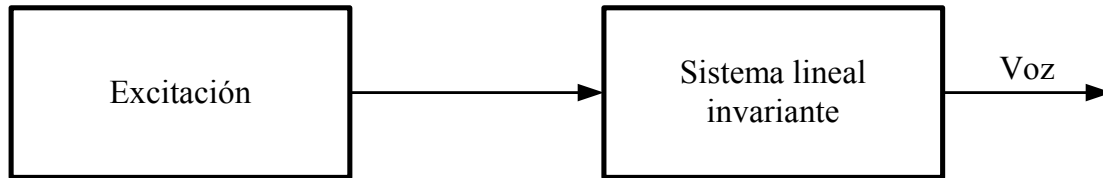
A título orientativo, cuando se lee un texto a una velocidad normal, unas 200 palabras/minuto, la información presente en el texto se transmite a una velocidad de unos 75 bit/seg. Sin embargo, hay que recordar que en la señal de voz se transmite más información que la que figuraba en el texto (emociones, entonación, énfasis, identidad del locutor, etc.).

Ya se ha señalado que la señal de voz, aunque muy distinta para cada uno de los sonidos, mantiene sus características durante períodos bastante grandes, correspondientes aproximadamente a los alófonos que se están pronunciando. Se puede intentar encontrar algunas de las características que definen estos sonidos, y transmitir las sólo al ritmo que estas cambian, cada pocas decenas de milisegundos.

Para definir y extraer esas características es preciso desarrollar un modelo de la señal. El modelo más empleado y aceptado está muy relacionado con la descripción que se ha hecho de la generación de la voz.

Se divide la señal de voz en intervalos, en los que mantiene sus características, y se supone que las propiedades de la señal en cada intervalo se extienden indefinidamente en el tiempo. En cada uno de estos intervalos se separan e independizan las características de la excitación y las del tracto vocal y la radiación, esto se muestra en la figura 38.

Figura 38. Modelo de producción de voz



Los métodos de codificación que utilizan propiedades de la señal de voz en vez de intentar regenerar la forma de onda, suelen denominarse genéricamente *Vocoder's*. Los parámetros con los que trabajan suelen estar relacionados con una descripción del espectro, aunque también hay esquemas híbridos que manejan algún tipo de señal residual, difícil de caracterizar, y que mejora notablemente la calidad del sistema. A continuación se describen algunos tipos de *Vocoder's*.

2.1.2.2.3.9.1. *Vocoder* de canal

Se pasa la señal por un conjunto de filtros estrechos, de manera que como mucho aparezca a la salida de cada uno un armónico. Se codifica y se transmite la amplitud de los armónicos, y la separación entre los mismos. En el *Vocoder* de fase, se transmite la información de fase y de amplitud de cada una de las señales de salida.

2.1.2.2.3.9.2. *Vocoder* de formantes

Se identifican y transmiten las frecuencias de los formantes y los anchos de banda, en lugar de la energía en un conjunto de bandas predefinidas.

2.1.2.2.3.9.3. *Vocoder* articulatorio

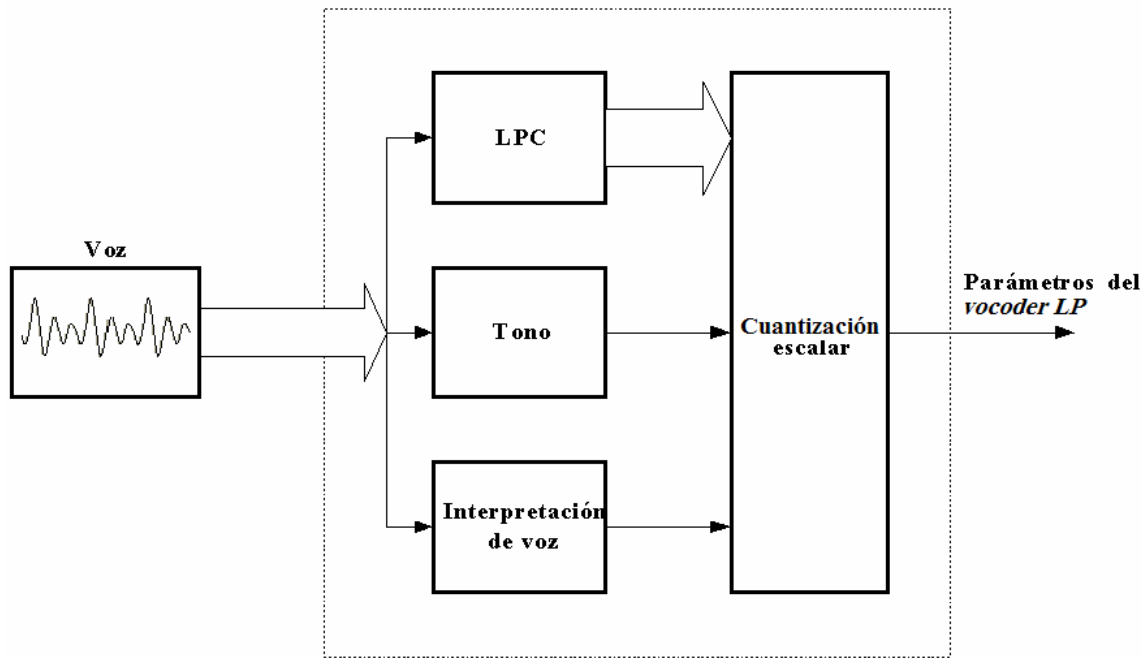
Se estiman y transmiten parámetros relacionados con el tracto vocal, como áreas de secciones del tracto vocal, coeficientes de acoplamiento de la cavidad nasal, y presión subglotal del aire y tensión de las cuerdas vocales.

2.1.2.2.3.9.4. *Vocoder* homomórfico

Según el modelo, se puede representar la voz como convolución de una señal de excitación con la respuesta impulsiva del tracto vocal. Si se pudieran recuperar estas dos componentes de la señal de voz, podrían transmitirse de manera eficiente. Esto es precisamente lo que realiza el filtrado homomórfico.

2.1.2.2.3.9.5. *Vocoder* LPC

En casi todos los *Vocoder*'s anteriores, la información transmitida sobre la función de transferencia del tracto vocal se obtiene muestreando alguna señal directamente relacionada con el espectro (el propio espectro, el *cepstrum*, ...). En los *Vocoder*'s *LPC* se caracteriza la función de transferencia por un modelo y unos coeficientes, que son los que se actualizan y transmiten en cada intervalo. El modelo es un filtro todo polos, y el ajuste de los coeficientes de filtro se hace según las técnicas de predicción lineal. De esta manera, el filtro captura la información sobre la envolvente espectral (formantes), y la excitación debería representar la estructura final del espectro. La figura muestra una forma general de este *Vocoder*.

Figura 39. Vocoder LPC

En el caso más sencillo, el modelo de excitación es una señal periódica de pulsos (caracterizada por el período de repetición de los pulsos), o ruido (normalmente ruido blanco *gaussiano*). Esta descripción es bastante pobre, y limita la calidad de estos sistemas. Surgen por tanto sistemas híbridos para representar la excitación: una versión filtrada de la voz original (*VEV*), una versión filtrada del residuo del análisis *LPC* (*RELP*), una señal de excitación seleccionada entre un amplio registro ya fijado (*CELP*), o una secuencia de pulsos, cuya amplitud y situación se ajusta para minimizar el error (definido sobre el espectro) entre la señal original y la reconstruida (*MPLPC*).

3. SISTEMA IRIDIUM

3.1. Constelaciones de satélites de orbita baja (*LEO*)

Los sistemas de comunicaciones vía satélite se clasifican según su órbita alrededor de la tierra en varios tipos: Satélites geoestacionarios (*GEO*), Satélites de órbita media (*MEO*), Satélites de órbita baja (*LEO*). Estas órbitas se muestran en la figura 40.

3.1.1. Satélites geoestacionarios (*GEO*)

Están en una órbita circular ecuatorial de altitud 33,786 Km. de altitud. Con pocos (basta con 3), sería suficiente para dar cobertura global a la tierra.

3.1.2. Satélites de órbita media (*MEO*)

Estos sistemas tienen típicamente una altitud de 9,600 a 14,500 Km., el período de su órbita es de varias horas y los ángulos de inclinación de las órbitas son 45°- 90°. Si nos fijamos en un punto sobre la superficie terrestre, el tiempo que el satélite permanece sobre el es del orden de varias horas, por lo que el traspaso entre células es menos frecuente y el retardo de propagación es mayor, teniendo en consideración esto son necesarios de 10 a 15 satélites para abarcar toda la tierra.

3.1.3. Satélites de órbita baja (*LEO*)

Son órbitas típicamente circulares, cuya altitud varía entre 725 - 1,450 Km. (450 - 900 millas), el período de su órbita varía entre 90 minutos y 2 horas, con ángulos de inclinación entre 45° y 90°.

En un punto concreto de la superficie terrestre, el tiempo en que el satélite permanece sobre él es de alrededor de 15 minutos (este es el tiempo en que un móvil es servido por su satélite, tras el cual otro satélite pasa a prestarle servicio), siendo necesarios para esto más de unos 40 satélites para la cobertura total.

Los sistemas *LEO* son similares a las redes celulares. La diferencia es que el tamaño del radio celular es mayor y que las células se mueven. Este movimiento es el que determina el intervalo de traspaso de una celda a otra.

Ejemplos de sistemas *LEO* son *GlobalStar* e *Iridium* que es el que se tratará más adelante.

Para el futuro se planea usar este tipo de sistema de baja altitud en los diferentes servicios teniendo éste sus ventajas y desventajas.

3.1.3.1. Ventajas

- Las pérdidas de propagación van a ser menores, usando entonces menores antenas y potencias de transmisión.
- Menores retardos de propagación.
- Actualmente este tipo de órbitas no están congestionadas.

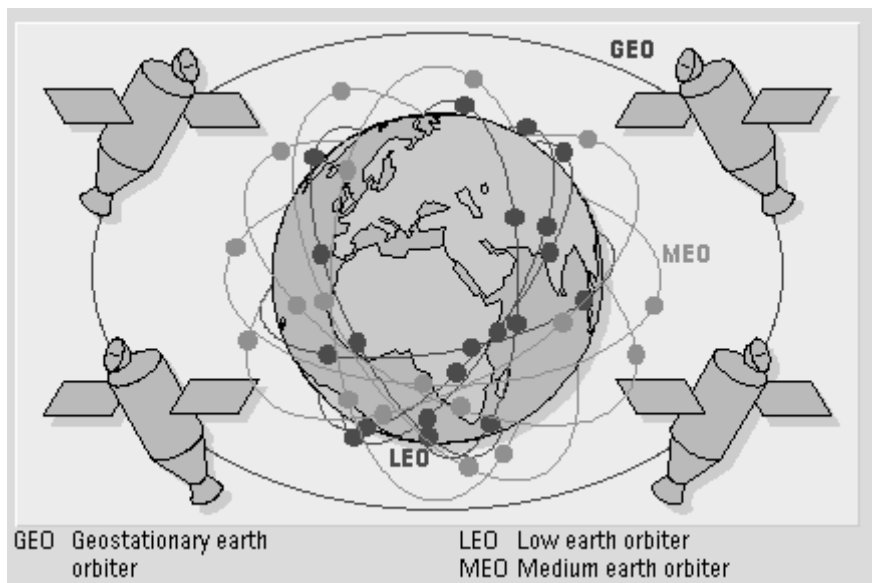
- La cobertura no será pobre para altas latitudes y el ángulo de elevación con que se va a ver el satélite desde estos puntos será alto, lo cual con llevará una mayor calidad de comunicación. Esto es importante sobre todo en entornos urbanos.
- Al haber más satélites, habrán más canales de comunicación disponibles y reutilizables, y por tanto, mayor número de usuarios.

3.1.3.2. Desventajas

- El diseño del sistema será más complejo porque:
 - ◆ el número de satélites necesarios para una cobertura global es mayor,
 - ◆ se deberá emplazar la constelación de satélites con mayor o menor proximidad a los anillos de *Van Allen*. Estos anillos son 2 toroides cuyo eje es el eje geomagnético y están poblados de partículas que dañan el satélite. En el caso de *Iridium*, la constelación va a estar por debajo del primer anillo de *Van Allen*.
- El mantenimiento del sistema va a ser mayor pues los satélites van a verse más afectados por la atmósfera y los anillos de *Van Allen*.
- La velocidad de desplazamiento relativo entre emisor y receptor en un escenario como éste va a ser muy elevada, por lo que:
 - ◆ el tema del *handover* (o *handoff*) va tener que ser tratado con ahínco. Las celdas de cobertura de un satélite van a desplazarse a gran velocidad sobre la superficie de la tierra, ocasionando sesiones de conexiones entre satélites que deberán ser llevadas a cabo de manera eficiente con objeto de que el sistema resultante tenga buenas prestaciones.

- ◆ El efecto *doppler* va a presentar otro problema a resolver. El tamaño de la velocidad de desplazamiento relativo provocará altas derivas en las frecuencias de trabajo.
- Mayor complicación para el control y posicionamiento de los satélites.

Figura 40. Órbitas de la Tierra



Fuente: *Iridium*

3.2. Funcionamiento del sistema *Iridium*

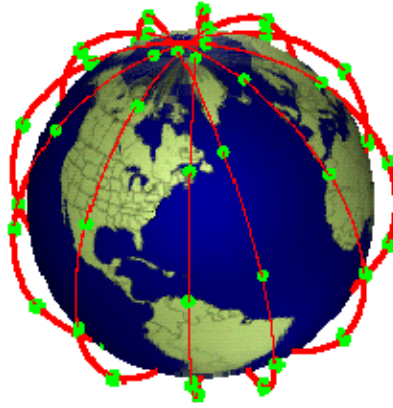
Es un sistema global de comunicaciones móviles que utiliza la estructura de red celular cuyas estaciones base se encuentran en el espacio en forma de 66 satélites pequeños e inteligentes en órbita baja (*LEO*) a una altura de 780Km. Esta altura orbital fue elegida para asegurarse que cualquier punto de la superficie terrestre quedaba a la vista de uno o más satélites, aunque inicialmente se pensó en un sistema con 77 satélites y de ahí su nombre *Iridium*, que se debe al elemento Iridio que tiene 77 electrones en torno a su núcleo atómico.

3.2.1. Componentes del sistema

3.2.1.1. Satélites

3.2.1.1.1. Órbitas

La constelación incluye 6 planos orbitales separados 30° con 11 satélites cada uno describiendo una órbita polar circular, como se muestra en la figura 41. Todos los satélites viajan en la misma dirección; es decir, los satélites de los 6 planos rotan en el mismo sentido en un lado de la tierra hacia el norte cruzan el polo y siguen viajando en la misma dirección en el otro lado de la tierra hacia el sur. Los 11 satélites de cada plano están equiespaciados en su órbita. Además los satélites de los planos 1, 3 y 5 recorrerán casi en fase sus órbitas; igualmente ocurre con los satélites de los planos 2, 4 y 6 que están desfasados con los planos 1, 2 y 3, de tal manera que dado un satélite del plano 'x' los satélites más próximos del plano 'x+1' a este último son 2 esto equivale a decir que el desfase es de 180° ; para poder prevenir una colisión de satélites en los polos, se permite cierta tolerancia cuando se dice que los satélites de los planos pares giran en fase, no es absolutamente cierto, además la inclinación orbital es de 86.4° en vez de 90° , con un período orbital de 100 minutos y 28 segundos a una velocidad de 28,000 Km./h. Los enlaces entre satélites se llevan a cabo en la banda Ka (23.18 a 23.38 Ghz.) proporcionando comunicaciones fiables y de alta velocidad entre los satélites vecinos y permitiendo un direccionamiento óptimo de las llamadas.

Figura 41. Sistema de satélites

Fuente: *Iridium*

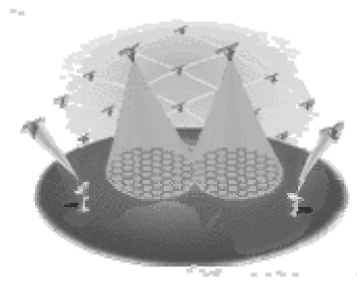
3.2.1.1.2. Celdas

El área de cobertura de un satélite se divide en celdas cada una de ellas iluminada por una antena multihaz (*multispot*). Cada satélite tiene una proyección de 4,700 Km. de diámetro de cobertura, la cual esta subdividida en 48 celdas originadas por 3 antenas de 16 *spot beams* proyectadas sobre la superficie terrestre pudiendo reutilizar frecuencias en diferentes celdas no adyacentes según un patrón que ha sido facilitado y que es conforme al *Seamless handover*, esto se muestra en la figura 42. La huella que deja cada uno de estos haces sobre la superficie terrestre tiene un diámetro aproximado de 660 Km. En el modelo de ensayo de *Iridium* todas las celdas tienen la misma forma en este caso hexagonal y el margen de enlace es de 16 dB. Los enlaces a tierra se llevan a cabo en banda L (1616 a 1626.5 Mhz, enlace de usuarios) y en banda Ka (para enlaces de pasarela y satélite; 19.4 a 19.6 Ghz para enlaces descendentes, y 29.1 a 29.3 Ghz para enlaces ascendentes).

La constelación de satélites y su proyección de celdas es algo análoga a la de un sistema telefónico celular, en la figura 43, se muestra el mapa de cobertura. En el caso de telefonía celular un juego de celdas estáticas da cobertura a un gran número de usuarios móviles; en el caso de *Iridium*, los usuarios se mueven a mucha menor velocidad que las celdas, ya que la velocidad de los satélites es altísima, en concreto, se ha calculado que cada 0.9 segundos el satélite que cubre varía, de manera que el usuario parece inmóvil y las celdas son las que se mueven. En esta situación, esto representa una ventaja: los *handoffs* son deterministas, en general no se producen con igual probabilidad hacia una de las 6 celdas adyacentes, sino que en muchos casos, el *handoff* se produce hacia la celda superior o la inferior depende de hacia donde esté girando el satélite, de norte a sur o viceversa.

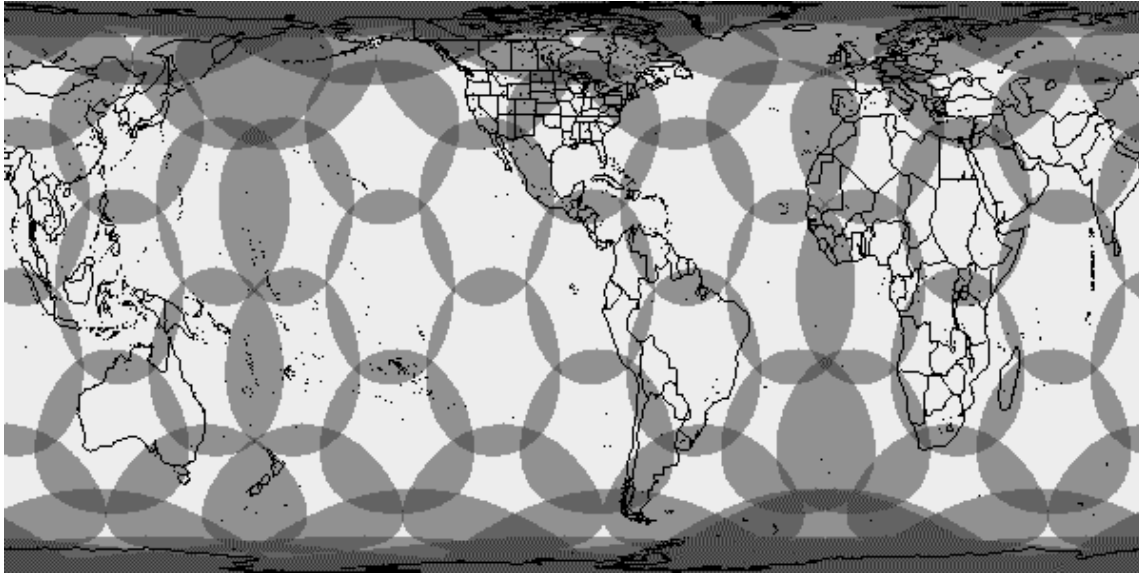
Cabe destacar que el enfoque de haces tan ajustado permite en recepción minimizar el eco y aumentar la calidad de las conexiones. Además permite el uso en recepción de antenas lo suficientemente pequeñas como para ser integradas en el terminal de bolsillo.

Figura 42. Haces de cada satélite *Iridium*



Fuente: *Iridium*

Figura 43. Mapa de cobertura del sistema *Iridium*



Fuente: *Iridium*

3.2.1.1.3 Bus

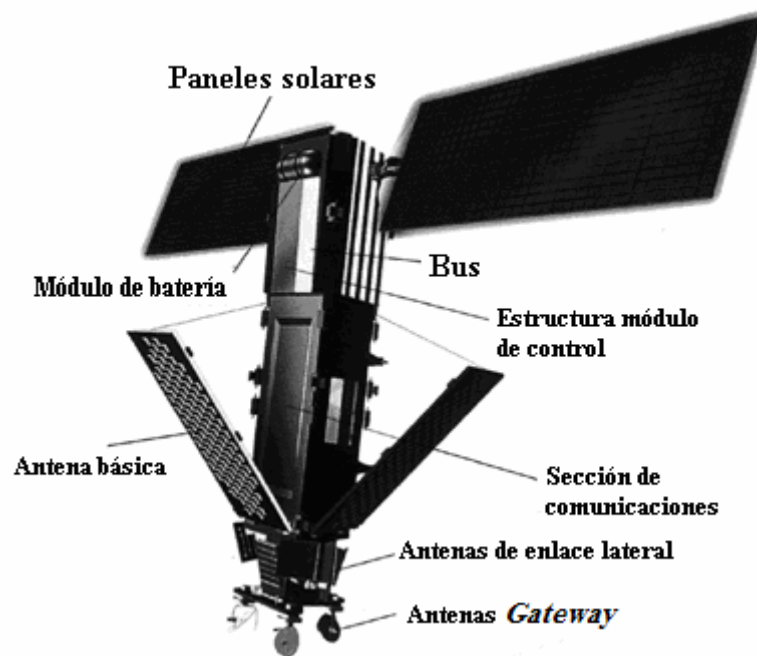
El bus es básicamente la nave, quitando la carga útil, es decir, hace de plataforma soporte para la carga útil electrónica y las antenas. Viene en tres piezas:

- Unidad modular del bus que será integrada con la carga útil de comunicaciones.
- Las placas solares.
- Módulo batería, conjunto radiador / acumulador.

3.2.1.1.4. Antenas

Los satélites *Iridium* darán una vuelta a la tierra pasando por los polos cada 100 minutos. Tendrán varias antenas que les permitirán comunicarse con los otros satélites visibles de la constelación, y con estaciones terrenas (puertas de enlace); la estructura de los satélites se muestra en la figura 44. Cada satélite tiene 13 antenas en total, distribuidas de la siguiente manera:

Figura 44. Estructura de los satélites *Iridium*



Fuente: *Iridium*

3.2.1.1.4.1. Antenas básicas de *phased array*

Estas sirven para la comunicación del satélite con el usuario, es una de las antenas más sofisticadas hoy en día para comunicaciones vía satélite. Cada satélite lleva 3 antenas de este tipo, cada una esta constituida por 106 ranuras y proyecta 16 haces de los 48 de cobertura que hay por satélite estos haces son fijos respecto del satélite. La antena está hecha en una placa de aluminio de 86 cm * 188 cm * 4 cm. Los circuitos electrónicos son de ArGa y van montados detrás de la placa que en total tiene un peso de 36 Kg. La empresa desarrolladora de esta antena fue *Raytheon*.

3.2.1.1.4.2. Antena de enlace secundario

Es la más pequeña y simple, está al final del satélite, y sirve para orientarlo después del despliegue o cuando se desorienta. Cada satélite lleva 2 antenas de este tipo. La empresa desarrolladora de esta antena fue “COM DEV”.

3.2.1.1.4.3. Antena de enlace lateral

Permiten las comunicaciones entre satélites. La antena de enlace lateral fija se usa para la comunicación en dirección Norte-Sur, la antena de enlace lateral móvil para la Este-Oeste. Están formadas por un arreglo de láminas planas, y cada satélite lleva 2 de cada tipo anteriormente mencionadas. La empresa desarrolladora de estas antenas es “COM DEV”.

3.2.1.1.4.4. Antena móvil de pasarela o de enlace vertical

Proporcionan una comunicación continua con una estación terrestre (puerta de enlace). A diferencia de las antenas de enlace lateral que consisten en un arreglo de láminas planas, la antena de enlace vertical (antena de puerta de enlace) es de diseño *full duplex* y combina el uso de reflectores parabólicos y láminas planas. “COM DEV” construyó este tipo de antenas de puerta de enlace y cada satélite lleva 4 de éstas.

3.2.1.1.5. Módulo de control

Es el que controla y direcciona la red. “Arinc Inc.” es una compañía de “Annapolis, Maryland” que diseñó el *software* que llevan los satélites *Iridium*, este se ha desarrollado para el control y dirección de la red. “IONA Technologies Ltd.” está compañía desarrollo el *software* para el sistema de control que lleva el nombre de “ORBIX” y permite que en el segmento de control espacial trabaje varias aplicaciones e interactuen entre ellas.

3.2.1.2. Puertas de enlace (*gateway*’s)

La conexión entre el sistema *Iridium* y las redes terrestres fijas se realizará a través de puertas de enlace. Estas puertas de enlace forman una red y tienen dos funciones principales:

- El almacenamiento de la información de abonado: validez de la cuenta y situación geográfica para saber a dónde dirigir las llamadas que se hagan.
- Interconexión con la red telefónica pública conmutada (*PSTN*).

Para la conexión con los satélites, las puertas de enlace cuentan con una antena parabólica de largo alcance de 10 pies de diámetro (operando en frecuencias de banda Ka en un rango de 19 a 29 Ghz). Cada puerta de enlace controla un solo satélite *Iridium* a la vez, los satélites que viajan a una velocidad aproximada de una órbita cada 100 minutos, son captados por la terminal terrestre cuando se encuentran a una elevación de 8° por encima del horizonte y son seguidos hasta llegar a la misma elevación sobre el horizonte opuesto. Esto significa que si un usuario estuviera situado sobre la línea del Ecuador, podría estar conectado con cada satélite una media de 10 minutos, antes de realizar *handover* al siguiente satélite.

Almacenarán la validez de la cuenta de cada abonado y la situación del usuario. El sistema *Iridium* seguirá la posición de cada usuario esta se almacenará en las puertas de enlace. La necesidad de almacenar la posición del teléfono viene del hecho de que se puede recibir una llamada en cualquier momento y el sistema debe saber en que celda se encuentra para poder enviar la llamada. Asimismo por ejemplo, el usuario destino es un usuario de *PSTN*, la puerta de enlace hará de interfaz entre la red de satélites *Iridium* y la *PSTN* a la que se deba acceder. El sistema debe ser capaz de localizar y de utilizar la *PSTN* en cuestión y de esto se encargarán las puertas de enlace. En la figura siguiente se muestra una estructura de una puerta de enlace.

Figura 45. Estructura de una puerta de enlace o gateway



Fuente: *Iridium*

3.2.1.3. Terminales del usuario

La necesidad de tener comunicaciones móviles globales a llevado a diseñar el sistema *Iridium* recomendado para usuarios que se desplazan de un estándar celular a otro (por ejemplo de Europa a Sudamérica) o zonas de cobertura terrestre inexistente. Para esos usuarios se diseñaron los teléfonos móviles *Iridium*, de poco más volumen que los *GSM*, y los buscapersonas o localizadores de prácticamente igual tamaño que los terrestres, y se optimizó el sistema de transmisión de voz, para cada canal de voz se espera utilizar un ancho de banda de 8 khz., con una velocidad de transmisión de 2.4 Kbps *full duplex*, que también se usarán en las puertas de enlace a las *PSTN* analógicas para recrear la señal de audio siendo el sistema enteramente digital. Estos y otros comunicadores se pondrán en contacto directo y recibirán mensajes de los satélites desde cualquier punto de la superficie terrestre, marítima o incluso en el aire, en la banda L (1616 a 1626.5 Mhz). *Iridium* permite al usuario establecer enlaces para los servicios de fax y datos que operarán a 2400 baudios. Los enlaces de usuario utilizarán una modulación *QPSK* con un esquema de multiplexación compatible con los sistemas celulares terrestres.

A continuación se da información general de los terminales de usuario:

- Unidades transportables o bien a mano
- Pueden contactar con otro terminal *Iridium* del planeta, así como interconectarse con los sistemas actuales de telefonía por medio de la estación de conmutación.
- Información de latitud, longitud, altitud y hora respecto al meridiano *Greenwich*.
- Teléfono dual *GSM-Iridium*
- Peso de modelo de “**Motorola**”: 453 gr.
- Peso de modelo de “**Kyocera**”: 400 gr.

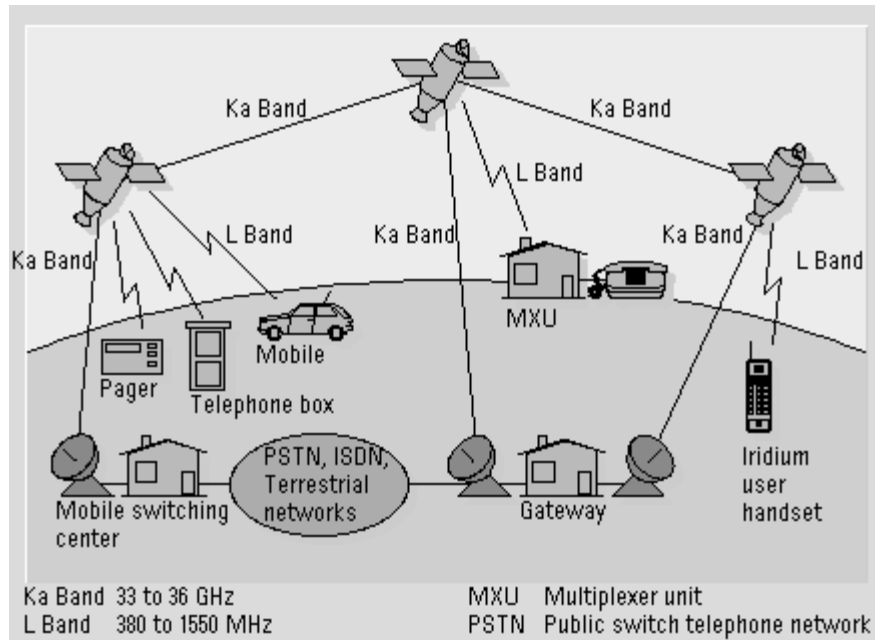
3.2.2. Asignación de frecuencias

El sistema *Iridium*, como otros sistemas *LEO*, consta de cuatro componentes básicos: una constelación de satélites no geoestacionarios *LEO*, terminales móviles de usuario, bases terrestres para control de los satélites y puertas de enlace.

A continuación se dan los parámetros de las bandas de frecuencias para los diferentes servicios y en la figura 46 se muestra la distribución de dichas bandas.

- Enlaces con los usuarios (*service links*): 1,616 - 1,6626.5 Mhz., (Banda L).
- Enlaces entre satélites (*ISL*) (*intersatellite transmissions*): 23,18 - 23,38 Ghz (Banda Ka).
- Enlaces de bajada satélite - pasarela: 19,4 - 19,6 Ghz (Banda Ka).
- Enlaces de subida pasarela - satélite: 29,1 - 29,3 Ghz (Banda Ka).
- Servicios especiales:
 - ◆ Canales de una sola dirección de alerta: 1626,270833 Mhz.
 - ◆ Para alertar a usuarios con terminales móviles terrestres de sólo recepción de la presencia de una llamada *paging*.
 - ◆ Las potencias de transmisión de las alertas son superiores a las de voz / datos normales permitiendo su recepción en condiciones más adversas donde los terminales terrenos no consiguen transmitir.
 - ◆ Transmisión de *pagings*: 1626,4375; 1626,395833; 1626,145833; 1626,104167 Mhz.
 - ◆ Su duración no excederá de 20,32 milisegundos.

Figura 46. Asignación de frecuencias



Fuente: Iridium

3.2.2.1. Enlace entre satélites (ISL)

Interconexión entre satélites, no limitándose a ser meros repetidores de señal.

- Ventajas:
 - ◆ Es posible encaminar el tráfico de larga distancia en el espacio, consiguiendo mayor autonomía, reduciendo el retardo de propagación y el coste de los enlaces terrestres.
 - ◆ Se puede dar servicios a zonas en las que no se ve ninguna pasarela.

- Inconvenientes:
 - ◆ Necesidad de antenas, transmisores, receptores y dispositivos de conmutación para esta función de *ISL*.
 - ◆ Estas antenas han de ser además orientables para apuntamiento, adquisición y seguimiento de otros satélites.
 - ◆ Mayor complejidad y coste de la carga útil.
 - ◆ No obstante, los inconvenientes se ve reducidos al tratarse de un sistema *LEO* de cobertura global, a diferencia de los sistemas *ICO*.

3.2.2.1.1. Tipos de enlaces entre satélites (*ISL*)

- Intraplano: interconexión de satélites en un mismo plano polar.
- Interplano: en distintos planos orbitales.
 - ◆ Es en este caso, donde van a haber más problemas de diseño.
 - ◆ El ángulo de apuntamiento de un satélite a otro en este caso es variable, siendo necesarias las antenas reorientables.
 - ◆ La conexión pasa por estados activo / inactivo.
 - ◆ Sin embargo, las tareas de TTC (telemetría, seguimiento y control) en los satélites son menos pesadas que en otros sistemas, por ejemplo *Leonet*, debido a que los ángulos de apuntamiento en *Iridium* varían dentro de unos márgenes menores. Una comparación de esto se muestra en la tabla III.

Tabla III. Comparación *ISL* entre *Leonet* e *Iridium*

<i>ISL</i>	<i>Leonet</i>		<i>Iridium</i>	
	Intraplano	Íterplano	Intraplano	Íterplano
Distancia (Km.)	15,000	4,550 a 18,460	4,030	3,270 a 4,480
Variación de la pérdidas (dB)	-----	12.2	-----	2.7
Apuntamiento horizontal (°)	0	-106 a 106	0	34 a 65
Apuntamiento vertical (°)	-36	-10 a -46	-16	-16

Fuente: *Iridium*.

3.2.3. Modulaciones y accesos múltiples que utiliza el sistema *Iridium*

3.2.3.1. Modulaciones

3.2.3.1.1. Cambio de fase (*PSK*)

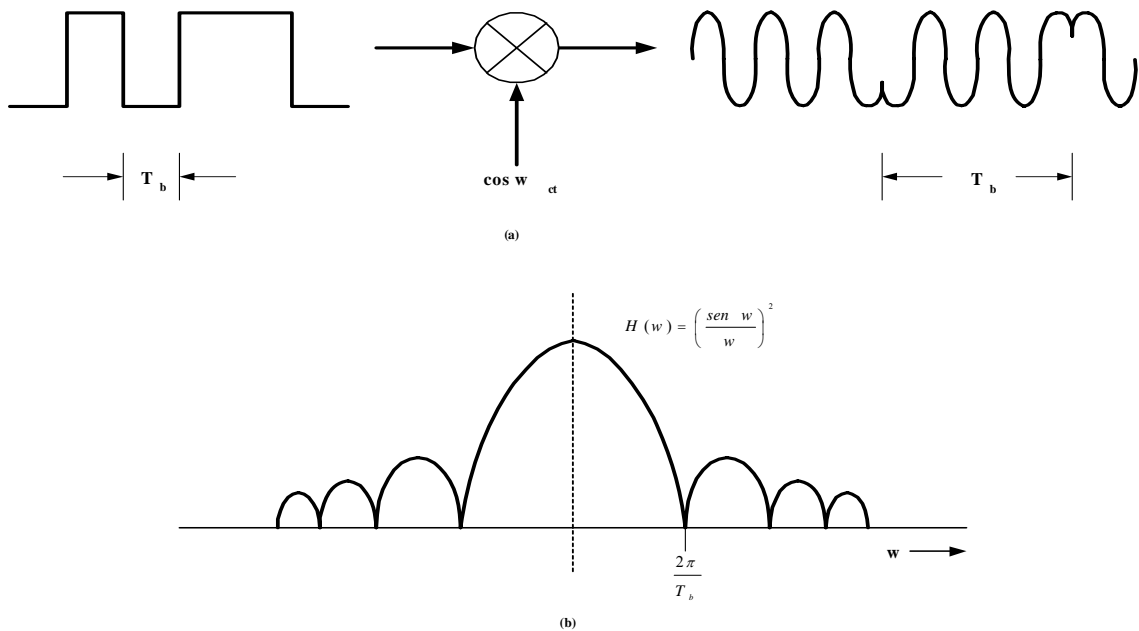
La forma mas simple de *PSK* es *PSK* binario (*BPSK*), los datos digitales modulan una portadora senoidal la salida modulada puede ser vista o asumida como dos posibles estados de fase (digamos 0 y π radianes) durante el periodo del *bit* (T_b), representando cualquiera el binario 0 o el binario 1. Esta forma de modulación es realmente idéntica a la modulación en amplitud con portadora suprimida y una señal modulante que tiene valores positivos y negativos (para binario 0 y 1). En tal modulación la amplitud de la portadora viene hacer negativa durante la parte negativa de la señal. Esto corresponde a la fase de π radianes de la portadora como se describe en la terminología *BPSK*. En el

dominio del tiempo, la portadora modulada aparece como una constante de envoltura senoidal, con rápidos cambios de fase a una velocidad llamada la velocidad de llaviado, dependiente de la velocidad de los datos digitales. En el dominio de la frecuencia la densidad espectral de potencia de la portadora modulada varía de acuerdo con

$$H(\omega) = \left| \frac{\text{sen } \omega}{\omega} \right|^2 \quad (3.1)$$

Como se ilustra en la figura 47, la mayoría de la energía de la señal modulada está contenida en el lóbulo mayor del ancho del cual depende del período del *bit* o de la velocidad de llaviado. El ancho de banda de la señal modulada está considerado para ser la porción del espectro contenido dentro del lóbulo mayor, los lóbulos menores del espectro se repiten indefinidamente a amplitudes cada vez más pequeñas, decreciendo a una tasa de $1/\omega^2$. Sin embargo, así como en FM, el espectro de las señales *PSK* moduladas es teóricamente infinito. En *BPSK*, imitando el ancho de banda va hacer aproximadamente igual a la tasa de *bits*, la energía de los lóbulos laterales que se pierde, con poco impacto en el desempeño. En la práctica las técnicas más sofisticadas de modulación se utilizan para restringir más la energía de la portadora modulada en el lóbulo mayor y reducir la potencia en los lóbulos laterales. Esto se logra procesando la señal digital de tal manera que las transiciones de datos sean lo menos abrupto posible resultando en más suaves transiciones en fase. Esto permite limitar un ancho de banda esencial y provee una operación más eficiente. La técnica más usada para este propósito es llamada *Minimum-Shift Key (MSK)*.

Figura 47. BPSK: a) En el dominio del tiempo; b) En el dominio de la frecuencia



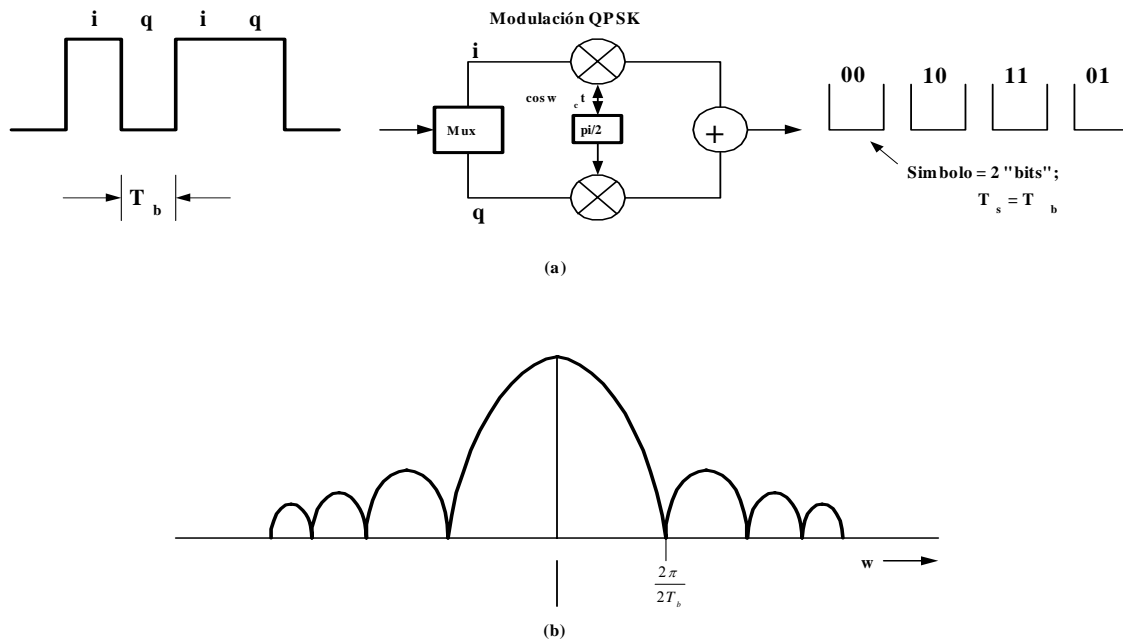
Fuente: Wilbur L. Pritchard, Henri G. Snyderhoud, Robert A. Nelson; *Satellite Communication Systems Engineering*; Página 336

3.2.3.1.2. Cambio de fase en cuadratura (QPSK)

Una mayor eficiencia en la utilización del ancho de banda de RF, puede obtenerse sin degradación en la tasa de error mediante la utilización de cuatro fases o PSK cuaternaria (QPSK) como se ilustra en la figura 48. Una señal modulada en QPSK se construye mediante la operación BPSK modulados en cuadratura. Los bits pares de la entrada se enrutan por el canal i (*in-phase*) y los bits impares se envían al canal Q (*quadrature*). La frecuencia portadora se alimenta directamente al modulador del canal i, pero se desfasa en 90° previo a entrar al modulador del canal de cuadratura. La salida de los dos canales se suman para formar la señal QPSK. Nótese que el estado de fase de

la salida del dispositivo sumador depende de ambos *bits* Q e i. Sin embargo el estado de salida para cada intervalo de señal (llamado un símbolo) depende de un par de *bits*. En el dominio de la frecuencia el espectro de potencia de nuevo toma el perfil $[(\text{sen } w) / w]^2$. Sin embargo, el ancho esencial de la señal *QPSK* es exactamente la mitad del necesario para la señal *BPSK* para la misma *rate* de *bits*. Esta reducción en el ancho de banda es el resultado del hecho, que la velocidad de llaviado a la salida del modulador se a reducido por un factor de 2 en el caso del *QPSK* comparado con el caso del *BPSK*.

Figura 48. QPSK: a) En el dominio del tiempo; b) En el dominio de la frecuencia



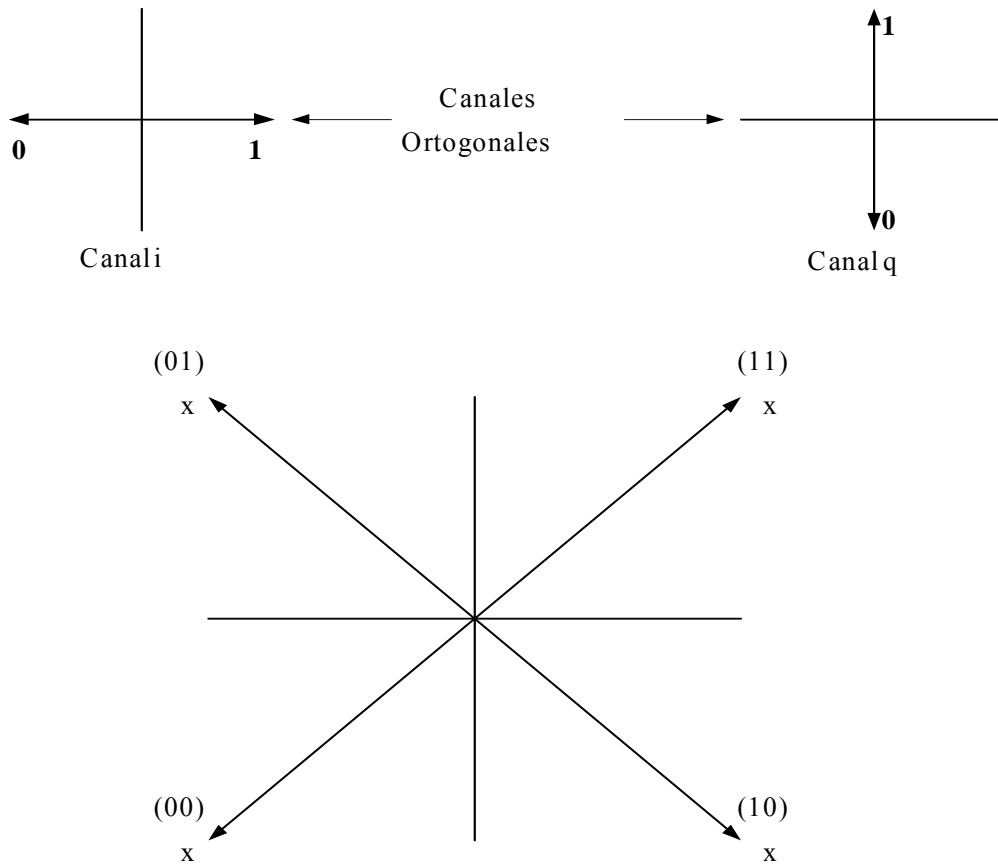
Fuente: Wilbur L. Pritchard, Henri G. Snyderhoud, Robert A. Nelson; *Satellite Communication Systems Engineering*; Página 340

Hay cuatro estados de fase que pueden asumirse para una señal *QPSK* como se ilustra en la figura 49. Nótese que cada estado de fase depende de un par de *bits*. Viéndolo vectorialmente, los *bits* del canal *i* operan en el eje horizontal en estados de fase 0 y π radianes; mientras que los *bits* del canal *Q* operan en el eje vertical en estados de fase $\pi/2$ y $3\pi/2$. El vector suma de la fase del canal *i* y de la fase del canal *Q* producen uno de los cuatro estados mostrados en el diagrama de fase. Dado que los canales *Q* e *i* son ortogonales en uno con respecto del otro, la ventaja de una detección coherente puede ser completamente realizada en canal independientemente. Esto significa que sobre el canal *i* solamente la componente de ruido enfasado puede causar un error, mientras que en el canal *Q* solamente la componente en cuadratura del ruido puede causar un error. Entonces la probabilidad de un *bit* errado en un canal es idéntica a la realizada en la operación *BPSK* a su correspondiente velocidad y sobre la base de la probabilidad de *bit* errado, ambos *BPSK* y *QPSK* se desempeñan idénticamente de acuerdo a la ecuación

$$P_b = \frac{1}{2} \left[\frac{2}{\sqrt{\pi}} \int_u^{\infty} e^{-v^2} dv \right] \quad (3.2)$$

Sin embargo el uso de *QPSK* provee una ventaja significativa comparada con *BPSK*, dado que *QPSK* puede obtener la misma tasa de error en un ambiente ruidoso dado mientras esta utilizando solamente la mitad de la banda requerida por *BPSK*.

Figura 49. Generación de estados de fase en QPSK



Fuente: *Wilbur L. Pritchard, Henri G. Snyderhoud, Robert A. Nelson; Satellite Communication Systems Engineering; Página 341*

3.2.3.1.3. Modulación por codificación de pulsos (*PCM*)

La señal *PCM* modulada en amplitud es todavía analógica; precisamente con un proceso de cuantificación y codificación se convertirá en una señal digital.

Durante el proceso de cuantificación se asigna a los impulsos de amplitud de la señal *PCM* un número limitado de intervalos de cuantificación discretos. Para ello se divide el margen de amplitudes de la señal en una cantidad igual de intervalos esto se muestra en las figuras 50 y 51. En el caso de que un valor de amplitud se encuentre situado entre los límites de un intervalo, se le asignará a este valor el intervalo mencionado. En el punto de recepción y para que el error de cuantificación permanezca lo menor posible, se reproduce este valor analógico que esté situado en la mitad (en el centro) de este intervalo de cuantificación.

Por codificación se sobreentiende un convenio por el cual cada uno de los intervalos individuales se numeran correlativamente con el código apropiado.

La secuencia de muestreo, cuantificación y codificación se resume bajo el concepto general de conversión analógico-digital (A/D).

Para la transmisión de los valores de amplitud se envía el código binario del intervalo correspondiente en forma de telegrama de impulsos. La diferenciación entre sí de los 8 niveles o escalones de amplitud elegidos como ejemplo en las figuras 50 y 51, requiere un código de 3 *bits*.

En el punto de recepción se genera, a partir de la decodificación del telegrama de impulsos, la señal analógica especificada [conversión digital-analógica (D/A)].

Figura 50. Transmisión PCM

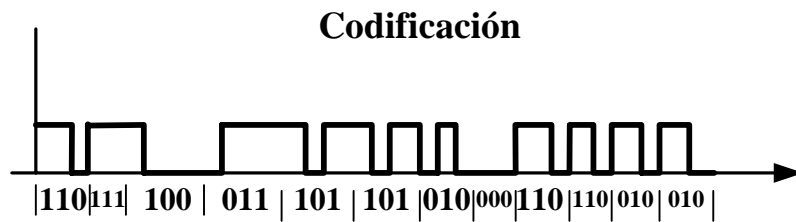
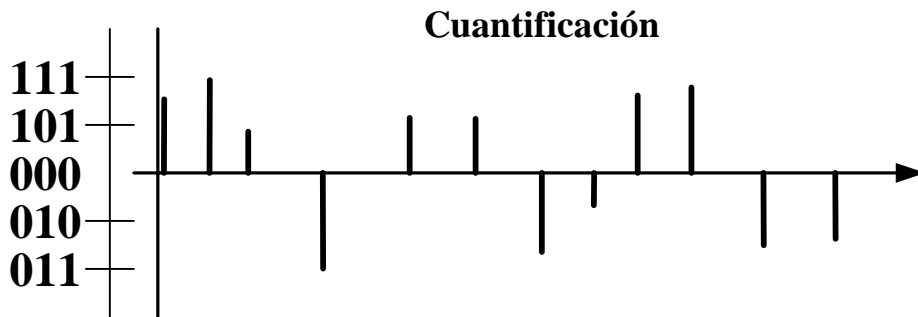
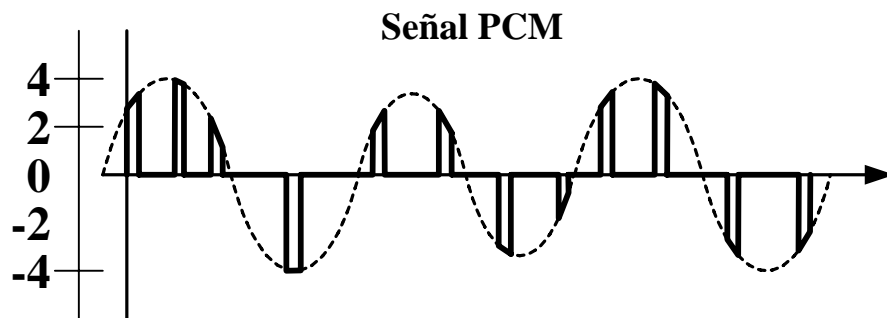
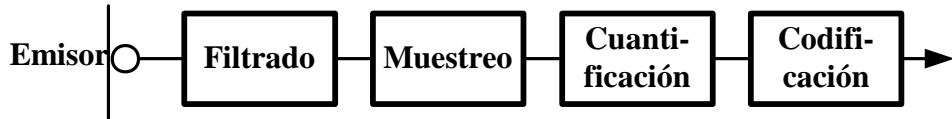
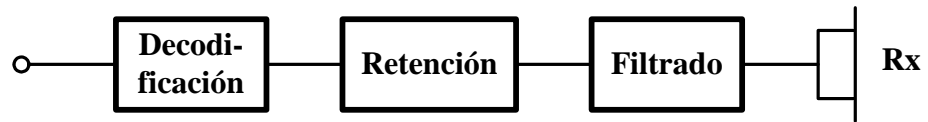
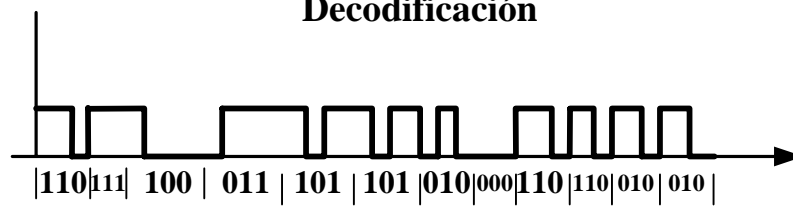


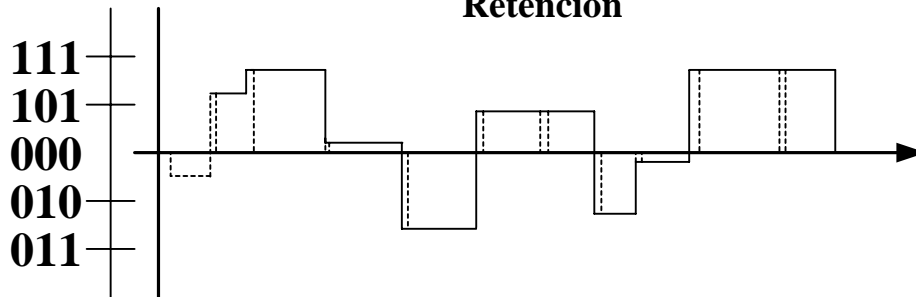
Figura 51. Recepción *PCM*



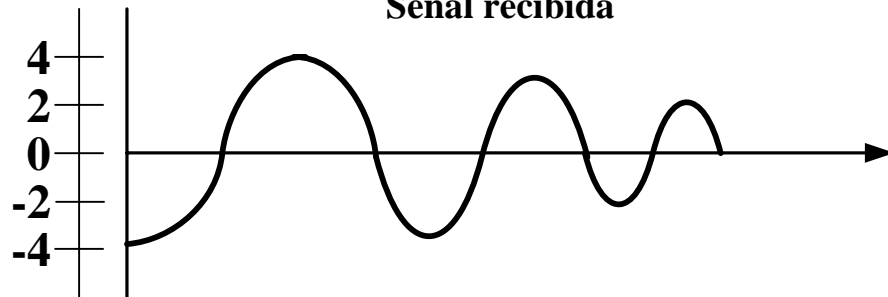
Decodificación



Retención



Señal recibida



3.2.3.2. Accesos múltiples

Un acceso múltiple está definido como la técnica mediante la cual, más de un par de estaciones terrenas pueden usar un satélite *transponder*. Esta es la técnica usada para explotar la ventaja geométrica de los satélites y es el núcleo de la red satelital.

3.2.3.2.1. Acceso múltiple por división de frecuencia (*FDMA*)

Estos sistemas canalizan un *transponder* utilizando portadoras múltiples. El ancho de banda asociado con cada portadora puede ser tan pequeño como lo requiera un canal de voz. *FDMA* puede usar transmisión analógica o digital en el modo continuo o como ráfagas. En la figura 52 se presenta un espectro compuesto de *FDMA*, en el *transponder*.

Figura 52. *FDMA*

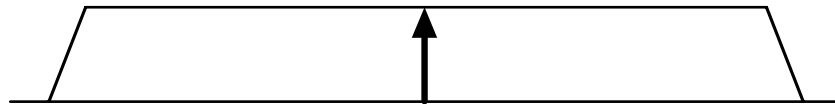


Fuente: Wilbur L. Pritchard, Henri G. Snyderhoud, Robert A. Nelson; *Satellite Communication Systems Engineering*; Página 360

3.2.3.2.2. Acceso múltiple por división de tiempo (TDMA)

Se caracteriza por el uso de una portadora simple digitalmente modulada por *transponder*, en donde el ancho de banda asociado por la portadora es típicamente todo el ancho del *transponder* esto minimiza la velocidad de transmisión para el *transponder*. La velocidad de transmisión de la portadora se comparte en tiempo entre un número de estaciones terrenas tales que la suma de la información de tráfico entre estaciones terrenas individuales nunca excedan la velocidad de la portadora. Sin embargo, cada estación terrena debe transmitir su información en ráfagas a la velocidad acordada de la portadora en espacios de tiempo preasignados y recurrentes. Claramente los requerimientos de tiempo que deben cumplirse en el sistema son crucialmente exigentes, particularmente a causa de que el satélite esta siempre ligeramente moviéndose en una forma ondulante. Sin embargo, la principal ventaja de TDMA es que esta realizada en arreglos de portadora simple por *transponder*, hay casos donde el ancho de banda de TDMA puede ser una fracción del ancho de banda del *transponder*. En la figura 53 se presenta el concepto básico de TDMA en el uso de satélites.

Figura 53. TDMA



Fuente: Wilbur L. Pritchard, Henri G. Snyderhoud, Robert A. Nelson; *Satellite Communication Systems Engineering*; Página 360

3.2.3.3. Arquitectura de la comunicación del sistema *Iridium*

Los enlaces de los servicios de comunicación de la banda L operan entre el rango de 1610 a 1626.5 MHz., el total de ancho de banda ocupado es de 5.15 MHz., la técnica de modulación que utilizan es del cambio de fase en cuadratura (*QPSK*). El máximo número de canales con el que cuenta es de 3,840, con polarización circular derecha. Ambos enlaces de subida y bajada operan en la misma banda, utilizando la duplexación por división de tiempo (con sus siglas en ingles TDD).

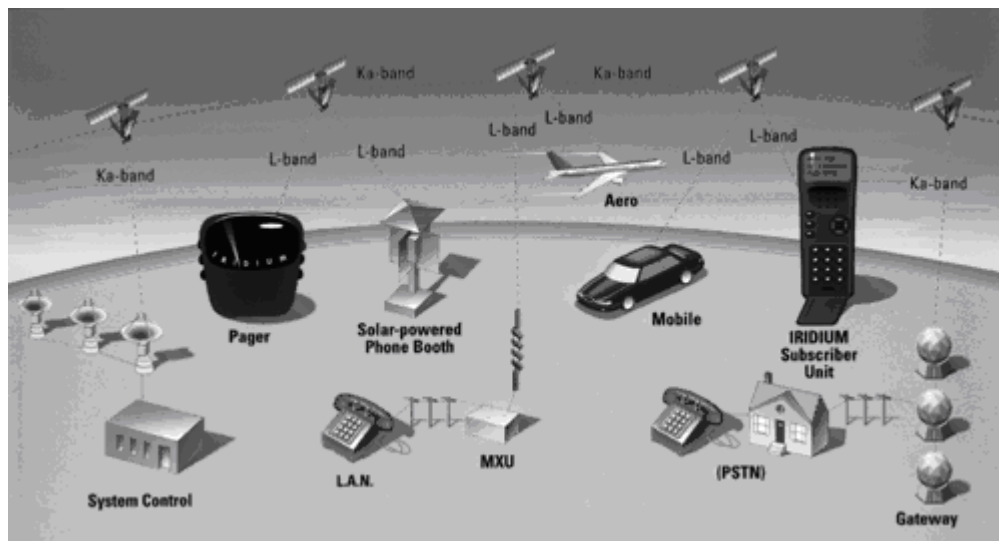
Toda la banda K de enlaces para arriba y para abajo y los cruzados están empaçados en *TDM/FDMA* usando modulación *QPSK*. Esos enlaces tienen una tasa FEC de $\frac{1}{2}$, utilizando una codificación convencional con decodificación *Viterbi*. Los datos codificados por las puertas de enlace y satélite son a razón de 6.25 Mbps., para un fácil control de enlace, con los enlaces cruzados de alta velocidad tienen una razón de datos de 25Mbps.

Ambos enlaces para arriba y para abajo ocupan un total de 100 MHz. de ancho de banda y usan una polarización circular derecha. Los enlaces intersatélitales utilizan 200 MHz. de ancho de banda con polarización vertical. Los enlaces cruzados utilizan como entrada las pasarelas, este proceso solo necesita que la información llegue a la estación designada por ellos. Los enlaces cruzados permiten que se requieran menos infraestructura en tierra, además de tener un menor costo el recorrido y permitir una cobertura global total (incluyendo los océanos).

3.2.4. Servicios

La oferta de servicios *Iridium* tienen como denominador común la conveniencia de la movilidad e integración tecnológica. La propuesta comercial de *Iridium* se basa en los siguientes servicios: voz, radiolocalización (*paging*), fax, datos; estos servicios se pueden obtener en cualquier punto de la Tierra (tierra, mar y aire) siempre que la altitud sea inferior a 185 Km. La figura 54 muestra los servicios prestados por el sistema *Iridium*.

Figura 54. Servicios del sistema *Iridium*



Fuente: *Iridium*

3.2.4.1. Voz

Por el modo de telefonar, el usuario de *Iridium* puede utilizar un sistema local terrestre (cuando esta disponible) o el *Iridium* que esta disponible las 24 horas del día, por medio del cual puede llamar, localizar o transmitir llamadas, la alta calidad de la transmisión provee un servicio de voz digital con muy pocos segundos de retraso no como los que se encontraron en sistemas de satélite geoestacionarios. Además, la banda L de enlaces celulares esta diseñada con canales fuertes. El acceso para la constelación de satélites *Iridium* es a través de las técnicas de acceso *FDMA/TDMA*, usando una antena de bajo contorno. La unidad del usuario utiliza baja potencia a comparación con los teléfonos actuales que utilizan satélites. El protocolo de sistemas como *TDMA*, combinado con el protocolo único de *Iridium* envía la voz al destino apropiado. *Iridium* para el servicio de voz ofrece los siguientes planes:

- Solo Satélite: ofrece comunicaciones móviles con terminales de mano a través de satélite, desde cualquier punto del globo.
- Universal: Con la ayuda de terminales duales celular satelital, el usuario podrá utilizar la constelación de satélites cuando su estándar celular no esté disponible y viceversa, asegurando un *roaming* bidireccional, cobertura global y las ventajas de ambas tecnologías.
- Ciudad a Ciudad: Sin utilizar la constelación de satélites, este servicio ofrece la posibilidad de efectuar *roaming* entre protocolos celulares terrestres como el *GSM*, *TDMA* y el *IS-41 (AMPS)*, en América principalmente). El usuario utilizará su suscripción *GSM* y los servicios ofrecidos por ella en el mundo *AMPS* y viceversa, conservando un único número personal y un único recibo con su proveedor de servicios.

3.2.4.2. Radiolocalización

El sistema *Iridium* ofrece la solución al actual problema de *roaming* en redes de radiolocalización, a través de la emisión de mensajes por satélite en las áreas suscritas por el usuario, incluyendo la posibilidad de búsqueda en todo el globo.

El sistema incluye un servicio de mensajes tanto numéricos con alfanuméricos, las unidades de radiolocalización tienen una forma similar a la de los radiolocalizadores que se utilizan hoy. La función de rescate del mensaje utiliza una multiplexación *FDMA/TDMA* para compartir la banda L con los enlaces de voz y otros servicios del *Iridium*. En la transmisión al radiolocalizador se puede alcanzar gran exactitud de los caracteres en el mensaje siendo este de gran alcance o no, según sea la posición de la unidad receptora.

3.2.4.3. Fax

El sistema *Iridium* tiene comunicación de mensajes de fax, los faxes son formateados en paquetes digitales de voz usando el mismo protocolo e interfaces de ruteo que las unidades del abonado. Estas unidades tienen incorporado un puerto de datos para que funcione de interfaz entre estas y una unidad de fax.

3.2.5. Comunicación pasó a paso del sistema

A continuación se resumen los pasos para una llamada usando el sistema *Iridium*.

- Cuando un teléfono se active se conectará al satélite más próximo.
- Gracias a la red de pasarelas podrá determinar la validez de la cuenta y situación del usuario.
- El usuario podrá realizar una llamada eligiendo entre las alternativas de transmisión celular terrestre o vía satélite, dependiendo de la compatibilidad y disponibilidad del sistema celular terrestre.
- En caso de no estar disponible el sistema celular del abonado, el teléfono comunicará automáticamente con el satélite.
- La llamada será transferida de satélite en satélite a través de la red hasta su destino sin saber el abonado donde se encuentra la persona a la que llama, un teléfono *Iridium* o una puerta de enlace del sistema *Iridium*.
- Las comunicaciones podrán involucrar varios satélites y puertas de enlace terrestres donde se almacena información general y la situación de los usuarios hasta su destino.

3.2.6. Mercado del sistema *Iridium*

El sistema *Iridium* tiene un presupuesto 4,600 millones de dólares, y se calcula que *Iridium* necesita alrededor de un millón de clientes para mantenerse, se piensa que las personas que pagarán este servicio son:

- Personas que tienen trabajos en lugares remotos sin un buen servicio telefónico.
- Representantes comerciales, periodistas y en general, personas cuyo trabajo exige movilidad en grandes áreas (algunas sin cobertura terrestre), pero con necesidad de estar localizados en todo momento.
- Personas de clase acomodada que pueden pagar por el servicio global, y que viven en zonas poco pobladas sin cobertura celular (por ejemplo zonas residenciales de lujo, terratenientes en grandes latifundios,...).
- Países del tercer mundo donde la cobertura telefónica (cable y celular) es muy pequeña, pero se necesita comunicación desde la capital con los jefes locales de áreas remotas.
- Empresas con constantes cambios de ubicación. Especialmente, aquellas que han de moverse por países (con estructuras anticuadas) donde el tiempo de espera a la instalación de una línea telefónica es de varios años.

4. RED SISTEMA INTEGRADO DE MENSAJERÍA

Después de describir e informarse de la teoría necesaria que involucran a los componentes básicos de este sistema integrado de mensajería, se procede a integrar estos, en un diseño de red que pueda cumplir los objetivos y aplicaciones especiales para que el sistema sea vanguardista.

Este capítulo pretende transmitir todos los pasos necesarios llevados a cabo en la integración de los componentes que forman el sistema integrado de mensajería. Como se sabe, al momento de diseñar una red, el punto de partida es la planificación y dimensionado de red y es por allí donde se empieza, ya que es uno de los procesos clave para el desarrollo del negocio de un operador de telecomunicaciones. Luego se procede a desarrollar el diseño de red, enfatizando en las funciones y características principales que debe llevar cada componente del sistema para que este sea óptimo. Y para finalizar, se realiza una evaluación del sistema por medio de un estudio de preinversión el cual contempla análisis económico el cual determina la factibilidad del sistema.

4.1. Planificación y dimensionado de red

Es innegable que una planificación y dimensionado inadecuados puede llevar al operador a situaciones de bloqueo, incapacidad para reaccionar ante los cambios del mercado y, por ende, a ser incapaz de sobrevivir en el entorno de la competencia actual. Es por ello que la planificación y dimensionado de red se encargan de encontrar las condiciones bajo las cuales se ofrece un servicio adecuado, a los subscriptores o abonados, a la vez que se hace un uso económico de las facilidades disponibles; logrando esto, con un análisis del tráfico telefónico.

La capacidad de un sistema depende de varios factores. Un factor importante es el número de canales disponibles para la voz y datos. Asignar el número correcto de canales para cada área específica requiere una aplicación de la teoría del tráfico telefónico. La teoría del tráfico telefónico utiliza estimaciones acerca de la conducta del suscriptor y de la manera como él utiliza el sistema; con estos estimados y modelos matemáticos se puede calcular el número requerido de canales. Se pueden deducir varios períodos en el cual el volumen de tráfico varió, estos regularmente pueden ser: a) variaciones de semana a semana, b) de día a día y c) de hora a hora. Estas variaciones se toman en la hora más activa o pico.

Antes de empezar a calcular las variables que interesan para llegar a saber el número de canales de voz, sería bueno definir qué es el tráfico telefónico y de allí partir hacia estas variables, las cuales son: relación tráfico por usuario; intensidad del tráfico telefónico; y grado de servicio.

El tráfico telefónico se mide en *erlangs*^[3], consiste en el conjunto de llamadas ofrecidas a un grupo de circuitos o canales de voz. Y un *erlang* es la ocupación de un circuito durante una hora. Con esto se puede ver la cantidad de tiempo que un usuario ocupa un circuito de voz o datos.

4.1.1. Relación tráfico por usuario

Basado en las definiciones anteriores, se dice que uno de los objetivos principales del tráfico telefónico es que la red esté diseñada para dar servicio durante la hora más activa o pico, en la parte más cargada del año. Teniendo en cuenta esto, se pretende predecir la demanda del tráfico a través de promedios, para lo cual se tiene que definir una hora activa o pico, un tiempo medio de conversación y un número promedio de llamadas en la hora activa o pico; para los cálculos estas estimaciones se tomarán en una hora activa estimada durante un período de muestra de una semana.

Basándose en las definiciones y estimaciones que se realizarán, se utilizará la ecuación siguiente para calcular la relación tráfico por usuario.

$$\bar{\varepsilon} = \sum_{k=1}^N (n_k / \Delta T) T_k \quad (4.1)$$

Donde $\bar{\varepsilon}$ es el tráfico promedio por usuario;

n_k es el número promedio de llamadas del usuario;

T_k es el tiempo promedio que el usuario utiliza para hacer una llamada o tiempo medio de conversación;

ΔT es el intervalo de tiempo donde el tráfico es más alto o llamado hora pico.

4.1.1.1. Hora activa o pico

Esta hora es definida desde el punto de vista de ingeniería como el período de 60 minutos consecutivos durante el cual el tráfico telefónico es el más alto. En la práctica sin embargo para simplificar las mediciones del tráfico, la hora pico es siempre una, comenzando en la hora, media hora o cuarto de hora, y es la más cargada de tales horas.

Como se sabe el tráfico tiende a variar dependiendo la hora, pero el interés es dimensionar la red en la hora más ocupada y ésta se considerará en el período de las 15:00 Hrs. a 15:30 Hrs.¹, con lo que se tiene una hora pico de 30 minutos o sea un $\Delta T = 30$ mins.

¹ Se tomo este período de referencia, dado que es un dato real usado por Telered.

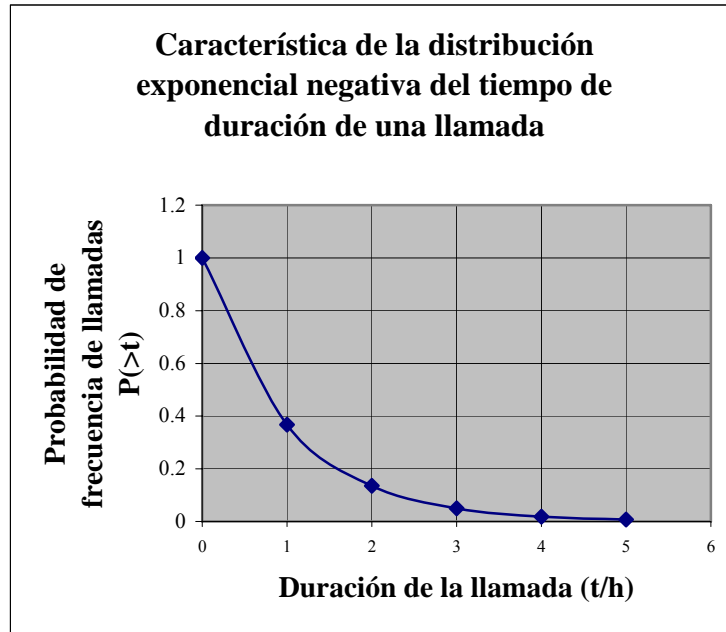
4.1.1.2. Tiempo medio de conversación

Ahora se necesita calcular el tiempo promedio que utiliza un usuario para hacer una llamada, pero antes de esto se puede hablar de algunas características del tiempo de duración de una llamada. Los tiempos de duración de una llamada se pueden clasificar en dos tipos:

1. Constantes: estos tiempos de duración se aplican en el caso de telefonistas y órganos de control como registradores, marcadores, etc., los cuales sólo intervienen el tiempo necesario para establecer la conexión.
2. Variables: estos se aplican a las conversaciones de abonados.

La distribución del tiempo de duración de una llamada tiene un comportamiento modelado por la distribución exponencial negativa con una probabilidad expresada por $P(>t) = e^{-t/h}$, mostrado en la figura 55. Esto indica que en un proceso *poissoniano* el tiempo que transcurre entre dos llamadas sucesivas llamado tiempo entre llamadas, se distribuye exponencialmente; en otras palabras, el intervalo de tiempo entre dos llamadas consecutivas tiene una distribución aleatoria.

Figura 55. Comportamiento *poissoniano* del tiempo de duración de una llamada.



Luego de un recordatorio de las características principales del tiempo de duración de una llamada se procede a calcular el tiempo medio de conversación. Lo que se hace es promediar los tiempos de cada llamada de un usuario en la hora pico durante cinco días, se toma durante cinco días asumiendo que los otros dos son fin de semana y el tráfico no tiene mucha demanda; por lo tanto se tiene:

$$\bar{T}_k = (T_1 + T_2 + T_3 + T_4 + T_5) / 5 \tag{4.2}$$

$$\bar{T}_k = (1 + 0 + 1.5 + 3 + 2.5) / 5 = 1.6 \text{ [min.]} \tag{4.3}$$

Como resultado se obtuvo que el tiempo medio de conversación estimado es de 1.6 minutos. Esto muestra que tiempos de duración de llamadas de 1 a 3 minutos son relativamente frecuentes mientras duraciones de llamadas de 10 ó más minutos ocurren con mucha menor frecuencia, esto se visualiza en la figura 55.

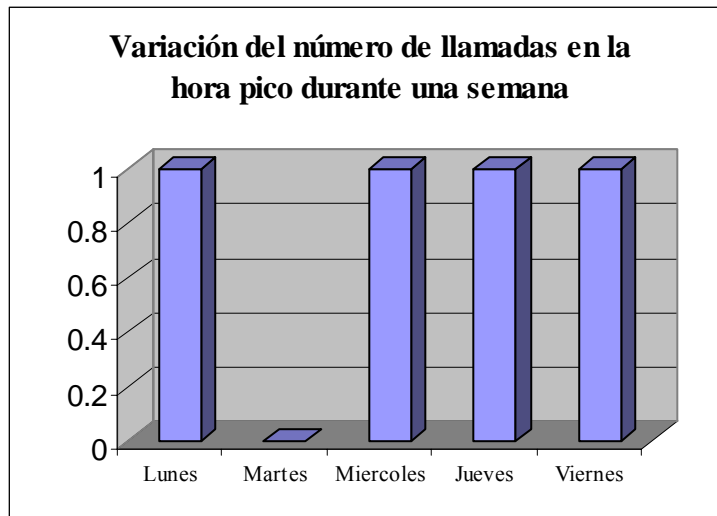
4.1.1.3. Número promedio de llamadas por usuario en la hora pico

La naturaleza aleatoria de la distribución del tráfico dentro de la hora está basada en la suposición de que los usuarios originen llamadas independientemente de otros usuarios. Esta suposición afecta mucho el comportamiento del tráfico, pero siempre se tiene un promedio del número de llamadas por usuario, este dato se obtiene promediando las llamadas hechas por el usuario en la hora pico durante una semana como se muestra en la figura 56, entonces se tiene:

$$\bar{n}_k = (n_1 + n_2 + n_3 + n_4 + n_5) / 5 \quad (4.4)$$

$$\bar{n}_k = (1+0+1+1+1) / 5 = 0.8 [\text{llamadas / usuario}] \quad (4.5)$$

Figura 56. Muestra de llamadas en la hora pico durante una semana.



Las ecuaciones anteriores proporcionan el número promedio de llamadas por usuario en la hora pico durante una semana y el cual es igual a 0.8 llamadas / usuario.

Después de haberse definido y estimado las variables de interés se procede a calcular la relación tráfico por usuario, ingresando los valores de variables encontradas en la ecuación (4.1).

Se puede escribir la ecuación (4.1) de la forma siguiente, para luego hacer el cálculo y obtener la relación tráfico por usuario ($\bar{\varepsilon}$):

$$\bar{\varepsilon} = (\bar{n}_k * \bar{T}_k) / \Delta T \quad (4.6)$$

$$\bar{\varepsilon} = (0.8 * 1.6) / 30 = 0.0427 [\text{erlang} / \text{usuario}] \quad (4.7)$$

4.1.2. Intensidad de tráfico telefónico

Dada la cantidad de tráfico generado por un usuario, se puede calcular el tráfico total del sistema y éste no es más que la intensidad de tráfico telefónico que en otras palabras representa el promedio de la densidad de tráfico durante un período de una hora y también representa:

- a. El promedio de llamadas originadas durante un período de tiempo igual al promedio del tiempo de ocupación; y
- b. El tiempo total, expresado en horas para cursar todas las llamadas.

Para el cálculo de la intensidad de tráfico sólo se multiplica la relación de tráfico por usuario (o el tráfico de un usuario) por el número total de usuarios que se esperan, ecuación (4.8), en este caso se estima que inicialmente el sistema estará diseñado para un número total de 10,000 subscriptores o abonados.

$$E = N * \bar{\varepsilon} \quad (4.8)$$

Donde: E es la intensidad de tráfico telefónico o el tráfico total del sistema

N es el número de usuarios

$\bar{\varepsilon}$ es la relación tráfico por usuario

$$E = 10,000 * 0.0427 = 427[\text{erlangs}] \quad (4.9)$$

Entonces, para 10,000 abonados se generará un tráfico total igual al calculado en la ecuación (4.9) que es de 427 *erlangs*.

4.1.3 Grado de servicio

El objetivo final es fijar unas estructuras y un dimensionado, de tal manera que se garantice un grado de servicio adecuado a la demanda de los usuarios. El grado de servicio se define como la probabilidad de que el usuario llamante pueda conseguir el establecimiento de su llamada en un instante determinado. En el caso de que en ese instante la red no disponga de recursos insuficientes para realizar este establecimiento, la llamada será rechazada; por lo que con esto es necesario abandonar el concepto ideal de servicio sin bloqueo, por lo que el usuario debe realizar otra vez la llamada o retrasarla.

Se puede calcular el grado de servicio como la relación entre el número de llamadas perdidas y el número total de llamadas ofrecidas o recibidas por la central; esta relación siempre se estima durante la hora activa o pico de la semana de muestra.

Para este caso se asume que se atienden 1500 llamadas diarias, de las cuales, 30 llamadas van a hacer desbordadas o perdidas, entonces se tiene:

$$GOS = \frac{\text{Número de llamadas perdidas}}{\text{Número total de llamadas ofrecidas}} \quad (4.10)$$

donde GOS es el grado de servicio.

$$GOS = \frac{30 * 5}{1500 * 5} = 0.04 \quad (4.11)$$

El resultado de este cálculo es de un grado de servicio igual al 4%, el cual es aceptable ya que está entre el rango del 2% al 5%. Aunque esto es aceptable es algo muy engañoso porque una calidad más alta de servicio es una probabilidad más alta de congestión.

Después de haber encontrado las variables de interés, siendo ellas: la intensidad de tráfico telefónico y el grado de servicio, se procede a averiguar el número de canales de voz que necesita el sistema y para eso se utiliza el concepto de *erlang B*, el cual se basa en el bloqueo de llamadas sin espera y abandono inmediato; donde el usuario no intenta hacer otra llamada y la cual realiza más tarde.

Siguiendo en el cálculo, se tiene $E = 427 \text{ erlangs}$ y $GOS = 4\%$, con estos valores se utiliza una tabla de tráfico de *erlang B* ó para que sea más fácil un calculador de tráfico de *erlang B*, los cuales proporcionan el número de canales de voz; se utilizó un calculador de *erlangs B* que dio un resultado igual a 426 canales de voz.

Luego de saber el número de canales de voz que el sistema requiere, se debe calcular el número de enlaces digitales de voz o *EI*'s. Se utiliza el *E1* debido a que puede transportar voz, datos y video, adecuándose más a esta aplicación por su alta capacidad y calidad de transmisión.

La trama *E1* esta compuesta de 32 canales *DS0* cada uno a 64 kb/s (8 bits/palabra *PCM* muestreada a 8000 muestras por segundo), estos 32 canales *DS0* son multicanalizados en conjunto para formar una trama de 256 *bits* (8 bits/palabra *PCM* x 32 (tiempos) = 256 *bits*). Así de esa forma, la trama *E1* opera a una velocidad de 2048 kbits/seg, que brindan la compatibilidad a la red con la *PSTN*.

La multicanalización para *E1* es por División de Tiempo (*TDM*) para manejar varios canales *PCM* (30 canales de voz) dentro del flujo de bits de 2.048 Mb/s. Para mantener la alineación de trama y sincronización, se maneja información en la ranura de tiempo 0 del *E1*.

Calculando el número de *EI*'s se tienen las ecuaciones siguientes:

$$\text{Número de } EI's = \frac{\text{Número de canales [canales]}}{30 [\text{canales} / EI]} \quad (4.12)$$

$$\text{Número de } EI's = \frac{426}{30} = 14.2 [EI's] \quad (4.13)$$

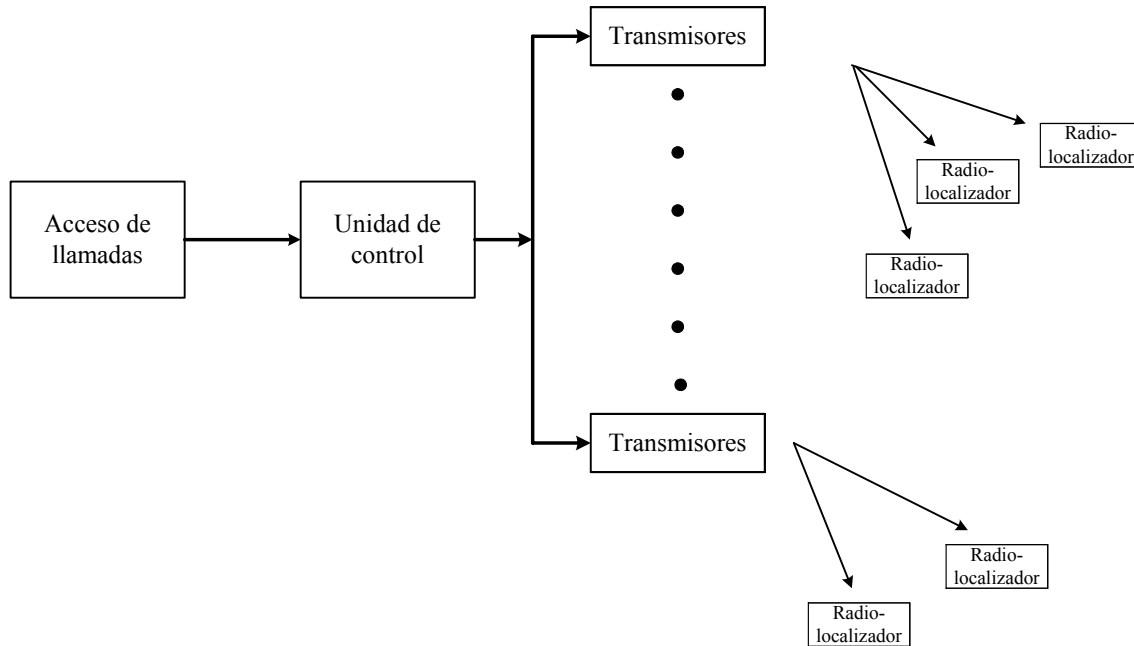
El resultado que proporciona la ecuación 4.13 es de 14.2 *EI*'s, para mejor conveniencia en costos se utilizarán 14 *EI*'s, que satisfacen la demanda que debe atender la capa de conmutación para prestar de buena forma los diferentes servicios a los 10,000 abonados. El costo del total de *EI*'s se verá más adelante en el análisis de costos de red.

4.2. Diseño de red Sistema Integrado de Mensajería

En este apartado, el más interesante, puesto que es aquí donde se integraran todos los componentes del sistema integrado de mensajería, se desarrollará un diseño completo por bloques principales de la red, para luego ir describiendo las funciones y características de los componentes de cada bloque; y por último se obtiene el diseño general del sistema y los servicios prestados.

Este sistema es adecuado para personas con un nivel alto de movilidad, con desplazamientos continuos y a quienes es necesario comunicar mensajes o decisiones urgentes o no, pero que deben ser conocidas por ellos. Es por esto que a continuación se presenta una solución integrada de mensajería que satisface estas necesidades del abonado.

Esta solución integrada no es más que una red de radiolocalización a la cual se le adaptan primero telefonía *IP* que es la voz transmitida sobre el protocolo *Internet (IP)*, esto permite a los usuarios establecer llamadas de voz, fax, datos, sobre conexiones *IP* en redes *Lan's*, *Intranets*, *Internet*, etc.; esto se explicará con más detalle en las siguientes páginas y segundo un sistema de reconocimiento de voz, que permite convertir la voz a texto. Sabiendo a grandes rasgos las principales características del sistema integrado de mensajería, se inicia con el diseño en bloques de la estructura del sistema, mostrado en la figura 57.

Figura 57. Diseño en bloques red Sistema Integrado de Mensajería.

4.2.1. Acceso de llamadas

Esta sección describe el primer bloque de la estructura del sistema, este es el acceso de llamadas. Este bloque está compuesto de las fuentes de entrada para acceder al sistema y la capa de conmutación que en este caso es un centro de llamadas y un controlador de conmutación virtual basado en la infraestructura de paquetes.

4.2.1.1 Fuentes de entrada

Para poder contar con un servicio muy eficiente, el sistema maneja los mensajes de una forma automática, evitando gran cantidad de operarios en el proceso y solo requiriendo supervisores técnicos de red.

Los elementos primarios en el bloque de acceso de llamadas del sistema son las fuentes de entrada, estas son importantes porque de ellas dependen los diferentes servicios que se puedan prestar a los abonados.

Las fuentes de entrada por la que el usuario puede conectarse al sistema pueden ser: a) por un teléfono a través de la *PSTN*, b) teléfonos *IP*, ya que se tiene telefonía sobre *IP*, c) *Internet* o red *IP* a través de un computador personal.

- a) Teléfono a través de la *PSTN*: el usuario llama a través de un teléfono de la *PSTN*, a un centro de llamadas de telefonía de paquetes (*PTCC*) el cual remplazará a una *PBX*, del *PTCC* se hablará más adelante.
- b) Teléfonos *IP*: en una infraestructura de *IP*, se pueden utilizar funciones de datos como el protocolo de configuración dinámica de *host*. Este protocolo permite que un dispositivo (un PC o un teléfono *IP*) reciba dinámicamente una dirección *IP*. Por lo tanto, si el usuario tiene un teléfono *IP* configurado, podrá comunicarse con el sistema entrando en el servidor de red, teniendo la ventaja de que a la red *IP* no le importa dónde se encuentre este.
- c) *Internet* o red *IP*: un servicio muy prometedor es el envío de texto o, aún más, archivos completos de una computadora al abonado, quien después podrá manipular los datos o el archivo en una computadora personal. Este servicio se puede llevar a cabo por medio de *Internet* o la red *IP*, el usuario se conecta a través de *Internet* o por medio de la red *IP* al servidor del sistema, para eso el usuario deberá tener una dirección *IP* configurada en su PC.

4.2.1.2. Conmutación virtual

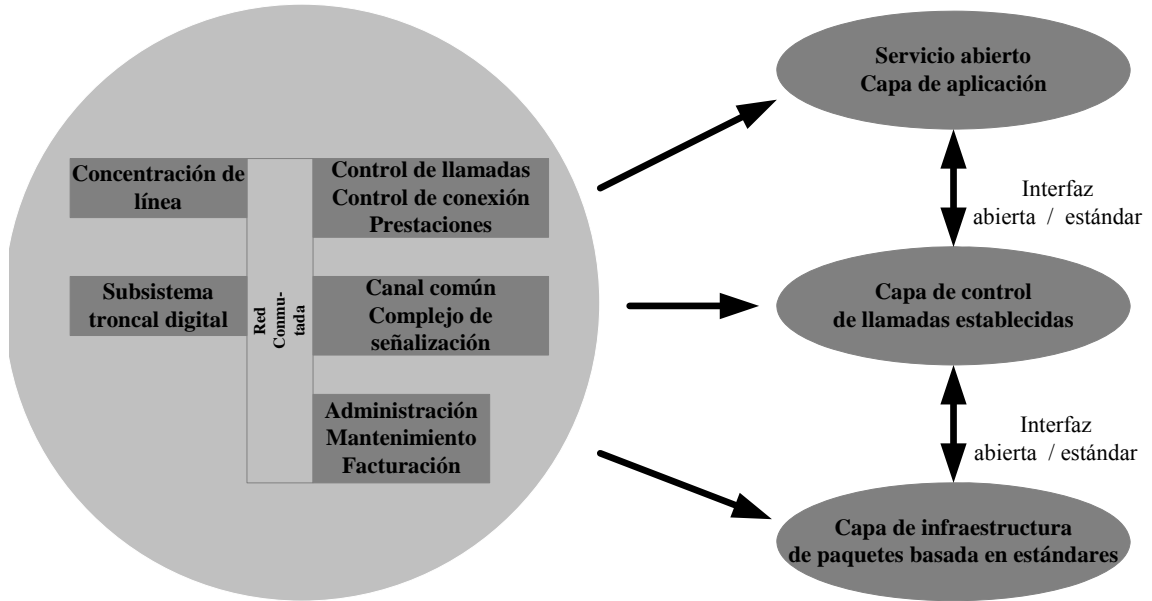
Optimizar recursos, es uno de los objetivos cada día de los sistemas, es por eso que en esta aplicación se desea transportar voz comprimida a 8 kb/s., con lo cual se tienen 240 canales de voz a 8 kb/s., por E1. Para lograrlo se utilizará voz sobre *IP*, que de ahora en adelante llamaremos *VoIP*, ya que la *VoIP* proporciona la capacidad de fragmentar la voz en pequeñas piezas, conocidas como muestras y colocarlas en un paquete *IP* (protocolo *Internet*); lo cual satisface las características que se necesitan.

Debido a que se transportará *VoIP*, se va a utilizar una nueva estructura de servicios de voz en paquetes en la capa de conmutación, en lugar de la ya conocida infraestructura de conmutación de circuitos basada en *TDM*.

La base de esta nueva estructura de paquetes es el modelo de circuito conmutado, el cual se convierte en un nuevo modelo de estándares abiertos de tres capas llamado conmutador virtual, mostrado en la figura 58. A continuación se presentan las tres capas:

- a) Capa de infraestructura de paquetes basada en estándares: esta se utiliza para la transmisión de voz, se basa en el protocolo de transporte en tiempo real *RTP* (*Real Time Transport Protocol*) que junto con la cabecera del protocolo del datagrama del usuario *UDP* (*User Datagram Protocol*) e *IP* proporcionan la marca de temporización u hora exacta en la cual se envió el paquete, se anota normalmente como *RTP / UDP / IP*. Básicamente el dispositivo de recepción utiliza la marca de temporización para determinar a la hora que se esperaba un paquete, si el paquete era correcto y si se recibió cuando se esperaba. Esta información ayuda a enmascarar cualquier problema de red como por ej. retraso, fluctuaciones de fase y pérdida de paquetes. El control del flujo de paquetes lo hace el *TCP / IP* (*Transmission Control Protocol*).

Figura 58. Conmutación de circuitos frente a conmutación de paquetes.



Fuente: Jonathan Davidson & James Peters. *Fundamentos de voz sobre IP*. Pág. 20.

- b) Capa de control de llamadas establecidas: dirige el control de llamadas, el cual es el proceso de tomar una decisión de enrutamiento o señalización sobre donde necesita ir una llamada y como hacer que la llamada ocurra, en el modelo de circuito conmutado esto se hace por el SS7 (Sistema de señalización 7). Utiliza protocolos de control de llamadas los cuales indican al flujo RTP donde terminar y donde empezar; también traducen una numeración telefónica en dirección IP. Como principales protocolos de esta capa se tienen: H.323, SGCP (*Simple Gateway Control Protocol*), IPDC (*Internet Protocol Device Control*), MGCP (*Media Gateway Control Protocol*), SIP (*Session Initiation Protocol*).

- c) Capa de aplicaciones de servicios abiertos: En esta capa se conectan los servicios o aplicaciones para que el usuario tenga una interfaz con el sistema, así esta infraestructura es aprovechada al máximo. Ejemplo de aplicaciones tenemos el centro de llamadas, llamadas en espera o desvío de llamadas, llamada por *Internet*, teniendo con todo esto un servicio de mensajería unificada.

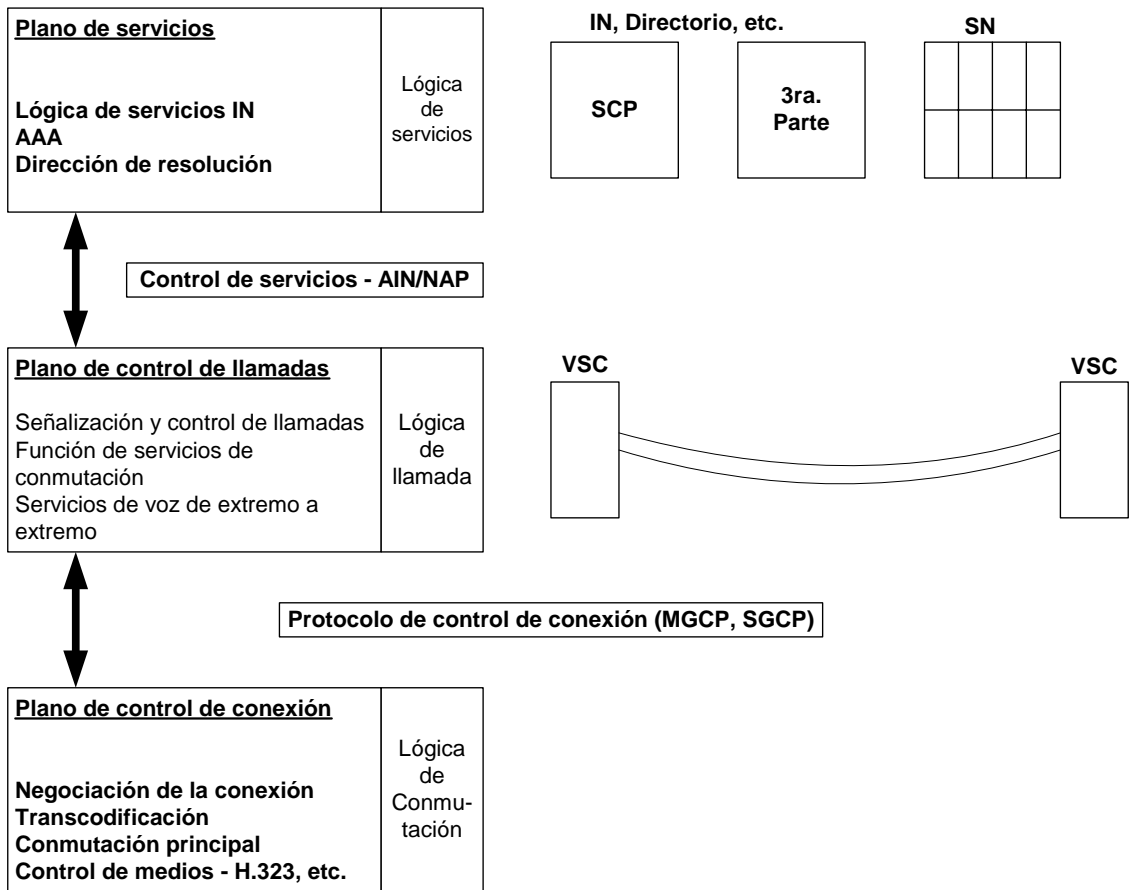
Ya que lo que se requiere es intercomunicar las redes de telefonía tradicional con las redes de datos, se utiliza un conmutador virtual que está basado en el nuevo modelo descrito anteriormente, este equipo empaqueta las conversaciones y las envía vía una *Lan Ethernet*, la cual tiene mayor calidad de servicio y seguridad.

El conmutador virtual, como en la telefonía de paquetes se basa en tres planos lógicos, control de conexión, control de llamadas y control de servicios, como el que se muestra en la figura 59. Cada plano representa un aspecto funcional diferente e interactúa con los otros planos a través de interfases abiertas bien definidas.

Estos planos están organizados jerárquicamente, con el plano de control de conexión en el nivel más bajo, este plano abarca las funciones de configurar, mantener y borrar rutas de voz a través de la red de paquetes. El plano siguiente es el de control de llamadas este se encarga de señalar, procesar y enrutar llamadas de voz y datos sobre la red de paquetes, sus funciones se asemejan a las que se encuentran en la lógica del procesamiento de llamadas de un conmutador *TDM*. Las funciones típicas del plano de control de llamadas incluyen el protocolo del sistema de señalización (*SS7*), el análisis y manipulación de dígitos, la selección de ruta, el seguimiento, las funciones basadas en conmutador y la comunicación con programas lógicos de servicios externos. Esta organización de los planos permite que el plano de control de llamadas sea independiente del transporte de paquetes de voz subyacente, permitiendo que se utilice el mismo plano de control de llamadas con *gateway* de medios (MG) orientado a la capa dos y/o tres del modelo *OSI* de referencia sobre *VoIP*, mostrado en la figura 60. Este

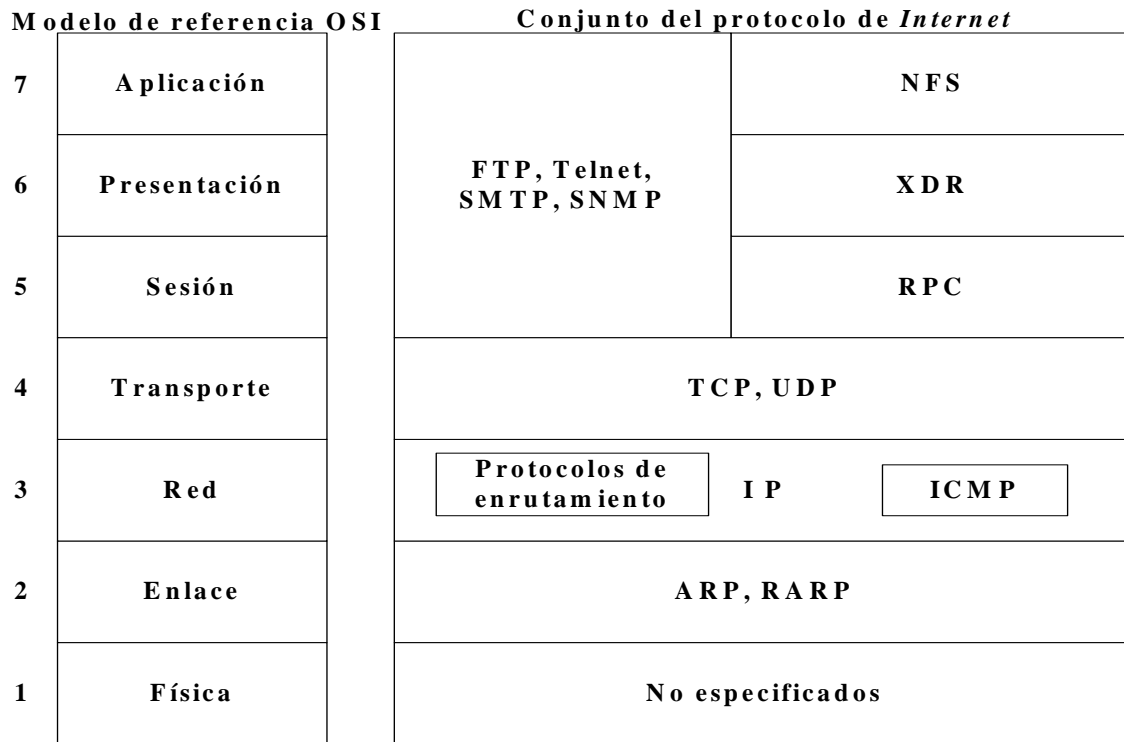
plano también comunica con el plano de servicios para proporcionar servicios flexibles mejorados, esta interfaz es típicamente un protocolo de red inteligente (IN) basado en estándares que funcionan sobre *TCAP* del sistema de señalización (*SS7*), esto se realiza teniendo puntos de control de servicio (*SCP*) o nodos de servicio.

Figura 59. Arquitectura de telefonía de paquetes abierta.



Fuente: Jonathan Davidson & James Peters. *Fundamentos de voz sobre IP*. Pág. 280.

Figura 60. Modelo de referencia OSI sobre VoIP



Fuente: *Jonathan Davidson & James Peters. Fundamentos de voz sobre IP. Pág. 146.*

El conmutador virtual se compone de los siguientes elementos:

- Controlador de conmutación virtual (VCS);
- Gateway de medios (MG);
- Punto de control de servicio (SCP);
- Nodo de servicios;
- H.323.

4.2.1.2.1. Controlador de conmutación virtual (VCS)

La función principal es el control de llamadas, básicamente funciona de la siguiente forma:

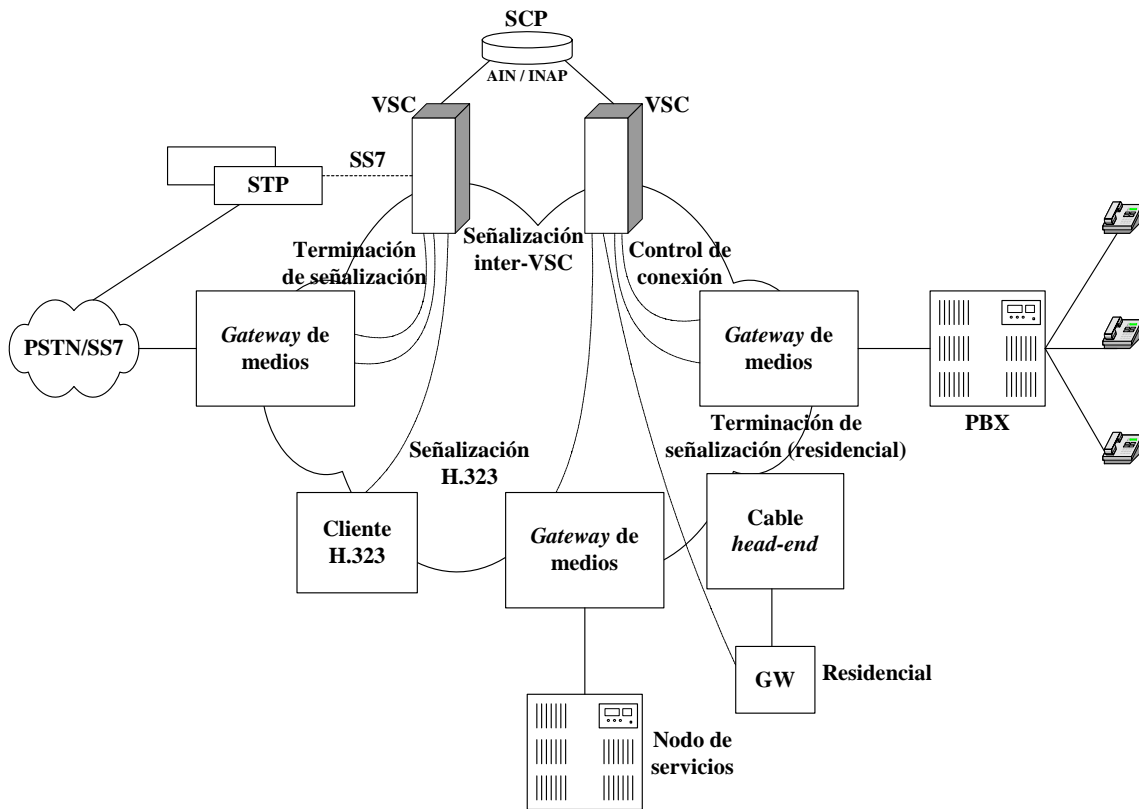
Se tiene una red existente *PSTN / SS7* que esta conectada a un punto de transferencia de conmutación (*STP, Switching Transfer Point*), y este a su vez está conectado al controlador de conmutación virtual (*VSC*) es aquí donde tiene lugar la señalización *SS7*. La red *PSTN / SS7* también está conectada a un *gateway* de medios (*MG*) que es un enlace troncal de menor señalización conocido como enlace troncal intermáquina (*IMT Inter. – Machine Trunk*). En el *MG* es donde los enlaces troncales de voz de 64 kbps son convertidos en paquetes y colocados en la red *IP*. El *VSC* y los *MG* también se comunican entre si, para indicar cuándo un *MG* debe configurar las conexiones *IP* y con que otros *MG* las deben configurar, este funcionamiento se puede observar en la figura 61.

A continuación se resumen sus principales funciones:

- Procesamiento de la señal de llamada, incluido el nivel tres de la *ISDN Q.931*, el nivel cuatro de la *SS7*, señalización *H.323* asociada a canal / multifrecuencia (*MF / CAS*), señalización de llamadas hacia dispositivos ubicados en *gateway's* residenciales. Traducción entre diferentes tipos de señalizaciones en distintos establecimientos de llamadas.
- Resolución de direcciones, enrutamiento de llamadas, administración de recursos, control de conexión y generación del registro detallado de llamadas.
- Funciones de acceso a los servicios para acceder a servicios que se ejecutan en plataformas de servidores externos (como *SCP* o nodos de servicio).

- Interfaces de administración que utilizan el protocolo *SNMP* para errores, rendimiento y configuración; herramientas de configuración basadas en la *web* y sistemas de administración de elementos.

Figura 61. Arquitectura de red de voz por paquetes



Fuente: Jonathan Davidson & James Peters. Fundamentos de voz sobre IP. Pág. 282.

4.2.1.2.2. Puerta de enlace o *gateway* de medios (*MG*)

A continuación se resumen las principales funciones de un *gateway* de medios (*MG*, *Media Gateway*):

- Terminación física de la instalación *TI/ E1 TDM* desde la *PSTN* o *PBX*.
- Comunicación con el *VSC* para la configuración y borrado de la llamada utilizando *SGCP* o *MGCP*.
- Cancelación de eco en la red de circuito conmutado.
- Equilibrado de los búferes de fluctuación de fase.
- Detección de actividad de voz, como la supresión de silencio y la generación de ruido de apaciguamiento.
- Compresión de voz utilizando las recomendaciones de la Unión Internacional de Telecomunicaciones (ITU), como G.711, G.723.1 y G.729.
- Generación de tono, que genera los tonos de marcado, ocupado, *ring back* y congestión.
- Transporte de marcación multifrecuencia (*DTMF*), que permite la utilización de señal de tonos para aplicaciones de correo de voz con codificadores que soportan la detección y transporte *DTMF*.
- Transcodificación de la ley μ y la ley a cuando se necesite.

4.2.1.2.3. Punto de control de servicio (*SCP*)

Es el responsable de procesar las peticiones de transacción y de devolver una respuesta. Una petición de transacción típica en el mundo de la voz es una conversión de número.

4.2.1.2.4. Nodo de servicios

Cuando se utilizan nodos de servicios, el VSC enruta normalmente una llamada al nodo de servicios para el procesamiento. El nodo de servicios aplica luego su propio tratamiento de la función específica a la voz y al flujo de datos y completa el enrutamiento de llamadas al destino deseado. Dependiendo de la función, el nodo de servicios puede permanecer en la ruta de la llamada o dar el control de la llamada de nuevo al VSC. Las aplicaciones típicas del nodo de servicios incluyen correo de voz y la marcación por voz.

4.2.1.2.5. H.323

Es una especificación de la ITU para transmitir audio, video y datos a través de una red de protocolo *Internet* (IP), incluida la propia *Internet*. El H.323 estándar dirige la señalización y control de llamadas, transporte y control multimedia y control del ancho de banda para conferencias punto a punto y multipunto.

El H.323 estándar consta de funciones y protocolos, como a continuación se muestra en la tabla IV.

Tabla IV. Componentes H.323 estándar

Función	Protocolo
Señalización de llamadas	H.225
Control de medios	H.245
Codificadores de audio	G.711, G.722, G.723, G.728, G.729
Codificadores de video	H.261, H.263
Compartir datos	T.120
Transporte de medios	RTP / RTCP

Fuente: Jonathan Davidson & James Peters. Fundamentos de voz sobre IP. Pág. 223.

El H.323^[4] consta de varios elementos que aquí solo se mencionarán y no se entrará en detalle, estos elementos utilizan los protocolos antes descritos en la tabla IV para proporcionar un flujo de llamada entre dos puntos. Los elementos del H.323 incluyen terminales, *gateway's*, *gatekeeper's*, y unidades de control multipunto (*MCU*, *Multipoint Control Units*).

Luego de haber revisado los componentes más importantes de un conmutador virtual, el cual se basa en una serie de elementos y estándares y se apoya en una serie de protocolos que cubre los distintos aspectos de la comunicación. A continuación se presenta un resumen de las funciones que realiza y protocolos utilizados en la proceso de una llamada, también se muestra en la tabla V la pila de protocolos en *VoIP*.

1. Direccionamiento:

- a. *RAS (Registration, Admisión and Status)*: permite a una estación H.323 localizar otra estación H.323 a través de un *Gatekeeper*.
- b. *DNS (Domain Name Service)*: Servicio de resolución de nombres en direcciones *IP* con el mismo fin que el protocolo *RAS* pero a través de un servidor *DNS*.

2. Señalización:

- a. Q.931 (Mensajes de control de llamadas): señalización inicial de llamada entre dos puntos, para el propósito de conectar, mantener y desconectar las llamadas.
- b. H.225 (Control de llamadas): señalización, registro, admisión, paquetización y sincronización del flujo de voz.
- c. H.225 (Control de medios): especifica mensajes de apertura y cierre de canales para flujos de voz.

3. Compresión de voz: para llevar a cabo esto se necesitan mensajes que intercambian de manera segura las capacidades entre dos terminales, estos indican las capacidades del terminal para transmitir y recibir audio, vídeo y datos. Para audio el intercambio de capacidades incluye codificadores de transcodificación de voz de la serie G.

- a. Requeridos: G.711 a 64 kbps y G.723 a 5.3 ó 6.3 kbps.
- b. Opcionales: G.728 a 16 kbps, G.729 a 8kbps y G.722 a 48, 56 y 64 kbps.

4. Transmisión de voz:

- a. *UDP (User Datagram Protocol)*: la transmisión se realiza sobre paquetes *UDP*, pues aunque *UDP* no ofrece integridad en los datos, el aprovechamiento del ancho de banda es mayor que con *TCP*.
- b. *RTP (Real Time Protocol)*: Maneja los aspectos relativos a la temporización, marcando los paquetes *UDP* con la información necesaria para la correcta entrega de los mismos en recepción.

5. Control de la Transmisión:

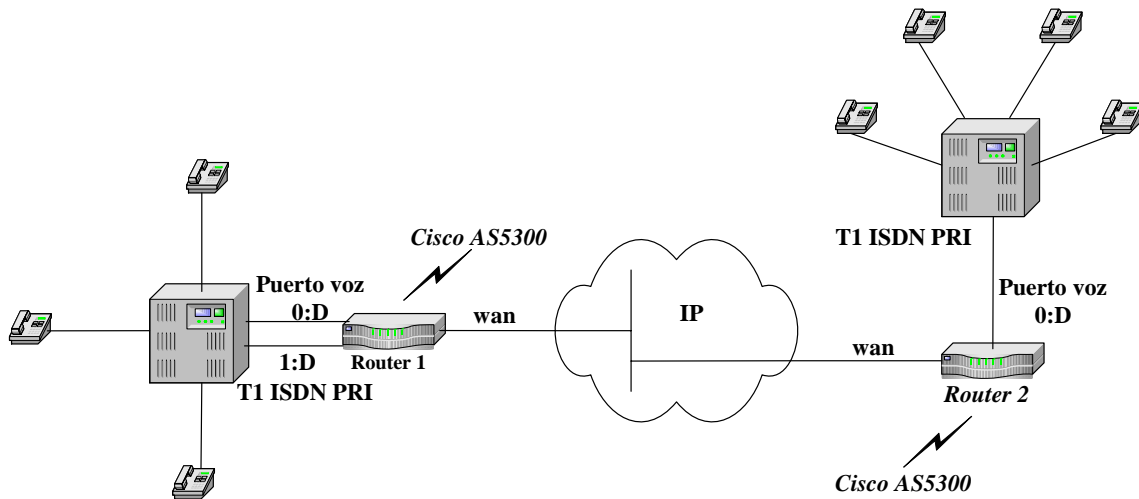
- a. *RTCP (Real Time Control Protocol)*: se utiliza principalmente para detectar situaciones de congestión de la red y tomar, en su caso acciones correctoras.

Tabla V. Pila de protocolos en VoIP

Establecimiento de llamada y control											
<table border="1" style="margin: auto;"> <tr> <td colspan="6" style="text-align: center;">Presentación</td> </tr> </table>						Presentación					
Presentación											
Direccionamiento		Compresión de audio G.711 ó G.723		DTMF							
RAS (H.225)	DNS	RTP / RTCP		H.245	Q.931(H.225)						
Transporte UDP				Transporte TCP							
Red (IP)											
Enlace											
Físico											

Después de haber visto el funcionamiento de *VoIP* y el controlador de conmutación virtual, para la interconexión de la red telefónica a *IP* se utilizará un equipo de *Cisco* llamado *Access Server 5300* que cumple con todas las características antes mencionadas, ya que incluye un módulo dedicado que realiza las tareas de *gateway* y viceversa; en la figura 62 se muestra un modelo de cómo se utiliza, este permite que en un ancho de 10 kbps se pueda transportar el equivalente a 64 kbps de una línea telefónica tradicional con buena calidad.

Figura 62. Modelo de interconexión de la red telefónica a la red IP.



A continuación se describe la configuración del equipo AS5850-G-SK para que pueda funcionar en esta red, se describen los componentes más importante, pero también hay que tomar en cuenta licencias para datos, *upgrades* y otros aspectos que son fundamentales en el costo y que se consideran a la hora de elegir el equipo.

- Equipo AS5850, interfaz Gigabit
- Cable de alimentación AC para AS5800 (120 / 240 VAC)
- Tres cables de alimentación AC
- Tarjeta *Router switch controller* para AS5850
- Interfaz Gigabit, con convertidor 1000Base-SX (longitud de onda corta, *Short Wavelength*)
- Proveedor *plus* IOS de la serie de “Cisco” 5850
- Administrador de puertos IOS serie AS5000
- IOS SS7 serie AS5000
- Tarjeta canalizadora para T3 más 216 puertos universales, serie AS5800
- Tarjeta de puerto 324, serie AS5800

4.2.1.3. Centro de llamadas de telefonía de paquetes (*PTCC Packet Telephony Call Center*)

Lo que se trata es que este sistema de mensajería sea vanguardista, y permitirle al usuario un servicio al cliente más rápido, es por eso, que en lugar de una PBX, se tiene un centro de llamadas por paquetes (*PTCC Packet Telephony Call Center*) el cual brinda una integración de la telefonía en la nueva red de soporte *web*, *VoIP* y comunicaciones unificadas.

Este centro de llamadas se compone de un procesador de llamadas y un *Call Center*, los cuales están conectados a un *Call Manager* a través de enlaces de telefonía y de computadoras, la figura 63 muestra los componentes y diseño de un centro de llamadas PTCC.

El procesador de llamadas permite que las llamadas de los usuarios sean respondidas automáticamente y que estos puedan introducir y obtener información básica; para esto se utilizan funciones mejoradas de sistemas como, *IVR* basados en *IP* (también conocidos como unidades de respuesta de voz *VRU*, [*Voice Response Units*]) y servicios de mensajería unificada como, fax a correo electrónico, voz a texto, texto a texto, para su enrutamiento final, lo cual hace que las llamadas sean manejadas con mayor rapidez.

El *Call Center* permite que teletrabajadores o agentes de llamadas respondan a las llamadas con interrogantes de los abonados y usuarios, teniendo en pantalla los datos de los mismos, de tal manera que el agente pueda manejar la llamada de una forma más rápida y eficientemente. También evita que al haber una transferencia de la llamada a un nuevo agente, este pueda hacer la misma pregunta; esto es posible gracias a que la información del primer agente es volcada en la pantalla del segundo cuando la llamada le es transferida. Es necesario notar que el abonado también puede hacer su consulta por la red en línea con el agente. El acceso a una infraestructura común otorga a todos las mismas condiciones y da al usuario una idea y sentimiento común, como muestra la figura 64.

Figura 63. Centro de llamadas de telefonía de paquetes.

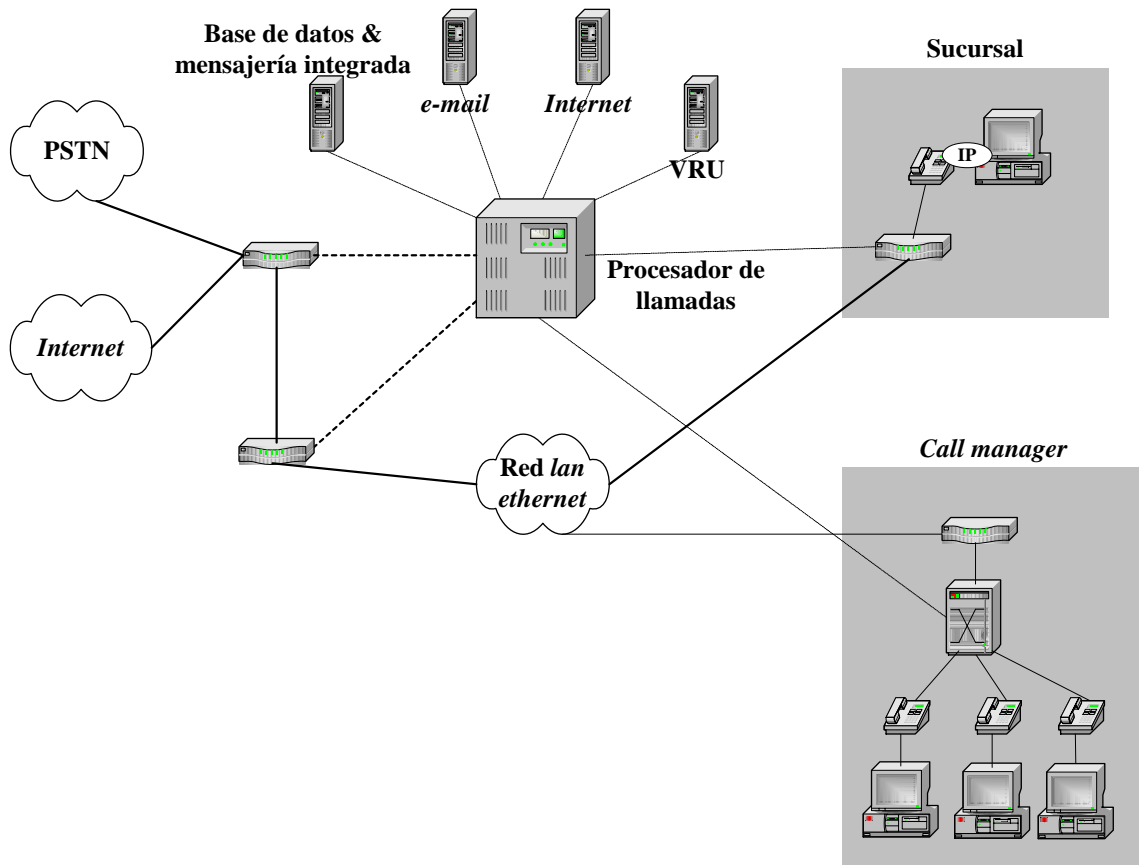
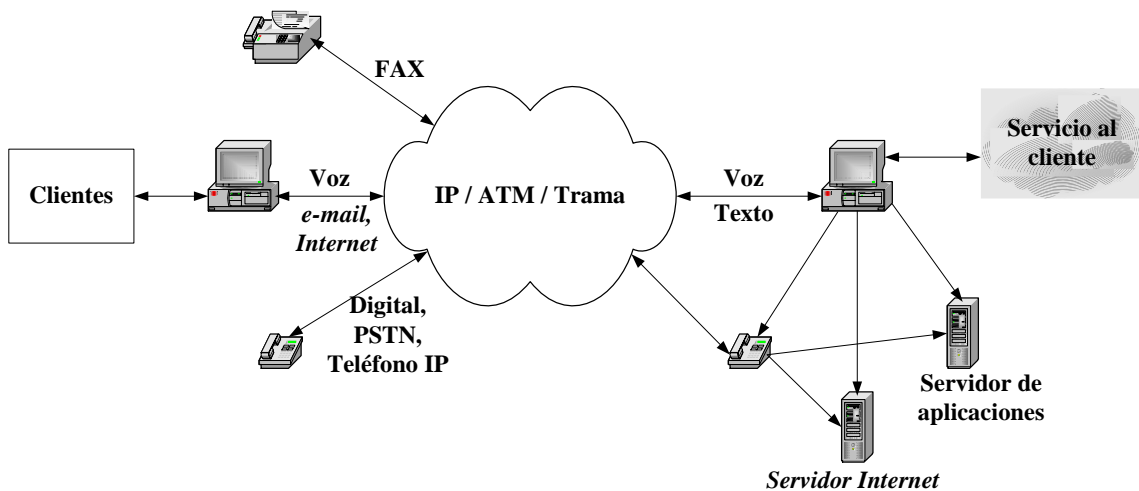


Figura 64. Infraestructura común para los agentes de llamadas.



4.2.2. Unidad de Control

La unidad de control constituye la interfaz inteligente entre la red de conmutación o fuentes de entrada y el propio sistema de mensajería. Su misión principal es la de recepción de mensajes de voz que luego son convertidos a texto para su codificación de acuerdo a un protocolo determinado situándolos en la cola y luego los envía a los transmisores para su difusión. Así también como el control y supervisión de la red radioeléctrica.

A grandes rasgos la unidad de control debe tener estas características:

- Una base de datos que contenga las informaciones generales de los abonados y sus relaciones con los servicios y zonas de coberturas de cada uno de ellos.
- Servidores de aplicaciones especiales de nuevas tecnologías, reconocimiento de voz, servicios de *Internet*, correo electrónico, etc.
- Base de datos de memorización de mensajes.
- Módulo de estadísticas de mensajes.
- Control y supervisión de la red de transmisión
- Control de encendido y apagado de los transmisores.
- *Software* de detección y diagnóstico de fallos.
- *Software* de codificación.
- Soporte del sistema por medio de una red *Lan*, ya que se necesita compartir recursos.

De acuerdo a las características antes mencionadas, en este apartado se describirá la arquitectura de la red *Lan* que soportará al sistema, las características de los servidores con funciones especiales, la terminal de paginación, protocolos del sistema, base de datos del sistema, y los concentradores o controladores de radiolocalización.

4.2.2.1. Arquitectura red *Lan*

Una red *Ethernet* es una buena elección para trabajar la arquitectura de la red *Lan*, debido al bajo costo y al alto desempeño. Además las redes *Ethernet* son compatibles con los protocolos *TCP / IP*, *UDP / IP* e *IP*, con lo cual ayuda en la aplicación de *VoIP*. Esta red tendrá sus propios servidores de aplicaciones especiales y de los cuales se hablará más adelante.

Ethernet es actualmente la arquitectura de redes más popular. Esta arquitectura de banda base utiliza la topología de bus o bus en estrella, suele transmitir a 10 Mbps y utiliza el método de acceso en *CSMA / CD* para regular el tráfico en el segmento principal del cable.

El medio de *Ethernet* es pasivo, es decir, su fuente de alimentación es el equipo y, por tanto, no falla a menos que el medio se corte físicamente o se termine incorrectamente. En la tabla VI se resumen las características generales de *Ethernet*.

Tabla VI. Características generales de la red *Lan Ethernet*

Topología tradicional:	Bus lineal
Otras topologías:	Bus en estrella
Tipo de arquitectura:	Banda base
Método de acceso:	<i>CSMA / CD</i>
Especificaciones:	IEEE 802.3
Velocidad de transmisión:	10 Mbps o 100 Mbps
Tipos de cables	Cable delgado, cable grueso, <i>UTP</i>

En este diseño se tendrá una red *Lan Ethernet* de 10BaseT, con una topología de bus en estrella, es una red *Ethernet* con una velocidad de transmisión de 10 Mbps, que suele utilizar un cable UTP (par trenzado no blindado) para conectar los equipos. Este tipo de red requiere que cada terminal de usuario tenga un adaptador de red, el cual es una tarjeta que se conecta en un puerto ISA o PCI. La longitud máxima de un segmento 10BaseT es de 100 metros y la longitud mínima de cable entre dos equipos es de 2.5 mts., este tipo de red puede dar servicio a 1024 equipos. En la tabla VII se resumen las características de la topología 10BaseT.

Tabla VII. Características topología 10BaseT

Categoría	Especificación
Cable:	Categorías 3, 4 o 5 UTP
Conectores:	RJ – 45 al final de los cables
Transceptor:	Cada equipo requiere uno; en algunas tarjetas ya están incorporados
Distancia del transceptor al concentrador:	100 metros como máximo
Total de equipos por red sin componentes de conectividad:	1024 según la especificación
Número máximo de segmentos conectados:	5 (utilizando 4 repetidores); solo tres segmentos pueden tener equipos conectados

A continuación se dan las características de los equipos o pc's de los usuarios de la red.

- Procesador: *Intel Pentium 4* ó su equivalente; a 2 GHz.
- Memoria: 128 MB.
- Disco duro: 20 GB.
- Unidades externas de almacenamiento: Disco 3.5 y *CD-ROM*
- Monitor: *SVGA* de 14"
- Sistema operativo: *Unix*
- Tarjeta de red: Gigabit (10/100/1000) *Ethernet*
- Periféricos: *Modem* a 9600, *mouse*, puerto serial

4.2.2.2. Servidores de aplicaciones especiales

Debido a que el sistema necesita compartir recursos, se utilizarán servidores especializados en aplicaciones, estos se optimizan para dar un servicio rápido y seguro en ciertas tareas importantes del sistema, como la base de datos, reconocimiento de voz, servicio de correo electrónico y servicio de *Internet*. En este apartado se darán las características que deben de tener estos servidores y en el caso de la base datos los campos necesarios para su desarrollo.

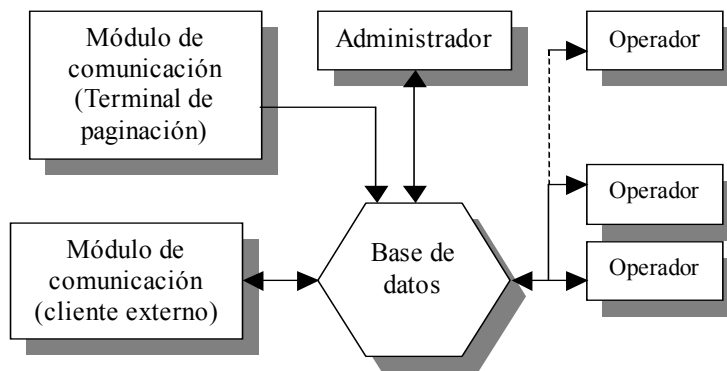
4.2.2.2.1. Servidor de base de datos y reconocimiento de voz

Este es el encargado de tener la base de datos y el software de reconocimiento de voz, para su fácil acceso.

4.2.2.2.1.1. Base de datos

Es la encargada de soportar todas las entradas de información, seguridad en la entrega de los datos y distribución en la red; y debido a la interoperabilidad del sistema esta debe tener interrelaciones entre diferentes elementos del mismo, este modelo se muestra en la figura 65.

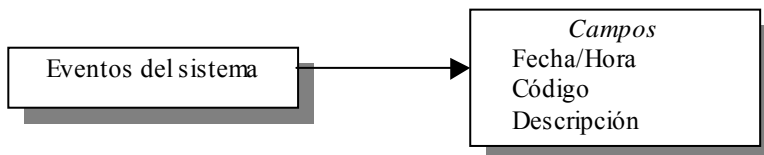
Figura 65. Interrelación de elementos principales de la red



La base de datos debe estar constituida por las siguientes partes:

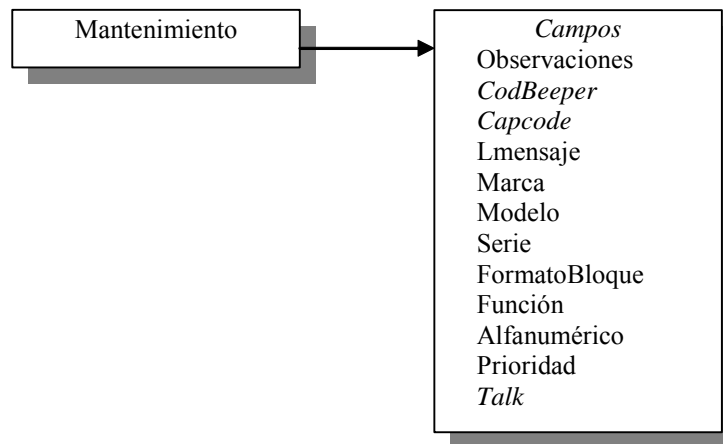
1. Eventos del sistema: esta parte se encarga del almacenar los acontecimientos del sistema, como son: a) entradas y salidas en los módulos cliente, administración y de comunicaciones, b) cambios en los datos de abonados y claves, y c) eliminación de información. La estructura se muestra en la figura 66.

Figura 66. Módulo eventos del sistema

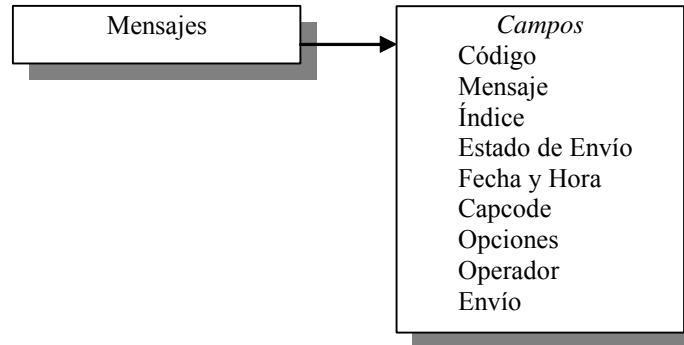


2. Mantenimiento: se encarga de los radiolocalizadores que están en reparación o mantenimiento, llevando un registro de estos equipos se logra tener un mejor control de daños. Los campos de esta parte se muestran en la figura 67.

Figura 67. Módulo de mantenimiento de radiolocalizadores



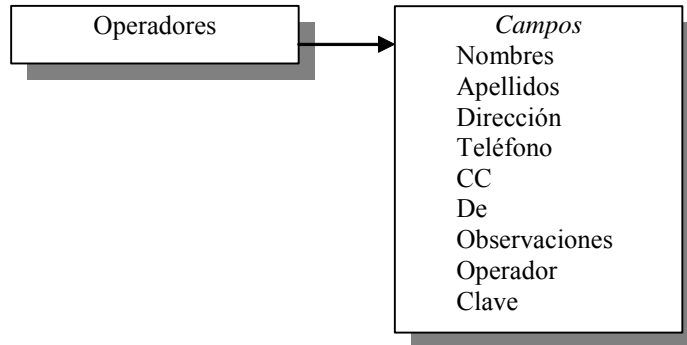
3. Mensajes: Es la encargada de almacenar los datos referentes a todos los mensajes que van a ser enviados a través del módulo de comunicación del terminal de paginación, los mensajes son grabados originalmente con el estado de envío igual a NO, esto significa que estos mensajes están esperando para ser enviados, el módulo de comunicaciones periódicamente busca los mensajes no enviados o con estado igual a NO para luego distribuirlos a través de la terminal de paginación. En esta tabla se guardan los siguientes datos: código del abonado, mensaje tipo alfanumérico, índice de la cola de mensajes (sirve para llevar un consecutivo de mensajes), estado de envío del mensaje, fecha y hora del envío, *capcode* del abonado, opciones (como son el formato de bloque, prioridad, *talk*, alfanumérico y función), operador quien envía el mensaje (por ser el sistema automático, solo se utilizará de emergente), y envío (cuantas veces a sido reenviado el mensaje); en la figura 68 se muestran los campos de este bloque.

Figura 68. Módulo de mensajes

4. Operadores: Esta información corresponde a las personas autorizadas (en caso de emergencia) para enviar mensajes a los abonados conectados al sistema, cada vez que se ingresa un mensaje para un abonado a la base de datos, se verifica el nombre o identificación del operador con su respectiva clave, si la identificación o clave no es correcta no se permite el acceso al sistema. En la figura 69 se muestran los campos de esta tabla. Los operadores cuando ingresen al sistema podrán realizar las siguiente labores:

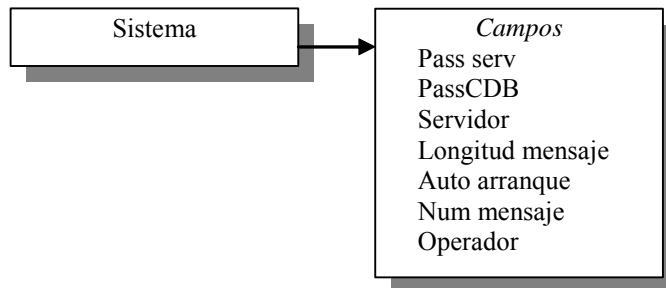
- Enviar mensajes a un abonado o grupo de abonados.
- Buscar información acerca del código, nombre, apellido de un abonado registrado en la base de datos.
- Reenviar mensajes que ya han sido transmitidos al abonado.

Figura 69. Módulo de operadores



5. Sistema: En este lugar la base de datos deposita información general acerca del sistema y cómo debe operar, hay datos tales como la clave del equipo servidor de mensajes, clave administrativa, indicación si el servidor de mensajes se encuentra activo, longitud de mensajes, autoarranque del servidor de mensajes. Los campos se muestra en la figura 70.

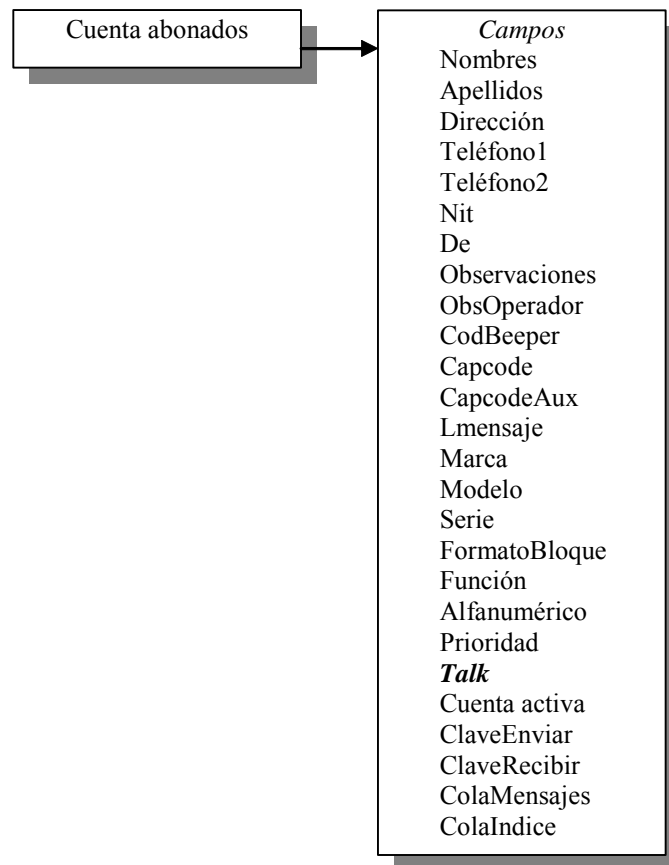
Figura 70. Módulo sistema



6. Abonados: En esta tabla se encuentra la información de los abonados del sistema, estos datos se pueden dividir en varias clases, primero se encuentran los datos generales del abonado: nombres, apellidos, teléfono1, teléfono2, Nit, observaciones y las observaciones del abonado para el operador; en una segunda parte se colocan los datos del radiolocalizador: código radiolocalizador, *capcode*, *capcode* auxiliar, longitud del mensaje, marca del radiolocalizador, modelo y serie. En la tercera parte que se encuentra dividido están los datos técnicos: formato de

bloque, función, alfanumérico, prioridad y *talk*; y en la última parte se encuentra información referente a la cuenta: cuenta activa, clave para enviar mensajes, clave para recibir mensajes, tamaño de la cola de mensajes y el índice de la cola de mensajes. En la figura 71 se muestran estos campos.

Figura 71. Módulo cuenta abonados



A continuación se describen los módulos del sistema con los cuales la base de datos tiene una interrelación:

- a. Módulo de administración: es el encargado de coordinar los diferentes módulos del sistema, además a esta parte se le ha asignado labores que solo competen al administrador del sistema.

Entre las tareas del administrador se encuentran:

- Permitir la adición, modificación o borrado de registros tanto de abonados, operadores, o radiolocalizadores en mantenimiento.
 - Crear estadísticas de envíos de mensajes como por ejemplo:
 - ◆ Mes del año: mensajes enviados por mes, durante un año.
 - ◆ Día del mes: mensajes enviados por día, durante un mes.
 - ◆ Días de la semana: mensajes enviados por día, durante una semana.
 - ◆ Hora del día: mensajes enviados en una hora, durante las veinticuatro horas del día.
 - Impresión de reportes o informes de los registros de la base de datos, se puede imprimir información de abonados, operadores, mantenimiento y estadísticas.
 - Llevar controles de los servidores conectados a la terminal de paginación.
- b. Módulo operador: el trabajo primordial es de distribuir los mensajes de los usuarios hasta la base de datos para luego ser enviados al abonado correspondiente, cabe aclarar que este trabajo solo se realizará en el momento de tener una emergencia cuando el sistema no pueda ser automático. Otro de los trabajos de este módulo es atender las consultas e interrogantes del abonado, sirviendo como un *Call Center*.

- c. Módulo de comunicación con la terminal de paginación: este módulo sirve de interfaz entre los mensajes a enviar y el protocolo del terminal de radiolocalización, entre las funciones de este módulo se encuentran:
- Verificar que el terminal de radiolocalización se encuentre presente, o sea que se encuentre conectado al equipo y encendido para su correcto funcionamiento.
 - Conectar al terminal de radiolocalización y colocar en la tabla sistema el registro servidor en SI, que da la opción de comenzar a enviar mensajes.
 - El servidor busca periódicamente mensajes sin enviar, los codifica en el protocolo que maneja el terminal de radiolocalización y los despacha, luego hace el cambio respectivo en la base de datos, o sea consigna en la base de datos que el mensaje ha sido enviado cambiando el valor del registro con estado de envío igual a SI.
 - El servidor también verifica que el mensaje no sea posfechado, o sea se pueden programar mensajes para que se envíen a una hora y fecha programada, esto se logra cuando el servidor compara la fecha actual con la del mensaje, si la fecha es mayor que la actual el servidor no envía este mensaje, solo se distribuye el mensaje cuando la fecha del mensaje sea menor o igual que la presente y el estado de envío es igual a NO.
- d. Módulo de comunicación cliente externo: debido a la estructura de esta red, se ofrece una conexión vía *Internet*, *módem* y por *IP*. De esta forma el usuario podrá enviar mensajes directos sin pasar a través de la red de conmutación. Este módulo está equipado con una interfaz de *software* el cual permite hacer la conexión con el servidor de *Internet*, un *módem*, etc.

Entre las funciones de este módulo se encuentran:

- Proveer un enlace constante entre los usuarios externos (quienes se conectan a través de *Internet, módem, o IP*) y el servidor de mensajes el cual se encarga de conectarse con la base de datos y enviar los mensajes.
- Verificar la integridad de los datos del abonado en el envío de mensajes.
- Manejo de seguridad de los clientes externos, así no habrán intromisiones de personas no autorizadas para el envío de mensajes.

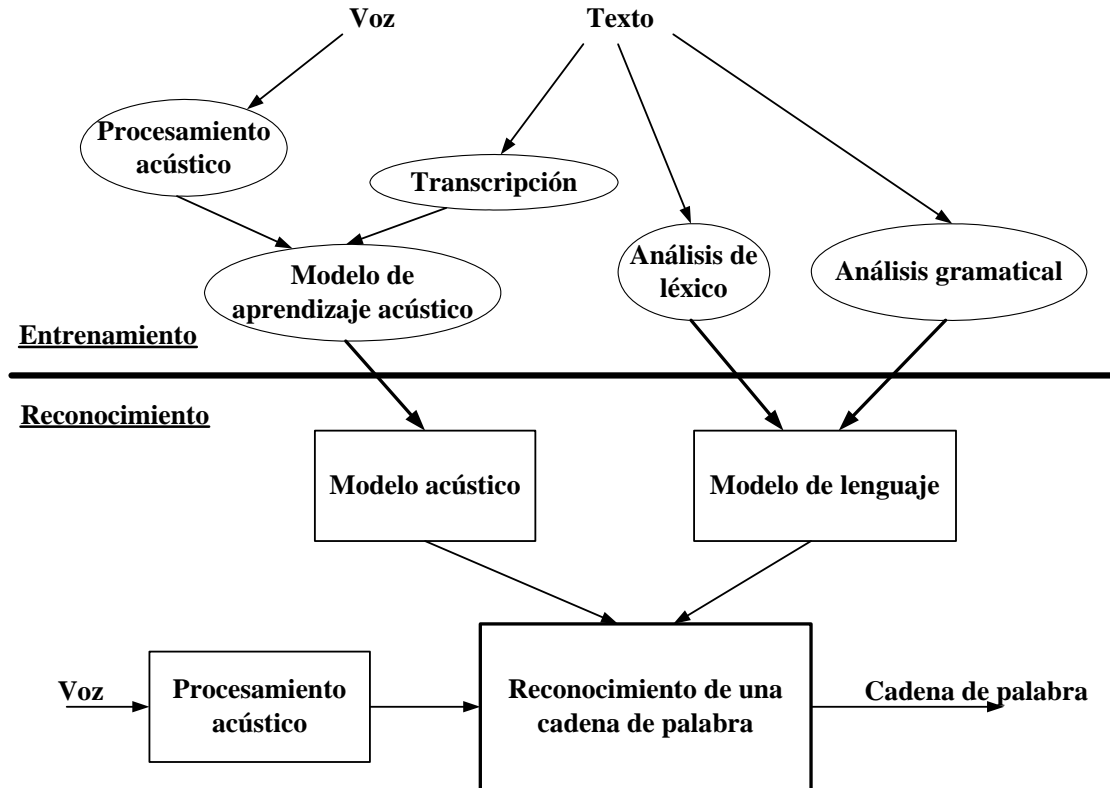
4.2.2.2.1.2. Reconocimiento de voz

En este apartado, se hablará más expresamente de las características que tiene un buen reconocedor de voz y que determinan su elección, así también de cómo interactúa con la base de datos; no se entrará en detalle ya que en el capítulo número dos se anotó extensamente de cómo este se produce y dota a las máquinas de la capacidad de recibir mensajes orales y convertirlos a texto.

Básicamente, el reconocimiento de voz es un proceso de clasificación de patrones, cuyo objetivo es clasificar la señal de entrada (onda acústica) en una secuencia de patrones previamente aprendidos y almacenados en unos diccionarios de modelos acústicos y de lenguaje. Este proceso de clasificación supone, en primer lugar que la señal de voz puede ser analizada en segmentos de corta duración y representar cada uno de los segmentos mediante su contenido frecuencial, de forma análoga al funcionamiento del oído. En segundo lugar, que mediante un proceso de clasificación podemos asignar a cada segmento o conjuntos consecutivos de segmentos una unidad con significado lingüístico. En tercer lugar, que mediante un procesador lingüístico podemos dar significado a las secuencias de unidades. Este último paso del sistema, supone incorporar al sistema de reconocimiento de voz, conocimiento acerca de la

estructura sintáctica, semántica y pragmática del lenguaje. Todas estas características se muestran en la figura 72.

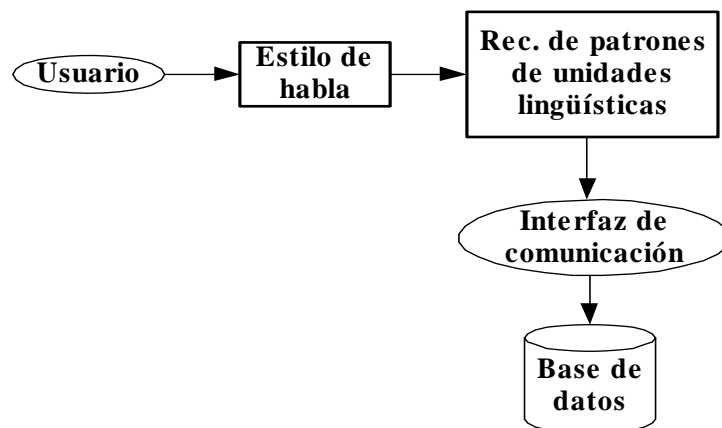
Figura 72. Características en el proceso de reconocimiento de voz.



En la figura 73 se muestra un diagrama del modelo de comunicación para reconocimiento de voz, el acceso a la información contenida en una base de datos comienza con la producción de un mensaje hablado por el usuario, pero utilizando una forma o estilo de habla restringido (en este época, esperando que en un futuro si se pueda dar el reconocimiento en un 100%). A partir de la señal de voz, un proceso de clasificación, basado en reconocimiento de patrones asociados a diferentes unidades lingüísticas (palabras, fonemas, sílabas, etc.), permite a una interfaz de comunicaciones proporcionar a la base de datos la información que el usuario necesite enviar. Cabe mencionar que en este momento sólo se encontró en el ámbito de las telecomunicaciones

tres aplicaciones que usan reconocimiento de voz, marcación por tono en la red privada, reconocimiento de letras (palabras aisladas independiente del locutor) y reconocedor de palabras aisladas dependiente del locutor aplicado a la marcación de números de teléfonos por voz.

Figura 73. Modelo de comunicación para el reconocimiento de voz



4.2.2.2.1.3. Características del *hardware* del servidor

Este servidor se estipula para la base de datos del sistema de radiolocalización y el software de reconocimiento de voz, a continuación se dan sus características:

- Procesador: *Intel Pentium 4* ó su equivalente; a 2 GHz.
- Memoria: 1.5 GB.
- Disco Duro: 20 GB.
- Unidades externas de almacenamiento: Disco 3.5 y *CD-ROM*.
- Monitor: *SVGA* de 14".
- Sistema Operativo: *Unix*
- Tarjeta de red: Gigabit (10/100/1000) *Ethernet*.
- Periféricos: *Módem* a 9600, ratón, puertos seriales y paralelos.

4.2.2.2. Servidor correo electrónico

Ya que uno de los servicios es la distribución de mensajes a través de la red, la función de este es guardar los correos electrónicos y luego descargarlos cuando sean requeridos, ya sea en un PC o en un radiolocalizador. Básicamente puede tener la siguiente configuración:

- Procesador: *Intel Pentium 4* ó su equivalente; a 2 GHz.
- Memoria: 1.5 GB.
- Discos Duros: de 20 GB, cada uno maneja 1500 usuarios.
- Unidades externas de almacenamiento: Disco 3.5 y *CD-ROM*.
- Monitor: *SVGA* de 14".
- Sistema Operativo: *Unix*
- Software de correo electrónico: *send mail*.
- Tarjeta de red: Gigabit (10/100/1000) *Ethernet*.
- Periféricos: *Módem* a 9600, ratón, puertos seriales y paralelos.

4.2.2.3. Servidor *Internet*

Este, básicamente sirve para prestar el servicio de página *web*, información general a abonados y usuarios de los servicios prestados y también se encarga de tomar datos de la página como son el código del abonado, mensaje, verificar que existan y luego distribuir el mensaje al radiolocalizador destinado por el usuario. A continuación se da la configuración que puede tener:

- Procesador: *Intel Pentium 4* ó su equivalente; a 2 GHz.
- Memoria: 1.5 GB.
- Disco Duro: 20 GB.

- Unidades externas de almacenamiento: Disco 3.5 y *CD-ROM*.
- Monitor: *SVGA* de 14”.
- Sistema Operativo: *Unix*
- Tarjeta de red: Gigabit (10/100/1000) *Ethernet*.
- Periféricos: *Módem* a 9600, ratón, puertos seriales y paralelos.

Cabe hacer notar que por seguridad para *Internet* se requiere de un *firewall* basado en *software*; este debe estar conectado a el servidor de correo electrónico y al servidor de *Internet* para lograr protegerlos.

4.2.2.3. Terminal de radiolocalización

En un sistema de radiolocalización la capacidad del sistema se determina por el tamaño de la terminal de radiolocalización. Tamaño se refiere a la cantidad de entradas y salidas, a la cantidad de memoria (cantidad de códigos que soporta). Este dimensionado es muy importante ya que esta terminal es la responsable de recibir, procesar y guardar la información del abonado a quien va dirigido el mensaje. Además valida los tipos de mensajes pedidos, determina la autenticidad del abonado y le remite los mensajes a los controladores de zona (red de distribución de mensajes).

La terminal de radiolocalización a utilizar es la *Unipage* es modular y expandible de 2,000 códigos hasta más de 2,000,000 códigos, se adapta muy bien al crecimiento de nueva tecnología y expansión de la empresa. A continuación se dan las características más importantes que tiene:

- Capacidad de más de 1,000,000 de abonados y 1,900 troncales.
- En las troncales de entrada acepta señalización por pulso, DTMF y MF.
- Formatos de localización *POCSAG* (“502”, “1200”, “2400”), GSC, 2 tonos MVS, RDS y *FlexTM*.

- Mensajes alfabéticos, numéricos, por tonos y voz.
- Soporta troncales análogas y digitales.
- Mensajes de voz: un sistema de mensajes versátil de voz y números puede proveer desde 6 hasta 600 Hrs. de almacenamiento de voz. También posee características adicionales como mensajes desechados, notificaciones, recordatorio de mensajes no contestados, saludos y anuncio de cambio de número.
- Transcripción alfabética: convierte todos los mensajes de voz a los mensajes alfabéticos a través de un operador para su posterior transmisión.
- Capacidad de red: ofrece toda la posibilidad de interconexión por red.
- Mantenimiento flexible y fácil: de diagnóstico local o remoto y capacidad estadística, mantenimiento en línea y una opción de sistema de redundancia.

4.2.2.4. Protocolos del Sistema Integrado de Mensajería

Los protocolos son el sistema nervioso del sistema, ya que como se definió antes, un protocolo, es un juego de reglas diseñado para facilitar de manera eficiente y creíble la transferencia de información sobre una red. Los protocolos decretan la capacidad, velocidad de señalización, tiempo de vida de la batería e integridad de los datos.

Esta red es de topología local, donde generalmente un terminal de radiolocalización, distribuye los mensajes en un área geográfica limitada, aquí no existe comunicación entre terminales de paginación y por consiguiente no se puede enviar mensajes de otros terminales de paginación.

En los apartados siguientes se darán los distintos tipos de protocolos que debe utilizar la red, pero a continuación se dan los requisitos mínimos que debe tener un protocolo de radiolocalización:

- Capacidad de direccionamiento.
- Velocidad de transmisión acorde con el tráfico.
- Alta protección frente a llamadas dirigidas a otro receptor.
- Métodos que permiten consumo bajo de energía a los receptores.
- Permitir la transmisión secuencial y cuasisíncrona.
- Decodificación sencilla en los receptores.
- Protección contra efectos de transmisión: interferencias, solapamiento, etc.
- Permitir convivir a varios fabricantes sobre un mismo sistema.
- Capacidad de transmisión de mensajes numéricos y alfanuméricos además de avisos.
- Protección contra la generación de mensajes incorrectos.

4.2.2.4.1. Protocolo de la *PSTN* al sistema de radiolocalización

En sistemas como este, el protocolo más comúnmente utilizado para enviar mensajes alfanuméricos del usuario al terminal de paginación es el denominado *TAP* (*Telocator Alphanumeric Protocol*). Este protocolo hace posible la entrada de datos desde dispositivos, tales como PC's, comunicaciones seriales dedicadas, y en este caso desde un teléfono digital por el cual se accesa a un *ivr* y luego a un reconocimiento de voz.

4.2.2.4.2. Protocolos de *Internet* y correo electrónico

Los protocolos a utilizar para la comunicación del servidor de *Internet* y correo electrónico con la terminal de radiolocalización son el *SMTP* (*Simple Mail Transfer Protocol*) y el *http* (*Hypertext Transfer Protocol*).

El *SMTP* es un protocolo usado para enviar correo electrónico por una red, con la limitante de que no define métodos para enviar otras formas de información y documentos adjuntos. El *http* es usado para enviar mensajes por medio de sitios de *Internet*, es implementado en los *gateway's* que son usados como portales a los sistemas de radiolocalización, requiere una conexión fiable y es soportado en redes de *TCP / IP*.

4.2.2.4.3. Protocolos sobre la infraestructura del sistema de radiolocalización

Sobre el protocolo que se debe de utilizar sobre la comunicación entre la terminal y el sistema de control del sistema de radiolocalización, existen dos protocolos pero se puede seguir usando el más conocido que es el *TNPP (Telocator Network Paging Protocol)*, este protocolo hace posible las comunicaciones entre estos los dos equipos, utilizando paquetes de información que son distribuidos convenientemente en la red.

El otro protocolo a mencionar es el *WMtp (Wireless Messanging Transfer Protocol)* es un protocolo basado sobre *IP*, es usado para cambiar mensajes binarios, numéricos y alfanuméricos entre la terminal de paginación y el sistema de control en un sistema de radiolocalización. Debido a las características del sistema es el más adecuado a utilizar.

4.2.2.4.4. Protocolos entre la infraestructura del sistema de radiolocalización y los radiolocalizadores

En este apartado, se hablará exclusivamente del protocolo de señalización a usar, el cual permite la comunicación entre la infraestructura y el radiolocalizador. Como se sabe, la velocidad de información a los radiolocalizadores es un factor crítico en la determinación de la capacidad (número de abonados que pueden atenderse) del canal de radio. Lo que resalta aquí es que los protocolos más rápidos permiten bases más grandes de abonados. En el capítulo uno se habló de los diferentes protocolos que existen, esta red se diseñará para que funcione con el protocolo *FlexTM*, este es más rápido y más confiable, y lo más importante es que puede atender más abonados por canal, también se puede resaltar que soporta datos binarios, numéricos y alfanuméricos.

El protocolo *FlexTM* se estructura por tramas donde cada trama tiene 1.875 segundos de longitud, un ciclo completo de radiolocalización se realiza con 128 tramas, en un periodo de 4 minutos (1.875 segs. / trama * 128 tramas / ciclo).

También se puede mencionar que los radiolocalizadores se pueden programar para que se activen con ciertas tramas especificadas. La vida de la batería depende de la frecuencia de los mensajes. Los mensajes se pueden programar en cada trama (128 veces en 4 minutos), o infrecuentemente cada ciclo de radiolocalización (1 vez cada 4 minutos). Esta frecuencia de los mensajes puede ser controlada por un valor de colapso C , entonces las tramas que los radiolocalizadores tienen programadas están determinadas por el valor de colapso según la expresión 2^C ; esto significa que los radiolocalizadores pueden activarse en cada trama, o todas las tramas, cada 4 tramas, cada 8 tramas,...., o cada 128 tramas.

El protocolo *FlexTM* define diferentes tipos de campos que contienen diferente información como: configuración y control del sistema, sincronización, dirección del radiolocalizador, datos del usuario y otros. Cada trama contiene una combinación de estos tipos de campos; donde los datos del usuario son parte de la información que se transmite.

Los datos son encapsulados en cadenas (ó *codewords*) de 32 bits (4 octetos) de longitud, donde solo 21 *bits* están habilitados para información, los otros 11 son usados para detección de errores. Los datos se condensan en cada cadena no importando cuantas cadenas se usan; es decir, se llena la información en los 21 *bits* habilitados, hasta que esté lleno, lleno este se continúa con otro y así se repite hasta que la información del usuario esté completa.

Los datos numéricos se codifican usando 4 *bits* por dígito, los alfanuméricos 7 *bits* por carácter y los hexadecimales y binarios se toman como *bits* individuales. Con esto se sabe que: a) el número máximo de mensajes numéricos que pueden enviarse en una trama es de 8 cadenas (*codewords*) conteniendo 41 dígitos, b) el número máximo de mensajes alfanuméricos que pueden enviarse en una trama es de 84 cadenas (*codewords*) conteniendo 252 caracteres, y c) el número máximo de mensajes hexadecimales y binarios que pueden enviarse en una trama es de 84 cadenas (*codewords*) conteniendo 1764 *bits*. Cabe mencionar que en los mensajes alfanuméricos y hexadecimales / binarios pueden enviarse en las subsiguientes tramas con una longitud máxima no especificada, que por lo regular el operador del sistema de radiolocalización limita el tamaño de los mensajes. En la tabla VIII se muestra los límites máximos de datos en el protocolo *FlexTM*.

Tabla VIII. Límites máximos de datos del protocolo *FlexTM*

	Mensaje numérico	Mensaje alfanumérico	Mensaje hex / binario
Máximo de cadenas por trama	8	83 en la 1ra. trama, 84 en las tramas subsiguientes	83 en la 1ra. trama, 84 en las tramas subsiguientes
Máximo de datos del usuario por trama	41 dígitos	249 caracteres en la 1ra. trama, 252 en las tramas subsiguientes	1743 <i>bits</i> en la 1ra. trama, 1764 en las tramas subsiguientes
Continuación de tramas	No	Si (sin límite en la trama)	Si (sin límite en la trama)
Máximo de datos del usuario impuestos por el protocolo	41 dígitos	Sin límite	Sin límite

Fuente: “Motorola”

En las siguientes tablas se dan algunas comparaciones de características importantes de los protocolos y demuestran el porqué de la escogencia de protocolo *FlexTM*. En la tabla IX se muestra una comparación de velocidad y número de abonados por canal de los distintos tipos de protocolo; luego en la tabla X se muestra una matriz de comparación de características de protocolos más recientes incluido el *FlexTM*.

Tabla IX. Comparación de características de los tipos de protocolo

Tipo de radiolocalizador	Tipo de codificación	Velocidad de Señalización del radiolocalizador	No. Abonados por canal
Tono y voz	2 tonos	4 segundos	1,500
Tono y voz	5/6 tonos	0.25 segundos	1,700
Únicamente tono	<i>POCSAG</i>	512 bps	313,920
Únicamente tono	<i>POCSAG</i>	1200 bps	737,280
Numérico	<i>POCSAG</i>	512 bps	62,784
Numérico	<i>POCSAG</i>	1200 bps	147,456
Numérico	<i>POCSAG</i>	2400 bps	294,336
Numérico	<i>FlexTM</i>	6400 bps	668,200
Alfanumérico	<i>POCSAG</i>	512 bps	10,464
Alfanumérico	<i>POCSAG</i>	1200 bps	24,525
Alfanumérico	<i>FlexTM</i>	6400 bps	131,000

Fuente: "Motorola"

Tabla X. Matriz de comparación de los protocolos más recientes

Protocolo	<i>POCSAG</i>	<i>ERMES</i>	<i>FlexTM</i>
Descripción	Tecnología de una vía, velocidad baja	Tecnología Europea de una vía	Tecnología de una vía, velocidad alta
Aplicación	Numérico y alfanumérico (4/7 bit)	Numérico y alfanumérico (4/7 bit, binario)	Numérico y alfanumérico (4/7 bit, binario y caracteres simbólicos)
Frecuencia de operación	Cualquier frecuencia habilitada para el sistema de radiolocalización	169, 425 a 169, 800 MHz	Cualquier frecuencia habilitada para el sistema de radiolocalización
Capacidad de <i>roaming</i>	No soporta	Si, entre los sistemas <i>roaming</i>	Soporta
Canal de salida	25 kHz	25 kHz	25 kHz
Velocidad de señal de salida	512, 1200 o 2400 bps	6250 bps	6250 bps

Fuente: “Motorola”

4.2.2.5. Sistema de control de radiolocalización

El sistema de control de radiolocalización (controladores) es el encargado de ejecutar varias tareas como: formación de colas de espera de lotes de mensajes, codificar, programar los mensajes recibidos de la terminal de radiolocalización para distribuirlos a los sitios transmisores de manera que se optimicen los medios de transmisión y el espectro de radio frecuencia. Algunos proveedores en la infraestructura del sistema combinan la función del sistema de control en la terminal de radiolocalización y otros lo hacen con los dos equipos separados.

El sistema de control o controlador a utilizar es el *C-Net* y el cual tiene las siguientes características:

- Controlador avanzado de redes RF.
- Enlaces de la terminal a la estación de base.
- *C-Net* se puede conectar a la estación base por medio de línea física, radio enlace, microonda o satélite.
- *C-net* reduce grandemente la necesidad de viajar a las estaciones base y permite actualizar su software sin tener que hacer el viaje.
- Soporta multiplexación de múltiples sistemas digitales radiolocalizadores montados en un enlace simple.
- Soporta señal de radiolocalización de más de 6,400 bps. y formatos de alta velocidad como *FlexTM*.
- Soporta corrección de error saliente para todos los datos de localización y en todos los formatos.
- Soporta autoalineamiento de modulación en retraso.
- Soporta regeneración de señales en los puntos de transmisión.
- Provee un sistema comprensible y equipo de diagnóstico.

- Actualiza simultáneamente los programas de la estación base vía el enlace del sistema.

4.2.2.6. Puertas de enlace (*gateway*'s)

Para crear enlaces entre el satélite y la red de radiolocalización y entre un proveedor de *Internet* y la red de radiolocalización es necesario contar con puertas de enlace o *gateway*'s, los cuales manejan la traslación protocolar y enrutan la información entre los sistemas externos y la infraestructura de la red.

4.2.2.6.1. Puertas de enlace (*gateway*) de *Internet*

La puerta de enlace o *gateway* de *Internet* es un enlace entre el proveedor de servicio de *Internet* y el servidor de mensajes, este no es más que un sistema de software diseñado para crear la interfaz entre los equipos codificadores e *Internet*.

Generalmente se envía un correo electrónico el cual se transmite directamente al radiolocalizador utilizando una página *web*. El sistema también puede enviar mensajes teniendo la base de datos conectada a una página en *Internet* en la cual se coloca la información del abonado y del mensaje para luego distribuirse adecuadamente a través de la red.

El equipo que se puede utilizar para hacer las tareas anteriores es el *router* de *Cisco* 2610, el cual posee entre sus características interfaz *Ethernet*, puerto serial E1PRI.

4.2.2.6.2. Puerta de enlace (*gateway*) satelital

El sistema integrado de mensajería, va a ser parte de una red nacional e internacional con la adición de un enlace con el satélite *Iridium*. Para tener esta conexión se necesita que la red este equipada con una interfaz satelital. La conexión entre el satélite *Iridium* y la red se realiza a través de un *gateway* y el más cercano esta situado en Florida. Cada *gateway* controla un solo satélite *Iridium* a la vez. Esto permite a un usuario enviar mensajes a un abonado que se encuentra en alguna parte del globo. Otra función que puede dar el satélite es la de sincronizar la transmisión simultanea de mensajes en la red nacional, por medio de su sistema *GPS*.

4.2.2.7. Medios de transmisión

Son los encargados del traslado de los mensajes desde la infraestructura de radiolocalización a las estaciones base. Los medios a utilizar son radios módem marca MDS modelo *LEDR 1400S*, los cuales tienen las características siguientes:

- Banda de operación 1350 a 1535 MHz.
- Potencia de transmisión: + 30 dBm.
- Sensibilidad en recepción: -98 dBm @ 50 kHz, 128 kbps y BER 10^{-6} .
- Algoritmo de corrección de error adelantado: *Reed Solomon*.
- Ancho de banda utilizado: 50 kHz., 64 –128 kbps.
- Interfaz digital RS530.
- Modulación configurable por *software QPSK, 16 QAM y 32 QAM*.
- Canal de servicio telefónico integrado.
- Canal de servicio de datos RS232 integrado.
- Administración y gestión por medio de panel frontal, terminal ASCII, *Telnet*, *browser* de *Internet* generico y agente SNMP.

- Alimentación –48 Vdc.

La distancia típica alcanzada por estos radios es de 40 Km. Para el enlace se utilizan antenas parabólicas de 3 pies de diámetro en ambos extremos.

4.2.3. Transmisión

La parte de transmisión es la encargada de aceptar y propagar los mensajes provenientes de la infraestructura de radiolocalización a los radiolocalizadores de los abonados. Reciben la señal del sistema de control o controlador, extraen la orden de disparo y modulan convenientemente a una portadora de radiofrecuencia.

La transmisión es diseñada para operar en un rango de frecuencias asignadas exclusivas para el sistema de radiolocalización de Guatemala, la cuales se pueden ver en la tabla XI. En estas frecuencias no hay refracción y las sombras de la señal radioeléctrica se producen en cara no iluminada del objeto que obstruye, por ejemplo una montaña; La peor obstrucción es la del radiohorizonte, debida a la curvatura terrestre. También se puede mencionar que la señal se refleja sobre la superficie de los objetos próximos, y las zonas de sombra se rellenan de reflexiones de la señal. Esto permite que en núcleos urbanos la recepción sea posible tanto directamente desde el transmisor como por el conjunto de diversos rayos reflejados.

Tabla XI. Frecuencias de operación sistema de radiolocalización en Guatemala

Servicio	Tx (MHz)	
<i>Paging</i>	931.000	932.000
<i>Paging</i>	929.000	930.000
<i>Paging</i>	460.000	470.000
<i>Paging</i>	156.8375	174.000
<i>Paging</i>	150.050	156.7625
<i>Paging two way</i>	901.000	902.000
<i>Paging two way</i>	940.000	941.000
<i>Paging two way</i>	930.000	931.000

Fuente: Superintendencia de telecomunicaciones

Otro aspecto importante a considerar es el área de cobertura de la ciudad principal y áreas adyacentes, en este caso son la Ciudad Capital y sus departamentos, del análisis de cobertura se toma la decisión de cuantos sitios transmisores debe tener la red de transmisión, para este se puede utilizar algún software que analice áreas de cobertura deseadas, estos *softwares* hacen su análisis en base al criterio de *Ohmori Okumura*^[5].

Cabe mencionar que se utilizará transmisión simultánea, en la cual la señal se envía a múltiples transmisores los cuales se sincronizan por un sistema *GPS*, para que esta llegue al mismo tiempo; esto evita que se tengan problemas por interferencias en la propagación por solapamiento de las señales provenientes de los transmisores próximos.

Después de considerar aspectos importantes como frecuencias y área de cobertura, los equipos encargados de la transmisión son núcleos, antenas, que juntos hacen las estaciones base y el satélite *Iridium* del que se habló en el capítulo 3 y en el apartado 4.2.2.6.2.

4.2.3.1. Núcleos

Básicamente los núcleos son el cerebro de la transmisión, reciben los mensajes de los controladores, los cuales son codificados según el protocolo usado en la transmisión, para luego transmitirse en el momento preciso indicado por el controlador. Por lo regular el sistema de radiolocalización maneja una tasa de transferencia de datos entre 1200 a 6400 bps en el canal de salida y el cual tradicionalmente, los protocolos de radiolocalización lo usan a 25 kHz.

A continuación se dan las características que tienen los núcleos:

- Control por microprocesador ofrece la máxima flexibilidad para cambios en el sistema y parámetros de operación.
- Un sistema de procesamiento digital de la señal, soporta cuatro niveles de operación y códigos de alta velocidad.
- Un diseño térmico óptimo y disipadores térmicos confiables para prolongar la vida de los componentes e incrementar la confiabilidad de la estación base.
- Fácil mantenimiento debido a su diseño completamente modular, panel frontal de acceso y módulos de componentes tipo gaveta.
- Trabaja con diferentes distribuciones de medios de transmisión: línea física, radio enlaces, microondas.
- Una amplia variedad de salidas de RF disponibles desde 20 hasta 350 vatios.
- Disponible en diferentes bandas de frecuencia: VHF, 280 MHz., 900 MHz., y UHF.

- Acomoda todos los formatos de radiolocalización: voz, 2 tonos, 5/6 tonos, *POCSAG*, *GSC* y formatos *FLEXTM* que pueden usarse solos o en combinación.

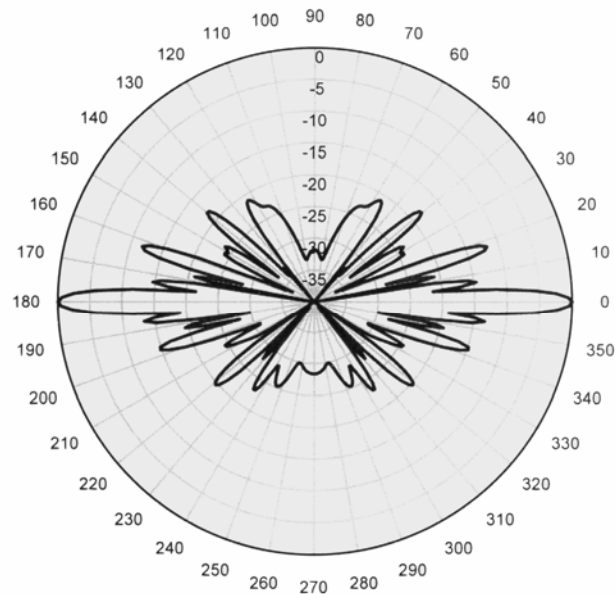
4.2.3.2. Antenas

El servicio de radiolocalización requiere penetración dentro de los edificios. A las frecuencias de la tabla XI. la penetración es buena, aunque se producen pérdidas. Las antenas suelen ser omnidireccionales, en la tabla XII y en la figura 74 se dan las especificaciones eléctricas y el patrón de radiación del antena, del modelo CT1D0F-0083-012 de “**Andrew**”, que se utilizará en el diseño, aunque se pueden utilizar de otras marcas también siempre y cuando cumplan con los requisitos.

Tabla XII. Especificaciones Eléctricas

Banda de frecuencia, MHz	806 a 869
Ganancia, dBd (dBi)	10 (12.2)
Elevación <i>Beamwidth</i> (grados)	4.5
Retorno de pérdidas, dB (VSWR)	>14.0 (<1.5)
Impedancia, ohms	50
Tipo de polarización	Simple / Vertical
Inclinación eléctrica vertical, grados	0
Potencia máxima de entrada, Watts	500
Protección	Tierra DC
Tipo de conector	7 – 16 DIN hembra
Posición del conector	Bajo

Fuente: “Andrew”

Figura 74. Patrón de radiación.

Fuente: "Andrew"

4.2.4. Recepción

El bloque de recepción lo componen radiolocalizadores, que son dispositivos de comunicación personal del abonado. En el siguiente apartado se da un resumen del tipo de radiolocalizador a utilizar, tomando en cuenta que en el capítulo uno se describieron ciertas características de funcionamiento y funciones de los mismos. Además se toma en cuenta los radiolocalizadores *Iridium* que son un complemento a los servicios prestados por este sistema integrado de mensajería.

4.2.4.1. Radiolocalizadores

Típicamente los radiolocalizadores operan en una banda de frecuencia limitada, la frecuencia se programa en el dispositivo y corresponde a la frecuencia que el portador es autorizado a usar; una de las consideraciones es el uso del protocolo *FlexTM* por lo que los dispositivos a usar deben de soportar dicho protocolo.

Los radiolocalizadores que utilizan el protocolo *FlexTM*, se clasifican con base en la forma en que se comunican, y tomando en cuenta que dicho protocolo acepta mensajes numéricos y alfanuméricos, entonces los posibles radiolocalizadores a utilizar son alfanuméricos por lo que a continuación se dan ciertas características de ellos y se comparan con los numéricos.

4.2.4.1.1. Radiolocalizadores numéricos

Estos radiolocalizadores, antes de los alfanuméricos eran los más usados en el mundo, ya que el contenido de información del mensaje es bastante completo (más que un tono), y la capacidad del canal es mucho mayor que en el caso de mensajes de voz. Dado que muchos radiolocalizadores pueden operar en un canal de radio, y casi todos los sistemas numéricos están automatizados (sin operadores), el costo de su utilización es muy bajo.

4.2.4.1.2. Radiolocalizadores alfanuméricos

Con frecuencia se ha dicho que “lo bueno es el enemigo de lo mejor”. Esto puede ser cierto cuando se compara la radiolocalización con despliegue numérico con la radiolocalización alfanumérica. Un despliegue numérico solamente muestra un número telefónico. Si el abonado no reconoce el número, no sabrá quien le esta llamando, que asunto puede ser o si es o no urgente. Por ejemplo, es muy errático recibir un mensaje numérico, salirse de la ruta que uno lleva, estacionarse, encontrar un teléfono, obtener las monedas necesarias y realizar la llamada, sólo para darse cuenta que el mensaje no era importante.

Un mensaje alfanumérico enviado apropiadamente le dice al abonado: quién llamó, por qué, dónde debe ir, cuándo, la dirección, la hora, etc. Es importante observar que no se requiere contestar una llamada con un radiolocalizador alfanumérico. Esta sola característica puede ahorrarle al abonado mucho tiempo y esfuerzo. En aquellos lugares del mundo donde no se tienen disponibles teléfonos públicos (para contestar una llamada), la característica alfanumérica ofrece una gran ventaja sobre los otros tipos de radiolocalizadores y es por eso que es la que se va a utilizar. Se presenta una tabla con características del radiolocalizador alfanumérico.

Tabla XIII. Características del radiocalizador alfanumérico

Características	Función	Modo de comunicación
	Sonido - Tono	Audible
	Luz intermitente	Visual
	Vibración	Mecánica
	Letras del alfabeto y puntuación	Mensaje completo para leer
Ventajas:	<ul style="list-style-type: none"> • El contenido de la información es el más alto de todos los dispositivos. • La señalización es rápida, por lo que la capacidad de canal de radio es grande. • Los mensajes son privados. • Los mensajes pueden ser salvados para una revisión posterior. • Notificación de noticias, información climatológica, deportes, finanzas, etc. 	
Desventajas:	<ul style="list-style-type: none"> • El costo es más elevado que el numérico. 	

Fuente: "Motorola"

Ya que se ha resuelto que se utilizarán radiocalizadores alfanuméricos con protocolo *FlexTM*, para este protocolo se tienen los siguientes modelos de radiocalizadores: *Advisor FLX32TM*, *Advisor Gold FLXTM*, *Advisor Pro FLXTM*, *Memo Jazz FLEXTM*, *ScriptorTM FLX2*, *ScriptorTM FLX4*, *Wordline FLXTM*.

4.2.4.1.3. Radiolocalizadores *Iridium*

Uno de los propósitos es que este sistema de mensajería integre el sistema *Iridium*, el cual provee un servicio en cualquier punto de la tierra (tierra, mar y aire) siempre y cuando la altitud sea inferior a 185 km, para aquellos momentos en los que se necesita estar localizable. El sistema *Iridium* incluye un servicio global de mensajes numéricos y alfanuméricos, los radiolocalizadores tienen una forma similar a la del radiolocalizador familiar que se utiliza hoy en día, con lo cual se convierten en una herramienta de radiolocalización que cabe en la palma de la mano.

La función de la decodificación del mensaje utiliza *FDMA / TDMA*, comparte los enlaces de la banda L de voz y otros servicios del *Iridium*. Al transmitir por el radiolocalizador se puede alcanzar gran exactitud de caracteres en el mensaje.

A continuación se dan algunas características de este radiolocalizador:

- Mensajes alfanuméricos.
- Pantalla iluminada *Optimax EL™*.
- Pantalla de cuatro líneas y 80 caracteres.
- Juegos de caracteres internacionales.
- Memoria para agenda telefónica.
- Indicador de batería.
- Reloj de dos horarios de uso (hora local y del lugar de residencia).
- Alarma de viaje.
- Tamaño: 77 x 72.3 x 22.5 mm (anchura x altura x profundidad).
- Peso: 118 gramos.
- Duración de la batería: 30 días aprox. (batería alcalina de 1.5 V AA).
- Memoria: 128 kB de *RAM*.

Una vez descrito cada bloque del Sistema Integrado de Mensajería, puede este representarse en un diagrama general donde se muestre cada componente, visualizándose en la figura 75. Teniendo en el diagrama la solución integrada de mensajería, se pueden definir los siguientes servicios que el sistema puede prestar.

- Servicio de radiolocalización con reconocimiento de voz: a través de un teléfono PSTN o teléfono IP al sistema, en la cual el mensaje del usuario es convertido a texto para su transmisión a un radiolocalizador o al correo electrónico del abonado, según la elección del usuario.
- Servicio de radiolocalización a través de *Internet* o correo electrónico: envío de mensajes de texto a un radiolocalizador o al correo electrónico del abonado.
- Servicio de radiolocalización vía satélite: envío de mensajes a un radiolocalizador *Iridium* en cualquier parte del globo.

Una vez se tiene una red *VoIP*, se puede empezar a ofrecer otros servicios de valor añadido y comunicaciones interactivas de multimedia, como son la llamada por *Internet* en espera (*ICW, Internet Call Waiting*) y segunda línea virtual (*V2L, Virtual Second Line*).

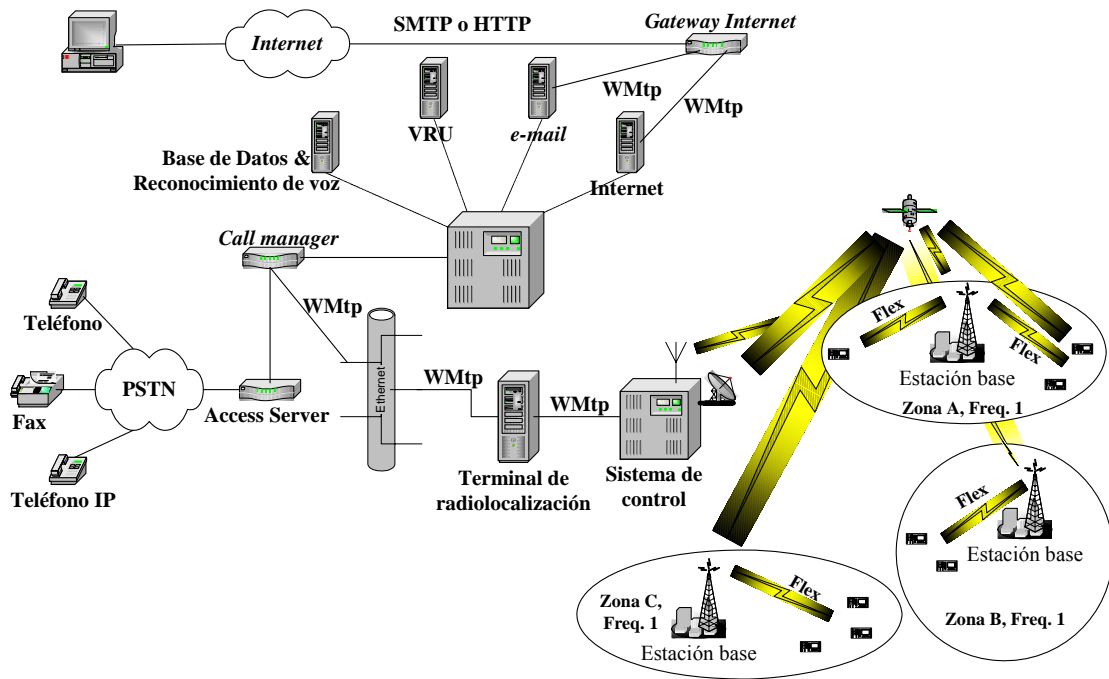
Básicamente, el *ICW* es un servicio que permite a los abonados recibir la notificación de una llamada de voz entrante en su PC, mientras están conectados a *Internet*; la notificación se realiza mediante una pantalla emergente y en ese momento pueden hacer lo siguiente:

- Enviar la llamada al correo de voz.
- Recibir la llamada en el PC utilizando *software H.323 (VoIP)*.
- Suspender la sesión de *Internet* y recibir la llamada en el teléfono tradicional (*PSTN*).
- Transferir la llamada a un teléfono celular.

- Ignorar la llamada, proporcionando una señal de ocupado o dejar de sonar.

El V2L, es un servicio sencillo, ya que permite a los usuarios de *Internet* hacer y recibir llamadas de teléfono a través de su servidor de servicios de *Internet* sólo cuando están conectados a través de su conexión a *Internet*.

Figura. 75. Diagrama general “Sistema Integrado de Mensajería”.



4.3. Evaluación “Sistema Integrado de Mensajería”

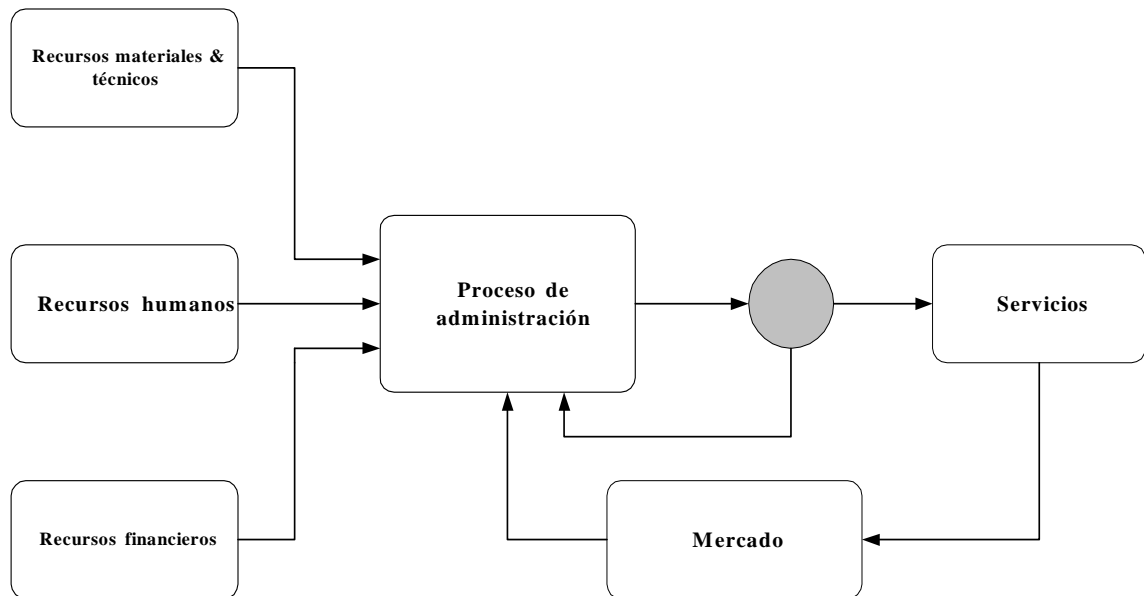
En este apartado, se hará un análisis del diseño “Sistema Integrado de Mensajería”, dicho análisis solo comprende el estudio de preinversión para el desarrollo del diseño, ya que los estudios de inversión y postinversión son: primero la ejecución del proyecto y segundo la etapa final donde deben cumplirse los objetivos que se determinaron en el estudio de preinversión; los cuales no se realizarán ya que este es un trabajo investigativo.

El estudio de preinversión permite determinar la bondad del proyecto en si mismo y permite justificar el proyecto ante potenciales inversionistas o financiadores (obtener financiamiento).

4.3.1. Estudio preinversión

Las empresas de telecomunicaciones son la base de nuestra actual sociedad, ya que proveen los servicios que el hombre moderno requiere para comunicarse.

Para cumplir este importante objetivo, la empresa debe contar con recursos, materiales, técnicos, humanos y financieros, los cuales deben ser administrados racionalmente para optimizar los resultados operativos y económicos de la organización, proyectando esta afirmación a un esquema, la empresa puede representarse como en la figura 76.

Figura 76. Esquema de administración de recursos.

Los recursos financieros son en general escasos ya que las posibilidades de utilización del dinero son mayores que la disponibilidad de capital. Luego se requiere de una serie de criterios y técnicas para distribuir racionalmente el recurso financiero para que el proyecto sea más rentable. De esta forma la evaluación económica mide las ventajas y desventajas de un proyecto, con el propósito de que los recursos disponibles sean asignados a un proyecto que sea factible.

La selección de un proyecto de telecomunicaciones no puede ser hecha arbitrariamente, sino que debe ser el resultado de una decisión cuidadosa, basada sobre consideraciones objetivas. Lo que se requiere son técnicas tales como métodos de evaluación, los que guían la selección de un curso particular de acción de entre varias alternativas, por métodos cuantificados.

Por la naturaleza y características propias este proyecto entrega un servicio de comunicación, con una función estratégica independiente, con origen en un sector económico privado. Por otro lado el objetivo para la elaboración de este proyecto (hipotéticamente) se origina para presentársele a posibles inversionistas.

El desarrollo de un proyecto de telecomunicaciones requiere generalmente pasar por una serie de etapas desde la idea, perfil hasta la puesta en marcha del servicio. El proceso general distingue una primera fase de concepción intelectual (preinversión), en la cual el proyecto se analiza por medio de técnicas de evaluación realizando primeramente un estudio de viabilidad o prefactibilidad, para luego continuar, con el estudio de factibilidad técnico – económico o evaluación de proyecto propiamente. En los siguientes apartados se describen estas etapas de preinversión.

4.3.1.1. Idea

El origen de este diseño, prevé un sistema donde se tengan varios servicios integrados que hagan fácil la comunicación a usuarios que necesiten estar localizables.

4.3.1.2. Perfil y Prefactibilidad

En muchas oportunidades los estudios de perfil y prefactibilidad no se realizan y se pasa de inmediato a la etapa de estudio de factibilidad, en este caso no se desarrollan por dar más énfasis al estudio de factibilidad. Sin embargo, hay que señalar que estos estudios son rápidos y breves en los cuales se analizan aspectos técnicos generales, condiciones de abastecimientos de materias primas y principalmente las características económicas del proyecto.

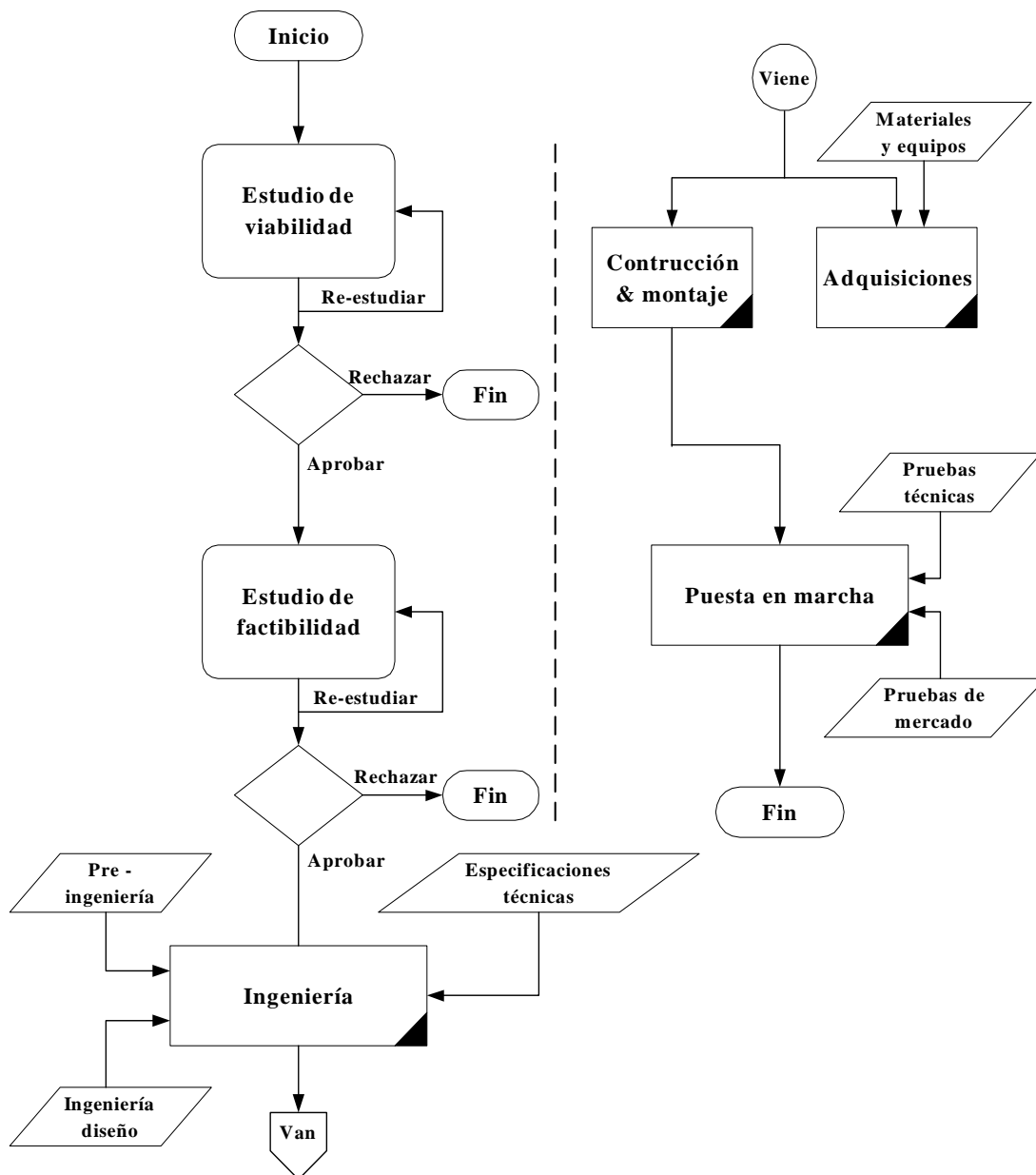
4.3.1.3. Factibilidad

El estudio de factibilidad técnico económico, se define como el conjunto de antecedentes que permiten analizar las ventajas y desventajas tanto técnicas, como económicas de asignar recursos a una unidad productiva para producir un servicio. Si la evaluación del proyecto concluye que es rentable, se pasa a la segunda fase de implementación física en la cual se realiza la ingeniería final del proyecto, su construcción, montaje y puesta en marcha, tal y como se aprecia en la figura 77, como se mencionó anteriormente esta segunda fase no se desarrollará.

La evaluación técnica económica es una herramienta analítica que responde a la pregunta sobre si un proyecto es o no rentable y si se recomienda su implementación a través de la asignación de los recursos financieros requeridos, o por el contrario, si se rechaza el proyecto analizado por no poseer factibilidad económica y por generar futuras pérdidas financieras; al no ser capaz de originar flujos que amorticen las inversiones requeridas y costos asociados al proyecto, todo esto es lo que se comprueba en las siguientes paginas.

Para cumplir los objetivos, la evaluación del proyecto sigue habitualmente una metodología que distingue tres etapas principales que darán las conclusiones esperadas en cuanto la factibilidad del proyecto, estas son: estudio de mercado, estudio técnico y estudio económico.

Figura 77. Implementación física.



4.3.1.3.1. Estudio de mercado

Este estudio da el comportamiento del consumidor y la posición de la competencia para determinar la demanda insatisfecha presente en un mercado, con el objeto de definir un plan de ventas de mediano y largo plazo que defina el respectivo nivel operativo del proyecto; el mercado de hoy exige un enfoque más agresivo por lo que se necesita un plan el cual se presenta a continuación.

- a. ¿Quiénes serán los clientes? Se tiene la certeza de que si se tiene una buena idea de quiénes son nuestros clientes del servicio de radiolocalización por segmentos comerciales, entonces aumentarán nuestras perspectivas de éxito si se buscan nuevos clientes precisamente dentro de estos segmentos. Si se sabe donde trabajan y como se ganan la vida nuestros clientes potenciales, entonces son mejores nuestras posibilidades de encontrarlos. Si se les puede encontrar y explicarles las ventajas de usar un radiolocalizador en su segmento comercial, es probable que muchos de ellos se conviertan en nuestros clientes. Observen el énfasis de lo expresado: un buen vendedor conocerá el negocio de su cliente y podrá demostrarle minuciosamente los beneficios de un radiolocalizador en ese ambiente comercial. Otra opción en el plan que puede servir para crear un perfil de quienes son clientes de radiolocalización es recurrir a compañías que se especializan en hacer encuestas de mercado, esto puede ser útil para penetrar un mercado ya establecido.

- 1° Radiolocalizador comercial: En el pasado, la mayoría de los abonados alrededor del mundo han sido usuarios comerciales que a veces dan un uso personal a sus radiolocalizadores. Si se tiene la información de qué clientes hay por segmentos comerciales, probablemente se puedan identificar y luego señalar a nuestros clientes potenciales en cada segmento comercial, ya que debe haber algo especial acerca de cada uno de estos negocios. De alguna

forma, un radiolocalizador le ayuda a estas personas a ser más efectivas en su trabajo.

2° Radiolocalizador personal: en la actualidad existe un mayor énfasis para usar un radiolocalizador para mantenerse en contacto con miembros de la familia. Si se tiene éxito en abrir el mercado del consumidor, se verá que el número de radiolocalizadores en uso se multiplicará por diez.

3° Radiolocalizador en líneas no tradicionales: En la actualidad los sistemas de información son muy importantes en la radiolocalización. Una variación muy prometedora de este tipo de servicio es el envío de texto o, aún más, de archivos de completos de una computadora al usuario, quien después podrá manipular los datos o correr el archivo en su computador. Obviamente, este servicio es de mucho más valor para el abonado.

b. ¿En que negocio se está? El negocio de radiolocalización, realmente proporciona un medio de comunicación personal a personas con un estilo de vida dinámico. Desde el punto de vista puramente de ventas, es más rentable vender el servicio que vender el radiolocalizador. La gente responde a la idea de estar en contacto con sus clientes ó bien, con sus superiores, el hecho de estar disponible brinda muchos beneficios palpables. Una vez que estén convencidos de que esto es algo bueno, la idea de andar con un radiolocalizador todo el tiempo será aceptada con mayor facilidad. Así que se vende el servicio, el servicio es enviarle mensajes a nuestros clientes. Estos mensajes son transmitidos a través de un canal de radio, así que lo que verdaderamente se vende son pedacitos de tiempo en el canal de radio. Nunca hay que olvidar que este canal de radio es el recurso más valioso. Una vez que la hora más ocupada del día en el canal de radio esté topada, será imposible añadir más clientes para el servicio de radiolocalización personal.

c. ¿Como promover nuestro servicio?

- 1º Promoción tradicional: el aliado tradicional es el anuncio en el periódico, esto aún funciona muy bien, especialmente en la sección comercial del periódico. La televisión probablemente resulta cara, a menos que se haya tomado la decisión de captar el mercado del consumidor. Las papeletas parecen funcionar en pueblos pequeños pero con frecuencia son ignoradas en la ciudad. Anuncios en la radio son efectivos, especialmente durante las horas en que las personas viajan a y desde sus trabajos. Demasiada publicidad puede llevar un negocio a la bancarrota, muy poca puede hacer que se estanque.
- 2º Distribución al por menor: el mejor negocio hoy en el mundo de radiolocalización es la distribución al por menor de radiolocalizadores; ya que estos se pueden tener en vitrinas donde los clientes pueden llegar y comprarlos.
- 3º Ventas directas: esta es todavía la mejor forma de captar cuentas, especialmente las cuentas vitales de clientes grandes corporativos, esto se llama mercadeo de empuje, se debe contar con buenos vendedores, bien entrenados, motivados y agresivos. Se debe ofrecer un proceso sencillo y placentero al cliente cuando éste quiera conseguir un radiolocalizador y así se publicitará el servicio dentro de la gente.

d. ¿Que tipo de gente se necesita?

- 1° Gerencia: se necesita un buen gerente, un líder, que pueda motivar a la gente y comprenda las exigencias de administrar bien un negocio y la labor de ventas.
- 2° Ventas: las tres primeras reglas para manejar exitosamente un equipo de ventas es: motivación, motivación y motivación. Los rubros que siguen pueden parecer obvios, pero es sorprendente ver que con frecuencia son ignorados. Comisiones: un buen plan de comisiones mantendrá al equipo de ventas continuamente ágil y atrayendo a nuevos clientes, y estos nuevos clientes son el pulso de la empresa. Apariencia: los vendedores deben de tener una presentación profesional ya que esto les ayudará a ser percibidos como tal. Actitud: este rubro es más complejo, posiblemente es el aspecto más difícil de la conducta humana con el cual se enfrenta siempre; se necesitan empleados con una buena actitud, el respeto a la dignidad humana y la cortesía de parte de los empleados no sucede automáticamente, debe darse desde arriba hasta abajo en la compañía, para así estos a los clientes y se tenga una satisfacción total hacia el cliente.
- 3° Gestión administrativa y operacional: a menudo se considera que esta parte del personal está detrás del telón y que es poco importante. Nada más lejos de la verdad, este grupo consiste de personas de apoyo y son vitales para la organización. Deben ser escogidos por su talento y personalidad, recibir entrenamiento sobre como operar el equipo y sobre la etiqueta a observar en sus relaciones telefónicas comerciales.
- 4° Tecnificación: es esencial que se disponga de personal técnico altamente entrenado y calificado. La confiabilidad del sistema no es tan sólo un asunto

de tener un buen negocio, es también un asunto de ética. Se debe de tener un personal técnico (ingenieros y técnicos) debidamente entrenado, con repuestos adecuados y equipo para realizar pruebas que deben mantener funcionando el sistema.

e. ¿Cómo planificar una expansión?

- 1º Capacidad del sistema: Se recomienda que el sistema siempre esté por lo menos un año delante de su ritmo de crecimiento, con una configuración inicial apta para los primeros cinco años. La buena planificación es una buena práctica comercial, hay que evitar no ignorarla.
- 2º Área de cobertura: El área de comercialización debe estar saturada con la señal de radiolocalización, debe penetrar edificios, alcanzar a los sitios de estacionamiento subterráneo y extenderse hasta afueras del área que se pretende abarcar. Una expansión planificada del área de cobertura debe ser un evento continuo hasta los límites de la licencia o concesión y debe incluirse en cada presupuesto anual.

4.3.1.3.2. Estudio técnico

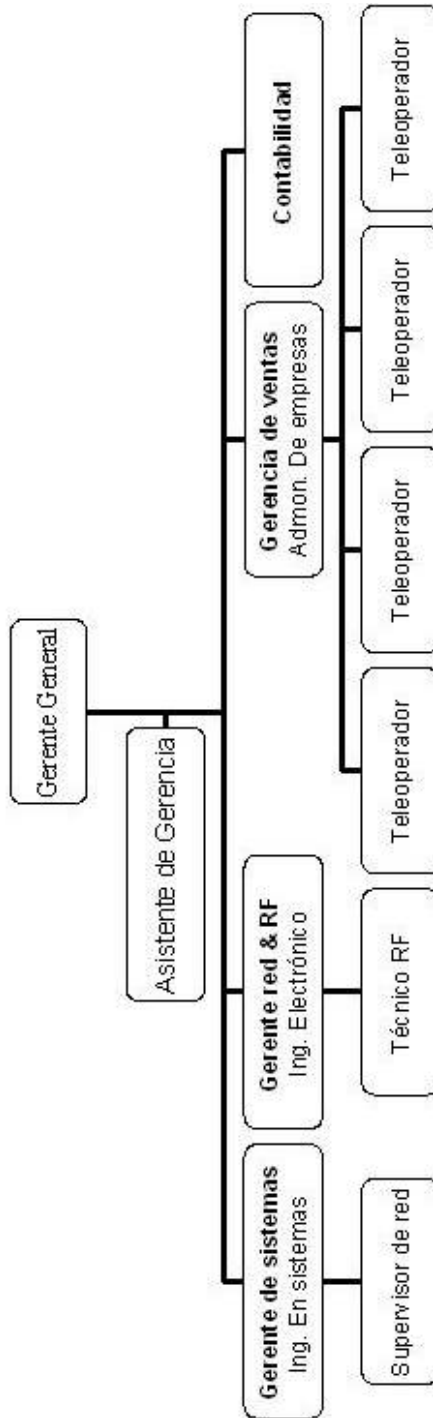
Conocido el nivel operativo determinado por el estudio de mercado, el proyecto se define técnicamente en cuanto a procesos y tecnologías a utilizar y organización de los recursos humanos, respecto a los procesos y tecnologías no se entrará en mucho detalle ya que estos aspectos se describieron en el apartado 4.2, esta etapa se desarrollo en base a una ingeniería básica y/o conceptual, a continuación se describirá en la tabla XIV la infraestructura requerida de la red a emplear y en la figura 78 la organización con el objetivo de conocer los costos de inversión y operación del proyecto.

Tabla XIV. Costos de equipos de la red

Cantidad	Nombre equipo	Costo unitario	Costo total	Costo total
14	EI's	\$. 2,000	\$. 28,000	Q. 224,000
1	<i>Access server</i>	\$. 146,700	\$. 146,700	Q.1,173,600
1	<i>Call manager</i>	\$. 2,000	\$. 12,000	Q. 96,000
1	<i>Red Lan Ethernet</i>	\$. 11,200	\$. 11,200	Q. 89,600
1	Servidor base de datos	\$. 1,000	\$. 1,000	Q. 8,000
1	Servidor correo electrónico	\$. 2,500	\$. 2,500	Q. 20,000
1	Servidor <i>Internet</i>	\$. 4,000	\$. 4,000	Q. 32,000
1	Base de datos	\$. 5,000	\$. 5,000	Q. 40,000
1	<i>Software</i> reconocimiento de voz	\$. 1,500	\$. 1,500	Q. 12,000
1	Terminal de radiolocalización	\$. 5,000	\$. 5,000	Q. 40,000
1	Control de radiolocalización	\$. 12,000	\$. 12,000	Q. 96,000
1	<i>Gateway Internet</i>	\$. 2,000	\$. 2,000	Q. 16,000
6	Radios	\$. 9,200	\$. 55,200	Q. 441,600
6	Núcleos	\$. 14,000	\$. 84,000	Q. 672,000
6	Antenas	\$. 1,200	\$. 7,200	Q. 57,600
6	Sitios estaciones base	\$. 25,000	\$. 150,000	Q. 1,200,00
	Totales		\$. 527,300	Q.4,218,400

Organigrama "Sistema Integrado de Mensajería"

Figura 78. Organización recursos humanos



4.3.1.3.3. Estudio económico

Conocidos los niveles de inversión en activos fijos y capital de trabajo, se está en condiciones de calcular el punto de equilibrio de la operación, por medio de la proyección de ingresos y egresos asociados al proyecto sujeto a evaluación. Se determinan los flujos netos correspondientes y se procede a calcular los indicadores de rentabilidad de acuerdo a los métodos de evaluación *VAN* y *TIR* o combinación de ambos.

El propósito final de cualquier estudio de preinversión consiste en establecer las ventajas y desventajas que ofrece la propuesta ante otras alternativas de inversión. O visto desde otro ángulo, se trata de establecer cuánto dinero se aporta a manera de inversión y cuánto se recupera a través de la operación normal del mismo; esto equivale, en últimas, a determinar la magnitud del efectivo entregado por inversionistas, con relación al efectivo que retorna como consecuencia del giro normal de operación de la empresa.

Con base en la información recogida a través de los estudios anteriores, se puede elaborar un balance de la situación inicial de la empresa: la inversión necesaria para la operación de la empresa se obtiene a través de los inversionistas que aportarán el 30% y el resto se cubrirán con un crédito, ofrecido por una institución bancaria que ofrezca una tasa de interés anual entre el 12% y el 24%. Permitirá adelantar las primeras inversiones, para activos fijos (inmuebles, equipo, infraestructura, etc.) se utilizan Q.4,198,400; para diferidos (estudios preliminares, gastos preoperativos, etc.) se emplean Q.32,000; y se asignan Q.60,000 para atender las necesidades de capital de trabajo, esto se muestra en la tabla XV.

Tabla XV. Balance inicial

Balance inicial					
Activo			Pasivo		
1	Caja	Q. 60,000	1	Obligaciones (70%)	Q. 3,017,280
2	Diferidos	Q. 32,000	2	Patrimonio (30%)	Q. 1,293,120
3	Activos fijos	Q. 4,218,400			
	Total	Q. 4,310,400		Total	Q. 4,310,400

Desde un principio vale la pena estimar las ventas, teniendo en cuenta la capacidad de operación de la empresa y los precios unitarios asignados a los servicios, se considerará un precio promedio de todos los servicios prestados siendo este de Q.480.00. Se puede observar que el primer año se dedica a la instalación de la empresa y, por lo tanto, no se prestan servicios; durante los próximos años, se considera que el proyecto está activo y paulatinamente se incrementarán los abonados hasta llegar a la óptima utilización de su capacidad instalada. En la tabla XVI se presenta la estimación del incremento de abonados y las ventas por períodos establecidos.

Tabla XVI. Estimación de ventas

Ventas						
Período	1	2	3	4	5	6
Abonados		1500	1750	2000	2250	2500
Ventas		Q.720,000	Q.840,000	Q. 60,000	Q.1,080,000	Q. 1,200,000

Como lo que interesa es obtener flujos de efectivo, entonces se diseña un modelo que permite alcanzar este objetivo. Primeramente se considera que para costos de operación se tendrá el primer año de operación Q.96,000.00 y los cuales se incrementan en promedio un 5% en cada año, debido a que crece la utilización de la capacidad instalada. Cabe mencionar que los costos de operación se clasifican en gastos de fabricación, gastos administrativos, gastos de ventas y gastos de financieros; discriminados en costos fijos y variables.

Otro factor importante a considerar es el tratamiento que se les da a las depreciaciones y a la amortización de diferidos, pues siendo contabilizados como costos no constituyen erogaciones reales de la empresa durante el periodo de operación. La tabla XVII ilustra su comportamiento.

Tabla XVII. Depreciaciones y diferidos

Depreciaciones y diferidos						
Período	1	2	3	4	5	6
Inv. Fijas depreciables	Q. 4,218,400					
Deprec. Anual (20%)		Q.843,680	Q.843,680	Q.843,680	Q.843,680	Q.843,680
Diferidos	Q. 32,000					
Amort. diferidos		Q. 8,000	Q. 8,000	Q. 8,000	Q. 8,000	Q. 8,000

Considerando, que los aportes de los socios no son suficientes para atender las necesidades financieras, se precisa acudir al crédito; tal como lo precisa la información del balance inicial se requieren Q.3,017,280 de crédito que se consiguen en las siguientes condiciones: pagadero en 5 años, con uno de gracia (no se amortiza a capital, pero si se pagan intereses) y con una tasa de interés anual del 19% sobre saldos. Los intereses se causan y se pagan en cada período dependiendo del monto del saldo del crédito vigente. En la siguiente tabla se muestra el manejo del crédito.

Tabla XVIII. Financiamiento

Manejo de crédito				
Período	Intereses (19%)	Amortización crédito	Saldo	Total
1			Q. 3,017,280	
2	Q. 573,283		Q. 3,017,280	Q. 573,283
3	Q. 573,283	Q. 754,320	Q. 2,262,960	Q. 1,327,603
4	Q. 429,962	Q. 754,320	Q. 1,508,640	Q. 1,184,282
5	Q. 286,642	Q. 754,320	Q. 754,320	Q. 1,040,962
6	Q. 143,321	Q. 754,320	Q. -	Q. 897,641
Totales	Q. 2,006,491	Q. 3,017,280	Q. -	Q. 5,023,771

A continuación se presenta la tabla XIX que permite recoger y organizar la información sobre la utilización inicial de los recursos, es el presupuesto de inversiones con financiamiento². Cabe mencionar que en el período séptimo, año de liquidación del proyecto, se estima que los activos fijos tendrán un valor comercial calculado en el 33.33% de su valor inicial, por lo tanto se recuperan Q.1,405,993 por este concepto al liquidar el proyecto.

² De la tabla XIX en adelante se encontraran algunas cantidades con signo negativo o con paréntesis, estas son cantidades negativas que significan inversión o gasto.

Tabla XIX. Programa de inversión

Presupuesto de inversiones (con financiamiento)							
Período	1	2	3	4	5	6	7
Flujo original de inversiones	-4,310,400	0.00	0.00	0.00	0.00	0.00	1,405,993
Recursos de crédito	3,017,280						
Amortización de crédito		0.00	- 754,320	- 754,320	- 754,320	- 754,320	
Flujo ajustado de inversión	-1,293,120		- 754,320	- 754,320	- 754,320	- 754,320	1,405,993

A continuación se presenta el programa de beneficios anuales, que con la utilización de financiamiento por parte de la empresa genera una disminución en los impuestos. A partir de la ventas se resta el monto de los costos de operación para obtener la utilidad antes de impuestos, a este valor se le descuentan los impuestos (34 %), para llegar a las utilidades después de impuestos; a esta cifra se le agregan las depreciaciones, amortización de diferidos y reserva legal para obtener finalmente el flujo de efectivo o flujo de ingresos de operación, el cual se presenta a continuación.

Tabla XX. Beneficios anuales

Programa de beneficios anuales					
Período	2	3	4	5	6
Ventas	Q. 720,000	Q. 840,000	Q. 960,000	Q.1,080,000	Q.1,200,000
(-)Costos de operación	Q. 96,000	Q. 100,800	Q. 105,840	Q. 111,132	Q. 116,689
Utilidades antes de impuestos	Q. 624,000	Q. 739,200	Q. 854,160	Q. 968,868	Q.1,083,311
Interés crédito a largo plazo	Q. 573,283	Q. 573,283	Q. 429,962	Q. 286,642	Q. 143,321
Margen ajustado antes de imp.	Q. 50,717	Q. 165,917	Q. 424,198	Q. 682,226	Q. 939,991
(-)Impuestos (34%)	Q. 17,244	Q. 56,412	Q. 144,227	Q. 231,957	Q. 319,597
(-) Utilidades después de impuestos	Q. 33,473	Q. 109,505	Q. 279,970	Q. 450,269	Q. 620,394
Reserva legal (10%)	Q. 3,347	Q. 10,951	Q. 27,997	Q. 45,027	Q. 62,039
Utilidad por distribuir	Q. 30,126	Q. 98,555	Q. 251,973	Q. 405,242	Q. 558,354
(+) Depreciación	Q. 843,680	Q. 843,680	Q. 843,680	Q. 843,680	Q. 843,680
(+) Amortización diferidos	Q. 8,000	Q. 8,000	Q. 8,000	Q. 8000	Q. -
(+) Reserva legal	Q. 3,347	Q. 10,951	Q. 27,997	Q. 45,027	Q. 62,039
Flujo de efectivo	Q. 885,153	Q. 961,185	Q.1,131,650	Q.1,301,949	Q.1,464,074

Como resultado del flujo ajustado de inversiones y del flujo de efectivo, se obtiene el flujo neto de efectivo, que da una idea clara de los valores que el proyecto retribuye en cada uno de los períodos y el cual queda como en la tabla siguiente.

Tabla XXI. Flujo neto de efectivo

Flujo neto de efectivo			
Período	Flujo ajustado de inversión	Flujo de efectivo	Flujo neto de efectivo
1	Q. (1,23,120)	Q. -	Q. (1,293,120)
2	Q. -	Q. 885,153	Q. 885,153
3	Q. (754,320)	Q. 961,185	Q. 206,865
4	Q. (754,320)	Q. 1,131,650	Q. 377,330
5	Q. (754,320)	Q. 1,301,949	Q. 547,629
6	Q. (754,320)	Q. 1,464,074	Q. 709,754
7	Q. 1,405,993	Q. -	Q. 1,405,993

Ahora que se ha definido el flujo neto de efectivo, se procede aplicar los métodos de evaluación que se basan en el valor temporal del dinero para medir su rendimiento y actúan sobre actualizaciones periódicas de los flujos netos determinados para cada año de operación, considerado en el horizonte de evaluación del proyecto. Los procedimientos de actualización que se emplearán, proporcionan una base para la comparación de los ingresos y costos que se produzcan en el futuro, reduciéndolos a un valor actualizado, que determinará la factibilidad del proyecto en estudio. A continuación se presentan los análisis con los métodos de evaluación económica.

4.3.1.3.3.1. Período de recuperación de la inversión

El objetivo es determinar el tiempo en que se recupera la inversión en el proyecto, el cual es de 4.49 años. Para determinarlo se utiliza la fórmula siguiente:

$$PR = \frac{\text{Inversión}}{\text{Utilidades netas}} = \frac{Q.4,310,400}{Q.960,000} = 4.49 \text{ años} \quad (4.14)$$

4.3.1.3.3.2. Valor actual neto o valor presente

Determina el valor del dinero en el tiempo, es decir, establece lo que valdría el día de hoy una suma de dinero a recibir en el futuro. Consiste básicamente en:

- a. Se escoge una tasa de interés igual al costo de capital o la tasa de interés a la cual el dinero es prestado para financiar el proyecto.
- b. Flujos netos de efectivo o utilidades del proyecto.
- c. Números de períodos o años, que se espera se produzcan los flujos netos de efectivo.

Luego de tener los datos anteriores se procede a calcular el valor presente en base a la ecuación siguiente:

$$VAN = \sum_{j=0}^n \frac{FN_j}{(I + r)^n} \quad (4.15)$$

donde: VAN = valor actual neto o valor presente
 FN_j = flujo neto en el año j
 n = período de evaluación

Aplicando la fórmula (4.15) para varios flujos de efectivo en los distintos períodos se puede elaborar la tabla del VAN siguiente:

Tabla XXII. Cálculo del VAN

Resultados VAN						
Período	Flujo ajustado de inversión	Flujos netos de efectivo	Factor VAN (19%)	Inflación (7.10%)	Factor con inflación	VAN flujo efectivo
1	-1,293,120		1.0000	0.0000	1.0000	-1,293,120
2	0	885,153	0.8403	0.9337	0.7846	694,516
3	0	206,865	0.7062	0.8718	0.6156	127,355
4	0	377,330	0.5934	0.8140	0.4830	182,269
5	0	547,629	0.4987	0.7600	0.3790	207,559
6	0	709,754	0.4190	0.7097	0.2974	211,070
7		1,405,993	0.3521	0.6626	0.2333	328,068
VAN (+)						457,716

Observando la tabla XII se tiene un $VAN > 0$, esto indica que el proyecto es factible, y quiere decir que la rentabilidad supera la tasa de actualización o *TIR* elegida.

Teniendo en cuenta que el Sistema Integrado de Mensajería se ha diseñado para manejar mensajes en forma automática, sin que un operador se involucre en el proceso; a medida de comparación si el sistema no fuera automático, se tiene una repercusión del costo económico de cada operador que se necesita para recibir mensajes en los costos de operación.

Si para el tráfico destinado se tienen 125 operadores con un sueldo de Q.1600.00 más prestaciones en turnos diurnos y nocturnos (para el cálculo esto no se toma en cuenta), entonces en cinco años se tiene un gasto en sueldos igual a Q.12,000,000 que aumenta los costos de operación; las tablas siguientes se presentan con los costos de operación aumentados y reflejas la repercusión de ellos.

Tabla XXIII. Beneficios anuales con operadores

Programa de beneficios anuales					
Período	2	3	4	5	6
Ventas	Q. 720,000	Q. 840,000	Q. 960,000	Q. 1,080,000	Q. 1,200,000
(-)Costos de operación	Q. 12,096,000	Q. 12,700,800	Q. 13,335,840	Q. 14,002,632	Q. 14,702,764
Utilidades antes de impuestos	Q.(11,376,000)	Q.(11,860,800)	Q. (12,375,840)	Q. (12,922,632)	Q. (13,502,764)
Interés crédito a largo plazo	Q. 573,283	Q. 573,283	Q. 429,962	Q. 286,642	Q. 143,321
Margen ajustado antes de imp.	Q.(11,949,283)	Q.(12,434,083)	Q. (12,805,802)	Q. 13,209,274)	Q. (13,646,084)
(-) Impuestos (34%)	Q. (4,062,756)	Q. (4,227,588)	Q. (4,353,973)	Q. (4,491,153)	Q. (4,639,669)
(-) Utilidades después de impuestos	Q. (7,886,527)	Q. (8,206,495)	Q. (8,451,830)	Q. (8,718,121)	Q. (9,006,416)
Reserva legal (10%)	Q. (788,653)	Q. (820,649)	Q. (845,183)	Q. (871,812)	Q. (900,642)
Utilidad por distribuir	Q. (7,097,874)	Q. (7,385,845)	Q. (7,606,647)	Q. (7,846,309)	Q. (8,105,774)
(+) Depreciación	Q. 843,680	Q. 843,680	Q. 843,680	Q. 843,680	Q. 843,680
(+) Amortización Diferidos	Q. 8,000	Q. 8,000	Q. 8,000	Q. 8000	Q. -
(+) Reserva legal	Q. (788,653)	Q. (820,649)	Q. (845,183)	Q. (871,812)	Q. (900,642)
Flujo de efectivo	Q. (7,034,847)	Q. (7,354,815)	Q. (7,600,150)	Q. (7,866,441)	Q. (8,162,736)

Tabla XXIV. Flujo neto de efectivo con operadores

Flujo neto de efectivo			
Período	Flujo ajustado de inversión	Flujo de efectivo	Flujo neto de efectivo
1	Q. (1,23,120)	Q. -	Q. (1,293,120)
2	Q. -	Q. (7,034,847)	Q. (7,034,847)
3	Q. (754,320)	Q. (7,354,815)	Q. (8,109,135)
4	Q. (754,320)	Q. (7,600,150)	Q. (8,354,470)
5	Q. (754,320)	Q. (7,866,441)	Q. (8,620,761)
6	Q. (754,320)	Q. (8,162,736)	Q. (8,917,056)
7	Q. 1,405,993	Q. -	Q. 1,405,993

Tabla XXV. Cálculo del VAN con operadores

Resultados VAN						
Período	Flujo ajustado de inversión	Flujos netos de efectivo	Factor VAN (19%)	Inflación 7.10%	Factor c/inflación	VAN flujo efectivo
1	-1,293,120		1.0000	0.0000	1.0000	-1,293,120
2	0	-7,034,847	0.8403	0.9337	0.7846	-5,519,695
3	0	-8,109,135	0.7062	0.8718	0.6156	-4,992,263
4	0	-8,354,470	0.5934	0.8140	0.4830	-4,035,547
5	0	-8,620,761	0.4987	0.7600	0.3790	-3,267,167
6	0	-8,917,056	0.4190	0.7097	0.2974	-2,651,927
7		1,405,993	0.3521	0.6626	0.2333	328,060
	VAN (-)					-21,431,659

Observando la tabla XXXV se tiene un $VAN < 0$, esto indica que el proyecto no es factible, y refleja la repercusión de optar por un sistema con operadores de mensajes.

4.3.1.3.3.3. Tasa interna de retorno (TIR)

Este método también se basa en el valor temporal del dinero y en los flujos de caja durante toda la vida útil del proyecto, da el retorno porcentual que en promedio anual rinde el proyecto. La TIR utiliza el concepto del valor presente, pero intenta evitar la elección arbitraria de una tasa de interés al evaluar una inversión propuesta. Se define, como aquella tasa que iguala el valor presente de los flujos netos de todos los años del horizonte de evaluación con la inversión inicial, esto se muestra en la ecuación (4.16). Para nuestro fin económico, representa la máxima tasa de interés que se puede pagar por un préstamo, que financie las inversiones necesarias para realizar el proyecto y cuyo resultado sea no obtener pérdidas ni ganancias. Para calcular la TIR se hacen aproximaciones sucesivas hasta resolver la igualdad $VAN = 0$; esto se muestra en la tabla XXVI.

$$I = \frac{FN_1}{(1+r)} + \frac{FN_2}{(1+r)^2} + \frac{FN_3}{(1+r)^3} + \dots + \frac{FN_n}{(1+r)^n} \quad (4.16)$$

Donde: I = Inversión inicial;
 N = Horizonte de evaluación;
 FN_1, FN_2, \dots, FN_n = Flujos netos de efectivo
 r = Tasa interna de retorno

Tabla XXVI. Cálculo TIR

Resultados VAN						
Período	Flujo ajustado de inversión	Flujos netos de efectivo	TIR (32.3429%)	Inflación (7.10%)	Factor con inflación	VAN flujo efectivo
1	-1,293,120		1.0000	0.0000	1.0000	-1,293,120
2	0	885,153	0.7556	0.9337	0.7055	624,494
3	0	206,865	0.5710	0.8718	0.4978	102,969
4	0	377,330	0.4314	0.8140	0.3512	132,511
5	0	547,629	0.3260	0.7600	0.2478	135,683
6	0	709,754	0.2463	0.7097	0.1748	124,067
7		1,405,993	0.1861	0.6626	0.1233	173,397
VAN						0

El retorno porcentual del proyecto en promedio es del 32.34297%, que se encuentra arriba del 19% del requerimiento del inversionista, por lo que se considera factible y aceptado.

CONCLUSIONES

1. El Sistema Integrado de Mensajería, es una red de radiolocalización con incorporación de voz sobre *IP*, que es una de las tres tecnologías que lo complementan. Proporciona al sistema una innovación en telefonía al comunicar las redes de telefonía tradicional con las de datos. Esto permite integrar una variedad de nuevos servicios apenas imaginados y no cayendo en obsolescencia, ya que en un futuro la red tendrá el mismo tamaño que el sistema telefónico actual.
2. Permite la unificación de estructura y la interoperabilidad entre los distintos fabricantes.
3. Ahorro de costos de comunicación, ya que las llamadas entre las distintas divisiones de la empresa no tendrán costo. La voz se integra sobre una *Intranet* como un servicio más de la red, tal como otros servicios informativos. A medida que se sale de un mercado regulado (voz) y se va hacia uno desregulado (datos), hay un ahorro significativo a causa de la competencia.
4. La voz sobre *IP* es más económica que la convencional porque el sistema de encaminamiento y conmutación es más eficiente que el de las grandes centrales telefónicas, que necesitan un circuito por cada conversación, mientras que en *IP* la información se trocea en paquetes y se pueden enviar varias conversaciones multiplexadas sobre un único circuito físico.

5. La incorporación del reconocimiento de voz al Sistema Integrado de Mensajería, se dificulta el día de hoy, ya que todavía no hay un sistema de reconocimiento de voz que sea independiente del hablante (solo prototipos de laboratorio). Los que fallan en precisión y robustez en la comunicación para diferentes usuarios y diferentes entornos. Uno de los inconvenientes de la adaptación de los sistemas de reconocimiento de voz sobre la red telefónica, es debido a las limitaciones del ancho de banda, sensibilidad frente al ruido y a la variabilidad de las líneas telefónicas (coexistencia de líneas analógicas junto con líneas digitales).
6. A pesar de los inconvenientes que hay, al no contar con un sistema de reconocimiento de voz donde se pueda pronunciar las palabras como habitualmente se hace, el Sistema Integrado de Mensajería en un futuro al incorporársele el reconocimiento sobre la red telefónica, ofrece ventajas respecto al método tradicional de comunicación entre el usuario y la maquina. Permite que la comunicación sea más rápida. Permite tener las manos libres, permite movilidad, accesibilidad a usuarios con discapacidades físicas
7. *Iridium* servicio alto costo, aporta novedades de diseño de sistemas de telecomunicaciones espaciales revolucionarias: como la conmutación a bordo de los satélites, que dotan de una gran ventaja tecnológica al sistema en aras de una mayor calidad de comunicación, técnicamente hablando es un sistema bastante equilibrado que complementa el servicio de radiolocalización en todo el mundo.
8. El Sistema Integrado de Mensajería, maneja los mensajes que van a ser enviados de una forma automática (sin operador que se involucre en el proceso). Fusiona diferentes tipos de servicios de información, por ejemplo, radiolocalización, correo electrónico, correo de voz, varios servicios de voz sobre *IP*. Da una versatilidad en el servicio de envío de mensajes; por ejemplo, texto a texto, voz a texto (con reconocimiento), texto a correo electrónico, mensajes vía *Internet*.

Recomendaciones

1. El automatismo del sistema, comprime muchos servicios en un solo dispositivo de uso muy generalizado, permitiendo ampliar su panorama de acción, a través del mercadeo y seminarios que ahonden su funcionalidad.
2. El sistema es viable desde el punto de vista económico, fundamentalmente debido a su independencia de la mano de obra y una vida útil de por lo menos cinco años.
3. Debido a que involucra tecnología de punta, permitiría en una considerable medida una mejor calificación del personal a cargo.
4. A pesar de la limitaciones actuales del reconocimiento de voz, es posible utilizar un identificador de voz dependiente del usuario (hay que personalizar la voz). Lo que implica tener la necesidad de educar al usuario en el uso del sistema.
5. Ponderar el hecho que sea factible económicamente, no necesariamente significa que sea exitoso, ya que el éxito de un nuevo sistema de comunicación, depende en gran medida del mercadeo y del entorno social donde se pretenda colocar. Como por ejemplo: Primero, diferencia entre entornos urbanos y rurales y Segundo, en los sistemas de telefonía celular, es un hecho frecuente que se compren aparatos que no tienen ninguna ventaja competitiva sobre otros, es más, tienen un costo más elevado. Sin embargo son símbolo de estatus; es decir existe un valor puramente psicológico asociado (prestigio).

REFERENCIAS BIBLIOGRÁFICAS

1. Dalbor, John. **Spanish Pronunciation: Theory and Practice**. (New York, USA: Holt, Rinehart and Winston, 1969).
2. X. HUANG, Y. ARIKI and M. JACK. **Hidden Markov Models for Speech Recognition**. (Edinburgh. U.K. Edinburgh Univ. Press 1990).
3. André Girard. **Routing and Dimensioning Circuit Switched Networks**. (Addison-Wesley, 1990),
4. Jonathan Davidson y James Peters. **Fundamentos de voz sobre IP**. (Primera edición; Madrid, España: Pearson Educación, S.A. 2001).
5. Ohmori, Okumura, Kawano and Fukuda. **Field strength and Its Variability in VHF an UHF Land Mobile Radio Service**. (Tokio, Japon: Review Tokyo Elec. Commun. Lab. Vol. 16, 1968).

BIBLIOGRAFÍA

1. Armbruster Pete y Mala Laurin. **THE IRIDIUM® NETWORK FOR GLOBAL PERSONAL COMMUNICATIONS.** Satellite Communications Group, Motorola Inc., USA, August, 1997.
2. Bates, Regis J. y Donald Gregory. **Óbice and Data Communications Handbook.** USA: McGraw-Hill, 1998.
3. Davidson Jonathan y James Peters. **Fundamentos de voz sobre IP.** 1ra. Edición España: PEARSON EDUCACIÓN, S.A., 2001.
4. Duarte Cordón, Julio Cesar. **Elaboración y Evaluación de Proyectos.** 1ra. Edición Guatemala: Dpto. de Publicaciones, Facultad de Ciencias Económicas, USAC., 1995.
5. Evans, John V. **New Satellites for Personal Communications.** Scientific American (USA): 70-77. April, 1998.
6. Gradecki, Joseph D. **The Virtual Reality Construction.** USA: Editorial John Wiley & Sons, Inc., 1995.
7. Hon, A.S. **An Introduction To Paging.** Motorola. 1993.
8. Huang, Xuedong y otros. **Spoken Language Processing.** USA: Prentice Hall, 2001.
9. Huidobro, José Manuel. **Todo sobre COMUNICACIONES.** 2da. Edición España: Editorial Paraninfo, 1999.
10. Hutcheson Jonathan y Mala Laurin. **Network Flexibility of the Iridium Global Mobile Satellite System.** Proceeding of the International Mobile Satellite Conference, Ottawa, IMSC, 1995.

11. ICE. **Curso Tráfico Telefónico**. Costa Rica: Oficina de Capacitación en Telecomunicaciones, 1988.
12. INTELSAT. **TECNOLOGIA DE ESTACIONES TERRENAS**. 3ra. Edición USA: INTELSAT Assistance and Development Program, 1987.
13. Jim, W. Roberts. **Traffic Theory and the Internet**. IEEE Communications Magazine. Enero 2001.
14. Johansson, Johny K y Ikujiro Nonaka. **Implacable**. Colombia: Editorial Norma, 1997.
15. Lee, William C. Y. **Mobile Cellular Telecommunications Systems**. USA: McGraw Hill, 1989.
16. Li Qi y otros. **Recent Advancements in Automatic Speaker Authentication**. IEEE Robotics & Automation Magazine. (USA)(6): 24-34. Marzo, 1999.
17. Madden Jeff y Stuart J. Stuple. **Fundamentos de Redes**. 2da. Edición Colombia: Microsoft® Press, 1998.
18. Maine, Kris y otros. **Overview of Iridium Satellite Network**. IEEE Western Conference, San Francisco, California, USA, 1995.
19. Miller Irwin y John E. Freud. **Probabilidad y Estadística para Ingenieros**. 3ra. Edición USA: Prentice-Hall, 1985.
20. Miranda Miranda, Juan José. **Gestion de Proyectos**. 4ta. Edición Colombia: Editora Guadalupe Ltda., 2002.
21. Motorola Flex™. **Motorola New High Speed Paging Protocol**. USA: Motorola, 1993.

22. Motorola. **Inc. Telecator Alphanumeric Protocolo**. Edición electrónica, Enero 1996.
23. Motorola. **Inc. Telecator Network Paging Protocol (TNPP)**. Edición electrónica, Septiembre 1996.
24. Motorola. **Issues and Guidelines Paging Systems Emphasis**. USA: Motorola, 2000.
25. Mundo Electrónico. **Telecomunicaciones Móviles**. México: Alfaomega & Marcombo, 1995.
26. Pelton, Joseph N. **Telecommunications for the 21st Century**. Scientific American (USA): 80-85. April, 1998.
27. Ph. D. Peter A. Swan. **A Revolution in Progress: IRIDIUM[®] LEO Operations Space Operations Lessons Learned and Current Experiences**. 1997 AIAA Defense & Space Programs, Conference & Exhibit Huns, Huntsville, Alabama, USA, September 23-25, 1997.
28. Poza M. J. y otros. **Design of an Isolated Word ASR for the Spanish Telephone Network**. Proceedings of International Conference on Signal Processing, Beijing, 1990.
29. Pritchard, Wilbur L. **Satellite Communication Systems Engineering**. 2da. Edición USA: Prentice Hall, 1993.
30. Rabiner, L. R. Y B. H. Juang. **An Introduction to Hidden Markov Models**. IEEE ASSP Magazine, Enero, 1986.
31. Rabiner, Lawrence y R. Schafer. **Digital Processing of Speech Signals**. USA: Prentice Hall, 1978.

32. Rabiner; Lawrence y Biing-Hwang Juang. **Fundamentals of Speech Recognition**. USA: Prentice Hall, 1993.
33. Rosenblatt, Alfred. **Satellites Free The Mobile Pone**. IEEE Spectrum. (USA)(35): 26-35. Marzo, 1998.
34. Sheth, Dinesh I y John S. Colonias. **The History of Paging: Technology Advances Paging**. USA. Telocator, Motorola, 1993.
35. Siemens. **Telecomunicación Digital Tomo 1**. España: Editorial Marcombo, 1988.
36. Taub Herbert y Donald L. Schilling. **Principles of Communications Systems**. 2da. Edición USA: McGraw Hill, 1986.
37. Weston, J. F. y E.F. Brigham. **Fundamentos de Administración Financiera**. 5ta. Edición México: Nueva Editorial Interamericana, 1982.
38. Witowsky, William E. **IP Telephone Design and Implementation Issues**. USA: Texas Instruments, 1998.
39. Zue, Victor. **Acoustic Segmentation and Phonetic Classification in the Summit System**. In proceedings of the International Conference on Acoustics, Speech and Signal Processing, Scotland, 1989.

FORMATOS DE 2 Y 5/6 TONOS

A.1. Formatos de 2 y 5/6 tonos

En estos formatos, una sucesión de tonos de audio se usan para identificar cada radiolocalizador. Como un ejemplo podemos usar tres tipos de tonos, f_1 , f_2 , f_3 , de los cuales se pueden combinar dos al mismo tiempo, para utilizar seis distintos tonos en el radiolocalizador así: f_1f_2 , f_1f_3 , f_2f_1 , f_2f_3 , f_3f_1 y f_3f_2 .

Esta secuencia de tonos no puede ser transmitida directamente al radiolocalizador por ciertas razones, la más importante es que las frecuencias de audio no puede viajar muy lejos y la otra es que se requieren antenas sumamente grandes. El método usual las comunicaría con el radiolocalizador a través de la modulación de frecuencia (FM) del transmisor de la frecuencia. Este transmisor utiliza una frecuencia muy alta, inaudible conocida como RF (Frecuencia de Radio). En la modulación de frecuencia, la frecuencia del transmisor varía de acuerdo con la secuencia de tono. Cada tono único, causará una sola variación de la frecuencia del transmisor y por consiguiente el receptor de la localización, haría un proceso inverso de la modulación y así identificar que tono de secuencia en particular esta siendo enviado.

El número total de tonos de audio utilizados en cada secuencia determinará el máximo de radiolocalizadores, los cuales mantienen el sistema. Por ejemplo, el localizador en 2 tonos donde cada secuencia esta comprendida de dos tonos, utilizando 60 tonos diferentes obtendríamos 3,540 direcciones distintas localizadas.

A.1.1. Formato de 2 tonos

Dos tonos de audio de secuencia, representando una única dirección del radiolocalizador, están siendo utilizados para modular el portador transmitido. Los dos tonos varían en frecuencia y duración. En el receptor de radiolocalización, la modulación del tono esta siendo descifrada, y si el transmisor de tonos marca el código (determinado por el enchufe en el filtro del receptor) de este receptor, el usuario se alertará.

El formato de 2 tonos puede mantener el tono de alerta solo en el tono de voz del radiolocalizador. La opción de protección de la batería también puede proveer este formato codificado.

Un ejemplo de este formato de 2 tonos utilizado en los radiolocalizadores de voz muestra lo siguiente: el tono A se transmite a uno por segundo seguido por el tono B que se entona después de tres segundos. Este espacio de 0.7 segundos permite que el radiolocalizador active el tono antes que el transmisor de voz inicie. Si el que llama no deja su mensaje, la terminal automáticamente insertará una brecha del interlocalizador de 1.3 segundos para impedir el paso al tono B, entonces el tono A estará siendo utilizado.

El formato de 2 tonos experimenta las siguientes desventajas:

- Exceso de tono largo y períodos de lagunas.
- Limitación del número de radiolocalizadores que el transmisor de frecuencia puede soportar.

Sin embargo, muchos sistemas de radiolocalización todavía pueden soportar este formato porque existen radiolocalizadores antiguos. Otro nuevo formato puede transmitir arriba de 60 veces más rápido y tener una alta capacidad de códigos.

A.1.2. Formato de 5/6 tonos

Este es una mejora del formato de 2 tonos. En este formato, la transmisión de la frecuencia modulada por un tono de preámbulo (opcional) y cinco o seis códigos de direcciones de tonos de secuencia.

Doce frecuencias serán utilizadas para representar los dígitos del 0 al 9, repetición del tono R y un tono especial X. Arriba de 100,000 radiolocalizadores con códigos de direcciones con números desde el 00,000 hasta el 99,999 para soportar estos tonos.

Para la alarma del radiolocalizador, se utilizan los cinco tonos correspondientes a cada número de código que esta transmitiendo la secuencia. El tono identificado con la letra R (repetir), se sustituye automáticamente por el segundo de dos dígitos idénticos sucesivos cuando quiera que aparezcan en la codificación de la dirección. Por ejemplo, una dirección con código 13387 se convierte en 13R87. Esto asegura que dos dígitos sucesivos idénticos no serán descifrados como un solo dígito. El tono X es utilizado para implementar llamadas a dos radiolocalizadores al mismo tiempo, por ejemplo, indicando que una configuración diferente debe aparecer.

El formato de tono 5/6 es aproximadamente 121/2 veces más rápido que el formato de radiolocalización de 2 tonos.

B. *HANDOVER*

B.1. Definición de *handover*

El problema que surge en los satélites móviles es que un abonado móvil con una llamada en progreso, puede cruzar la frontera de la celda adyacente; en este caso, un nuevo canal debe ser asignado a la llamada en la nueva celda (celda destino), para evitar una terminación forzada de la llamada a este procedimiento se le llama *handover*.

Algunas técnicas de prioridad dan privilegio al servicio de *handover* a expensas de nuevas llamadas y se basan en:

- Canales de guarda usados exclusivamente para servir a intentos de *handover*.
- Peticiones de encolamiento de *handover*.
- Técnicas de realojamiento de canales (solo en el caso de reparto dinámico).

En los sistemas móviles, se pueden necesitar varios *handover* para completar una única llamada. Entonces, el estudio del sistema se expresa en términos de distribución del número de *handover* implicados en la duración de una única llamada y la probabilidad media de caída de una llamada. O sea, se trata de ver la probabilidad de que la llegada de una nueva llamada no pueda ser completamente servida a causa de un bloqueo inicial o un *handover* insatisfactorio.

B.2. Modelo de movilidad

El análisis del sistema depende generalmente de la técnica de administración del *handover* usada. Según la proposición de medida de la calidad de los enlaces, un *handover* debe realizarse para evitar las terminaciones forzadas de una llamada, asociadas al intervalo de tiempo durante el cual el nivel de potencia recibida desde el haz *spot* de la celda actual del abonado móvil y el haz *spot* de su celda destino están incluidas entre dos umbrales apropiados. Estos son:

- El umbral más alto está delimitado en el punto donde la potencia recibida desde el *spot beam* iluminando la celda destino, comienza a incrementarse a expensas de la reducción de la potencia desde el haz *spot* iluminando la celda actual del abonado móvil.
- El umbral más bajo es el punto donde la potencia recibida por el abonado móvil desde el *spot beam* actual está en el mínimo aceptable.

El área sobre la cual un abonado móvil puede recibir una señal con un aceptable nivel desde los dos *spot beam* adyacentes se llama área de *overlap*.

El área de *overlap* es un parámetro que afecta directamente a la realización del sistema. Obviamente un área más ancha produce una probabilidad de fallo de *handover* más baja, bajo condiciones específicas de carga de tráfico. La extensión de estas áreas dependen de:

1. de la dirección del movimiento del abonado móvil relativo al satélite.
2. de las características de las antenas del satélite.
3. de las condiciones de propagación.

Se llama celda fuente a la celda donde la llamada del abonado móvil empieza y todas aquellas celdas alcanzadas por el abonado móvil asociado durante el tiempo de vida de la llamada se les llama celda de tránsito. Entonces teniendo en cuenta la configuración de la red celular de satélite el área de *overlap* es:

$$R'*(2 - \sqrt{3}) = 0.31 * R' \quad (\text{B.1})$$

B. 3. Distribución de *handover*

En los sistemas *LEO* se necesita distinguir entre dos clases de *handover* dependiendo de la arquitectura e implementación del sistema:

- *Handover* celda a celda (*handover* haz a haz).
- *Handover* satélite a satélite.

Esta distinción se debe al hecho de que el procesamiento posterior implicado en estos *handovers* pueden ser distintos. Se define una zona como término genérico tanto para celda, haz o satélite. El área de la zona está claramente determinada por su radio el cual se define por el radio de la celda, haz o satélite.

La distribución del número de *handovers* zona a zona durante la duración de una llamada está como función de:

1. El radio de la zona r
2. La velocidad orbital del satélite *LEO*, V_{orb}
3. La duración media de la llamada $1/\mu$

Debido a la complejidad del sistema subyacente, deben asumirse algunas hipótesis para simplificarlo:

- La velocidad es constante en una dirección fija durante la duración entera de la llamada.
- Sólo están permitidos cuatro direcciones con movimientos ortogonales.
- Cuando ocurre un *handover*, la celda destino es la celda vecina en la dirección del movimiento del satélite.

El cálculo de la distribución del número de los *handovers* para una primera llamada arbitraria condiciona a la probabilidad de que una llamada atraviese al menos n *handovers*, llamada $P \geq nH$. Esta probabilidad puede ser calculada mediante condiciones adicionales para todas las direcciones del usuario (el cual es bastante determinista en el sistema *Iridium*, ya que el satélite tiene un movimiento de Sur - Norte y por tanto se considera que el abonado móvil sólo atraviesa las celdas superiores a la actual). Para cada dirección, primero se calcula la probabilidad de que un abonado cruce n límites de zonas en un tiempo fijo t ; entonces $P \geq nH$ se obtiene por darse cuenta del hecho de que para una llamada con tiempo de servicio t , si el tiempo para el límite n -ésimo cruzado es más pequeño o igual a t , asegura al menos n *handovers* antes de que la llamada termine. Por lo tanto:

$$P_{\Sigma nH} = \int_0^{\infty} P_{[T_n \leq t]} s(t) dt \quad (\text{B.2})$$

donde $P[\dots]$ es la probabilidad de que la n -ésima frontera la hayamos cruzado antes del instante t , y $s(t)$ es la función de la densidad de probabilidad de tiempo de servicio. Para simplificar la notación se define:

$$\alpha = \frac{r \cdot \mu}{V_{orb.}} \quad (B.3)$$

(este parámetro caracteriza a la movilidad del abonado). Después de algunas manipulaciones algebraicas se puede obtener para n impares ($n = 2m + 1$):

$$\begin{aligned} P_{(2m+1)H(\alpha)} &= e^{-\sqrt{3}m\alpha} \left[\frac{-2}{9\alpha^2} (1 - e^{-\sqrt{3}\alpha}) + \frac{\sqrt{3}}{9\alpha} (1 + e^{-\sqrt{3}\alpha}) \right] + e^{-2\sqrt{3}m\alpha} \left[\frac{\sqrt{3}}{9\alpha} (1 - e^{-\sqrt{3}\alpha})^2 \right] \\ &+ e^{-3m\alpha} \left[\frac{-2}{\sqrt{3}\alpha^2} (e^{-\alpha} - e^{-2\alpha}) \right] \end{aligned} \quad (B.4)$$

Análogamente, para n pares ($n = 2m > 0$) se tiene:

$$\begin{aligned} P_{(2m)H(\alpha)} &= e^{-\sqrt{3}m\alpha} \left[\frac{1}{9\alpha^2} (e^{\sqrt{3}\alpha} - e^{-\sqrt{3}\alpha}) \right] + e^{-2\sqrt{3}m\alpha} \left[\frac{\sqrt{3}}{9\alpha} (e^{\sqrt{3}\alpha} - e^{-\sqrt{3}\alpha})^2 \right] \\ &+ e^{-\sqrt{3}m\alpha} \left[\frac{1}{\sqrt{3}\alpha^2} (e^{-2\alpha} - e^{-2\alpha}) \right] \end{aligned} \quad (B.5)$$

De las ecuaciones anteriores se deriva la siguiente expresión para el número medio de *handovers* (\bar{h}):

$$\bar{h} = \frac{(3 + 2\sqrt{3})}{9\alpha} \approx \frac{0.7182 V_{orb.}}{r\mu} \quad (B.6)$$

Se puede observar que el número medio de *handovers* es igual a las veces en que una llamada de duración promedio cruza un número medio de fronteras de zonas en una unidad de tiempo por el abonado móvil. Además, el número medio de zonas limítrofes cruzadas por unidad de tiempo es una función lineal de r/V_{orb} .

B.4. Distribución de los canales ocupados

Esta distribución se define como la distribución del tiempo gastado por los abonados en una frecuencia de canal en una celda dada. El tiempo de ocupación de canal de una cierta celda sólo corresponde a la fracción de la duración total de la llamada durante el cual el abonado se localiza en esa celda en particular. Primero se calcula el tiempo medio de ocupación de un canal basado en la hipótesis de que la distribución del tiempo de ocupación de canal sigue una distribución exponencial y entonces se utiliza la prueba *Kolmogorov-Smirnov* para comprobar la utilización correcta de esta hipótesis.

Se asume que la duración de la llamada está exponencialmente distribuida con el parámetro $1/\mu$ igual a la inversa de la duración media de una llamada. La aleatorización se adopta en cada inicio de llamada y en cada *handover*. Estas presunciones llevan a la hipótesis de que la distribución exponencial puede proporcionar una mejor adaptación para la distribución de la ocupación de los canales. El parámetro de la distribución exponencial apropiada, denotada por θ es igual a la suma de la inversa de la duración media de una llamada y la media del tiempo consumido por un abonado en una celda dada:

$$\theta = \mu + \bar{h}_c \mu \quad (\text{B.7})$$

Para medir la diferencia total entre los resultados simulados y los modelos exponenciales, se usa la prueba llamada *Kolmogorov-Smirnov (K-S)*. Esta prueba permite evaluar la hipótesis de que un muestreo de datos sea sacado de una distribución continua específica, en este caso distribución exponencial.

La prueba de *Kolmogorov-Smirnov*, es una prueba paramétrica que se utiliza para diferencias entre distribuciones acumuladas. La prueba unimuestral se refiere a la concordancia entre una distribución acumulada observada de valores muestrales y una función de distribución continua determinada; es, pues, una prueba de bondad de ajuste. La prueba bimuestral está relacionada con la conformidad entre dos distribuciones acumuladas observadas; prueba la hipótesis de que dos muestras independientes provienen de distribuciones continuas idénticas y es sensible a diferencias de la población con respecto a localización, dispersión o sesgo.

La prueba de *Kolmogorov-Smirnov* unimuestral es en general más eficiente que la ji-cuadrada para bondad de ajuste en muestras pequeñas, y puede emplearse en muestras muy pequeñas donde la prueba ji-cuadrada no se aplica. La prueba unimuestral se fundamenta en la diferencia absoluta máxima D entre los valores de la distribución acumulada de una muestra aleatoria de tamaño n y una distribución teórica determinada.

La prueba está realizada mediante:

1. El cálculo de la diferencia máxima de la distribución entre la simulación y el modelo ajustado exponencialmente D_n (donde n es el tamaño de muestreo de la prueba) y
2. Comparándolo con valores críticos D_{Δ} es igual al tiempo medio de ocupación en una celda, para un nivel significativo dado. Se acepta que la hipótesis original es nula si $D_n < D_{\Delta}$ y la rechaza de otro modo.

Para realizar esta prueba se diseñaron y probaron simulaciones. Las presunciones usadas en la simulación se resumen en:

1. Se asume que todas las celdas generan el mismo tráfico, por ejemplo, todas las celdas tienen la misma velocidad de inicialización. Puede verse, que en el caso de celdas de tráfico equivalente, el tráfico medio que deja una celda, es decir, entregado a otras celdas debe compensar exactamente el tráfico medio entrante a esa celda, es decir, pasado por otras celdas. Esta presunción permite considerar sólo el tráfico generado por una celda, y no el tráfico generado por el sistema entero.
2. Este principio devuelve las llamadas de *handover* a la llamada inicial tantas veces como necesite mantener una llamada dentro de la celda inicial hasta su terminación. Después de cada *handover* o en la iniciación de una llamada, empieza un nuevo período de ocupación de canal. Cada período termina en el siguiente *handover* o cuando la llamada termina. Esta presunción nos libera del requisito de seguir el tráfico del sistema entero y así poder centrarse simplemente en una celda.
3. Se asigna una posición aleatoria del usuario de una distribución uniforme en el área de la celda. Las coordenadas angulares y radiales del usuario móvil, están relacionadas con el centro de la celda.
 - La componente radial se asume que sea una variable distribuida uniformemente en $[0,r]$ y
 - La componente angular se toma como una variable distribuida uniformemente en $[0, 2\pi]$.
 - La dirección inicial del usuario también se elige como una distribución uniforme en $[0, 2\pi]$.

El número total de llamadas varía en la simulación desde 3100 a 10000, y la duración media de la llamada se elige que sea de 2, 5, 10, 15 y 20 minutos. Cada celda contribuye a un tiempo de ocupación de canal de acuerdo al número de *handovers* que haya realizado.

B.5. Probabilidad de caída de una llamada

Bajo la presunción del caso de celda con tráfico equivalente, el número medio de *handovers* desde una celda a otras celdas es compensado por el número medio de *handovers* desde otras celdas iguales a ella. Por tanto, el número de llamadas a una celda iguala al número de llamadas generadas por la celda. Se asume que el proceso de llegada de una llamada sigue una distribución de *Poisson* con parámetros λ y θ . Anteriormente se describió que el tiempo gastado por una llamada en una celda dada sigue una distribución exponencial con parámetro θ . Dado que tanto el tiempo de llegadas de las llamadas y el tiempo que dura una llamada, siguen distribuciones exponenciales, cada celda puede ser modelada como una cola M/M/K/K, donde K denota el número de canales por celda. Si una llamada establecida requiere un canal o un *handover* llega cuando todos los canales están ocupados, la llamada simplemente se bloquea o cae. La probabilidad de llamada bloqueada se denota por P_b y se da:

$$P_b = \text{Pr ob.}[K \text{ canales están ocupados}] = \left[\sum_{n=0}^K (\lambda\theta)^n \right]^{-1} (\lambda\theta)^K \frac{1}{K!} \quad (\text{B.8})$$

Como el número de *handovers* implicados en cada llamada (incluidos los *handovers* haz a haz y satélite a satélite) es \bar{h}_c , la probabilidad de caída de una llamada se da con:

$$P_a = 1 - (1 - P_b)^{h_c} = 1 - \left[1 - \left(\sum_{n=0}^K (\lambda\theta)^n \frac{1}{n!} \right)^{-1} (\lambda\theta)^K \frac{1}{K!} \right]^{h_c} \quad (\text{B.9})$$

B.6. Tipos de *handovers* según el tipo de acceso

Para mantener la tasa de éxito de *handovers*, y mantener la alta calidad de servicio cuando un terminal entra en el área de *overlap* de dos celdas, se ha propuesto una macrodiversidad, la cual mantiene un enlace de comunicación entre el abonado móvil y las múltiples estaciones base. Según la técnica de acceso se puede distinguir dos propuestas de *handover*:

- *Seamless handover* en TDMA
- *Soft handover* en CDMA

B.6.1. *Seamless handover* en TDMA

Esta es una elegante propuesta que no interrumpe las llamadas y necesita solo un intercambio mínimo de señalización. *Seamless handover* puede ser implementado mediante la descentralización de varias funciones de los sistemas de satélites móviles.

En GSM, la reelección de celda corre a cargo casi completamente del abonado móvil, entonces cada abonado móvil desocupado servido por una cierta celda a, automáticamente reelecciona una nueva celda b tan pronto como perciba que el nivel de potencia recibida de la portadora correspondiente a la celda b es más grande que el nivel de potencia de la portadora correspondiente a la celda a; cabe destacar que el abonado móvil ni siquiera informa a la red fija sobre la celda que reelecciona, a menos que cambie de área de localización.

Contrariamente, en GSM, los *handovers* corren a cargo de la red fija (estaciones base y centros de conmutación para servicios móviles). Esta elección se hizo porque, por una parte, la mayoría de las capacidades de una llamada en progreso del abonado móvil se atendían en la llamada correspondiente (y por lo tanto no puede usarse para procesar los niveles de potencia recibidos relacionados con varias portadoras) y, por otra parte (las técnicas tradicionales de *TDMA* proveen la reserva de una cierta ranura de tiempo para la duración entera de la llamada) en cada *handover* es necesario consultar el mapa de asignación de las ranuras-canal almacenadas en la estación base, que sirve a la nueva celda; además, los procedimientos de *handover* requieren alcanzar el correcto alineamiento de la trama respecto a la nueva estación base.

Tales dificultades pueden ser superadas en el entorno de la tercera generación donde, por una parte, se espera que los usuarios móviles sean provistos con capacidades mayores de proceso y, por otra parte, una apropiada divulgación del control por las estaciones base (estaciones fijas en Tierra, en los sistemas de satélite) pueda informar al abonado móvil sobre la carga de la celda actual.

Un posible procedimiento de *handover* el cual descentraliza la mayoría de las funciones de control del abonado móvil, para mantener un mínimo de intercambios de señalización, basada en la técnica *PRMA*, puede trabajar de forma similar al procedimiento de reelección de celda en *GSM*. De todos modos, en este caso, el abonado móvil debe comprobar los paquetes de control transmitidos por la estación base que sirve a la nueva celda candidata b, para detectar el porcentaje de ranuras de tiempo disponible en las portadoras asignadas a la celda b. Si, en una de estas portadoras, llamada c, tal porcentaje es mayor que un cierto umbral (seleccionado, sobre consideraciones estadísticas, para mantener una tolerable probabilidad de colisión), se autoriza al abonado móvil a realizar el *handover* hacia la celda b. Así, conmuta su transmisor a la portadora c, mientras está todavía recibiendo de la celda antigua a, hasta

que la red fija se de cuenta de el *handover* y proporcione un enrutamiento del tráfico avanzado hacia la nueva celda b.

B.6.2. *Soft handover* en CDMA

Soft handover no supone ningún cambio de frecuencia en el sintonizador del receptor. Claramente, la implementación de esta clase de *handover* requiere la reutilización de un mismo modulo de CDMA en haces *spot* adyacentes.

En el caso de que el sistema de conmutación del satélite (*payload*) sea transparente, todas las señales en un modulo CDMA dado debe ser enrutado como si fuese una única señal; entonces, todas las señales en un modulo dado CDMA son enrutados por el mismo haz *spot*, el que está asociado con ese modulo CDMA. Cada estación fija en tierra, implicada en una llamada con un abonado móvil dado, debe transmitir sobre un modulo asociado con el haz *spot* donde el destino abonado móvil esté actualmente vagando (donde la estación fija destino está emplazada). Debería estar claro que este arreglo no pertenece a *soft handover*, ya que adyacentes *spot beam* están asociados con módulos diferentes de CDMA.

En el caso de que el sistema de conmutación del satélite se regenere e incluya un procesador de a bordo, cada señal en un cierto módulo CDMA puede enrutarse a cualquier haz *spot* independientemente de las otras señales en el mismo módulo. Entonces, el mismo módulo CDMA puede ser eficientemente reutilizado en adyacentes haces *spot*, entonces, gracias a un apropiado enrutamiento a bordo (*on-board routing*), un mismo módulo CDMA transporta diferentes señales en los distintos haces *spot* (es decir, en un haz *spot* dado solo transporta las señales actualmente direccionadas al haz *spot*). Este arreglo no pertenece a la implementación del *soft handover*. De hecho, la selección del módulo CDMA solo se puede llevar a cabo en una llamada establecida (de acuerdo con el procedimiento esbozado al final de esta sección); una llamada establecida

en un cierto módulo CDMA permanece en tal módulo durante su duración entera, a pesar de los haces *spot* que el abonado móvil cruza durante la llamada.

En el caso de que *soft handover* sea posible, puede introducirse el concepto de macrodiversidad. Una llamada en progreso abonado móvil la cual va a cambiar de celda de satélite puede simultáneamente conectarse con más de una celda satélite; el conjunto de tales celdas satélites se refieren al conjunto activo. Típicamente, un abonado móvil el cual pasaba desde un haz *spot* a otro puede ser conectado con dos celdas satélites. En el enlace de subida gracias al receptor de barrido (*rake* receptor) el usuario móvil puede combinar juntas las señales provenientes de todas las estaciones fijas en tierra en su conjunto activo; igualmente, en el enlace de bajada, las señales recibidas por las distintas estaciones fijas en tierra relacionadas con las celdas satélite en el conjunto activo de un cierto abonado móvil, pueden combinarse juntas (la red fija debe estar provista para esta misión).

B.7. Técnicas de reparto

B.7.1. Reparto fijo de canales (FCA)

En la técnica FCA, un conjunto de canales está permanentemente asignado a cada celda. El mismo conjunto de canales se reutiliza en celdas a una distancia D . La técnica básica FCA implica que un intento de llamada a una celda pueda sólo ser servida por un canal disponible perteneciente al conjunto de canales asignados a cada celda. Si no hay canales disponibles la llamada se bloquea y se pierde. El número de canales permanentemente repartidos a cada celda se deriva de:

$$S = \frac{M}{K} \quad (\text{B.10})$$

donde M = número de recursos del sistema y K = factor de reutilización.

Se llama $A(x)$ al conjunto de canales disponibles para la celda x en el instante de llegada de una llamada a x . Cuando llega una llamada debe servirse en la celda x , si resulta $A(x)$ distinto de cero, el mejor canal para el reparto se selecciona sobre la base de la evaluación de la función coste $C_x(i)$, i perteneciente a $A(x)$, de acuerdo al siguiente criterio de mínimo coste:

$$K = \frac{D^2}{3R^{12}} \quad (\text{B.11})$$

Si lo cumplen más canales, se lleva a cabo una elección aleatoria. Se puede distinguir dos alternativas según cómo se evalué la función coste. Estas son: DCA1, DCA2.

El canal i , el cual queda encerrado en un mínimo número de celdas pertenecientes a $I(x)$, se reparte a la celda x . Entonces, la función coste del reparto se define como :

$$C_x(i'') = \min_{i \in A(x)} \{C_x(i)\} \quad (\text{B.12})$$

donde

$$C_x(i) = \sum \{U_k(i)\} \quad i \in A(x) \quad \forall K \in I(x)$$

$$C_x(k, i) = U_k(i) + 2(1 - q_k(i)) \quad \forall K \in I(x)$$

Esta técnica sólo interesa para la optimización local de la asignación del canal. Esto conlleva a una ineficiente distribución de canales en la red, en particular bajo condiciones de tráfico muy adversas, es decir, la distancia de reutilización de canales es mucho mayor que D . Esto permite un incremento de llamadas bloqueadas, incluso mayor que en FCA. Para resolver este problema, se propuso la técnica DCA2.

Comienza con una técnica de reparto fija de canales, de acuerdo con la distancia de reutilización D . Se llama $F_d(x)$ al conjunto de canales repartidos según FCA a la celda x . Se fuerza a que el reparto fijo asegure una distribución de canales entre las celdas de la red, con la distancia mínima posible de reutilización. Con la técnica DCA2, aunque sea una técnica de reparto dinámico de canales, siempre que sea posible se seleccionan los canales pertenecientes a $F_d(x)$.

Entonces, la contribución al coste del reparto del canal i perteneciente a $A(x)$, debido a la interferencia de la celda k perteneciente a $I(x)$, $C_x(k,i)$, puede expresarse como:

$$qK(i) = \begin{cases} 0, & \text{si } i \in F_d(K) \\ 1, & \text{el resto} \end{cases} \quad (\text{B.13})$$

Al definir la función coste $C_x(k,i)$, relacionada con el estado del canal i en la celda k perteneciente a $I(x)$, el primer término tiene en cuenta la disponibilidad del canal i en la celda k , mientras que el segundo término representa el canal i perteneciente al conjunto óptimo para la celda k (es decir $F_d(x)$). Además de esto, se ha elegido ponderar el segundo término con un factor 2, para seguir tanto como sea posible la distribución de canales FCA. $C_x(k,i)$ sólo puede adoptar cuatro valores diferentes:

$$C_x(K, i) = \begin{cases} 0, & \text{si } i \notin A(K) \text{ y } i \notin F_D(K) \\ 1, & \text{si } i \notin A(K) \text{ y } i \in F_D(K) \\ 2, & \text{si } i \in A(K) \text{ y } i \notin F_D(K) \\ 3, & \text{si } i \in A(K) \text{ y } i \in F_D(K) \end{cases} \quad (\text{B.14})$$

Entonces el costo total se obtiene como:

$$C_x(i) = q_x(i) + \sum \{C_x(k, i)\} \quad (\text{B.15})$$

Donde:

$$\begin{aligned} K &\in I(x) \\ \forall i &\in A(x) \end{aligned}$$

donde el término $q_{x(i)}$ se introduce para tener en cuenta que es preferible repartir en x un canal i perteneciente al conjunto de canales de x ($F_D(x)$). En particular, $q_{x(i)}$ se ha incluido sólo para distinguir entre las diferentes situaciones con coste igual por las celdas interferentes, para repartir en x , si es posible, un canal perteneciente a $F_D(x)$.

B.7.2. Asignación dinámica de canales (*Dynamic Channel Allocation*)

:DCA

En una estrategia tradicional de asignación fija de canales (*FCA*), las portadoras están semipermanentemente asignadas a las celdas del satélite según las previsiones de tráfico. La reducción del radio de los haces *spot* y/o el movimiento rápido de los satélites no *GEO* suponen variaciones imprevisibles temporales y geográficas de la densidad de tráfico de la celda del satélite gran varianza del tráfico que ofrece una celda de un satélite. Entonces, para aprovechar más eficientemente los recursos del satélite, la asignación portadora-celda del satélite debe ser actualizada en tiempo real. Por otro

lado, si una entidad central de una estación de control de red (*NCS*) está encargada de la actualización de las asignaciones, se necesita una multitud de intercambios de información de señalización.

Una de las soluciones a los problemas expuestos anteriormente es la implementación de la estrategia *DCA* con control distribuido; por una parte, la asignación de portadoras a las celdas es actualizada dinámicamente en tiempo real, para poder atender la demanda, cambiante incesantemente, de canales de comunicación. Y por otro lado, las decisiones las toman las *FES* (Estaciones fijas de tierra) y no una *NCS*, reduciéndose así los intercambios de información de señalización y aumentando la robustez del sistema, el sistema de satélites ya no depende de la *NCS*.

La estrategia *DCA* prevé que tan pronto como los canales disponibles en cierta celda de satélite es decir, la suma de los canales disponibles sobre todas las portadoras asignadas actualmente a esta celda de satélite es inferior a un umbral dado T_a , la *FES* intenta una adquisición de portadora; igualmente, tan pronto como el número de canales disponibles es mayor que un umbral T_r , la *FES* libera una portadora posiblemente después de efectuar un cierto número de *handovers* intracelulares para liberar dicha portadora.

Para poder evitar decisiones conflictivas en la adquisición de portadora, cada *FES* tiene asignado un intervalo de tiempo diferente de decisión dentro de una trama de decisión de duración T ; se permite a cada *FES* intentar la adquisición de portadora solo al principio de su intervalo de tiempo es decir, cada T segundos. La presencia de los ya mencionados umbrales compensa los inconvenientes causados por los retardos en la adquisición de portadora. El intervalo de decisión debe ser suficientemente largo como para permitir que las decisiones tomadas por una *FES* puedan ser comunicadas a todas las demás *FES* vía un canal de control dedicado terrestre.

Las *FES* pueden descubrir portadoras disponibles o algunas portadoras libres bien por el enlace con otra *FES* vía un canal de control dedicado terrestre, bien en caso de que el enlace de control (*feeder link*) enlace entre la *FES* y el satélite y viceversa tenga un haz global, monitoreando el nivel de potencia recibido de algunas portadoras.

Las *FES* pueden seguir varias estrategias para seleccionar la portadora a adquirir entre el juego de portadoras disponibles, o para seleccionar la portadora a liberar de entre las que tiene asignadas la celda; la Asignación Dinámica Geométrica de Canales (*GDCA*) realiza funciones muy eficientes en la que cada *FES* trata de mantener el mismo juego de portadoras (como en *FCA*) siempre que así lo permita el tráfico ofrecido; esto significa que las portadoras con más alta prioridad para una *FES* 'a' serán las primeras que intentará recuperar y las últimas que estará dispuesta a liberar son aquellas que en un esquema *FCA* habrían estado semipermanentemente asignadas a la *FES* 'a'.