



1

**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

FACULTAD DE INGENIERIA

**HACIA LA TERCERA GENERACION DE
COMUNICACIONES MOVILES**

TESIS

**QUE PARA OBTENER EL TITULO DE:
INGENIERO MECANICO ELECTRICISTA
AREA ELECTRICA ELECTRONICA**

PRESENTA:

YOLOTZIN AGUIRRE FIGUEROA

ASESOR DE TESIS:

DR. VICTOR GARCIA GARDUÑO



CIUDAD UNIVERSITARIA, D.F. 2002

**TESIS CON
FALLA DE ORIGEN**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

PAGINACIÓN

DISCONTINUA

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: Yoletzin Aguilar

Figueras

FECHA: 2/02 octubre/2002

FIRMA: 

ESTA FENSA
DE LA BIBLIOTECA

LIBRERIA DE ORN
ESTRUC

**Por la fuerza y
apoyo espiritual
para lograr esto, se
que estas muy
orgullosa.**

**A mis hermanos y
familia que me
apoyan
incondicionalmente.**

**A mi asesor por su
inagotable
paciencia y
dedicación.**

**Por los bellos
momentos que me
das.**

ÍNDICE

INTRODUCCIÓN GENERAL	1
PARTE I	
ESTADO DEL ARTE.	1
CAPÍTULO 1	
REDES.	1
1.1 INTRODUCCIÓN.	1
1.1.1 Algunas aplicaciones de las Redes de Cómputo.	1
1.1.1.1 Uso de las Redes en Empresas.	1
1.1.1.2 Uso de las Redes en Particulares.	1
1.2 REDES DE AREA LOCAL.	2
1.2.1 Tipos Principales de Redes.	3
1.2.1.1 Punto a Punto.	3
1.2.1.2 Redes Basadas en Servidor.	3
1.2.2 Topologías de Red.	4
1.2.2.1 Bus.	4
1.2.2.2 Anillo.	4
1.2.2.3 Estrella.	5
1.2.3 Tipos de Cables.	5
1.2.3.1 Cable Coaxial.	5
1.2.3.2 Par Trenzado.	6
1.2.3.3 Fibra Óptica.	7
1.2.4 Modelos de Referencia.	8
1.2.4.1 Modelo de Referencia OS.	9
1.2.4.2 Modelo de Referencia TCP/IP.	15
1.2.5 Protocolos de Comunicación.	18
1.2.5.1 ¿Cómo trabajan los protocolos?.	18
1.2.5.2 Protocolos Dentro de una Arquitectura de Capas.	19
1.2.5.3 Pilas Estándar.	19
1.2.5.4 Estructura de los Protocolos.	20
1.2.6 Métodos de Control de Acceso al Medio.. . . .	22
1.2.6.1 CDMA/CD.	22
1.2.6.2 CDMA/CA.	23
1.2.6.3 Token Passing.	23
1.2.6.4 Prioridad Según Demanda.	23
1.2.7 Arquitecturas de Red.	24
1.2.7.1 Ethernet.	24
1.2.7.2 Token Ring.	25
1.3 REDES DE AREA EXTENSA.	25
1.3.1 Ampliación de Redes de Área Local.	25

1.3.1.1	Repetidores..	25
1.3.1.2	Puentes.	26
1.3.1.3	Enrutadores..	26
1.3.1.4	Gateways.	27
1.3.2	Transmisión de Redes de Área Amplia.	28
1.3.2.1	Conectividad Analógica.	28
1.3.2.2	Conectividad Digital.	29
1.3.3	Protocolos de Comunicación.	30
1.3.3.1	X.25.	30
1.3.3.2	Frame Relay.	31
1.3.3.3	Modo Transferencia Asíncrono (ATM).	31
1.3.3.4	Red Digital de Servicios Integrados (ISDN).	33
1.3.3.5	Interfaz de Datos Distribuidos por Fibra (FDDI).	33
1.4	REDES INALÁMBRICAS.	34
1.4.1	Aplicaciones de los sistemas inalámbricos.	34
1.4.2	Tipos de redes inalámbricas.	34
1.4.2.1	Redes de Área Local.	35
1.4.2.2	Redes de Área Local Extendida.	37
1.4.2.3	Computación Móvil.	38
1.4.3	Métodos de Control de Acceso al Medio..	39
1.4.3.1	CDMA.	40
1.4.3.2	CSMA/CD.	41
1.4.3.3	CSMA/CA.	42
1.4.3.4	TDMA.	43
1.4.3.5	FDMA.	44
 CAPÍTULO 2		
SISTEMA DE COMUNICACIONES.		
2.1	INTRODUCCIÓN.	46
2.1.1	Codificación.	46
2.1.2	Clasificación de los códigos.	46
2.1.3	Propiedades de los códigos	47
2.2	ESTRUCTURA GENÉRICA DE UN SISTEMA DE COMUNICACIÓN.	48
2.3	CODIFICACIÓN FUENTE.	49
2.2.1	Codificación fuente para una fuente discreta sin memoria.	49
2.2.2	Codificación Shannon-Fano.	50
2.2.3	Codificación de Huffman.	51
2.2.4	Codificación por Transformada Coseno Discreta.	52
2.2.5	Codificación por compensación de movimiento..	54
2.4	CODIFICACIÓN DE CANAL.	56
2.4.1	Tipos de codificación.	56
2.3.1.1	Códigos de paridad.	56
2.3.1.2	Códigos m entre n.	56
2.3.1.3	Códigos de bloque lineal y sistemático.	57
2.3.1.4	Código de Hamming.	59
2.3.1.5	Códigos cíclicos.	60

	2.3.1.6 Códigos cíclicos sistemáticos.	60
	2.3.1.7 Códigos de redundancia cíclica.	62
	2.3.1.8 Códigos convolucionales.	62
	2.4.2 Proceso de codificación.	63
2.5	NORMAS DE COMPRESIÓN DE VIDEO.	65
	2.5.1 JPEG.	65
	2.5.2 H.261.	70
	2.5.3 H.263.	76

CAPÍTULO 3

SEGUNDA GENERACIÓN DE COMUNICACIONES MÓVILES.

3.1	INTRODUCCIÓN.	80
3.2	SERVICIOS PROPORCIONADOS POR GSM.	81
3.3	ARQUITECTURA DE LA RED GSM.	82
	3.3.1 Estación móvil.	83
	3.3.2 Subsistema de estación base.	83
	3.3.3 Subsistema de red.	83
3.4	ASPECTOS DEL ENLACE DE RADIO.	84
	3.4.1 Acceso múltiple y estructura del canal.	85
	3.4.1.1 Canales de tráfico.	86
	3.4.1.2 Canales de control.	86
	3.4.1.3 Estructura de estallido.	86
	3.4.2 Codificación del discurso.	87
	3.4.3 Codificación de canal y modulación.	87
	3.4.4 Ecuación de canal y modulación.	87
	3.4.5 Salto de frecuencia.	88
	3.4.6 Transmisión discontinua.	88
	3.4.7 Recepción discontinua.	88
	3.4.8 Control de potencia.	89
3.5	Aspectos de la red.	89
	3.5.1 Administración de recursos de radio.	90
	3.5.1.1 Handover.	90
	3.5.2 Administración de movilidad.	91
	3.5.2.1 Actualización de ubicación.	91
	3.5.2.2 Autenticación y seguridad.	92
	3.5.3 Administración de comunicación.	92
	3.5.3.1 Direccinamiento de llamadas.	93

PARTE II

TERCERA GENERACIÓN DE COMUNICACIONES MÓVILES. 94

CAPÍTULO 4

	TELEFONÍA MÓVIL.	94
4.1	INTRODUCCIÓN.	94
4.2	GENERACIONES DE LA TELEFONÍA INALÁMBRICA.	96

4.3	VÍAS DE EVOLUCIÓN HACIA IMT-2000.	97
4.3.1	Evolución de GSM a UMTS.	98
4.3.1.1	HSCSD.	99
4.3.1.2	GPRS.	101
4.3.1.3	EDGE.	102
4.3.2	EVOLUCIÓN DE CdmaOne a Cdma2000.	103
4.3.2.1	IS-95 B.	103
4.3.2.2	Cdma2000 1X.	104
4.3.2.3	Cdma2000 1XEV.	104
4.4	ESTÁNDARES DE COMUNICACIÓN IMT-2000.	105
4.4.1	Avances en la estandarización de los sistemas 3G.	106
4.4.2	Sistemas 3G/IMT-2000.	108
4.4.2.1	UMTS.	108
4.4.2.2	Cdma2000.	109
CAPÍTULO 5		
ARQUITECTURAS DE TERCERA GENERACIÓN.		
5.1	TECNOLOGÍA DE TERCERA GENERACIÓN..	110
5.1.1	Tasas de transmisión.	110
5.1.1.1	Alta movilidad.	110
5.1.1.2	Movilidad total.	110
5.1.1.3	Movilidad limitada.	110
5.1.2	Nodos de red.	110
5.1.2.1	Red de radio.	110
5.1.2.2	Controlador de radio de red.	110
5.1.2.3	Núcleo de red.	111
5.1.2.4	Conexión central de red.	111
5.1.2.5	Cambios del sistema.	112
5.1.3	Aplicaciones 3G específicas.	112
5.1.3.1	Audio.	112
5.1.3.2	Voz sobre IP.	113
5.1.3.3	Imágenes fijas.	113
5.1.3.4	Imágenes en movimiento.	114
5.1.3.5	Entorno de casa virtual.	114
5.1.3.6	Agentes electrónicos	115
5.1.3.7	Descarga de software.	115
5.1.4	Generadores de tráfico en 3G.	116
5.1.5	Terminales 3G.	117
5.2	ARQUITECTURA BASADA EN DSP PARA COMUNICACIONES 3G.	118
5.2.1	Tendencia a la baja potencia de los DSPs.	120
5.2.2	Coprocesadores.	122
5.3	ARQUITECTURA PROGRAMABLE PARA OFDM-CDMA.	125
5.3.1	OFDM-CDMA.	127
5.3.2	OFCDMA.	130
5.3.3	Transceptor programable OFDM-CDMA.	133
5.3.4	Características de implantación.	135

INTRODUCCIÓN GENERAL

El objetivo de la presente tesis es dar a conocer un panorama general del paso evolutivo de las nuevas tecnologías de comunicaciones, partiendo desde el concepto de una red de computadoras, hasta llegar a las comunicaciones móviles inalámbricas conocidas como tercera generación, donde se tratará un poco la arquitectura empleada para la implantación de esta nueva generación de comunicaciones, pudiéndose observar todos los cambios que deberán existir para su puesta en marcha.

La tesis será tratada a partir de la definición y explicación de las redes de cómputo, las cuales surgen con la idea principal de compartir información, en cuanto fueron existiendo avances en la tecnología, surge la manera de comunicarse a través de Internet, en donde se tiene la red de redes, y existe plena comunicación desde cualquier punto de la tierra. En la actualidad es muy común la transmisión de voz, datos y video en tiempo real por Internet. Como otra forma de comunicación se tiene la telefonía móvil, en la cual, al inicio únicamente existía transmisión de voz, posteriormente datos y en la actualidad surge la nueva generación de comunicaciones móviles en donde éstas convergen hacia un mismo punto, en la transmisión de voz, datos y video en tiempo real desde un teléfono móvil y estando en movimiento. En esta nueva generación de comunicación encontramos la combinación del acceso móvil de alta velocidad con el protocolo basado en Internet, en donde a partir de una llamada se puede conectar a esta red y así realizar transmisión multimedia y videoconferencia, todo esto en tiempo real. La naturaleza de esta tecnología es la razón por la cual se inicia la tesis, explicando los fundamentos de redes.

Este documento se encuentra dividido en tres partes, dentro de la primera parte se encuentran tres capítulos, en el primero se hablará de las redes de cómputo, en donde se tratarán los principales tipos de redes tanto LAN, WAN como Inalámbricas, analizando las distintas topologías de red, métodos de acceso al medio, arquitecturas y protocolos de comunicación para los tres tipos de redes. Al tener un panorama general acerca de las redes, se hablará en el capítulo dos de un sistema de comunicación, con todas sus partes, analizaremos desde lo que es la codificación fuente, codificación de canal y las principales normas de compresión para la transmisión de video. En el capítulo tres se hablará acerca de la segunda generación de comunicaciones móviles, específicamente de GSM para tener un punto de partida para la explicación de la tercera generación de comunicaciones móviles, poder visualizar todos los cambios y avances de una tecnología a otra.

Dentro de la segunda parte de la tesis tendremos dos capítulos, en el capítulo cuatro se hablará de la evolución de la telefonía móvil, se podrá observar como las comunicaciones móviles han evolucionado a partir de la primera generación en donde se contaban con dispositivos analógicos, y únicamente se transmitía voz, la segunda generación en donde ya se contaban con dispositivos digitales en donde

además de voz también se puede transmitir texto, y se pueden establecer múltiples conversaciones. En la actualidad, se están empezando a desplegar sistemas de lo que se ha denominado generación 2.5 que harán de puente entre los de segunda generación y la telefonía móvil de tercera generación (la UMTS). La cual ofrecerá grandes velocidades de conexión, por lo que se espera que se convierta en la forma más habitual de acceso a Internet. Permitirá la transmisión de todo tipo de comunicaciones: voz, datos, imágenes, video.

Se puede observar que al inicio únicamente podíamos hacer y recibir llamadas de manera local o regional sin contar con casi ningún servicio adicional y existían distintos estándares de comunicación por lo cual no se podía comunicar un teléfono de una localidad con otro que perteneciera a otra localidad en donde se contara con un estándar distinto, con lo cual era imposible pensar en el roaming internacional. Uno de los principales retos a vencer a lo largo de esta evolución de las comunicaciones móviles es la estandarización, para que pueda existir comunicación a nivel mundial manejando un único estándar, mismo que se esta resolviendo en la tercera generación de comunicaciones móviles.

Otro tópico importante es la velocidad de transmisión, así como dentro de las redes de computadoras ha ido aumentando considerablemente la velocidad de transmisión entre computadoras, también se ha notado un gran aumento en la velocidad de Internet para poder manejar voz, datos y video. En la telefonía móvil también se puede observar un notable aumento en la velocidad de transmisión, el cual resulta de gran importancia debido a que podremos contar con acceso a este tipo de comunicación y así establecer videoconferencia por medio de un móvil. Todos estos avances en las comunicaciones nos llevarán a tecnologías no imaginadas, como lo puede ser el tener dispositivos electrónicos en refrigeradores y ellos serán los encargados de comunicarse directamente al supermercado y solicitar todos los productos faltantes.

Dentro del último capítulo se hablará de las distintas arquitecturas empleadas en la tercera generación de comunicaciones móviles, las tasas de transmisión a determinada velocidad de movimiento, terminales de tercera generación, arquitectura basada en DSP para comunicaciones móviles, arquitectura programable para OFDM-CDMA y las antenas inteligentes.

Por último se presentarán las conclusiones y se describirá brevemente cual es la perspectiva actual de las comunicaciones móviles, hacia dónde van y cuáles son las limitantes tanto tecnológicas como económicas.

Con base a todo lo anterior podremos tener un panorama general hacia la tercera generación de comunicaciones móviles.

PARTE I ESTADO DEL ARTE

CAPÍTULO 1 REDES

1.1 INTRODUCCIÓN

En el nivel más elemental, una red consiste en dos equipos conectados entre sí mediante un cable de tal forma que pueden compartir información. Todas las redes, no importa lo sofisticadas que sean, parten de este sencillo sistema. Aunque la idea de dos equipos conectados mediante un cable no parece extraordinaria, visto en el tiempo ha sido el mayor logro en el mundo de las comunicaciones.

Las redes surgen de la necesidad de compartir información en el momento oportuno. Las computadoras son herramientas extraordinarias que las organizaciones utilizan para producir datos, hojas de cálculo, gráficos y todo tipo de información, pero no permitían compartir rápidamente los datos que producían. Sin una red, sería necesario que los usuarios imprimieran los documentos para que otros usuarios pudieran editarlos o utilizarlos. Una situación muy común en una organización sin redes es el proporcionar discos a los demás usuarios para que copien los archivos que necesitan a su equipos. De esta manera, si cualquier persona modifica un documento no hay forma de fusionar los cambios. Esto se llamaba, y se sigue llamando, trabajar en un entorno autónomo.

1.1.1 Algunas Aplicaciones de las Redes de Cómputo

Se pueden diferenciar claramente dos tipos de usos o usuarios de las redes de cómputo, las cuales son: el profesional, que se da normalmente en una empresa, y el particular, que generalmente tiene lugar en la residencia habitual de usuario.

1.1.1.1 Uso de las Redes en Empresas

Prácticamente cualquier empresa que tenga varias computadoras hoy en día tiene una red local que las interconecta. Si la empresa dispone de varias sedes u oficinas dispersas dispondrá típicamente de una red local (LAN: Local Area Network) en cada una de ellas y de un medio de interconexión de dichas redes locales a través de enlaces telefónicos, fibra óptica, infrarrojos, etc. (también llamados accesos WAN). La red o redes permiten acceder a información importante y actualizada de manera rápida y confiable.

Por otro lado, la existencia de redes de cómputo permite a la empresa tener duplicado su servidor de base de datos, o cualquier otra información vital, de forma que en caso de falla del software, hardware, o destrucción física del servidor la información no se vea afectada, así los usuarios pueden seguir trabajando con el servidor de respaldo sin notar la falla. Esto se traduce en una mayor fiabilidad del sistema, aspecto imprescindible en algunas empresas. Por supuesto para que el sistema en su conjunto sea altamente fiable es preciso duplicar no sólo el servidor de archivos, sino la propia red de forma que no haya ningún elemento importante susceptible de falla.

La red en las empresas permite compartir recursos, tales como periféricos de elevado costo (impresoras láser, escáneres, plotters, filmadoras, etc.), o programas con el consiguiente ahorro de espacio en disco y sencillez de actualización.

Otra utilidad importante de la red en las empresas es como medio de comunicación entre sus empleados; el correo electrónico es el servicio básico, pero otros más avanzados se están implantando tales como la videoconferencia o las aplicaciones que permiten compartir un documento entre varios usuarios trabajando desde computadoras distintas.

Hasta este momento únicamente se ha tratado el aspecto de una red dentro de una empresa (Intranet). Dicha red puede conectarse hacia el exterior, ya sea directamente por medio de una línea telefónica o por medio de un *firewall*, es decir, un *gateway* intermedio que permita controlar el acceso para evitar problemas de seguridad. Cuando la red de la empresa se conecta al exterior (generalmente a Internet) aparecen una serie de nuevas aplicaciones que le dan aun mayor utilidad, entre las que cabe destacar: actividades de mercadotecnia, soporte en línea, herramientas de comunicación, etc.

Algunas empresas encuentran en Internet una manera económica de interconectar sus oficinas remotas, evitando así la contratación de líneas propias de larga distancia, pero esto con ciertas reservas, sobretodo de seguridad.

1.1.1.2 Uso de las Redes en Particulares

Las aplicaciones básicas de redes de cómputo por particulares tiene tres objetivos fundamentales:

- ↓ Acceso a información
- ↓ Comunicación
- ↓ Entretenimiento

El acceso a información actualmente se centra en el acceso a Internet y sobre todo a servidores Web. En torno a esto han aparecido multitud de servicios para

diversos fines, tales como teletrabajo, telecompra, teleenseñanza, telemedicina, etc.

La comunicación tiene lugar tanto a nivel individual (correo electrónico) como en grupos (listas de distribución, grupos de noticias, etc.). Esto incluye no solo información textual, sino también multimedia: sonido, imagen y vídeo. Además de estas aplicaciones asíncronas, en las que los participantes no han de coincidir en el tiempo, existen otras en las que sí han de coincidir, como las que permiten utilizar la computadora como un teléfono o videoconferencia para hablar con un usuario remoto a través de Internet, por medio de un módem.

El uso con fines de entretenimiento es la gran aplicación de las redes de cómputo, debido a la existencia de una gran variedad de juegos que se pueden jugar por Internet, sin la necesidad de tener instalado el software en la computadora.

1.2 REDES DE AREA LOCAL (LAN)

1.2.1 Tipos Principales de Redes

1.2.1.1 Punto a Punto

En una red punto a punto no hay servidores dedicados únicamente a compartir recursos. Todos los equipos son semejantes y por lo tanto se les conoce como "iguales". De esta forma, cada equipo funciona a la vez como cliente y servidor, y no se requiere un administrador responsable de la red, sino que cada usuario es administrador de su propio equipo. El usuario de cada equipo determina qué recurso de su equipo compartirá a través de la red.

1.2.1.2 Redes Basadas en Servidor

En un entorno con más de 10 usuarios, probablemente no resultaría adecuada una red punto a punto, con todos los equipos actuando como servidores y como clientes al mismo tiempo. Por lo tanto, la mayoría de las redes tiene servidores dedicados. Un servidor dedicado es un equipo que sólo funciona como servidor y no se utiliza como cliente o como estación de trabajo. Los servidores son peticiones de los clientes de la red para garantizar la seguridad de los archivos y de los directorios. Las redes basadas en servidor se han convertido en el modelo estándar de red.

A medida que aumenta el tamaño y el tráfico en la red, será necesario más de un servidor en la misma. Distribuir las tareas entre varios servidores asegura que cada tarea se realizará de la forma más eficiente posible.

1.2.2 Topologías de Red

Cuando hablamos de topología de una red, hablamos de su configuración. Esta configuración recoge tres campos: físico, eléctrico y lógico.

El nivel físico y eléctrico se puede entender como la configuración del cableado entre máquinas o dispositivos de control o conmutación. Cuando hablamos de la configuración lógica tenemos que pensar en como se trata la información dentro de nuestra red, como se dirige de un sitio a otro o como la recoge cada estación.

Clasificación de a cuerdo a la configuración física:

1.2.2.1 Bus

"Las redes de bus emplean la filosofía del broadcast. El bus es formado por medio de un cable donde van conectados todos los dispositivos. La información desde los dispositivos es transmitida a todos los dispositivos de manera simultánea pero únicamente es aceptada la información por el dispositivo en el cual corresponde la dirección de destino." [1]

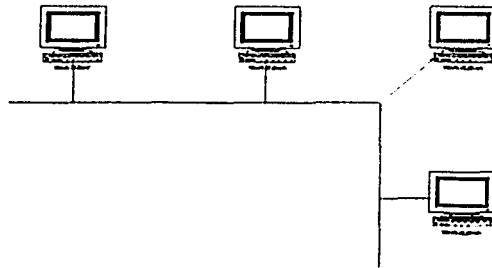


Figura 1.1. Red con topología de bus

1.2.2.2 Anillo

"Esta red consiste en un número determinado de dispositivos interconectados entre sí, formando un anillo. Cada dispositivo recibe la información la regenera y la retransmite a la máquina siguiente." [1]

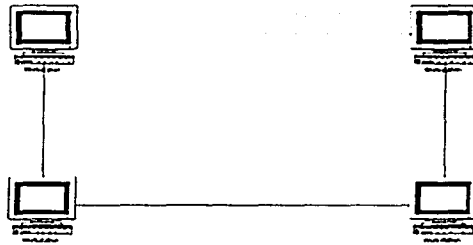


Figura 1.2. Red con topología de anillo

1.2.2.3 Estrella

"Los dispositivos de red son conectados por medio de enlaces punto a punto a un nodo central o computadora." [1]. Esta configuración está siendo eliminada poco a poco, puesto que todo el sistema se centra en la estación central y si por algún motivo fallara, todo el sistema se vería afectado.

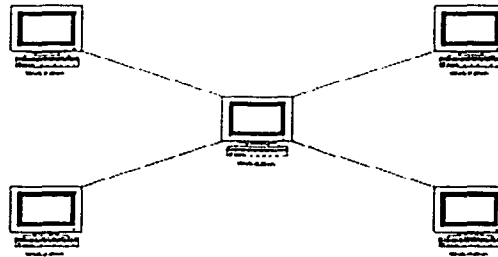


Figura 1.3. Red en estrella

1.2.3 Tipos de Cables

1.2.3.1 Cable Coaxial

En los primeros años, el cable coaxial era el que más se utilizaba en el cableado de redes. Había muchas razones para ello. En primer lugar, era relativamente barato, ligero y flexible, y era fácil trabajar con él. Era tan popular que se convirtió en la mejor forma de realizar una instalación compatible segura y fácil.

El cable coaxial más simple está formado de un núcleo de cobre rodeado por un dieléctrico, una malla metálica que tiene la función de blindaje y la cubierta exterior. La capa de dieléctrico y la malla metálica forman un blindaje doble. No obstante, en entornos sujetos a altas interferencias es necesario un blindaje cuádruple. Un aislamiento cuádruple está formado por dos capas de dieléctrico y dos capas de blindaje de malla metálica.

El blindaje protege la señal transmitida contra las interferencias de señales externas no deseadas o ruidos, de manera que no se introduzcan en el conductor interno y distorsionen la señal original.

El núcleo del cable coaxial es del encargado de transportar la señal eléctrica que contiene la información. Este núcleo puede ser macizo o trenzado. El núcleo macizo generalmente se fabrica de cobre.

El núcleo está rodeado por una capa de aislante que lo separa de la malla metálica. La malla metálica hace la función de tierra y protege al núcleo frente a ruidos e interferencias. El cable está cubierto por una capa de material no conductor, que normalmente está hecha de goma, teflón o plástico.

Cable coaxial delgado

"El cable fino es un coaxial flexible con un grosor aproximado de 6 mm y una resistencia de 50 ohms" [2]. Debido a que es flexible y es fácil trabajar con él, puede utilizarse en cualquier tipo de red. El cable fino se conecta directamente a la tarjeta adaptadora de red de los equipos.

Con cable coaxial fino se pueden transportar señales a distancias de hasta 185 metros sin que la señal sufra una atenuación apreciable.

Cable coaxial grueso

"El cable grueso es un cable relativamente rígido de un diámetro aproximado de 12 mm. A menudo se le denomina Ethernet estándar, ya que fue el primer tipo de cable que se utilizó en redes Ethernet" [2]. El núcleo es más grueso que el del cable fino. Entre más grueso sea un cable podrá transportar señales a más distancia con igual atenuación. Esto quiere decir que el cable grueso puede llevar las señales a mayor distancia que el cable fino. "El cable grueso puede transportar señales hasta 500 metros" [4]. Por lo tanto, debido a la posibilidad de poder soportar transmisiones de datos a largas distancias, es habitual que el cable grueso se utilice como cable principal para conectar varias redes pequeñas de cable fino.

1.2.3.2 Par trenzado

El cable de par trenzado más sencillo está formado por dos conductores de cobre enrollados entre sí por una cubierta aislante. Existen dos tipos de par trenzado: par trenzado no blindado o sin blindar (UTP) y par trenzado blindado (STP).

Normalmente los cables de par trenzado contienen varios pares de cable agrupados juntos en el interior de una misma cubierta o revestimiento. El número de pares varía de un cable a otro. "El enrollamiento de un conductor sobre otro anula los efectos de la interferencia de las señales de los pares adyacentes y de otras fuentes externas, tales como motores, relés o transformadores" [2].

Par trenzado no blindado (UTP)

El cable UTP que corresponde a la especificación 10BaseT es el más popular de los cables de par trenzado y se está convirtiendo rápidamente en el cableado de redes de área local más utilizado. La longitud máxima soportada por segmento es de 100 metros.

El cable UTP consta de dos conductores de cobre aislados. Dependiendo de cada aplicación particular, existen distintas normas que indican la densidad de trenzado por unidad de longitud.

El par trenzado puede ser usado tanto en comunicaciones digitales como analógicas y todas sus características son directamente proporcionales a la selección del cable.

La EIA/TIA (Electronics Industries Association / Telecommunication Industry Association) ha dividido el par trenzado en varias categorías dependiendo de sus características:

Categoría	Velocidad de Transmisión	Características
1	>1 Mbps	Hilo telefónico, transmisión de voz pero no de datos
2	>4 Mbps	Cuatro pares trenzados
3	>10 Mbps	Cuatro pares trenzados con 13 vueltas por metro
4	16 Mbps	Cuatro pares trenzados
5	>100 Mbps	Cuatro pares trenzados de hilo de cobre

Par trenzado blindado (STP)

El cable STP utiliza una cubierta de malla metálica de mayor calidad y protección que la del cable UTP. Además se utiliza una lámina metálica fina para aislar cada par y el trenzado de los pares entre sí. Estas características proporcionan al cable STP una excelente inmunidad a las interferencias externas.

Esto significa que el cable STP es menos susceptible a interferencias eléctricas y permite velocidades y distancias de transmisión mayores que el cable UTP.

1.2.3.3 Fibra óptica

"La ventaja de este medio de transmisión se basa en la frecuencia que tiene la luz, unos 10^8 MHz por lo que el ancho de banda en un sistema de transmisión óptica es muy grande. Para un bit con valor 1, un pulso de luz, para un bit con valor 0,

bastaría la ausencia de luz" [2]. Este sistema, no se ve afectado por ningún tipo de interferencia y casi la única desventaja es el hecho de no poder empalmar fácilmente cables para conectarlos a nuevos nodos.

La fibra óptica está formada por un cilindro finísimo de cristal, llamado núcleo, rodeado por una capa concéntrica de cristal, conocida como revestimiento. En ocasiones está hecho de plástico. El plástico es mucho más fácil de instalar, pero no se puede utilizar para transmitir las señales de luz a tanta distancia como con el cristal.

Cada hilo de cristal permite la transmisión en un sentido, por esta razón un cable contiene dos hilos con cubiertas separadas, uno de los cuales transmite y el otro recibe. Una capa de refuerzo de plástico rodea cada hilo de cristal, mientras que la cubierta de kevlar le proporciona resistencia.

1.2.4 Modelos de Referencia

Existen dos arquitecturas de redes más importantes en la actualidad, correspondientes a los protocolos OSI (Open Systems Interconnection) y TCP/IP (Transmission Control Protocol/Internet Protocol). La arquitectura es una identidad abstracta, más general que los protocolos o las implementaciones concretas en que luego se materializan éstos. Típicamente para cada capa de una arquitectura existirán uno o varios protocolos, y para cada protocolo habrá múltiples implementaciones. Las implementaciones cambian continuamente; los protocolos ocasionalmente se modifican o aparecen otros nuevos que coexisten con los anteriores o los dejan anticuados; sin embargo una vez definida una arquitectura ésta permanece esencialmente intacta y muy raramente modificada.

1.2.4.1 Modelo de Referencia OSI [3]

Después de la especificación de SNA por parte de IBM cada fabricante importante definió su propia arquitectura de redes; así la evolución de los productos de comunicaciones estaba garantizada, pero no se había resuelto el problema de la interoperabilidad entre diferentes fabricantes. Debido a la posición de hegemonía que IBM disfrutaba en los años 70 y principios de los ochenta la compatibilidad con IBM era un requisito necesario, por lo que la mayoría de los fabricantes tenían implementaciones de los protocolos SNA para sus productos, o estas estaban disponibles a través de terceros. Así, la forma más sencilla de interconectar dos equipos cualesquiera era conseguir que ambos hablaran SNA.

En 1977 la ISO consideró que esta situación no era la más conveniente, por lo que entre 1977 y 1983 definió la arquitectura de redes OSI con el fin de promover la creación de una serie de estándares que especificaran un conjunto de protocolos independientes de cualquier fabricante. Se pretendía con ello no favorecer a ninguno a la hora de desarrollar implementaciones de los protocolos

correspondientes, cosa que inevitablemente habría ocurrido si se hubiera adoptado alguna de las arquitecturas existentes, como la SNA de IBM o DNA (Digital Network Architecture) de Digital.

La aportación más importante de la iniciativa OSI ha sido precisamente su arquitectura. Esta ha servido como marco de referencia para describir multitud de redes correspondientes a diversas arquitecturas, ya que la arquitectura OSI es bien conocida en entornos de redes, y su generalidad y no dependencia de ningún fabricante en particular le hacen especialmente adecuada para estos fines. Por este motivo generalmente a la arquitectura OSI se le denomina Modelo de Referencia OSI, o también OSIRM (OSI Reference Model). Por extensión hoy en día se utiliza a menudo el término modelo de referencia para referirse a una arquitectura de red como lo es el Modelo de Referencia TCP/IP, el Modelo de Referencia ATM, etc.

El subsistema de comunicación total se descompone en varias capas, cada una de las cuales realiza una función bien definida. Desde el punto de vista conceptual, puede considerarse que estas capas efectúan una de dos funciones genéricas: funciones dependientes de la red o funciones orientadas a aplicaciones. Esto da pie a tres entornos operativos bien definidos:

- 1) **El entorno de red**, que se ocupa de los protocolos y normas relacionados con los diferentes tipos de redes de comunicación de datos subyacentes.
- 2) **El entorno OSI**, que comprende el entorno de red y añade otros protocolos y normas orientados a aplicaciones que permiten a los sistemas de extremo (computadoras) comunicarse entre sí abiertamente.
- 3) **El entorno de sistemas reales**, que se construye sobre el entorno OSI y se ocupa del software y servicios propietarios de un fabricante, que han sido creados con el fin de ejecutar una tarea de procesamiento de información distribuida específica.

Tanto los componentes dependientes de la red como los orientados a las aplicaciones (independientes de la red) del modelo OSI vienen implementados en varias capas. Las fronteras entre las capas y las funciones que realiza cada capa se escogieron en base con la experiencia adquirida durante la actividad de estandarización anterior.

Cada capa desempeña una función bien definida en el contexto del subsistema de comunicación global, y opera según un protocolo definido (conjunto de reglas) intercambiando mensajes (tanto datos de usuario como información de control adicional) con una capa par (similar) correspondiente a un sistema remoto. Cada capa tiene una interfaz bien definida con las capas que están inmediatamente

arriba y debajo de ella. La implementación del protocolo de una capa específica es independiente de todas las demás capas.

La estructura lógica con que cuenta el modelo de referencia OSI está constituida por siete capas de protocolos, como se muestra en la figura 1.4. Las tres capas inferiores (1 a 3) dependen de la red y se ocupan de los protocolos relacionados con la red de comunicación de datos que está usando para enlazar dos computadoras. Las tres capas superiores (5 a 7) están orientadas a las aplicaciones y se ocupan de los protocolos que permiten interactuar a dos procesos de aplicación de usuario final, casi siempre a través de una diversidad de servicios que ofrece el sistema operativo local. La capa de transporte intermedia (capa 4) oculta a las capas superiores, orientadas a las aplicaciones, los detalles del funcionamiento de las capas inferiores, dependientes de la red. En esencia, utiliza los servicios provistos por estas últimas para ofrecer a las capas orientadas a las aplicaciones un servicio de intercambio de mensajes independiente de la red.

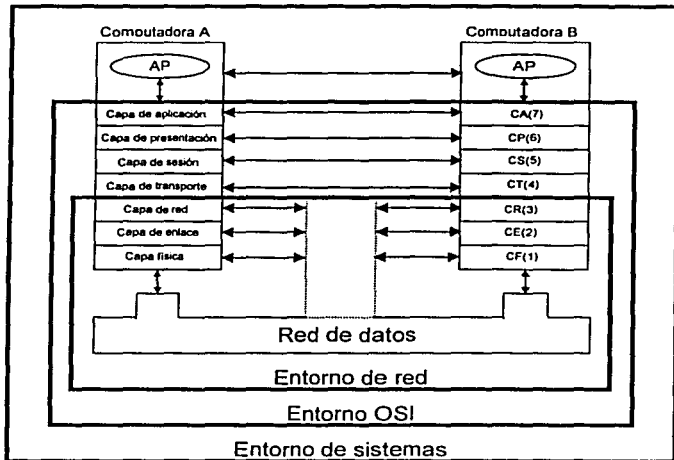


Figura 1.4. Estructura global del modelo de referencia OSI

La función de cada capa se especifica formalmente como un protocolo que define el conjunto de reglas y convenciones con las cuales la capa se comunica con una capa par en otro sistema (remoto). Cada capa proporciona un conjunto definido de servicios a la capa inmediatamente superior, y además se vale de los servicios

provistos por la capa inmediatamente inferior para transportar a la capa par remota las unidades de mensaje asociadas al protocolo. Por ejemplo, la capa de transporte ofrece a la capa de sesión que está arriba de ella un servicio de transporte de mensajes independiente de la red, y usa el servicio provisto por la capa de red que está debajo de ella para transferir el conjunto de unidades de mensaje asociadas al protocolo de transporte a una capa de transporte par de otro sistema.

Cada capa se comunica con una capa par similar de un sistema remoto siguiendo un protocolo definido, pero en la práctica las unidades de mensaje de protocolo resultantes de la capa se transfieren por medio de los servicios provistos por la siguiente capa inferior. Las funciones básicas de cada capa se muestran a continuación.

1. La Capa Física

Esta capa transmite los bits entre dos entidades (nodos) directamente conectadas. Puede tratarse de un enlace punto a punto o de una conexión multipunto (una red broadcast, por ejemplo Ethernet). La comunicación puede ser dúplex, semi-dúplex o simplex. Si la información se transmite por señales eléctricas se especifican los voltajes permitidos y su significado (1 ó 0) y análogamente para el caso de fibra óptica. Se especifican las características mecánicas del conector, la señalización básica, etc.

Como ejemplos de la capa física se puede mencionar la norma EIA RS-232-C, utilizada por los puertos COM de las computadoras personales. Las normas de redes locales incluyen en sus especificaciones la capa física (IEEE 802.3 o Ethernet, IEEE 802.5 o Token Ring, ISO 9314 o FDDI, etc.)

Muchas de las normas que existen en la capa física se refieren a la interfaz utilizada para conectar un computadora con un módem o dispositivo equivalente, que a través de una línea telefónica se conecta con otro módem y computadora en el extremo opuesto. Este es el caso por ejemplo de las normas EIA RS-232-C antes mencionado.

2. La capa de enlace

La principal función de la capa de enlace es ofrecer un servicio de comunicación confiable a partir de los servicios que se reciben de la capa física, también entre dos entidades contiguas de la red. Esto supone que se realice detección y posiblemente corrección de errores. A diferencia de la capa física, que transmitía los bits de manera continua, la capa de enlace transmite los bits en grupos denominados tramas (frames) cuyo tamaño es típicamente de unos pocos cientos a unos pocos miles de bytes. Si el paquete recibido de la capa superior es mayor que el tamaño máximo de trama, la capa física debe encargarse de fragmentarlo, enviarlo y recomponerlo en el lado opuesto. En caso de que una trama no haya

sido transmitida correctamente se deberá enviar de nuevo; también debe haber mecanismos para reconocer cuando una trama se recibe duplicada. Generalmente se utiliza algún mecanismo de control de flujo, para evitar que un transmisor rápido pueda "afectar" a un receptor lento.

Las redes basadas en servidor utilizan funciones especiales de la capa de enlace para controlar el acceso al medio de transmisión, ya que éste es compartido por todos los nodos de la red. Esto añade una complejidad a la capa de enlace que no está presente en las redes basadas en líneas punto a punto, razón por la cual en las redes basadas en servidor la capa de enlace se subdivide en dos subcapas: la inferior, denominada subcapa MAC (Media Access Control) se ocupa de resolver el problema de acceso al medio, y la superior, subcapa LLC (Logical Link Control) cumple una función equivalente a la capa de enlace en las líneas punto a punto.

Ejemplos de protocolos de la capa de enlace incluyen ISO 7776, la capa de enlace de CCITT X.25, RDSI, LAP-D, ISO HDLC. Como ejemplos de protocolos de la subcapa MAC se pueden citar los de IEEE 802.3 (Ethernet), IEEE 802.5 (Token Ring), ISO 9314 (FDDI). El protocolo de subcapa LLC de todas las redes basadas en servidor IEEE 802.2.

3. La capa de red

La capa de red se ocupa del control de la subred. Esta capa que tiene "conciencia" de la topología de red, y se ocupa de decidir por que ruta va a ser enviada la información; la decisión de la ruta a seguir puede hacerse de forma estática, o de forma dinámica basándose en información obtenida de otros nodos sobre el estado de la red.

De forma análoga a la capa de enlace la capa de red maneja los bits en grupos discretos que aquí reciben el nombre de paquetes; motivo por el cual a veces se le llama capa de paquete. Los paquetes tienen tamaños variables, pudiendo llegar a ser muy elevados, sobre todo en protocolos recientes, para poder aprovechar eficientemente la elevada velocidad de los nuevos medios de transmisión. Por ejemplo en TCP/IP el tamaño máximo de paquete es de 64 Kb, pero en el nuevo estándar, llamado Ipv6, el tamaño máximo puede llegar a ser de 4 Gb.

Entre las funciones de la capa de red cabe destacar, aparte de la ya mencionada de elegir la ruta a seguir, el control del tráfico para evitar situaciones de congestión. En el caso de ofrecer servicios con QoS el nivel de red debe ocuparse de reservar los recursos necesarios para poder ofrecer el servicio prometido con garantías. También debe ser capaz de efectuar labores de contabilidad del tráfico en caso necesario (por ejemplo si el servicio se factura basándose en la cantidad de datos transmitidos).

La capa de red es la más importante en redes de conmutación de paquetes (tales como X.25 o TCP/IP). Algunos ejemplos de protocolos utilizados en la capa de red son los protocolos de nivel paquete.

En las redes de tipo broadcast el nivel de red es casi inexistente, ya que desde un punto de vista topológico podemos considerar que en una red broadcast los nodos están interconectados todos con todos, por lo que no se toman decisiones de encaminamiento.

4. La capa de transporte

La capa de transporte es la primera que se ocupa de comunicar directamente nodos terminales, utilizando la subred como un medio de transporte transparente gracias a los servicios obtenidos de la capa de red. Por esta razón se le ha llamado históricamente la capa host-host. También se suele decir que es la primera capa extremo a extremo.

La principal función de la capa de transporte es fragmentar de forma adecuada los datos recibidos de la capa superior (sesión) para transferirlos a la capa de red, y asegurar que los fragmentos lleguen y sean recompuestos correctamente en su destino.

En condiciones normales la capa de transporte solicita a la capa de red una conexión diferente por cada solicitud recibida de la capa de sesión, pero puede haber razones de costo que aconsejen multiplexar diferentes conexiones en la capa de sesión sobre una sola conexión en la capa de red o, inversamente, razones de rendimiento pueden requerir que una conexión solicitada por la capa de sesión sea atendida por varias conexiones en la capa de red; en ambos casos la capa de transporte se ocupará de hacer la multiplexación mas adecuada de forma transparente a la capa de sesión.

La capa de transporte establece el tipo de servicio que recibe la capa de sesión, y en último extremo los usuarios. Este podría ser por ejemplo un servicio libre de errores que entrega los mensajes en el mismo orden en que se envían; también podría ser un servicio de datagramas, es decir, mensajes independientes sin garantía en cuanto al orden de entrega o falta de confirmación de la misma, o un servicio broadcast o multicast en que los paquetes se distribuyen a múltiples destinos simultáneamente.

Salvo el caso de transmisiones multicast o broadcast el nivel de transporte se ocupa siempre de una comunicación entre dos entidades, lo cual le asemeja en cierto sentido al nivel de enlace. Por esto existen grandes similitudes entre ambas capas en cuestiones tales como el control de errores o control de flujo.

Ejemplos de protocolos de transporte incluyen el CCITT X.224, también llamado protocolo de transporte OSI TP4 (Transport Protocol 4). En Internet existen dos protocolos de transporte: TCP y UDP.

5. La capa de sesión

La capa de sesión es la primera que es accesible al usuario, y es su interfaz la más básica con la red. Por ejemplo, mediante los servicios de la capa de sesión un usuario podría establecer una conexión como terminal remota de otro computadora. En un sistema multiusuario la capa de sesión se ocupa de ofrecer un SAP a cada usuario para acceder al nivel de transporte.

6. La capa de presentación

En todas las capas anteriores únicamente ocurre un intercambio de bits entre dos usuarios ubicados en dos computadoras diferentes. Se ha hecho de manera fiable y entregando los datos a la sesión, es decir al usuario, pero sin tomar en cuenta el significado de los bits transportados. La capa de presentación se ocupa de realizar las conversiones necesarias para asegurar que dichos bits se presenten al usuario de forma esperada.

7. La capa de aplicación

La capa de aplicación comprende los servicios que el usuario final está acostumbrado a utilizar en una red telemática, por lo que a menudo los protocolos de la capa de aplicación se denominan servicios. Dado que se crean continuamente nuevos servicios, existen muchos protocolos para la capa de aplicación, uno o más por cada tipo de servicio.

Ejemplos de protocolos estándar de la capa de aplicación son el CCITT X.400, X.420, X.500, FTAM, SMTP, FTP, HTTP, etc.

Transmisión de Datos en el Modelo OSI

En síntesis podemos describir la transmisión de datos en el modelo OSI de la siguiente forma; la capa de aplicación recibe los datos del usuario y les añade una cabecera, constituyendo así la PDU (Protocol Data Unit) de la capa de aplicación. La cabecera contiene información de control propia del protocolo en cuestión. La PDU es transferida a la capa de aplicación en el nodo de destino, la cual recibe la PDU y elimina la cabecera entregando los datos al usuario.

En realidad la PDU no es entregada directamente a la capa de aplicación en el nodo de destino, sino que es transferida a la capa de presentación en el nodo local a través de la interfaz; esto es una cuestión secundaria para la capa de aplicación, que ve a la capa de presentación como el instrumento que le permite hablar con su homólogo en el otro lado.

A su vez la capa de presentación recibe la PDU de la capa de aplicación y le añade una cabecera propia, creando la PDU de la capa de presentación. Esta

PDU es transferida a la capa de presentación en el nodo remoto usando a la capa de sesión como instrumento para la comunicación, de manera análoga a lo ya descrito para la capa de aplicación.

En el caso mas general cada capa añade una cabecera propia a los datos recibidos de la capa superior, y construye así su PDU. La capa homóloga del nodo de destino se ocupará de extraer dicha cabecera, interpretarla, y entregar la PDU correspondiente a la capa superior.

En algunos casos la cabecera puede no existir. En el caso particular de la capa de enlace además de la cabecera añade una cola al construir la PDU (trama) que entrega a la capa f

Las funciones básicas de cada una de las capas se resumen en la figura 1.5.

1.2.4.2 Modelo de Referencia TCP/IP [4]

La aproximación adoptada por los diseñadores de TCP/IP fue mucho más pragmática que la de los autores del modelo OSI. Mientras que en el caso de OSI se emplearon varios años en definir con sumo cuidado una arquitectura de capas donde la función y servicios de cada una estaban perfectamente definidas, y donde posteriormente se planteó desarrollar los protocolos para cada una de ellas, en el caso de TCP/IP la operación fue a la inversa; primero se especificaron los protocolos, y luego se definió el modelo como una simple descripción de los protocolos ya existentes. Por este motivo el modelo TCP/IP es mucho más simple que el OSI. También por este motivo el modelo OSI se utiliza a menudo para describir otras arquitecturas, como por ejemplo la TCP/IP, mientras que el modelo TCP/IP nunca suele emplearse para describir otra arquitectura que no sea la propia.

En el modelo TCP/IP se pueden distinguir cuatro capas: host-red, Internet, transporte y aplicación.

↓ La capa host-red

Esta capa engloba realmente las funciones de la capa física y la capa de enlace del modelo OSI. El modelo TCP/IP no contiene gran información respecto a ella, salvo que debe ser capaz de conectar el host a la red por medio de algún protocolo que permita enviar paquetes IP. Se puede decir que para el modelo TCP/IP esta capa se comporta como una "caja negra". Cuando surge una nueva tecnología de red (por ejemplo ATM) uno de los primeros datos que aparece es un estándar que especifica de que forma se pueden enviar sobre ella paquetes IP; a partir de ahí la capa Internet ya puede utilizar esa tecnología de manera transparente.

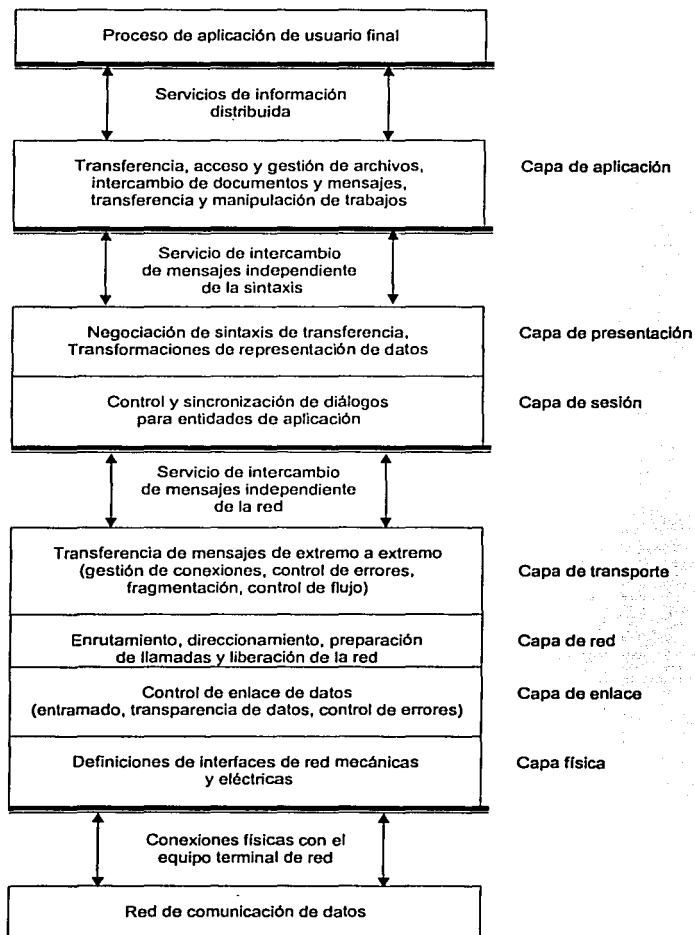


Figura 1.5. Resumen de las capas de Modelo OSI

↓ **La capa Internet**

Esta capa es el "corazón" de la red. Su papel equivale al desempeñado por la capa de red en el modelo OSI, es decir, se ocupa de encaminar los paquetes de la forma más conveniente para que lleguen a su destino, y de evitar que se produzcan situaciones de congestión en los nodos intermedios. Debido a los requisitos de robustez impuestos en el diseño, la capa Internet da únicamente un servicio de conmutación de paquetes no orientado a conexión. Los paquetes pueden llegar desordenados a su destino, en cuyo caso es responsabilidad de las capas superiores en el nodo receptor la reordenación para que sean presentadas al usuario de forma adecuada.

A diferencia de lo que ocurre en el modelo OSI, donde los protocolos para cada capa intervienen en la descripción del modelo, la capa Internet define un formato de paquete y un protocolo, llamado IP (Internet Protocol), que se considera el protocolo oficial de la arquitectura.

↓ **La capa de transporte**

Esta capa recibe el mismo nombre y desarrolla la misma función que la cuarta capa del modelo OSI, consistente en permitir la comunicación extremo a extremo (host a host) en la red. Aquí se definen dos protocolos, en primer lugar tenemos el TCP, el cual ofrece un servicio CONS fiable, con lo que los paquetes llegan ordenados y sin errores. TCP se ocupa también del control del flujo extremo a extremo, para evitar que por ejemplo un host rápido saturé a un receptor más lento. Ejemplos de protocolos de aplicación que utilizan TCP son el SMTP (Simple Mail Transfer Program, correo electrónico) y el FTP (File Transfer Program).

El otro protocolo de transporte es UDP (User Datagram Protocol) que da un servicio CLNS no fiable. UDP no realiza control de errores ni de flujo. Una aplicación típica donde se utiliza UDP es la transmisión de voz y vídeo en tiempo real; aquí el retardo que introduciría el control de errores. Otro ejemplo de aplicación que utiliza UDP es el NFS (Network File System); aquí el control de errores y de flujo se realiza en la capa de aplicación.

↓ **La capa de aplicación**

Esta capa desarrolla las funciones de las capas de sesión, presentación y aplicación del modelo OSI. La experiencia ha demostrado que las capas de sesión y presentación son de poca utilidad, debido a su escaso contenido, por lo que la aproximación adoptada por el modelo TCP/IP parece más acertada.

La capa de aplicación contiene todos los protocolos de alto nivel que se utilizan para ofrecer servicios a los usuarios. Entre estos podemos mencionar tanto las tradicionales, que existen desde que se creó el TCP/IP: terminal virtual (TelNet), transferencia de ficheros (FTP), correo electrónico (SMTP) y

servidor de nombres (DNS), como los más recientes, servicio de noticias (NNTP), el Web (HTTP), el Gopher, etc.

1.2.5 Protocolos de Comunicación

Los protocolos son un conjunto de reglas y procedimientos para la comunicación. El protocolo de comunicación determina el formato de los paquetes que transportan la producción real de datos de un dispositivo de red a otro. El servidor y cada estación de trabajo o dispositivo de la red deben manejar un protocolo común para poder comunicarse.

Existen tres puntos importantes que deben tomarse en cuenta cuando se piensa en protocolos en un entorno de red:

Existe una gran cantidad de protocolos. Mientras que cada protocolo permite comunicaciones básicas, éstos tienen diferentes propósitos y cumplen diferentes tareas. Cada protocolo tiene sus propias ventajas y restricciones.

Algunos protocolos trabajan en varias capas del modelo OSI. La capa en la cual el protocolo trabaja describe su función.

Por ejemplo, un protocolo que trabaja en la capa física, significa que el protocolo en esa capa asegura que el paquete de datos pase a través de la tarjeta de red y que sea enviado al cable de red.

Muchos protocolos pueden trabajar juntos, a esto se le conoce como pila (stack) de protocolos o suite de protocolos.

Así como la red incorpora las funciones en cada capa del modelo OSI, diversos protocolos también trabajan juntos en diferentes niveles en un stack único de protocolo. Los niveles en el stack de protocolos mapean o tienen correspondencia con las capas del modelo OSI. Trabajando juntos, los protocolos describen el funcionamiento completo del stack y sus capacidades.

1.2.5.1 ¿Cómo trabajan los protocolos?

La operación técnica completa de transmitir datos a través de la red debe de dividirse en pasos sistemáticos y discretos, en cada paso ciertas acciones se llevan a cabo, mismas que no pueden realizarse en otro paso. Cada paso tiene sus propias reglas y procedimientos o protocolo.

Los pasos deben de llevarse a cabo en un orden consistente que sea el mismo en cada computadora dentro de la red. En la computadora que envía los datos, estos pasos deben llevarse a cabo de arriba a abajo. En la máquina receptora estos pasos deben realizarse de abajo hacia arriba.

Equipo emisor

En esta computadora, el protocolo:

- Divide los datos en secciones pequeñas, llamadas paquetes, que el protocolo puede manejar.
- Añade información direccionada a los paquetes, entonces la computadora destino de la red sabrá que la información le pertenece.
- Prepara los datos para la transmisión actual a través de la tarjeta de red y los envía al cable de red.

Equipo receptor

En la computadora receptora el protocolo realiza la misma serie de pasos a la inversa de la computadora que envía los datos. La computadora receptora:

- Toma los paquetes de información del cable de red. y los trae a través de una tarjeta de red.
- Separa los paquetes de información de toda la información transmitida que fue añadida por la computadora de envío.
- Copia los datos de los paquetes en un buffer para reensamblarlos y pasa la información reensambalada a la aplicación en una forma utilizable.

Tanto la computadora de envío como la de recepción necesitan realizar cada paso del mismo modo para que no exista pérdida de información en el proceso de transmisión.

1.2.5.2 Protocolos dentro de una arquitectura de capas

En una red, varios protocolos tienen que funcionar juntos para asegurar que los datos:

- ↓ Se preparen
- ↓ Se transfieran
- ↓ Se reciban
- ↓ Se actúe sobre ellos

El trabajo de los distintos protocolos debe coordinarse de forma que no haya conflicto u operaciones incompletas. La respuesta a este esfuerzo de coordinación se denomina colocación de niveles.

1.2.5.3 Pilas de protocolos

Una pila de protocolos es una combinación de protocolos. Cada nivel especifica un protocolo distinto para manejar una función o un subsistema del proceso de comunicación. Cada nivel tiene su propio conjunto de reglas.

Capa	Definición
Aplicación	Inicia una solicitud o acepta una solicitud
Presentación	Añade formato, despliega e información encriptada al paquete

Sesión	Añade información para determinar cuando se envió el paquete
Transporte	Añade información de manejo de errores
Red	Se añade al paquete secuenciación e información direccionada
Unión de datos	Añade información que checa los errores y prepara datos para que se vayan dentro de la conexión física
Física	El paquete es enviado como una cadena de bits

Figura 1.6. El modelo OSI muestra los niveles de protocolos

Al igual que en el modelo OSI, los niveles inferiores de la pila especifican cómo pueden los fabricantes hacer que sus equipos se conecten con los equipos de otros fabricantes. Los niveles superiores especifican las reglas para dirigir las sesiones de comunicaciones y la interpretación de las aplicaciones. Cuanto más arriba estén en la pila, más sofisticadas resultan las tareas y sus protocolos asociados.

1.2.5.3 Pilas estándar

La industria informática ha diseñado varias pilas como modelos estándar de protocolos. Los más importantes incluyen:

- ↓ El conjunto de protocolos OSI
- ↓ La arquitectura de redes de sistemas (SNA: Systems Network Architecture) de IBM
- ↓ DECnet de Digital
- ↓ NetWare de Novell
- ↓ AppleTalk de Apple
- ↓ El conjunto de protocolos Internet, TCP/IP

Existen protocolos en cada nivel de estas pilas haciendo el trabajo especificado para ese nivel. Sin embargo, las tareas de comunicaciones que las redes necesitan realizar se asignan a protocolos que funcionan como uno de los tres tipos de protocolo. Estos tipos de protocolo coinciden aproximadamente con el modelo OSI:

- ↓ Aplicación
- ↓ Transporte
- ↓ Red
- ↓ Protocolos de aplicaciones

Los protocolos de aplicaciones funcionan en el nivel superior del modelo OSI

1.2.5.4 Estructura de los protocolos

"Un paquete de protocolo se compone básicamente de una porción de encabezado y de los datos. La porción de encabezado incluye información acerca

de la dirección de red del remitente, la dirección de destino, la extensión y tipo de paquete y una suma verificadora que asegura una entrega apropiada. Un ejemplo lo podemos encontrar en la a figura 1.7, el cual muestra un paquete de protocolo IPX simplificado" [5].

Verificación de suma	Extensión del paquete	Tipo de paquete	Dirección fuente	Dirección de destino	DATOS
----------------------	-----------------------	-----------------	------------------	----------------------	-------

Figura 1.7. Un paquete de protocolo IPX simplificado

Los paquetes son transportados a través de la red en tramas, las cuales son generadas por el hardware para red.

"Las diversas normas de protocolo para las LAN, que se ocupan de las capas física y de enlace en el contexto del modelo de referencia OSI, son las definidas en IEEE 802 (febrero de 1980). Esta norma define una familia de protocolos, cada uno de los cuales tiene que ver con un tipo de método de MAC específico. En la figura 1.8 se ilustran algunas normas IEEE y su relación con el modelo de referencia OSI." [3]

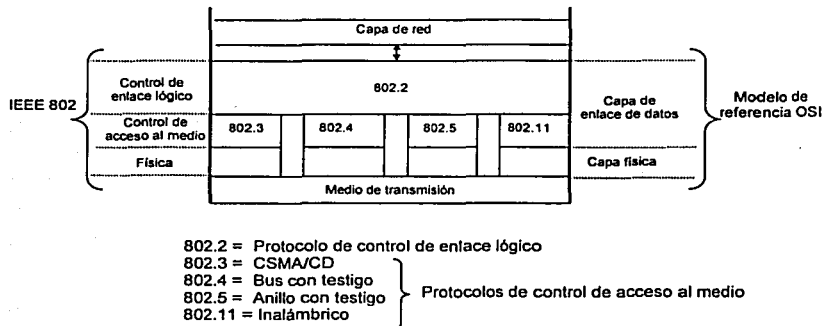


Figura 1.8. Conjunto de protocolos IEEE 802

Dentro de las LAN's hay una serie de particularidades que es interesante citar para entender como funcionan en el ámbito de comunicaciones. Las tramas transmitidas llevan marcado su origen y destino, no hay conmutadores intermedios ni tampoco existe conmutación ni encaminamiento. El control de errores, de flujo, el direccionamiento y el secuenciamiento los realiza la capa de enlace (nivel 2), el hecho de que exista un enlace que interconecta todas las estaciones y que no haya encaminamiento, hace pensar que la capa tres, no es necesaria. Esto es

cierto si tenemos una LAN aislada, pero en el momento que deseemos conectar con otra red o con otra agrupación de redes entonces la capa de red (nivel 3) cobra relevancia.

1.2.6 Métodos de Acceso al Medio

1.2.6.1 Acceso Múltiple con Detección de Portadora y Detección de Colisiones [3]

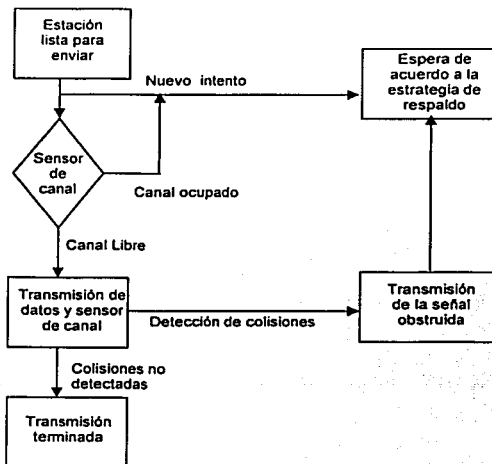
Con este método de acceso (CSMA/CD), cada equipo de la red, incluyendo los clientes y los servidores, comprueban el cable en busca de tráfico de red.

1. Un equipo "detecta" que el cable está libre, que no exista tráfico en el cable.
2. El equipo puede enviar datos.
3. Si existen datos en el cable, ningún otro equipo puede transmitir hasta que los datos lleguen a su destino y el cable esté libre de nuevo.

Si dos o más equipos están enviando datos exactamente al mismo tiempo, habrá una colisión de datos. Cuando esto ocurre, los dos equipos implicados dejan de transmitir durante un periodo aleatorio de tiempo y luego intentan retransmitir.

La capacidad de detección de colisiones es el parámetro que impone una limitación de distancia en CSMA/CD. Debido a la atenuación, el mecanismo de detección de colisiones no es efectivo más allá de los 2500 metros. Los segmentos no pueden detectar las señales más allá de esa distancia y, por tanto, no pueden darse cuenta de que un equipo en el otro extremo alejado de una red grande está transmitiendo. Si más de un equipo transmite datos sobre la red, se producirá una colisión de datos que dañará los datos.

En el siguiente diagrama de flujo se puede observar el proceso:



1.2.6.2 Acceso Múltiple con Detección de Portadora y Prevención de Colisiones [3]

"El acceso múltiple con detección de portadora y prevención de colisiones (CSMA/CA) no es tan popular como CSMA/CD ni como el token passing. En CSMA/CD, cada equipo indica su intención de transmitir antes de que transmita los datos en realidad. De esta forma, los equipos detectan cuándo se puede producir una colisión y pueden evitar las colisiones de transmisión.

Sin embargo, difundir la intención de transmitir datos incrementa la cantidad de tráfico sobre el cable y hace más lento el funcionamiento de la red. Puesto que CSMA/CA es un método de acceso más lento, goza de menos aceptación que CSMA/CD"

1.2.6.3 Token Passing [3]

En token passing, un tipo especial de paquete llamado token circula alrededor del anillo del cable de un equipo a otro. Cuando cualquier equipo del anillo de será enviar a través de la red, debe esperar un token libre. Cuando se detectan un token libre, el equipo puede tomar el control del mismo.

El equipo puede transmitir datos ahora. Los datos se transmite en tramas y la información adicional, como direccionamiento, se adjunta a la trama en forma encabezados y colas.

Mientras un equipo está usando el token, los demás equipos no pueden transmitir datos. Puesto que sólo un equipo puede usar el token a la vez, no hay lucha por transmitir, ni colisiones y no se pierde tiempo esperando a que los equipos envíen los tokens de nuevo debido al tráfico de red sobre el cable.

1.2.6.4 Prioridad según demanda [3]

Prioridad según demanda es un método de acceso relativamente nuevo, diseñado para el estándar Ethernet de 100 Mbps llamado 100VG-AnyLAN. Lo aprobó y estandarizó el IEEE en 802.12.

Este método de acceso se basa en el hecho de que los concentradores y nodos finales son los dos componentes que forman todas las redes 100VG-AnyLAN. Los concentradores administran el acceso a la red realizando búsquedas "round-robin" de peticiones para enviar desde todos los nodos de la red. El concentrados es responsable de anotar todas las direcciones, vínculos y nodos finales, y comprobar que todos están funcionando. De acuerdo con la definición de 100VG-AnyLAN, un nodo final podría ser un equipo, un puente, un enrutador o un conmutador.

Al igual que en el caso de CSMA/CD, dos equipos pueden provocar una colisión transmitiendo exactamente al mismo tiempo. Sin embargo, con la prioridad según demanda, es posible implementar un esquema en el que ciertos tipos de datos

recibirán prioridad si existiera lucha por transmitir. Si el concentrados o repetidor recibe dos peticiones al mismo tiempo, se atenderá primero a la petición de mayor prioridad. Si las dos peticiones son de la misma prioridad, se atenderá a ambas peticiones alternando entre las dos.

1.2.7 Arquitecturas de Red

1.2.7.1 Ethernet

En la siguiente tabla se resumen los principios básicos de Ethernet.

Topología tradicional	Bus lineal
Tipo de transmisión	Banda base
Método de acceso	CSMA/CD
Especificaciones	IEEE 802.3
Velocidad de transmisión	10 Mbps o 100 Mbps
Tipos de cables	Coaxial delgado, grueso y UTP

Estándares IEEE de 10 y 100 Mbps

Ethernet	Características
10-BASE-5	Es la única especificada en el estándar original ANSI/IEEE de 1985. Ethernet gruesa (coaxial grueso), la velocidad de transmisión es de 10 Mbps usando la codificación Manchester en banda base. Los segmentos tienen como máximo una distancia de 500 metros, no pudiendo conectar más de 100 estaciones en un segmento. Los equipos se conectan a la red a través de un transconductor.
10-BASE-2	Ethernet fina (coaxial fino de 50 ohms), también llamada "Cheapernet" la velocidad de transmisión es de 10 Mbps y los segmentos tienen como máximo 185 metros, pudiendo conectar los equipos a la red directamente aunque solo permite 30 nodos por segmento.
10-BASE-T	Alcanza una velocidad de 10 Mbps sobre un par trenzado no apantallado, cuya longitud máxima es de 100 metros. Su topología es un bus con forma de estrella.
100-BASE-X	Llamada Fast-Ethernet (Ethernet rápida), alcanza una velocidad máxima de 100 Mbps, puede ir con UTP (Par trenzado sin apantallar), STP (UTP apantallado) o con fibra óptica.
100VG-AnyLAN	Es una tecnología que combina elementos de Ethernet y Token Ring. Construida con una topología en estrella y velocidad de transmisión de 100 Mbps sobre cable de par trenzado de categoría 3, 4 y 5 así como fibra óptica. Es una especificación 802.12 del IEEE.

1.2.7.2 Token Ring

En la siguiente tabla se resumen los principios básicos de Ethernet.

Topología tradicional	Anillo
Transmisión	Banda base
Método de acceso	Token passing
Especificaciones	IEEE 802.5
Velocidad de transmisión	4 Mbps o 16 Mbps
Tipos de cables	UTP y STP (1, 2 y 3 de IBM)

1.3 REDES DE AREA EXTENSA

1.3.1 Ampliación de Redes de Área Local

En la medida en que las organizaciones crecen, también crecen sus redes. Las redes LAN tienden a crecer más allá de sus diseños originales. Siendo esto evidente cuando el cable comienza a estar ocupado todo el tiempo por el tráfico de la red.

Llegará un momento en el que cada administrador necesitará aumentar el tamaño de su red o mejorar su rendimiento. Las redes no se pueden ampliar simplemente agregando más equipos y más cables. Cada topología o arquitectura tiene sus límites. Sin embargo, hay componentes que se pueden instalar para incrementar el tamaño de una red dentro de su entorno existente. Estos componentes pueden:

- ↯ Segmentar las LAN de manera que cada segmento se convierta en su propia LAN
- ↯ Unir dos LAN separadas
- ↯ Conectarse con otras LAN y con otros entornos para unirlos y formar una red mayor.

Los componentes que permiten lograr estos objetivos son:

1.3.1.1 Repetidores

"Los repetidores trabajan en el nivel físico del modelo OSI para regenerar las señales de red y volver a enviarlas a otros segmentos." [6]

El repetidor recibe la señal débil de un segmento, la regenera y la pasa al siguiente segmento. Para poder pasar exitosamente los datos entre los segmentos a través del repetidor, los paquetes y los protocolos del Controlador de Enlace Lógico (LLC) tienen que ser iguales en cada segmento. Esto quiere decir que un repetidor no permite la comunicación, por ejemplo, entre una LAN 802.3 y una LAN 802.5.

Los repetidores no traducen ni filtran nada. Para que un repetidor funcione, los segmentos que éste tiene tienen que tener el mismo método de acceso. Un repetidor no puede conectar un segmento que utilice CSMA/CD con un segmento que use Token Passing. Esto es, no puede traducir un paquete Ethernet a un paquete Token Ring.

Los repetidores envían cada bit de datos desde un segmento de cable hacia otro, incluso si los paquetes de datos están mal formados o no están destinados para su uso en la red. Esto supone una dificultad debido a que un segmento puede ocasionar problemas a todos los demás segmentos. Los repetidores no actúan como filtros para restringir el flujo del tráfico que origina problemas.

1.3.1.2 Puentes

Al igual que los repetidores, los puentes pueden unir segmentos o grupos de trabajo en las LAN. Sin embargo, un puente también puede dividir una red para aislar el tráfico o los problemas.

"Los puentes trabajan en el nivel de enlace del modelo OSI" [6]. Debido a que se encuentran en este nivel, toda la información de los niveles superiores del modelo OSI no estará disponible para ellos. Por lo tanto, los puentes no distinguen entre los protocolos. Los puentes simplemente pasan todos los protocolos a través de la red. Por ello, cada equipo tiene que determinar en forma individual los protocolos que puede reconocer.

Los puentes se basan en el principio de que cada nodo de la red tiene su propia dirección, y reenvían los paquetes basándose en la dirección del nodo de destino.

Los puentes tienen algún grado de inteligencia debido a que son capaces de aprender a reenviar los datos. Cuando el tráfico pasa a través del puente, la información acerca de las direcciones de los equipos se almacena en la memoria RAM del puente. El puente usa su memoria RAM para construir una tabla de enrutamiento basándose en las direcciones origen.

Inicialmente, la tabla de enrutamiento de los puentes está vacía. A medida que los nodos van transmitiendo paquetes, la dirección origen se copia en dicha tabla. Con esta información de direcciones, el puente sabe cuáles son los equipos que se encuentran en cada segmento de la red.

1.3.1.3 Enrutadores

En un entorno formado por varios segmentos de red con protocolos y arquitecturas diferentes, un puente tal vez no sea lo más adecuado para asegurar la rápida comunicación entre todos los segmentos. Una red tan compleja necesita un dispositivo que no sólo conozca la dirección de cada segmento, sino que también pueda determinar la mejor ruta para enviar los datos y filtrar el tráfico hacia el segmento local.

"Los enrutadores trabajan en el nivel de red del modelo OSI"[6]. Esto significa que pueden conmutar y enrutar paquetes entre múltiples redes. Esto lo hacen intercambiando información específica de los protocolos entre las distintas redes. Los enrutadores leen la compleja información de direcciones de red del paquete y debido a que trabajan en un nivel superior de los puentes, tienen acceso a información adicional.

La tabla de enrutamiento contiene las direcciones de red. Sin embargo, las direcciones de los anfitriones se pueden guardar dependiendo del protocolo que esté ejecutando la red. El enrutador usa una tabla para determinar la dirección destino de los datos entrantes. La tabla contiene la siguiente información:

- ↓ Todas las direcciones de red conocidas
- ↓ Cómo conectarse con otras redes
- ↓ Las posibles rutas entre los enrutadores

El enrutador selecciona la mejor ruta para los datos basándose en las rutas disponibles y tomando en cuenta la más óptima.

1.3.1.4 Gateways

Las puertas de enlace o *gateways* hacen posible la comunicación entre las diferentes arquitecturas y los entornos. Vuelven a empaquetar y convierten los datos que van desde un entorno hacia otro, de manera que cada uno de los entornos pueda entender los datos de los otros. Las puertas de enlace vuelven a empaquetar la información para que coincida con los requisitos del sistema de destino. Las puertas de enlace tienen la capacidad de cambiar el formato de un mensaje para que se ajuste al programa de aplicación del extremo receptor de la transferencia.

Las puertas de enlace son específicas para cada tarea, lo cual significa que están dedicadas a un tipo particular de transferencia. La puerta de enlace toma los datos de un entorno, elimina su antigua pila de protocolos y los vuelve a empaquetar con la pila de protocolos de la red de destino.

Para procesar los datos, la puerta de enlace o *gateway*:

- ↓ Descapsula los datos entrantes a través de toda la pila de protocolos de la red.
- ↓ Encapsula los datos salientes con la pila de protocolos de la otra red para permitir la transmisión.

"Algunas puertas de enlace o *gateways* usan los siete niveles del modelo OSI, pero normalmente realizan la conversión del protocolo en el nivel de aplicación, sin embargo, el nivel de funcionalidad varía ampliamente entre los distintos tipos de puertas de enlace o *gateways*" [6].

1.3.2 Transmisión en Redes de Área Amplia (WAN)

La mayoría de las redes WAN son combinaciones de LAN y de otros tipos de componentes de comunicaciones conectados a través de enlaces de comunicación llamados enlaces WAN. Algunos de los enlaces son:

- a) Redes de conmutación por paquetes
- b) Cable de fibra óptica
- c) Transmisores de microondas
- d) Enlaces por satélite
- e) Sistemas coaxiales de televisión por cable

Los enlaces WAN, así como las conexiones telefónicas de área amplia, son demasiado costosos y complejos como para que las empresas privadas los puedan adquirir, implementar y mantener. Por lo tanto, normalmente se obtienen en renta por parte de los proveedores de estos servicios.

La comunicación entre diversas LAN requiere una de las siguientes tecnologías de transmisión:

- a) Analógica
- b) Digital
- c) Conmutación de Paquetes

1.3.2.1 Conectividad analógica

La misma red que utiliza su teléfono está disponible para comunicar los equipos. El nombre de esta red mundial es: red telefónica pública conmutada (PSTN). En el entorno informático, la PSTN, que ofrece líneas telefónicas de marcación con calidad de voz, puede verse como un gran enlace WAN.

El hecho de que la PSTN fuera diseñada para la comunicación de voz la hace lenta, y las líneas analógicas de acceso telefónico requieren módems que las pueden volver incluso más lentas. Ya que la PSTN es una red de circuito conmutado, las conexiones no tienen una calidad consistente. La calidad de una sesión de comunicación dependerá de los circuitos que intervengan en esa sesión particular.

Líneas Analógicas Dedicadas

A diferencia de las líneas de acceso telefónico que tienen que volver a abrirse cada vez que se van a usar, las líneas analógicas dedicadas (o rentadas) proporcionan un enlace de comunicación preparado para su uso. Una línea analógica dedicada es más rápida y confiable que una conexión de acceso telefónico. También es relativamente costosa porque la portadora dedica recursos a la conexión rentada aunque ésta no se utilice.

1.3.2.2 Conectividad Digital

Las organizaciones que necesiten un entorno de transmisión más seguro y rápido pueden optar por las líneas de servicio de datos digitales (DDS: Digital Data Service). Las DDS proporcionan comunicaciones síncronas punto a punto a 2.4 y, 4.8, 9.6 o 56 Kbps. Los circuitos digitales punto a punto son circuitos dedicados que se ofrecen a través de varios proveedores de servicios de telecomunicaciones. El proveedor garantiza el ancho de banda full-dúplex (completamente bidireccional) al establecer un enlace permanente entre cada extremo.

La razón principal por la que los clientes usan líneas digitales es porque el 99 por ciento de los casos proporcionan una transmisión libre de errores. Las líneas digitales están disponibles en varias formas, incluyendo DDS, T1, T3, T4 y 56 conmutado.

Ya que DDS emplea la comunicación digital, no necesita módems. En su lugar, DDS envía los datos desde un puente o enrutador a través de un dispositivo llamado unidad de servicio de canal/unidad de servicio de datos (CSU/DSU: Channel Service Unit/Data Service Unit). Este dispositivo convierte las señales digitales estándar que el equipo genera en señales digitales (bipolares) que forman parte del entorno de comunicaciones síncrono. También contiene circuitos electrónicos para proteger la red del proveedor de servicios DDS.

56 Conmutado

Las empresas telefónicas de servicios locales y de larga distancia ofrecen este servicio de acceso telefónico digital de LAN a LAN que transmite datos a 56 Kbps. El 56 conmutado sólo es una versión de circuito conmutado de la línea DDS a 56 Kbps. La ventaja de este sistema es que se usa bajo solicitud, lo cual elimina el costo de las líneas dedicadas. Cada equipo que emplea este servicio deberá contar con un CSU/DSU que pueda llamar a otro sitio que también utilice un servicio 56 conmutado.

Conmutación por Paquetes

El paquete de datos original se divide en otros paquetes y cada uno se etiqueta con una dirección destino además de que incluye otro tipo de información. Esto hace posible enviar cada paquete por separado a través de la red.

En la conmutación por paquetes, los paquetes se retransmiten a través de las estaciones de una red de computadoras por la mejor ruta disponible entre el origen y el destino.

Cada paquete se transmite en forma separada. Dos paquetes procedentes del mismo paquete de datos original pueden seguir caminos totalmente diferentes para alcanzar el mismo destino. Las rutas de datos de los paquetes individuales dependen de la mejor ruta que esté disponible en cada momento.

Incluso aunque cada paquete pueda viajar a través de una ruta diferente y aunque los paquetes que componen un mensaje puedan llegar en momentos diferentes o fuera de secuencia, el equipo receptor puede volver a componer el mensaje original.

El tamaño del paquete se mantiene pequeño. Si se presenta un error en la transmisión, la retransmisión de un paquete pequeño será más sencilla que la de un paquete grande. Asimismo, los paquetes pequeños mantienen ocupado al nodo durante periodos más cortos.

Circuitos Virtuales

Muchas redes de conmutación por paquetes usan circuitos virtuales. Estos son circuitos compuestos de una serie de conexiones lógicas entre el equipo emisor y el receptor. El circuito no es un cable real, sino el ancho de banda asignado bajo solicitud, en oposición a un enlace físico y permanente entre dos estaciones.

La conexión se realiza después de que ambos equipos intercambian información y acuerdan los parámetros de comunicación que establecen y mantienen la conexión. Estos parámetros incluyen el tamaño máximo del mensaje y la ruta que van a tomar los datos.

Los circuitos virtuales incorporan parámetros de comunicación para asegurar la fiabilidad. Entre ellos se incluyen:

- ↓ Confirmaciones
- ↓ Control de flujo
- ↓ Control de errores

Los circuitos virtuales pueden durar tanto como la conversación (temporales) o mientras los dos equipos que se comunican estén en funcionamiento (permanentes).

Circuitos Virtuales Conmutados (SVC:Switched Virtual Circuit)

En los SVC, la conexión entre los equipos sigue una ruta específica a través de la red. Se dedican recursos al circuito y la ruta se mantiene hasta que la conexión termina. También se conocen como conexiones de un punto a múltiples puntos.

Circuitos Virtuales Permanentes (PVC:Permanent Virtual Circuit)

Los PVC son similares a las líneas rentadas en el sentido de que son permanentes y virtuales. La diferencia es que el cliente sólo paga por el tiempo de uso de la línea.

1.3.3 Protocolos de Comunicación

1.3.3.1 X.25 [3]

X.25 es un conjunto de protocolos incorporados a una red de conmutación por paquetes. La red de conmutación por paquetes se compone de servicios de

conmutación que se establecieron originalmente con la intención de conectar terminales remotas hacia sistemas anfitriones grandes.

Una red X.25 de conmutación por paquetes emplea conmutadores, circuitos y enrutadores disponibles, con el fin de proporcionar la mejor ruta en cada momento. Debido a que estos componentes cambian rápidamente dependiendo de la necesidad y de la disponibilidad, a veces se les caracteriza como nubes. Las nubes indican una situación siempre cambiante o la carencia de un conjunto estándar de circuitos.

Las primeras redes X.25 trabajaban sobre líneas telefónicas para transmitir los datos. Éste era un medio poco confiable que producía muchos errores, así que X.25 incorporó una comprobación de errores extensa. Debido a la corrección de errores y a las retransmisiones, X.25 puede parecer lento.

El conjunto actual de protocolos X.25 define la interfaz entre un anfitrión (u otro dispositivo en modo de paquete síncrono) y la red de datos pública (PDN: Public Data Network) sobre un circuito dedicado o rentado. Dicha interfaz es en realidad una interfaz de equipo terminal de datos/equipo de comunicaciones de los datos (DTE/DCE: Data Terminal Equipment/Data Communication Equipment).

Entre los ejemplos de DTE se incluyen los siguientes:

Un equipo anfitrión con una interfaz X.25

Un ensambladora/desensamblador de paquetes (PAD: Packet Assembler/Disassembler) que recibe caracteres asíncrono desde una terminal de baja velocidad y los ensambla en paquetes para transmitirlos a través de la red. El PAD también desensambla los paquetes recibidos desde la red de tal forma que los datos se puedan entregar a las terminales como caracteres.

Una puerta de enlace o *gateway* entre el PDN y una LAN o una WAN

Para estos tres ejemplos de DTE, la mitad DCE de cada DTE/DCE es la PDN.

1.3.3.2 Frame Relay [3]

A medida que las comunicaciones de red se mueven hacia entornos digitales y de fibra óptica, parecerá que las nuevas tecnologías requieren una menor comprobación de errores que los primeros métodos analógicos de conmutación por paquetes.

Frame relay es una tecnología avanzada rápida, digital y de conmutación por paquetes de longitud variable. Con esta tecnología, los diseñadores han dejado de lado muchas funciones de recuento y comprobación que se emplean en X.25, pero que no son necesarias en un entorno de fibra óptica, el cual es confiable y seguro.

Frame relay es un sistema punto a punto que usa un PVC para la transmisión de tramas de longitud variable en el nivel de enlace. Los datos viajan desde una red

por una línea digital rentada, hasta llegar al conmutador de datos de la red frame relay. Después, pasan por la red frame relay y llegan a su red de destino.

Las redes frame relay están alcanzando popularidad debido que son mucho más rápidas que otros sistemas de conmutación para realizar las operaciones básicas de conmutación por paquetes. Esto se debe a que frame relay utilizan un PVC de tal forma que la ruta completa entre los extremos es conocida. No se presenta la necesidad de dispositivos de retransmisión de tramas que realicen fragmentación o reensamblado, ni que proporcionen la mejor ruta.

Las redes frame relay también le pueden ofrecer a los suscriptores el ancho de banda que necesiten, con lo que les permiten efectuar casi cualquier tipo de transmisión.

La tecnología frame relay requiere un enrutador o puente con capacidad para frame relay, con lo cual es posible transmitir los datos sobre la red con éxito. Un enrutador que permite la retransmisión de tramas necesitará al menos un puerto WAN para la conexión con la red frame relay y otro puerto para la LAN.

1.3.3.3 Modo de Transferencia Asíncrono (ATM) [3]

El modo de transferencia asíncrono es un desarrollo avanzado de la conmutación por paquetes que proporciona transmisión de datos a altas velocidades, con el fin de enviar paquetes de un solo tamaño sobre redes LAN o WAN de banda ancha y de banda base. ATM puede aceptar voz, datos, fax, video en tiempo real, audio con calidad de CD, imágenes y transmisión de datos de múltiples Mb.

En 1988, el CCITT definió a ATM como parte de la red digital de servicios integrados de banda ancha (BISDN: Broadband Integrated Services Digital Network).

ATM es un métodos de retransmisión de celdas de banda ancha que envía los datos en celdas de 53 bytes en lugar de hacerlo longitud variable. Estas celdas consisten en 48 bytes de información de la aplicación y en cinco bytes adicionales de datos del encabezado ATM. Por ejemplo, ATM dividiría un paquete de 1000 bytes en 21 tramas de datos y colocaría cada trama de datos en una celda. El resultado es una tecnología que transmite un paquete consistente y uniforme.

Los equipos de red pueden conmutar, enrutar y mover tramas de tamaño uniforme con mayor rapidez que las tramas de tamaño aleatorio. La celda de tamaño estándar y consistente emplea los búferes de forma eficiente y reduce el trabajo requerido para procesar los datos entrantes. El tamaño un informe de la celda también ayuda al diseño del ancho de banda de la aplicación.

Teóricamente, ATM puede ofrecer velocidades e transmisión de hasta 1.2 Gbps. Sin embargo, actualmente ATM sufre limitaciones en su velocidad debido a que

las fibras ópticas sólo pueden alcanzar hasta 622 Mbps. La mayoría de las tarjetas ATM comerciales transmitirán los datos aproximadamente a 155 Mbps.

1.3.3.4 Red Digital de Servicios Integrados (ISDN) [3]

ISDN es una especificación de conectividad digital entre varias LAN que acepta: voz, datos y video.

Uno de los objetivos originales de los desarrolladores de ISDN era unir casas y oficinas utilizando cables telefónicos de cobre. El primer plan de la instalación de ISDN intenta convertir los circuitos telefónicos analógicos existentes en circuitos digitales. Este plan se está llevando a cabo a nivel mundial.

El servicio básico de ISDN divide el ancho de banda disponible en tres canales de datos. Dos de éstos mueven los datos a 64 Kbps, y el tercero a 16 Kbps.

Los canales de 64 Kbps se conocen como canales B. Pueden transportar voz, datos o video. El canal más lento, de 16 Kbps, se llama canal D. Este canal lleva señalización y la información de la administración del enlace. El servicio básicos de ISDN para los escritorios se llama 2B+D..

Un equipo conectado a un servicio ISDN puede usar los dos canales B de forma conjunta resultando en un flujo de datos a 128 Kbps. Si las dos estaciones de cada extremo también admiten compresión, se puede conseguir una mayor eficiencia de procesamiento.

1.3.3.5 Interfaz de Datos Distribuidos por Fibra (FDDI) [3]

Esta interfase se basa en una topología de doble anillo, con método de acceso por paso de testigo, pudiendo alcanzar velocidades de 100 Mbps, con una distancia entre nodos de hasta 2 Km. Las redes de fibra óptica gozan de las ventajas propias de este medio, como son la fiabilidad, la elevada velocidad y la inmunidad al ruido. Además, puede utilizarse este medio para la transmisión de señales de video y/o audio.

En la topología de doble anillo, uno de ellos (llamado primario) se utiliza para la transmisión de datos, mientras que el otro (llamado secundario) permanece como reserva para el caso de que se produzca alguna falla.

Según la norma FDDI, existen tres entidades que pueden conectarse al anillo de fibra óptica:

- ↓ Estaciones de conexión única, SAS (Single Attachment Station). Son las que se conectan únicamente al anillo primario.
- ↓ Estaciones de conexión doble, DAS (Dual Attachment Station). Son las que se conectan a ambos anillos, primario y secundario.
- ↓ Concentrador, (HUB). Son unos dispositivos de conexión que disponen de múltiples puertos de conexión.

En una red FDDI se pueden conectar un máximo de 500 estaciones, pudiendo llegar a tener la circunferencia del anillo hasta 200 km. La razón por la que existe una distancia máxima no es otra que la de reducir al mínimo el tiempo de latencia, es decir, el intervalo durante el cual los datos se encuentran viajando por el anillo.

1.4 REDES INALÁMBRICAS

En los tipos de redes que se han analizado hasta ahora, los medios físicos de transmisión han sido el cable de par trenzado o bien el cable coaxial. Un costo importante asociado a estas redes es el de instalar el cableado físico. Además, si se modifica la disposición de las computadoras interconectadas, se puede incurrir en un costo similar al de la instalación original para cambiar el plan del cableado. Esta es una de las razones por las que han aparecido redes inalámbricas, esto es, redes que no utilizan cables físicos como medio de transmisión.

Una segunda razón es la aparición de las terminales manuales y de las computadoras portátiles. Los avances tecnológicos han hecho que tales dispositivos puedan compararse cada vez más a muchas computadoras estáticas.

1.4.1 Aplicaciones de los Sistemas Inalámbricos

La dificultad en la implementación de redes de cable es un factor que hace que los sistemas inalámbricos ganen aceptación. Estos sistemas pueden resultar especialmente útiles para conectar en red a:

- ↓ Zonas de gran tránsito tales como pasillos o áreas de recepción.
- ↓ Usuarios que se encuentran constantemente en movimiento dentro del edificio de trabajo, por ejemplo médico y enfermeras en hospitales.
- ↓ Áreas y edificios aislados.
- ↓ Departamentos en los que la ubicación física de los puestos de trabajo cambia constantemente.
- ↓ Edificios de valor histórico, en los que el cableado implica grandes complicaciones.

1.4.2 Tipos de Redes Inalámbricas

Las redes inalámbricas se pueden dividir en tres categorías según su tecnología:

- ↓ Redes de área local
- ↓ Redes de área local extendida
- ↓ Computación móvil

La principal diferencia entre estas categorías radica en los medios de transmisión. Las redes de área local y las redes de área local extendida utilizan transmisores y receptores que pertenecen a la misma compañía en la que funciona la red. La computación móvil utiliza portadoras públicas tales como AT&T, MCI, Sprint y las

compañías telefónicas locales y sus servicios públicos para transmitir y recibir señales.

1.4.2.1 Redes de Área Local

Una red inalámbrica característica tiene la apariencia y funciona como una red de cable excepto por el medio de transmisión. En cada equipo se instala una tarjeta adaptadora de red equipada con un transductor y los usuarios se conectan a la red como si estuvieran cableados.

Los transductores, a menudo denominados puntos de conexión, emiten y reciben las señales de los equipos circundantes, y unen los equipos sin cables con la red cableada.

Estas redes inalámbricas de área local utilizan pequeños transductores de pared para conectarse con la red cableada. Los transductores establecen una conexión vía radio con los dispositivos portátiles de la red. Estas redes no son verdaderamente inalámbricas, ya que utilizan los transductores para conectarse con una red de área local cableada estándar.

Las redes inalámbricas utilizan cuatro técnicas de transmisión de datos:

a) Infrarrojos

Las redes inalámbricas funcionan utilizando un haz de luz infrarroja para el transporte de las señales. Este sistema necesita generar señales muy intensas en el haz infrarrojo, debido a que las señales débiles se ven afectadas por la luz de otras fuentes tales como ventanas.

Este método permite la transmisión a altas velocidades, debido a su gran ancho de banda. Una red de infrarrojos normalmente puede trabajar a velocidades de 10 Mbps.

Existen cuatro tipos de redes de infrarrojos:

Redes de Visión Directa Tal como lo indica su nombre, este tipo de red sólo funciona si el receptor y el transmisor disponen de espacio libre entre ellos para que el haz llegue de uno a otro sin ningún obstáculo.

Redes de Infrarrojos Dispersos Esta tecnología se basa en que la luz se dispersa y se refleja en las paredes y en los techos hasta llegar eventualmente al receptor. Tiene un área efectiva de unos 30 metros y una velocidad limitada debido al reflejo en los distintos elementos de la construcción.

Redes de Reflectores En este tipo de redes, los transductores ópticos de los equipos emiten hacia un lugar determinado, desde el que se retransmite hasta el punto de destino.

Redes Ópticas de Banda Ancha Esta versión para redes inalámbricas de infrarrojos de área local proporciona servicios de banda ancha. Es capaz de trabajar con grandes cantidades de información, cubriendo los requerimientos de aplicaciones multimedia, con prestaciones similares a las de redes de cables.

Aunque la velocidad y utilidad de las redes de infrarrojos despiertan cierto interés, estos sistemas tienen limitado su alcance a unos 30 metros. Además están sujetos a la interferencia de fuentes de luz intensa, habituales en las oficinas.

b) Láser

La tecnología láser es similar a la tecnología de infrarrojos en el sentido de que en ambas es necesario una línea directa de visión. Cualquier persona o cosa que se interponga en el haz láser lo interrumpirá, bloqueando la transmisión.

c) Radio en Banda Estrecha (una única frecuencia)

Se trata de un sistema similar a la radiodifusión sonora. El usuario sintoniza su receptor y su emisor a una determinada frecuencia. No requiere visión directa porque cubre superficies de 500 metros cuadrados. No obstante, debido a que se utilizan altas frecuencias, la señal no puede atravesar acero ni estructuras armadas.

Los usuarios pueden adquirir este sistema a través de proveedores de servicios tales como Motorola. El proveedor deberá cumplir todas las normas legales exigidas por la Comisión Federal de Comunicaciones (FCC: Federal Communications Commission). El sistema es relativamente lento, con velocidades en torno a 4.8 Mbps.

d) Radio en todo el Espectro de Frecuencias

Este sistema, también llamado radio en banda ancha, envía las señales dentro de un intervalo de frecuencias. Se evitan así los problemas del sistema de banda estrecha.

Las frecuencias disponibles se dividen en canales o saltos. Los adaptadores de este sistema sintonizan un determinado canal durante un periodo preestablecido, y a continuación cambian a otro canal. Una secuencia de saltos determina la temporización. Todos los equipos de la red están sincronizados con esta temporización de canales o saltos. Este tipo de envío de señales proporciona cierta seguridad "integrada" ya que el algoritmo de salto de frecuencia de la red tendría que ser "reconocido" para intervenir el flujo de datos.

Para mejorar la seguridad y evitar las escuchas no autorizadas, tanto el equipo emisor como el receptor codifican la transmisión.

La velocidad típica de 250 Kbps convierte a este método en el más lento de todos. No obstante, algunas implementaciones de radio en todo el espectro de frecuencias permiten velocidades de 4 Mbps en distancias de más de 3 kilómetros en exteriores y 150 metros en el interior.

Ésta es un área en la que la tecnología de hoy en día ofrece verdaderas redes inalámbricas. Por ejemplo, dos o más equipos con adaptadores Xircom CreditCard Netwave, y un sistema operativo del tipo de Microsoft Windows 98 o Microsoft Windows NT, pueden trabajar en una red de área local punto a punto sin necesidad de cables. Si lo que tiene es una red basada en Windows NT Server, puede conectar la red anterior con ésta agregando un punto de acceso inalámbrico (Netwave Access Point) en uno de los equipos de la red basada en Windows NT Server.

e) Transmisión Punto a Punto

Este sistema de comunicaciones no pertenece claramente a las definiciones que estamos realizando sobre redes. Utiliza la tecnología punto a punto para transferir datos de un equipo a otro, a diferencia de los sistemas que se están analizando, los cuales conectan varios equipos y periféricos entre sí. Sin embargo, se cuenta con elementos adicionales tales como transconductores simples y de anfitrión, los cuales se pueden utilizar para conectar equipos aislados, o que pertenecen a redes, para crear una red de transmisión de datos inalámbrica.

Esta tecnología incluye transferencia inalámbrica de datos en serie que cumple Con lo siguiente:

- ↓ Utiliza un enlace punto a punto por radio para transmitir de manera rápida y sin errores.
- ↓ Atraviesa paredes, techos y suelos.
- ↓ Permite velocidades desde 1.2 hasta 38.4 Kbps, en distancias de hasta 70 metros en interiores y 500 metros en el exterior o hasta 536 metros en redes de visión directa.
- ↓ Este tipo de sistema se utiliza para transferir datos entre equipos, o entre éstos y otros dispositivos como impresoras o lectores de código de barras.

1.4.2.2 Redes de Área Local Extendida

Otros tipos de elementos inalámbricos permiten trabajar en entornos de redes de área local extendida de manera similar a sus equivalentes cableados. Un puente inalámbrico de red de área local, por ejemplo, puede conectar dos redes separadas por hasta 4.5 kilómetros.

a) Conectividad Inalámbrica Multipunto

El elemento llamado puente inalámbrico proporciona un método sencillo para conectar edificios sin necesidad de emplear cables. Del mismo modo, un puente arquitectónico proporciona un camino más fácil entre dos puntos con alguna separación física, el puente inalámbrico tiende una conexión de datos entre dos edificios. Por ejemplo, el AIRLAN/Bridge Plus utiliza tecnología de radio en todo el espectro de frecuencias para crear una estructura inalámbrica y conectar oficinas más allá del alcance de las redes de área local. Según las condiciones, esta distancia puede alcanzar los 5 kilómetros.

El costo de este elemento se justifica porque elimina la necesidad de utilizar líneas alquiladas.

b) Puente Inalámbrico de Larga Distancia

Si el puente inalámbrico no tiene el alcance suficiente, la organización deberá plantearse la utilización de un puente inalámbrico de larga distancia. Este también utiliza todo el espectro de frecuencia para permitir puentes Ethernet y Token Ring consiguiendo distancias de hasta 38 kilómetros.

Al igual que en puente inalámbrico original, el costo de un puente de larga distancia se justifica porque elimina la necesidad de utilizar una línea T1 o conexiones por microondas. T1 es un estándar que identifica una línea de servicio digital que permite transmisiones a 1,544 Mbps. Puede transportar tanto voz como datos.

1.4.2.3 Computación Móvil

Las redes inalámbricas móviles implican el uso de señales telefónicas y de servicios públicos para transmitir y recibir señales utilizando:

- a) Comunicación a través de paquetes por radio
- b) Redes celulares
- c) Estaciones satelitales

Los empleados cuando viajan pueden utilizar esta tecnología con equipos portátiles o PDA (asistente personal digital) para el intercambio de correo electrónico, archivos o cualquier otro tipo de información.

Aunque el sistema ofrece ventajas, resulta bastante lento. Las velocidades de conexión van desde los 8 hasta los 19.2 Kbps. Estas velocidades son incluso inferiores cuando se incluye corrección de errores.

La computación móvil incluye adaptadores inalámbricos para utilizar con los sistemas de telefonía celular y así conectar los equipos portátiles a la red cableada. Los equipos personales utilizan pequeñas antenas para comunicarse con las torres de radio del área circundante. Los satélites de órbitas bajas pueden recibir las señales de baja potencia de los dispositivos de red móviles y portátiles.

a) Comunicación a través de Paquetes por Radio

Este método divide los datos de la transmisión en paquetes, similares a otros paquetes de red que incluyen:

- ↓ La dirección de origen
- ↓ La dirección de destino
- ↓ Información para la corrección de errores

Los paquetes se envían a un satélite que los retransmite. Únicamente los dispositivos con la dirección de destino indicada pueden recibir los paquetes transmitidos.

b) Redes Celulares

Los paquetes de datos celulares (CDPD) utilizan la misma tecnología y en ocasiones los mismos equipos que la telefonía celular. Ofrecen transmisión de datos a través de redes analógicas de voz intercaladas con llamadas cuando el sistema no está ocupado. Es una tecnología bastante rápida que sufre retrasos de fracciones de segundo lo cual la vuelve adecuada para transmisiones en tiempo real.

Al igual que otras redes inalámbricas, tiene que haber una manera de conectarse a la red cableada existente. Nortel de Mississauga en Ontario, Canadá, es una compañía que fabrica una unidad de interfaz Ethernet (EIU) que proporciona este tipo de conexión.

c) Estaciones Satélite

Los sistemas de microondas son ideales para interconectar edificios en áreas de tamaño limitado, como puede ocurrir en campus universitarios o parques industriales.

En la actualidad las microondas son el sistema más extendido en los Estados Unidos para transmisiones en largas distancias. Resultan muy adecuadas para comunicar dos puntos con visión directa como:

- ↓ Vínculos satélite- Tierra-
- ↓ Entre dos edificios.
- ↓ Atravesar grandes espacios abiertos y llanos, tales como mares, lagos o desiertos.

Un sistema de microondas consta de:

- ↓ Dos transductores de radio: uno que emite (es la estación emisora) y otro que recibe (estación receptora)
- ↓ Dos antenas direccionales apuntando una a la otra para conseguir la comunicación de las señales entre los transductores. Estas antenas se suelen instalar en torres para proporcionarles mayor margen sobre cualquier tipo de obstáculo que pudiera interferir en la comunicación.

1.4.3 Métodos de Control de Acceso al Medio

Tanto el radio como el infrarrojo operan en un medio de difusión, es decir, todas las transmisiones son recibidas por todos los receptores que están dentro del campo de cobertura del transmisor. En consecuencia, así como necesitamos recurrir aun método de MAC con las LAN por cable de medio compartido, para

asegurarnos de que sea un solo transmisor el que esté usando el medio, también se necesita un método de MAC con las redes inalámbricas. Los esquemas que más se usan para ello son:

1.4.3.1 CDMA

El acceso múltiple por división de código (CDMA: Code Division Multiple Access) se aplica específicamente a los sistemas de radio de espectro disperso. Para sus modos de operación los esquemas tanto por secuencia directa como por salto de frecuencia se basan en una secuencia pseudoaleatoria única de dispersión/salto.

En virtud de ello, "es posible asignar una secuencia pseudoaleatoria distinta a cada nodo, y todos los nodos pueden conocer el conjunto completo de secuencias. Para comunicarse con otro nodo, el transmisor sólo tiene que seleccionar y usar la secuencia pseudoaleatoria del destinatario considerado." [3] Esto hace posible la realización concurrente de múltiples comunicaciones entre diferentes pares de nodos.

En la práctica, esto sólo es factible con los sistemas por salto de frecuencia, pues en los que son por secuencia directa puede presentarse un fenómeno denominado efector cercano-lejano. Esto sucede cuando está operando un segundo transmisor. Por ejemplo el nodo X de la figura 6, que está más cerca del destinatario, el nodo A, que el otro participante en la comunicación por el proceso de antidispersión en el nodo A, la mayor cercanía de la señal de interferencia (dispersa) puede hacer que ésta tenga mayor potencia que la señal requerida del nodo B y haga que el receptor del nodo A no acepte la transmisión. Esto se conoce también como efecto de terminal oculta.

En contraste, con el salto de frecuencia, como los dos transmisores están cambiando constantemente de canales, es muy baja la probabilidad de que ambos operen en el mismo canal al mismo tiempo, que puede reducirse aún más si se planean con mucho cuidado las secuencias de salto. La desventaja de ambos esquemas, radica en que todos los nodos deben conocer la secuencia pseudoaleatoria de todos los demás nodos, cosa que no es fácil gestionar en una red inalámbrica.

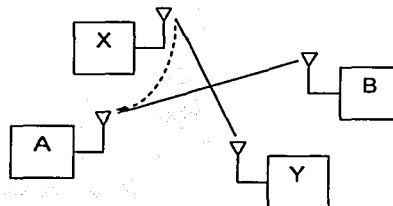


Figura 1.9. CDMA

1.4.3.2 Acceso Múltiple con Detección de Portadora y Detección de Colisiones (CSMA/CD)

“En este esquema, cuando un nodo tiene una trama para transmitir, lo primero que hace es generar una secuencia binaria pseudoaleatoria corta, denominada peine y la añade al preámbulo de la trama. A continuación, el nodo realiza la operación normal de detección de portadora y suponiendo que el medio está inactivo, procede a transmitir la secuencia de peine. Para cada 1 binario de la secuencia, el nodo transmite una señal durante un intervalo de tiempo corto, pero cuando la secuencia incluye un 0 binario, el nodo cambia al modo de recepción.” [3] Si un nodo detecta la transmisión de una señal mientras está en el modo de recepción, se elimina a sí mismo de la competencia por el canal y espera hasta que el (los) otro (s) nodo (s) haya (n) transmitido su trama. El principio de funcionamiento del esquema se ilustra en la figura 1.10.

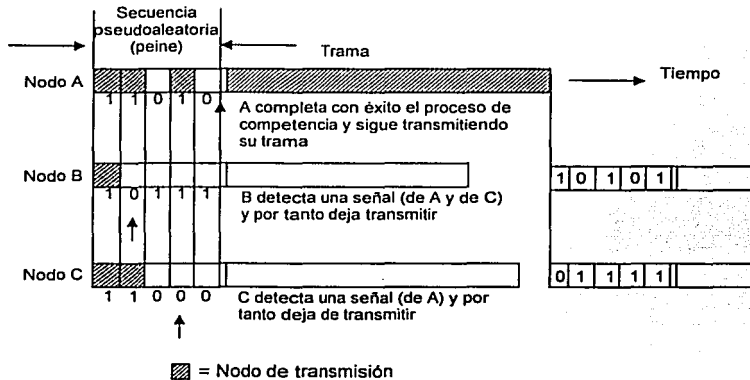


Figura 1.10. CSMA/CD

En este ejemplo, tres nodos A, B y C están compitiendo por el canal y los códigos pseudoaleatorios generados por cada uno son los que se muestran. Puesto que el primer bit de la secuencia es un 1 binario para todos los nodos, ningún nodo está escuchando y por tanto la transmisión no es detectada. Durante el segundo intervalo del peine los nodos A y C siguen transmitiendo, pero el nodo B ahora está en el modo de recepción, así que detecta una señal y deja de competir por el canal en este punto. Durante el tercer intervalo, como el nodo B está inactivo y tanto el nodo A como el C están en el modo de recepción, ninguno de estos dos detectará una señal. En el cuarto intervalo el nodo A está transmitiendo y el nodo C está en el modo de recepción, así que este último detectará una señal y dejará

de competir. Ahora sólo queda el nodo A que, después de completar con éxito el resto del proceso de competencia, procede a transmitir la trama que tiene en espera.

La eficiencia del esquema depende del número de bits de la secuencia pseudoaleatoria, ya que si dos nodos generan la misma secuencia puede ocurrir una colisión. En la práctica es probable que el número de nodos en competencia, en un momento dado, sea bajo, así que la longitud del peine puede ser relativamente corta. Además, como hay un límite máximo para la velocidad con que un transceptor de radio (o de infrarrojo) puede conmutar entre los modos de transmisión y de recepción, por lo regular un microsegundo, con un peine más corto se reduce la duración del periodo de resolución de competencia (detección de colisiones).

1.4.3.3 Acceso Múltiple con Detección de Portadora y Prevención de Colisiones (CSMA/CA)

También existe otra adaptación de CSMA/CD, que se llama CSMA con prevención de colisiones (CSMA/CA: CSMA with collision avoidance). Su principio de funcionamiento se ilustra en la figura 1.11.

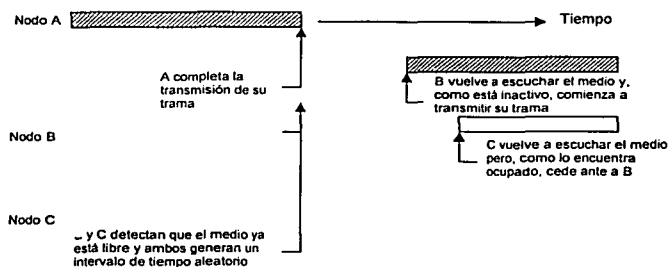


Figura 1.11. CSMA/CA

"Como puede verse, en lugar de iniciar la transmisión de una trama inmediatamente después de que el medio queda inactivo, el nodo espera un lapso aleatorio adicional corto, y sólo si el medio sigue inactivo después de este intervalo comienza a transmitir." [3] De esta manera, si hay otros nodos en espera, el nodo que calcula el tiempo más corto obtendrá acceso primero y los nodos restantes cederán ante él. Una vez más, la eficiencia del esquema es una función del número de incrementos de tiempo y por tanto de bits en la secuencia pseudoaleatoria, que tiene el periodo de evitación de colisiones máximo.

Surge otro problema por resolver cuando se utiliza radio (e infrarrojo) porque no hay ninguna garantía de que el interlocutor (nodo) considerado esté en contacto de radio con el nodo de origen. Así, aunque el algoritmo de CSMA/CA o de CSMA/CD asegure que un nodo obtendrá acceso al medio, el o los destinatarios de la trama tal vez nunca lo reciban porque están fuera del alcance de radio. Es por esta razón que en el protocolo de MAC se incorpora un procedimiento de saludo adicional, anterior al método MAC básico. Como la intención es que este procedimiento sirva a los diferentes tipos de métodos de MAC, se le ha denominado protocolo de MAC inalámbrico de fundamento distribuido (DFW MAC: distributed foundation wireless MAC).

Siempre que una unidad portátil necesita enviar una trama, lo primero que hace es enviar un mensaje/trama de control corto de solicitud de transmisión (RTS: request to send), ya sea a su PAU o a otra unidad portátil, mediante uno de los métodos de MAC que descrito anteriormente: CSMA/CD o CSMA/CD. El mensaje de control RTS contiene la dirección de MAC de las unidades tanto de origen como de destino. Si el destinatario recibe este mensaje y está en condiciones de recibir una trama, difundirá un mensaje/trama de contestación preparado para transmitir (CTS: clear to send) con el mismo par de direcciones aunque en orden inverso. Como alternativa, si el receptor no está preparado para recibir una trama, devolverá una contestación de receptor ocupado (RRxBUSY). Si la respuesta es positiva, la unidad solicitante transmitirá la trama en espera (DATA) que, si es recibida correctamente por el destinatario, se contestará con un mensaje de confirmación (ACK: acknowledgement) positivo. Por otro lado, si la trama está alterada, el destinatario devolverá un mensaje de confirmación negativa (NAK) y el origen tratará de enviarlo otra vez. Este procedimiento va a repetirse hasta que se cumpla un número previamente definido de reintentos.

1.4.3.4 Acceso Múltiple por División de Tiempo (TDMA)

En la figura 1.12 se muestra el principio de funcionamiento del acceso múltiple por división de tiempo (TDMA: time division multiple access). "Con este método, cada (nodo) transmisor tiene un intervalo/ranura de tiempo específico y, una vez alcanzada la ranura de tiempo de un transmisor, éste transmite por el ancho de banda completo durante el periodo (fijo) de la ranura." [3] Casi siempre la duración de cada ranura de tiempo es corta y se escoge de modo que la probabilidad de ocurrencia de errores en ese lapso sea baja. El periodo de trama/ciclo está determinado por la duración de cada ranura y del número de ranuras de transmisión manejadas.

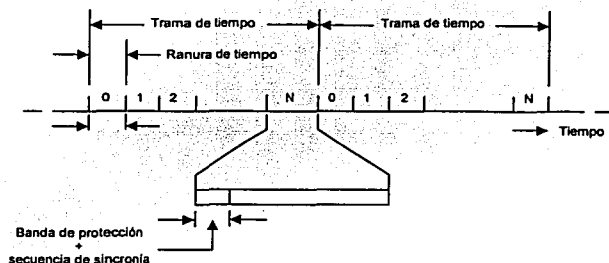


Figura 1.12. TDMA

Por lo regular se emplea TDMA cuando hay una sola estación (base) a través de la cual se realizan todas las transmisiones. A cada terminal portátil en el campo de cobertura de la estación base se le asigna una ranura de tiempo específica, o, lo que es más común, se proporciona una ranura de tiempo independiente (de señalización) para que cada uno de los dispositivos portátiles pueda solicitar a la estación base una ranura de tiempo libre en caso de tener una trama por transmitir. Las transmisiones desde la estación base a las portátiles se realizan ya sea en un modo difundido mediante una ranura de tiempo específica, con la dirección del destinatario en la cabecera de la trama transmitida, o en una ranura de tiempo específica establecida por medio del canal de señalización. Como alternativa, para controlar cada ranura de tiempo puede emplearse una subranura de señalización independiente incorporada.

Como se aprecia en la figura 1.12, hay una banda de protección y una secuencia de sincronía al principio de cada ranura de tiempo. La banda de protección compensa los diversos retardos de propagación entre el conjunto distribuido de portátiles y la estación base, en tanto que el intervalo de sincronización permite al receptor ya sea portátil o estación base, sintonizarse con el transmisor antes de recibir el contenido de la ranura.

1.4.3.5 Acceso Múltiple por División de Frecuencia (FDMA)

El principio de funcionamiento del acceso múltiple por división de frecuencias (FDMA: frequency division multiple access) se ilustra en la figura 1.13. El método FDMA principalmente se aplica a los sistemas de radio y al igual que TDMA, requiere una estación base para controlar su funcionamiento. "Con el FDMA, el ancho de banda de frecuencias total asignado se divide en varias sub-bandas de frecuencias o canales, que, en principio, son similares al espectro disperso por salto de frecuencia. Sin embargo, en el caso del FDMA, una vez que se ha

asignado un canal de frecuencia específico, se le utiliza durante todo el periodo de transmisión de una trama." [3] Lo normal es que los canales de frecuencia se asignen por demanda mediante un canal de señalización aparte.

En general, la estación base de un sistema FDMA es más compleja que la de los sistemas TDMA, razón por la cual estos últimos son más comunes. También se emplean esquemas híbridos que utilizan FDMA para obtener varios canales de frecuencia, que entonces se manejan con TDMA.

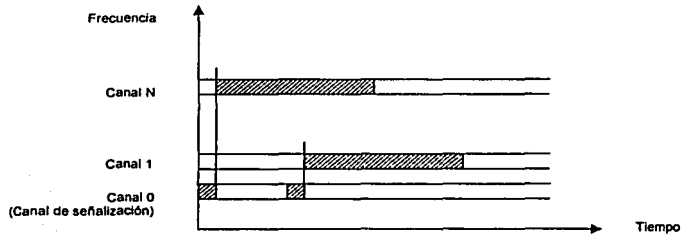


Figura 1.13. FDMA

CAPÍTULO 2 SISTEMA DE COMUNICACIONES

2.1 INTRODUCCIÓN

La información que transmite un mensaje no está relacionada con su longitud. Podemos tener dos mensajes con distinta longitud y que transmitan la misma información. El concepto de información está muy relacionado con el concepto de probabilidad, cuanto más probable es un mensaje menos información contiene.

En nuestro caso contamos con los símbolos de un alfabeto fuente que son transmitidos por el emisor. Cada uno de estos símbolos tiene asociada una probabilidad. El contenido en información de cada uno de los símbolos se define como:

$$I = \log_2 \frac{1}{p_i} \quad [\text{Bits de información}] \quad (1)$$

Un concepto muy ligado al de cantidad de información es el concepto de entropía.

La entropía es una forma de evaluar la calidad del dispositivo codificador. Se define como el valor medio de la información por símbolo:

$$H = - \sum_i p_i \log_2 p_i \quad [\text{Bits/mensaje}] \quad (2)$$

2.1.1 Codificación [7]

La codificación consiste en establecer una correspondencia entre cada uno de los símbolos de un alfabeto fuente y una secuencia de símbolos de un alfabeto destino. Al alfabeto destino se le denomina alfabeto código y a cada una de las secuencias de símbolos de este alfabeto que se corresponda con un símbolo del alfabeto fuente se denomina palabra de código, como se muestra en la figura 2.1.

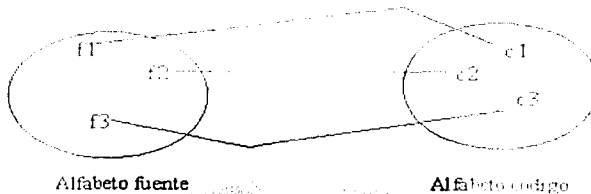


Figura 2.1 Codificación

CAPÍTULO 2 SISTEMA DE COMUNICACIONES

2.1 INTRODUCCIÓN

La información que transmite un mensaje no está relacionada con su longitud. Podemos tener dos mensajes con distinta longitud y que transmitan la misma información. El concepto de información está muy relacionado con el concepto de probabilidad, cuanto más probable es un mensaje menos información contiene.

En nuestro caso contamos con los símbolos de un alfabeto fuente que son transmitidos por el emisor. Cada uno de estos símbolos tiene asociada una probabilidad. El contenido en información de cada uno de los símbolos se define como:

$$I_i = \log_2 \frac{1}{p_i} \quad \text{[Bits de información]} \quad (1)$$

Un concepto muy ligado al de cantidad de información es el concepto de entropía.

La entropía es una forma de evaluar la calidad del dispositivo codificador. Se define como el valor medio de la información por símbolo:

$$H = -\sum_i p_i \log_2 p_i \quad \text{[Bits/mensaje]} \quad (2)$$

2.1.1 Codificación [7]

La codificación consiste en establecer una correspondencia entre cada uno de los símbolos de un alfabeto fuente y una secuencia de símbolos de un alfabeto destino. Al alfabeto destino se le denomina alfabeto código y a cada una de las secuencias de símbolos de este alfabeto que se corresponda con un símbolo del alfabeto fuente se denomina palabra de código, como se muestra en la figura 2.1.

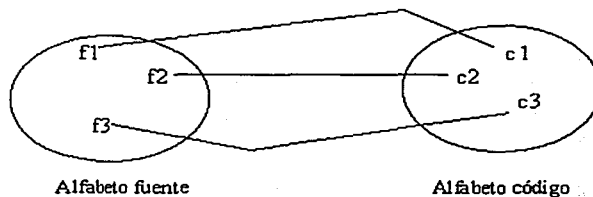


Figura 2.1. Codificación

El alfabeto fuente contiene los símbolos originales que se quieren codificar. El alfabeto código contiene las palabras de código equivalentes en que se codificarán los símbolos originales. Estas palabras de código son aptas para ser transmitidas por un sistema de comunicaciones.

2.1.2 Clasificación de los códigos

Código bloque: es aquel código en el que todas las palabras de código correspondientes a cada símbolo del alfabeto fuente tienen la misma longitud. Dentro de estos códigos podemos distinguir:

- ↓ Código singular: a cada símbolo del alfabeto fuente le corresponde una única palabra de código.
- ↓ Código no singular: a cada símbolo del alfabeto fuente le corresponde dos o más palabras de código.

Código compacto o de longitud variable: se busca que a cada símbolo del alfabeto fuente le corresponda una palabra de código de longitud mínima según algún criterio de minimización dado.

2.1.3 Propiedades de los códigos

Longitud media: Cada palabra de código asignada a cada símbolo del alfabeto fuente tiene una longitud l_k . A partir de aquí se define la longitud media de un código como:

$$\bar{L} = \sum_{x=1}^n p_x l_x \quad (3)$$

La longitud media representa el número medio de bits por símbolo del alfabeto fuente que se utilizan en el proceso de codificación

Eficiencia: A partir del concepto de longitud media la eficiencia de un código se define como:

$$\eta = \frac{L_{\min}}{\bar{L}} \quad (4)$$

Siendo:

$$\bar{L} \geq L_{\min} \quad (5)$$

Para calcular L_{\min} es necesario tener en cuenta el primer teorema de Shannon o teorema de la codificación de la fuente: Dada una fuente discreta de entropía H , la longitud media de la palabra de código está acotada inferiormente por H . Teniendo esto en cuenta L_{\min} se fija como el valor de la entropía con lo que la eficiencia puede escribirse como:

$$\eta = \frac{H}{L} \quad (6)$$

Redundancia: Se denomina redundancia de un código a la información superflua o innecesaria para interpretar el significado de los datos originales. Se define como:

$$\sigma = 1 - \eta \quad (7)$$

2.2 ESTRUCTURA GENÉRICA DE UN SISTEMA DE COMUNICACIÓN [8]

El propósito de un sistema de comunicaciones es transmitir información desde un emisor hasta un receptor a través de un canal. En la figura 2.2 se puede observar un esquema genérico de un sistema de comunicaciones.

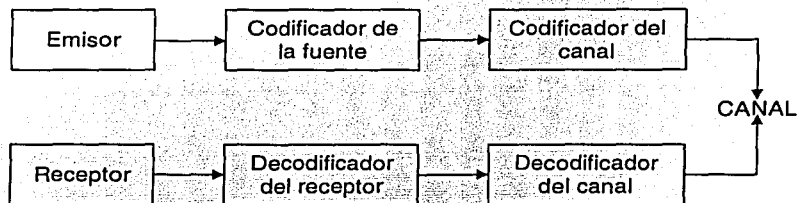


Figura 2.2. Esquema genérico de un sistema de comunicaciones

El emisor es una fuente discreta de información desde la que se emiten los distintos símbolos del alfabeto fuente que se quieren transmitir.

Los símbolos emitidos por la fuente llegan al codificador de la fuente donde son transformados en símbolos de un código binario más adecuado para ser transmitido a través de un canal de comunicaciones.

Opcionalmente estos símbolos codificados pueden ser comprimidos con el objetivo de reducir su tamaño para conseguir una transmisión más rápida.

Durante la transmisión de los símbolos a través del canal pueden producirse alteraciones de los mismos debidas a la presencia de ruido en el canal. A estas alteraciones se las denomina errores. Por ello, antes de enviar los símbolos codificados a través del canal, se realiza una nueva codificación orientada a que el receptor pueda detectar y corregir los errores producidos en el canal.

En la recepción se realiza un proceso inverso. En primer lugar se realiza una decodificación del canal para detectar y corregir los posibles errores que contengan los símbolos recibidos a través del canal.

A continuación se procede a una posible descompresión de los símbolos en el caso de haber sido comprimidos en la fuente.

Por último se realiza una decodificación en la que los símbolos codificados se transforman en los símbolos originales que fueron transmitidos por el emisor.

2.3 CODIFICACIÓN FUENTE [9]

Introducción

La codificación fuente está orientada a representar eficientemente una señal en formato digital, es decir, minimizar el número de bits necesario para su transmisión de forma eficiente. A su vez la codificación fuente puede ser sin pérdidas (por entropía) o con pérdidas; esto significa que se pueda reconstruir la señal original de forma exacta en recepción, como es el primer caso, o que se produzcan ciertas pérdidas.

2.3.1 Codificación fuente para una fuente sin memoria discreta

Interpretando el trabajo de Shannon, se puede observar que la codificación fuente, es el proceso por el cual la salida de una fuente q es convertida a secuencia binaria para la transmisión vía canales binarios. Cuando una fuente sin memoria discreta genera q símbolos con la misma probabilidad con un porcentaje de tasa de información de $R = R_s \log_2 q$, todos los símbolos transportan la misma cantidad de información y una eficiente señalización toma la forma de transmisión binaria a una tasa de R bps. Cuando las probabilidades de los símbolos no son iguales, la mínima tasa de fuente requerida para que se disminuya la distorsión, se reduce a la siguiente ecuación.

$$R = R_s H < R_s \log_2 q \quad (8)$$

Entonces la transmisión para símbolos altamente probables lleva poca información y por lo tanto asignando $\log_2 q$ números de bits hace que no se use el canal eficientemente. Por lo que el teorema de codificación fuente de Shannon sugiere que por medio de utilizar un codificador fuente antes de la transmisión, la eficiencia del sistema con fuentes de símbolos con la misma probabilidad puede ser arbitrariamente acercado.

La eficiencia de codificación puede ser definida como el radio la tasa de información fuente y el porcentaje de la tasa de salida del codificador fuente.

2.3.2 Codificación Shannon-Fano

El algoritmo de codificación Shannon-Fano es basado en el concepto de codificación de mensajes frecuentes utilizando códigos de palabras cortos y los poco frecuentes por códigos de palabras largos, reduciendo la longitud del mensaje. Este algoritmo es parte virtual de repartir todos los tratados con teoría de la información, como por ejemplo, el trabajo de Carlson. En base al algoritmo 1 obtenemos la longitud promedio del código de palabra es dado por medio de medir la longitud de cualquier código de palabra por su probabilidad, basándonos en la tabla 2, obtenemos el resultando lo siguiente:

$$(0.27+0.2)*2+(0.17+0.16)*3+2*0.06*.4+2*0.04*.4 \approx 2.73[\text{bit}]$$

La entropía de la fuente es:

$$\begin{aligned} H &= -\sum_i p_i \log_2 p_i \\ &= -(\log_2 10) \sum_i p_i \log_{10} p_i \\ &= -3.322[0.27 \log_{10} 0.27 + 0.2 \log_{10} 0.2 + 0.17 \log_{10} 0.17 \\ &= + 0.16 \log_{10} 0.17 + 0.16 \log_{10} 0.16 + 2 * 0.06 \log_{10} 0.06 \\ &\quad + 2 * 0.04 \log_{10} 0.04] \\ &\approx 2.691 \text{ [bits/símbolo]} \end{aligned}$$

Algoritmo 1

1. Los símbolos fuente $S_0 \dots S_7$, primero son ordenados en forma descendente de acuerdo a la probabilidad de ocurrencia.
2. Posteriormente los símbolos son divididos en dos subgrupos de tal forma que las probabilidades de los subgrupos sean tan cercanas a cada otra como sea posible. Esto es simbolizado por la división horizontal en la tabla 2.
3. Cuando localizamos el código de palabras para representar los símbolos fuente, asignamos ceros lógicos al grupo superior y unos lógicos al grupo inferior en la columna apropiada denominada pasos de codificación.
4. Si existe mas de un símbolo en el subgrupo, este método se continúa hasta que ya no haya divisiones.
5. Finalmente, la variable longitud del código de palabra es la que se manda al canal.

Símbolo	Probabilidad	Pasos de Codificación				Código de palabra
		1	2	3	4	
S_0	0.27	0	0			00
S_1	0.20	0	1			01
S_2	0.17	1	0	0		100
S_3	0.16	1	0	1		101
S_4	0.06	1	1	0	0	1100
S_5	0.06	1	1	0	1	1101
S_6	0.04	1	1	1	0	1110
S_7	0.04	1	1	1	1	1111

Tabla 2.1. Ejemplo de codificación de Shannon-Fano basado en la figura 1.6

A partir de que el promedio de la longitud de palabra (2.73 bits/símbolo) es muy cercana a la entropía (2.691 bits/símbolo), se predice una alta eficiencia de codificación, la cual se puede calcular por medio de :

$$E = \frac{2.691}{2.73} \approx 98.6\%$$

La asignación directa BCD (binary coded decimal) es de 3 bits/símbolo y nos da una eficiencia de:

$$E \approx \frac{2.691}{3} \approx 89.69\%$$

En conclusión, se puede decir, que la codificación de Shannon-Fano permite crear un conjunto de mapeos únicos e invertibles a un conjunto de código de palabras, lo cual facilita una transmisión más eficiente de la fuente de símbolos que una representación BCD directa.

2.3.3 Codificación de Huffman

El algoritmo de codificación de Huffman (HC) es describe en el algoritmo 2, mientras que en la tabla 3 podemos observar un ejemplo práctico del mismo. Se utilizan las mismas probabilidades de símbolos que en el ejemplo de Shannon-Fano, pero el algoritmo permite asignación de diferentes códigos de palabra. No obstante, la eficiencia de código es idéntica a la del algoritmo de Shannon-Fano.

Algoritmo 2

1. Ordenar las probabilidades de símbolos p_i en orden descendente y considerarlos como nodos, así como se muestra en la tabla 3.
2. Mientras existan más de dos nodos, fusionar los dos nodos que tengan la probabilidad mas baja y asignar 0/1 a la parte superior e inferior respectivamente.
3. Leer los bits de transición asignados en los subconjuntos desde arriba hasta abajo en orden de obtener los códigos de palabra.

Símbolo	Probabilidad	Paso 1 y 2		Paso 3 y 4		Grupo	Código
		Código	Prob.	Código	Prob.		
S_0	0.27					S_0	-
S_1	0.20					S_1	-
S_2	0.17			0	0.33	S_{23}	0
S_3	0.16			1			1
S_4	0.06	0	0.12	0	0.20	S_{4567}	00
S_5	0.06	1		0			01
S_6	0.04	0	0.08	1			10
S_7	0.04	1		1	11		

Símbolo	Probabilidad	Paso 5 y 6		Paso 7		Código de palabra
		Código	Prob.	Código	Prob.	
S_{23}	0.33	0	0.6	0	1.0	00
S_0	0.27	1		01		
S_1	0.20	0	0.4	1		10
S_{4567}	0.20	1		11		

Tabla 2.2. Ejemplo de codificación de Huffman basado en el algoritmo 2.

2.3.4 Codificación por Transformada Coseno Discreta [9]

La Transformada Coseno Discreta (DCT) para compresión de imágenes es una técnica de codificación con pérdida. Los bloques de píxeles son de tamaño 8x8 píxeles, éstos son procesados con el objeto de obtener su representación en el dominio de la frecuencia, y se logra por medio de la Transformada Coseno Discreta. Donde los componentes de frecuencia con valor mínimo son descartados, quedando únicamente los componentes de mayor energía.

Aplicaciones

Las técnicas de compresión con pérdidas de imagen son usadas para reducir la cantidad de información necesaria para representar una imagen, sacrificando calidad.

Implementación

1. Se divide la imagen en bloques de tamaño de 8X8 pixeles.
2. Se aplica la Transformada Coseno Discreta a cada bloque
3. Se descartan los coeficientes de baja energía, como se muestra en la figura 2.3

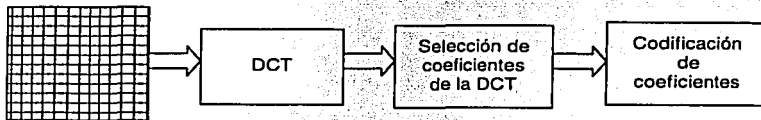


Figura 2.3. Esquema general de compresión por DCT

La DCT para compresión de imágenes causa degradación en la intensidad de los bloques, que se incrementa proporcionalmente con la tasa de compresión. Las técnicas de compresión de imágenes por la DCT típicamente proporcionan tasas de compresión aproximadas de 10:1 sin degradación importante.

La Transformada Coseno Discreta se define como:

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos[(2x+1)ux/2N] \quad (9)$$

Para $u=0,1,2,\dots,N-1$. De igual manera, la DCT Inversa está definida como:

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos[(2x+1)ux/2N] \quad (10)$$

Para $x=0,1,2,3,\dots,N-1$. En ambas ecuaciones α está definida como:

$$\alpha(u) = \begin{cases} \sqrt{1/N} & \text{Para } u=0 \\ \sqrt{2/N} & \text{Para } u=1,2,\dots,N-1 \end{cases}$$

La Transformada Coseno Discreta correspondiente en dos dimensiones es:

$$C(u) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos[(2x+1)u\pi] \cos[(2y+1)v\pi] / 2N \quad (11)$$

Para $u, v=0, 1, 2, \dots, N-1$. De igual manera, la DCT Inversa está definida como:

$$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v) C(u,v) \cos[(2x+1)u\pi] \cos[(2y+1)v\pi] \quad (12)$$

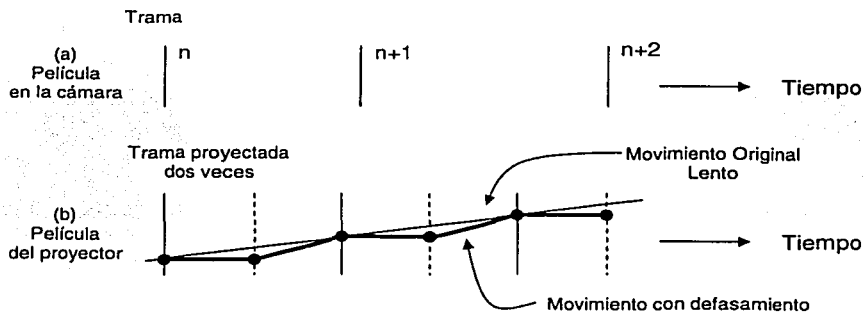
2.3.5 Codificación por compensación de movimiento [9]

La forma en la cual el movimiento del ojo evita el desalineamiento es fundamental para la calidad de percepción de las imágenes de televisión. Muchos procesos necesitan manipular imágenes en movimiento de esta misma manera para evitar las obvias dificultades de procesamiento con respecto a una trama fija de referencia.

Procesos de éste tipo son referidos como compensación de movimiento, contando con un proceso por separado el cual tiene medido el movimiento. Estimación de movimiento es literalmente un proceso, el cual analiza campos sucesivos y determina como los objetos se mueven de un campo al siguiente. Es una importante y capaz tecnología porque la función es comparable a la acción del ojo humano.

A fin de ver porque la compensación de movimiento es esencial para ciertos procesos, es necesario considerar la manera en la cual varios sistemas de formación de imagen manejan el movimiento. Esto puede hacerse mostrando la posición de imágenes sucesivas sobre el eje del tiempo. De hecho hay una considerable variedad de formas en la cual el eje del tiempo es manejado, y con esto guiar algunas potenciales incompatibilidades.

La figura 2.4 (a) muestra el más simple eje del tiempo, el de la cinta donde tramas enteras son simultáneamente expuestas o muestreadas a 24 Hz típicamente



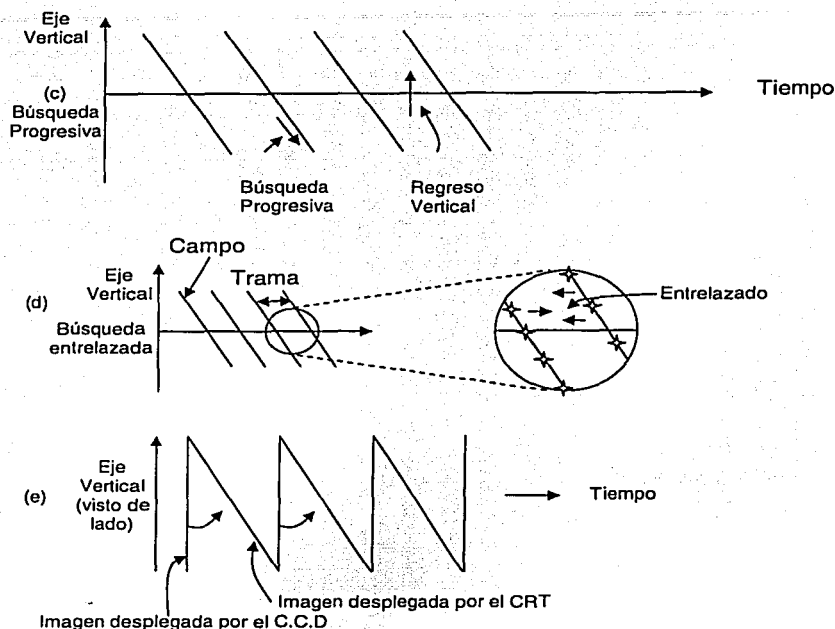


Figura 2.4. Características de espacio-tiempo de una película

El resultado es que la imagen está efectivamente en ángulos rectos al eje del tiempo. Durante la filmación, alguno de los períodos de trama es requerido para el corrimiento de la cinta, y el obturador se cierra mientras esto ocurre. Por lo tanto, la apertura temporal o exposición es un poco más corta que el período de trama y cuando son proyectadas, produciendo un parpadeo con una frecuencia de 48Hz. El resultado con un objeto en movimiento es que el movimiento no es apropiadamente representado y se observa una sobreexposición. La figura 2.4 (b) muestra el origen de la sobreexposición. La figura 2.4 (c) muestra el eje del tiempo en búsqueda progresiva de cámaras de televisión (por ejemplo, tubos de cámaras) y pantallas de rayos catódicos (CRT). La búsqueda vertical toma una parte substancial del período de trama y así la imagen se inclina con respecto al eje del tiempo. Como la cámara y la pantalla tienen la misma inclinación, esta inclinación no tiene efecto en la representación del movimiento.

En las cámaras CCD, la imagen es muestreada al mismo tiempo que la cinta, y así los campos están en ángulos rectos al eje del tiempo. Si es proyectada en un

CRT, el eje del tiempo es deformado; como que las cosas son proyectadas más tarde que cuando fueron enviadas a la parte baja de la pantalla. Los objetos moviéndose transversalmente tienen velocidad vertical dependiente, como se muestra en la figura 2.4 (e).

2.4 Codificación de canal [10]

Introducción

La codificación del canal consiste en 'mapear' (añadir redundancia) la secuencia de datos entrante en una secuencia de entrada al canal y realizar el 'mapeo' inverso a la salida del canal en una secuencia de datos tal que los efectos del ruido estén minimizados.

La introducción de redundancia en la codificación del canal tiene como finalidad mejorar la fiabilidad de la transmisión.

2.4.1 Tipos de codificación

2.4.1.1 Códigos de paridad

Se basan en añadir un bit a la secuencia a enviar de forma que el número de "1's" sea par o impar, dependiendo del tipo de paridad:

- ↓ Si la paridad es par, el número final de "1's" debe ser par
- ↓ Si la paridad es impar, el número final de "1's" debe ser impar

Es un código sistemático. Sólo será detector (vuelta atrás) y detectará los errores producidos en un número impar de bits. Si por ejemplo se producen 2 bits erróneos, este código no los detectará.

Ejemplo

Paridad par: 10111011
Paridad impar: 10111010

2.4.1.2 Códigos m entre n

Un código m entre n se caracteriza porque todas las palabras de código tienen la misma longitud de m bits, de los cuales, n bits son "1's". Es un código sistemático. Sólo será detector (vuelta atrás) y detectará los errores producidos en un número impar de bits. Si por ejemplo se producen 2 bits erróneos, este código no los detectará.

Ejemplo

3 "1's" entre 5 bits: 011 01
 100 11
 111 00

2.4.1.3 Códigos de bloque lineal sistemático

Un código de bloque lineal sistemático será capaz de detectar $X-1$ bits erróneos, donde X viene dado por la distancia de Hamming mínima entre 2 palabras cualesquiera del código, y será capaz de corregir $(X-1)/2$ bits erróneos.

Estos códigos cumplen la siguiente propiedad:

La suma módulo-2 de dos palabras del código da lugar a otra palabra de código.

En la especificación de estos códigos se utiliza la siguiente notación:

(n, k)

↓ n es el tamaño de la palabra codificada

↓ k es el tamaño del mensaje original. Estos k bits se envían sin alterar

↓ los $n-k$ bits restantes son los bits de paridad. Es la redundancia mediante la cual se detectan y corrigen los errores.

La forma de una palabra de código de un código de bloque lineal sistemático es la siguiente:

$$m_0 m_1 \dots m_{k-1} b_0 b_1 b_2 \dots b_{n-k-1} \quad (13)$$

donde $m_0 \dots m_{k-1}$ son los bits del mensaje original y $b_0 \dots b_{n-k-1}$ son los bits de paridad que se añaden como redundancia.

De esta forma, podemos expresar una palabra de código como:

$$C_0 C_1 \dots C_{n-k-1} C_{n-k} \dots C_{n-1}$$

El cálculo de los bits de paridad se realiza de la siguiente forma:

$$b_i = P_{0i} m_0 + P_{1i} m_1 + \dots + P_{k-1i} m_{k-1}$$

donde los P_{ij} deben ser tales que la matriz generadora del código tenga filas independientes y las ecuaciones de paridad sean iguales.

Para realizar la codificación se utiliza una notación matricial. Consideraremos la palabra original, la palabra formada por los bits de paridad y la palabra de código como vectores:

$$\begin{aligned} m &= (m_0 m_1 \dots m_{k-1}) \\ b &= (b_0 b_1 \dots b_{n-k-1}) \\ c &= (c_0 c_1 \dots c_{n-1}) \end{aligned} \quad (14)$$

También se utiliza la matriz de coeficientes:

$$P = \begin{pmatrix} P_{00}P_{01}\dots P_{0,x-k-1} \\ P_{10}P_{11}\dots P_{1,x-k-1} \\ \dots\dots\dots \\ P_{k-1,0}\dots P_{k-1,x-k-1} \end{pmatrix}$$

Para realizar la codificación se utiliza la matriz generadora:

$$G = [I_{k,k} \mid P]$$

Siendo $I_{k,k}$ la matriz identidad de tamaño $(k \times k)$.

De esta forma podemos obtener cada palabra de código a partir de cada palabra de mensaje original realizando la siguiente multiplicación:

$$c = m * G \tag{15}$$

Para realizar la decodificación, en destino se recibe un vector (c) de tamaño n y lo que se puede hacer es repetir la operación realizada en la codificación: se toman los primeros k bits y se calcula la redundancia usando la matriz generadora y se comprueba si la redundancia obtenida es igual a la redundancia recibida.

Otra opción más eficiente es la basada en el concepto de síndrome.

En el proceso de decodificación basado en el síndrome se utiliza la matriz de chequeo de paridad, que se define de la siguiente forma:

$$H = [I \mid P^t] \tag{16}$$

H tiene la propiedad de que sólo las palabras de código verifican que al multiplicarlas por H^t el resultado es el vector nulo. Esta propiedad será utilizada para la detección y corrección de errores.

A cada palabra que el receptor recibe a través del canal la denominaremos palabra recibida y la denominaremos r . Una palabra recibida la podemos expresar como:

$$r = c + e \tag{17}$$

Donde c es la palabra de código enviada por el emisor y e es una palabra de error.

Cada componente e_i de la palabra de error podrá valer 1 si hay un error en esa posición y 0 si no lo hay.

El receptor para realizar la codificación utiliza la matriz H para calcular el vector de síndrome de error a partir de la palabra recibida. El vector de síndrome de error se obtiene de la siguiente forma:

$$s = r \cdot H^t \tag{18}$$

El vector de síndrome tiene tantos elementos como bits de paridad se estén usando.

El vector de síndrome sólo depende de la secuencia de error y no de la palabra de código transmitida.

Si en la transmisión no se ha producido un error, el síndrome es el vector nulo:

$$s = r \cdot H^t = 0 \quad (19)$$

Si se ha producido un error la multiplicación de la palabra recibida por H^t nos da un vector que es igual a una de las filas de H^t . La posición que ocupa esa fila es la posición donde hay un error.

Notas acerca de estos códigos:

Para la detección y corrección de errores simples, la matriz H debe cumplir:

1. Todas las columnas de la matriz H deben ser diferentes. Esta condición hace que se pueda localizar la posición del error.
2. Ninguna de las columnas de H pueden ser todas ceros. Esta condición se debe a que el síndrome es el vector nulo cuando no hay error.

Cuando se quieren corregir más de un error, por ejemplo dos errores, la matriz H debe verificar:

1. Todas las columnas de la matriz H deben ser diferentes
2. Ninguna de las columnas de H pueden ser todas ceros
3. La suma de las columnas dos a dos debe ser diferente

(Si la matriz H tiene 5 columnas hay 10 sumas diferentes que se pueden hacer. Lo que dice esta propiedad es que todas ellas deben dar resultados diferentes).

Esta última condición complica el cálculo de códigos correctores de dos bits y en la práctica no se suelen utilizar. Cuando hay más de un error se pedirá una repetición de la secuencia al emisor (vuelta atrás).

2.4.1.4 Código de Hamming

Un código de Hamming es un código de bloque lineal. Al igual que en los códigos de bloque lineales sistemáticos, podemos denotar un código de Hamming mediante un par (n,k) . Sin embargo los valores de n y k deberán verificar una serie de condiciones:

- ↯ n es la longitud de la palabra de código
- ↯ k es el número de bits de datos de la palabra original sin codificar
- ↯ El número de bits de paridad será $m = n - k$, pero deberá cumplirse la siguiente relación entre la longitud de la palabra de código y el número de bits de paridad:

$$n = 2^m - 1 \quad \text{con } m \geq 3 \quad (20)$$

- ↯ Según esto también se cumplirá la siguiente relación entre el número de bits de datos y el número de bits de paridad:

$$k = 2^m - m - 1 \quad (21)$$

Por lo tanto, a cada palabra original la añadiremos unos bits de paridad para obtener la palabra de código, de forma que estos bits de paridad sirvan posteriormente para encontrar y corregir errores que se produzcan en la transmisión.

Cada uno de los bits de paridad que añadimos a una palabra original va a afectar a unas determinadas posiciones de la nueva palabra de código, de forma que tomarán un valor adecuado para que se cumpla el criterio de paridad (par o impar) preestablecido en las subcombinaciones afectadas por cada uno de estos bits de paridad.

En el proceso de decodificación, el receptor recibe una palabra de un código de Hamming, y deberá comprobar si es correcta o no, y en el caso de que no fuera correcta deberá comprobar en que bit se produjo el error y corregir ese error.

Para comprobar si la palabra recibida es correcta, el receptor debe utilizar los bits de paridad de la palabra y hacer con ellos un control de paridad. Para realizar el control de paridad creamos una palabra que tendrá un bit por cada uno de los bits de paridad utilizados. Cada uno de los bits de esta palabra tomará el valor 0 o 1 dependiendo de si el número de unos de las posiciones de la palabra de código afectadas por su correspondiente bit de paridad cumplen o no el criterio de paridad establecido. Interpretando la combinación resultante en binario natural tendremos dos posibilidades:

- ⬇ Que sea un 0, lo cual quiere decir que no se han producido errores en la transmisión
- ⬇ Que se corresponda a un número distinto de 0, lo cual quiere decir que durante la transmisión ha variado el bit situado en la posición indicada por ese número.

Una vez obtenida la palabra de código correcta, basta con quitar los bits de paridad para obtener la palabra original enviada por el emisor.

2.4.1.5 Códigos Cíclicos

Los códigos cíclicos son una subclase de los códigos de bloque lineales.

Son fáciles de codificar y cumplen las siguientes propiedades:

- ⬇ **Linealidad:** la suma módulo-2 de dos palabras del código es otra palabra del código.
- ⬇ **Cíclicos:** cualquier desplazamiento cíclico de una palabra del código también pertenece al código.

Al igual que en los códigos de bloque lineales sistemáticos y en los códigos de Hamming, denotaremos un código cíclico mediante un par (n,k) , donde n es la longitud de las palabras de código y k es la longitud de una palabra original.

Para el manejo de estos códigos se utiliza una notación polinómica, de forma que una palabra de código $C = (c_0, \dots, c_{n-1})$ la interpretaremos como un polinomio, y cada uno de los bits de la palabra de código será uno de los coeficientes de este polinomio:

$$C(x) = c_0 + c_1x + \dots + c_{n-1}x^{n-1} \quad (22)$$

A su vez, una palabra original $m = (m_0, \dots, m_{k-1})$ la interpretaremos como el polinomio:

$$m(x) = m_0 + m_1x + \dots + m_{k-1}x^{k-1} \quad (23)$$

Para generar $C(x)$ a partir de $m(x)$ se usa el polinomio generador $g(x)$ que es un factor de $x^n + 1$. Su grado es $n-k$. La obtención de la palabra codificada se hace de la siguiente forma:

$$C(x) = m(x) * g(x) \quad (24)$$

Así tenemos que un código cíclico queda perfectamente determinado por su polinomio generador.

Tal y como hemos planteado estos códigos hasta el momento, los códigos cíclicos no son sistemáticos.

Para realizar el control de errores se utiliza el polinomio de chequeo de paridad, que es un polinomio de grado k tal que:

$$g(x) \cdot H(x) = x^n + 1 \quad (25)$$

2.4.1.6 Códigos Cíclicos Sistemáticos

La palabra de código de un código sistemático está formada por los bits de la palabra original y una serie de bits de paridad.

Consideraremos la palabra de mensaje como un polinomio:

$$m(x) = m_0 + m_1x + \dots + m_{k-1}x^{k-1} \quad (26)$$

A su vez consideraremos que los bits de paridad forman una palabra que también interpretaremos en forma de polinomio:

$$b(x) = b_0 + b_1x + \dots + b_{n-k}x^{n-k-1} \quad (27)$$

Por lo tanto la palabra de código será:

$$C(x) = x^{n-k}m(x) + b(x) \quad (28)$$

Teniendo en cuenta todo esto, los pasos para obtener el código cíclico sistemático son:

1. Multiplicar la palabra original $m(x)$ por x^{n-k}
2. Dividir $x^{n-k}m(x)$ por $g(x)$, denotando el resto por $b(x)$. El resto es el polinomio con los bits de paridad
3. La palabra codificada será:

$$c(x) = b(x) + x^{n-k}m(x) \quad (29)$$

Al igual que en los códigos de bloque lineales sistemáticos, para realizar el control de errores en la decodificación se utiliza el síndrome.

La palabra recibida por el receptor la denotaremos por:

$$r(x) = r_0 + r_1x + \dots + r_{n-1}x^{n-1} \quad (30)$$

Para calcular el síndrome utilizaremos el polinomio generador, con el cual realizaremos la siguiente división:

$$r(x) \quad \overline{) \quad g(x)}$$

$$S(x) \quad q(x)$$

$$\text{Luego } r(x) = q(x) \cdot g(x) + S(x) \quad (31)$$

$S(x)$ es el polinomio de síndrome, y será de un grado $n-k-1$ o menor.

El error y el síndrome coinciden, por lo tanto, si no se producen errores en la transmisión, el síndrome valdrá 0.

Si se produce un error en los bits de paridad se puede corregir sumando el síndrome a la palabra recibida.

2.4.1.7 Códigos de Redundancia Cíclica

Son un tipo de códigos cíclicos especialmente buenos para la detección de errores:

- ↓ Se diseñan para detectar muchas combinaciones de errores.
- ↓ La implementación práctica es sencilla. Son los que se usan en la práctica.

2.4.1.8 Códigos Convolutionales

Se diferencian de los códigos de bloque en su forma estructural y las propiedades para corregir errores.

Los códigos de bloque suelen tener limitada la capacidad de corrección de errores alrededor de 1 o 2 símbolos erróneos por palabra de código. Estos códigos son buenos para utilizar en canales con baja probabilidad de error.

Los códigos convolutionales son adecuados para usar sobre canales con mucho ruido (alta probabilidad de error).

Los códigos convolutionales son códigos lineales, donde la suma de dos palabras de código cualesquiera también es una palabra de código. Y al contrario que con los códigos lineales, se prefieren los códigos no sistemáticos.

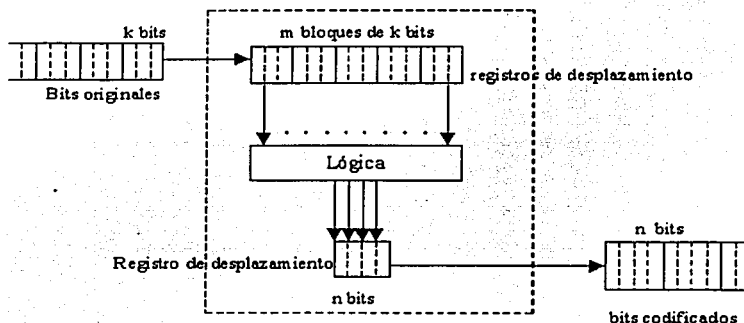
El sistema tiene memoria: la codificación actual depende de los datos que se envían ahora y que se enviaron en el pasado. Un código convolucional queda especificado por tres parámetros (n, k, m) :

n es el número de bits de la palabra codificada
 k es el número de bits de la palabra de datos
 m es la memoria del código o longitud restringida

2.4.2 Proceso de codificación

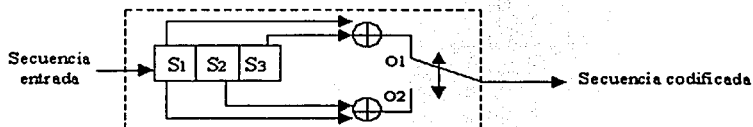
El proceso de codificación de estos códigos se realiza utilizando un dispositivo lógico en el codificador.

Ejemplo: Codificador convolucional $(4, 3, 5)$



La palabra codificada se obtendría como el resultado de realizar una serie de operaciones lógicas entre determinados bits que están almacenados en los registros intermedios.

Ejemplo: Codificador convolucional $(2, 1, 3)$



↪ El conmutador con las dos entradas hace el papel de un registro de desplazamiento de dos estados.

- ↳ El código convolucional es generado introduciendo un bit de datos y dando una revolución completa al conmutador.
- ↳ Inicialmente se supone que los registros intermedios contienen ceros.

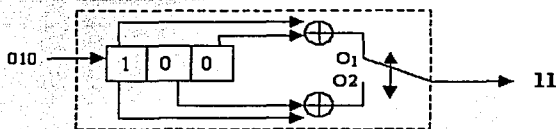
En este ejemplo la palabra codificada se obtiene como resultado de sumas módulo-2 entre los bits indicados que están almacenados en los registros intermedios.

Las secuencias de salida para el código anteriormente descrito:

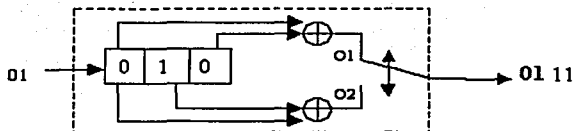
Entrada (S3,S2,S1)	Salida (O1,O2)
000	00
001	11
010	01
011	10
100	10
101	01
110	11
111	00

Como ejemplo del funcionamiento de este codificador, supongamos que se quiere enviar la secuencia de bits 0101 (donde los bits más a la derecha son los más antiguos). El proceso de codificación es el siguiente:

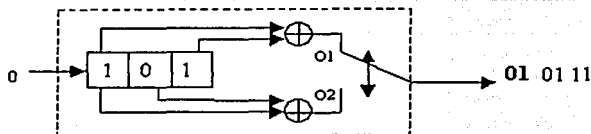
- Se introduce el primer bit de la secuencia en el codificador:



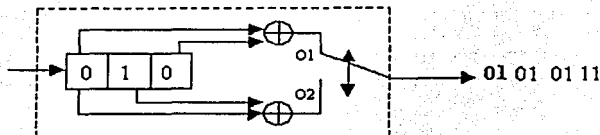
- Se introduce el segundo bit de la secuencia en el codificador:



- Se introduce el tercer bit de la secuencia en el codificador:



- Se introduce el cuarto bit de la secuencia en el codificador:



Al final del proceso de codificación obtenemos que la secuencia codificada es 01010111.

2.5 NORMAS DE COMPRESIÓN DE VIDEO [11]

2.5.1 JPEG

Imágenes Fijas

JPEG es un mecanismo estandarizado de compresión de imágenes.

JPEG se diseñó para comprimir ya sea imagen a color o en escala de grises. Este sistema de compresión trabaja bastante bien en fotografías y materiales similares, pero no lo hacen tan bien con escritura, caricaturas o dibujos de líneas. JPEG maneja solamente imágenes fijas, pero existe un estándar relacionado llamado MPEG para imágenes en movimiento.

JPEG es un "sistema que pierde información", lo que significa que la imagen descomprimida no es lo suficientemente parecida a la imagen original con la que se inició el proceso. Existen algoritmos de compresión de imágenes que tienen menos pérdida, pero JPEG logra una compresión mucho mayor que la que se logra con métodos con menor pérdida. JPEG es un estándar diseñado para explotar las limitaciones conocidas del ojo humano, haciendo énfasis en el hecho de que los pequeños cambios de color se perciben mucho menos precisos que aquellos pequeños cambios de brillantez. Así pues, JPEG está diseñada para comprimir imágenes que se vean bien a los ojos de los humanos. Por otro lado, si se pretende analizar las imágenes por medio de computadoras o equipos

similares, todos estos pequeños errores generados por JPEG pueden llegar a ser un verdadero problema, aún cuando ellos sean invisibles al ojo.

El estándar de compresión de imágenes JPEG no es simplemente un algoritmo, sino un paquete de técnicas de compresión de imágenes. Bajo este estándar de compresión de imágenes, se pueden seleccionar las características de la compresión. Existen 4 modos principales de compresión bajo JPEG:

- 1.- Sin pérdidas. La imagen se reproduce de manera idéntica a la original. Todos los demás modos tienen pérdidas.
- 2.- Secuencial. Este modo codifica la imagen en el orden en que fue "escaneada". Este es el modo en el que se procesan las imágenes generalmente.
- 3.- Progresivo. En este modo, primero se transmite una imagen burda que se puede desplegar rápidamente, y ésta se va definiendo de manera progresiva.
- 4.- Jerárquico. En este modo, la imagen se codifica de a diferentes resoluciones. El usuario puede seleccionar la resolución de visualización de la imagen.

La figura 2.5 muestra un diagrama representativo de la arquitectura de JPEG.

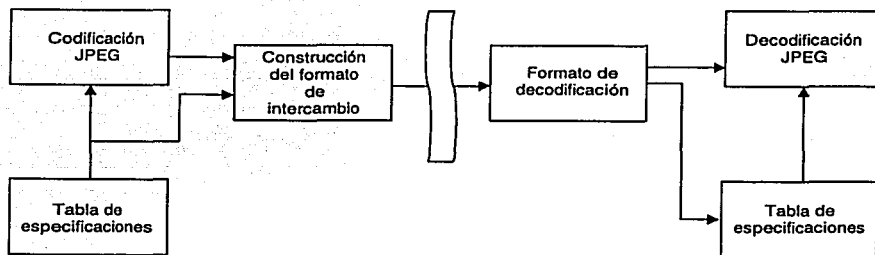


Figura 2.5. Arquitectura de JPEG

La imagen se codifica bajo el control de una o más tablas de especificaciones. La información ya codificada se envía, y en el otro extremo con un proceso inverso, se recupera la imagen original. Todas la codificaciones con JPEG (excepto en el modo sin pérdidas) se basan en la Transformada Coseno Discreta.

Cada componente de la imagen original se divide en bloques de 8 por 8 píxeles, a los que se les aplica el proceso TDC. Esto nos da como resultado un coeficiente de DC y 63 componentes de mayor frecuencia, cada uno de los cuales se cuantifica de acuerdo a una tabla de 64 valores de cuantificación.

Esta tabla de cuantificación, es la primera de las tablas de especificaciones del algoritmo, la cual debe ser proporcionada por el usuario y será parte de la cadena de datos que se transmitirá o recibirá. El usuario puede controlar la relación entre el grado de compresión y la calidad de la imagen dependiendo de las especificaciones de ésta tabla. El contenido de la tabla de cuantificación corresponde con el tamaño de los escalones que se utilizarán; el coeficiente DCT simplemente se divide por el tamaño del correspondiente escalón y el resultado se redondea al número de bits especificado. De esta manera, un gran valor de cuantificación resulta en una cuantificación burda y esto significa que valores grandes se redondearán a CERO.

Después de la cuantificación, el coeficiente de DC es codificado diferencialmente contra el coeficiente de DC del bloque de imagen precedente. Los otros 63 coeficientes cuantificados se ubican en un orden de zig-zag y se codifican estadísticamente de acuerdo, ya sea al método de HUFFMAN o a un método aritmético de codificación que se describe en el estándar. En cualquiera de los casos se requiere de una segunda tabla de especificaciones para definir la codificación estadística. Nuevamente, el usuario puede variar ésta tabla para controlar el rendimiento de la compresión.

El último paso de la codificación es ensamblar los datos comprimidos y las tablas de especificaciones en una cadena de bits unidimensional de acuerdo al formato de intercambio, que se muestra en la figura 2.6.

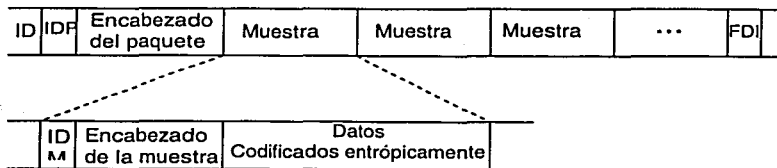


Figura 2.6. Formato típico de intercambio de información en JPEG secuencial.

Los datos de la imagen codificada para un solo cuadro de imagen se envían en un bloque, que (para formatos TDC) contiene todos los componentes de la imagen de manera intercalada. Las tablas de especificaciones se envían primero de tal manera que estén disponibles para la descompresión de los paquetes de datos que llegarán posteriormente. Los encabezados de los paquetes de información contienen los detalles del formato de la imagen y de su descompresión.

Unas de las grandes ventajas de JPEG es que el grado de pérdida de información se puede variar ajustando los parámetros de compresión. Esto significa que un creador de imágenes puede sacrificar el tamaño del archivo por una buena calidad de imagen de salida. También pueden realizar archivos "extremadamente"

pequeños si no interesa mucho la poca calidad de la imagen. Este estándar es útil para aplicaciones tales como archivos de índices de imágenes. Por otro lado si la calidad de la imagen a la salida no es la adecuada a los valores preestablecidos de compresión, ésta se puede mejorar hasta que sea satisfactoria, pero teniendo en cuenta que la compresión será mucho menor.

Otro aspecto importante de JPEG es que los decodificadores pueden sacrificar la velocidad de decodificación por la calidad de la imagen, usando aproximaciones más rápidas pero por otro lado más imprecisas que los cálculos requeridos. JPEG comprime simétricamente; lo que significa que la compresión y la descompresión toman la misma cantidad de tiempo. Algunos visualizadores logran un incremento notable de velocidad con estos métodos.

Existen dos razones importantes por las que se usa JPEG: para hacer los tamaños los archivos de imagen más pequeños y para almacenar información de color de 24 bits/pixel en lugar de información de 8 bits/pixel.

El obtener que los archivos de imágenes sean mas pequeños, es un gran logro para la transmisión de imágenes a través de una red y para el almacenamiento de bibliotecas de imágenes. Con JPEG se tiene la posibilidad de comprimir un archivo de color de 2 Mbytes a, aproximadamente 100 Kbytes, con lo que se logra un gran ahorro tanto en medio de almacenamiento como en tiempo de transmisión. JPEG puede, fácilmente, proporcionar una compresión de datos de color de 20:1. Si se comparan los archivos GIF y JPEG, la relación de tamaños es normalmente más de 4: 1.

Si los programas de visualización no soportan directamente JPEG, entonces se tiene que convertir el archivo JPEG en una imagen visible en otro formato. Aún cuando el visualizador sea capaz de ver imágenes JPEG, la decodificación y la visualización de imágenes JPEG toma mayor tiempo que la visualización de una imagen con un formato simple, tal como el GIF. Así pues, el usar JPEG esencialmente implica un sacrificio de tiempo/espacio. Pero esto puede llegar a ser mejor cuando se esta hablando de un sistema de transmisión a través del teléfono o a través de una red, ya que el ahorro en tiempo debido a la transferencia de archivos mas pequeños puede ser mucho mayor que el tiempo empleado para descomprimir el archivo.

La segunda ventaja fundamental de JPEG es que ésta almacena la información a pleno color 24 bits/pixel (16 millones de colores). El formato GIF, que es el otro formato ampliamente utilizado, solamente puede almacenar imágenes de 8 bits/pixel (256 o menos colores). Actualmente el formato GIF está siendo reemplazado por el formato JPEG, como lo fueron, hace algunos años desplazados formatos en blanco y negro. Aún más, JPEG es mucho más utilizado que GIF para el intercambio de imágenes entre las personas que tienen una gran diversidad de equipos de despliegue, ya que esto evita la predefinición de cuantos

colores se tienen que utilizar. Así pues, el uso de JPEG en lugar del GIF, es mucho más apropiado como formato estándar.

Mucha gente se espanta del término "compresión con pérdidas", pero no existe formato de imágenes digitales que puedan contener toda la información que reciben nuestros ojos. La desventaja real de la compresión con pérdidas, es que si se comprime y descomprime una imagen en repetidas ocasiones, en cada ocasión que se hace este procedimiento existe una pequeña pérdida de calidad. Esto puede ser un severo problema sobre todo en algunas aplicaciones, pero afortunadamente para la mayoría de ellas no.

JPEG comprime imágenes bastante bien en realidad, cuando se utiliza con cierto tipo de imágenes, tales como fotografías. Para imágenes a pleno color, la información no comprimida es normalmente de 24 bits/píxel. El método de compresión con menos pérdidas mejor conocido puede comprimir la información aproximadamente con una relación promedio de 2:1. A este tipo de relaciones de compresión si se quisieran guardar 30 segundos de vídeo en un medio de almacenamiento como un disco duro, de requeriría un espacio en disco aproximado de 400 MB.

JPEG puede lograr compresión de 10:1 a 20:1 "sin pérdida de información visible", logrando con esto una reducción de almacenamiento efectivo a 1 ó 2 bits/píxel. Actualmente se pueden alcanzar compresiones de 30:1 a 50:1 con algunos pequeños defectos moderados, y cuando se requieren imágenes de muy poca calidad tales como presentaciones preliminares o archivos de índices, se puede lograr, razonablemente bien, una compresión de 100:1. Una imagen comprimida con JPEG a 100:1 ocupa el mismo espacio que una imagen del tamaño de la uña del dedo gordo a pleno color y contiene información mucho más detallada que esa imagen. La desventaja con JPEG (y otros algoritmos de compresión) es que entre más se comprime el vídeo éste se ve peor. Además de esto, cada vez que se comprime el vídeo con JPEG a menos de una sexta parte de su tamaño original (6:1), entonces se van a ver pequeños cuadros, bordes no deseados, y colores no correspondientes a los originales en las imágenes.

Las imágenes en escala de grises no se pueden comprimir a factores tan altos. Debido a que el ojo humano es mucho más sensitivo a las variaciones de brillantez que a las variaciones de color, JPEG puede comprimir datos de color con una relación mayor que lo que puede hacer con información de brillantez (escala de grises). Un archivo con JPEG de escala de grises generalmente es aproximadamente de 10% a 25 % más pequeño que un archivo con JPEG a pleno color de una calidad visual similar. Debido a que los datos en escala grises antes de comprimirse son de solamente de 8 bits/píxel o una tercera parte de tamaño de los datos a color, la relación de compresión calculada es menor. El umbral de pérdida visible por compresión es alrededor de 5:1 para imágenes en escala de grises.

El valor exacto al cual un error se convierte en visible, depende de las condiciones de visualización. Entre más pequeño es un pixel independiente, es mucho más difícil de percibir el error; por lo que los errores son mucho más visibles en monitores de computadora (a aproximadamente 70 puntos/pulgada) que en un monitor de alta resolución a color (de 300 puntos/pulgada o más). Así pues, una imagen de mayor resolución puede tolerar mayor compresión. Las relaciones de compresión mencionadas arriba son relaciones típicas para unidades de visualización típicas.

Actualmente existen muchos muy diferentes métodos de compresión ellos conocidos como "JPEG", El método comúnmente usado es el JPEG original (o su variante JPEG progresivo). El mismo estándar

2.5.2 H.261

Códec para los Servicios Audiovisuales a $n \times 384$ kbps.

El CCITT considerando:

- ↓ Que existe una demanda significativa del servicio de videoconferencia.
- ↓ Que la transmisión digital a la velocidad de los canales H_n o sus múltiplos hasta la velocidad primaria puede proporcionar los circuitos necesarios para satisfacer esta demanda.
- ↓ Que la existencia de diferentes jerarquías y diferentes normas de televisión en diferentes partes del mundo complica los problemas relativos a la especificación de las normas de transmisión y codificación en las conexiones internacionales.

El objetivo básico del CCITT es recomendar soluciones únicas para las conexiones internacionales, reconociendo que el objetivo consiste en proporcionar interfuncionamiento entre los códecs a $n \times 384$ kbps y los códecs a $m \times 64$ kbps definidos en las recomendaciones de la serie H. El interfuncionamiento se hará tomando como base $m \times 64$ kbps; los valores de m están estudiándose.

Esta recomendación describe los métodos de codificación y decodificación de los servicios audiovisuales a las velocidades de $n \times 384$ kbps, donde n toma los valores de 1 a 5.

La figura 2.7 muestra un diagrama de bloques resumido del códec.

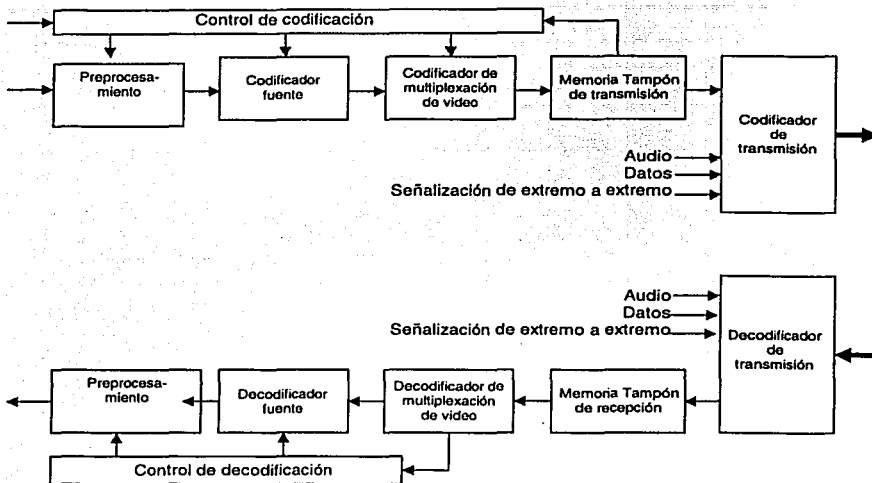


Figura 2.7. Diagrama de bloques resumido del códec.

Entrada y salida de video

Para poder abarcar con una sola recomendación la utilización en las regiones de 625 y 525 líneas entre las mismas, las imágenes se codifican en un formato intermedio común. Las normas de señales de televisión de entrada y salida, que puede, por ejemplo, ser compuestas o separadas en componentes, analógicas o digitales no son objeto de recomendación, como tampoco lo son los métodos para realizar cualquier conversión necesaria de y hacia el formato de codificación intermedio.

Frecuencia de muestreo

Las frecuencias se muestrean a un múltiplo entero de la velocidad de línea de video. El reloj de muestreo y el reloj de red digital son asíncronos.

Algoritmo de codificación fuente

Se adopta una combinación de predicción interimágenes para utilizar redundancia temporal y codificación de la transformada de la señal restante para reducir la redundancia espacial. El decodificador tiene capacidad de compensación de movimiento, permitiendo la incorporación facultativa de esta técnica en el codificador.

Codificador Fuente

El codificador fuente opera con imágenes no entrelazadas que ocurren 30,000/1001 (aproximadamente 29.97) veces por segundo. La tolerancia en la frecuencia de imagen es de +/- 50 ppm.

Las imágenes se codifican mediante las componentes de luminancia y las dos de diferencia de color (Y, CR y CB). Estas componentes y los códigos que representan sus valores muestreados se definen en la Recomendación 601 del CCIR.

Negro = 16

Blanco = 235

Diferencia de color nula = 128

Diferencia de color máxima = 16 y 240

Estos valores son nominales y el algoritmo de codificación funciona con los valores de entrada comprendidos entre 0 y 255.

Para la codificación, la estructura de muestreo de luminancia es de 288 líneas por imagen, 352 elementos de imagen por línea en una disposición ortogonal. El muestreo de cada una de las dos componentes de diferencia de color es de 144 líneas, 176 elementos de imagen por línea, ortogonal. Las muestras de diferencia de color se sitúan de manera que sus fronteras de bloque coincidan con las fronteras de bloque de luminancia, como se muestra en la figura 2.8. La zona de imagen cubierta por estos números de elementos de imagen y líneas tiene una relación de 4:3 y corresponde a la porción activa de la entrada de vídeo de norma local.

X		X	X		X	X		X
	0			0			0	
X		X	X		X	X		X
X		X	X		X	X		X
	0			0			0	
X		X	X		X	X		X
X		X	X		X	X		X
	0			0			0	
X		X	X		X	X		X

X Muestra luminancia
0 Muestra crominancia
---- Borde de bloque

Figura 2.8. Posicionamiento de las muestras de luminancia y crominancia

Algoritmo de codificación de fuente de video

Los principales elementos del algoritmo de codificación de fuente de video son la predicción, la transformación de bloque, la cuantificación y la clasificación.

El error de predicción (modo INTER) o la imagen de entrada (modo INTRA) se subdivide en 8 elementos de imagen por 8 bloques de línea que se segmentan como transmitidos o no transmitidos.

Los criterios de elección del modo y la transmisión de un bloque no son objeto de recomendación, y pueden variar dinámicamente como parte de la estrategia de control de la velocidad de datos. Los bloques transmitidos se transforman y los coeficientes resultantes se cuantifican y se codifican con longitud variable.

Compensación de movimiento

La compensación del movimiento es facultativa en el codificador. El decodificador aceptará un vector para cada bloque de 8 elementos de imagen por 8 líneas.

Un valor positivo de la componente horizontal o vertical del vector de movimiento significa que la predicción está formada a partir de elementos de imagen de la imagen anterior, situados espacialmente a la derecha o debajo de los elementos de imagen objeto de predicción.

Los vectores de movimiento están limitados de manera que todos los elementos de imagen por ellos referenciados estén dentro de la zona de imagen codificada.

Filtro de bucle

El proceso de predicción puede modificarse mediante un filtro especial bidimensional que actúa sobre los elementos de imagen de un bloque predicho.

El filtro es separable en funciones unidimensionales horizontal y vertical. Ambas son no recursivas con coeficientes de $1/4$, $1/2$, $1/4$. En los bordes del bloque, donde uno de los puntos de toma caería fuera del bloque, el elemento de imagen periférico se utiliza para dos puntos de toma. La precisión aritmética se conserva totalmente redondeando a valores enteros de 8 bits en la salida del filtro bidimensional. Los valores de cuya parte fraccionaria es un medio se aproximan al valor superior.

Dispositivo de transformación

Los bloques transmitidos se codifican mediante una transformada de coseno discreto bidimensional separable de 8×8 . La entrada de la transformada directa y la salida de la transformada inversa tienen 9 bits.

Recorte

Para evitar las amplitudes de distorsión de cuantificación de coeficiente de transformada que causen desbordamiento aritmético en los bucles del codificador y decodificador, se insertan funciones de recorte. Además de las funciones de

recorte aplicadas en la transformada inversa, se aplica una función de recorte, tanto en el codificador como en el decodificador, a la imagen reconstruida que se forma sumando la predicción y el error de predicción modificados por el proceso de codificación. Esta función de recorte actúa sobre los valores de elementos de imagen resultantes inferiores a 0 ó superiores a 255, cambiándolos a 0 y 255 respectivamente.

Control de Velocidad de datos

Las secciones en que pueden variarse parámetros para controlar velocidad de generación de datos de video codificados incluyen procesamiento antes del codificador fuente, el cuantificador, criterio de significado de bloque y submuestreo temporal. Las proporciones de tales medidas en la estrategia de control global no son objeto de recomendación.

Al ser invocado, el submuestreo, temporal se realiza descartando imágenes completas. Las figuras interpoladas no se colocan en la memoria de imágenes

Actualización Forzada:

Esta función se realiza forzando la utilización del modo INTRA del algoritmo de codificación.

Históricamente, H.261 fue desarrollada mucho antes que JPEG. En diciembre de 1984, el Grupo de Estudio XV del CCITT (equipo y sistemas de transmisión) estableció un grupo de "Especialistas en codificación para telefonía visual". El desarrollo de éste, estándar de transmisión de video para una baja tasa de transmisión de servicios ISDN ha pasado por varios estados. Al principio la meta fue diseñar un esquema de transmisión para tasas de transmisión de $e(m)$ veces 384 kb/s, donde $e(m)$ estuvo entre 1 y 5. Posteriormente en tasas de $e(n)$ veces 64 kbps donde $e(n)$ de 1 a 5 fueron considerados. Sin embargo, luego fue hecha en 1989 la recomendación final del CCITT H.26120.

De hecho, la sede H de teleservicios audiovisuales, es un grupo de estándares (o recomendaciones) que consisten en:

H.221 estructura de trama.

H.230 control síncrono de trama.

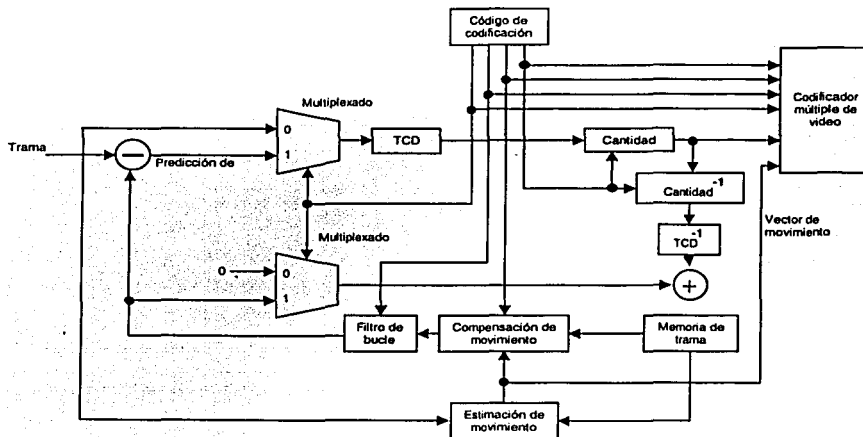
H.242 comunicación entre terminales audiovisuales.

H.320 sistemas y equipo terminal

H.261 códec de video.

Los códecs JPEG línea base y H.261 utilizan ambas técnicas DCT (Discrete Cosine Transform) y VLC (Variable-Length Code). La principal diferencia entre el esquema de compresión JPEG y H.261 es que JPEG codifica cada trama

individualmente, mientras que H.261 realiza codificación entre tramas. En H.261 la compensación de movimiento es realizada basada en bloques para registrar las diferencias entre tramas, las cuales son entonces codificadas en DCT. Aquí, el dato del cuadro en la trama anterior puede ser usado para predecir los bloques de imagen en la trama actual. Como resultado solamente las diferencias, típicamente de pequeña magnitud entre los bloques previamente desplazados y el bloque actual son transmitidos.



2.9. Codificador fuente H.261

Hay algunas características interesantes o consideraciones de diseño en H.261.

Primero, define esencialmente solo el decodificador. Sin embargo, el codificador el cual no es completa y explícitamente especificado por el estándar, se espera sea completamente compatible con el decodificador.

Segundo, porque H.261 es propuesto para comunicaciones en tiempo real, usa solamente la previa y más cercana trama como predicción para reducir el retardo de codificación.

Tercero, trata de balancear las complejidades del hardware del codificador y del decodificador, dado que ambos son necesarios para aplicaciones de videoteléfono en tiempo real. Otros esquemas de codificación, como el vector de cuantificación (VQ), pueden tener un decodificador un poco más simple, pero un codificador muy complejo.

Cuarto, H.261 es un compromiso entre el funcionamiento del codificador, requerimiento de tiempo real, complejidad de implementación, y robustez del sistema. El movimiento compensado codificado en DCT es un algoritmo maduro, y después de años de estudio, bastante general y robusto que puede manejar varios tipos de cuadros.

Quinto, las estructuras de codificación finales y parámetros son orientados más hacia aplicaciones de baja tasa de bit. Esta selección es lógica, porque la estructura del codificador y parámetros de codificación es más crítica para el desempeño del códec a muy bajas tasas de bit. A tasas de bit más altas, los valores de parámetros más bajos que el óptimo no afectan mucho el rendimiento del códec.

H.261 especifica un grupo de protocolos que comprime cada ráfaga de bits, y un grupo de operaciones que cada decodificador estándar compatible debe ser capaz de realizar. La actual implementación de hardware del códec y la estructura del codificador puede variar drásticamente de un diseñador a otro.

2.5.3 NORMA H. 263

Transmisión en línea de señales no telefónicas.

Codificación de video para la comunicación a baja velocidad binaria

Esta recomendación especifica una representación codificada, que se puede usar para comprimir el componente de imagen en movimiento de los servicios audiovisuales a velocidades binarias bajas.

El decodificador tiene capacidad de compensación de movimiento, lo que permite la incorporación opcional de esta técnica en el codificador. En la compensación de movimiento se utiliza la precisión de mitad de pixel, mientras que en la recomendación H.261 se utiliza la precisión de pixel entero y un filtro de bucle. Se emplea la codificación de longitud variable para la transmisión de los símbolos.

La Recomendación UIT H.263, es un híbrido de la predicción entre imágenes que utiliza la redundancia temporal y la codificación con transformada de la señal restante para reducir la redundancia espacial. La configuración básica del algoritmo de codificación de fuente de video se basa en la Recomendación UIT H.261, a continuación en la figura 3 aparece el diagrama de bloques resumido del códec.

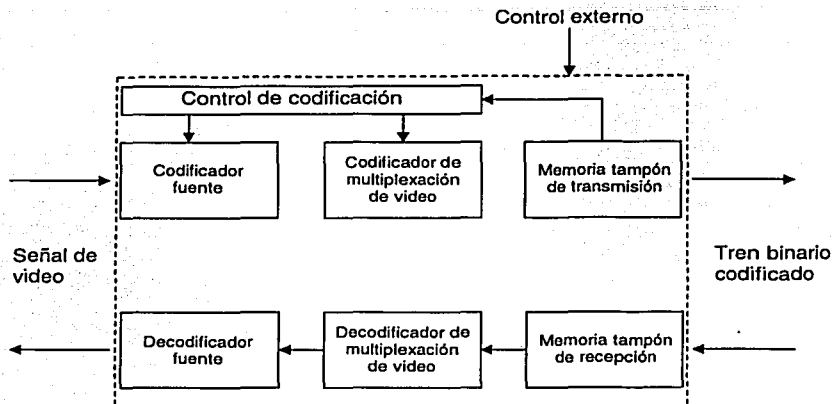


Figura 2.10. Diagrama de bloques resumido del códec de vídeo

Para poder abarcar con una sola recomendación la utilización dentro entre regiones que emplean normas de televisión de 625 y 525 líneas el codificador de fuente actúa sobre imágenes basadas en un formato común (CIF, Common intermediate format).

El codificador de vídeo proporciona un tren binario digital auto contenido que se puede combinar con otras señales con multifacilidades, (como se define en la norma H.223). Y el codificador de vídeo efectúa el proceso inverso. Las imágenes se muestrean a un múltiplo entero de la frecuencia de línea de vídeo. El reloj de muestreo y el reloj de red digital son asíncronos. Además del algoritmo de codificación de fuente de vídeo básico, se incluyen cuatro opciones de codificación negociables para mejorar el funcionamiento, estas opciones se pueden utilizar juntas o por separado.

Vector de Movimiento sin restricción.

En este modo opcional, los vectores de movimiento están autorizados para apuntar fuera de la imagen. Los pixeles de borde se utilizan como predicción de los pixeles "no existentes". Con este modo se consigue una ganancia significativa cuando hay movimiento a través de los bordes de la imagen, especialmente con los formatos de imagen más pequeños. Además, este modo incluye una extensión de la gama de vectores de movimiento de modo que se puedan utilizar vectores de movimiento mayores, lo que es especialmente útil en el caso de movimiento de la cámara

Codificación Aritmética basada en Sintaxis

En este modo facultativo, se utiliza la codificación aritmética en vez de la codificación de longitud variable. La relación señal a ruido (SNR) y las imágenes reconstruidas serán las mismas, pero se producirá un número considerable menor de bits.

Predicción avanzada

En este modo opcional, se utiliza la compensación de movimiento de bloques superpuestos (OBMC, Overlapped Block Motion Compensation) para la parte de luminancia de las imágenes P. En algunos macrobloques de la imagen, se utilizan cuatro vectores de 8x8 en vez de un vector de 16x16. El codificador decidirá el tipo de vector que se utilizará. Cuatro vectores utilizan más bits, pero dan una predicción mejor. La aplicación de este modo permite, en general un perfeccionamiento considerable. Más especialmente, se consigue una ganancia subjetiva porque OBMC resulta en menos dispositivos de bloqueo. transmitan únicamente los encabezados GOB no vacíos. Estas señales se transmiten por medios externos por ejemplo, Recomendación H.245.

Trama PB

Una trama PB consiste en dos imágenes que se codifican como una unidad E: I nombre "PB" proviene de los tipos de imagen de la recomendación H-262 en la que hay imágenes P e imágenes B. Por

nsiguiente una trama PB consiste en una imagen P, que se predice a partir de la anterior imagen P decodificada, y una imagen B, que se predice a partir de la anterior imagen P decodificada y la imagen P que esta siendo decodificada en ese momento.

El reloj de transmisión se proporciona por medios externos. La velocidad binaria de video puede ser variable. En esta recomendación no se imponen restricciones a la velocidad binaria de video estas restricciones las determinará la terminaj o la red. El codificador controlará su tren binario de salida para cumplir los requisitos de decodificador ficticio de referencia. Los datos de video se proporcionarán en cada ciclo de reloj válido. Esto se consigue mediante relleno de tipo de macrobloque y patrón de grupo modificado (MCBCP, macroblock type and codec block pattern) (véase las figuras 3.3.11 y 3.3.12) o, cuando se utiliza corrección de errores hacia adelante, también mediante tramas de relleno de corrección de errores hacia adelante.

El tratamiento de los errores se suministra por medios externos (recomendación H.223). Si no se proporciona por medios externos (por ejemplo, Recomendación H.221), se puede utilizar la alineación de trama y la corrección de errores opcional.

El decodificador puede enviar una instrucción para codificar uno o más GOB de la imagen próxima en modo INTRA con los parámetros de codificación adecuados para evitar el desbordamiento de la memoria tampón, Y también puede enviar una

instrucción para que se transmitan únicamente los encabezados GOB no vacíos. Estas señales se transmiten por medio externo por ejemplo, Recomendación H.245.

Codificador Fuente

El codificador de fuente trabaja con imágenes no entrelazadas que se presentan 30,000/1,001 (aproximadamente 29.97) veces por segundo. La tolerancia de frecuencia de imagen es de ± 50 ppm.

Las imágenes se codifican como una componente de luminancia y dos componentes de diferencia de color (Y, CR, CB). Estas componentes y los códigos que representan sus valores muestreados son los que define la recomendación 601 de CCIR.

Negro = 16

Blanco = 235

Diferencia de Color Nula = 128

Diferencia de Color Máxima = 16 y 240

Estos valores son nominales y el algoritmo de codificación funciona con valores de entrada comprendidos entre 1 y 254. Hay cinco formatos de imagen normalizados: sub-QCIF, QCIF, CIF, 4 CIF, y 16 CIF. Para cada uno de estos formatos de imagen, la estructura de muestreo de luminancia dx pixeles por línea y dy líneas por imagen en una disposición ortogonal. El muestreo de cada una de las dos componentes de diferencia de color se efectúa a dx/2 pixeles por línea y dy/2 líneas por imagen ortogonal. Los valores de dx, dy, dx/2 y dy/2 se ofrecen en la tabla 1 para cada uno de los formatos de imagen.

En cada uno de los formatos de imagen, las muestras de diferencia de color se colocan de forma tal que sus límites de bloque coincidan con los límites de bloque de luminancia. La relación de aspecto de píxel en la misma para cada uno de estos formatos de imagen, es la misma que se define para QCIF y CIF en la recomendación H.261: $(4/3) * (288/352)$. El área de la imagen cubierta por todos los formatos de imagen, salvo el sub-QCIF, tiene una relación de aspecto de 4:3.

Formato de imagen	Número de píxeles de luminancia	Número de líneas de luminancia Dy	Número de píxeles de crominancia dx/2	Número de líneas de crominancia dy/2
sub-QCIF	128	96	64	48
QCIF	176	144	88	72
CIF	352	288	176	144
4 CIF	704	576	352	288
16 CIF	1408	1152	704	576

Tabla 2.3. Número de píxeles por línea y números de líneas de cada formato de imagen de H.263.

Todos los decodificadores deberán poder funcionar con sub-QCIF. Los medios externos, como la recomendación H.245 indican los formatos que puede manipular el codificador. Para un panorama completo de los posibles formatos de imagen y los algoritmos de codificación de video, por ejemplo la recomendación H.324.

X		X	X		X	X		X
	0				0			0
X		X	X		X	X		X
X		X	X		X	X		X
	0				0			0
X		X	X		X	X		X
X		X	X		X	X		X
	0				0			0
X		X	X		X	X		X

X Muestra luminancia
0 Muestra crominancia
---- Borde de bloque

Figura 2.11. Posición de las muestras de luminancia y crominancia

Colocando un número mínimo de imágenes no transmitidas, entre las imágenes transmitidas se proporcionarán los medios para restringir el periodo máximo de transmisión de imagen de los codificadores.

Algoritmo de codificación de fuente de video

Los principales elementos son la predicción, la transformación de bloques y la cuantificación.

Grupos de bloques y macrobloques

Cada imagen se divide en grupos de bloques (GOB, Group of blocks). Un grupo de bloques comprende $k \cdot 16$ líneas, según el formato de imagen $k=1$ para sub-QCIF, QCIF y CIF, $k=2$ para 4CIF, $k=4$ para 16CIF. El número de GOB por imagen es de 6 para sub-QCIF, 9 para QCIF y 18 para CIF, 4 CIF y 16 CIF.

Cada GOB se divide en macrobloques. Un macrobloque se relaciona con 16 píxeles por 16 líneas de Y, y con 8 píxeles 8 líneas correspondientes espacialmente de C_B y C_R . Además, un macrobloque consiste en cuatro bloques de luminancia y los dos bloques de diferencia de color correspondientes espacialmente. Cada bloque de luminancia o de crominancia se relaciona con 8 píxeles por 8 líneas de Y, C_B o C_R .

La numeración de macrobloques se efectúa mediante un barrido horizontal de las filas de macrobloque de izquierda a derecha, iniciando por la línea de

macrobloque superior y terminando por la línea de macrobloque en orden de macrobloque creciente.

Los criterios de elección del modo y la transmisión de un bloque no son objeto de recomendación y pueden variar dinámicamente como parte de la estrategia de control de la codificación. Los bloques transmitidos se transforman y los coeficientes resultantes se cuantifican y se codifican en entropía.

Predicción

La predicción se efectúa entre imágenes y se puede aumentar por compensación de movimiento. El modo de codificación en el que se aplica la predicción se denomina INTER, el modo de codificación se denomina INTRA cuando no se aplica la predicción. El modo de codificación INTRA se puede señalar en el nivel de la imagen (INTRA para imágenes I o INTER para imágenes P) o el nivel de macrobloque en las imágenes P. En el modo de trama PB opcional, las imágenes B están siempre codificadas en modo INTER. Las imágenes B se predicen parcialmente de forma bidireccional.

Compensación de movimiento

El codificador aceptará un vector por macrobloque o, si se utiliza el modo de predicción avanzada de H.263 aceptará uno o cuatro vectores por macrobloque para la adaptación de vectores de movimiento destinado a la predicción de macrobloque B.

Un valor positivo del componente horizontal o vertical del vector significa que la predicción se forma a partir de los píxeles de la imagen referenciada que están espacialmente a la derecha o por debajo de los píxeles que se predicen.

Los vectores de movimiento están restringidos de modo que todos los píxeles referenciados por ellos se encuentren dentro de la zona de imagen codificada, salvo cuando se utiliza el modo de vector de movimiento sin restricción y/o el modo de predicción avanzada.

Cuantificación

El número de cuantificadores es 1 para el primer coeficiente de bloque INTRA y 31 para todos los demás coeficientes, excepto el primero de los bloques INTRA. No se definen los niveles de decisión. El primer coeficiente de los bloques INTRA es nominalmente, el valor de la transformada en continua (cc) cuantificado uniformemente con un paso de valor 8. Cada uno de los restantes 31 cuantificadores utilizan niveles de reconstrucción de espaciado igual con una zona muerta central alrededor de cero y un paso de-valor par entre 2 y 62.

Control de la codificación

Se pueden cambiar varios parámetros para controlar la velocidad de generación de datos de video codificado. Se trata del proceso efectuado antes del proceso del codificador de fuente, el cuantificador, el criterio de significación de bloque y el

submuestreo temporal. Las proporciones de estas medidas en la estrategia de control global no son objeto de esta recomendación. Al ser invocado, el muestreo temporal se ejecuta mediante el descarte de imágenes completas.

Un decodificador puede indicar su preferencia por un compromiso específico entre la resolución espacial y la temporal de la señal de video. El codificador señalará su compromiso por defecto al comienzo de la llamada e indicará si es capaz de responder a las solicitudes del decodificador de cambiar este compromiso. Estas señales se transmiten por medios externos, como se puede observar en la recomendación H.245.

Actualización Forzada

Esta función se realiza forzando la utilización del modo INTRA del algoritmo de codificación. El esquema de actualización no se define.

Para controlar la acumulación de errores por desajuste de la transformada inversa, cada macrobloque se codificará en modo INTRA al menos una vez cada 132 veces cuando se transmiten los coeficientes de este macrobloque en imágenes P.

Alineación en byte de los códigos de comienzo

La Alineación en byte de los Códigos de Comienzo se consigue insertando una palabra de código de menos de 8 bits cero antes del código del comienzo, de manera que el primer bit del código del comienzo es el primer bit más significativo de un byte. Por consiguiente un código de comienzo esta alineado en byte si la posición de su bit más significativo es un múltiplo de 8 bits del primer bit del tren binario de video. Es obligatorio que los códigos de comienzo de todos los tipos de imagen estén alineados en byte, y los códigos de comienzo de GOB y EOS no están sujetos a esta obligación.

Sintaxis y semántica

El múltiplex de video tiene una estructura jerárquica de cuatro capas, de arriba hacia abajo, esas capas son:

- ↓ Imagen
- ↓ Grupo de Bloques
- ↓ Macrobloque
- ↓ Bloque

El bit más significativo se transmite primero. Es el bit 1 situado en el extremo izquierdo de las tablas de bits de esta recomendación. A menos que se especifique otra cosa, todos los bits no utilizados o de reserva se fijan a 1, Los bits de reserva no se utilizarán hasta que la UIT no especifique sus funciones.

CAPÍTULO 3 SEGUNDA GENERACIÓN DE COMUNICACIONES MÓVILES

3.1 INTRODUCCIÓN

A inicio de los años 80, los sistemas de telefonía celular analógicos estaban experimentando un rápido crecimiento en Europa, particularmente en Escandinavia y en Inglaterra, así como en Francia y Alemania. Cada ciudad desarrollaba su propio sistema, lo cual lo hacía incompatible en equipamiento y operación con los demás. Esta era una situación no deseable, porque no solamente hubo limitación en la operación del equipo móvil dentro de los límites nacionales, que dentro de Europa fueron cada vez menos importantes, pero también hubo un mercado limitado para cada tipo de equipo, entonces economías a escala y subsecuentes ahorros no fueron realizados.

En 1982 la Conferencia de Postes y Telégrafos (CEPT: Conference of European Posts and Telegraphs) formó un grupo de estudio llamado Grupo Móvil Especial (GSM: Groupe Spécial Mobile) para estudiar y desarrollar un sistema móvil público europeo. El sistema propuesto debía de cubrir cierto criterio:

- ✚ Buena calidad de discurso (recepción de la llamada)
- ✚ Bajo costo en la terminal y servicio
- ✚ Soporte para roaming internacional
- ✚ Habilidad para soportar terminales handheld
- ✚ Tener soporte de nuevos servicios y medios (crecimiento)
- ✚ Eficiencia espectral
- ✚ Compatibilidad con ISDN

En 1989, la responsabilidad de GSM fue transferida al Instituto de Estándares de Telecomunicaciones Europeo (ETSI: European Telecommunication Standards Institute), la Fase I de las especificaciones de GSM fueron publicadas en 1990. El servicio comercial fue iniciado a mediados de 1991, y para 1993 había 36 redes GSM en 22 países. Aunque en Europa se estandarizó, GSM no sólo se considera como un estándar europeo. Más de 200 redes GSM (incluyendo DCS1800 y PCS1900) son operacionales en 110 países alrededor del mundo. A inicio de 1994 ya habían 1.3 millones de suscriptores a nivel mundial y a finales de octubre de 1997, había crecido a más de 55 millones. Con América del Norte se hace una entrada tardada en el campo del servicio móvil, con un derivado de GSM llamado PCS1900, estos sistemas móviles existen en cada continente, y la sigla GSM ahora declina las posiciones para el Sistema Global para Comunicaciones Móviles.

Los diseñadores de GSM escogieron un sistema digital no probado (para ese momento), como oposición a los sistemas celulares analógicos como AMPS en Estados Unidos y TACS para Inglaterra. Estos diseñadores tenían fe en que los avances en algoritmos de compresión y procesadores de señales digitales permitirían el cumplimiento del criterio original y la mejora incesante del sistema en términos de calidad y costo. En más de 8000 páginas de recomendaciones de GSM, se intenta permitir la flexibilidad y la innovación competitiva entre los proveedores, proporcionando la regularización para garantizar la integración apropiada entre los componentes del sistema. Esto es hecho proporcionando descripciones funcionales y de interfase para cada una de las identidades funcionales definidas en el sistema.

3.2 SERVICIOS PROPORCIONADOS POR GSM

Desde el principio, los proyectistas de GSM quisieron compatibilidad de ISDN en los términos de los servicios ofrecidos y el control de la señal usada. Sin embargo, las limitaciones de transmisión por medio de radio, en términos de ancho de banda y costo, no permiten al estándar ISDN canal-B con tasa de transmisión de 64kbps que sea prácticamente alcanzado.

Usando las definiciones de ITU-T, los servicios de la telecomunicación pueden ser divididos en servicios de portador, teleservicios y servicios suplementarios. El teleservicio más básico apoyado por GSM es la telefonía. Así como con otras comunicaciones, el discurso es codificado digitalmente y transmitido a través de la red GSM como una cadena digital. Hay también un servicio de emergencia, donde el proveedor de este servicio más cercano es notificado marcando tres dígitos (similar al 911).

Una variedad de servicios de datos es ofrecido. Los usuarios de GSM pueden enviar a tasas arriba de 9600 bps, a usuarios con un viejo plan de servicio telefónico POTS (Plain Old Telephone Service), ISDN, Red de Datos Pública por Conmutación de Paquetes y Red de Datos Pública por Conmutación de Circuitos utilizando una variedad de métodos de acceso y protocolos como lo son X.25 o X.32. Debido a que GSM es una red digital, un módem no es requerido entre el usuario y la red GSM, a pesar de que un módem de audio es requerido dentro de la red de GSM para interoperar con POTS.

“Otros servicios de datos incluye el facsímil del Grupo 3, como se describe en la recomendación T.30 de la ITU-T, la cual es soportada por medio del uso de una adaptador de fax apropiado. Una única característica de GSM, no encontrada en sistemas analógicos antiguos es el Servicio de Mensajes Cortos (SMS: Short Message Service). SMS es un servicio bidireccional para mensajes alfanuméricos cortos (hasta 160 bytes). Para SMS punto a punto, un mensaje puede ser enviado a otro suscriptor y el reconocimiento del receptor es provisto al emisor. SMS también puede ser utilizado en forma de broadcast, para enviar mensajes como

actualizaciones de tráfico o actualizaciones de noticias. Los mensajes también pueden ser almacenados en SIMS." [12].

Se proporcionan servicios suplementarios sobre teleservicios o servicios de portador. En la especificación actual (Fase I), se incluyen varias formas de reenvío de llamada (como reenvío de llamadas cuando el suscriptor móvil es inalcanzable por la red), y retención de llamadas entrantes y salientes, por ejemplo cuando te encuentras en otro país. Se proporcionan muchos servicios suplementarios adicionales en la especificación de la Fase II, como, identificación de llamada, llamada en espera, y multiconversación.

3.3 ARQUITECTURA DE LA RED GSM

Una red GSM está compuesta de varias entidades funcionales cuyas funciones e interfaces son específicas. La figura 3.1 muestra el esquema de una red GSM genérica. La red GSM puede ser dividida en tres partes importantes. La Estación Móvil, llevada por el suscriptor. El Subsistema de la Estación Base, controla el enlace de radio con la Estación Móvil. El Subsistema de la Red, la parte principal es el Centro de Conmutación de servicios Móviles (MSC: Mobile Services Switching Center), realiza la conmutación de llamadas entre los usuarios móviles, y entre los usuarios de la red móviles y fijos. El MSC también maneja la movilidad de operaciones de administración. No mostrado es el centro de mantenimiento y operación, el cual vigila el funcionamiento apropiado y arreglo de la red. El subsistema de estación móvil y de estación base comunicados a través de una interfase Um, también conocido como interfase aérea o enlace de radio. El subsistema de estación base se comunica con el Centro de Conmutación de servicios Móviles a través de una interfase A.

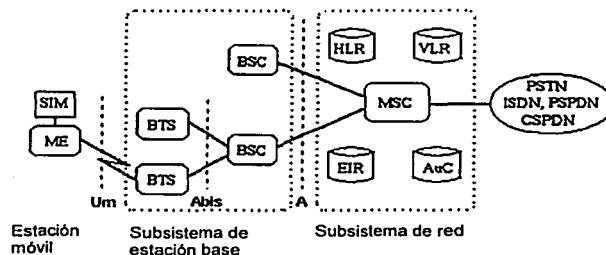


Figura 3.1. Arquitectura general de una red GSM

3.3.1 Estación móvil

La estación móvil (MS) consiste de equipamiento móvil (la terminal) y una tarjeta inteligente llamada Módulo de Identidad del Suscriptor SIM (Subscriber Identity Module). El SIM proporciona movilidad personal, de tal manera que el usuario pueda tener acceso a servicios suscritos de una terminal específica. Por medio de insertar un SIM dentro de otra terminal GSM, el usuario será capaz de recibir llamadas a dicha terminal, hacer llamadas desde dicha terminal y recibir otros servicios suscritos.

El equipo móvil es identificado singularmente por la Identidad de Equipo Móvil Internacional (IMEI: Internacional Mobile Equipment Identity). El SIM contiene la Identidad del Suscriptor Móvil Internacional (IMSI: Internacional Mobile Subscriber Identity), utilizada para identificar al suscriptor con el sistema, una llave secreta para autenticación y otra información. El IMEI y el IMSI son independientes por lo que permiten movilidad personal. El SIM puede protegerse contra el uso no autorizado por una contraseña o el número de identidad personal.

3.3.2 Subsistema de estación base

El subsistema de estación base está compuesto de dos partes, la Estación de Tranceptor Base (BTS: Base Transceiver Station) y el Controlador de Estación Base (BSC: Base Station Controller). Estos se comunican a través de la interfase estandarizada Abis, permitiendo (como el resto del sistema) operación entre componentes hechos por distintos proveedores.

La BTS hospeda los tranceptores de radio que definen una celda y maneja los protocolos del enlace de radio con la Estación Móvil. En una extensa área urbana, potencialmente habrá un gran número de BTSs desplegados, así los requerimientos para un BTS son: fiabilidad, portabilidad y costo mínimo.

El BSC maneja los recursos de radio para uno o mas BTSs. Este maneja la configuración del canal de radio, frecuencia, y handovers. El BSC es la conexión entre la estación móvil ay el MSC.

3.3.3 Subsistema de red

El componente central del Subsistema de la Red es el MSC. Actúa como un nodo de conmutación normal de PSTN o ISDN y adicionalmente proporciona toda la funcionalidad requerida para manejar un suscriptor móvil como registro, autenticación, actualización de ubicación, handover y redireccionamiento de llamada a un suscriptor foráneo. Estos servicios son proporcionados en conjunto con gran cantidad de identidades funcionales, que en conjunto forman es subsistema de red. El MSC proporciona la conexión a redes fijas (como lo es PSTN e ISDN). Señalización entre identidades funcionales en el subsistema de

res utilizan SS7 (Signalling System Number 7), utilizado para señalización troncal en ISDN y ampliamente utilizado en redes públicas actuales.

El Registro de Ubicación de Casa (HLR: Home Location Register) y el Registro de Ubicación de Visitante (VLR: Visitor Location Register), en conjunto con el MSC, proporcionan el redireccionamiento de llamada y la capacidad de roaming para GSM. El HLR contiene toda la información administrativa para cada suscriptor registrado en la correspondiente red GSM, con la ubicación actual del móvil. La ubicación del móvil es típicamente en la forma de dirección de señalización del VLR asociado con la estación móvil. Existe un HLR por red GSM, a pesar de que debe ser implementado como una base de datos distribuida.

El VLR contiene información administrativa seleccionada desde el HLR, necesaria para control de llamada y provisión del servicio suscrito, para cada móvil actualmente localizado en el área geográfica controlada por el VLR. A pesar de que cada identidad funcional puede ser implementada como unidad independiente, todos los fabricantes de equipo de conmutación para implementar el VLR junto con el MSC, de tal manera que el área geográfica controlada por el MSC correspondiente a tal controlador por el VLR, esto simplificando los requerimientos de señalización. El MSC no contiene información acerca de una estación móvil particular, esta información es almacenada en registros de ubicación.

Los otros dos registros se usan para la autenticación para propósitos de seguridad. El Registro de Identidad de Equipo (EIR) es un banco de datos que contiene una lista de todo el equipo móvil válido en la red, donde cada estación móvil es identificada por su Identidad de Equipo Móvil Internacional (IMEI). Un IMEI es marcado como inválido si se ha informado como robado o no ha sido aprobado. El Centro de la Autenticación (AuC) es un banco de datos protegido que guarda una copia de la llave confidencial guardada en el SIM de cada suscriptor que se usa para la autenticación y encriptación sobre el canal de radio.

3.4 Aspectos del enlace de radio

La Unión Internacional de Telecomunicaciones (ITU) que administra la asignación internacional de frecuencia de radio (entre muchas otras funciones), asignó la frecuencia de 890-915 MHz para el uplink (estación móvil a estación base) y 935-960 MHz para downlink (estación base a estación móvil) para las redes móviles en Europa. Puesto que este rango ya estaba utilizándose en los inicios de los 1980s por los sistemas analógicos, el CEPT reservó los últimos 10 MHz de cada banda para la red de GSM que todavía estaba desarrollándose GSM esta ubicada en un ancho de banda de 2x25 MHz.

3.4.1 Acceso Múltiple y Estructura del Canal

Puesto que la frecuencia de radio es un recurso limitado compartido por todos los usuarios, un método debe inventarse para dividir el ancho de banda, entre tantos usuarios como sea posible. El método escogido por GSM es una combinación de TDMA y FDMA. La parte de FDMA involucra la división a través de frecuencia de un ancho de banda (máximo) en 124 frecuencias portadoras espaciadas a 200 kHz. Una o más frecuencias portadoras son asignadas a cada estación base. Cada una de estas frecuencias portadoras es entonces dividida en tiempo, utilizando el esquema TDMA. La unidad fundamental de tiempo en este esquema TDMA es llamado "*burst period*" (periodo de estallido) y dura 15/26 ms. 8 periodos de estallido son agrupados en una trama TDMA (120/26 ms aproximadamente), el cual forma una unidad básica para la definición de canales lógicos. Un canal físico es un periodo de estallido por trama TDMA.

Los canales son definidos por el número y posición de sus correspondientes periodos de estallido. Todas estas definiciones son cíclicas, y el modelo entero se repite cada 3 horas aproximadamente. Los canales pueden ser divididos en canales dedicados que se asignan a una estación móvil y canales comunes que son utilizados por estaciones móviles en modo *idle*.

3.4.1.1 Canales de tráfico

Un canal de tráfico (TCH) se usa para llevar el tráfico de datos y el discurso. Se definen canales de tráfico utilizando una multitrama de 26 tramas o grupos de 26 tramas TDMA. La longitud de una multitrama de 26 tramas es de 120ms, que es cómo la longitud de un periodo de estallido se define (120 ms divididos entre 26 tramas, divididos en 8 periodos de estallido por trama). Fuera de las 26 tramas, 24 son utilizadas para tráfico, 1 es utilizada para SACCH (Show Associated Control Channel) y 1 actualmente no se utiliza. TCHs para *uplink* y *downlink* están separadas en tiempo por 3 periodos de estallido, de tal manera que la estación móvil no tenga que transmitir y recibir simultáneamente, simplificándose la electrónica. Además de éstos TCHs a tasa completa (full rate) también hay THCs definidos a tasa media (half-rate).

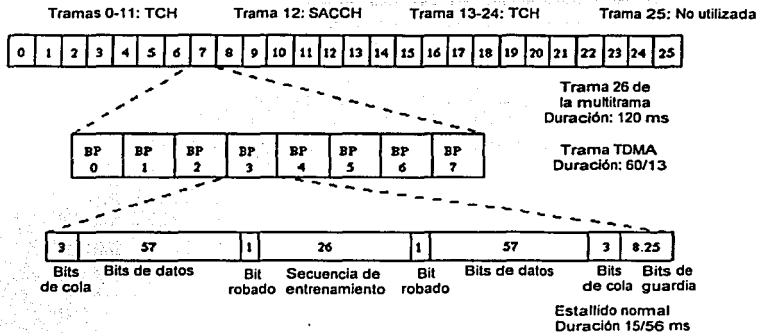


Figura 3.2. Organización de estallidos, TDMA idea, y multiframe para el discurso y datos

3.4.1.2 Canales de control

Canales comunes pueden ser accedidos tanto para modo *idle* y modo dedicado. Los canales comunes pueden ser utilizados por móviles en modo *idle* para intercambiar información de señalización requerida para cambiar a modo dedicado. Móviles en modo dedicado monitorean las estaciones base para handover y otra información. Los canales comunes son definidos dentro de una multiframe de 51 tramas, de tal manera que móviles utilizando estructura multiframe de 26 tramas, también TCH puede monitorear canales de control. Los canales comunes incluyen:

- ↓ BCCH (Broadcast Control Channel)
- ↓ FCCH (Frequency Correction Channel)
- ↓ RACH (Random Access Channel)
- ↓ PCH (Paging Channel)

3.4.1.3 Estructura de Estallido

"Hay cuatro tipos diferentes de estallido usados para la transmisión en GSM. El Estallido normal se usa para llevar datos y señalización. Tiene una longitud total de 156.25 bits. El estallido F utilizado en FCCH y el estallido S utilizado en SCH tiene la misma longitud que un estallido normal, pero una estructura interna diferente, lo cual lo diferencia del estallido normal (permitiendo sincronización). El estallido de acceso es mas corto que el estallido normal y es utilizado únicamente en RACH." [13]

3.4.2 Codificación de discurso

GSM es un sistema digital, y el discurso inherentemente analógico, tiene que ser digitalizado. El método empleado por ISDN, y para sistemas telefónicos para multiplexación de líneas de voz sobre troncales de alta velocidad y líneas de fibra óptica es PCM (Pulse Coded Modulation). La cadena de salida desde PCM es de 64kbps, la cual es muy elevada para ser enviada sobre un enlace de radio. La señal de 64 Kbps, a pesar de que es simple de implementar, contiene mucha redundancia. El grupo GSM estudió gran cantidad de algoritmos de codificación de discurso en base a la calidad de discurso subjetiva y complejidad antes de arribar a RPE/LPC (Regular Pulse Excited/Linear Predictive). Básicamente, información desde previas muestras, las cuales no cambian muy rápido, es utilizado para predecir la muestra actual. Los coeficientes de la combinación lineal de las muestras previas mas la forma codificada del residuo, la diferencia entre la muestra predicha y la actual, representa la señal. El discurso es dividido en muestras de 20 ms, cada una de las cuales es codificada como 260 bits, dando en total una tasa de transmisión de 13 kbps. Esto es lo que se conoce como codificación de discurso a tasa completa.

3.4.3 Codificación de canal y modulación

Debido a la interferencia electromagnética natural y artificial, el discurso codificado o señal de datos transmitida sobre la interfase de radio debe ser protegida de errores. GSM utiliza codificación convolucional e interpolación de bloques para llevar a cabo esta protección.

Quando se realiza una nueva llamada (recall) el discurso produce 260 bits por bloque para cada 20 ms. De la comprobación subjetiva, se encontró que algunos momentos de este bloque eran más importantes para calidad del discurso percibida que otros. Los bits son así divididos en tres clases:

- ↓ Clase Ia 50 bits - más sensible a los errores
- ↓ Clase Ib 132 bits - ligeramente sensible a los errores
- ↓ Clasificación II 78 bits – menos sensible a los errores

3.4.4 Ecuilización Multiruta

En el rango de 900 MHz, las ondas de radio se emanan a cualquier parte, construcciones, montañas, carros, aviones, etc. Esta gran cantidad de señales reflejadas, cada una con diferente fase, puede llegar a una antena. La ecuilización es utilizada para extraer la señal deseada. Este trabaja por medio de encontrar como una señal conocida transmitida es modificada por medio de alimentación multiruta y construyendo un filtro inverso para extraer el resto de la señal deseada.

3.4.5 Salto de Frecuencia

La estación móvil ya tiene que ser una frecuencia ágil, lo cual significa que se puedan mover entre transmisión, recepción y monitorear los slots de tiempo dentro de una trama TDMA, los cuales normalmente están en distintas frecuencias. GSM hace uso de su inherente agilidad de frecuencias para implementar salto de frecuencia baja, donde el móvil y el BTS transmiten cada trama TDMA en una portadora de frecuencia distinta. El algoritmo de salto de frecuencia es difundido en el Canal de Control de Broadcast. Debido a que la alimentación multiruta es dependiente de la frecuencia de portadora, el salto de frecuencia baja ayuda a corregir el problema. Además la interferencia co-canal es en efecto aleatoria.

3.4.6 Transmisión Discontinua

La minimización de interferencia co-canal es una característica en cualquier sistema celular, debido a que permite mejor servicio para un tamaño de célula dado o el uso de células mas pequeñas, esto incrementando la capacidad del sistema. Transmisión discontinua (DTX: Discontinuos Transmition) es un método que toma ventaja del hecho de que una persona habla menos del 40% del tiempo en una conversación normal [22], por medio de apagar el transmisor durante periodos de silencio. Un beneficio adicional de DTX es que la potencia se conserva en la unidad móvil.

El componente más importante DTX es la Detección de Actividad de Voz. Este debe distinguir entre voz y entradas de ruido, tarea que no es tan trivial como parece, considerando ruido de fondo. Si una señal de voz es malinterpretada como ruido, el transmisor se apaga y un efecto muy molesto llamado truncamiento es escuchado por el receptor. Si por el otro lado, ruido es malinterpretado como señal de voz muy seguido, la eficiencia del DTX es dramáticamente decrementada. Otro factor para considerarse que cuando el transmisor se apaga, se escucha un silencio total en el receptor, naturaleza digital de GSM. Para asegurar el receptor de que la conexión no se ha muerto, ruido aceptable es creado al final de la recepción por medio de tratar de emparejar las características del ruido de fondo transmitido al final.

3.4.7 Recepción Discontinua

Otro método utilizado para conservar la potencia en la estación móvil es la recepción discontinua. El canal de paginación utilizado por la estación base para señalar una llamada entrante, es estructurado dentro de sub-canales. Cada estación móvil requiere escuchar únicamente su propio sub-canal. En el tiempo entre sub-cnales de paginación sucesivos, el móvil puede pasar a modo de hibernación, donde casi no se usa potencia.

3.4.8 Control de Potencia

Existen cinco clases de estaciones móviles definidas, de acuerdo a la potencia pico del transmisor, clasificadas en 20, 8, 5, 2 y 0.8 watts. Para minimizar la interferencia co-canal y conservar la potencia, tanto los móviles como la BTS operan a un nivel de baja potencia que mantendrá una calidad de señal aceptable. Niveles de potencia pueden ser hacia arriba y abajo de 2 dB desde la potencia pico para mantener un mínimo de 20 mW.

La estación móvil mide la calidad de la señal (basada en Bit Error Ratio), y pasa la información a la BSC, el cual automáticamente decide cuando el nivel de potencia debe ser cambiado. El control de potencia debe ser manejado cuidadosamente, debido ha que existe la posibilidad de inestabilidad.

3.5 ASPECTOS DE LA RED

Asegurando la transmisión de voz o datos de una calidad dada encima del enlace de radio es sólo parte de la función de una red móvil celular. Un móvil GSM puede funcionar transparentemente nacionalmente e internacionalmente, el cual requiere que el registro, autenticación, redireccionamiento de llamadas y funciones de actualización de ubicación. Además, el área geográfica cubierta por la red es dividida en células necesita la implementación de un mecanismo handover. Estas funciones son desempeñadas por el Subsistema de Red, principalmente utilizando Parte de Aplicación Móvil (MAP: Mobile Application Part) contruidoa en la parte superior del protocolo de Sistema de Señalización No7.

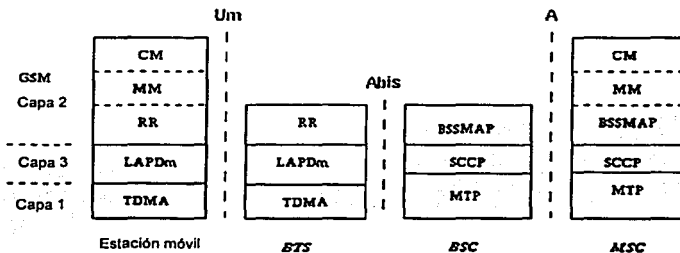


Figura 3.3. Estructura del protocolo de señalización de GSM

"El protocolo de señalización en GSM es estructurado dentro de tres capas generales, dependiendo de la interfase, como se muestra en la figura 3.3. La capa 1 es la capa física, la cual utiliza la estructura de canal sobre la interfase aérea. La capa 2 es la capa de enlace, a través de la interfase Um, la capa de enlace es una

versión modificada del protocolo LAPD utilizada en ISDN, llamada LAPDm. La capa 3 del protocolo de señalización GSM se divide en 3 subcapas, Administración de recursos de radio, Administración de movilidad y Administración de conexión. Señalización entre diferentes identidades en la parte fija de la red, como lo es entre el HLR y el VLR, es realizada a través del MAP. "[14]

3.5.1 Administración de recursos de radio

La capa de administración de recursos de radio (RR) vigila el establecimiento de un enlace, tanto de radio como fijo, entre la estación móvil y el MSC. Los componentes funcionales principales involucrados son la estación móvil, y el Subsistema de Estación Base, así como el MSC. La capa de RR se preocupa por la dirección de una sesión RR[16] que es el tiempo que un móvil está en modo dedicado, así como la configuración de canales de radio incluyendo la asignación de canales dedicados.

Una sesión RR siempre es comenzada por una estación móvil a través del procedimiento de acceso, o para una llamada saliente, o en respuesta a un mensaje de paginación. Los detalles del acceso y procedimientos de paginación, como cuando un canal dedicado se asigna al móvil, y la estructura de paginación de sub-canal, son manejados en la capa RR. Además, maneja la administración de características de radio como control de potencia, transmisión discontinua, recepción, y avance del cronometrado.

3.5.1.1 Handover

En una red celular, los enlaces de radio y fijos requeridos no son permanentemente asignados por duración de llamada. Handover y handoff como se conoce en Norte América, es la conmutación de una llamada saliente a un canal o célula diferente. La ejecución y medida requerida por el hand over desde una estación base funciona en la capa RR.

Hay cuatro tipos diferentes de handover en el sistema de GSM que involucra transferencia de una llamada entre:

- ↓ Canales (time slots) en la misma célula
- ↓ Células (BTS) bajo el control del mismo BSC
- ↓ Células bajo el control de diferentes BSCs, pero perteneciendo al mismo MSC
- ↓ Células bajo el control de diferentes MSCs

Los primeros dos tipos de handover, llamados handovers internos involucran únicamente una BSC. Para ahorrar ancho de banda de señalización, son administrados por el BSC sin involucrar al MSC, excepto para notificarle que se ha completado el handover. Los otros dos tipos de handover, llamados handovers externos, son manejados por el MSC involucrado, Un aspecto importante de GSM es que el MSC original, el anchor MSC, permanecen responsables de casi todas

las funciones de llamada, con la excepción de los subsecuentes handovers entre BSC bajo el control del nuevo MSC, llamado relay MSC.

Handovers pueden ser iniciados tanto por el móvil o por el MSC. Durante el idle de los time slots, el móvil escanea el canal de control de broadcast de arriba de 16 células vecinas, y forma una lista de las seis candidatas para un posible handover, basado en la señal recibida. Esta información es pasada al BSC y MSC, al menos una vez por segundo, y es utilizada por el algoritmo de handover.

El algoritmo para cuando una decisión de handover se debe tomar no está especificada en las recomendaciones GSM. Existen dos algoritmos básicos utilizados, ambos estrechamente ligados con el control de potencia. Esto es debido a que el BSC usualmente no conoce si la calidad de señal pobre en de alimentación multiruta o si el móvil tuvo que moverse a otra célula. Esto es especialmente cierto en células urbanas pequeñas.

3.5.2 Administración de Movilidad

La capa de Administración de Movilidad (MM) se construye en la parte superior de la capa de RR, y maneja las funciones que surgen de la movilidad del suscriptor, así como la autenticación y aspectos de seguridad. La administración de ubicación se preocupa por los procedimientos que permiten al sistema saber la situación actual de la estación móvil de tal manera que el redireccionamiento de llamadas entrantes pueda ser completado.

3.5.2.1 Actualización de Ubicación

Un móvil encendido es informado de una llamada entrante por un mensaje de paginación enviado sobre el canal PAGCH de una célula. Un extremo sería para paginar cada célula en la red para cada llamada, lo cual es obviamente una pérdida de ancho de banda. El otro extremo sería para el móvil para notificar el sistema, vía mensajes de actualización de ubicación para ser enviado exactamente a una célula, pero podría ser mucha pérdida por el gran número de mensajes de actualización. La solución utilizada por GSM es agrupar células en áreas de ubicación. Actualización de mensajes son requeridos cuando se mueve entre áreas de ubicación y las estaciones base son paginadas en células de su ubicación actual.

Los procedimientos de actualización de ubicación y el subsecuente redireccionamiento de llamada utiliza MSC y dos registros de ubicación HLR y VLR. Cuando una estación móvil es encendida en una nueva área de ubicación o que se mueve a una nueva área de ubicación o a diferentes operadores PLMN, se debe registrar con la red para indicar su ubicación actual.

Por razones de fiabilidad, GSM también tiene procedimientos de actualización de ubicación. Si un HLR o un MSC/VLR falla, para tener cada móvil registrado

simultáneamente para proporcionar la actualización de la base de datos puede causar una sobrecarga. Por lo que, la base de datos es actualizada como eventos de actualización de ubicación. La habilitación de actualización periódica y el periodo de tiempo entre actualizaciones periódicas es controlada por el operador. Si un móvil no se registra después del periodo de actualización, pierde el registro.

3.5.2.2 Autenticación y seguridad

A partir de que el medio de radio puede ser accedido por cualquiera, autenticación de usuarios para probar que ellos son los que deben ser, es un elemento importante para una red móvil.

La autenticación involucra dos identidades funcionales, el SIM y el Centro de Autentificación (AuC). A cada suscriptor se le da una llave se crea, una copia de la cual esta almacenada en el SIM y la otra en el AuC. Durante el proceso de autenticación, el AuC genera un número aleatorio que es enviado al móvil. Tanto el móvil y el AuC entonces utilizar un número aleatorio, en conjunto con la llave secreta del suscriptor y el algoritmo de cifrado llamado A3, para generar respuesta firmada (SRES) que es enviada de regreso al AuC. Si el número enviado por el móvil es el mismo que el calculado por el AuC, el suscriptor es autenticado [16]

El mismo número aleatorio inicial y la llave suscrita también son utilizadas para calcular la llave de cifrado utilizando el algoritmo llamado A8. Esta llave de cifrado, en conjunto con el número de trama TDMA, utiliza el algoritmo A5 para crear una secuencia de 114 bits es XORed con los 114 bits del estallido (los dos bloques de 57 bits).

Otro nivel de seguridad es desempeñado por el mismo equipo móvil, en oposición al suscriptor móvil. Como se mencionó antes, cada terminal GSM es identificada por un número IMEI único. Una lista de IMEIs en la red es almacenado en el EIR. El estatus regresado en respuesta a la petición del IMEI al EIR es uno de los siguientes:

Listado-blanco: la terminal es permitida para conectarse con la red.

Listado-gris: la terminal es bajo observación desde la red por posibles problemas.

Listado-negro: la terminal ha sido reportada como robada o no es del tipo aprobado. La terminal no es permitida para conectarse a la red.

3.5.3 ADMINISTRACIÓN DE COMUNICACIÓN

La capa de Administración de Comunicación (CM) es responsable del control de llamada (CC), servicio de administración suplementario, y administración de servicios de mensajes cortos. Cada un de estos puede ser considerado como una subcapa separada dentro de la capa CM. El control de llamadas intenta seguir los procedimientos ISDN especificados en Q.931, a pesar de que enrutar a un suscriptor móvil foráneo (roaming) es único de GSM. Otras funciones de la subcapa

CC incluye establecimiento de llamadas, selección del tipo de servicio y liberación de llamadas.

3.5.3.1 Direccionamiento de Llamadas

A diferencia del direccionamiento de llamadas en redes fijas, donde las terminales están cableadas casi permanentemente a una oficina central, un usuario GSM puede tener comunicación a nivel nacional como internacional. El número de directorio marcado para alcanzar al suscriptor móvil es llamado Suscriptor móvil ISDN (MSISDN: Mobile Subscriber ISDN), el cual es definido por el plan E.164. Este número incluye código de ciudad y código de destino nacional, el cual identifica al operador del suscriptor. Los primeros dígitos de número del suscriptor puede identificar el HLR del suscriptor dentro del PLMN local.

La terminación entrante de una llamada móvil es dirigida a la función Gateway MSC (GMSC). El GMSC es básicamente un interruptor el cual es capaz de interrogar al HLR del suscriptor para obtener la información de direccionamiento, y este contiene una tabla que enlaza MSISDN a su HLR correspondiente. Una simplificación es tener el manejo de un GSMC en un PLMN específico. Se debe notar que la función GMSC es distinta desde la función MSC, pero es usualmente implementada en un MSC.

La información de direccionamiento que es regresada al GMSC es el MSRN (Mobile Station Roaming Number), que también esta definida por el plan de numeración E.164. MSRN's son relacionados con el plan de numeración geográfica y no asignada a suscriptores y no es visible para suscriptores. En la figura 4 se muestra el esquema completo de direccionamiento de llamadas.

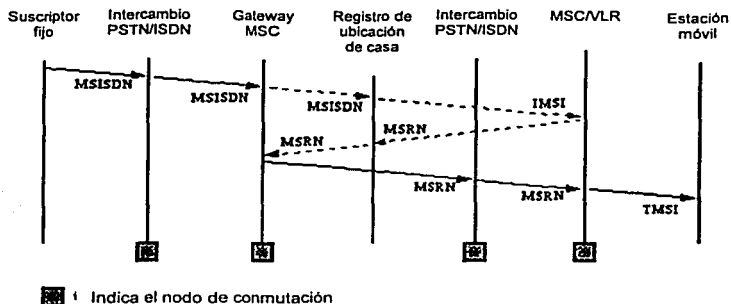


Figura 3.4. Direccionamiento de una llamada a una terminal móvil.

PARTE II

TERCERA GENERACIÓN DE COMUNICACIONES MÓVILES

CAPÍTULO 4

TELEFONÍA MÓVIL

4.1 INTRODUCCIÓN

Martin Cooper fue el pionero en esta tecnología, a él se le considera como el padre de la telefonía celular al introducir el primer radioteléfono en 1973 en los Estados Unidos mientras trabajaba para Motorola; pero no fue hasta 1979 en que aparece el primer sistema comercial en Tokio Japón por la compañía NTT (Nipon Telegraph & Telephone Corp.)

En 1981 en los países Nórdicos se introduce un sistema celular similar al sistema avanzado de teléfono móvil (AMPS: Advanced Mobile Phone System). Por otro lado, en los Estados Unidos gracias a que la entidad reguladora de ese país adopta reglas para la creación de un servicio comercial de telefonía celular, en octubre de 1983 se pone en operación el primer sistema comercial en la ciudad de Chicago. A partir de entonces en varios países se diseminó la telefonía celular como una alternativa a la telefonía convencional alámbrica. La tecnología inalámbrica tuvo gran aceptación, por lo que a los pocos años de implantarse se empezó a saturar el servicio, por lo que hubo la necesidad de desarrollar e implementar otras formas de acceso múltiple al canal y transformar los sistemas analógicos a digitales para darle cabida a más usuarios. Para separar una etapa de la otra, a la telefonía celular se ha categorizado por generaciones.

Las tecnologías inalámbricas están teniendo mucho auge y desarrollo en estos últimos años, una de las que ha tenido un gran desarrollo ha sido la telefonía celular, desde sus inicios a finales de los 70s ha revolucionado enormemente las actividades que realizamos diariamente. Los teléfonos celulares se han convertido en una herramienta primordial para la gente común y de negocios, las hace sentir más seguras y las hace más productivas.

A pesar que la telefonía celular fue concebida para la voz únicamente, debido a las limitaciones tecnológicas de esa época, la tecnología celular de hoy en día es capaz de brindar otro tipo de servicios tales como datos, audio y video con algunas limitaciones, pero la telefonía inalámbrica del mañana hará posible aplicaciones que requieran un mayor consumo de ancho de banda.

"Actualmente, la telefonía móvil e Internet captan el mayor interés dentro del mundo de las telecomunicaciones y la informática, y prueba de ello es el crecimiento experimentado en el número de usuarios que optan por utilizar estos dos servicios como se muestra en la figura 4.1. Así, Internet crece aun ritmo superior al 100% anual mientras que la telefonía móvil lo hace a un ritmo entre el 40-60%, cifras espectaculares frente al crecimiento de la telefonía fija que no supera en los países mas avanzados el 5 o 10%. A finales del año 2000 existían 650 millones de usuarios de telefonía móvil en todo el mundo y se prevé alcanzar los mil millones en el 2003. " [17]

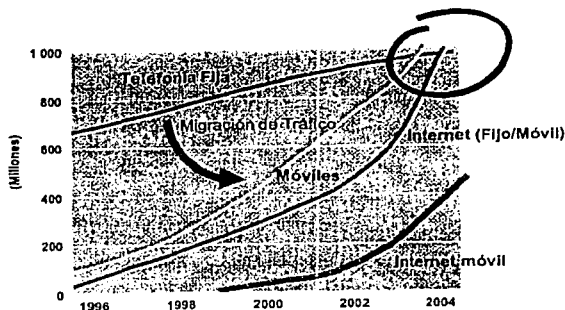


Figura 4.1. Crecimiento de usuarios

Pero este crecimiento, hasta el momento, no ha satisfecho las expectativas que se tenían, principalmente por las limitaciones de las redes de comunicaciones y la ausencia de servicios de acceso móvil. Por un lado, los sistemas de comunicaciones móviles de primera generación (analógicos) y segunda generación (digitales) no estaban concebidos para funcionar a nivel mundial sino más bien a nivel nacional, y en el mejor de los casos, regional. A esto se suma la gran variedad de estándares de comunicaciones móviles existentes hoy día, distribuidos por todo el mundo como GSM, CDMA, AMPS, TDMA y DECT. Estas tecnologías no son compatibles entre sí, por lo que la utilización de una única terminal en estas redes es imposible, y así, el soñado roaming internacional no es posible en su totalidad.

En este momento, las comunicaciones móviles se encuentran en un punto de inflexión, donde la movilidad a escala internacional, el acceso a Internet y las aplicaciones multimedia sobre terminales móviles cobran una gran importancia. Prueba de ello son las cifras bilionarias que se han pagado en algunos países de Europa por las nuevas licencias y el interés mostrado por la industria y grupos financieros en participar en este prometedor negocio que aunque tardará unos pocos años en desarrollarse, es el futuro a mediano plazo.

4.2 GENERACIONES DE LA TELEFONÍA INALÁMBRICA

La primera generación (1G) de telefonía móvil hizo su aparición en 1979, se caracterizó por ser analógica y estrictamente para voz. La calidad de los enlaces de voz era muy baja, baja velocidad (2400 bauds), transferencia de celdas muy imprecisa, baja capacidad (basada en FDMA), no tenía seguridad y los teléfonos estaban diseñados para uso en vehículos. AMPS fue el principal estándar de la primera generación y se desarrolló entre 1982 y 1992. El sistema analógico empleado todavía en Europa, el TACS (Total Access Communications System), se basa en el AMPS.

La segunda generación 2G arribó hasta 1990 y a diferencia de la primera se caracterizó por ser digital. El sistema 2G utiliza protocolos de codificación más sofisticados y son los sistemas de telefonía celular usados en la actualidad como GSM (apoyado por la ETSI), cdmaOne (apoyado por la ANSI), TDMA también llamada D-AMPS por ser la versión digital de AMPS y PDC (Personal Digital Communication) utilizado en Japón.

Los protocolos empleados en los sistemas 2G soportan velocidades de información más altas para voz pero limitados en comunicaciones de datos. Se pueden ofrecer servicios auxiliares tales como datos, fax y SMS (Short Message Service). La mayoría de los protocolos de 2G ofrecen diferentes niveles de encriptación. En los Estados Unidos y otros países se le conoce a 2G como PCS (Personal Communication Services).

A la generación intermedia entre la 2G y la 3G se le denomina 2.5G y corresponde a mejoras tecnológicas en las redes 2G actuales, muchos de los proveedores de servicios de telecomunicaciones (carriers) se moverán a las redes 2.5G antes de entrar masivamente a 3G. La tecnología 2.5G es más rápida y más económica para actualizar a 3G.

La generación 2.5G ofrece características extendidas para ofrecer capacidades adicionales que los sistemas 2G tales como GPRS, HSCSD, EDGE, IS-136B, IS95B, entre otros. Los carriers Europeos y de Estados Unidos se han estado moviendo a 2.5G a los largo del 2001.

La tercera generación (3G) es tipificada por la convergencia de la voz y datos con acceso inalámbrico a Internet, aplicaciones multimedia y altas transmisiones de datos. Los protocolos empleados en los sistemas 3G soportan más altas velocidades de información enfocados para aplicaciones más allá de la voz tales como audio (MP3), video en movimiento, videoconferencia y acceso rápidos a Internet, sólo por nombrar algunos. Las redes empezaron a operar en Japón por NTT DoCoMo en el 2001 y en Europa y parte de Asia han empezado a operar en el 2002, posteriormente Estados Unidos y otros países.

Los sistemas 3G alcanzarán velocidades de hasta 384 Kbps permitiendo una movilidad total a usuarios viajando a 120 km/h en ambientes exteriores y alcanzará una velocidad máxima de 2 Mbps permitiendo una movilidad limitada a usuarios caminando a menos de 10 km/h en ambientes estacionarios de corto alcance o en interiores. Entre las tecnologías contendientes de la tercera generación se encuentran UMTS (Universal Mobile Telephone Service), cdma2000, IMT-2000, ARIB[3gpp], UWC-136, entre otras.

El impulso de los estándares de la 3G está siendo apoyado por la ITU (International Telecommunications Union) y a este esfuerzo se le conoce como IMT-2000 (International Mobile Telephone).

La cuarta generación 4G es un proyecto a largo plazo que será 50 veces más rápida en velocidad que la tercera generación. Se planean hacer pruebas de esta tecnología hasta el 2005 y se espera que se empiecen a comercializar la mayoría de los servicios hasta el 2010.

4.3 VÍAS DE EVOLUCIÓN HACIA IMT-2000

Con el fin de aprovechar al máximo los beneficios de las inversiones hechas en los sistemas móviles actuales, es conveniente determinar la manera en la que estos sistemas pueden evolucionar hacia IMT-2000. Esto facilitará también la introducción de IMT-2000 y permitirá un mayor grado de reutilización de la infraestructura de redes existentes.

Este aspecto se está considerando sobre la base de que los sistemas previos a IMT-2000 pueden poseer ya algunas características y admitir desarrollos posteriores que permitan su evolución hacia IMT-2000. Cabe señalar también que este enfoque puede ser el más apropiado para los sistemas que funcionan en bandas de frecuencia próximas a las bandas identificadas para IMT-2000, como es el caso del sistema GSM.

En este sentido, la industria global de telecomunicaciones ha reducido en general el número de normas de tercera generación, respetando al mismo tiempo las normas existentes. Así, se han logrado dos hechos importantes: la convergencia de TDMA/136 y GSM, y la convergencia de modos CDMA.

La convergencia de TDMA/136 y GSM comienza con GPRS, que crea una arquitectura de la red común y comparte componentes de red de radio y terminales, por lo que su introducción en las redes resulta fácil y económica. La convergencia de modos CDMA crea una sola familia de acceso radio CDMA de tercera generación: DS-WCDMA y MC-WCDMA en modo FDD.

Los modos WCDMA de secuencia directa DS-WCDMA (también llamados de espectro expandido) son los modos principales aceptados por la UIT para UMTS. El modo multiportadora (MC-WCDMA) es sobre todo para la evolución de cdmaOne/cdma2000.

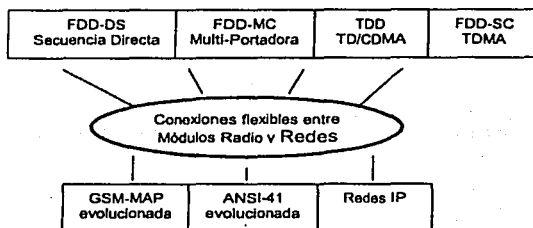


Figura 4.2. Convergencia de estándares de comunicación

Para que puedan trabajar juntas las normas de segunda generación con las de tercera generación se debe dar funcionalidad de interoperabilidad al nivel de red (protocolos) y al nivel terminal (terminales multi-entorno y multi-modo).

4.3.1 Evolución de GSM a UMTS

Existen básicamente cuatro sistemas de transmisión que juegan el papel de evolución desde GSM hacia la tercera generación de comunicaciones móviles (3G) que son:



Requerimientos	HSCSD	GPRS	EDGE	UMTS
Microteléfono	No tiene capacidad para paquetes de datos	Microteléfonos permitidos en redes GPRS habilitadas y a 9.6kbps en redes GSM utilizando teléfonos en modo CSD-dual	Microteléfonos trabajarán a 384kbps en redes habilitadas para EDGE, en redes habilitadas para GPRS y a 9.6kbps en redes GSM utilizando teléfonos en modo CSD-tri	Nuevos microteléfonos 3G trabajarán únicamente en estaciones base.

Infraestructura	No tiene capacidad para paquetes de datos	Nuevo backbone de paquetes	Modificaciones del backbone	Modificaciones del backbone
-----------------	---	----------------------------	-----------------------------	-----------------------------

La ruta de evolución de datos GSM siempre requiere nueva infraestructura de red y nuevos teléfonos. "Cada uno de los servicios de datos GSM desde HSCSD, WAP, GPRS, EDGE y 3G requieren de nuevos microteléfonos. Los microteléfonos de tercera generación, no trabajarán en estaciones base EDGE o WCDMA. De cualquier forma, las terminales multibanda GSM/3G, GSM/GPRS, GSM/EDGE estarán disponibles. Por el lado de la infraestructura, los operadores de red GSM deben hacer nuevas inversiones en estaciones base para GPRS, EDGE y 3G. Una vez que el backbone de GPRS se implementado, la evolución de 3G requiere únicamente mejoras con rutas de actualización de definidas ya desarrolladas por los vendedores de infraestructura." [18]

La evolución de las redes GSM hacia UMTS se basa en tres tecnologías

4.3.1.1 HSCSD

El sistema de conmutación de circuitos de datos a alta velocidad (HSCSD: High Speed Circuit Switched data) maneja múltiples frecuencias o canales de radio para el establecimiento del estándar de radio móvil GSM programado para comunicación de voz.

La GSM soporta a un usuario por canal, por ranura de tiempo. La HSCSD, proporciona a un usuario independiente acceso simultáneo a canales múltiples (hasta cuatro) al mismo tiempo. Como tal, existe un intercambio directo entre una velocidad mayor y el costo asociado de utilizar más recursos de radio, ya que es muy caro para los usuarios finales pagar por llamadas múltiples y simultáneas.

Asumiendo un rango estándar de transmisión de 14.4 kbps de HSCSD, utilizando cuatro ranuras de tiempo con HSCSD, se permiten velocidades teóricas de hasta 57.6 kbps. Esto es ampliamente equivalente a proporcionar el mismo rango de transmisión que el disponible en más de un canal ISDN B. Algunos Centros Móviles de Enlace (MSCs) están limitados hasta 64 kbps como entrada máxima, esta restricción se elimina con GPRS.

En redes donde la HSCSD es desplegada, el GPRS sólo deberá asignarse con una prioridad de tercer grado, después de la voz que tiene una prioridad en primer grado y la HSCSD con un segundo grado. En teoría la HSCSD puede ser adelantada por llamadas de voz ya que las llamadas HSCSD pueden reducirse a un canal si las llamadas de voz buscan ocupar estos canales. La HSCSD no interrumpe el servicio de disponibilidad de voz, pero si afecta al GPRS. Aún siendo adelantado, es difícil ver que la HSCSD pueda ser desplegada en redes ocupadas y conferir una experiencia agradable al usuario, es decir, un rango de

alta información continua. Por lo tanto parece ser que la HSCSD debe desplegarse en redes que comienzan o aquellas que tengan suficiente capacidad desde que relativamente no es tan costoso desplegarla y que puede transformar algunos canales disponibles en corrientes rentables. La HSCSD es más fácil de implementar en redes móviles que en GPRS porque algunas de las soluciones de venta de GSM requieren únicamente una actualización del software de estaciones base y no hardware nuevo. Este no es el caso con redes D-AMPS y algunas soluciones de venta de GSM.

Existe un par de razones por las cuales la HSCSD debiera ser el portador preferido para ciertas aplicaciones cuando es comparado con GPRS. El hecho de que paquetes asociados pueden ser enviados en diferentes direcciones para llegar al mismo destino, en teoría debería hacer que la transmisión fuera más robusta ya que existen varias maneras de alcanzar el resultado final. Sin embargo, esta naturaleza de transmisión en paquete significa que los paquetes están sujetos a retraso variable y algunos pudieran perderse. Mientras que la retransmisión en paquete se incorpora a los estándares GPRS, naturalmente este proceso tomará tiempo y en el caso de aplicaciones tales como vídeo transmisión podrá causar imágenes con baja calidad.

Otra aplicación preferida para HSCSD podría ser el hecho de que mientras GPRS es complementario para comunicarse con otras redes basadas en paquete, tales como Internet, la HSCSD podría ser la mejor forma de comunicarse con otros medios de comunicación de circuito abierto, tales como PSTN y ISDN.

Una dificultad técnica potencial con HSCSD surge porque en un ambiente de múltiples ranuras de tiempo (timeslots), la transferencia de la llamada dinámica entre diferentes celdas en una red móvil ("handover") se complica, a no ser que las mismas ranuras estén disponibles de fin a fin a través de la duración de la llamada CSD.

"ESTI publicó la especificación para HSCSD como parte del estándar GSM Fase 2+ (como GSM 03.34), casi un año antes que el GPRS. A pesar de esto, los vendedores de infraestructura celular tales como Nokia y Ericsson han mostrado mayores demostraciones del servicio GPRS que HSCSD. A finales de 1998 y principios de 1999, casi una docena de operadores de red enviaron órdenes para HSCSD a vendedores tales como Nokia y Ericsson. Sin embargo, los servicios no comerciales de HSCSD han sido lanzados comercialmente desde mediados de 1999" [19].

En mayo de 1999, Ericsson anunció que no tenía planes para fabricar terminales que soportaran HSCSD. Por el contrario, se estaba enfocando en dispositivos GPRS. En junio de 1999, Nokia anunció su Tarjeta Telefónica 2.0, consistente en una tarjeta GSM para PC con soporte telefónico reconstruido para uso simultáneo de hasta cuatro canales CSD y que permitía la transmisión de información hasta 43.2 Kbps sin compresión de información. La Tarjeta Telefónica de Nokia 2.0 se

diseñó para ser comercialmente disponible a finales de 1999. El teléfono por medio de tarjeta de datos pudiera ser el factor que mejor forma para una terminal HSCSD, ya que el voltaje que consume ranuras múltiples es significativo y por lo menos la tarjeta telefónica puede derivar su fuente de poder desde una laptop a la cual está típicamente conectada.

En resumen, se puede decir que los vendedores de infraestructura celular no están tan comprometidos con HSCSD como lo están con GPRS.

4.3.1.2 GPRS

El servicio general de paquetes de radio (GPRS: General Paquet Radio Service) permite transmisión de datos por conmutación de paquetes en lugar de por conmutación de circuitos a alta velocidad en radio móvil GSM. GPRS es una expansión de GSM.

La red de acceso de radio (RAN: Radio Access Network) permanece idéntico, excepto por la actualización de software y nuevos módulos de hardware para comunicaciones orientadas a paquetes. Un paso de desarrollo significativo hacia 3G, por lo cual GPRS se clasifica como tecnología 2.5G. En lugar de tener una conexión permanente virtual y relativamente cara, los contenidos digitales (texto, imagen, software, etc.) son convertidos en pequeños paquetes de datos comprimidos y codificados. Estos paquetes de datos entonces son enviados al receptor.

Una distinción es hecha entre diferentes niveles de calidad de servicio (QoS: Quality of Service), cumplen con diferentes requerimientos y situaciones de uso. Menos aplicaciones críticas en tiempo como descargar archivos multimedia no requeridos inmediatamente se pueden administrar lo suficientemente bien con bajo nivel QoS que las llamadas para situaciones donde los datos debes ser intercambiados en tiempo real, sin retraso. Con el estándar GPRS también los proveedores de radio deben decidir que tanta capacidad hacer disponible para servicios de datos. El uso excesivo de datos fácilmente puede sobrecargar la red GSM y ocasionar cuellos de botella en comunicaciones de voz.

Los GPRS facilitan las conexiones instantáneas mientras que la información puede ser enviada o recibida inmediatamente como sea necesario. La conexiones por módem no son necesarios. Esto es porque los usuarios GPRS algunas veces se refieren al término siempre conectado. La disponibilidad es una de la ventajas de los GPRS cuando se comparan con los datos por conmutación de circuitos. Alta disponibilidad es una característica muy importante para aplicaciones críticas como los es la autorización remota de una tarjeta de crédito donde podría no ser aceptada y mantener al proveedor esperando tiempo extra.

Velocidad

"Con el estándar GPRS, el envío ya no esta basado en la longitud del tiempo de transmisión, en su lugar esta basado en el volumen de intercambio de datos o el

tipo de servicio. Esto significa que las terminales siempre pueden estar encendidas y despliegan los mensajes entrantes sin retraso. Dependiendo de la codificación tasas de transmisión mas elevadas pueden ser desarrolladas por ranura de tiempo, velocidades teóricas de hasta 171.2 Kbps son desarrolladas con GPRS utilizando las ocho ranuras de tiempo al mismo tiempo". [20] Esto es acerca de tres veces más rápido que las velocidades posibles actualmente en redes de comunicaciones mixtas y diez veces más rápido que los servicios de datos por conmutación de circuitos en redes GSM.

Acceso a los servicios

Para utilizar GPRS, los usuarios requieren específicamente:

- ✚ Un teléfono móvil o terminal que soporte GPRS (los teléfonos GSM existentes no soportan GPRS)
- ✚ Una suscripción a un red de telefonía móvil que soporte GPRS
- ✚ El uso de GPRS debe ser habilitado para dicho usuario. Acceso automático a los GPRS debe ser permitido por algún operador de red móvil.
- ✚ Conocimiento de como enviar y recibir información GPRS utilizando el modelo específico de teléfono móvil, incluyendo configuración de software y hardware, lo cual crea el requerimiento de un proveedor de servicio.
- ✚ Un destino para enviar o recibir información a través de GPRS. Se debe tener una dirección de Internet, a partir de que GPRS esta diseñado para hacer totalmente disponible a Internet para usuarios móviles.

4.3.1.3 EDGE

Tasa de datos mejorada para evolución GSM (EDGE: Enhanced Data Rates from GSM) es otro estándar de datos móviles de alta velocidad. Permite tasas de transmisión más rápidas y es también otro paso intermedio desde GSM hacia la tecnología UMTS. "Las mejoras de desempeño en términos de tasas de transmisión es desarrollada por medio de un nuevo sistema de modulación. Cuando este tipo de modulación es también aplicado en combinación con técnicas de codificación es posible desarrollar tasas de transmisión de hasta 384kbps. EDGE fue formalmente llamado GSM384. Esto significa una tasa binaria de 48Kbps por ranura de tiempo. De cualquier forma esta tasa máxima puede únicamente ser ofrecida sobre un periodo corto de tiempo." [21]

EDGE fue desarrollado inicialmente por Ericsson para operadores de red móvil que no obtuvieron el espectro UMTS. EDGE ofrece a los operadores GSM la oportunidad de ofrecer servicios a velocidades que están cerca de aquellos disponibles en redes UMTS.

EDGE también proporciona una ruta de evolución desde GPRS a UMTS para la implementación de los cambios de modulación que serán necesarios para la implementación UMTS. La idea detrás de EDGE es de tasas de datos más elevadas en la actual portadora GSM de 200kHz por el cambio del tipo de modulación utilizado.

HSCSD y GPRS están basados en algo conocido como GMSK (Gaussian minimum-shift keying), el cual cubre únicamente un incremento moderado de la tasa de datos binarios por ranura de tiempo. Es el esquema de modulación que utiliza GPRS que limita su velocidad de datos a 115.2kbps.

Por otro lado EDGE está basado en un nuevo esquema de modulación que permite una tasa binaria mayor a través de una interfaz aérea. Esta técnica de modulación es conocida como 8PSK (eight-phase-shift keying). Este automáticamente adapta a circunstancias de radio y por lo tanto ofrece su tasas más altas en buenas condiciones de propagación cerca del site de la estación bases. Este cambio de modulación desde GMSK a 8 PSK es el cambio central con EDGE, el cual prepara el mundo GSM (TDMA en general) para UMTS.

4.3.2 Evolución de CdmaOne a Cdma2000

En el caso de las redes basadas en cdmaOne (IS-95A) de banda estrecha, existentes en los Estados Unidos y otros países de su área de influencia, la transición hacia IMT-2000 consta de los siguientes pasos.



Requerimientos	IS-95A	IS-95B	IMT-2000 CDMA Multi-carrier 1X (MC 1X)	IMT-2000 CDMA Multi-carrier 3X (MC 3X)
Microteléfono	Microteléfonos del estándar 95B trabajarán en las redes 95B, 1X y 3X a 14.4Kbps en teléfonos en modo individual	Microteléfonos trabajarán en redes 95A a 14.4Kbps y 95B, 1X y sistemas 3X a 64kbps en teléfonos en modo individual	Microteléfonos 1X trabajarán en redes 95 ^a a 14.4kbps, redes 95B a 64Kbps y redes 1X y 3X a 144kbps en teléfonos en modo individual	Nuevos microteléfonos 3X trabajarán en redes 95 ^a a 14.4Kbps, redes 95B a 64kbps, redes 1X a 144Kbps y redes 3X a 2Mbps en teléfonos en modo individual
Infraestructura	Estándar	Nuevo software en controladores de estación base	1X requiere nuevo software en backbone y nuevas tarjetas de canal en estaciones base	Modificaciones del backbone y nuevas tarjetas de canal en estaciones base

4.3.2.1 IS-95B

La norma IS-95B mejora las velocidades de 64 a 115kbps agregando a los 8 canales de tráfico CDMA 14.4kbps y asignándolo a un móvil el tiempo que dure su

operación. Conveniente para acceso a Internet y aplicaciones que requieran velocidades media, particularmente en áreas de bajo tráfico (suburbano/rural).

4.3.2.2 Cdma2000 1X

"Cdma2000 1X también conocida como 1XRTT, IS-95C o Cdma2000-fase-1, emplea un canal de 1.25MHz de ancho de banda y ofrece una velocidad nominal de 144kbps para aplicaciones móviles y estacionarias, conveniente para requerimientos superiores en áreas de alto tráfico, pero no llega a soportar los servicios 3G" [23]. También se cuenta con otra fase intermedia llamada Cdma2000 1xEV.

El estándar IS-2000 (Cdma2000 1X) fue publicado por la Asociación de Industrias de Telecomunicaciones (TIA: Telecommunications Industries Association). 1X ofrece aproximadamente el doble de capacidad para voz que cdmaOne. Las transmisiones de datos promedio son de 144kbps.

1X se refiere a la implementación de cdma2000 dentro del espectro existente para las portadoras de 1.25MHz de cdmaOne. El término técnico se deriva de $N=1$, es decir, el uso de la misma portadora de 1.25MHz de cdmaOne y el 1X significa una vez 1.25MHz.

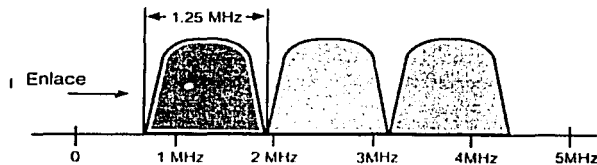


Figura 4.3. Espectro de frecuencias

Cdma2000 1X puede ser implementado en un espectro existente o en un nuevo espectro designado.

4.3.2.3 Cdma2000 1xEV

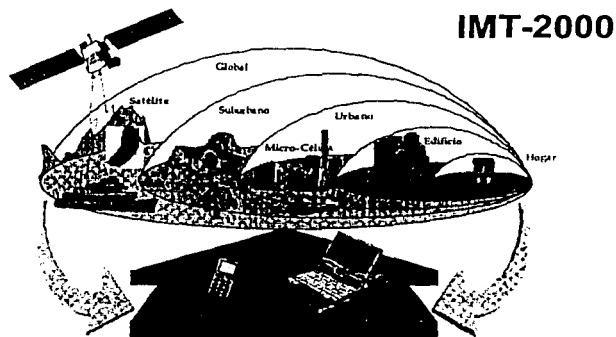
La evolución de cdma2000 más allá de 1X es ahora llamado cdma2000 1xEV. 1xEV será dividido en dos pasos: 1xEV-DO y 1xEV-DV. 1xEV-DO significa 1X únicamente evolución de datos (Evolution Data Only). 1xEV-DV significa evolución de voz y datos 1X (Evolution Data and Voice). Ambas proveen pasos para proveer servicios avanzados cdma2000 utilizando el estándar de portadora de 1.25MHz.

1xEV-DO estará disponible para los operadores durante el 2002. 1xEV-DV estará disponible uno y medio o dos años después que 1xEV-DO.

4.4 ESTÁNDARES DE COMUNICACIÓN IMT-2000

IMT-2000 es un estándar de la ITU definido por un grupo de recomendaciones de la serie M, F, G y Q. Agrupa una familia de sistemas con capacidades y servicios 3G cuya puesta en servicio en la Unión Europea y otros países, como es Japón, está prevista para finales del 2001 y principios del 2002, sujeta a consideraciones de mercado y disponibilidad de terminales en cantidades masivas para atender la demanda.

IMT-2000 entendida sobre la base de sistema de tercera generación y su futura evolución, viene a consolidar y unificar los diversos e incompatibles ambientes móviles de hoy a una infraestructura de red y radio capaz de ofrecer un amplio rango de servicios a escala global. Proporciona acceso a servicios de telecomunicaciones prestados por las redes fijas de telecomunicaciones (RTC y/o RDSI) y a otros servicios específicos de los usuarios móviles. IMT-2000 abarca una gama de servicios y terminales móviles, enlazados a redes terrenas o satelitales y las terminales pueden ser diseñadas para uso móvil o fijo, para ambientes tanto profesionales como domésticos, públicos o privados.



Para asegurar el éxito de los servicios 3G, se ha de proporcionar a los usuarios una comunicaciones muy eficientes, con una alta velocidad y calidad y, además, fáciles de utilizar. Los sistemas de 3G deben ofrecer:

- ✦ Transmisión simétrica/asimétrica de alta fiabilidad.
- ✦ Uso de ancho de banda dinámico, en función de la aplicación.
- ✦ Velocidades binarias mucho más altas: 144kbps en alta movilidad, 384kbps en espacios abiertos y 2Mbps en baja movilidad.
- ✦ Soporte tanto de conmutación de paquetes (IP) como de circuitos.

- ✚ Soporte IP para acceso a Internet (navegación www), videojuegos, comercio electrónico y video y audio en tiempo real.
- ✚ Diferentes servicios simultáneos en una sola conexión.
- ✚ Calidad de voz como en la red fija.
- ✚ Soporte radioelectrónico flexible, con utilización más eficaz del espectro, con bandas de frecuencias comunes con todo el mundo.
- ✚ Personalización de los servicios, según perfil de usuario.
- ✚ Servicios dependientes de la posición (localización) del usuario.
- ✚ Incorporación gradual en coexistencia con los sistemas actuales de 2G.
- ✚ Roaming, incluido el internacional, entre diferentes operadores y tipos de redes.
- ✚ Ambientes de funcionamiento marítimo, terrestre y aeronáutico.
- ✚ Capacidad de terminales telecargables, multibanda y multientorno.
- ✚ Economía de escala y un estándar global y abierto que cubra las necesidades de un mercado de masas.
- ✚ Provisión de un ambiente local virtual VHE: el usuario podrá recibir el mismo servicio independiente de su ubicación geográfica.

Existen razones evidentes que explican la necesidad de introducir la 3G: por una parte está la capacidad de las redes móviles actuales que permiten albergar un número limitado de usuarios, con un patrón de consumo similar al actual, y en cuanto se sobrepase la congestión de la red se manifiesta de manera insoportable para los usuarios; por otra parte, tenemos el incremento del tráfico motivado por la sustitución del tráfico fijo por el móvil, en cuanto el costo de las llamadas se reduzca y los hábitos de los usuarios se modifiquen, necesitándose entonces más espectro; y por último, por la aparición de nuevos servicios, muchos de ellos personalizados, donde la convergencia con Internet y el aumento de aplicaciones multimedia significará un aumento significativo de tráfico, tanto es así que los analistas estiman que este supondrá en torno a un 30% de tráfico adicional en tan solo dos o tres años.

4.4.1 Avances den la estandarización de los sistemas 3G

La ITU inició el camino de estandarización de IMT-2000 cuando estableció el Interim Working Party 8/13 en 1985 (en la ITU-R los estudios los realizó la Comisión de Estudios 8 y en la ITU-T la Comisión de Estudio 11). Esta iniciativa fue llamada Future Public Land Movil Telecommunications Systems (FPLMTS), que en 1996 fue denominada IMT-200, para hacerla más fácil de pronunciar, a la vez que hacía referencia al año en que se esperaba la definición final del estándar IMT-200.

En 1998 la ITU denominó RTT (Radio Transmisión Technology) a las tecnologías que harían de interfaz aire entre las estaciones base y las terminales móviles. En junio de ese año, la ITU había recibido 15 propuestas (10 terrestres y 5 satelitales), siendo evaluadas por grupos independientes especiales, los que presentaron los informes en septiembre del mismo año.

Organismo	Propuesta	Descripción
ARIB Japón TTA	WCDMA	CDMA de banda ancha
TR45.5 EUA	Cdma2000	CDMA de banda ancha (IS95)
ETSI SMG2 Europa	UTRA	Acceso Radio Terrestre UMTS
TIA TR 45.3 EUA	UWC-136	Comunicaciones Universales Inalámbricas
ETSI Europa	DECT	Comunicaciones Inalámbricas Digitales Mejoradas
CATT China	TD-SCDMA	CDMA síncrono por división de tiempo
TIA TR46.1 EUA	WIMS	CDMA inalámbrico de banda ancha para multimedia
T1P1-ATIS EUA	NA	CDMA norteamericano de banda ancha
TTA Corea del Sur	CDMA I	DS-CDMA síncrono multibanda
TTA Corea del Sur	CDMA II	DS-CDMA asíncrono
TTA Corea del Sur	SAT CDMA	49 satélites LEO a 2 km
ESA Europa	SW-CDMA	CDMA de banda ancha por satélite
ESA Europa	SW-CTDMA	CDMA/TDMA de banda ancha por satélite
ICO Global	ICO RTT	10 satélites MEO a 10 km
Comm. Inmarsat	Horizon	10 satélites Horizons

Las especificaciones técnicas de las RTT terrestres fueron aprobadas en la WRC-2000 y se definieron como sigue:

- ↘ IMT-2000 CDMA Direct Spread (UTRA W-CDMA)
- ↘ IMT-2000 CDMA Multi-Carrier (cdma2000)
- ↘ IMT-2000 CDMA TDD (UTRA TD-CDMA)
- ↘ IMT-2000 TDMA Single-Carrier (UWC-136)

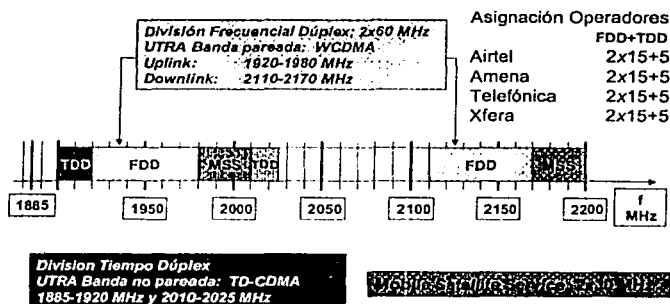
En la Conferencia Mundial de Radio WRC-92 se definió un rango de 230MHz de espectro radioeléctrico, sin asociarlo a ninguna tecnología, en las bandas de 1885-2025 MHz y 2110-2200 MHz identificadas para los servicios públicos IMT-2000, incluyendo las componentes por satélite y terrestres. Las bandas para el componente satélite de estos sistemas están comprendidas entre: 1980-2010 y 2200 MHz.

Esta decisión se basó en un modelo en el cual los servicios de telefonía se consideraban la fuente de tráfico principal (165 MHz), lo que se puede ver que es del todo insuficiente, dada la evolución de las aplicaciones en las redes móviles.

En particular, la aparición de Internet, Intranet, el correo y el comercio electrónico y los servicios de video ha elevado considerablemente las esperanzas del usuario en cuanto a la interoperabilidad de las de redes y las terminales y por ende el ancho de banda del canal móvil. Esta mayor demanda de servicios de banda ancha, junto con el aumento de la penetración de las comunicaciones móviles en la población mundial, creó la necesidad de un espectro adicional para IMT-2000,

que la ITU a través del Grupo de Trabajo 8/11 de la ITU-R calculó en 160 MHz sobre las bandas ya identificadas para la componente terrestre IMT-2000 en 1992. Finalmente, en la WRC-2000 celebrada en Estambul del 8 e Mayo al 2 de Junio del 2000, se aprobaron los requerimientos de espectro adicionales para IMT-2000. La decisión proporciona 3 bandas, como se ve en la figura. A continuación se describen las bandas.

- ↘ Banda 1 GHz (806-960 MHz)
- ↘ Banda 1.7 GHz (1710-1885 MHz). Frecuencia en la que funciona actualmente la mayoría de los sistemas de segunda generación, para facilitar la evolución con el tiempo de estos sistemas de tercera generación.
- ↘ Banda 2.5 GHz (2500-2690 MHz). Estas frecuencias compiten la banda de la gama 2 GHz ya identificada para IMT-2000.



4.4.2 Sistemas 3G/IMT-2000

Las distintas interfaces aire propuestas ante la Unión Internacional de Telecomunicaciones están basadas en CDMA que se acompañan de tres modalidades de operación, cada una de las cuales podría perfectamente funcionar sobre la red base de GSM (GSM-MAP) y sobre la red base de cdmaOne (IS-41).

Los distintos entes involucrados en los sistemas 3G han propuesto, básicamente, dos sistemas de tercera generación: UMTS y cdma2000.

4.4.2.1 UMTS

Las siglas UMTS son la abreviación de Sistema Universal de Telecomunicaciones Móviles (Universal Mobile Telecommunications System) y constituye la visión europea de sistemas con capacidades 3G como parte de la familia de estándares

IMT-2000. UMTS es la evolución lógica de la comunicad GSM a la tercera generación, por lo que está siendo mayoritariamente adoptado en la Unión Europea.

En enero de 1998, el Instituto de Estándares de Telecomunicaciones Europeo (ETSI: European Telecommunications Standards Institute) adoptó la tecnología W-CDMA (Widebanda CDMA) en modo FDD con provisión para TDD como la tecnología apropiada para Acceso de Radio Terrestre UMTS (UTRA: UMTS Terrestrial Radio Access), cada una diferente, pero basada en tecnologías similares. Es por esta razón que W-CDMA y UTRA se usan como términos intercambiables.

La combinación de los dos modos (FDD y TDD) ofrece la oportunidad de obtener la mayor eficiencia del mismo sistema bajo cualquier condición (urbana, suburbana, interiores y exteriores)

UMTS, en la componente terrestre, tiene una estructura jerárquica, esto es, está compuesta por tres tipos de celdas: Macro Celda, Micro Celda y Pico Celda con un mínimo de 5 MHz de ancho de banda por Celda.

La Macro Celda tiene radios desde 1 km hasta 35 km y se destinan para ofrecer cobertura rural y carreteras para vehículos u otros objetos que se mueven a alta velocidad (transmisión de datos de 114 kbps). La Micro Celda tiene radios desde 50 m hasta 1 km. Ofrecen servicios a usuarios fijos o que se mueven lentamente con elevada densidad de tráfico (urbana) con velocidades de 384 kbps. Las Pico Celdas tiene radios hasta a 50 m. Ofrecen coberturas localizadas en interiores, con velocidades del orden de los 2 Mbps.

4.4.2.2 Cdma2000

Cdma2000 es un sistema de banda ancha que perfecciona el actual estándar móvil digital de segunda generación cdmaOne (IS-95 CDMA). Permite al usuario obtener mayores velocidades de transmisión de datos y un uso más eficaz del espectro de radio que las técnicas de radio existentes en la actualidad.

Cdma2000 es una tecnología de interfaz de radio de banda ancha compatible con IMT-2000 que podrá utilizarse en todo el continente americano, China, Rusia y la región de Asia-Pacífico, áreas que poseen en este momento redes cdmaOne.

Operando en modo TDD y/o FDD, cdma2000 ofrece velocidades desde 1.2Kbps hasta 2Mbps y soporte para canales de 1.25-3.75, 7.5-11.25 y 15 MHz con una o múltiples portadoras.

Los 3.6864 Mcps (cps/chip determina el grado de ensanchamiento del espectro) de cdma2000 proporciona una capacidad superior al sistema en desarrollos de 10 y 20 MHz. No así cuando se trata de desarrollos de 5 MHz donde los 4.096 Mcps de W-CDMA proporcionan un mejor rendimiento. Cdma2000 además agrega una banda de guardia de 640KHz por lado para protección contra interferencia en canales adyacentes (interferencia co-canal). Cdma2000 opera con sincronismo entre el móvil y la estación base.

CAPÍTULO 5 ARQUITECTURAS DE TERCERA GENERACIÓN

5.1 TECNOLOGÍA DE TERCERA GENERACIÓN

5.1.1 Tasas de transmisión

La Unión Internacional de Telecomunicaciones ha establecido requerimientos mínimos de velocidad que el estándar IMT-2000 debe soportar. Estos requerimientos están definidos de acuerdo a la movilidad involucrada cuando se hacen llamadas. La tasa de transmisión estará disponible sobre 3G dependiendo del entorno en el cual se realiza la llamada:

5.1.1.1 Alta Movilidad

144 Kbps para usuarios móviles rurales. Esta tasa de transmisión esta disponible para entornos en los cuales los usuarios 3G están viajando a más de 120km/h en áreas abiertas, como lo pueden ser usuarios que están viajando a dichas velocidades en un tren y quieren utilizar sus terminales 3G.

5.1.1.2 Movilidad Total

385 Kbps para personas que viajan a una velocidad menor de 120km/h en entornos urbanos al aire libre.

5.1.1.3 Movilidad Limitada

Al menos 2 Mbps con baja movilidad a una velocidad menor de 10km/h en áreas cerradas y entornos al aire libre. A esta tasa de transmisión es a la que hacen referencia cuando se ilustra el potencial de la tecnología 3G.

5.1.2 Nodos de Red

Las redes 3G requerirán nuevos elementos de radio y núcleos de red:

5.1.2.1 Red de Radio

Se necesita una nueva interfaz aerea para 3G. Requerirá nuevos sistemas de Estación Base (BSSs). Específicamente, los cambios requeridos en los BSS son:

- ↓ El nuevo acceso por radio a redes 3G
- ↓ Comprenderá de un RNC (Radio Network Controller) y un Nodo B.

5.1.2.2 Controlador de Red de Radio (RNC)

El RNC reemplazará los Controladores de Estación Base. El RNC incluirá soporte para conexiones existentes y proveerá una eficiente conexión por paquetes con el dispositivo de paquetes núcleo de red (SSGN o equivalente). El RNC desempeña funciones de control de redes de radio que incluyen el establecimiento y fin de la llamada, manos libres, administración de recursos de red y control de energía.

5.1.2.3 Núcleo de la Red

El núcleo de la red 3G será una evolución de los GPRS o los núcleos de sistemas de red equivalentes a 2.5G. Los nodos de GPRS son SGSN (Serving GPRS Support Node) y GGSN (Gateway GPRS Support Node). Van a ser requeridas actualizaciones a móviles y sistemas transitorios de conmutación para liberación de paquetes.

Una nueva infraestructura de red para 3G son los MGW (Media Gateway) que es la frontera entre diferentes redes para procesar los datos de los usuarios finales como los son la codificación y decodificación de datos, convención de protocolos y calidad de servicio.

La capa de enlace también proporciona acceso a conmutadores centrales y redes que no son móviles como Televisión por cable. En soluciones de algunos fabricantes MGW estará controlado remotamente por medio de MSC (Mobile Switching Centre) y servidores GSN lo que nos lleva a un GPC (Gateway Control Protocol). El grupo de estudio ITU 16 y IETF Megaco H.248 están trabajando para asegurar que el Protocolo de Control de Gateway sea un protocolo estándar abierto.

Operadores de red existentes pueden entonces actualizar su Centro de Conmutación Móvil y GSNs para implementar 3G o para alternativamente implementar un MGW individual que es controlado desde parte del servidor de un nodo actualizado 2G.

5.1.2.4 Conexión Central de Red

La red de radio será conectada al núcleo de la red por medio de una conexión central de red, permitiendo acceso de banda ancha e interconexión para los suscriptores. La conexión central 3G puede utilizar cualquier tecnología de transporte pero debe estar basado en tecnologías por conmutación de paquetes como lo es ATM (Asynchronous Transfer Mode) e IP (Internet Protocol). La conexión de red central esta construida como una malla de ruteo de IP o nodos conmutados ATM interconectados por medio de enlaces punto a punto. Tecnologías como IP sobre ATM podría utilizar conmutación ATM para multiplexar el tráfico IP. La arquitectura IP sobre ATM soporta tráfico de voz sobre IP. Muchos fabricantes prefieren una red puramente IP mientras que otros prefieren un híbrido de ATM/IP para garantizar la calidad del servicio.

Alternativamente IP sobre SONET/SDH es una solución de conexión central que elimina la capa ATM por medio del establecimiento de enlaces punto a punto entre ruteadores de IP directamente sobre anillos SONET/SDH los cuales corren sobre DWDM (Dense Wavelength División Multiplexing) que habilita la transmisión a velocidades del orden de los Terabits por segundo (Tbit/s) en un ancho de banda agregado.

5.1.2.5 Cambios del Sistema

Por supuesto, plataformas y sistemas como lo son los centros de servicios de valor agregado, gateways, sistemas de facturación, sistemas de Red Inteligentes así como sistemas parecidos serán actualizados. Una vez más, esto es como una evolución desde los centros de servicios de datos como GPRS donde se introdujo todo nuevo como elementos de direccionamiento de paquetes.

También se requerirán cambios en el personal, especialistas en las aplicaciones, administradores de alianzas, administradores de sectores de Internet así como desarrolladores de aplicaciones sobre redes 3G.

5.1.3 Aplicaciones 3G Específicas

Existen gran cantidad de aplicaciones que se habilitarán debido al ancho de banda, el cual es de banda ancha que estará disponible en 3G. Estas aplicaciones incluyen:

5.1.3.1 Audio

Audio o video se podrá transferir, almacenar y ejecutar desde Internet o ejecutar las aplicaciones en el momento que se transfieren pero sin almacenarlas. Los contenidos serán transferidos utilizando diferentes tipos algoritmos de compresión como los que proporciona Microsoft o Real Networks o el protocolo MP3. MP3 es un algoritmo de compresión y descompresión el cual fue inventado en Alemania en 1987 y aprobado por el Grupo de Expertos en Movimiento de Imágenes como parte de la Organización Internacional de Estandarización en 1992.

En 3G, los archivos MP3 serán cargados sobre un medio aéreo directamente al teléfono vía un servidor dedicado. Existen numerosos modelos de negocios para permitir que tanto proveedores de red como propietarios de materiales MP3 se beneficien. Se espera que la integración de la telefonía móvil con productos que utilizamos día a día se extienda en los próximos cuatro años para proporcionarnos transmisión de datos, voz, video, música y acceso a Internet en cualquier momento, dentro de la siguiente generación de dispositivos móviles.

Los teléfonos con MP3 integrado de Samsung y los teléfonos con módulos MP3 de Ericsson han sido demostrados a finales de 1999 y se espera que sean comerciales en México dentro de unos años.

La era en la que se pueden cargar contenidos multimedia desde Internet sobre telecomunicaciones estacionarias y enlaces por cable para PCs es únicamente el inicio y es en gran medida dependiente del ancho de banda, la calidad de la imagen y disponibilidad del servicio es inversamente proporcional a la cantidad de ancho de banda disponible.

Podemos cargar desde Internet archivos de música o video a nuestro dispositivo móvil sobre redes estacionarias los cuales podrán ser vistos o escuchados

posteriormente, en este caso no tendríamos audio y video en tiempo real sobre las redes móviles.

Desde que los archivos de video con sonido tienen un gran tamaño, es necesaria una alta velocidad en los servicios móviles de datos para habilitar las aplicaciones móviles de audio. El gran ancho de banda es uno de los principales atractivos de 3G para aplicaciones móviles multimedia.

5.1.3.2 Voz sobre IP

Otra aplicación de audio para 3G es Voz sobre IP (VoIP), la habilidad para enrutar las llamadas telefónicas sobre Internet para proporcionar servicios de telefonía de voz como llamadas locales desde cualquier parte del mundo. Con 3G y tecnologías 2.5G con altas tasas de transmisión como lo es EDGE, VoIP estarán disponibles por primera vez en los teléfonos móviles. Para hacer una llamada de voz, Voz sobre IP puede pasar a ser de un servicio alternativo a un servicio regular.

De cualquier forma VoIP no reemplazara a los servicios estándar de voz debido a que los servicios VoIP demandan un gran ancho de banda y se requiere conmutación a alta velocidad para minimizar la pérdida de paquetes.

5.1.3.3 Imágenes Fijas

Las imágenes fijas como los son fotografías, dibujos, cartas, tarjetas postales, tarjetas de felicitación, presentaciones y paginas web estáticas serán enviadas y recibidas sobre redes móviles así como se hace a través de redes de telefonía estacionarias.

Existen dos variables que afectan el uso de estas aplicaciones, el ancho de banda y el tiempo. La velocidad, el ancho de banda y poco tiempo es requerido para transmitir imágenes. Esta es la razón por la cual la transmisión basada en imágenes en lugar de información textual no es una aplicación móvil tan popular, se toma mucho tiempo en efectuarse la transmisión debido a la velocidad que se tiene disponible.

Una vez capturadas las imágenes, pueden ser enviadas directamente a sitios de Internet permitiendo publicaciones en tiempo real. El tamaño del archivo para dibujos depende de la resolución y tipo de compresión. Típicamente, cada dibujo esta entre los 50K y 100K en el formato JPEG. Esta puede ser transmitida rápidamente utilizando datos de paquetes móviles.

La transmisión de imágenes fijas son aplicaciones más recomendadas para transmisión en servicios menores de paquetes de datos como lo son los GPRS.

Mientras que un dibujo pinta miles de palabras y esta cantidad de texto puede ser soportada fácilmente por los GPRS, se espera que una imagen simple pueda ser utilizada en su lugar.

5.1.3.4 Imágenes en Movimiento

Enviar imágenes en movimiento en entornos móviles tiene gran cantidad de aplicaciones como los son las cámaras en construcciones para observar el acceso, enviar imágenes de pacientes desde ambulancias a los hospitales, etc. Aplicaciones de videoconferencia en donde la gente se puede reunir sin ir a la locación físicamente es otra de las aplicaciones a las imágenes en movimiento que es similar a compartir documentos, trabajo en equipo. Se ha comentado que los fabricante no requieren video y que los consumidores no los necesitan. De cualquier forma Internet se esta convirtiendo en entorno mas multimedia, 3G será capaz de desplegar imágenes y acceder servicios web.

La transmisión de imágenes en movimiento es una de las aplicaciones que GPRS y terminales 3G y vendedores de infraestructura rutinariamente y repetidamente proponen como área de aplicación obligatoria que habilitara mejores tasas de transmisión de datos. No es incorrecto hacer esto, debido a que en algunas demostraciones con tecnologías actuales se ha confirmado que un megabyte de datos transmitido sobre un enlace aéreo utilizando Microsoft NetMeeting para establecer una videoconferencia no tan fácilmente libera broadcast de imágenes con calidad de video. De cualquier forma aplicando técnicas de compresión se deben permitir transmisiones de imágenes con calidad de video más aceptables utilizando un ancho de banda de 64Kbps.

Las imágenes como los son dibujos y tarjetas postales serán las aplicaciones mas significativas para GPRS, imágenes en movimiento no tendrán la suficiente calidad al inicio para elevar la comunicación por el medio. Los usuarios podrán gastar todo el tiempo para ajustar el tamaño de la imagen en su pantalla tratando de trabajar con lo que están viendo.

Por esta razón, una vez mas, 3G elevara el ancho de banda para permitir gran calidad en la transmisión de imágenes sobre redes móviles. Se puede notar que el video y aplicaciones de transmisión de imágenes migraran a portadoras 3G tan pronto como estén disponibles.

5.1.3.5 Entorno de Casa Virtual (VHE)

El Servicio de Telefonía Móvil Unieversal (UMTS) que ocasionalmente es mencionado en los folletos de vendedores se conoce como VHE (Virtual Home Environment), este es un servicio que permite a los consumidores tener acceso similar al de casa con una vista común y sentir sus servicios desde cualquier ciudad. VHE es entonces un pequeño conjunto del total de usuarios de teléfonos móviles. Contando cada individuo con un dispositivo móvil con un número telefónico asociado en lugar de que el número corresponda a la ubicación.

VHE también puede permitir algunos otros servicios por medio de colocar el Modulo de Identidad Universal (UIM) en cualquier terminal, no solamente en dispositivos móviles sino en cualquier otro si las tarjetas magnéticas son mas ampliamente utilizadas que en la actualidad. Debido a que al querer tener acceso

de cualquier otra terminal se requiere de una tarjeta magnética para la validación del usuario. Esto surge a partir de la idea de que el correo electrónico y el acceso a Internet día a día esta siendo mas utilizado sin que exista alguna limitante del la ubicación del usuario la cual va a poder ser desde cualquier parte del mundo.

5.1.3.6 Agentes Electrónicos

Los agentes electrónicos es la tecnología que se piensa que jugara un papel importante en el trabajo móvil en el futuro. Los agentes nos servirán para obtener búsqueda y tareas en Internet y obtener reportes de los propietarios. Estoa es una manera eficiente de obtener información en movimiento.

Los Agentes Electrónicos están definidos como programas móviles que irán a lugares en la red para obtener instrucciones de los propietarios. Pueden ser pensados como extensiones de gente que los despacha. Los Agentes contienen programas que rastrean las redes de comunicaciones liberando y recibiendo mensajes o viendo información o servicios.

La terminal 3 G les dará a los propietarios más control sobre sus vidas que los teléfonos móviles. Esto serán Asistentes, Secretarias y Administradores electrónicos, como se puede observar tendremos ayuda vía Internet para cualquier actividad. El tipo de control que los Ingleses esperan dentro de los próximos 10 años es que las terminales móviles nos despierten, nos lean los correos, ordenen los productos del supermercado, nos informen la ruta de trabajo mas conveniente, nos recuerden llamadas y nos realicen traducciones.

5.1.3.7 Descarga de Software

En el siglo XXI, será mas común cargar el software electrónicamente como lo es Internet en lugar de comprarle como productos en cajas en las tiendas. Esto es como aplicaciones de transferencia de archivos que involucra cargar el software por si mismo. Se va a requerir, por ejemplo, el WinZip o Adobe Acrobat para leer los archivos y puedan ser cargados desde redes 3G a terminales 3G.

Cargar el software de la red tiene varias ventajas:

- ↓ Ambiente agradable: no se necesitan empaquetar para transportarlos a tiendas.
- ↓ Rápido y conveniente: los productos que pueden cargar se liberan directamente a la computadora. Arribaran en minutos, no en días.

Tiempos de descarga

Los tiempos de carga dependen de la velocidad del modem y tamaño de la aplicación. Típicamente el tiempo de carga varia entre los 10 minutos y dos horas.

En la tabla siguiente se puede observar el tiempo de descarga de una aplicación de 5MB

Velocidad de la Conexión	Tiempo de Carga
Conexión T1	30 segundos
Conexión ISDN	12 minutos
Modem 28.8	104 minutos

A partir de que 3G este disponible, dado el ancho de banda que maneja, este será el principal portador para descargar el software.

5.1.4 Generadores de tráfico en 3G

Con todos los servicios nuevos, en el caso de los negocios es importante estimar cuáles aplicaciones se implantarán en ésta tecnología. Las principales aplicaciones que utilizarán la Tercera Generación serán:

Posición	Aplicación	Portadora
1	Voz sobre IP (VoIP)	3G
2	Imágenes en movimiento	3G
3	Acceso remoto a LANs	GPRS /3G
4	Transferencia de archivos	3G
5	Descarga de software	3G
6	Web Browsing	GPRS /3G
7	Audio	GPRS /HSCSD /3G
8	Compartición de documentos y trabajo en equipo	GPRS /3G
9	Automatización de Casa	GPRS /3G
10	Agentes electrónicos	GPRS /3G
11	Autorización Dinámica	GPRS /3G
12	Entorno de casa virtual	3G

La primera de las aplicaciones listada será la más popular debido a la expansión de redes de telefonía estacionarias pero que anteriormente no estaba totalmente disponible sobre redes móviles. El Internet y el correo electrónico ahora están ocupando su lugar. La 3G permitirá que sean totalmente libres de cables y disponibles en cualquier lugar.

Las aplicaciones que se encuentran en la parte inferior carecen de popularidad en el mundo de las comunicaciones estacionarias pensando en la disponibilidad de soluciones de software específicas.

Estas aplicaciones son técnicamente factibles, los servicios de transmisión de datos a alta velocidad como lo es 3G, el volumen de uso es dependiente de factores comerciales como lo es el precio.

Se espera que 3G incorpore carga de software basada en volumen de manera que únicamente el dato enviado sea cargado para crear la manera de extender el uso a lo largo de los dispositivos compatibles con 3G.

5.1.5 Terminales 3G

Existen varias tendencias en las terminales 3G como se lista a continuación:

- ↳ Mayor y mejor tecnología en pantallas, las pantallas tendrán mejor colores, serán brillantes y tendrán un mayor área de pantalla que muchos teléfonos de hoy en día.
- ↳ El video será parte central en la demostración de la tecnología, aplicaciones multimedia es una de las partes mas importantes en que piensa 3G. La videoconferencia es una aplicación que anticipa muchos de los nuevos conceptos de terminales.
- ↳ Fabricantes de electrónica y teléfonos móviles convergen, como lo son las cámaras construidas dentro de los teléfonos móviles. La mayoría de estos dispositivos incluyen cámaras en miniatura.
- ↳ El factor mas popular que se muestra en los dispositivos 3G es el video en mano (video palm), un dispositivo que puede ser sostenido en una mano y que soporta aplicaciones de video con pantallas pequeñas, medianas y grandes.
- ↳ Casi todos los dispositivos nos son familiares, se podrá utilizar el teléfono para varias actividades y de diferentes maneras, pero probablemente se vea más similar a los teléfonos móviles actuales.

Las redes en 3G de banda ancha habilitan la opción de tener multimedia en dispositivos móviles. Cuando las redes estén en armonía y los clientes sean fuertes, las Tres Dimensiones de la Tercera Generación serán suficientes a nivel de aplicaciones y servicios, lo cual no era posible en las redes móviles actuales.

A continuación se muestra un conjunto de dispositivos con el concepto 3G de NTT DoCoMo, un operador de red Japonés. En la figura 5.1 se muestra un video teléfono construido sobre una barra flexible y en la figura 5.2 se muestra un teléfono visor que se utiliza como lentes y se puede ver como un televisor.



Figura 5.1. Videoteléfono

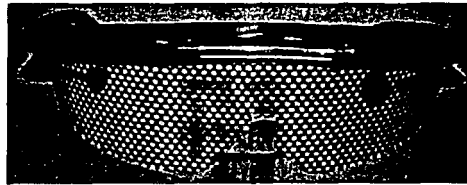


Figura 5.2. Teléfono visor

5.2 ARQUITECTURA BASADA EN DSP PARA COMUNICACIONES MÓVILES

Un diagrama de bloques del funcionamiento del sistema GSM se muestra en la figura 5.3. Se reconoce un modelo clásico de comunicaciones digitales con compresión de señales, corrección de errores, encriptación, modulación y ecualización.

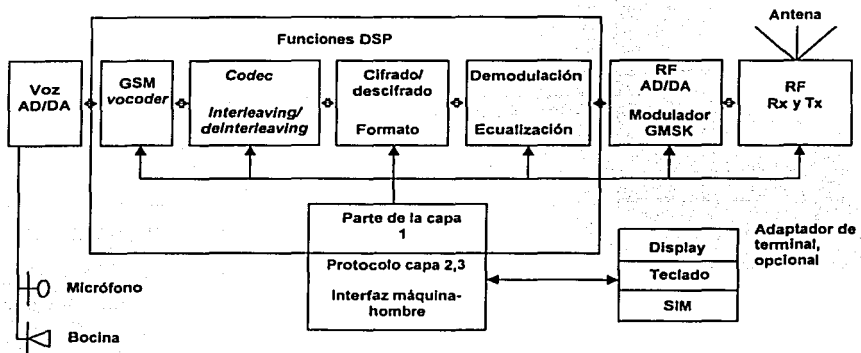


Figura 5.3: Diagrama de bloques de un teléfono GSM

El codificar utilizando GSM fase 1 comprime la señal a 13 KPSS utilizando la técnica de codificación predictiva lineal de excitación de pulsos rectangulares con predicción a largo plazo (RPE-LTP: Rectangular Pulse Excited Linear Predictive Coding With Longterm Prediction) como parte de la especificación GSM 06-10. El codificador de voz es la parte de la arquitectura con la cual la mayoría de los ingenieros están de acuerdo en que deba ser realizada en un Procesamiento

Digital de Señales (DSP: Digital Signal Processing) y en algunos diseños el DSP fue incluido principalmente para la codificación de voz. A medida en que los DSP se hicieron más potentes, empezaron a tomar parte en otras tareas de la capa física hasta que todas las funciones del cuadro de funciones DSP fueron incluidas, como se puede observar en la figura 5.3.

Los DSP no fueron lo suficientemente significativas para que los diseñadores de sistemas los implantaran. Después de 1994, un DSP era lo suficientemente potente para desempeñar todas las funciones DSP, convirtiéndose en la única solución. Para mejorar el consumo de energía y el tamaño del tablero, gran cantidad de DSP como el Motorola 56652 y plataformas digitales de banda base TI's, integran un microcontrolador RISC para manejar el protocolo y tareas de la interfaz de usuario para liberar al DSP de las tareas del algoritmo de comunicación. El particionamiento actual más popular de GSM se muestra en la figura 5.4.

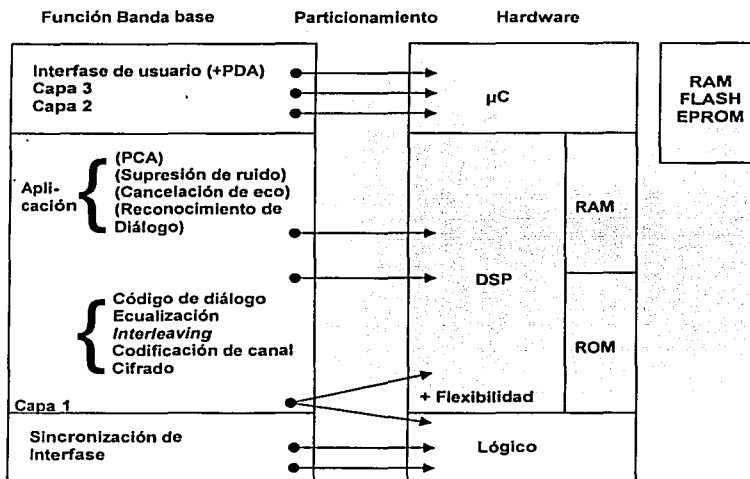


Figura 5.4. Función de particionamiento GSM

También es cierto que los teléfonos GSM han evolucionado, gradualmente han cambiado desde las funciones simples del teléfono y han permitido un incremento en la fracción del MIPS del DSP utilizado por algunos otros como la capa física.

Esta evolución se muestra en la figura 5.5. Con la llegada de aplicaciones inalámbricas de datos y el incremento de ancho de banda para 3G, se espera que esta tendencia se acelere.

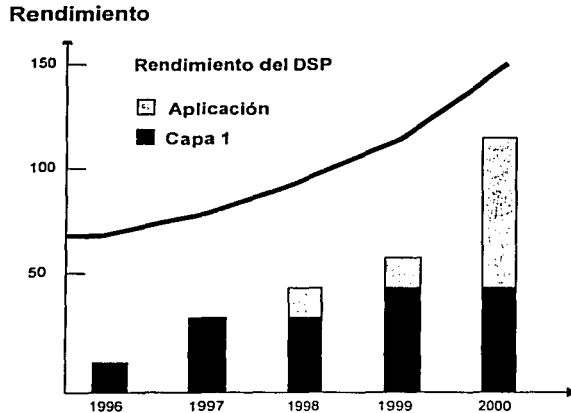


Figura 5.5. Capa1 y aplicación MIPS contra tiempo

5.2.1 Tendencia a la baja potencia de los DSPs

La sección digital de banda base es crítica para el éxito de los microteléfonos, los DSPs programables son esenciales para proporcionar costo efectivo, ruta de actualización flexible para la variedad de estándares en evolución. Arquitectura, diseño y mejoras en los procesos están produciendo una nueva generación de procesadores que proporcionan alto rendimiento mientras mantienen baja disipación de energía, necesaria para aplicaciones que demandan mayor energía.

Antes de evaluar la disipación de energía de los DSP utilizando miliwatts por millón MuAccs (MmuAcc), donde MuAcc consiste en dos funciones operacionales de memoria, desempeñando una acumulación multiplicada y almacenando el resultado nuevamente en memoria. La disipación de energía sigue la tendencia a disminuir a la mitad cada 18 meses. Así como la industria cambió de la segunda generación a la tercera generación inalámbrica, se puede observar el porcentaje de la capa física MIPS que reside en el DSP que va del 100% en tecnología para GSM hasta acerca de un 10% para acceso múltiple por división de código de banda base (WCDMA: Wideband Code Division Multiple Access). De cualquier

forma la tendencia, a lo largo de arquitecturas más eficientes y conjunto de instrucciones mejoradas, implica que el DSP dentro de algunos años será capaz de implementar en su totalidad la capa física WCDMA con el mismo consumo de energía que los teléfonos GSM.

Desde que el uso de DSPs utilizan lógica estática, el principal consumo de energía será cargado y descargado en los capacitores cuando el dispositivo este programado. Esta disipación de energía dinámica esta dada por:

$$Potencia = \alpha C * V_{swing} V_{sup\ plv} * f$$

Donde α es el número de veces en que el nodo interno lleva a cabo un ciclo por cada ciclo de reloj y V_{swing} es usualmente igual a $V_{sup\ plv}$. La energía dinámica para dicho chip es la suma de esta potencia sobre todos los nodos del circuito. A partir de que la potencia es igual al voltaje al cuadrado, disminuir el voltaje tiene el impacto más significativo en la potencia. Por ejemplo, disminuir el voltaje desde 3.3V a 1.8V decremента la disipación de potencia en un factor de 3.4, de cualquier forma, si la tecnología es constante, disminuir el suministro de voltaje también decremента el rendimiento. Por lo que escalar la tecnología (lo cual decremента la capacitancia) y escalar el suministro de energía son combinados para mejorar el rendimiento mientras que decremента el total de consumo de energía del DSP[3]. El TMS320C54x es capaz de 120 MmuAccs a 1.8 V y 0.72 mW/MmuAcc en 0.18 μ en tecnología CMOS.

Si se utilizara el TIC54x como ejemplo de la evolución de los DSP que es optimizado para aplicaciones inalámbricas. De cualquier forma, a pesar de que los DSP fueron diseñados para aplicaciones inalámbricas, encuentran mayor mercado en otras aplicaciones de baja potencia como los módems, cámaras digitales y reproductores de audio de estado sólido. El núcleo utiliza una arquitectura que incorpora 3 bases de memoria (2 lectura y 1 escritura) y un bus de memoria programable acoplado con 2 generadores de direcciones de datos y un generador de dirección de programa. Esta arquitectura permite alto ancho de banda en memoria y habilita múltiples operaciones, resultando en pocos ciclos para completar la misma función.

Otra estrategia utilizada para diseñadores de DSP es agregar instrucciones que a pesar de ser justamente genéricas para ellos mismos, permiten implementación de algoritmos importantes para aplicaciones inalámbricas.

En el C54x, una de las entradas a unidades lógicas aritméticas (ALU: Arithmetic Logic Unit) pueden ser tomadas desde un *40-bit barrel shifter*, permitiendo al procesador realizar escalamiento numérico, extracción de bits, aritmética extendida, prevención de sobrecarga. El intercambiador y detector exponente habilitan la normalización de ciclos individuales de valores y codificación exponencial para soporte de aritmética de punto flotante para codificación de voz.

Una unidad de comparación, selección y almacenamiento contiene un acelerador que para decodificación de canal reduce el Viterbi "butterfly update" a cuatro ciclos. Esta unidad generalmente proporciona aceleración para cualquier código convolucional basado en intercambios individuales de registro, cuyas cuentas se utilizan para todos los códigos en aplicaciones inalámbricas.

Otra tendencia en la evolución de DSP es hacia los procesadores VLIW que soportan un entorno de programación amigable basado en compiladores. Los procesadores VLIW utilizan computación de instrucciones paralelas explícitas (EPIC: Explicitly Parallel Instruction Computing) con predicción y especulación para ayudar a los compiladores. Los procesadores también son estáticamente programables, implementación de múltiples ediciones para explotar el paralelismo a nivel de instrucciones inherentes de gran cantidad de aplicaciones de DSP.

A pesar de que las aplicaciones de procesamiento de la capa física en el microteléfono aparentemente no es tan cercano, estos dispositivos permiten compilaciones de código muy eficientes en niveles altos, lo cual reduce la necesidad de algoritmos de codificación a nivel de ensamblado para DSP específicos. La tendencia de comunicaciones inalámbricas va conducida hacia sistemas abiertos que harán este tipo de DSP mucho más enfocados hacia procesadores multimedia en los microteléfonos.

La administración de energía es muy importante en los DSP de baja potencia, y el C54x utiliza una estrategia de administración de energía híbrida con reloj local automático y tres modos de controladora idle para apagar el CPU únicamente, el CPU y los periféricos o el apagado del dispositivo completo los cuales preservaran los estados de la memoria. Un PLL (Phase-Locked Loop) digital basado en un generador de reloj y un multiplicador permite al usuario optimizar la frecuencia y la energía para sus aplicaciones. En general, estas técnicas permiten a un DSP no diseñado para funciones específicas, optimizar el uso de energía para dicha función proporcionando un nivel de energía cercano para esa o algún diseño ASIC dedicado.

5.2.2 Coprocesadores

Los coprocesadores pueden complementar la función de los DSPs programables en la implantación de una plataforma flexible 3G.

Se utilizará el estándar WCDMA de ETSI (ARIB/European Telecommunications Standards Institute) para ilustrar dicha función[6].

A pesar de que se mencionó que en un estándar inalámbrico dado, la cantidad de la capa física desempeñada por un DSP se incrementará hasta un valor cercano al 100% como un estándar maduro, todavía existe el problema de implantación de un nuevo estándar con los DSPs actuales. Generalmente porque los estándares están dirigidos para que sea posible la implementación de ASIC para un punto de

energía y costo dado, un nueva definición de estándar no puede ser implantado en un DSP individual. Para una terminal de transmisión de voz WCDMA, si hacemos un conteo de operaciones requeridas, únicamente el 10% es compatible para implementación en un DSP actual, esto siendo esencialmente la operación de funciones en datos a tasa de símbolo en oposición a tasa de chip. Pero una solución mezclando funciones podría ser un opción de alto riesgo debido a la falta de flexibilidad, especialmente en un nuevo estándar. Por lo que los diseñadores de sistemas están enfrentando el problema balanceando los requerimientos de energía y flexibilidad. Una de las soluciones a este problema es una arquitectura basada en coprocesadores con un único dispositivo programable como núcleo.

El mundo de los coprocesadores se dividen en libremente acoplados LLC (Loosely Coupled Coprocessor) y fuertemente acoplados TCC (Tightly Coupled Coprocessor), los cuales están definidos por el porcentaje relativo para completar una instrucción en el DSP. Con un TCC el DSP iniciará una tarea en el coprocesador que completa en pocos ciclos de instrucción. Una tarea iniciada en un LCC se ejecutará para múltiples ciclos de instrucciones antes de que requiera nueva interacción con el DSP.

El TCC generalmente tendrá una interfaz específica con el núcleo del DSP y tendrá acceso a algunos registros dentro de dicho núcleo. A partir de que cada tarea únicamente tomo algunos ciclos, éste involucrará únicamente una pequeña cantidad de datos. Además de que la programación paralela de tareas en el DSP y el TCC serán muy complicados, desde que el DSP interrumpirá sus tareas después de algunos ciclos para servir el TCC. El DSP generalmente se congelará durante la operación del TCC. El TCC es por lo tanto una instrucción definible por el usuario que proporciona mejoras en la potencia y velocidad para pequeñas tareas donde no existen cuellos de botella a través del DSP. Un TCC también debe tener una tarea muy específica y relativamente pequeña comparada con el DSP. Con el tiempo, la función del TCC será absorbida dentro del DSP, por medio de reemplazarlo con código en un DSP de baja potencia más rápido o absorbiendo la función del TCC dentro del núcleo del DSP y dándole instrucciones específicas.

Los LCCs son más análogos para una llamada de subrutina que una instrucción. Como ellos desempeñan gran cantidad de operación sin la intervención del DSP, generalmente operan en largos conjuntos de datos. A diferencia del TCC, el LCC tendrá que ejecutarse en paralelo con el DSP para el completo funcionamiento. Esto significa que el programador tendrá que ser más cuidadoso con la programación de las instrucciones LCC. De cualquier forma desde que el LCC tiene contacto mínimo con el DSP, la principal ventaja del LCC es que resuelve el problema del ancho de banda de bus que puede ocurrir cuando la entrada de datos al sistema es muy elevada. En cualquier caso, el ancho de banda se convierte en un cuello de botella para desempeñar la operación porque los datos son almacenados al final del bus desde las unidades computacionales. El LCC remueve estos cuellos de botella por medio de tener las unidades operacionales

locales para datos específicos para una clase de operaciones. En el tiempo el DSP evolucionará a un punto en donde el ancho de banda del bus y la potencia operacional sea suficiente para tareas del LCCs y la pseudo rutina implementada por medio del LCC que se convertirá en una subrutina real.

El concepto de LCC aplica fácilmente de la velocidad electrónica al límite de velocidad de símbolo para el sistema CDMA. En la capa física podría todavía desempeñar gran cantidad de tareas de procesamiento a velocidad de símbolos como lo es la recuperación de la regulación de tiempo, frecuencia, estimación de canal y ubicación digital. Las tareas de procesamiento a velocidad electrónica como lo es estimación de dispersión de la ruta de retraso y la adquisición puede ser cedida al coprocesador que es diseñado para realizar dichas tareas eficientemente.

Una posibilidad es un procesador correlacional, el cual lleve a cabo tareas comunes para la ruta de estimación de retraso. El procesador también puede desempeñar algunas simples pero elevadas tareas MIPS que ocurren directamente en el límite a tasa de símbolo. Ejemplos de estos son promedios para estimación de canal. De cualquier forma, el DSP aún selecciona el tipo de promedio que deberá ocurrir y cómo los post procesos de datos para producir la estimación de canal final.

El sistema es completamente programable en el dominio de procesamiento a tasa de chip CDMA. El DSP puede escoger ubicar una porción del MIPS para un usuario con 6 rutas. Alternativamente, puede reubicar los mismos MIPS a múltiples usuarios con múltiples rutas. Además de permitir diferentes usuarios para diferenciar y mejorar las soluciones WCDMA por medio de software, esto permite que el mismo sistema sea reprogramado para desempeñar sistemas de demodulación WCDMA, GSM, CDMA2000 y algunos otros basados en CDMA. También proporciona una plataforma común tanto para terminales multimedia y de voz. Un diagrama de bloques simplificado de este coprocesador se muestra en la figura 5.6. Se puede observar que el coprocesador es conectado directamente al usuario final para remover los datos a velocidad electrónica del bus. Otra característica importante es que la instrucción y el buffer de salida están mapeados a la memoria para permitir acceso flexible al coprocesador, por el DSP.

La decodificación es otra área que se puede beneficiar por las aplicaciones LCCs. La decodificación a tasa de voz Viterbi es fácilmente desempeñada por los DSPs actuales, pero los requerimientos a alta tasa de datos en 3G hacen que la decodificación se difícil para efectuarse de manera programable.

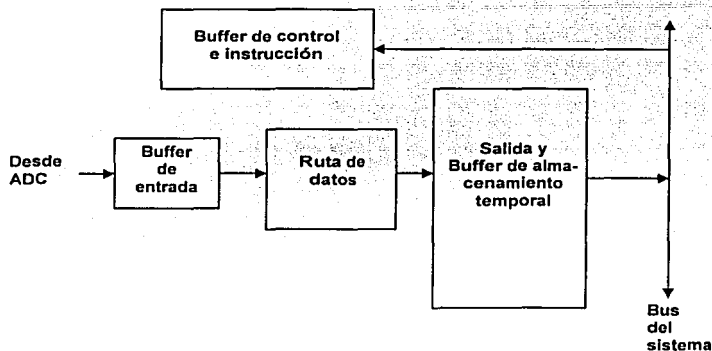


Figura 5.6. Coprocesador correlacional para procesamiento a tasa binaria en WCDMA

5.3 Arquitectura programable para OFDM-CDMA

Combinando multiplexión ortogonal por división de frecuencia (OFDM: Orthogonal Frequency Division Multiplexing) y tecnologías CDMA hace posible la realización del software de radio. Basado en el cuadro de trabajo unificado conocido como OFCDMA, se desarrolla una estructura programable de transconductores OFDM-CDMA. Por medio de ajustar los parámetros del sistema sin cambiar la arquitectura fundamental de hardware y software, varios escenarios de sistemas pueden ser implementados, lo cual puede servir como la fundación para diseñar el software de radio.

En múltiples sistemas y en entornos de múltiples estándares inalámbricos, el software de radio definido es una tecnología prometedora para cumplir con las necesidades en todos los aspectos. Como CDMA ha captado la atención de la tercera generación de sistemas inalámbricos principalmente por su alta capacidad sobre sistemas de comunicación celular TDMA y FDMA, su realización para CDMA de banda ancha invoca gran interés en el desarrollo de comunicaciones inalámbricas. De cualquier forma, el desempeño para CDMA de banda ancha para múltiples interfaces de acceso (MAI: Multiple Access Interface) e interfaces Inter símbolo (ISI: Inter Symbol interface), ambas para alimentación de múltiples rutas. Para utilizar mejor el espectro y los recursos de radio para servicios multimedia

con una tasa de error (BER: Bit Error Rate) satisfactoria, CDMA de múltiple portadora y OFDM es ampliamente considerado para aplicaciones multimedia inalámbricas.

La transmisión multiportadora, también conocida como OFDM es utilizada para combatir la fluctuación en la potencia de la señal de radio. En un sistema de transmisión de múltiples portadoras, el ancho de banda del canal disponible es dividido en múltiples subcanales como los símbolos de datos modulados por diferentes subportadoras pueden ser transmitidos en paralelo. En orden para hacer funcional el mayor ancho de banda disponible, el espectro de subcanales adyacentes es permitido para coincidir sin la interferencia entre canales, así como una manera de que todas las formas de onda del comportamiento de la información de los subcanales sean ortogonales en algún intervalo de tiempo. En un transmisor OFDM la cadena de datos originales con duración simbólica T' es convertida de serial a paralelo (S/P) y dividida en J grupos de subcadenas con duración simbólica T . La mayor eficiencia del ancho de banda es desarrollado por medio de $T = JT'$. De cualquier forma, el aprovechamiento usual es para diseñar $T = JT' + T_g$, donde T_g es el tiempo de espera. OFDM puede disminuir la fluctuación en la potencia de la señal de radio si T_g es mayor que el retraso de dispersión del canal desde que las formas de onda de los subcanales permanecen ortogonales en JT' , lo cual es entonces utilizado para demodulación en lugar de T . Los sistemas OFDM son usualmente implementados utilizando transformada de Fourier por dos razones, primero es programable en implementaciones digitales y segundo, por la existencia de algoritmos rápidos para la transformada de Fourier. En implementaciones digitales, un prefijo cíclico es utilizado con duración equivalente al tiempo de espera T_g . La mayor desventaja de OFDM es su alta sensibilidad a variaciones de frecuencia y de formas de onda. Sin embargo, la precisión de datos y linealidad del amplificador RF puede ser tomado en consideración.

Modernas comunicaciones multimedia inalámbricas adoptando CDMA usualmente ocupan un gran ancho de banda, resultando en una posible alimentación de frecuencia selectiva para técnicas de transmisión de una sola portadora. La transmisión multiportadora, la cual espera para evitar la selección de frecuencias, ha sido propuesta para una mejor eficiencia del ancho de banda por medio de minimizar la separación entre portadoras adyacentes. Esto reduce considerablemente la complejidad de la ecualización o de la estructura de recepción RAKE. Por otro lado, OFDM combinado con CDMA puede resultar en un fino particionamiento de los recursos de radio en el dominio del tiempo, frecuencia y código, entonces mayor asignación de recursos de radio podrían posiblemente proporcionar un esquema de asignación de ancho de banda.

5.3.1 OFDM-CDMA

Generalmente tres tipos de sistemas para combinar OFDM y CDMA son presentados:

- ↓ Multiportadora-CDMA (MC-CDMA)
- ↓ Secuencia directa de multiportadora (DS)-CDMA (MC-CDE-CDMA)
- ↓ Multitono (MT)-CDMA

MC-CDMA.- el transmisor MC-CDMA distribuye la cadena de datos originales sobre diferentes subportadoras utilizando un código de dispersión en el dominio de la frecuencia. La figura 5.7a muestra el transmisor MC-CDMA para los usuarios en un esquema de transferencia de fase binaria (BPSK: Binary Phase Shift Keying). La cadena de datos originales con duración simbólica T' es convertida de serial a paralelo para J grupos de cadenas de datos con duración simbólica $T = JT'$. La secuencia $(c_k^1 c_k^2 c_k^3 \dots c_k^M)$ constituye un código de dispersión para k usuarios. Entonces M ramas idénticas de cadenas de datos para cada grupo son distribuidas por el código de dispersión. Finalmente portadoras adyacentes iguales $\frac{1}{T}$ para satisfacer la condición de ortogonalidad con una mínima separación de frecuencia.

MC-DS-CDMA.- el transmisor MC-DS-CDMA distribuye la cadena de datos convertidos de serial a paralelo utilizando un código dado de dispersión en el dominio del tiempo, entonces el espectro para cada subportadora es ortogonal con una mínima frecuencia de separación. La figura 5.7b muestra el transmisor MC-DS-CDMA para los usuarios en un esquema BPSK. La cadena de datos originales con duración simbólica T' es la conversión de serial a paralelo para J grupos de cadenas de datos con duración simbólica $T = MJT'$. Entonces M ramas idénticas de cadenas de datos convertidos de serial a paralelo para cada grupo son distribuidos por la misma forma de onda $(c_k(t) = c_k^1 c_k^2 c_k^3 \dots c_k^{N_{MD}})$, donde N_{MD} denota la ganancia del procesamiento. Finalmente, cada rama modula una subportadora para transmisión. La separación entre portadoras adyacentes iguales $\frac{N_{MD}}{T}$ para satisfacer la condición de ortogonalidad con una mínima separación de frecuencia.

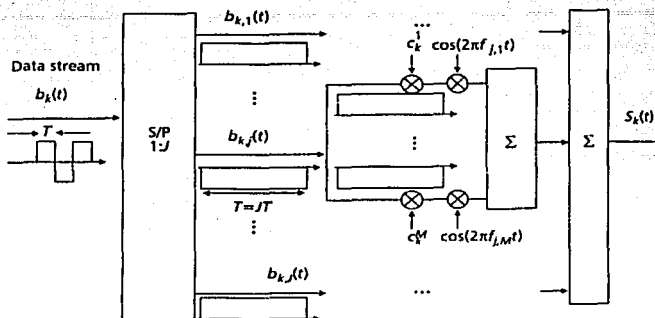


Figura 5.7a. Transmisor para el usuario k para MC-CDMA

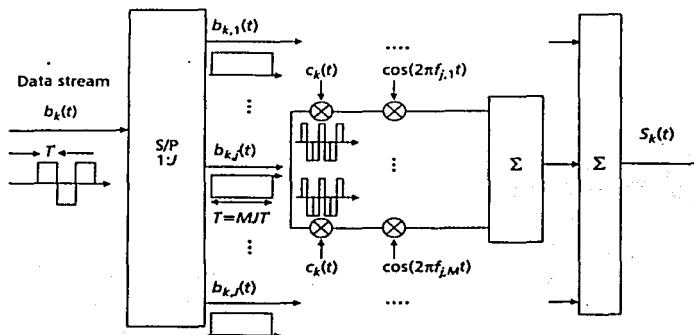


Figura 5.7b. Transmisor del usuario k para MC-DS-CDMA

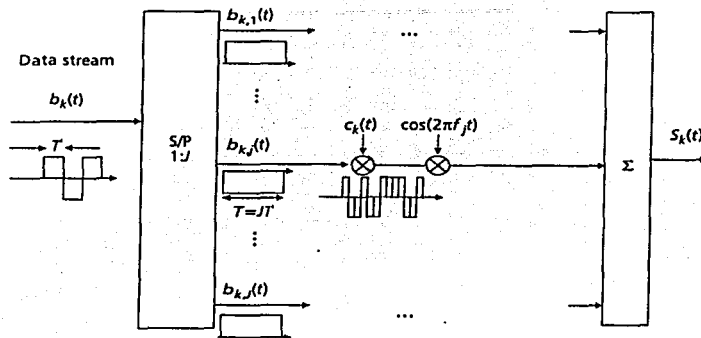


Figura 5.7c. Transmisor para k usuarios para MT-CDMA

MT-CDMA.- el transmisor MT-CDMA distribuye las cadenas de datos convertidas de serial a paralelo utilizando un código de dispersión dado en el dominio del tiempo, por lo que el espectro para cada subportadora antes de distribuirse es ortogonal con la mínima separación de frecuencia. Después de esto, el espectro para cada subportadora aumenta la posibilidad de la eficiencia espectral. La figura 5.7c muestra el transmisor MT-CDMA para k usuarios para el esquema BPSK. La cadena de datos originales con duración simbólica T' convertida de serial a paralelo para J grupos con duración simbólica $T = JT'$. La cadena de datos convertida de serial a paralelo para todos los grupos son distribuidos por medio de la forma de onda $(c_k(t) = c_k^1 c_k^2 c_k^3 \dots c_k^{N_{MT}})$, donde N_{MT} es la ganancia del procesamiento. Y finalmente cada grupo modula una subportadora para transmisión. La separación entre portadoras adyacentes iguales $\frac{1}{T}$ para satisfacer la condición de ortogonalidad con una mínima separación de frecuencia para la dispersión.

El propósito de M ramas idénticas de información MC-DS-CDMA es para incrementar la diversidad de frecuencias. Si la separación de portadora entre los canales adyacentes es lo suficientemente grande, proporcionarían alimentación no correlacional y diversidad de frecuencias. Por lo que M debe ser lo suficientemente grande para asegurar la dimensión de diversidad de frecuencias, pero lo suficientemente pequeño para mantener alimentación no correlacional. El principio de diseño para un grupo multiplexado J después de la conversión S/P para tres esquemas es hacer cada frecuencia de subcanal no selectiva. Por lo que, J debe ser lo suficientemente grande para hacer cada frecuencia de subcanal no selectiva bajo una complejidad permisible. A partir de que la

transmisión multiportadora dispersa el total de la energía de la señal en múltiples canales, bajo las condiciones de no tener tono piloto y el mismo desempeño como usando información de un subcanal completo de una portadora, la complejidad de adquisición y sincronización para este tipo de sistemas crece con el número de subportadoras.

MC-CDMA es preferencial cuando todas las subportadoras coherentemente moduladas porque la dispersión en el dominio de la frecuencia requiere una portadora segura y recuperación de fase.

Cuando una modulación no coherente es llevada a cabo en aplicaciones MC-DS-CDMA y MT-CDMA tienen mejor desempeño. MT-CDMA cubre parcialmente el espectro más denso entre subportadoras que MC-DS-CDMA para una mejor eficiencia espectral. Esta propiedad tiene la ventaja de acomodar más usuarios desde una secuencia mayor de pseudo ruido (PN: Pseudo-Noise), puede ser adoptada cuando el ancho de banda disponible es utilizado. De cualquier forma, MT-CDMA es susceptible a dos tipos de interferencias, interferencia interportadora y multiportadora. Algunas técnicas para mejorar el BER₁, como ecualización y detección de múltiples usuarios deben ser consideradas para el diseño del sistema.

5.3.2 OFCDMA

Tres tipos de sistemas OFDM-CDMA pueden ser descritos utilizando tramas de trabajo unificadas para el desarrollo de OFCDMA. Una arquitectura programable basada en OFCDMA puede acomodar el tráfico en aplicaciones prácticas de software de radio. La figura 5.8a describe la estructura del transmisor. La cadena de datos originales para el usuario k es la conversión S/P a grupos de cadenas de datos y M ramas idénticas para cadenas de datos para cada grupo serán posteriormente distribuidas para la forma de onda: $c_{k,m}(t) = (c_{k,m}^1 c_{k,m}^2 \dots c_{k,m}^M)$. Finalmente cada rama modula la subportadora $\cos(2\pi f_j t)$ para $m = 1, 2, \dots, M$, $j = 1, 2, \dots, J$, la mínima separación entre portadoras es $f_0 = \frac{1}{T_o}$, que es

$$f_{a,b} - f_{c,d} = \frac{1}{T_o}(a - c) + \frac{J}{T_o}(b - d).$$

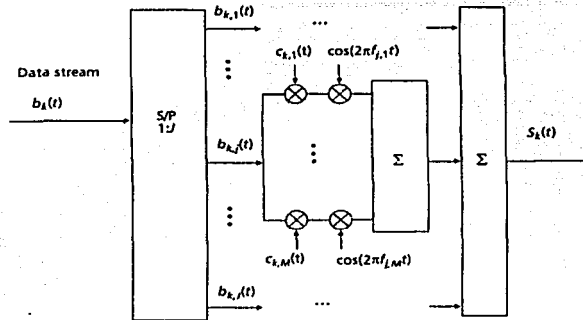


Figura 5.8a. Transmisor del usuario k para OFCDMA

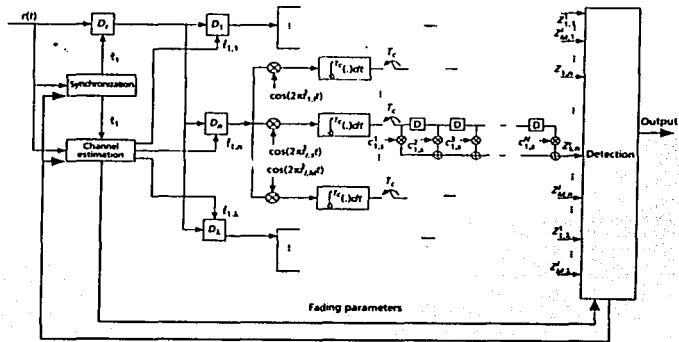


Figura 5.8b. Receptor RAKE del usuario 1 para OFCDMA

El receptor RAKE para OFCDMA es descrito en la figura 5.8b sin la pérdida de generalidad. La señal recibida $r(t)$ es sincronizada por retraso de $\hat{\tau}$ y entonces es bloqueada por el receptor RAKE por retraso de $\hat{\tau}_{1,n}$ para $n=1,2,\dots,L$ respectivamente. Donde $\hat{\tau}$ es la ruta de retraso estimada y para $n=1,2,\dots,L$ son los retrasos estimados para un canal de alimentación multiruta L -ray. En el

n dígito, la señal sincronizada primero es demodulada para cada subportadora de MJ , donde $\hat{f}_{r,s}$ es el estimador de frecuencia para cada subportadora. Entonces es alimentado un circuito con un intervalo de tiempo T_c , la duración del chip para firma de la forma de onda y muestreado al periodo de T_c .

Los datos muestreados entonces son pasados a una línea de retraso con coeficientes de carga $c_{r,s}^n$ para $n = 1, 2, \dots, N$ en orden de dispersión. Se reajustan los filtros de la línea de retraso a un periodo de $T = NT_c$. Finalmente, las salidas de los filtros de la línea de retraso para todas las subportadoras en todos los dígitos, $Z'_{r,s,n}$ para $r = 1, \dots, J$ $s = 1, \dots, M$, $n = 1, \dots, \lambda$, son posteriormente procesadas para detección de señal por medio de permitir parámetros de canal de estimación. Si la probabilidad de detección es lo suficientemente alta, las salidas detectadas son de ayuda para sincronización y estimación de canal. Por lo tanto, la detección de salidas es retroalimentada por sincronización y estimación de canal. Se pueden adoptar otras técnicas, como lo es el uso de una secuencia predefinida para sincronización o utilizar un tono piloto para estimación de canal.

El concepto OFCDMA puede describir los tres tipos de sistemas: MC-CDMA, MC-DS-CDMA y MT-CDMA. La estructura del transmisor para MC-CDMA descrito en la figura 5.7 puede ser implementado utilizando OFCDMA. La forma de onda generada es designada tal que $c_{k,m}(t) = c_k^m$ donde $c_k^1 c_k^2 \dots c_k^M$ constituye la secuencia PN para k usuarios, $T = JT'$ y $T_o = T$.

La estructura OFCDMA puede ser adoptada dentro de los transmisores descritos en la figura 5.8b. La forma de onda generada es tal que cada subportadora para k usuarios tengan la misma forma de onda, que es $c_{k,m}(t) = c_k(t) = (c_k^1 c_k^2 \dots c_k^{M_{usr}})$.

Además si $T_o = \frac{T}{N_{MID}}$, el transmisor para MC-DS-CDMA es un caso especial de la figura 5.9a por medio de colocar la misma forma de onda, lo cual implica una simplificación en la implantación real.

Finalmente OFCDMA puede ser transformada MT-CDMA por medio de tomar únicamente una rama en cada grupo J , que es $M = 1$, $c_{k,1}(t) \equiv c_k(t) = (c_k^1 c_k^2 \dots c_k^{M_{usr}})$ y $T = JT'$. Colocando $T_o = T$, la señal transmitida es reducida a la señal transmitida por MT-CDMA. Por simplicidad se define únicamente una rama para cada grupo de la cadena de datos convertidos de S/P y se define $C_{k,1} \equiv c_k(t)$.

seleccionadas G son retenidas. Esta operación es el bloque denotado como intercambio de ventanas. Finalmente la secuencia de la muestra G retenida es

$$\text{pasada a un filtro analógico paso bajas con un ancho de banda } \frac{MJ - 1 + \frac{T_o}{T_c}}{2T_o}.$$

La figura 5.8b muestra la arquitectura del receptor programable RAKE para OFCDMA. Los detalles de cada dígito es mostrado en la figura 5.9b. La tasa de datos (data rate) de los datos muestreados es $f_s = \frac{G}{T_c}$. La rama paralela G es completada con ceros y redondeando de acuerdo a las ramas paralelas después de conversión S/P desde los datos muestreados con tasa de datos $f_s = \frac{G}{T_c}$.

Entonces las muestras paralelas GH son posteriormente muestreadas por los puntos GH de la transformada inversa de Fourier y entonces la primer muestra MJ de la secuencia es retenida. Finalmente, N salidas consecutivas de MJ ramas paralelas multiplexadas a la entrada de la línea M de retraso para dispersión. Los coeficientes para la línea de retraso son $c_{1,n}^n$, para $n = 1, 2, \dots, N$. Por último, las salidas desde esos filtros de la línea de retraso son las variables previstas para la decisión de detección.

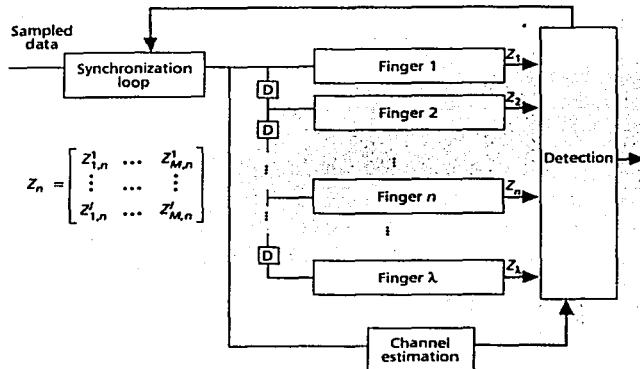


Figura 5.9b. Receptor RAKE digital del usuario 1 para OFCDMA

La arquitectura descrita en la figura 5.9a y 5.9b puede ser utilizada para implementar MC-CDMA, MC-DS-CDMA y MT-CDMA, respectivamente. El cambio de parámetros se puede observar en la tabla 1.

5.3.4 Características de Implantación

La arquitectura de software puede ser realizada por la estructura de hardware mostrado en la figura 4. Algoritmos de operación pueden ser almacenados en ROM. Basados en la selección del sistema, parámetros apropiados y subrutinas pueden ser escogidos y ajustados para desempeñar funciones de transmisión recepción por medio de un controlador apropiado en el controlador y ejecutando apropiadamente funciones del procesador de señal digital (DSP: Digital Signal Processor). El bloque del filtro paso bajas (LPF: Low Pass Filter) en la figura 3ª consiste de un convertidor analógico digital (DAC: Digital Analog Converter) y un LPF analógico. Es una característica importante si un DAC puede ser aplicado a todo el esquema multiportador CDMA. Se puede observar que la entrada de datos del LPF es diferente para los 3 esquemas. Para acomodar los tres esquemas por completo, se puede diseñar un DAC que opere a una tasa de datos mayores que esa en todos los esquemas y ajustar el ancho de banda para el LPF analógico. La entrada de datos para el LPF para cada esquema debe ser hecho igual por medio de técnicas de procesamiento digital de señales, como lo es la interpolación, muestreo y filtrado. Como para el ADC, se puede seleccionar con la misma tasa de muestreo como tasa de datos para el DAC mencionado anteriormente y muestrear las salidas siguientes al LPF digital para cubrir la menor tasa de datos requerida en cada esquema. De esta forma, es posible utilizar un DAC/ADC para ajustar los 3 esquemas.

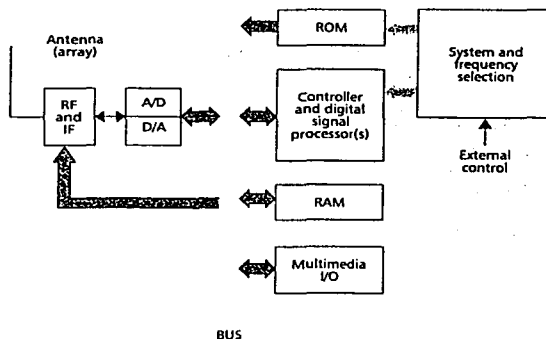


Figura 5.10. Estructura general de hardware para realizar el transceptor propuesto.

Para la selección RF, se deben considerar dos factores importantes; la asignación de la banda de la señal y el ancho de banda del transmisor RF. Si el ancho de banda del transmisor RF es permitido para ser lo suficientemente grande, mayor que todos los esquemas y la asignación de la banda de la señal es arreglada, únicamente se requiere una función RF. En general, la estructura programable OFCDMA ajusta funciones RF multibanda hasta el ancho de banda suficiente.

Para mejorar el rendimiento del sistema, algunas técnicas bien conocidas pueden ser incluidas, como modelo de señalización de ayuda, control de predicción de errores (FEC: Forward Error Control) y ecuación.

Las técnicas de modelo de señalización de ayuda pueden ser implementadas de manera flexible en sistemas multiportadores para mejorar el rendimiento del enlace. FEC puede ser aplicado también antes, después o entre dispersión y multiplexión por división de frecuencia. Ecuación puede ser realizada como un filtro en cada subportadora de nuestro receptor sin dificultad.

El mejor diseño tiene retraso real en la operación del sistema y puede ser incorporado dentro de la estructura programable propuesta. Algunas técnicas de DSP también pueden ser útiles para mejorar el rendimiento. También es necesario considerar un método para eliminar la interferencia causada por la alimentación de múltiples rutas.

Este puede ser mostrado de acuerdo a que OFCDMA puede ser descrito en términos de configuración síntesis/análisis de transformación de subbandas. Algunas ventajas de optimización propuestas en procesamiento de señales multitasa puede ser útil para el esquema OFCDMA.

Finalmente basado en la arquitectura programable propuesta, se pueden resumir las subrutinas básicas para desempeñar funciones de transmisión recepción OFDM-CDMA, como:

- ↓ Modulación y demodulación
- ↓ Completado por ceros
- ↓ Transformada inversa de Fourier
- ↓ Filtrado
- ↓ Multiplexión
- ↓ Filtrado de la línea de retraso
- ↓ Sincronización
- ↓ Filtrado paso bajas
- ↓ Intercambio de ventanas

5.4 Tecnología de antenas inteligentes. Aplicación a UMTS

Una antena inteligente (en la Figura 5.11 se muestran dos prototipos) es aquella que en vez de disponer de un diagrama de radiación fijo, en vez de disponer de un diagrama de radiación fijo, es capaz de generar o seleccionar haces muy directivos enfocados hacia el usuario deseado, e incluso adaptarse a las condiciones radioeléctricas en cada momento. Existen varias formas de implementar un sistema radiante con estas características, como se detalla a continuación.

La implantación de antenas inteligentes en una red de comunicaciones móviles se limita, en principio, a las estaciones base (o Nodos B en UMTS), debido a que necesariamente se deben emplear sistemas radiantes de mayor tamaño (*Arrays* de varios elementos). Esto tiene la ventaja añadida de que pueden introducirse las antenas inteligentes de forma transparente para los usuarios, que no tendrán que cambiar de terminal para beneficiarse de esta tecnología.



Figura 5.11. Prototipos de antena inteligente de la firma Allgon.

Las características de las antenas inteligentes de tener unos haces de radiación con una mayor directividad, hace pensar en las siguientes ventajas potenciales. [29], [30].

- ⚡ Incremento en la zona de cobertura: Dado que la ganancia es mayor que en el caso de antenas omnidireccionales o sectorizadas, para igual potencia transmitida, la señal se podría recibir a una mayor distancia. Este hecho podría permitir reducir el número de estaciones base necesaria para cubrir una zona, siempre y cuando no sea el tráfico el factor limitante.
- ⚡ Reducción de la potencia transmitida: La mayor ganancia de la antena permitirá incrementar la sensibilidad de la estación base, por lo que los móviles podrán transmitir con menor potencia, ahorrando batería. De igual modo, gracias a la ganancia del *array*, es posible que la estación base transmita igual

potencia, pese a que cada elemento del *array* esté radiando una potencia muy inferior.

- ↓ Reducción de la propagación multitrayecto. Debido a la menor dispersión angular de la potencia radiada por la estación base, se reducirá el número de trayectos múltiples que alcanzarán al móvil (mejorando así las características de dispersión de retardo del canal). Esto permitiría relajar los requisitos exigidos al ecualizador del terminal móvil. En UMTS, si bien se puede aprovechar la recepción multitrayecto mediante el uso de receptores Rake, cuando los móviles se desplacen a gran velocidad esta ventaja puede ser muy significativa.

En el caso del enlace ascendente, la antena inteligente de la estación base podría discriminar las componentes multitrayecto de la señal recibida desde el móvil, o incluso explotarlas (mediante receptores 2D-Rake, por ejemplo). Esto dependerá de la configuración de antena escogida.

- ↓ Reducción del nivel de interferencia: La mejor selectividad espacial de la antena permitirá a la estación base discriminar las señales de usuarios interferentes a favor de la señal del móvil deseado (en el caso del móvil ascendente), y también reducir el nivel de potencia transmitida en las direcciones de esos otros usuarios (en el caso del enlace descendente). De cualquier modo, se conseguiría aumentar la relación C/I , lo cual tiene dos consecuencias fundamentales:

- a) Una mejora en la C/I implica directamente una mejora en la tasa de error (BER), lo que hace que la calidad del servicio aumente.

- b) La reducción de la C/I puede explotarse directamente (mediante técnicas de multiplexación espacial) o indirectamente (realizando un plan de frecuencias más ajustado, en el caso de GSM) para aumentar la capacidad del sistema.

- ↓ Mejora de la seguridad. Gracias a que la transmisión entre la estación base y la terminal móvil es direccional, no será posible que un equipo ajeno intercepte la comunicación, a menos que se sitúe en la misma dirección en la que apunta la antena.
- ↓ Introducción de nuevos servicios: Puesto que la red podría tener acceso a información acerca de la posición de los móviles, es posible pensar en servicios tales como radiolocalización en llamadas de emergencia, tarificación geográfica, publicidad de servicios cercanos, información en lugares turísticos, gestión avanzada de flotas, etc.

No obstante, la implantación de estas antenas en la red móvil no está exenta de inconvenientes, como los que se detallan a continuación.

- ↓ Mayor complejidad de los transceptores: En comparación los sistemas radiantes convencionales, los sistemas de antenas inteligentes son mucho más complejos y difíciles de diseñar. Será necesaria una cadena de transmisión/recepción independiente para cada elemento del *array*, y todas ellas deberán estar balanceadas y calibradas en tiempo real. Además, es imprescindible el uso de potentes procesadores (DSPs, por ejemplo) para ejecutar los algoritmos de optimización, conformación de haz, detección del ángulo de llegada, etc.
- ↓ Mayor complejidad de los procedimientos de gestión. El hecho de que exista un haz de radiación enfocado hacia cada usuario implica que las funciones de red deben revisarse, en particular, las que afectan a la gestión de recursos de radio (RRC) y a la gestión de movilidad (MM). Por ejemplo, algunos procedimientos que pueden verse afectados son lo de selección y reselección de celda, establecimiento de conexiones, handover, paging, tec.
- ↓ Cambios en los métodos de planificación. La introducción de un sistema de antena inteligente implicará tener muy en cuenta sus características, a la hora de realizar la planificación de la red celular. En particular, habrá que contar con el aumento de alcance, la eliminación de fuentes de interferencia, el seguimiento angular de los usuarios, etc.

5.4.1 Implantación de un sistema de antena inteligente

La característica básica que se busca en el diseño de un sistema de antenas inteligentes es la capacidad de seleccionar espacialmente a los distintos usuarios. Existen varias formas de implementar un sistema con esta capacidad, las cuales se describen a continuación:

- ↓ Haz conmutado. Es la técnica más simple. El sistema radiante genera varios haces fijos, cada uno de ellos apuntando en una dirección distinta, de modo que entre todos se cubre toda la zona deseada (un sector o celda). La inteligencia del sistema se encarga de seleccionar el haz que mejor servicio da a cada usuario en particular, en función de algún parámetro de control (mayor nivel de potencia recibida, mejor SNR y mejor C/I).

Esta técnica no garantiza que el móvil se encuentre en la dirección de máxima radiación del haz que le da servicio, ni que las señales interferentes se vean notablemente reducidas (ya que siempre es posible que alguna entre por uno de los lóbulos secundarios). De hecho, sería posible recibir una señal interferente por un punto del diagrama de radiación con mayor ganancia que la señal deseada, empeorando apreciablemente las prestaciones del sistema.

Una versión más avanzada de esta técnica consistiría en seleccionar con un haz la señal deseada y con otros algunas de sus componentes multitrayecto, de forma que puedan procesarse todas con un receptor Rake, como se muestra en la figura 5.12.

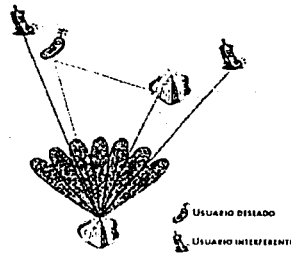


Figura 5.12. Antena de haz conmutado

- ✦ Haz de seguimiento. Esta técnica es más compleja que la anterior. Requiere el uso de un array progresivo (phased array); es decir, un array en el que se pueden controlar electrónicamente las fases con las que se alimentan los distintos elementos, de modo que puede modificarse a voluntad la dirección en la que apunta el lóbulo principal de la antena. A su vez, es necesario utilizar algún algoritmo de detección de la dirección de llegada (DoA), de modo que pueda reorientarse dinámicamente el haz para apuntar al usuario deseado.

Con esta técnica si se puede garantizar que el usuario se encuentra iluminado en todo momento por el lóbulo principal y con máxima ganancia (dentro de las limitaciones de los algoritmos que se empleen). Sin embargo, tampoco puede evitarse que las interferencias entren por algún lóbulo secundario del diagrama de radiación.

Para aprovechar las señales multitrayecto sería necesario detectar y seguir con otros haces dichas componentes y luego procesarlas con un receptor Rake, como se muestra en la figura 5.13.

- ✦ Haz adaptativo. Este sería el máximo nivel de inteligencia con que se podría dotar al sistema. En este caso, la salida de cada elemento del array se pondera con un factor de peso cuyo valor se asigna dinámicamente, de modo que se conforma el diagrama de radiación para maximizar algún parámetro de la señal (por ejemplo, la SINR). De este modo, el diagrama sintetizado habitualmente presentará un lóbulo principal en la dirección del usuario deseado, lóbulos secundarios en las direcciones de las componentes

multitrayecto (si se quieren procesar con un receptor Rake) y mínimos (e incluso nulos) de radiación en las direcciones de las fuentes de interferencia. Como se muestra en la figura 4.



Figura 5.13. Antena de haz de seguimiento

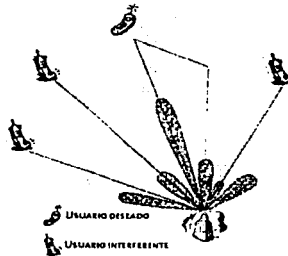


Figura 5.14. Antenas de haz adaptativo.

No siempre será posible eliminar toda la interferencia, ya que el número de fuentes interferentes que se pueden suprimir está directamente relacionado con el número de elementos de la antena.

Esta técnica requiere el uso de complicados algoritmos, tanto para la detección de las señales deseada e interferente como para la optimización de los pesos que conforman el haz. Estos algoritmos suelen conllevar una gran carga

computacional, mientras que deben procesarse en tiempo real, por lo que supone una seria limitación.

Existen tres métodos de introducir esta tecnología en una red de comunicaciones móviles, en función del grado de aprovechamiento de la selectividad espacial que ofrece:

1. Receptor de alta sensibilidad (HSR). Esta configuración consiste en utilizar antenas inteligentes sólo en el enlace ascendente. De este modo, gracias a la mayor directividad de la antena, se consigue mejorar la sensibilidad global de la cadena de recepción de la estación base. Esto supo varias ventajas.

En primer lugar, al mejorar la sensibilidad en el enlace ascendente, aumentará la extensión de la zona de cobertura. Esta mejora podría llegar a ser tan grande como para que fuera el enlace descendente el más restrictivo a la hora de calcular la cobertura de una estación base.

En segundo lugar, la mayor ganancia de la antena significa también que los móviles más cercanos podrían emitir con menor potencia manteniendo la calidad del enlace, con el consiguiente ahorro de baterías.

Por último, se lograría una mejora de la relación C/I, lo que implicaría menor tasas de error y una mejor calidad. No podría emplearse la mejora en la C/I para incrementar la capacidad de un sistema CDMA, ya que dicha mejora sólo está presente el enlace de subida y no en el de bajada.

2. Rechazo de interferencia por filtrado espacial (SFIR). En esta configuración se emplean antenas inteligentes tanto en el enlace ascendente como en el descendente, con lo cual se consigue aprovechar la mejora por selectividad espacial en ambas direcciones.

En este caso, la mejor que se experimenta en la C/I, además de reducir la BER del sistema, puede explotarse directamente para aumentar la capacidad de un sistema CDMA como es UMTS. Esto también podría lograrse indirectamente en GSM, si se hace un plan de frecuencias más ajustado: al ser menor la distancia de reutilización, puede aumentar el número de portadoras por estación base.

3. Acceso múltiple por división espacial (SDMA). Esta sería la configuración más compleja, pues consiste en aprovechar al máximo las propiedades de selectividad espacial de las antenas de ambos enlaces para ubicar simultáneamente a varios usuarios en el mismo canal. Es decir, que podría haber varios usuarios utilizando al mismo tiempo la misma frecuencia y el mismo código de scrambling (o el mismo timeslot en GSM), estando discriminados únicamente por su posición angular respecto de la estación base.

En este caso, el aumento en la capacidad se produce de forma directa, debido a que se ha añadido una nueva dimensión para la gestión del espectro.

La introducción de SDMA supondría la necesidad de contar también con complicados sistemas de gestión de usuarios, de asignación de canales, etc.

La aplicación de SDMA a UMTS es bastante dudosa: as ser un sistema CDMA existen gran cantidad de usuarios compartiendo simultáneamente la misma frecuencia y que se distinguen sólo por su código scrambling. Por tanto, sería muy complejo implementar un sistema capaz de diferenciar a cada usuario por su situación espacial, además de poco necesario, ya que los códigos de scrambling producen una separación suficiente y existen códigos de sobra para todos los usuarios.

5.4.2 Algoritmos de conformación de haz

A continuación se describen brevemente los algoritmos que deben implementarse a la hora de diseñar un sistema de antenas inteligentes basado en haces adaptativos, que son los más complejos y, a la vez, los más interesantes.

En primer lugar, supongamos que existen K usuarios, cada uno de los cuales transmite una señal $s_k(t)$. La señal del usuario k -ésimo sufre varias reflexiones y llega a la estación base a través de L_k trayectos distintos, cada uno de los cuales tiene una amplitud y una fase distintas ($\alpha_{ki} \cdot e^{j\varphi_{ki}} \equiv \alpha^{ki}$), y llega con un retardo (τ_{ki}) y una dirección ϕ_{ki} diferentes. En la estación base tenemos un array de N antenas, que, para simplificar, supondremos que están alineadas y equiespaciadas. En estas condiciones, la señal recibida en el elemento n -ésimo del array es:

$$u_n(t) = \sum_{k=1}^K \sum_{i=1}^{L_k} a_n(\varphi_{ki}) \cdot \alpha_{ki} s_k(t - \tau_{ki}) + n_n(t)$$

Donde:

$a_n(\phi_{ki}) = \exp(jknd \cdot \cos(\phi_{ki}))$ es la fase con la que llega la señal al elemento n -ésimo.

$k = 2\pi / \lambda$ es el número de onda para la frecuencia de trabajo, cuya longitud de onda es λ .

d es la distancia entre los elementos del array.

$n_n(t)$ es el ruido captado por el elemento n -ésimo del array.

Para sintetizar un diagrama de radiación con una cierta distribución de lóbulos, es necesario multiplicar la salida de cada uno de los elementos por un peso complejo $w_n(t)$ y luego sumarlas todas. Utilizando una notación vectorial, la señal a la salida del array será:

$$y(t) = \mathbf{w}(t)^H \cdot \mathbf{u}(t) = \mathbf{w}(t)^H \cdot \left(\sum_{k=1}^K \sum_{n=1}^{L_k} \mathbf{a}_n \cdot a_{ki} \cdot s_k(t - \tau_{ki}) + \mathbf{n}(t) \right)$$

Donde los vectores \mathbf{w} , \mathbf{a} y \mathbf{n} son vectores columna, con los elementos w_n , a_n y n_n correspondientes ($1 \leq n \leq N$), y el operador H significa que está transpuesto y conjugado.

El problema consiste entonces en encontrar un vector de pesos \mathbf{w} capaz de sintetizar un diagrama de radiación que nos sea útil para apuntar a un usuario. Si nuestro sistema dispone de N_R cadenas receptoras (es decir, es capaz de establecer hasta N_R comunicaciones simultáneas con otros tantos usuarios), habrá que calcular a la vez N_R vectores de pesos, aplicarlos a las salidas de los elementos del array y obtener así las N_R salidas (una para cada receptor).

El cálculo del vector \mathbf{w} que mejor se adapte a nuestras necesidades (según queramos maximizar la SNR, aprovechar la diversidad de trayectos, etc) puede realizarse mediante algoritmos muy diversos.

Existen básicamente tres tipos de algoritmos para el cálculo del vector de pesos, los cuales se diferencian en el modo en que seleccionan al usuario deseado: [31].

1. Técnicas con referencia temporal. En este caso, el receptor genera localmente una estimación de la señal del usuario que desea captar (obtenida usualmente a partir de una secuencia de entrenamiento o de un código scrambling). Entonces, el algoritmo encontrará el valor de \mathbf{w} que minimiza el error entre la señal de salida del array, $y(t)$ y la estimación $d(t)$. La solución óptima, por ejemplo es:

$$\mathbf{w}_{opt} = R^{-1} \cdot \mathbf{p}$$

Donde $R = E \begin{bmatrix} r & r_n \\ u & u \end{bmatrix}$ es la matriz de autocorrelación de la señal de entrada y

$\mathbf{p} = E \begin{bmatrix} r \\ u \cdot d(t)^* \end{bmatrix}$ es la correlación cruzada entre la señal a la entrada y la señal

deseada. No siempre es sencillo o viable obtener estas matrices de correlación, por lo que existen varias técnicas que tratan de aproximar este valor óptimo. Dos soluciones clásicas son la MMSE y la LS. La diferencia entre ambas es que la MMSE encuentra el w óptimo entre todo el conjunto de posibles realizaciones (para lo cual hay que suponer un entorno estacionario), mientras que la LS lo encuentra en un número finito de muestras (snapshots), por lo que no necesita suponer estacionariedad.

La solución encontrada con estas técnicas maximiza la SNR media (puede ser la media de conjunto en MMSE o una media temporal en LS). Con esto se consigue dirigir el haz hacia el usuario deseado, reduciéndose el nivel de ruido a la entrada. Además, impone nulos de radiación en las señales interferentes y de multitrayecto que lleguen no correlacionadas con la señal deseada. Las señales interferentes correlacionales entran por las antenas, pero se suman en contrafase a la salida, y el multitrayecto correlacionado se recibe por haces secundarios y puede explotarse con un receptor Rake convencional.

2. Técnicas con referencia espacial. Estas técnicas no requieren el uso de secuencias de entrenamiento, pero, en cambio, necesitan estimar la DoA del usuario deseado y de los interferentes (lo que supone el uso de otros complejos algoritmos para hacer dicha estimación).

La ventaja de éstas frente a las anteriores es que no necesitan conocer los estadísticos de la señal y convergen más deprisa, al no tener que invertir grandes matrices. En cambio, son muy sensibles a los posibles errores de apuntamiento (debidos a una mala calibración de los elementos del *array* o a desviaciones en los algoritmos de cálculo de DoA), ya que si la dirección en la que se dirige el haz no coincide con la de la señal que se busca, ésta puede ser eliminada como una interferencia más.

Como ejemplo, cabe citar las técnicas MaxSNR (que maximiza la SNR a la salida del *array*, si se conocen ciertas estadísticas del ruido) y LCMV (que minimiza la varianza de la señal de salida, sujeta a algunas restricciones).

3. Técnicas con referencia ciega. En este tipo de técnicas, más complejas, lo que se explota es alguna característica conocida de la señal deseada, como alguna modulación, algún tipo de cicloestacionariedad, etc. Por ejemplo, podemos citar la técnica CMA, en la que se fuerza una solución de envolvente constante (interesante cuando la señal deseada verifique esta característica, como es el caso de UMTS, que utiliza modulación QPSK).

Habitualmente, el algoritmo que se diseña para una aplicación no se encuadra exactamente en uno de estos tipos, sino que puede utilizar algún esquema híbrido para mejorar sus prestaciones. Algunos algoritmos interesantes en este sentido pueden ser:

- ↓ Procesado haz-espacio. Consta de dos etapas, la primera genera una serie de salidas asociadas a distintos haces fijos, las cuales, en la segunda etapa, se procesan según algún criterio establecido.
- ↓ Conformador de banda ancha. Emplea una serie de líneas de retardo y de filtros FIR con los que sintetiza la respuesta en frecuencia de cada haz.
- ↓ Conformador en el dominio de la frecuencia. Realiza una FFT a la salida de cada antena, de modo que se calcula el vector de pesos óptimo para cada componente espectral. Después se realiza una IFFT para recuperar la señal de banda ancha.
- ↓ Conformador digital. Se toman muestras a la salida de cada elemento del array, que luego se procesan (mediante pesos y retardos) para generar los haces deseados. Si se desea conseguir precisión en el apuntamiento del haz, es necesario muestrear mucho más allá del criterio de Nyquist.
- ↓ Método de la auestructura. Separa los subespacios de señal deseada y ruido utilizando los autovectores de la matriz de autocorrelación de la señal recibida.

A la hora de utilizar alguno de estos algoritmos, ha de tenerse en cuenta que o es muy eficiente calcular la solución completa del problema cada vez que se desee actualizar el valor de w (normalmente, cada trama: 10 ms en UMTS). En lugar de eso, es más razonable limitarse únicamente a la diferencia entre el valor que tenía en el instante anterior y el valor que se supone que debería tener en el actual (algoritmos adaptativos). Un algoritmo así permitiría, además, capturar a nuevos usuarios y seguir a aquellos que se desplazan. Entre los más interesantes destacan el método del gradiente conjugado, la regla III de Madaline (basado en redes neuronales), el RLS (que necesita secuencia de entrenamiento) y el LS-CMA (una extensión no lineal del LS que aprovecha las características del algoritmo CMA).

Existen algunos algoritmos diseñados explícitamente para sistemas CDMA. Estos algoritmos deben ser adaptativos y multiusuario; es decir, deben extraer la señal de todos los usuarios simultáneamente (por lo tanto, calcularán uno distinto para cada usuario). Unos son con referencia ciega, pero los más sofisticados utilizan como referencia temporal los propios códigos de *scrambling* de cada usuario, para distinguirlos.

La aplicación de un algoritmo u otro en el diseño de una antena inteligente depende del compromiso al que se desee llegar entre la funcionalidad del método (robustez, flexibilidad, etc) y su eficiencia computacional (sobre todo, teniendo en cuenta que se desea conseguir que funcione en tiempo real). Esto es cierto cuando hablamos del enlace ascendente, sin embargo en el caso del enlace

descendente, existen además ciertas consideraciones que se deben tomar en cuenta.

El enlace descendente es la estación base la que transmite y por tanto debe conocer de algún modo la dirección en la que se encuentran los usuarios para poder iluminarlos con los haces. Puesto que en el enlace ascendente ya se han calculado los pesos que apuntan los haces hacia los usuarios, podría parecer razonable utilizar esos mismos pesos para conformar el diagrama de radiación en el enlace de bajada. Sin embargo, en general esto no será así, dado que las bandas de frecuencias para *uplink* y *downlink* están lo bastante separadas como para que la respuesta del canal sea distinta en cada caso y, por tanto, cambie la distribución espacial, temporal y espectro de los multitrayectos y de las interferencias.

En estos casos es más razonable utilizar algún algoritmo para calcular la DoA del usuario deseado y de los interferentes en el enlace ascendente, de modo que luego pueda sintetizarse un diagrama de radiación atendiendo a las posiciones calculadas de los usuarios. Esta síntesis puede hacerse entonces utilizando las técnicas clásicas de síntesis de diagramas de radiación.

CAPÍTULO 6 CONCLUSIONES

El presente documento se desarrolló en 2 partes, divididas en cinco capítulos: en la primera parte (capítulos 1, 2 y 3), se hizo una descripción detallada de las redes, en donde se pueden notar claramente tres tipos de redes: LAN (Redes de Área Local), WAN (Redes de Área Extensa) y redes Inalámbricas. Dentro de las redes se encuentran distintas topologías que nos indican la forma de interconexión de los equipos, éstos se puede llevar a cabo a través de distintos medios físicos, como son: alámbricos, cable coaxial, par trenzado, fibra óptica y los inalámbricos, como enlaces satelitales, microondas y radio, el empleo de determinado medio de transmisión depende de la distancia entre los dispositivos que se van a conectar, así como de la velocidad a la cual se quiere transmitir.

Se analizaron los principales modelos de referencia, los cuales indican la forma en que se estructura la información para ser transmitida, siendo los principales modelos el OSI y el TCP/IP, también se cuenta con los protocolos de comunicación, los cuales indican el lenguaje con el cual se comunican las máquinas, siendo el protocolo TCP/IP el que presenta mayor demanda y tiende a ser el único protocolo que se emplee en un futuro, debido a que es universal y por medio del cual se pueden comunicar equipos de distintos fabricantes y actualmente se utiliza para transmisión de voz en la telefonía celular.

Se explicaron los métodos de control de acceso al medio, los cuales nos indican de que forma se va a entregar la información de un equipo a otro dentro de la red, independientemente del tipo de red.

En el segundo capítulo ya teniendo el panorama general de las redes se procedió a describir un sistema de comunicaciones, que en este caso sería nuestro teléfono móvil, transmitiendo voz, datos y video, por lo cual es necesario tener conceptualizado, de manera general la codificación de fuente y de canal, codificación fuente, codificación de canal así como las normas de compresión de video y poder analizar este sistema como parte de una gran red, donde tendremos comunicación entre teléfonos móviles, teléfonos móviles y equipos de cómputo, lo cual llevará a tener una interconexión total de todos los equipos visualizándonos como una gran nube en donde se tendrá comunicación total entre todo tipo de dispositivo que sea capaz de transmitir información.

La segunda generación es de suma importancia, debido a que es la tecnología intermedia hacia la tercera generación (3G) de comunicaciones móviles y a partir de esta tecnología es como se podrá migrar hacia 3G. Es por eso que en el tercer capítulo se describió detalladamente la segunda generación de comunicaciones móviles, específicamente se habló de GSM, su arquitectura, sus enlaces de radio

y aspectos de la red, ya que es el estándar que se está adoptando en México y el empleado en gran parte del mundo.

En la segunda parte, se describieron todas las generaciones previas a 3G y el camino evolutivo hacia la UMTS, se pudo observar que se están siguiendo principalmente dos rutas de evolución, las cuales van a converger hacia un mismo estándar, este fenómeno sucede de esta forma debido a que al inicio de la telefonía móvil se adoptaron distintos estándares de comunicación, dependiendo de la ubicación geográfica y la idea principal de 3G es converger hacia un mismo estándar de comunicación, con lo cual se lograría el roaming internacional y con esto poder comunicar los teléfonos móviles de cualquier parte del mundo, este ideal está presentando ciertas dificultades debido a la gran diversidad de estándares que se manejan actualmente a lo largo de los distintos continentes.

En el capítulo 5 se analizó la tecnología empleada en 3G, que es un término genérico que cubre un rango de las tecnologías con una red que ofrece servicios inalámbricos, incluyendo cdma2000, UMTS, GPRS, WCDMA y EDGE. Se denotó el aumento en la velocidad de transmisión, se describe la manera en que estará constituida la red 3G así como aplicaciones específicas de dicha tecnologías. También se describió la arquitectura basada en DSP para comunicaciones 3G, la arquitectura programable OFDM-CDMA y las antenas inteligentes. Es importante describir las arquitecturas empleadas en 3G porque deberán existir cambios significativos en la tecnología, tanto del teléfono móvil, como en los nodos de red y protocolos.

Se puede observar que actualmente, la telefonía móvil e Internet captan el mayor interés dentro del mundo de las telecomunicaciones y la informática, y prueba de ello es el crecimiento experimentado en el número de usuarios que optan por utilizar estos dos servicios como se muestra en la figura 1. Así, Internet crece a un ritmo superior al 100% anual mientras que la telefonía móvil lo hace a un ritmo entre el 40-60%, cifras espectaculares frente al crecimiento de la telefonía fija que no supera en los países mas avanzados el 5 o 10%. A finales del año 2000 existían 650 millones de usuarios de telefonía móvil en todo el mundo y se prevé alcanzar los mil millones en el 2003. Con el 70% de la ocupación del mercado global empleando GSM, contra el 30% de CDMA y otras.

La adopción de los sistemas 3G ha sido lenta en sus inicios debido a los plazos involucrados en los procesos de licenciamiento y despliegue de las redes, como lo ilustra la figura 2. No obstante se prevé una adopción más rápida en países en desarrollo.

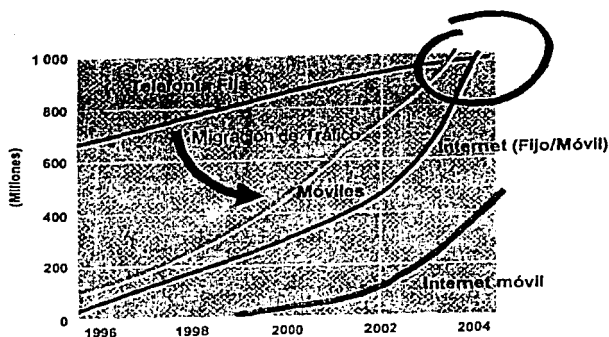


Figura 1. Crecimiento de usuarios

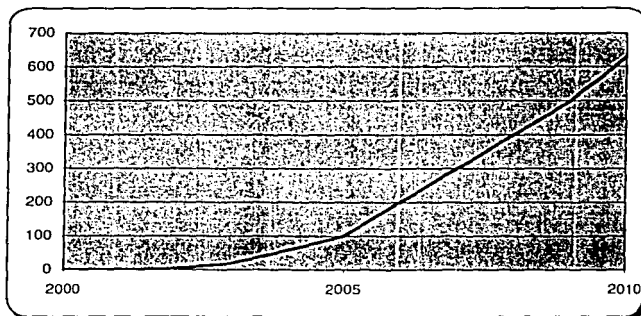


Figura 2. Suscriptores Mundiales de 3G

Pero este crecimiento, hasta el momento, no ha satisfecho las expectativas que se tenían, principalmente por las limitaciones de las redes de comunicaciones y la ausencia de servicios de acceso móvil. Por un lado, los sistemas de comunicaciones móviles de primera generación (analógicos) y segunda generación (digitales) no estaban concebidos para funcionar a nivel mundial sino más bien a nivel regional, y en el mejor de los casos, nacional. A esto se suma la gran variedad de estándares de comunicaciones móviles existentes hoy día, distribuidos por todo el mundo como GSM, CDMA, AMPS, TDMA y DECT. Estas tecnologías no son compatibles entre sí, por lo que la utilización de una única

terminal en estas redes es imposible, y así, el soñado roaming internacional no es posible en su totalidad.

Con lo cual se puede notar que la tercera generación de comunicaciones móviles no es lo que se había esperado, se requieren de grandes inversiones para poder implantar en su totalidad esta tecnología y no se han tenido los ingresos esperados para su adopción.

En este momento se despliegan sistemas de lo que se ha denominado generación 2.5 (HSCSD, GPRS, EDGE) que harán de puente entre los de segunda generación y la telefonía móvil de tercera generación. Esta última responde a un intento de estandarizar las comunicaciones móviles a nivel mundial, aunque ya están empezando a surgir pequeñas diferencias entre EEUU y el resto del mundo. Tiene como fin ofrecer grandes velocidades de conexión, por lo que se espera que se convierta en la forma más habitual de acceso a Internet. Se permite la transmisión de todo tipo de comunicaciones: voz, datos, imágenes, vídeo, radio, etc.

Hasta que la tercera generación se extienda (para lo que aún pueden quedar un par de años, los sistemas 2.5 supondrán un puente entre los de segunda generación y la UMTS).

En Europa, los operadores se están gastando grandes sumas de dinero en adquirir las licencias UMTS, con la esperanza de que será la tecnología que haga explotar las comunicaciones. Pero mientras esto ocurre, los que poseen sistemas 2G ya piensan en evolucionar a GPRS o EDGE.

Un aspecto muy importante es la capacidad de las baterías, debido a que al tener los móviles de tercera generación con las características antes mencionadas, existirá un gran consumo de energía, así como debe existir una mayor integración de los circuitos y el uso de antenas inteligentes. La manera en que se podrá tener acceso a Internet es haciendo una conexión a una central, misma que proporcionará el acceso hacia la red de redes. Los teléfonos móviles deberán contar con un software específico.

Considerando que la tercera generación de comunicaciones móviles se irá implantando poco a poco y que en algún momento se estará utilizando en conjunto tanto los móviles de segunda generación como los de tercera generación hasta que se emplee en su totalidad la tercera generación.

Cabe mencionar que ya se está planteando una cuarta generación de comunicaciones móviles en donde se podrá notar un gran incremento en las velocidades de transmisión. La velocidad es un factor muy importante a considerar, debido a que se requiere de una alta velocidad para establecer una videoconferencia, la cual es en tiempo real y para que no se note cierto retraso en el video, ya que se estará transmitiendo voz y video de manera simultánea y la

información debe llegar simultáneamente. Aunque resulta imposible eliminar por completo esto, debido a que siempre existirá cierto retraso en la transmisión aunque cada día será menos notable.

En la figura 3 se puede observar uno de los teléfonos de tercera generación que estará en el mercado a lo largo del mundo.

NTT DoCoMo, Inc. anunció que estará en el mercado el teléfono Lookwalk TM, el cual permite efectuar videoconferencias con otras terminales compatibles.

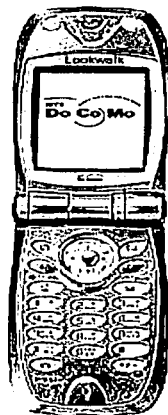


Figura 3. Teléfono 3G NTT DoCoMo

El nuevo Lookwalk P751v también proporciona acceso a Internet y servicio de correo electrónico. Una cámara integrada CMOS con 110,000 píxeles habilita la inmediata transmisión de imágenes fijas o en movimiento como archivos por medio de correo electrónico hacia otros teléfonos o computadoras. Lookwalk P751v viene con una pantalla de TFT LCD de 2.2", con 262,144 colores y puede acceder a aproximadamente 60,000 sitios de Web escritos en HTML versión 2.0 (los usuarios deben tener un contrato aparte para servicios de Internet como DoCoMo's).

Teléfono a color Cyon de LG (Modelo: LG-KH5000) con cámara para servicio CDMA 2000 1x EV-DO, viene con una pantalla de TFT LCD de 2.2", la pantalla es giratoria y cuenta con una cámara de alta calidad CMOS de 110,000 píxeles los

usuarios pueden utilizar la cámara integrada para tomar fotos sin ninguna restricción, estas fotos pueden ser enviadas vía correo electrónico, permitiendo libre intercambio y transmisión de fotos de alta calidad a cualquier hora y en cualquier parte. El móvil también soporta una velocidad de descarga de archivos de 2.4 Mbps y en un futuro proporcionará el servicio VOD (Videos Bajo Demanda).



Figura 4. Teléfono a color Cyon de LG (Modelo: LG-KH5000)

Con lo descrito podemos concluir que el objetivo de la tesis de dar a conocer un panorama general del paso evolutivo de las nuevas tecnologías de comunicaciones, partiendo desde el concepto de una red de computadoras, hasta llegar a las comunicaciones móviles inalámbricas conocidas como tercera generación, se ha cumplido.

Y a manera de conclusión final se presenta la situación actual tanto en México como en el resto del mundo.

Situación actual en el mundo

En España, Ericsson implementará más de 700 estaciones base, haciendo que el lanzamiento de la red UMTS de Telefónica Móviles sea una de las más grandes en el mundo. El contrato cubre un período de tres años y el lanzamiento comercial completo por parte del operador, dependerá de varios factores, incluyendo la cantidad disponible de terminales.

El contrato además confirma la posición de Ericsson como proveedor líder del Grupo Telefónica así como también su posición como proveedor líder de redes inalámbricas en España, Alemania y en el mundo. Ericsson tiene aproximadamente un 40% de participación en el mercado a nivel global tanto en GSM como en WCDMA. El valor del contrato global se estima en aproximadamente 400 millones de euros.

En Alemania, Ericsson como abastecedor principal, proveerá un proyecto llave en mano incluyendo provisión, instalación y puesta en marcha. Telefónica Móviles

opera en Alemania bajo la marca registrada Quam. Quam es un 57.2% de Telefónica Móviles y 42.8% de Sonera.

La posición de las empresas europeas Nokia y Ericsson es apoyar la evolución de GSM: GPRS en 2.5G, y Edge y W-CDMA en 3G. Sin embargo, ambas empresas aclaran que desarrollan todas las tecnologías. La explicación de su inclinación principal se debe a la velocidad y eficiencia con la que pueden transmitir voz y datos. Nokia indica que en el caso de CDMA los proveedores no cuentan con una solución real para llegar a 3G.

Empresas como AT&T Wireless decidieron moverse al estándar GPRS en octubre del 2000 porque están interesados en hacer un cambio más rápido para llegar a 3G.

En W-CDMA no habrá una central como las que hoy se conocen en el campo de la telefonía. Por medio del RAN la central virtual se encontrará en Internet.

La posición de Nortel es que son desarrolladores de todas las tecnologías y que en el caso de CDMA y GSM, la empresa canadiense no ve las ventajas tecnológicas sino la aceptación comercial, aunque GSM es un estándar semimundial y prevalece.

Todos los proveedores están de acuerdo en apoyar estándares abiertos para que pueda darse la complementación en infraestructura y así lograr que los operadores no se vean obligados a comprarle a un solo proveedor.

Si el futuro es Internet y esta red es la más abierta que se ha creado, el Internet móvil tiene que ser igual de abierto.

En el caso de la evolución de CDMA, existe ya la versión 1xrtt (2.5G), que tiene una velocidad de transmisión en datos de 300K en el sistema móvil, y ya se encuentra en operación con Telstra en Australia y Sprint en Estados Unidos. En el caso de México, Unefón se encuentra listo para la 1xrtt, igual que Telgua en Guatemala.

Los acuerdos básicos en modulación de 3 G indican que W-CDMA y CDMA2000 tienen el mismo estado básico de velocidad en la codificación y que las diferencias entre ambas tecnologías sólo son "una pelea por detalles".

Si bien Europa pudo estandarizar GSM, no ha ganado la batalla definitiva para que la evolución de esta tecnología se generalice por todo el mundo; sin embargo, cada vez más operadores se inclinan por este estándar de comunicación a nivel mundial.

La consultoría Analysys opina que los operadores virtuales de redes móviles (MVNO, por su acrónimo en inglés) de 3G se apresurarán a llegar a un mercado masivo y acelerarán los movimientos necesarios para acercarse más a esta generación de servicios móviles. Según la consultora, se espera que el mercado de suscriptores móviles en servicios de 3G sea de 480 millones para 2006 y de más de mil millones para 2010. Existiendo gran potencial para aquellos que se dedican a desarrollar contenido y servicios de comercio electrónico sobre el estándar de 3G.

Lo fundamental en 3G es la variedad y la calidad de los servicios que ofrecen.

Situación actual en México

Por parte de las operadoras en México, ya se empiezan a ver soluciones para que CDMA trabaje sobre GSM. Éste es el caso de Iusacell, compañía que ofrecerá a sus clientes la tarjeta inteligente R-IUM, un aditamento que estará disponible a finales del 2002 y que permitirá interoperabilidad entre las plataformas que se usan en México y las del continente Europeo.

Con tecnología de Tercera Generación, México estará en posibilidad de tener conexiones móviles más rápidas a Internet, liberándose de conexiones lentas, del equipo incómodo y de puntas de acceso a través de inmuebles. Así, 3G permitirá nuevas maneras de comunicarse, de tener acceso a la información, de dirigir un negocio y de aprender.

Esta vez fueron las diferencias políticas –y no tecnológicas– las que llevaron al mundo a estar nuevamente en desacuerdo. Y aunque esto no desatará una guerra nuclear ni mucho menos, los protagonistas están librando su batalla para conquistar un mayor territorio en el planeta.

El origen del conflicto es el estándar de comunicación para servicios móviles –3G– y los contendientes son bien conocidos como proveedores de infraestructura: Lucent, Ericsson, Nokia y Motorola. La solución, sin embargo, no puede tardar más de dos años.

La tercera generación (3G), como la llaman en la Unión Internacional de Telecomunicaciones (UIT), pudo ser ese remanso tecnológico en el que los continentes tuvieron un solo estándar de comunicación móvil. Todo indica, que este ideal no será alcanzado, y aquel sueño de viajar por todo el mundo con un solo teléfono pasará al olvido hasta nuevo aviso.

Este controvertido estándar permitirá la convergencia no sólo de voz, mensajes cortos y correo electrónico, sino de una serie de servicios que, poco a poco, cambiarán la forma de vivir, comunicarse y hacer negocios. Luego serán videos, compras, navegación y aplicaciones personalizadas.

Sin embargo, llegar a la 3G no es un proceso corto ni barato. Dotar a México de la infraestructura y las terminales (teléfonos) necesarias para prestar esos servicios requerirá una inversión multimillonaria, repartida entre los operadores y los usuarios.

Cambiar de 2G a 3G no se hace por decreto. Para empezar se debe abrir el proceso de licitación de las frecuencias, esta vez denominadas UMTS (Sistema de Telecomunicaciones Móvil Universal), que serán el vehículo para este nuevo modelo de comunicación móvil. Antes de esto, en México habrá que alcanzar el primer nivel, 2.5G, en sus versiones GPRS o CDMA2000 1x, que ya empieza a darse a conocer de este lado del Atlántico gracias al acuerdo entre Telcel y Ericsson.

Como se ve, el problema es el rezago tecnológico, porque todavía siguen vigentes las generaciones analógicas, que en otros países ya son cosa del pasado. De hecho, la era digital sólo tiene dos años en México, cuando en Europa cumplirá una década. Como quien dice, México sigue viviendo en el siglo pasado. Si bien es cierto que los proveedores aseguran que la tecnología ya se encuentra disponible, la verdad es que esta nueva era de comunicación no llegará hasta que las autoridades del sector (Secretaría de Comunicaciones y Transportes y Comisión Federal de Telecomunicaciones) liciten las frecuencias, las operadoras las compren y construyan e instalen la infraestructura necesaria. Un buen paquete para Telcel, Iusacell, Pegaso, Unefon y las que se interesen. ¿Alguien ha invitado a Telefónica Móviles de España?

El "cuándo" es la pregunta que abrumba a los impacientes usuarios mexicanos. "pronto", responden tanto autoridades como proveedores y operadores. La Cofetel argumenta que la demora en decidir las frecuencias que se utilizarán en México, obedece a su deseo de buscar armonía en el espectro radioeléctrico de la zona fronteriza. Esto evitaría interferencias y facilitaría la prestación transparente de servicios entre México y Estados Unidos. México como ha hecho en otras ocasiones, se pondrá en línea con la decisión de su vecino, independientemente de las recomendaciones de la ITU y la Comisión Interamericana de Telecomunicaciones (Citel).

A los fabricantes esta elección no parece preocuparles, porque, según afirman, tienen la tecnología y sólo necesitarán adaptarla a las frecuencias escogidas en cada país.

Guillermo Mairena, director de Redes de Trabajo de Nokia, explica que aún así, la mayoría de los proveedores tiene alguna preferencia por las tecnologías de 3G: Nokia favorece a GSM, W-CDMA y Edge, que son las que tienen vigencia actual en Europa. El funcionario afirma que en México se ha exagerado el potencial de 3G, y que algunos anuncios de estos servicios crean expectativas que no se alcanzarán. "Aquí hay barreras tecnológicas, psicológicas y de uso -agrega-

pues los usuarios creen que encontrarán la panacea a sus problemas o su sueño de comunicación”.

Para Ericsson, uno de los principales proveedores de infraestructura de 2G, impulsor de W-CDMA y dueño de patentes para desarrollo de sistemas CDMA2000, no existe batalla alguna porque sus soluciones dan como resultado un negocio de ganar-ganar. Al respecto, Raúl Lucido de la Parra, vicepresidente de Planeación Estratégica y Mercadotecnia de Ericsson México, asegura que las diferencias no son tan pequeñas como algunos operadores lo quieren hacer ver.

El tipo de servicios que se ofrecerán a los usuarios dependerá de la velocidad de transmisión de información – la cual varía de acuerdo a la tecnología que se implemente– y de la facilidad de homogeneizar las bandas en una sola frecuencia, por lo menos para que las terminales funcionen en cualquier parte si se llega al acuerdo de establecer un solo estándar en el mundo.

Motorola, en la voz de Javier Lizárraga, director de soluciones y cuentas clave de telecomunicaciones globales, señala que la empresa está desarrollando productos en ambos estándares (W-CDMA y CDMA2000), con la vista fija en el concepto de integración. Esta compañía ya tiene productos en el mercado 2.5G, los cuales se venden en Europa.

Victor Cervantes, director general de Nortel México, espera que la Cofetel no seleccione una tecnología demasiado específica porque eso provocaría que algunos proveedores tuvieran menor volumen de distribuidores, ocasionando que el usuario pague más. En cuanto a otras tecnologías como TDMA, indica que ésta no se encuentra moribunda, pero tampoco se desarrolla ni crece. Telcel es uno de los operadores que todavía la utilizan.

Sea como fuere, no hay mucho por hacer. En tanto la Cofetel no abra las subastas, todo serán meras especulaciones. Y cuando finalmente ocurra, todavía habrá que ver cuánto tiempo tomarán las operadoras para ofrecer los servicios al público. Será muy bueno entonces, pero por ahora hay que agradecer que los sencillos celulares digitales completen exitosamente una llamada.

BIBLIOGRAFIA

- [1] Ian Glover, Peter Grant;
Digital Communications;
Ed. Prentice Hall, 1998;
pp. 661-666
- [2] Fred J McClimans;
Communications Wiring and Interconnection;
Ed. Mc. Graw Hill, 1992;
pp. 117, 42, 43, 56
- [3] Fred Halsall;
Comunicación de Datos, Redes de Computadoras
Sistemas Abiertos;
Ed. Addison Wesley, 1998;
pp. 13-18, 273-277, 298-302, 305, 358, 450
- [4] Gilbert Held;
Data Communications Networking Devices;
Ed. Wiley, 4a ed., 1998;
pp. 202-204
- [5] Guy Yost;
Aprendiendo Netware 4.1;
Ed. Prentice Hall, 1997;
pp. 37
- [6] Microsoft Corporation;
Networking Essentials;
Course 578, 1996;
pp. 551,555,563,573, 589
- [7] <http://www.isa.cie.una.es/proyectos/codec>
- [8] Rodger E. Ziemer, Roger L. Peterson;
Introduction to Digital Communication;
Macmillan Publishing Company; 1992;
pp. 5-7, 328-329
- [9] Lajos Hanzo, Peter J. Cherman, Jürgen Street;
Wireless Video Communications, Second to Third
Generation and Beyond;
IEEE Series on Digital & Mobile Communication;
pp. 9-14, 357-381,650, 660

- [10] Bernard Sklar;
Digital Communication Fundamentals an Applications;
Prentice Hall, 1998;
pp. 279-274, 288, 298-300, 345, 385
- [11] <http://www.itu.int>
- [12] David M. Balstor;
Cellular Radio System;
Artech House, 1993;
pp. 7-11
- [13] Michel Mouly, Marie-Bernadette Pautet;
The GSM System for Mobile Communications;
Published by the authors, 1992;
63-71
- [14] Jan A. Andestad;
Network Aspects of the GSM System;
EUROCON 88; Junio 1988
- [15] Moe Rahnema;
Overview of the GSM system and protocol architecture;
IEEE Communications Magazine, April 1993
- [16] <http://www.gsmworld.com>
- [17] <http://www.cibertele.com/publicaciones/umts.pdf>
- [18] <http://www.mobiledataevolution.com/gsm.asp>
- [19] <http://www.mobiledataevolution.com/hscsd.asp>
- [20] <http://www.mobilegprs.com>
- [21] <http://www.mobledataevolution.com/edge.asp>
- [22] <http://www.ericsson.com/technology>
- [23] <http://3g.com>
- [24] <http://www.cdg.org/3gpavilion>
- [25] <http://www.hpl.hp.com/techreports/2001/HPL-2001-146.pdf>
- [26] <http://buffy.eecs.berkeley.edu/~linnartz/MCCDMA.htm!>

- [27] Alan Gatherer, Trudy Dtetzler, Mike Maman Edgar Auslander;
Arquitectura basada en DSP para comunicaciones móviles;
IEEE Communications, Vol. 38, No1;
Enero 2000; pp. 84-89

- [28] Kwang-Chen, Shan-Tsung Wu;
Arquitectura programable para OFDM-CDMA;
IEEE Communications, Vol. 37, No11;
Noviembre 1999; pp 76-82

- [29] P.H. Lehne, M. Petersen;
An overview of smart antenna technology for mobile
Communications Systems. IEEE Communications, vol. 2 no4,
Oct./Nov/Dic. 1999; pp. 2-12

- [30] L.C. Godara;
Applications of antenna arrays to mobile communications;
IEEE Communications, Vol. 85, No7;
Julio 1999; pp. 31-35

- [31] J. C. Liberti, T.S. Rappaport;
Smart antennas for wireless communications : Is-95 and
Third Generation CDMA applications;
Prentice Hall, 1999;
pp. 253-257, 315-338

WHITE PAPERS

- [1] The new path from GSM to UMTS "Soft Evolution" of mobile networks;
Agosto 2001; Siemens.

- [2] Evolution of mobile wireless network; Diciembre 2000; Motorola

- [3] Data on 3G, An Introduction to third generation; Febrero 2000; Lifestreams

- [4] Yes "2G"; Febrero 2001; Mobile Streams