



UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

UIT-T

SECTOR DE NORMALIZACIÓN
DE LAS TELECOMUNICACIONES
DE LA UIT

Apéndice I
Rec. G.764

(11/95)

**ASPECTOS GENERALES DE LOS SISTEMAS
DE TRANSMISIÓN DIGITAL**

GUÍA DE PAQUETIZACIÓN

**Apéndice I a la
Recomendación UIT-T G.764**

(Anteriormente «Recomendación del CCITT»)

PREFACIO

El UIT-T (Sector de Normalización de las Telecomunicaciones) es un órgano permanente de la Unión Internacional de Telecomunicaciones (UIT). Este órgano estudia los aspectos técnicos, de explotación y tarifarios y publica Recomendaciones sobre los mismos, con miras a la normalización de las telecomunicaciones en el plano mundial.

La Conferencia Mundial de Normalización de las Telecomunicaciones (CMNT), que se celebra cada cuatro años, establece los temas que han de estudiar las Comisiones de Estudio del UIT-T, que a su vez producen Recomendaciones sobre dichos temas.

La aprobación de Recomendaciones por los Miembros del UIT-T es el objeto del procedimiento establecido en la Resolución N.º 1 de la CMNT (Helsinki, 1 al 12 de marzo de 1993).

El Apéndice I a la Recomendación UIT-T G.764 ha sido preparado por la Comisión de Estudio 15 (1993-1996) del UIT-T y fue aprobada por el procedimiento de la Resolución N.º 1 de la CMNT el 13 de noviembre de 1995.

NOTA

En este Apéndice, la expresión «Administración» se utiliza para designar, en forma abreviada, tanto una administración de telecomunicaciones como una empresa de explotación reconocida de telecomunicaciones.

© UIT 1996

Es propiedad. Ninguna parte de esta publicación puede reproducirse o utilizarse, de ninguna forma o por ningún medio, sea éste electrónico o mecánico, de fotocopia o de microfilm, sin previa autorización escrita por parte de la UIT.

ÍNDICE

	<i>Página</i>
I.1	Introducción 1
I.2	Antecedentes históricos..... 1
I.3	Resumen de aspectos de diseño 2
I.4	Reconstrucción de las señales de habla..... 2
I.5	Igualación del retardo..... 2
I.5.1	Retardo ciego..... 3
I.5.2	Indicación temporal absoluta..... 3
I.5.3	Indicación temporal relativa 4
I.6	Resistencia a los errores 4
I.6.1	Sensibilidad del habla a los errores de bits..... 5
I.6.2	Cálculo de la CRC sobre parte de la trama..... 5
I.6.3	Cálculo de la CRC en la totalidad de las tramas..... 5
I.7	Control de la congestión..... 5
I.7.1	Control global..... 6
I.7.2	Control local..... 6
I.8	Pérdida de paquetes..... 8
I.8.1	Caso de señales de voz 8
I.8.2	Caso de datos en banda vocal..... 9
I.8.3	Estrategias de relleno (voz) 9
I.9	Elección del tamaño de los paquetes..... 10
I.9.1	Consideraciones sobre el habla 10
I.9.2	Consideraciones sobre los errores de bits..... 10
I.9.3	Tráfico integrado 10
I.9.4	Algoritmo de la Recomendación G.764 11
I.10	Aspectos de compresión..... 11
I.10.1	Algoritmos de codificación del habla..... 11
I.10.2	Interpolación digital del habla 11
I.11	Señalización orientada al canal 12
I.12	Ampliaciones..... 12
I.12.1	Facsímil 12
I.12.2	Sincronización voz/vídeo 13
I.12.3	Interfaz entre la RTPC y las LAN 13
I.12.4	Ampliación a nuevos algoritmos 14
I.13	Resumen..... 14
Referencias.....	14

Guía de paquetización

(Ginebra, 1995)

(Este apéndice no es parte integrante de la presente Recomendación)

I.1 Introducción

En este apéndice se resume la situación actual de la paquetización del habla en la Comisión de Estudio 15 del UIT-T en el periodo de estudios 1992-1996 y en la CE XV del CCITT en el periodo de estudios 1988-1992. Esta situación puede cambiar en el futuro.

En este apéndice no se contemplan las siguientes cuestiones:

- 1) los distintos tratamientos dentro de la red que pueden darse a los paquetes según su asignación a clases prioritarias. No obstante se considera, en general, que la red debería dar prioridad al habla sobre los datos digitales para reducir el retardo y la variabilidad del retardo y para independizar el tráfico de ráfagas del tráfico en tiempo real;
- 2) la calidad de servicio para distintas clases de tráfico.

Este apéndice tiene como finalidad:

- explicar los distintos aspectos que afectan a la paquetización del habla;
- proporcionar una visión general de las técnicas y consideraciones correspondientes al transporte de la voz paquetizada utilizadas en la Recomendación G.764;
- difundir información sobre asuntos de interés a los diseñadores y realizadores de equipos de voz paquetizada y a los proveedores de servicio que los utilizan.

I.2 Antecedentes históricos

Tradicionalmente se han establecido los servicios de voz en la red telefónica pública con conmutación (RTPC) (denominada también red de área extendida o WAN, *wide area network*) empleando un método orientado a circuitos. El progreso de las técnicas de transporte de paquetes [por ejemplo la Recomendación X.25/X.75, internet, tecnología de paquetes de banda ancha, retransmisión de tramas y modo de transferencia asíncrona (ATM)], ha incentivado la investigación de nuevas técnicas para el transporte del habla.

Los sistemas paquetizados pueden aprovechar la naturaleza discontinua del tráfico para multiplexar distintos tipos de tráfico (por ejemplo voz, datos, vídeo) procedente de diversos usuarios de forma que éstos puedan compartir de una manera dinámica los recursos de conmutación y de anchura de banda de transmisión. La paquetización facilita la integración de los diferentes tipos de tráfico a fin de conseguir una utilización más eficaz de los recursos de conmutación y de anchura de banda disponibles. La paquetización proporciona más flexibilidad que los métodos orientados a circuitos debido a que la cabecera del paquete contiene la información de control necesaria que identifica, por ejemplo, el tipo de tráfico y, cuando proceda, el método de codificación.

Los trabajos relativos a la paquetización de la voz comenzaron en el CCITT a mediados del periodo de estudios 1984-1988 en el Grupo de Trabajo XVIII/8 prosiguiendo, durante el periodo de estudios 1989-1992, en el Grupo de Trabajo XV/2. Los trabajos continúan actualmente en el UIT-T para paquetes de banda ancha y redes ATM con el objetivo de proporcionar una base homogénea de paquetización de la voz con o sin compresión del habla e interpolación del habla para facilitar el interfuncionamiento de equipos procedentes de diversos fabricantes en aplicaciones de telecomunicación.

Los trabajos desarrollados en el CCITT han fructificado en el protocolo de paquetización vocal de la Recomendación G.764 y sus ampliaciones en la Recomendación G.765. Estos protocolos son compatibles en la capa de enlace con los protocolos LAPD y LAPF de la RDSI especificados en las Recomendaciones Q.921 y Q.922, respectivamente.

I.3 Resumen de aspectos de diseño

Utilizando como referencia la torre de protocolo OSI, los aspectos principales de diseño para el desarrollo de un protocolo de paquetización del habla son:

- 1) Capa 1 (capa física) – La característica es si la interfaz física se ajustará al de las redes telefónicas públicas (Recomendaciones G.703/G.704) o al de otras redes de área local tales como las IEEE 802.2, 802.3 u 802.9, etc.
- 2) Capa 2 (capa de enlace) – Algunos de los aspectos son:
 - a) si la capa lógica será compatible con los protocolos RDSI (LAPD/LAPF) o tendrá la misma estructura que la correspondiente a la LAN;
 - b) como afrontar la pérdida de tramas;
 - c) resistencia frente a los errores.
- 3) Capa 3 (procedimientos que deben utilizarse con los tráficos de voz digitalizada y de datos en banda vocal) – Los aspectos son:
 - a) la variabilidad del retardo de los paquetes del habla;
 - b) el transporte de la señalización asociada al canal.
- 4) Aspectos de capas superiores que afectan al codificador del habla y al tipo de compresión utilizada.

Un protocolo de paquetización del habla tiene los siguientes requisitos:

- En el extremo receptor debe reconstruirse la señal del habla a partir de paquetes que llegan a intervalos irregulares (o, en algunas arquitecturas, desordenados).
- El protocolo debe ser resistente frente a los errores de línea.
- Debe ofrecer un método sencillo para el control de la congestión de la red.
- Debe especificar procedimientos en el extremo de terminación para corregir la pérdida de paquetes o los retardos excesivos.
- Debe transportar la señalización asociada al canal.
- Si se utiliza la interpolación de la voz para eliminar los intervalos de silencio deberá especificar el nivel al que se reinyecta el ruido en el extremo terminal.

I.4 Reconstrucción de las señales de habla

Para conseguir una buena calidad de voz, el extremo terminal debe reconstruir un flujo de bits de voz continuo y entregarlo a intervalos regulares a pesar de la variación de los tiempos de llegada de los paquetes. Esto requiere:

- 1) conservar la temporización relativa de la información dentro de una ráfaga de voz; y
- 2) igualación del retardo.

En la Recomendación G.764 se utiliza un número secuencial de paquete para codificar la temporización relativa de la información dentro de una ráfaga de voz. El primer paquete de una ráfaga de voz tiene siempre el número secuencial 0 y los paquetes subsiguientes de la misma ráfaga tienen números comprendidos entre 1 y 15, con repliegue a 1. Los puntos extremos terminales utilizan el número secuencial del paquete para:

- 1) determinar el primer paquete de una ráfaga de voz; y
- 2) detectar la pérdida de paquetes. La determinación del primer paquete es útil para la igualación del retardo y puede ser necesaria en algunos algoritmos de codificación del habla tales como el de la Recomendación G.728. En la cláusula siguiente se trata de la igualación del retardo.

I.5 Igualación del retardo

Los retardos de la comunicación por paquetes tienen dos componentes: retardo fijo y retardo variable [1]. El retardo fijo se debe a la propagación de la señal por los enlaces de transmisión así como a los retardos de tratamiento fijos en los puntos extremos de origen y de terminación y dentro de la red. Los efectos de las variaciones del retardo de propagación para un trayecto determinado se suponen despreciables.

En la paquetización del habla los retardos de tratamiento fijos constan de las siguientes componentes:

- 1) retardo de paquetización durante el cual se almacenan las muestras del habla para su tratamiento ulterior;
- 2) si se utiliza la interpolación del habla para suprimir intervalos de silencio, tiempo de conmutación del detector del habla [2];
- 3) retardo algorítmico de extremo a extremo debido a la codificación y a la decodificación del habla – Este retardo depende del método de codificación; por ejemplo es igual a 125 μ s para la PCM, 250 μ s para los algoritmos de modulación por impulsos codificados diferencial adaptativa (MICDA) de las Recomendaciones G.726 y G.727 siendo inferior a 2 ms para el algoritmo de predicción lineal con excitación por código de bajo retardo (LD-CELP, *low delay code excited linear predictor*) de la Recomendación G.728;
- 4) cualquier retardo añadido en el extremo terminal para compensar la fluctuación de temporización debida a la variabilidad del retardo – Este retardo añadido se denomina restitución.

Los retardos variables se deben en principio a la puesta en cola y tratamiento de los paquetes. Dependen de las características de la ruta de cada paquete, del número de saltos (nodos), del tipo y velocidad de cada enlace y de la intensidad de tráfico.

El tráfico de voz requiere un retardo reducido y uniforme. En la Recomendación G.114 (1993) se examinan los efectos de los retardos de extremo a extremo sobre la calidad de una conversación. Las variaciones del retardo que pueden ser aceptables para la transmisión de datos digitales afectan generalmente a las pausas entre palabras y sílabas por lo que son perjudiciales para la conversación vocal. Los datos disponibles sugieren que la variación del intervalo entre ráfagas de voz debe ser inferior a 200 ms para evitar la degradación subjetiva de la calidad del habla. En aplicaciones tales como la videotelefonía en la que se transmiten informaciones de vídeo y de audio que deben estar sincronizadas, ha de tenerse también en cuenta el efecto del retardo variable sobre esta sincronización.

Hay varios métodos para compensar la variabilidad del retardo en la red. Tales técnicas comprenden:

- 1) el retardo ciego;
- 2) la indicación temporal absoluta; y
- 3) la indicación temporal relativa. Estos métodos producen un incremento del retardo de almacenamiento efectivo de extremo a extremo y, en consecuencia, del retardo total de extremo a extremo.

I.5.1 Retardo ciego

En el método del retardo ciego se añade siempre, en el extremo terminal, un retardo de almacenamiento fijo al primer paquete de una ráfaga de voz. Este retardo se corresponde con el máximo retardo variable esperado.

La ventaja del método del retardo ciego es su sencillez, lo que resulta ventajoso cuando la voz transmitida es tal que los retardos variables son del orden de una fracción de milisegundo (por ejemplo redes de área local o redes de banda ancha a 150 Mbit/s). En estos casos, resulta adecuado un retardo de restitución fijo del orden de 10 ms para eliminar la fluctuación de retardo de extremo a extremo [3].

Para conexiones de larga distancia por la RTPC, este método puede exigir retardos excesivos, de manera que el retardo total de extremo a extremo rebasaría los límites de calidad de funcionamiento para retardos de red estipulados en la Recomendación G.114. Por ejemplo, si el primer paquete ya ha sufrido la variación del retardo más desfavorable, el retardo variable total que debe añadirse será igual al doble de este valor más desfavorable [1]. Debido a que este método no compensa totalmente la variabilidad del retardo, pueden ser variables las pausas entre palabras, lo que puede degradar la calidad subjetiva del habla. Puede perturbarse el tráfico en la banda de frecuencias vocales incluido el facsímil demodulado. Valores de retardo total superiores a 500 ms pueden provocar desconexiones prematuras de llamadas facsímil sobre todo cuando hay eco [4].

I.5.2 Indicación temporal absoluta

Este es el método que se utiliza en las redes de datagramas. La cabecera del paquete incluye un campo para una indicación temporal que representa tiempo real. La indicación temporal tiene una resolución suficiente para permitir la detección precisa de la fluctuación de paquetes y abarcar el valor más desfavorable del tiempo de tránsito de un paquete por la red. Por consiguiente, los paquetes que llegan desordenados pueden ordenarse correctamente y almacenarse en el receptor utilizando la información de indicación temporal [5]. Este enfoque de datagramas exige también la sincronización del reloj entre el transmisor y el receptor de forma que el retardo de cada paquete entrante pueda compararse con los anteriores suponiendo que el retardo de la red es fijo.

I.5.3 Indicación temporal relativa

En el método de la indicación temporal relativa, se obtiene una estimación del tiempo de reproducción para el primer paquete de una ráfaga de voz, para todos los paquetes de señalización y para el primer paquete que siga a un paquete perdido. Se utiliza este tiempo para ajustar el retardo de los paquetes restantes de esta ráfaga transmitidos por ese circuito virtual.

El retardo variable acumulado correspondiente a un paquete se almacena en el campo de indicación temporal de la cabecera del paquete [1]. Cada nodo de red agrega a esta indicación temporal el tiempo que emplea en servir un paquete antes de enviarlo, utilizando como referencia su reloj local.

Se denomina restitución al retardo variable admisible máximo de un circuito virtual. Se define la restitución para un circuito virtual determinado. Una vez que se ha efectuado la estimación del tiempo de reproducción, se sitúan los paquetes según el orden de su número de secuencia en el almacén de reproducción y se mantienen durante la siguiente duración:

$$\text{tiempo antes de la reproducción} = \text{retardo de restitución} - \text{valor de la indicación temporal}$$

El extremo terminal debe, en consecuencia, almacenar los paquetes de voz que llegan antes de su tiempo de reproducción programado y reproducirlos a intervalos regulares. Los paquetes cuyo campo de indicación temporal rebasa el retardo de restitución se consideran tardíos y se abandonan.

El método de la indicación temporal relativa es más simple que el método de la indicación temporal absoluta I.5.2 cuando se emplean circuitos virtuales para asegurar que los paquetes permanecen en secuencia. Este método no depende de la sincronización de reloj entre los puntos extremos y sitúa la función de temporización en la capa 3. Además puede utilizarse la medición del retardo de cada paquete para detectar la congestión en la red e invocar seguidamente estrategias de gestión de la sobrecarga.

En el caso de una conferencia entre múltiples ubicaciones, la introducción del retardo de restitución mejora la dinámica conversacional porque asegura la sincronización de la reproducción del habla de todas las partes.

Una desventaja del método de la indicación temporal relativa es la complejidad añadida a los nodos de la red. Cada nodo de red debe distinguir entre paquetes de voz y de datos, actualizar el campo de indicación temporal y recalcular la CRC. Otra desventaja es que el retardo de extremo a extremo se incrementa siempre en la restitución, incluso cuando no hay congestión de red. Por último, en comparación con el método del retardo fijo, el campo de indicación temporal añade alguna tara de protocolo.

En general, la selección del valor para la restitución es un compromiso entre la aceptación de retardos excesivamente largos y la eliminación de muchos paquetes. Como un circuito virtual emula la conexión del circuito físico, el retardo de restitución, típicamente, es más alto para el tráfico de datos digitales que para el tráfico en banda vocal. Para una conexión punto a punto los estudios de simulación han puesto de manifiesto que, para la velocidad primaria, un retardo de restitución de 10 ms asegurará que se eliminará menos de la décima parte del 1% de los paquetes de voz en un nodo determinado a causa de un retardo excesivo, incluso para cargas de datos que alcancen el 60 al 70% de la anchura de banda ofrecida [6]. Para permitir cierto margen de seguridad puede fijarse el retardo de restitución en torno a 20-30 ms por salto. Pueden utilizarse cifras más conservadoras (por ejemplo 70 ms) sobre todo cuando en el trayecto existe demodulación/remodulación facsímil, a fin de tener en cuenta el tiempo de procesamiento de la demodulación facsímil y las restricciones de temporización impuestas por el protocolo T.30.

El retardo de restitución se añade solamente una vez en el extremo de terminación. En consecuencia, si se conectan adosadas dos redes G.764 el retardo se añade en el último nodo únicamente. Esta propiedad ayuda al planificador/operador de la red a mantener el retardo unilateral global dentro del límite de 400 ms establecido por la Recomendación G.114.

La Recomendación G.764 establece el valor máximo de la restitución en 200 ms para tener en cuenta el tráfico en banda vocal y el tráfico de datos digitales. Los enlaces por satélite introducen un retardo unilateral igual a 250 ms, en tanto que las extensiones terrenales agregan retardos de unos 50 ms. En consecuencia, cuando se utiliza un enlace por satélite de un solo salto una restitución de unos 50 ms puede producir un retardo unilateral total de unos 350 ms con un porcentaje de pérdidas de paquetes muy reducido. Para el habla la restitución se limita generalmente a unos 70 ms para un trayecto de tres o más nodos o cuando se utiliza la demodulación facsímil de la Recomendación G.765.

I.6 Resistencia a los errores

La resistencia a los errores implica dos aspectos:

- 1) robustez del codificador de voz; y
- 2) robustez del protocolo.

En esta cláusula se examina la cuestión de como conseguir que el protocolo sea más robusto frente a los errores de línea. En particular, se explica porque es mejor calcular la CRC sobre parte de una trama en vez de sobre toda la trama. De esta forma no se eliminará la trama si se producen errores de bits en el campo no protegido. Debe observarse que los nuevos algoritmos de codificación pueden ser más resistentes a los errores de ráfagas (por ejemplo borrado de tramas o pérdidas de paquetes) que la MIC o la MICDA.

Debe contemplarse con precaución la utilización de canales paquetizados de 64 kbit/s sin restricciones para algunas clases de servicios RDSI de banda estrecha tales como los videoteléfonos en presencia de errores. El sistema de paquetes transforma errores aleatorios en ráfagas de errores [7].

I.6.1 Sensibilidad del habla a los errores de bits

Los errores de línea son una fuente de pérdida de paquetes (denominada también *corrupción de tramas* en las comunicaciones sin hilos). El tráfico de voz es más sensible a las pérdidas de paquetes que a los errores de bits [8], [9]. Por consiguiente, es preferible entregar un paquete con algunos bits deteriorados que descartarlo o intentar recuperar los bits originales por retransmisión. Esto se aplica al caso de la voz codificada con MIC así como a la voz comprimida si un error de bit en una muestra de voz tiene poca influencia sobre la calidad de reproducción del habla. Este hecho ha quedado ya bien establecido para el habla transportada en redes HDLC [10] y redes TCP/IP [11].

I.6.2 Cálculo de la CRC sobre parte de la trama

La Recomendación G.764 ha introducido una nueva trama HDLC denominada «información no numerada» con una verificación de cabecera o UIH para limitar el cálculo de la CRC al campo de dirección (para asegurar una entrega correcta), al campo de control (para garantizar la validez del tipo de trama) y a la cabecera de la capa 3. La ISO ha generalizado el ámbito de las aplicaciones a todos aquellos casos en los que la entrega puntual de la información es más importante que la integridad de la información que se transfiere (por ejemplo voz paquetizada o transferencia de información actualizada periódicamente o aplicaciones vídeo) [12]. La longitud de los bits protegidos se determina mediante negociación en forma de parámetro de sistema por intercambio de tramas XID en el establecimiento del enlace o por administración.

I.6.3 Cálculo de la CRC en la totalidad de las tramas

Algunas implementaciones especiales de voz paquetizada prefieren el empleo del método tradicional de cálculo de la CRC en la totalidad de la trama. Esto tiene una influencia mínima sobre los dispositivos HDLC que carecen de la capacidad de cálculo de la CRC sobre parte de la trama.

La desventaja de realizar el cálculo de la CRC sobre toda la trama es que cualquier error de bit provocará la eliminación de la totalidad de la trama. En este caso, la aplicación de la voz paquetizada se restringe a situaciones en las que la tasa de errores es muy reducida. Esta limitación excluirá la aplicación del protocolo en muchas redes públicas y privadas en todo el mundo, a menos que se acepte cierta degradación.

I.7 Control de la congestión

La congestión puede producirse por la naturaleza estadística del tráfico o por acciones iniciadas por la red. El método más utilizado habitualmente para el control del flujo de la señal de voz es el del «bloqueo de la llamada», es decir impedir la iniciación de nuevas llamadas en periodos de ocupación. Los mecanismos de control de flujo orientados a datos no se adaptan a las redes de voz debido a que no puede obligarse a los locutores a que reduzcan el ritmo de su conversación. Además, las pausas largas o retardos variables debidos a las acciones de control de flujos orientadas a datos pueden reducir la inteligibilidad de la voz [13], [14].

En redes de paquetes integradas puede haber dos tipos de técnicas de control de la congestión:

- Control global – La congestión se resuelve mediante procedimientos a escala de sistema.
- Control local – El nodo congestionado resuelve la congestión localmente reduciendo la carga ofrecida ya sea mediante el acortamiento de la longitud de los paquetes o reduciendo el número de éstos. El tipo de tratamiento depende de si la longitud del paquete es variable o fija (como en los sistemas ATM) y de si el codificador de la voz puede sustentar pérdidas de bloques o pérdidas de paquetes.

I.7.1 Control global

Con este método se evita el establecimiento de nuevas llamadas. En los equipos de multiplicación de circuitos, con independencia de la tecnología utilizada, puede conseguirse el control mediante la comunicación con la central de cabecera según las reglas del «control de carga dinámico» definido en los Anexos A/Q.50 o B/Q.50. En las redes privadas puede utilizarse para ejecutar el control de la llamada la interfaz entre una PBX y el equipo de paquetización.

I.7.2 Control local

El control local de la congestión puede realizarse abandonando elementos de un paquete denominados «bloques» o, alternativamente, eliminando paquetes completos. El efecto de eliminar bloques de un paquete es el de reducir la longitud del paquete, en tanto que el efecto de abandonar paquetes es reducir el número de paquetes que deben transmitirse. La reducción resultante de las memorias intermedias de voz libera cierta anchura de banda para tráfico no compresible al costo de una degradación controlada de la calidad de voz [6]. Por consiguiente, el control local depende de la aptitud del codificador de voz para funcionar con menos bits y/o soportar pérdidas de paquetes.

El control local permite que cada nodo funcione de forma independiente, de manera que no tiene que coordinarse con otros nodos en la dirección regresiva. Esto permite una respuesta instantánea a la congestión del tráfico en cualquier nodo que se adapta a las características del tipo de tráfico sin modificaciones explícitas de los mensajes de control entre todos los nodos del trayecto de los circuitos virtuales.

I.7.2.1 Algoritmos jerarquizados con bits eliminables

Los algoritmos con bits eliminables constituyen una clase de algoritmos cuyas palabras código están formadas por unos bits de núcleo o bits primarios y algunos bits de ampliación o bits secundarios. Los bits primarios se utilizan en el codificador y en el decodificador en tanto que los bits secundarios se utilizan para reducir el ruido de cuantificación en la señal reconstruida. Los bits primarios deben llegar al decodificador para evitar errores de seguimiento mientras que los bits secundarios pueden, en su caso, suprimirse acortando de este modo los paquetes para reducir la congestión.

Los métodos basados en la eliminación de bits proporcionan una forma flexible de reducción de la congestión en cualquier nodo de la red sin necesidad de intercambiar mensajes de control en el trayecto hacia atrás de la conexión (hacia el extremo de origen). Existen actualmente dos algoritmos normalizados de eliminación de bits, la MIC de la Recomendación G.711 y la MICDA de la Recomendación G.727. Este último se denomina también algoritmo jerarquizado porque existe un bucle de realimentación que enlaza el codificador con el decodificador. En este caso, se construye el decodificador de forma que los niveles del cuantificador para las velocidades mayores incluyen los niveles del cuantificador para velocidades reducidas. Por consiguiente pueden eliminarse los bits secundarios cuando exista congestión vocal sin modificar el cuadro de cuantificación [15].

Algunas propuestas recientes han puesto de manifiesto que puede lograrse una codificación jerarquizada en la CELP [16], [17], [18], [19], [20], [21]. Una versión inicial de un codificador LD-CELP jerarquizado deducida de la Recomendación G.728 funciona a 8 kbit/s, 14 kbit/s y 20 kbit/s y puede proporcionar una calidad telefónica para su velocidad de funcionamiento más elevada [22]. Este codificador tiene un retardo de procesamiento total de unos 12-18 ms y es resistente frente a errores de canal para una tasa de errores de bits media igual a 10^{-4} .

El inconveniente de los métodos de codificación jerarquizada es que introducen cierta degradación en la calidad de la voz en comparación con los algoritmos no jerarquizados que funcionan a la misma velocidad de bits. Sin embargo, con un buen diseño del codificador puede hacerse imperceptible la diferencia de calidad subjetiva entre un codificador jerarquizado y el codificador de velocidad fija correspondiente (por ejemplo, Recomendaciones G.727 y G.726, funcionando ambos con 4 bits/muestra). Otra desventaja es que a medida que se reduce la velocidad de bits disminuye también el rendimiento de la transmisión debido a que la tara permanece constante mientras que se descartan algunos bits de usuario.

I.7.2.1.1 Métodos de eliminación de bloques

La técnica de la eliminación de bloques es, idealmente, muy idónea para su empleo con codificadores de voz «jerarquizados» ya que asigna distintas prioridades a los diversos bits del tren de bits codificados. En el proceso de paquetización pueden, entonces, ensamblarse los bits de diversas prioridades en «bloques» separados de forma que los nodos de la red pueden aportar los paquetes suprimiendo ciertos bloques predefinidos para reducir la congestión (con una influencia mínima en la calidad de la voz).

En el algoritmo de codificación jerarquizada MICDA de las Recomendaciones G.727 y G.764 se especifica una ordenación de las muestras de forma que los bits más significativos de todas las muestras están situados en el mismo bloque, van seguidos de los siguientes bits más significativos y así sucesivamente. En consecuencia, cada bloque contiene bits de la misma importancia para las 128 muestras. El bloque 1 contiene la totalidad de los bits más significativos, el bloque 2 los siguientes bits significativos, etc. El bloque núcleo contiene información suficiente para la decodificación de la voz aunque la calidad completa únicamente se logra cuando se utilizan los bloques de núcleo y de ampliación.

Un indicador de eliminación de bloque situado en la cabecera del paquete señala el número máximo de bloques que pueden eliminarse y el número de bloques susceptible aún de eliminación. La eliminación de bloques puede tener lugar en cualquier nodo, lo que permite que los nodos intermedios puedan reducir la velocidad de bits y aliviar la congestión eliminando bits de los paquetes de voz sin tener que comunicarse con el nodo de origen.

Los resultados de simulación y de pruebas prácticas han demostrado que, en el caso de la codificación MICDA, el método de bloques ofrece una degradación más gradual en condiciones de sobrecarga que la eliminación de paquetes [23], [24], [25], [26], [27], [28], [7]. Cuando el número medio de bits por muestra se reduce a 3,97 bits/muestra, el valor de la MOS se reduce al 0,36 aproximadamente, con respecto al valor correspondiente a MIC. Para la misma relación de compresión el efecto de eliminación de paquetes es más drástico y el valor de la MOS cae en 1,36 aproximadamente, con un 0,54% de pérdida de paquetes. Para una tasa de pérdidas de paquetes de 2,2% o el 5,3% la diferencia entre los métodos de eliminación de bloques y de eliminación de paquetes es de una MOS de 1,4 y 2, respectivamente.

La ventaja del método de eliminación de bloques es que permite una decisión rápida y descentralizada. Asimismo la reducción de la calidad de la voz es gradual.

Los inconvenientes del método de eliminación de bloques son:

- 1) aumento de la complejidad de los nodos de la red;
- 2) únicamente hay un algoritmo jerarquizado normalizado;
- 3) el método podría no ser eficaz para velocidades de bits reducidas debido a que la tara del paquete puede ocupar una porción sustancial de la anchura de banda con respecto a la cabida útil (a menos que se multiplexen diversas llamadas o tipos de tráfico en la misma trama).

I.7.2.1.2 Método de eliminación selectiva de paquetes

Este método se basa en la posibilidad de eliminación de tramas (paquetes o células señaladas). El concepto es dividir las tramas (paquetes) en tramas (paquetes) de elevada prioridad y tramas (paquetes) de baja prioridad. Las tramas (paquetes) de alta prioridad contienen los bits más significativos de la palabra de código (o bits primarios) en tanto que las tramas (paquetes) de baja prioridad contienen los bits menos significativos (o bits secundarios) [13], [29], [30], [31]. Los bloques de ampliación se sitúan en las tramas (paquetes) de baja prioridad, mientras que los bloques núcleo se ubican en las tramas (paquetes) de alta prioridad, de forma que puedan eliminarse los bloques de ampliación cuando haya congestión.

Se han efectuado propuestas de voz codificada con MICDA sobre redes ATM o redes de retransmisión de tramas de datos únicamente. Por ejemplo, para la retransmisión de tramas de datos únicamente, se utiliza el bit «DE» para etiquetar las tramas eliminables [32].

Las ventajas principales del método de eliminación selectiva de paquetes son:

- 1) Simplicidad y eficacia del tratamiento de la congestión de la red.
- 2) Compatibilidad con los métodos utilizados en las redes de retransmisión de tramas de datos únicamente.
- 3) Facilidad relativa de realización.

Con el algoritmo existente este método adolece de algunos inconvenientes tales como:

- Se contemplan dos niveles de reducción de congestión, aunque en el caso de la Recomendación G.764 hay al menos tres niveles. En otras palabras, la tasa de codificación puede fluctuar entre dos extremos, por ejemplo entre 32 kbit/s y 16 kbit/s para la codificación MICDA jerarquizada (4,2) de la Recomendación G.727 es decir con 4 bits de código divididos en 2 bits primarios y 2 bits secundarios. En contraste, la codificación de la Recomendación G.764 permite una degradación más paulatina con una reducción de la tasa de codificación de 32 kbit/s a 24 kbit/s y a 16 kbit/s. Los resultados de pruebas prácticas han demostrado que el funcionamiento hasta una tasa de bits media de 3,0 bit/muestra es muy satisfactorio [7].

- Pueden introducirse retardos y procesamientos mayores debidos a que las muestras de la voz se ordenan en paquetes diferentes.
- Se utiliza más anchura de banda de transmisión para la tara de los paquetes.
- No existen garantías con este método de que los bloques de núcleo y los bloques de ampliación permanezcan sincronizados, lo cual es una desventaja importante. Por ejemplo, si se eliminan los bloques de núcleo debido a un fallo de la CRC pueden aún llegar al decodificador los bloques de ampliación. En el caso de la MICDA esto puede provocar el desajuste del decodificador respecto al codificador ya que la predicción de MICDA se basa en los bloques de núcleo. Ello provocará una degradación sustancial de la calidad del habla hasta que se produzca la resincronización. La codificación de la Recomendación G.764 asegura que no llegarán bloques de ampliación sin sus bloques de núcleo.
- Por último, en la codificación de la Recomendación G.764, se descartan los paquetes retardados mientras que en el método de Yin y otros [29] se transmiten los paquetes de alta prioridad con independencia de la condición de congestión, descartándose los paquetes de baja prioridad.

I.7.2.2 Métodos orientados a la clase

En estas propuestas, [33] y [34], el tráfico de voz se clasifica según los modelos de producción de la voz (sonidos explosivos, fricativos, vocales, etc.). Cada clase se codifica con algoritmo distinto, asignándosele una prioridad de entrega diferente. El objetivo es ajustar el tratamiento óptimo de la voz a las características de cada segmento. En contraposición, la codificación de la Recomendación G.764 codifica siempre la voz con la velocidad máxima especificada en la cabecera del paquete y deja la determinación de la prioridad del paquete a un mecanismo interno de la red.

I.8 Pérdida de paquetes

Como se ha explicado hasta ahora hay varias causas de pérdida de paquetes (borrado de tramas):

- 1) degradaciones de la transmisión;
- 2) retardo excesivo; y
- 3) congestión.

Los efectos de la pérdida de paquetes dependen del tipo de tráfico como se explica seguidamente.

I.8.1 Caso de señales de voz

La pérdida de paquetes se percibe como pausas que pueden producir una degradación importante de la calidad de funcionamiento. Hasta ahora, los algoritmos del CCITT no se habían diseñado explícitamente para que fueran resistentes a las pérdidas de paquetes. Por ejemplo, con la MIC y la MICDA si el porcentaje de tramas eliminadas rebasa el 1% puede degradarse sustancialmente la calidad de la voz [9], [10], [35]. En estos casos, la eliminación de paquetes no produce degradaciones sustanciales siempre que el porcentaje de paquetes perdidos sea inferior al 1%.

Se han diseñado algunos algoritmos MICDA no normalizados con el objetivo de mejorar su resistencia para transmisiones inalámbricas en interiores [36]. También uno de los requisitos del codificador UIT a 8 kbit/s es el de soportar aproximadamente un 3% de corrupciones de tramas (o pérdidas de paquetes) para el canal radioeléctrico.

Además, en marzo de 1994, la AT&T presentó modificaciones al decodificador de la Recomendación G.728 de forma que pudiera manejar mutilaciones de tramas en el tren de bits recibidos [37]. Ello se basaba en la extrapolación de algunos parámetros cuando «se mutila una trama» (es decir se pierde un paquete). Las reglas de decodificación especiales y simplificadas únicamente producen efectos cuando hay una pérdida de paquetes, lo cual asegura una compatibilidad completa con la Recomendación G.728 en condiciones normales de funcionamiento. Se mejora sustancialmente la resistencia a la pérdida de paquetes (mutilaciones de trama) sin que se incremente la complejidad del decodificador como se ha demostrado mediante resultados de simulaciones, utilizando errores en ráfagas obtenidos mediante las herramientas de la librería de instrumentos de soporte lógico CCITT/UIT. En el Cuadro I.1 que sigue se indican algunos resultados para paquetes correspondientes a 10 ms.

MOS para la LD-CELP robusta

Codificador	Tasa de pérdida de paquetes	MOS	
		Voz IRS	Voz No-IRS
MICDA a 32 kbit/s	0%	3,90	3,78
G.728	0%	3,90	3,79
G.728 ampliada	1%	3,95	3,82
	3%	3,77	3,59

Estos resultados ponen de manifiesto que hay modos de compensar con éxito los efectos de las pérdidas de paquetes los cuales no producirán una degradación de la MOS para una pérdida de paquetes del 1%, siendo la degradación del 0,2 aproximadamente en la escala de la MOS para un 3% de pérdidas de paquetes. Esto contrasta con las propiedades de la MICDA en la que debe evitarse a toda costa la pérdida de paquetes. Parece, por lo tanto, que puede ampliarse la Recomendación G.728 para tolerar más pérdidas de paquetes (aunque debe evitarse la eliminación incontrolada de paquetes). Queda en estudio la cuestión de como aprovechar estas características para simplificar el diseño de los protocolos de paquetización de la voz.

I.8.2 Caso de datos en banda vocal

Los datos en banda de frecuencias vocales son más susceptibles a los errores de bits aleatorios que la voz [7]. En estos casos cuando se pierde un paquete el módem receptor debe entrar en un procedimiento de recuperación si detecta una pérdida de portadora. El temporizador de desconexión de pérdida de portadora en un módem típico tiene valores comprendidos entre 100 ms y 20 s aproximadamente. Si expira este temporizador, el módem generalmente se desconectará y eliminará la llamada.

Si el módem está conectado a una red, el resultado dependerá también de los procedimientos internos que aplique la red para recuperarse de una pérdida de portadora. Por ejemplo, si la red reacciona prematuramente a una indicación de pérdida de portadora procedente del módem, puede desconectar el circuito virtual aun cuando el módem esté ejecutando el procedimiento de recuperación.

Dependiendo de las características del módem, incluso un único error de bit en la cabecera puede provocar esta reacción en cadena. Obviamente esto es una desventaja de la transmisión en modo paquetes comparada con la transmisión en modo circuito cuando los errores aleatorios son importantes. En contraste, para errores en ráfagas el modo paquete es superior al modo circuito.

I.8.3 Estrategias de relleno (voz)

Para compensar las pausas en la señal de voz recibida debidas a paquetes perdidos o eliminados se han propuestos diversas estrategias de relleno:

- 1) Sustitución del paquete eliminado con muestras de amplitud 0 [35]. Esto se denomina en las publicaciones técnicas *sustitución de los silencios*.
- 2) Relleno de las pausas con ruido inyectado denominado *ruido de confort*. Se denomina este método *interpolación de ruido k*.
- 3) Compensación de los paquetes eliminados por interpolación – Puede emplearse este método con la ampliación del algoritmo de la Recomendación G.728 propuesta por la AT&T [37].
- 4) Repetición del último paquete recibido – Si se clasifica la voz en diversas clases y se envía cada clase en un paquete diferente. El algoritmo de repetición puede variar con la clase del paquete anterior y con el número de secuencia [33], [34].
- 5) Extracción de segmentos de longitud igual a un paquete de los últimos paquetes recibidos (o de los que preceden y siguen al paquete perdido) que corresponden a los milisegundos anteriores a la detección de la pérdida del paquete – Se utilizan entonces estos segmentos para reemplazar el segmento ausente (*adaptación de estructura*) [38], [39].
- 6) Repetición de la última forma de onda de tono vocal para segmentos correspondientes a sonidos vocales y utilización del paquete anterior en los demás casos [38].

La sustitución de silencio o la interpolación de ruido son técnicas sencillas. La repetición de paquetes requiere menos memoria y procesamiento de la señal que la replicación de la forma de onda del tono vocal y la adaptación de estructura. Estos últimos métodos tienen la ventaja de que mantienen la continuidad de la fase en las fronteras entre paquetes.

I.9 Elección del tamaño de los paquetes

La elección del tamaño de los paquetes depende de varios factores, como son: retardo de paquetización, resistencia a los errores, retardo de extremo a extremo, fluctuación de fase, tara, etc. Cuando se contemplan diversos tipos de tráfico los requisitos para cada uno de esos tipos de tráfico pueden ser opuestos a los de otros.

I.9.1 Consideraciones sobre el habla

Los paquetes correspondientes a intervalos de voz de duración inferior a 50 ms proporcionan resistencia a la pérdida de paquetes [1], [5] (la longitud óptima se sitúa en la gama 16-32 ms) [8], [35], [40], siendo la tara más pequeña cuanto mayor es la longitud del paquete. En el algoritmo de la Recomendación G.764 se recopilan las muestras de voz en un periodo de 16 ms, lo que corresponde a 128 muestras para una velocidad de muestreo de 8 kHz.

Para un intervalo de paquetización determinado, el tamaño real del paquete, en octetos, dependerá de la velocidad de bits del algoritmo de codificación. En el caso de codificación jerarquizada, el tamaño se adapta a las condiciones de la red variables con el tiempo. En los sistemas ATM, se admite una tara mayor correspondiendo las muestras de la voz a 6 ms almacenados en 47 ó 48 octetos.

Para codificadores de voz de velocidad de bits reducida, por ejemplo 8 kbit/s, el tamaño no puede ser mayor que el correspondiente a los sistemas de la Recomendación G.764 o ATM para mantener el volumen de la tara dentro de límites razonables.

I.9.2 Consideraciones sobre los errores de bits

Según las normas de la Recomendación G.826 y ANSI/IEEE C37.1-1979 [41], las tasas de errores de bit pueden rebasar 10^{-4} durante intervalos breves. Teniendo en cuenta este factor en el cálculo de la probabilidad de error residual y considerando el mecanismo de transparencia de la HDLC con inserción de bits y sincronización de banderas, es posible demostrar que la longitud de la trama no deberá rebasar 500 octetos para una probabilidad de error residual inferior a 10^{-6} con tasas de errores de bits iguales a 10^{-4} [42].

I.9.3 Tráfico integrado

Las tramas largas requieren un mayor tiempo de transmisión que las tramas cortas. Por consiguiente, cuando se mezclan tramas largas y cortas estas últimas sufrirán un mayor retardo medio (para una utilización del enlace determinado) en comparación con el caso en que no hay tramas largas presentes [43], [44]. Esto se debe a que las tramas largas requieren un tiempo de transmisión mayor que las tramas cortas. Esto puede verse en la siguiente simulación que presenta dos escenarios distintos de mezclas de tramas de distintos tamaños en diferentes proporciones [43].

En el primer caso, la mitad de las tramas tiene una longitud igual a 74 octetos siendo la longitud de la otra mitad igual a 256 octetos. Para una utilización del enlace del 90% la longitud media de la cola de espera es 742,5 octetos y el límite de retardo para el percentil del 0,1% es del 7226,536 octetos.

Si la mezcla de tráfico varía según los siguientes perfiles:

Longitud de la trama (octetos)	Proporción
74	40%
256	20%
512	20%
1024	10%
1500	5%
2048	5%

para la misma utilización del enlace la longitud de la espera media resulta ser 2083 octetos (es decir unas tres veces mayor) y el límite del retardo del 0,1% es igual a 36171,875 (es decir unas cinco veces mayor). Obviamente, las tramas más cortas resultan sustancialmente retardadas en presencia de tramas muy largas. Cuando, para algunos tipos de tráfico, resultan admisibles los retardos grandes, la voz (y en menor grado el facsímil) plantea unos requisitos más estrictos sobre la variabilidad del retardo.

I.9.4 Algoritmo de la Recomendación G.764

La Recomendación G.764 estipula que la longitud máxima del campo de información debe ser 490 octetos. Por consiguiente, las tramas más largas como las admitidas en la Recomendación Q.922 deben segmentarse en el extremo de origen y recomponerse en el extremo de terminación.

I.10 Aspectos de compresión

Un sistema de transmisión digital típico codifica la voz dos veces. La primera codificación convierte la voz de origen del dominio analógico al dominio digital utilizando un método de codificación especial que implica la compresión para eliminar redundancia. Se emplea la segunda codificación para asegurar que la información binaria codificada cumple las limitaciones impuestas por el método de transmisión, por ejemplo para garantizar la densidad de unos. En consecuencia, existen dos métodos generales para el uso eficiente de la anchura de banda de transmisión disponible:

- 1) comprimir la señal de origen empleando un algoritmo de codificación de voz apropiado que elimine la redundancia en el material de origen; y
- 2) supresión de los periodos de inactividad, esto es eliminar los intervalos de silencio del habla o al menos codificarlos a una velocidad de bits inferior a la utilizada para la codificación de la voz.

I.10.1 Algoritmos de codificación del habla

Puede codificarse el tráfico de la voz empleando alguno de los diversos algoritmos de codificación normalizados. Existen varios métodos de codificación para señales telefónicas de banda estrecha recomendados por el CCITT/UIT que proporcionan calidad telefónica, a saber: MIC de la Recomendación G.711 (64 kbit/s), MICDA de la Recomendación G.726 (a 40 kbit/s y 32 kbit/s), algoritmos de codificación de 4 bits de la Recomendación G.727 (esto es a 32 kbit/s) y de la Recomendación G.728 (16 kbit/s). Cada uno de estos algoritmos introduce un retardo algorítmico de extremo a extremo distinto. Como se ha indicado anteriormente, el retardo algorítmico de extremo a extremo de los algoritmos G.726 y G.727 es 250 μ s y para el algoritmo G.728 es inferior a 2 ms.

El algoritmo de la Recomendación G.764 es compatible con los algoritmos MIC y MICDA, si bien la realización inicial de la ATM parece favorecer los sistemas MIC (sin interpolación digital del habla).

I.10.2 Interpolación digital del habla

Puede lograrse la reducción adicional de la velocidad de bits necesaria para la transmisión de la palabra eliminando los periodos de inactividad, es decir suprimiendo los intervalos de silencio del habla o codificando el silencio con una velocidad de bits inferior a la utilizada para el habla. Inicialmente se utilizó la interpolación del habla con asignación de tiempo (TASI, *time assignment speech interpolation*) para los sistemas analógicos por cable submarino, aplicándose subsiguientemente a los sistemas digitales de transmisión por satélite. Posteriormente, se combinó la interpolación digital de la palabra con la codificación MICDA de tasa variable [45] para aumentar el rendimiento de utilización del trayecto de transmisión.

A fin de reconocer que se está transmitiendo el habla se requiere un detector de la palabra muy sensible [2], [46]. La detección de la palabra se basa en diversas medidas tales como la energía a corto plazo, tasas de cruces por cero y secuencias de bits de signo. Dependiendo del tipo de diseño, se utiliza un tiempo de conmutación para extender el intervalo de la voz y evitar el recorte de la palabra. Esta extensión reduce obviamente la ganancia en anchura de banda por lo que generalmente se recomienda que el detector de voz funcione de tal forma que el periodo de actividad de voz que determine no rebase el periodo de actividad real en más del 5%. Por ejemplo, si el periodo de actividad de voz real es del 38% la actividad de voz medida en la interfaz de paquetes deberá ser inferior al 43%.

Debe resaltarse que la calidad del detector de voz en el extremo de origen es uno de los factores más importantes que determina la calidad global de la voz. Si el detector de voz no detecta correctamente la voz podría recortar el comienzo de las ráfagas de voz produciendo de este modo una importante degradación de la calidad del habla. Al contrario si el detector de voz es demasiado sensible dejará pasar más intervalos de silencio reduciéndose la ganancia de compresión.

En el extremo terminal puede insertarse ruido de confort o relleno de ruido en vez del silencio para minimizar las discontinuidades entre el ruido de fondo asociado a la voz y el silencio. El nivel de relleno de ruido puede especificarse en un campo de la cabecera del paquete. Es necesario realizar una selección cuidadosa de la potencia de ruido para evitar el problema del «bombeo de ruido» que consiste en un contraste desagradable entre el ruido de fondo durante el periodo de silencio y el ruido de fondo durante las ráfagas de voz [47].

Debe observarse que en la RTPC se requiere aplicar el control del eco siempre que se utilice la interpolación digital de la palabra debido a que el detector de voz puede considerar erróneamente las señales de eco como señales de voz y reducir la ganancia de compresión [48], [49]. Esto es independiente de la necesidad de cancelación del eco para compensar el retardo debido a la paquetización o para evitar el recorte de la voz debido al efecto de «congelación» [50], [51].

Todo protocolo debe incorporar algún método para distinguir entre las pausas debidas al silencio y las pausas debidas a paquetes extraviados/descartados. Es necesario efectuar esta distinción si se pretende realizar alguna acción de recuperación para sustituir los paquetes perdidos. En el algoritmo de la Recomendación G.764 se efectúa esta distinción mediante *bit-más* o bit-M lo que permite que el transmisor señalice si un paquete forma parte de una ráfaga o es el último paquete de una ráfaga. El bit-M se pone a 1 para todos los paquetes salvo el último de una ráfaga, en cuyo caso se ajusta a 0.

Utilizando el mecanismo del bit-M, el extremo terminal puede distinguir la ausencia de paquetes en el momento de la reproducción debidos al silencio en la fuente de la ausencia debida a la pérdida de paquetes en la transmisión. Cuando termina el tratamiento de un paquete y no dispone del siguiente, el extremo de terminación verifica el bit-M del último paquete. Si el valor de este bit es 0, la ausencia de un paquete indica que se trata de una pausa real en la transmisión en el extremo de origen del nodo de transmisión. Si el valor del bit-M del último paquete es 1, la ausencia indica un paquete tardío o perdido que no ha llegado. Este procedimiento permitirá al extremo terminal el empleo de procedimientos de relleno apropiados para los paquetes de voz.

En el caso de un locutor único, como sucede en telefonía celular/inalámbrica pueden emplearse técnicas de codificación adaptable para conseguir una reducción gradual de la velocidad de bits media asignando bits variables a las distintas partes del habla. De esta forma se asigna la velocidad de codificación más baja a las pausas de la conversación y la velocidad más elevada a la conversación activa.

Si se codifica el ruido de fondo mediante la velocidad más baja de un codificador de velocidad de bits variable, hay pocas probabilidades de que se recorten las partes iniciales de las palabras. En consecuencia, los requisitos de calidad de funcionamiento del detector de voz son menos estrictos en este caso. Esto podría ser ventajoso en aplicaciones de mayor dificultad cuando la interpolación de la palabra es menos efectiva como, por ejemplo, cuando el ruido ambiente es elevado.

I.11 Señalización orientada al canal

El protocolo de paquetización debe también retransmitir de forma transparente toda la señalización asociada al canal (señalización ABCD).

La información que debe transferirse es de dos tipos: «transición de señalización» y «refresco». Se produce una transición de señalización cuando cambia el estado de señalización; ello incluye la presencia de alarmas. La información de refresco se envía sobre una base periódica para indicar que el enlace permanece aún activo y para la recuperación de los efectos de pérdida de paquetes de transición de señalización.

Para cada conexión en banda de frecuencias vocales con señalización asociada al canal, la Recomendación G.764 especifica una dirección de capa 2 para el tráfico de banda de frecuencias vocales y otra para el tráfico de señalización. Este control fuera de banda es coherente con el método RDSI de separación de los planos de control y de usuario (definido en la Recomendación I.320). Actualmente es posible la integración de los procedimientos de control y de señalización para todas las fuentes, en tanto que la transferencia de las funciones de control y de señalización permanece diferenciada de la transferencia del habla del usuario o tramas de datos.

I.12 Ampliaciones

En esta cláusula se analizan las ampliaciones existentes y futuras del protocolo de paquetización de la voz básico.

I.12.1 Facsímil

Se ha ampliado el marco de la Recomendación G.764 al tráfico facsímil de la Recomendación G.765. Se han efectuado pruebas para evaluar la calidad de transmisión de los aparatos facsímil del Grupo III a través de enlaces por satélite con velocidad de datos intermedia (IDR, *intermediate data rate*)¹⁾ de Intelsat cuando se adapta la Recomendación G.764 para la transmisión facsímil de conformidad con la Recomendación G.765. En condiciones de prueba (es decir sin degradaciones de cola en las extensiones nacionales) el tráfico facsímil paquetizado conforme con el procedimiento de

¹⁾ IDR es el servicio digital de velocidad de datos intermedia de Intelsat que permite la transmisión de la voz, datos en banda vocal, datos digitales o vídeo con velocidades de información comprendidas entre 64 kbit/s y 45 Mbit/s. Los canales IDR utilizan aleatorizadores V.35, dispositivos de corrección de errores sin canal de retorno (FEC) de tasa 3/4 con decodificación con decisión ponderada y modulación QPSK [52].

modulación/remodulación de la Recomendación G.765 proporciona una mejora sustancial de la calidad de funcionamiento para la transmisión orientada a circuitos en enlaces IDR, en términos de páginas sin errores si no se utiliza la corrección de errores sin canal de retorno (FEC, *forward error correction*) propia del servicio [53] y [54]. Además, los resultados para las páginas con muchos errores indican que la remodulación paquetizada es mejor incluso para una BER de 10^{-5} . La concentración de los efectos de ráfagas de errores IDR en un único paquete junto con la estructura de control distribuida de los sistemas de paquetes reduce la frecuencia de la exposición a los errores y explica la gran ventaja asociada con la transmisión paquetizada.

I.12.2 Sincronización voz/vídeo

Cuando se transmite una señal vídeo junto con una señal de voz y tal vez de datos, a través de la RTPC, el protocolo de paquetización puede tener que contemplar la sincronización entre las informaciones de vídeo y de audio. Un resultado reciente indica que un retardo del canal de voz de ± 80 ms con respecto al canal vídeo puede tener un efecto mínimo o nulo sobre la inteligibilidad. La inteligibilidad parece reducirse con una desincronización superior a $+280$ ms o -160 ms con relación al canal vídeo [55].

I.12.3 Interfaz entre la RTPC y las LAN

El tráfico de una LAN puede integrarse con la RTPC a través de una o más interfaces en las nuevas redes que están apareciendo. Para mantener una calidad elevada es importante apreciar las diferencias entre las condiciones de funcionamiento de la RTPC y las LAN que pueden afectar a la paquetización de la voz, por ejemplo:

- 1) Las redes de telecomunicaciones públicas tienen planes de sincronización para evitar el desbordamiento o relleno insuficiente de las memorias intermedias [56]. En una LAN no se necesita una sincronización tan estricta, disponiéndose de otros mecanismos para controlar las memorias intermedias en diversas estaciones. Todo diseño de protocolos multimedia (voz/datos digitales/vídeo) deberá especificar como puede sincronizarse el receptor con los transmisores de distintos tipos de medios.
- 2) En una LAN son típicas tasas de errores de bits de 10^{-9} o mejores. En una conexión por la RTPC general que abarque varios países no siempre puede lograrse esa tasa de errores de bits. En consecuencia, cuando se transmita un servicio de voz de banda ancha de una LAN a una red de área amplia, el protocolo LAN deberá tener en cuenta esta diferencia entre las tasas de errores de bits.
- 3) La transmisión del tráfico de LAN a través de la RTPC puede, potencialmente, provocar grandes retardos y fluctuaciones de fase en el receptor (o pérdida de paquetes) durante los periodos de congestión y de tráfico elevado. Los efectos de la congestión en una LAN, por ejemplo pérdida o demora de los paquetes, se compensan a menudo mediante protocolos de capa superior tales como el TCP/IP para tráfico de datos digitales. Por ejemplo el TCP/IP puede invocar estrategias de retransmisión que son válidas para los datos pero no resultan adecuadas para la voz interactiva (desde un punto de vista de calidad subjetiva).
- 4) Por último si el tráfico LAN es ATM (esto es MIC de 64 kbit/s sin interpolación digital de la palabra) la anchura de banda requerida en la RTPC puede ser excesiva.

En resumen, las diferencias entre el tráfico por la RTPC y el tráfico por las LAN pueden deberse a:

- 1) las condiciones en la red de área extendida que son más diversificadas y menos predecibles que las de una LAN; y
- 2) que es relativamente más sencillo añadir capacidad a una LAN si la congestión es un problema de esa LAN.

En consecuencia, las limitaciones en la paquetización de la voz dentro de una LAN pueden ser más estrictas que las correspondientes a la red de área extensa. Si bien el empaquetamiento directo del tráfico de la LAN es la propuesta más simple, este método no tiene en cuenta las diferencias anteriores entre la naturaleza de las redes de área extendida y las LAN. Es muy probable que se requiera una interfaz para la conversión de protocolo entre las diversas LAN y la RTPC.

I.12.4 Ampliación a nuevos algoritmos

Sobre la base de la información disponible puede establecerse que puede diseñarse un algoritmo LD-CELP jerarquizado para su utilización en equipos de compresión para redes fijas y móviles (inalámbricas/celulares). Empleando un único algoritmo jerarquizado «universal» puede lograrse una integración «concordante» de numerosos servicios y evitarse diversos problemas ocasionados por la utilización de codificaciones no jerarquizadas y/o incompatibles.

Esto supondría, para el usuario final, la optimización del empleo de la anchura de banda manteniendo la calidad del servicio y posibilitando el despliegue rápido de líneas de acceso, así como un número mínimo de codificaciones/decodificaciones sobre una base de extremo a extremo. Otras ventajas serían la administración centralizada y armonizada de la totalidad de la red.

Como la red del futuro será más dinámica y sus fronteras aparecerán menos definidas (desde un punto de vista de ingeniería) deberán modificarse las normas para el control del acceso sin que resulte afectada la calidad del servicio. Por ejemplo un usuario móvil puede desear establecer comunicación, en un momento determinado, con un usuario fijo a través de la RTPC. Unos momentos después el mismo usuario puede desear efectuar una llamada a otro usuario inalámbrico/celular en una ubicación diferente o incluso en un continente distinto.

Obviamente, debido a la naturaleza dinámica de las complejidades lo más sencillo sería minimizar el número de codificaciones/decodificaciones en un trayecto de llamada. Además, puede ser necesario efectuar la compresión para aumentar el número de llamadas que pueden transmitirse a través de medios de transmisión escasos. Si el usuario permanece en el dominio de paquetes puede evitarse el problema de la codificación en cascada y lograrse una degradación paulatina en condiciones de sobrecarga utilizando, en cada enlace, el mismo algoritmo jerarquizado [57].

Durante el periodo de estudios final del CCITT (1988-1992), se realizó un gran esfuerzo para investigar como podrían evitarse las conexiones en cascada. Se intercambiaron coordinaciones entre las CE XV y II sobre las implicaciones de la utilización de un circuito equipado con DCME en una conexión. Se llegó a la conclusión de que «esta situación origina problemas importantes no solamente en lo relativo al encaminamiento, sino en lo que respecta a la calidad de servicio, gestión de la red, diseño de la red, planificación y operaciones» [58]. Tales problemas podrían reducirse, si no suprimirse totalmente, en el dominio de los paquetes. En consecuencia, no pueden pasarse por alto las implicaciones de las posibles conexiones en cascada sobre las operaciones y reglas de funcionamiento de la red cuando se inicien actividades de desarrollo de nuevos codificadores.

De alcanzarse un acuerdo universal sobre un codificador jerarquizado de velocidad reducida, ello mejoraría enormemente la calidad del servicio en banda de frecuencias vocales en una red de telecomunicaciones descentralizada.

I.13 Resumen

En este apéndice se han resumido y revisado las actividades de normalización de la paquetización de la voz en el CCITT (actualmente UIT-T). Los temas que quedan en estudio son:

- 1) paquetización de la voz en sistemas ATM; y
- 2) ampliaciones de la Recomendación G.764 para nuevos algoritmos normalizados de calidad telefónica.

Referencias

- [1] MONTGOMERY (W.A.): Techniques for packet voice synchronization, *IEEE J. Select. Areas Commun.*, Vol. SAC-1, N.º 6: 1022-1028, diciembre 1983.
- [2] MIEDEMA (H.), SCHACHTMAN (M.G.): TASI quality – Effect of speech detectors and interpolation. *The Bell System Technical Journal*, Vol. XLI, N.º 4, páginas 1455-1473, julio 1962.
- [3] SRIRAM (K.), MCKINNEY (R.S.), SHERIF (M.H.): Voice packetization and compression in broadband ATM networks, *IEEE Journal Selected Areas in Communication*, special issue on Teletraffic Analysis of ATM Systems, Vol. 9 (3): 294-305, abril 1991. Also in part in the *Proc. of the ITC Seminar*, Morristown, NJ, 9-11 de octubre, 1990.
- [4] Sector de Normalización de las Telecomunicaciones de la UIT: Impact of transmission delay on Group 3 facsimile, *USA Contribution*, Questions: 4, 8, 9, 13/15, COM 15/D.26, septiembre 1993.
- [5] WEINSTEIN (C.J.), FORGIE (J.W.): Experience with speech communication in packet networks, *IEEE J. Select. Areas Commun.*, Vol. SAC-1, N.º 6: 963-980, diciembre 1983.

- [6] SRIRAM (K.): Dynamic bandwidth allocation and congestion control schemes for voice and data multiplexing in wideband packet technology, *ICC'90*, Vol. 3, páginas 1003-1009, Atlanta, GA, 15-19 de abril 1990.
- [7] MONTEILLET (A.), SHERIF (M.H.), RUBIN-JAFFE (K.), BAILLY (R.), ERDREICH (M.): Expérimentation sur site d'un système de multiplication de circuits par mise en paquets, *Annales des Télécommunications*, Vol. 47: (5-6), 175-179, 1992.
- [8] GRUBER (J.), STRAWCZYNSKI (L.): Subjective effects of variable delay and speech loss in dynamically managed voice systems, *GLOBECOM'82*, Vol. 2: F.7.3.1-F.7.3.5, Miami, FL, noviembre-diciembre de 1982.
- [9] GRUBER (J.G.), LE (N.H.): Performance requirements for integrated voice/data networks, *IEEE J. Select. Areas Commun.*, Vol. SAC-1, N.º 6: 981-1005, diciembre 1983.
- [10] SHERIF (M.H.): Overview of wideband packet protocols. *Proceedings of the 9th International Conference on Digital Satellite Communications ICDSC 9*, Session A4, páginas 135-143, Copenhagen, 18-22 de mayo, 1992.
- [11] CASNER (S.), DEERING: First IETF Internet Audiocast, *Connexions Newsletter*, páginas 10-17, 1992.
- [12] ISO/CEI JTC 1/SC6/WG1: The Evolution of HDLC: Past! Present! Future!, *WGI Convener (R.I.G. Prince), Project JTC 1.06.16, Doc. No. 337r*, US Contact: American National Standards Institute, 11 West 42nd Street, NY, NY 10036.
- [13] BIALLY (T.), GOLD (B.), SENEFF (S.): A technique for adaptive voice flow control in integrated packet networks, *IEEE Trans. Commun.*, Vol. COM-28: 325-333, marzo 1980.
- [14] BIALLY (T.), MCLAUGHLIN (A.J.), WEINSTEIN (C.J.): Voice communications in integrated digital voice and data networks, *IEEE Trans. Commun.*, Vol. COM-28: 1478-1490, septiembre 1980.
- [15] SHERIF (M.H.), BOWKER (D.O.), BERTOCCI (G.), ORFORD (B.A.), MARIANO (G.A.): Overview and performance of CCITT/ANSI embedded ADPCM algorithms, *IEEE Trans. Comm.*, Vol. 41, N.º 2, 391-399, febrero 1993.
- [16] De IACOVO (R.D.), SERENO (D.): Embedded CELP coding for variable bit-rate between 6.4 kbit/s and 9.6 kbit/s, *ICASSP 91 International Conference on Acoustics, Speech and Signal Processing*, Vol. 1, páginas 681-684, Toronto, Ontario, Canadá, 14-17 de mayo, 1991.
- [17] DAVIDSON (G.), GERSHO (A.): Multiple stage vector excitation coding of speech waveforms, *ICASSP 88 International Conference on Acoustics, Speech and Signal Processing*, Vol. 1, páginas 163-166, NY, 11-14 de abril, 1988.
- [18] LE GUYADER (A.), LOZACH (B.), MOREAU (N.): Embedded algebraic CELP coders for wideband speech coding, *Signal Processing VI Theories and Applications, Proceedings of EUSIPCO-92 Sixth European Signal Processing Conference*, Vol. 1, páginas 527-530, Brussels, Belgium, 24-27 de agosto, 1992, J. Vandewable, R. Boch, M. Moonen et A. Osterlinch (eds.), Elsevier, Amsterdam, 1992.
- [19] LE GUYADER (A.), BOURSICAUT (E.): Embedded wideband VSELP speech coding with optimized codebooks, *Proceedings of IEEE workshop on speech coding for Telecommunications: Speech coding for the network of the future*, páginas 15-16, Sainte-Adèle, Quebec, Canadá, 13-15 de octubre, 1993.
- [20] TANIGUCHI (T.), OTHA (Y.), YAO (J.): Structure codebook for variable-rate CELP coding, *Proceedings of IEEE workshop on speech coding for Telecommunications: Speech coding for the network of the future*, páginas 87-88, Sainte-Adèle, Quebec, Canadá, 13-15 de octubre, 1993.
- [21] LEUNG (T.W.), LE BLANC (W.P.), MAHMOUD (S.A.): Speech coding over frame relay networks, *Proceedings of IEEE workshop on speech coding for Telecommunications: Speech coding for the network of the future*, páginas 75-76, Sainte-Adèle, Quebec, Canadá, 13-15 de octubre, 1993.
- [22] USA: Embedded LD-CELP, UIT-T SG 15 Contribution D.225, Geneva, Switzerland, 16-27 de mayo, 1994.
- [23] BOWKER (D.O.), ARMITAGE (C.B.): Performance issues for packetized voice communication, *Proc. Nat. Comm. Forum*, 41 (3): 1087-1092, 1987.
- [24] BOWKER (D.O.), DVORAK (C.A.): Speech transmission quality of Wideband Packet Technology, *GLOBECOM'87*, Vol. 3: 1887-1889, Tokyo, Japan, 15-18 de noviembre, 1987.
- [25] DRAVIDA (S.), SRIRAM (K.): End-to-end performance models for variable bit rate voice over tandem links in packet networks, *IEEE INFOCOM'89*, Vol. 3: 1617-1622, Ottawa, Canadá, 23-24 de abril, 1989.
- [26] KARANAM (V.R.), SRIRAM (K.), BOWKER (D.O.): Performance modeling of variable bit rate voice in packet switched networks, *GLOBECOM'88*, Vol. 3: 1617-1622, Hollywood, FL, 28 de noviembre-1 de diciembre, 1988.

- [27] SRIRAM (K.), LUCANTONI (D.M.): Traffic smoothing effects of bit dropping in a packet voice multiplexer, *IEEE INFOCOM'88*, páginas 759-770, New Orleans, LA, marzo 1988.
- [28] SHERIF (M.H.): Overview of wideband packet protocols, *Int. J. Sat. Comm.*, Vol. 10, 261-268, 1992.
- [29] YIN (N.), LI (S.Q.), STERN (T.E.): Congestion control for packet voice by selective packet discards, *GLOBECOM'87*, Vol. 3: 1782-1786, Tokyo, Japón, 15-18 de noviembre, 1987.
- [30] YIN (N.), STERN (T.), LI (S.Q.): Performance analysis of a priority-oriented system, *IEEE INFOCOM'87*, páginas 856-862, San Francisco, marzo 1987.
- [31] KASHORDA (M.), JONES (E.V.): Simultaneous voice/data transmission over a single 64 kbit/s ISDN bearer channel, *Proceedings of the Third IEE Conference on Telecommunications*, páginas 281-285, Heriot-Watt University, Edinburgh, UK, 1991.
- [32] Dowty Communications Ltd.: Packetized voice over frame relay: Issues and considerations, *Contribution to the frame relay forum*, mayo 1992. Contact: Ted Hatala, Cray Communications, Mayze House, Westmead, Swindon, Wiltshire SN5 7UQ, England.
- [33] Da SILVA (L.A.), PETR (D.W.), FROST (V.S.): A class-oriented replacement technique for lost speech packets, *IEEE INFOCOM'89*, Vol. 3: 1098-1105, Ottawa, Canadá, 23-24 de abril, 1989.
- [34] PETR (D.W.), Da SILVA Jr. (L.A.), FROST (V.S.): Priority discarding of speech in integrated packet networks, *IEEE J. Select. Areas Commun.*, Vol. 7, N.º 5: 644-656, junio 1989.
- [35] JAYANT (N.S.), CHRISTENSEN (S.W.): Effects of packet losses in waveform coded speech and improvements due to an odd-even sample-interpolation procedure, *IEEE Trans. Commun.*, Vol. COM-27, N.º 2, 101-109, febrero 1981.
- [36] GOULD (K.W.), COX (R.V.), JAYANT (N.S.), MELCHNER (M.J.): Robust speech coding for the indoor wireless channel, *AT&T Technical Journal*, Vol. 72, N.º 4, páginas 64-73, julio-agosto 1993.
- [37] AT&T: G.728 decoder modifications for frame erasure concealment, document N.º AH-5-10, ITU-T Rapporteur's meeting, Tokyo, Japon, marzo 1994. Contact: Mr. S. Hayashi, NTT, Human Interface Laboratories, Speech Laboratory, 9-11 Midori-cho 3 Chome, Musashino-shi, Tokyo 180, Japón.
- [38] GOODMAN (D.J.), LOCKHART (G.B.), WASEM (O.J.), WONG (W.C.): Waveform substitution techniques for recovering missing speech segments in packet voice communications, *IEEE Trans. Acoust. Speech, Signal Processing*, Vol. ASSP-34, N.º 6: 1440-1448, diciembre 1986.
- [39] WASEM (O.J.), GOODMAN (D.J.), DVORAK (C.A.), PAGE (H.G.): The effect of waveform substitution on the quality of PCM packet communications, *IEEE Trans. Acoust. Speech, Signal Processing*, Vol. ASSP-36, N.º 3: 342-348, marzo 1988.
- [40] MINOLI (D.): Optimal packet length for packet voice communication, *IEEE Trans. Commun.*, Vol. COM-27, N.º 6: 607-611, marzo 1979.
- [41] American National Standard ANSI/IEEE C37.1-1979: Definition, specification, and analysis of manual automatic, and supervisory station control and data acquisition, Sponsored by the *Substations Committee of the IEEE Power Engineering Society*, 15 de diciembre, 1977.
- [42] HOEFFELMAN (J.): Performance de la procédure de télécommunication HDLC et adaptation aux exigences d'une haute sécurité de transmission, *Onde Electrique*, Vol. 60, N.º 11, 52-60, 1980.
- [43] SRIRAM (K.), SHERIF (M.H.): Frame size requirements in frame relay networks, contribution N.º T1S1.1/90-185 (T1S1.2/90-128), 17 de abril, 1990.
- [44] SHERIF (M.H.), BOSSE (M.P.): Les paquets de bande élargie: une nouvelle technique de transmission, *Annales des Télécommunications*, 46, N.º 7-8, páginas 392-407, 1991.
- [45] YATSUZUKA (Y.): High-gain digital speech interpolation with adaptive differential PCM encoding, *IEEE Trans. on Commun.*, Vol. COM-30, N.º 4, 750-761, abril 1982.
- [46] YATSUZUKA (Y.): Highly sensitive speech detector and high-speed voiceband data discriminator in DSI-ADPCM systems, *IEEE Trans. on Commun.*, Vol. COM-30, N.º 4, 739-750, abril 1982.
- [47] DVORAK (C.A.), ROSENBERGER (J.R.): Deriving a subjective testing methodology for digital circuit multiplication and packetized voice systems, *IEEE J. Select. Areas Commun.*, Vol. 6, N.º 2: 235-241, febrero 1988.
- [48] ANSI T1.508-1992: Loss Plan for Evolving Digital Networks.

- [49] Committee T1 Report No. 27: A Technical report on echo cancelling, Prepared by *TIA1.6 Working Group on Specialized Signal Processing*, Washington, D.C., noviembre 1993.
- [50] KIHLSSTRÖM (L.O.), VIKLUND (N.), BRIDGEWATER (D.): The development of an open-network digital circuit multiplication equipment system, *International Journal of Satellite Communications*, Vol. 8, 461-490, 1990.
- [51] SHERIF (M.H.), CLARK (R.J.), FORCINA (G.P.): CCITT/ANSI Voice Packetization Protocol, *International Journal of Satellite Communications*, Vol. 8, 429-436, 1990.
- [52] Performance Characteristics for Intermediate Data Rate (IDR) Digital Carriers, *Intelsat Earth Station Standards, IESS-308 (Revision 6B)*, diciembre 1992.
- [53] SHERIF (M.H.), CUEVAS (E.), LIEBERT (T.), DIMOLITSAS (S.): Improvement of the Transmission Quality of Group III Facsimile on IDR Satellite Links, Proceedings of the *First IEEE Symposium on Global Data Networking*, páginas 101-111, Le Caire, Egipte, 12-15 de diciembre, 1993.
- [54] SHERIF (M.H.), CUEVAS (E.), LIEBERT (T.), DIMOLITSAS (S.): Transmission Quality of Group III Facsimile on IDR Satellite Links with Circuit Multiplication Equipment, *International Journal of Satellite Communications*, Vol. 12, 167-180, 1994.
- [55] Royal PTT, The Netherlands: Effects of desynchronization of vision and speech on the perception of speech: Preliminary results, *Delayed Contribution D.81 to CCITT SG XII. Q.27/XII and 31/XII*, Brasilia, 6-13 de septiembre, 1991.
- [56] REY (R.F.), Technical Editor: Engineering and operation in the Bell System, *AT&T Bell Laboratories*, 1983.
- [57] AT&T: Cellular applications of PCME, *ITU-T White Contribution COM 15-36-E*, julio 1993.
- [58] Speech processing techniques in the international digital network, *CCITT Liaison statement from WP II/1, Q.6/II (Routing), COM II-R8*, 24.10-3.11.89, Ginebra, Suiza.