VNIVERSIDAD NACIONAL AVFNOMA DE MEXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA

FACULTAD DE INGENIERÍA

"CIFRADO DE VOZ MEDIANTE UN SISTEMA INCRUSTADO EN TERMINALES ANALÓGICAS DE LA RTPC."

TESIS

QUE PARA OPTAR POR EL GRADO DE :

MAESTRO EN INGENIERÍA.

INGENIERÍA ELÉCTRICA-TELECOMUNICACIONES

P R E S E N T A :

RAFAEL MORALES SEVILLA

TUTOR:

M. EN I. FERNANDO LEPE CASILLAS

2008

Jurado Asignado:

Presidente	Dr. Federico José Kulhmann Rodríguez
Secretario	Dr. Francisco García Ugalde
Vocal	M. en I. Fernando Lepe Casillas
1er. Suplente	Dr. Carlos Rivera Rivera
2do. Suplente	Dr. Miguel Moctezuma Flores

Ciudad Universitaria, Distrito Federal, Méx.

TUTOR DE TESIS:

M. en I. Fernando Lepe Casillas.

_____

DEDICATORIAS.

AL GRAN INGENIERO DE LA

VIDA, POR PERMITIRME

TERMINAR ESTA META.

A MI ESPOSA SANDY Y A MI

HIJO EMILIO POR SER EL

MOTIVO DE MI VIDA. LOS

AMO.

A MI MADRE, LA GRAN

MUJER QUE GRACIAS A SU

CARIÑO Y EJEMPLO, HE

LOGRADO MIS OBJETIVOS.

AGRADECIMIENTOS.

QUIERO AGRADECER DE UNA MANERA ESPECIAL AL PROFESOR FERNANDO LEPE CASILLAS, QUIEN ADEMÁS DE SER MI ASESOR DE TESIS, FUE UNA PERSONA HONESTA Y SINCERA QUE SIEMPRE ME DIO ACERTADOS CONSEJOS.

ÍNDICE.

PAGINA

Jurado…… *ii*

Dedicatorias…….... *iii*

Agradecimientos…….. *iv*

Índice…… *v*

Introducción…… 1

CAPÍTULO 1

MARCO DE REFERENCIA.

1.1. Planteamiento del problema..... 3

1.2. Hipótesis. 4

1.3. Objetivos. 5

1.3.1. Objetivo general. 5

1.3.2. Objetivos particulares..... 5

CAPÍTULO 2

MARCO TEÓRICO CONCEPTUAL.

2.1. Comunicaciones de voz y datos en la red telefónica bàsica (**RTB**)…….. 7

2.1.1. El teléfono y el lazo local. …… 10

2.1.2. Establecimiento de enlace entre dos estaciones conectadas a la misma central…….…….... 14

2.1.3. Transmisión de datos en la red telefónica básica…….. 15

2.2. Seguridad de la información a través del cifrado..... 17

2.2.1. Sistemas de cifrado. 17

2.2.2. Encubridores de bloque y de flujo…… 22

2.2.3. Algoritmo Estándar de Cifrado 2001, **AEC-2001** (**AES**)..... 25

2.3. Compresión de la voz…… 37

2.3.1. Modelo básico de codificadores de voz mediante filtros y señales de excitación……..... 41

2.3.2. Codificador de predicción lineal con señal de excitación del filtro seleccionada según un código (***CELP***: *Code Excited Linear Prediction* 47

2.3.3. El codificador **CPLE *FS-1016* 4.8 kbps**…….. 51

CAPÍTULO 3

DESARROLLO.

3.1. Desarrollo del sistema de cifrado de voz…….. 53
3.1.1. Consideraciones generales……... 53
3.1.2. Propuesta de desarrollo del sistema de cifrado…… 54
3.2. Hardware del sistema de cifrado…… 56
3.3. Descripción de la tarjeta de desarrollo *DSP56858EVM*…….. 59
3.3.1. *DSP56858* de 16 bits…….. 59
3.3.2. Interfaces paralela y de comunicación serial ***SCI*** del *DSP56858*…….. 62
3.3.3. Puerto RS-232 para comunicación serial de la tarjeta *DSP56858EVM* 64
3.3.4. El codificador-decodifcador *CS4218*…… 65
3.3.5. Puertos de Propósito General de Entrada y/o Salida (***GPIO:*** *General Purpose Input Output*) …….. 65
3.4. El módem *Multitech MT5600*…… 66
3.4.1. Conexión del módem a la tarjeta *DSP56858EVM*…… 67
3.5. El codificador-decodificador *CS4218* y su conexión al *DSP56858*……. 70
3.5.1. Características…… 70
3.5.2. La interfaz Serial Síncrona Mejorada *(**ESSI**: Enhanced Synchronous Serial Interface)* del DSP56858…… 71
3.5.3. Conexión del codificador-decodificador al puerto ***ESSI*** del *DSP56858*. 73
3.6. Programación para la comunicación del *DSP56858* con el codificador-decodificador de voz y el módem…… 77
3.6.1. Comunicación de datos entre el codificador-decodificador y el **PSD**…. 77
3.6.1.1. Transmisión de datos del codificador-decodificador de voz al **PSD** (modo ***DMA***)…… 80
3.6.1.2. Transmisión de datos del **PSD** al codificador-decodificador de voz (modo ***DMA***)…… 82
3.6.2. Comunicación de datos entre el **PSD** y el módem…… 83
3.6.2.1. Transmisión…….. 84
3.6.2.2. Recepción…….. 89

3.7. Algoritmos de compresión y cifrado…… 92
3.7.1. Compresión-cifrado…… 92
3.7.1.1. Compresión **CPLE**…… 93
3.7.1.2. Cifrado…… 94
3.7.2. Descifrado-descompresión…… 97
3.7.2.1. Descifrado **AEC-2001**…… 98
3.7.2.2. Decompresión…… 100

CAPÍTULO 4

Resultados..... 101

CAPÍTULO 5

Conclusiones y recomendaciones……..... 111

Glosario…… 116
Referencias…….... 124

INTRODUCCIÓN.

La información es un recurso valioso que las personas, empresas y gobiernos requieren para tomar decisiones adecuadas. El avance de la tecnología ha permitido que todo tipo de información esté disponible en cualquier momento y de una manera fácil, incrementándose la demanda de transferencia de la información a través de las diferentes redes de telecomunicaciones. Sin embargo, con estos avances también han surgido personas que utilizan o manipulan la información para fines delictivos, lo que obliga a mantener su privacidad e integridad en todo el trayecto sobre el cual se transmite.

La información que se busca proteger es diversa tal como la incluida en las transacciones bancarias, datos confidenciales de personas, empresas o gobiernos, bases de datos, conversaciones telefónicas, entre otras. La voz que se transmite a través de la **Red Telefónica Básica RTB** (***PSTN:*** *Public Switched Telephone Network*)[1] es un tipo de información que requiere protección en su privacidad con el fin de evitar que personas no autorizadas escuchen las conversaciones telefónicas.

La importancia de la seguridad telefónica es motivo suficiente para proponer un sistema que cifre la voz transmitida a través de la **RTB** y que pueda ser incrustado en los teléfonos receptor y transmisor conectados al lazo local de esta red.

Al inicio del presente trabajo se analiza el problema de la seguridad en las conversaciones telefónicas, los componentes principales de la **RTB***,* la transmisión de la voz y se propone en forma general, una solución para proporcionar seguridad durante una conversación entre usuarios.

Posteriormente, se exponen los distintos conceptos relacionados con la infraestructura de la red telefónica básica, el funcionamiento del teléfono, el cifrado de la información, el **Algoritmo Estándar de Cifrado AEC-2001** (***AES:*** *Advanced Encryption Standard*) que en la actualidad es uno de los más seguros, y la compresión de la voz con el **Codificador con Predicción Lineal con Señal de Excitación del Filtro Seleccionada Según un Código** (***CELP:*** *Code Excited Linear Prediction*).

Una vez que se cuenta con las bases teóricas, se describe la propuesta de seguridad de la voz, la programación de los algoritmos de cifrado y codificación de la voz, la tarjeta de desarrollo para implantar el sistema, y la programación del módem para la transmisión de la información. También se describe el hardware adicional necesario para la transmisión de la voz cifrada y codificada a través de la red telefónica básica.

Por último, se muestran los resultados obtenidos en las distintas pruebas que se realizaron durante el desarrollo del trabajo, además de las conclusiones y recomendaciones para mejorar la presente investigación.

En esta tesis, los acrónimos en inglés se escriben con mayúsculas, itálicas y en negritas, los acrónimos en español en mayúsculas y negritas. Asimismo, las palabras en idioma diferente al español con letras itálicas.

Capítulo

1

MARCO DE REFERENCIA.

1.1. PLANTEAMIENTO DEL PROBLEMA.

Recientemente en México se han conocido situaciones de espionaje telefónico, como el caso de la conversación entre el gobernador del Estado de Puebla Mario Marín y el “rey de la mezclilla” Kamel Nacif, en donde quedó al descubierto la vulnerabilidad de las comunicaciones telefónicas (si bien este caso realmente sirvió para que la sociedad se percatara de la corrupción que impera en México). Y si este suceso no puso en riesgo la estabilidad del país, no se puede descartar que el espionaje telefónico pueda utilizarse para dañar al Estado.

Por la red telefónica básica, transita información de particulares, empresas y gobiernos, susceptible de intercepción para su empleo en actividades ilícitas. Un ejemplo muy importante de comunicación telefónica sensible al espionaje, es la que se da al acceder a servicios bancarios vía telefónica.

La **Red Telefónica Básica** (**RTB)** carece de esquemas de seguridad que protejan las conversaciones telefónicas entre usuarios, por lo que es relativamente fácil intervenir teléfonos en cualquier punto de la **RTB** (aparatos telefónicos, líneas de abonado, conmutadores) sin requerir equipos sofisticados y que pueden interceptar y obtener la voz de una manera clara y entendible. Incluso existen en el mercado grabadoras diseñadas para conectarse a la línea telefónica y obtener la voz en tiempo real; usualmente, estos equipos son colocados en los registros telefónicos sin que lo noten los usuarios.

1.2. HIPÓTESIS.

Para proteger una conversación entre dos personas, la seguridad proporcionada a la voz debe estar presente en los teléfonos, línea telefónica, circuitos y conmutadores de la red telefónica básica, de tal forma que una conversación entre dos usuarios no pueda ser decodificada o entendida por un tercero no autorizado.

Si se construyen dos sistemas idénticos que puedan ser incrustados en los aparatos telefónicos, uno en el receptor y otro en el transmisor, que se adapten fácilmente al lazo local, que permitan ambos realizar la conversión analógica-digital y digital-analógica, compresión y descompresión, cifrado y descifrado de las señales de audio, con los siguientes requerimientos:

1) **Proceso de establecimiento automático de la comunicación entre ambos teléfonos.**

2) **El sistema permitirá establecer una conversación segura entre dos personas sin que pueda ser descifrada por otras.**

3) **Los procesamientos que deban sufrir las señales de voz en los sistemas incrustados deberán ejecutarse lo suficientemente rápido como para permitir que la conversación se lleve a cabo en tiempo real, con una calidad de voz aceptable, es decir, inteligible y con retardo imperceptible.**

Se proporcionará confidencialidad de la información que se transmite sobre la RTB cuando se establece una conversación telefónica entre dos usuarios.

1.3. OBJETIVOS.

1.3.1. OBJETIVO GENERAL.

Construir un sistema de cifrado de voz que pueda ser incrustado en los aparatos telefónicos, que además de proporcionar seguridad, permita una transmisión de voz en tiempo real a través de la **RTB**, con una calidad aceptable e inteligible.

1.3.2. OBJETIVOS PARTICULARES.

- Determinar una solución general y aproximada al problema de seguridad de voz en conversaciones telefónicas factible de ser construida, partiendo del funcionamiento de la **RTB y** la forma de transmisión de la voz.

- Conocer el funcionamiento de cada uno de los componentes que integran la red telefónica, los distintos métodos de cifrado, además de otras formas de protección de la voz, los procesos que se aplican a la voz para su transmisión en forma digital, y la transmisión de datos sobre la **RTB**.
- Determinar el procesador micro y periféricos más adecuados para la construcción del sistema incrustado (*"embedded"*) de cifrado de voz con el menor costo.
- Programar los algoritmos de cifrado y codificación de voz, además de transmisión de datos sobre la línea telefónica, en un computador personal utilizando el lenguaje C para realizar simulaciones.
- Construir el sistema de cifrado con el procesador micro seleccionado y demás componentes necesarios, implantando los distintos algoritmos sobre dicho procesador.

Capítulo

2

MARCO TEÓRICO CONCEPTUAL.

2.1. COMUNICACIONES DE VOZ Y DATOS EN LA RED TELEFÓNICA BÁSICA (RTB).

El sistema telefónico ha sido, y es en la actualidad, el medio de comunicación más importante y extendido a nivel mundial. Permite la comunicación de millones de personas mediante conversaciones en tiempo real y con una aceptable calidad de voz. Originalmente, la redes telefónicas fueron diseñadas para transmitir voz analógica, pero con el surgimiento de Internet y el avance de las tecnologías de las telecomunicaciones y de la información, estas redes evolucionan y se convierten en redes digitales con capacidad para transmitir voz digital, imágenes, datos y fax, pero prevalece el servicio telefónico tradicional o **Servicio de Telefonía Básica** (***POTS:*** *Plain Old Telephone Service*) [1], el cual sigue empleando teléfonos para transmitir

voz analógica, y que seguramente seguirá funcionando por muchos años debido a su bajo costo y a la capacidad de transmisión ya instalada.

La **Red Telefónica Básica** (***PSTN:*** *Public switched telephone network*) [1], se compone principalmente de los siguientes elementos:

- Lazos locales.
- Troncales.
- Oficinas de cambios o centrales telefónicas.

En la **RTB** cada teléfono se conecta a una oficina de cambios (la local) a través de un par de hilos de cobre trenzados, cable al que se le conoce como lazo local. La central telefónica local a su vez se interconecta con otras centrales de diferentes tipos y cuya función es canalizar la comunicación de acuerdo a un número marcado en el teléfono (figura 2.1).

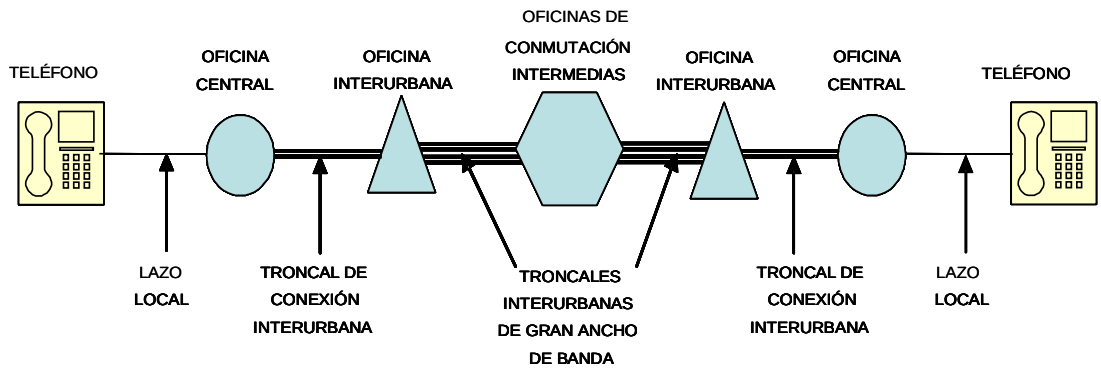


Figura. 2.1.- Ruta típica de un circuito para una llamada de media distancia. Tomada de [2].

Todo tipo de central consta de un equipo de conmutación, que permite seleccionar el puerto de salida adecuado para comunicarse con el teléfono al que se desea llamar, y de equipos que transmiten señales de unas centrales a otras.

Si un abonado telefónico conectado a una oficina central llama a otro abonado conectado a la misma oficina central, la conmutación dentro de la oficina establece una conexión directa entre los dos lazos locales. Si el teléfono al que se llama esta conectado a otra oficina central, el procedimiento de conexión es diferente. Cada oficina central tiene varias líneas salientes a otros centros de conmutación cercanos, llamados **oficinas interurbanas.** Estas líneas salientes se llaman **troncales de conexión interurbanas.** Si el teléfono que llama (origen) y el que es llamado (destino) no tienen una oficina interurbana en común, el enlace se establece recurriendo a centrales que están en un nivel más alto de la jerarquía de centrales. Existen oficinas primarias, seccionales y regionales que forman parte de una red que conecta a las oficinas interurbanas. Este tipo de oficinas se conectan entre sí mediante **troncales interurbanas** de ancho de banda grande.

También hay redes modernas con centrales que no están organizadas en estructura jerárquica, es decir, “redes ajerárquicas”, sin embargo esto no repercute en el desarrollo de esta tesis.

2.1.1. El teléfono y el lazo local.

Un teléfono permite originar y recibir llamadas telefónicas, ésta sencilla operación consta de varias funciones, siendo las más importantes:

- Realiza una petición del uso del sistema telefónico por el simple hecho de levantar el microteléfono (la manija que contiene el auricular y el micrófono).
- Indica que el sistema está listo para usarse cuando recibe un “tono”, conocido como “tono de invitación a marcar”.
- Envía a la central local el “número de directorio” correspondiente al teléfono destinatario.
- Indica al usuario el estado de la llamada mediante la recepción de distintos tonos (ocupado, timbrado, entre otros)
- Indica un tipo de llamada entrante al teléfono llamado mediante diferentes tipos de timbrado.
- Convierte las ondas sonoras de la voz, en señales eléctricas para su transmisión a través del sistema telefónico hasta otro teléfono, que convierte las señales eléctricas en ondas sonoras.

- Realiza automáticamente los ajustes necesarios para compensar los cambios de la potencia que se le suministra, y en especial, compensa los efectos de las diferentes longitudes del lazo local de tal forma que el usuario no se preocupa de la distancia entre su predio y la central local.
- Permite la señalización entre el sistema telefónico y el usuario. Todas las señales entre el aparato telefónico y la central que no son la voz u otras señales externas de micrófono a auricular, constituyen la señalización del lazo local. Tales son la solicitud de servicio (descolgar), terminar (colgar), los bitonos, o pulsos, que indican los números marcados, y los tonos escuchados correspondientes a las distintas notificaciones que da el sistema telefónico al usuario.

En el lazo local la comunicación es bidireccional simultánea (*full duplex*). El lazo local es un medio físico y nunca es en sí mismo digital; más, cuando transporta sólo señales digitales, se dice que la línea del subscriptor es digital (***DSL***: *digital subscriber line*).

En la figura 2.2 se muestra un diagrama simplificado de un teléfono convencional y las conexiones del lazo local hacia la central telefónica.

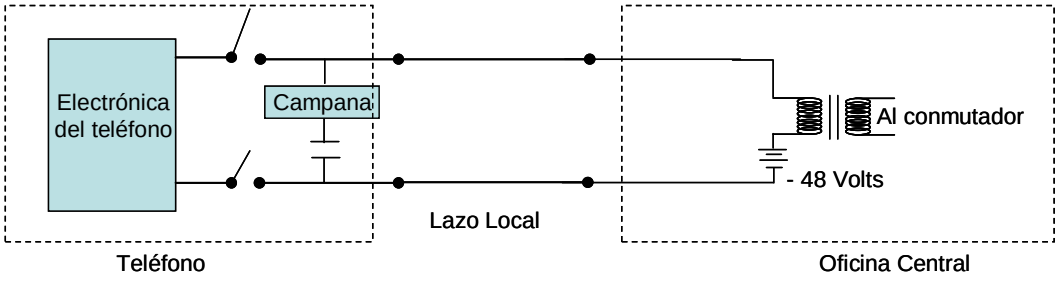


Figura 2.2.- Diagrama esquemático de las conexiones entre el teléfono, lazo local y la oficina central. Adaptada de [5].

La central telefónica aplica un voltaje de corriente directa a través del par trenzado de la línea del teléfono. En las condiciones de circuito abierto, este voltaje es normalmente -48 volts en uno de los hilos con respecto al otro hilo, el cual presuntamente esta a tierra. Como los aparatos telefónicos actuales están provistos casi siempre de protección para errores de polarización, es menos importante ahora distinguir los hilos.

Cuando se descuelga el auricular se conecta la electrónica del teléfono a la línea telefónica. La carga de los circuitos provoca que fluya corriente en el lazo local y que el voltaje dentro del teléfono descienda a veces casi hasta 5 volts, según la longitud del lazo.

Para cada teléfono conectado a la central telefónica, ella provee un grupo de circuitos básicos que alimentan al teléfono y proporcionan todas las funciones de señalización local como son: timbrado, tono de marcación y supervisión de la marcación. A estos circuitos se les denomina interfaz del suscriptor o **circuitos de interfaz con la línea del suscriptor** (***SLIC:*** *Suscriber Line interface Circuits*). La **interfaz de línea** proporciona 7 funciones básicas llamadas ***BORSCHT*** (acrónimo de las funciones en inglés *battery, overvoltge protection, ringing, supervision, coding, hybrid y test*: batería, protección de sobrevoltaje, llamada, supervisión, codificación, hibridación de direcciones y prueba) [3].

La resistencia total de todos los componentes del lazo local, incluyendo la resistencia del teléfono, debe ser lo bastante pequeña como para permitir que fluya una corriente suficiente entre el teléfono y el equipo de conmutación. La corriente de operación de un teléfono está entre 24 y 60 mA, con un valor óptimo de 35 mA. La resistencia de los teléfonos actuales es aproximadamente de 600 Ω [4].

2.1.2. Establecimiento del enlace entre dos estaciones conectadas a la misma central.

Una oficina de cambios tiene varios conmutadores que automáticamente establecen la conexión entre un puerto de entrada y uno de salida. Supongamos que la conexión ya fue establecida.

Si el teléfono destino está “descolgado” cuando existe un intento de conexión, la oficina de cambios genera un **tono de ocupado** que es enviado al teléfono que realiza la llamada. Si el receptor está “colgado”, se envía una **señal de timbrado** al teléfono destino para indicar que el teléfono origen desea establecer una comunicación; al mismo tiempo la central envía un **tono de notificación de timbrado** al teléfono llamador para indicar que el teléfono llamado está timbrando. Cuando se descuelga el microteléfono en el teléfono llamado en respuesta al timbrado, se cierra el circuito de su lazo local y fluye una corriente en el teléfono destino. La central detecta la corriente y entonces suprime ambas: las señales de **timbre** y **notificación de timbrado**. Ahora la conversación puede llevarse acabo. En muchos sistemas telefónicos la llamada termina sólo cuando el extremo origen cuelga el microteléfono, con objeto de permitir al usuario llamado cambiar de extensión. Al terminar la llamada, la oficina de cambios libera la conexión implantada de las líneas a través de sus conmutadores.

La banda de paso en un sistema telefónico reservado para un canal de voz es de 4000 Hz., sin embargo, no toda esta banda es utilizada para la transmisión de la voz. La banda base asignada a la voz va de 300 a 3400 Hz. Esta banda base es suficiente para un sonido con **calidad telefónica** (*telephone voice grade*).

2.1.3. Transmisión de datos en la red telefónica básica.

La **RTB** fue diseñada para la transmisión de voz, y no para la transmisión de datos. En particular, las frecuencias muy bajas se rechazan para evitar la interferencia de las líneas de potencia, es decir, se rechazan los 60 o 50 Hz, y junto con ellos, la componente de directa. Por tanto, los pulsos de tope plano, que contienen la mayor cantidad de su energía en directa, no pasan.

Cuando se requiere enviar pulsos rectangulares sobre el lazo local, es necesario trasladarlos en frecuencia utilizando para ello un módem telefónico. Las señales analógicas, ya sean voz o pulsos modulados, se convierten a formato digital en la oficina de cambios para transmitirlos sobre las troncales, las cuales tienen anchos de banda y repetidores adecuados. En el otro extremo de la oficina de cambios se realiza la conversión inversa – digital a analógico – para recorrer el circuito local destino, el cual es de banda angosta. Además

de que el módem permite la transmisión de datos, un módem moderno tiene capacidades añadidas para disminuir el efecto de los principales problemas que sufren las señales transmitidas sobre la red telefónica: atenuación, distorsión y ruido.

La parte del modulador del módem traslada la señal digital a una banda de frecuencias adecuada para transmisión por un canal telefónico en la red **RTB**, y la parte del demodulador recibe la señal del canal y la reconvierte a su formato digital original.

En la figura 2.3 se muestran dos equipos que se conectan a través del lazo local para la transmisión de datos. La salida de datos del equipo terminal de datos es en paralelo, luego se convierte a un flujo de datos en serie a través de un dispositivo llamado **Transmisor-Receptor Asíncrono Universal TRAU** (***UART:*** *Universal Asynchronous Receiver Transmitter*). El flujo serial de bits consiste de pulsos rectangulares sucesivos, por lo que dicho flujo debe modularse. Un módem proporciona esta conversión y acoplamiento a la línea telefónica. En la parte receptora, el módem convierte la señal recibida a un flujo serial de pulsos rectangulares y estos a una forma paralela por el transceptor **TRAU** para presentar carácter por carácter al bus de la computadora destinataria.

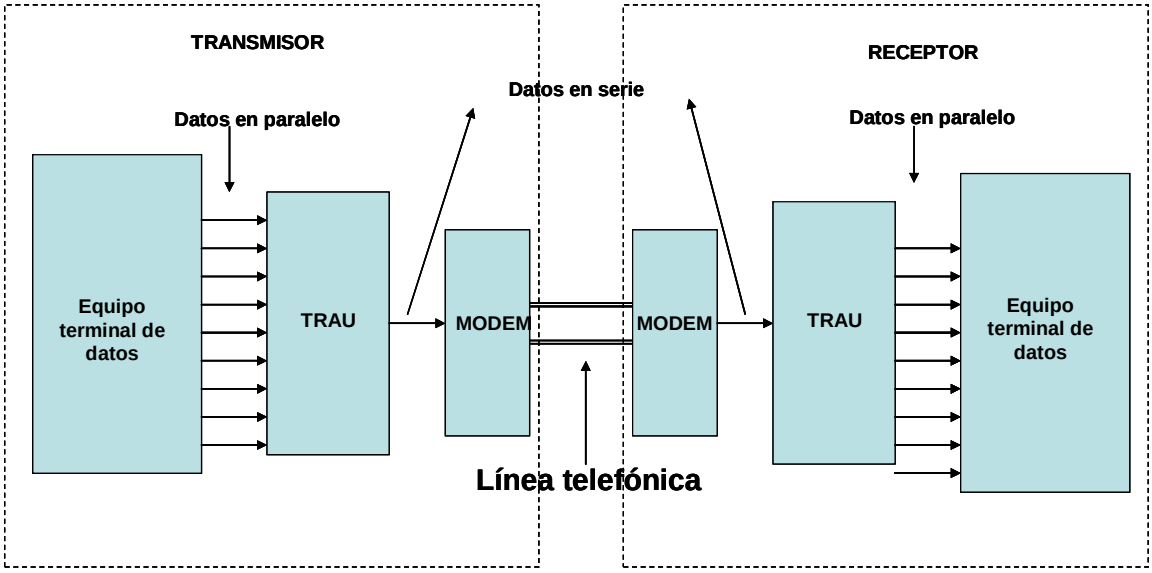


Figura 2.3.- Transmisión de datos sobre el lazo local. Tomada de [5].

2.2. SEGURIDAD DE LA INFORMACIÓN A TRAVÉS DEL CIFRADO.

2.2.1. Sistemas de cifrado.

La criptología es el término que describe la ciencia de las comunicaciones secretas: se deriva del griego *kryptos* y *logos,* que significan “oculto” y “palabra”, respectivamente [6]. La criptología se divide en criptografía y criptoanálisis. La primera tiene que ver con las transformaciones de un mensaje en forma codificada a través del cifrado, y la recuperación del mensaje original mediante el descifrado. El mensaje original que se va a cifrar recibe el nombre de **texto en claro,** y el resultado que produce el cifrado se conoce como **criptograma o texto cifrado.**

Los sistemas criptográficos persiguen los siguientes objetivos [7]:

- Confidencialidad.- Consiste en mantener la información fuera del alcance de personas sin autorización, para que no puedan entender o interpretar el mensaje.
- Integridad de datos.- Se refiere a asegurar que el mensaje no sea modificado por medios accidentales o deliberados durante el tránsito.
- Autentificación.- Es un servicio relacionado con la identificación. Se refiere a la validación de la fuente del mensaje.
- Impedimento de retracción.- Servicio que previene la negación de compromisos o acciones, es decir, el agente no puede “echarse atrás” o “meter reversa”.

Un sistema criptográfico convencional se sustenta en el uso de una pieza simple de información privada y necesariamente secreta conocida como la **clave;** por tanto, la criptografía convencional se conoce como **criptografía de clave simple, criptografía de clave secreta o criptografía de clave simétrica.** En esta forma de criptografía, la clave es conocida por el cifrador (emisor) y el descifrador (receptor), pero no otros; luego de que se cifra el mensaje,

resulta casi imposible realizar el descifrado sin el conocimiento de la clave.

Para entender el proceso de cifrado simétrico, se tomará como base un sistema como el de figura 2.4.

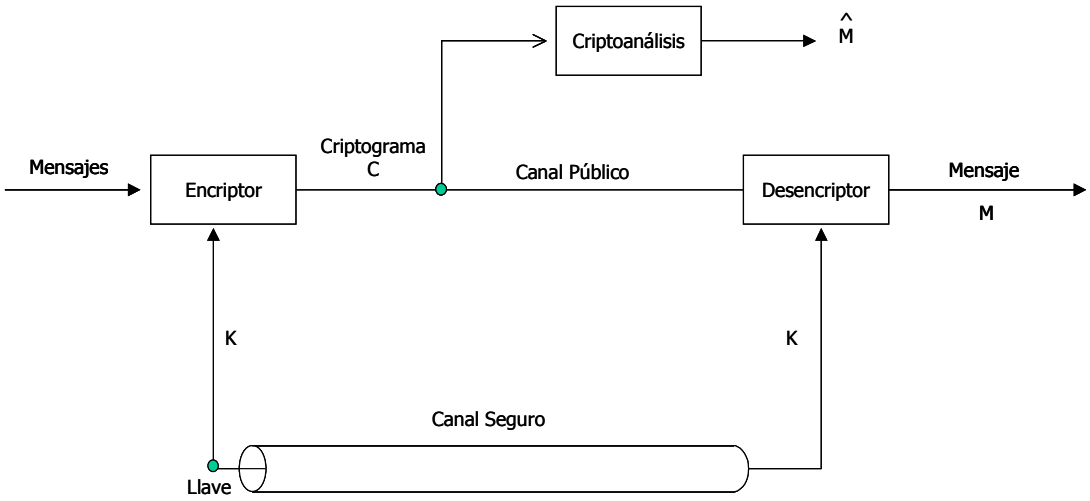


Figura 2.4.- Modelo de un canal criptográfico. Adaptada de [16].

En el modelo de canal criptográfico, un mensaje *M*, es cifrado usando una transformación o algoritmo invertible, E_K, la cual produce un criptograma *C*.

$$C = E_K(M) \text{ ----------(1)}$$

El criptograma es transmitido a través de un canal inseguro o canal público. Cuando un receptor autorizado recibe el criptograma *C*, lo descifra usando la transformación inversa, $D_K = E^{-1}_K$, con lo cual obtiene el mensaje original.

$$D_K(C) = E^{-1}_K(E_K (M)) = M \text{ ----------(2)}$$

El parámetro *K* se refiere a un conjunto de símbolos o caracteres llamados **claves o llaves**, estas claves especifican una transformación de cifrado en particular E_K. Originalmente, la seguridad del esquema de criptografía dependía de la confidencialidad de todo el proceso de cifrado, esto es, las transformaciones invertibles E_K y las llaves *K* utilizadas en cada transformación, pero actualmente las transformaciones *E*, pueden ser públicas, debido a que la confidencialidad de los mensajes, depende de la privacidad de las claves *K* utilizadas. Estas claves son distribuidas a los usuarios autorizados a través de un canal seguro.

La criptografía de clave pública o criptografía asimétrica, es un método criptográfico que emplea un par de claves para el envío de mensajes cifrados, una clave pública que puede ser conocida por cualquier persona, y una clave privada, la cual debe mantenerse en secreto por su propietario.

Los sistemas de cifrado de clave pública se basan en funciones cuya evaluación es fácil, pero despejar la función inversa resulta extremadamente difícil, a menos que se conozca cierto dato de la función, por ejemplo, es fácil multiplicar dos números primos entre sí para obtener el producto, pero es difícil factorizar un número en sus componentes primos; sin embargo, si se conoce un factor, es inmediato calcular el otro.

En general, el propietario de la clave privada es el mismo que difunde la clave pública, dado que dichas claves están apareadas y son generadas simultáneamente por el mismo algoritmo. Sin embargo, conocer el método que genera las dos claves al mismo tiempo, no acelera el cálculo de la claves privada correspondiente a una clave pública dada, cálculo que en sí le toma demasiado tiempo al criptoanalista espía, del orden de meses o años.

El sistema de criptografía pública se puede expresar como $P = D_{K2}(E_{K1}(P))$ donde *K1* es la clave pública y *K2* es la clave privada (figura 2.5). El transmisor utiliza la clave pública *K1* para cifrar el mensaje *P* y transmite la información cifrada al receptor, el cual podrá descifrar el mensaje sólo si posee la clave privada *K2.*

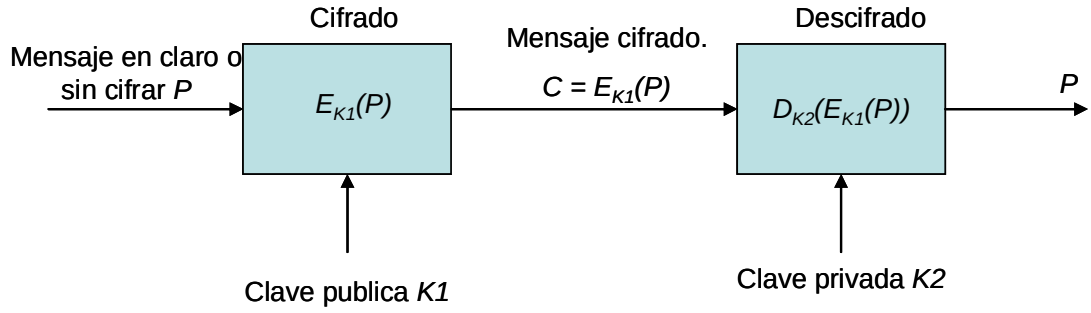


Figura 2.5.- Criptografía .de clave pública. Tomada de [19].

2.2.2. Encubridores de bloque y de flujo.

Los sistemas criptográficos pueden dividirse en dos clases: **encubridores de bloque y encubridores de flujo.** Los primeros operan de una manera combinatoria sobre grandes bloques de texto en claro, en tanto que los de secuencia o de flujo procesan el texto por caracteres o por bits. En un cifrado de bloques el texto en claro se divide en bloques, cada uno de los cuales suele integrarse con un número fijo de bits. Los bloques sucesivos de texto en claro se cifran utilizando la misma clave secreta, de manera independiente (figura 2.6).

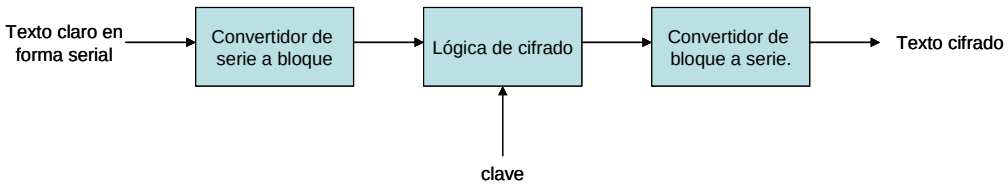


Figura 2.6.- Diagrama de un cifrado de bloque. Tomada de [8].

Los encubridores de bloque operan con una transformación fija aplicada a grandes bloques de datos de texto en claro, en un esquema de bloque por bloque. En contraste, un encubridor de flujo opera con base en una transformación variable en el tiempo aplicada a bits individuales del texto en claro. Los encubridores de flujo más populares son los llamados **encubridores de secuencia aditivos binarios,** como el que se muestra en la figura 2.7.

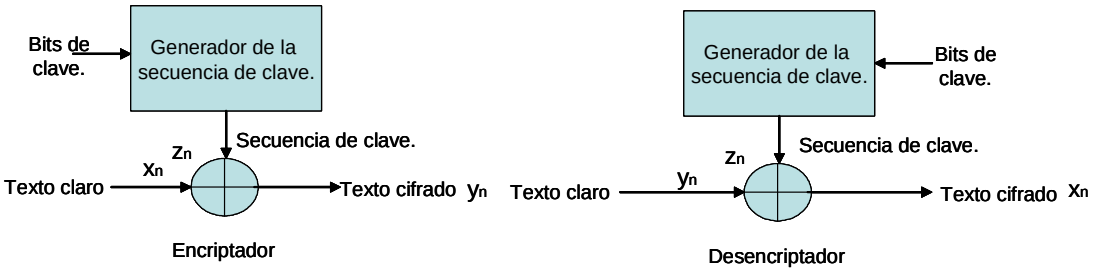


Figura 2.7.- Cifrador de Secuencia aditivo binario. Tomada de [8].

En un encubridor de este tipo la clave secreta se usa para controlar **un generador de secuencia de clave** que emite una secuencia binaria, llamada «**secuencia de clave»** cuya longitud es mucho mayor que la de la clave. Consideremos que x_n, y_n *y* z_n denotan los bits de texto en claro, el bit de texto cifrado y el bit de la secuencia de clave en el tiempo *n* respectivamente. Los bits de texto cifrado se determinan entonces mediante una suma módulo 2 de los bits de texto en claro y los bits de la secuencia clave:

$$y_n = x_n \oplus z_n \ , \quad n = 1, 2,\ldots, N \quad \text{---------- (3)}$$

donde *N* es la longitud de la secuencia de clave. Debido a que la suma y la sustracción en la aritmética módulo 2 son exactamente iguales, tenemos:

$$x_n = y_n \oplus z_n \ , \qquad n = 1, 2,\ldots, N \quad \text{---------- (4)}$$

por lo tanto, en los encubridores de secuencia es posible utilizar dispositivos idénticos para efectuar el cifrado y descifrado (Fig. 2.5); la clave secreta se elige de acuerdo con alguna distribución de probabilidad.

En los encubridores de bloque un pequeño cambio en un bloque de entrada de texto claro produce un cambio mayor en la salida resultante. Esta propiedad de propagación de errores de los encubridores de bloque es valiosa en la autentificación en el sentido que hace improbable que un enemigo modifique los datos cifrados, a menos de que conozca la clave. En cambio, un cifrado de flujo no tiene propagación de errores. El descifrado de un bit distorsionado en el texto cifrado afecta solamente al bit correspondiente de la salida resultante. Los encubridores de flujo son por lo general más adecuados para la transmisión de datos por canales de comunicación

propensos a errores; se usan en aplicaciones donde las altas velocidades de datos son un requerimiento.

2.2.3. Algoritmo estándar de cifrado 2001, AEC-2001 (*AES*).

Para proporcionar seguridad en la información, existen varios algoritmos y sistemas criptográficos, algunos de los cuales ya han quedado obsoletos puesto que han sido atacados exitosamente. Sin embargo, se han inventado nuevos algoritmos criptográficos que resisten los ataques basados en el uso de computadoras de alto desempeño.

En 1997, El Instituto Nacional de Estándares y Tecnología de los Estados Unidos de América (por sus siglas en inglés ***NIST***, *National Institute of Standards and Technology*) anunció una iniciativa para desarrollar un nuevo estándar de cifrado. El proceso de selección del nuevo estándar fue abierto a todo el público y cualquiera podía registrar un candidato. La agencia ***NIST*** del gobierno estadounidense no realizó la evaluación de la seguridad y eficiencia de los algoritmos, pero invitó a toda la comunidad de criptografía a que emplearan los distintos ataques posibles a los algoritmos de cifrado candidatos [9].

En este contexto, los algoritmos seleccionados por la agencia ***NIST*** como finalistas para el nuevo estándar de cifrado (*Advanced Encryption Standard*) fueron: *MARS, RC6, RIJNDAEL, SERPENT* y *TWOFISH* [2]. La agencia ***NIST*** anunció el 2 de 0ctubre del 2001, que el algoritmo Rijndael fue seleccionado como el «**nuevo algoritmo estándar de cifrado, o nuevo AEC**» o ***AES*** (*Advanced Encryption Standard*), debido a su seguridad, desempeño, eficiencia, flexibilidad y facilidad de implementación. Como este algoritmo fue publicado como estándar el 26 de noviembre de 2001, y como goza de amplia aceptación internacional, aquí se le llamará **AEC-2001.**

Todas las operaciones usadas en el algoritmo **AEC-2001** son operaciones de cambio de posición de los octetos (*bytes*) y operaciones sobre el campo finito ***GF*** (2^{8}). El tamaño de las entradas para el algoritmo **AEC-2001**, tanto para el texto como para la clave varía entre 128 y 256 bits considerando un incremento de 32 bits. La agencia ***NIST*** sólo estandarizó la versión con un tamaño de texto de 128 bits y un tamaño de clave de 128, 192 y 256 bits.

El funcionamiento del algoritmo **AEC-2001** se divide en dos partes, la primera en el proceso de cifrado y la segunda en el proceso de generación de las subclaves (figura 2.8). Como se ha mencionado, **AEC-2001** funciona con bloques de datos de 128 bits y longitudes de claves de 128, 192, 256 bits.

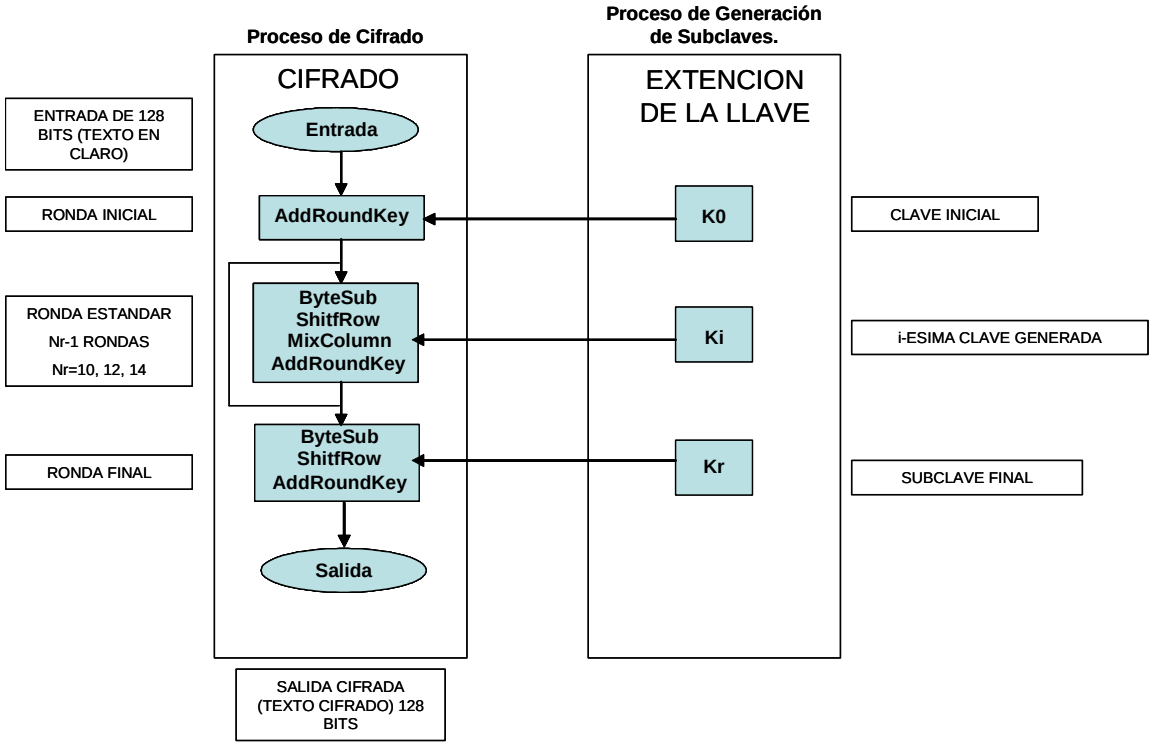


Figura 2.8.- Esquema sobre el funcionamiento del algoritmo **AEC-2001**. Tomada de [17]

Además **AEC-2001** tiene 10, 12 ó 14 vueltas ó rondas (*Round*), respectivamente, cada iteración de **AEC-2001** consiste en la aplicación de una ronda estándar que consiste de 4 transformaciones básicas, la última ronda es especial y consiste de 3 operaciones básicas, añadiendo siempre una ronda inicial. Por otro lado tenemos el proceso de claves o extensión de la clave, en donde el número de subclaves generadas es el número de rondas que aplica el algoritmo [10].

Las transformaciones básicas que aplica el **AEC-2001** son: **Sustitución de octetos (***ByteSub*), **Corrimiento circular de renglones** (*ShiftRow)*, **Mezclado de columna por columna** (*MixColumns*), y **Adición de llave propia de la ronda** (*AddRoundKey*).

El algoritmo **AEC-2001** interpreta al bloque de entrada de 128 bits, como una matriz 4x4 de entradas de octetos, si el bloque es de 192 bits se agregan 2 columnas más, si lo es de 256 se agregan 4 columnas más, por lo tanto, el número de columnas es determinado por el tamaño del bloque: si el tamaño del bloque es *N*, el número de columnas N_b es igual a *N/32* (figura 2.9). El resultado intermedio del proceso de cifrado se conoce como **estado** de un bloque de código. El estado es representado por un arreglo rectangular de octetos el cual siempre contiene 4 filas llamada **matriz de estado**.

$$\begin{matrix} \mathbf{a_{00}} & \mathbf{a_{01}} & \mathbf{a_{02}} & \mathbf{a_{03}} \\ \mathbf{a_{10}} & \mathbf{a_{11}} & \mathbf{a_{12}} & \mathbf{a_{13}} \\ \mathbf{a_{20}} & \mathbf{a_{21}} & \mathbf{a_{22}} & \mathbf{a_{23}} \\ \mathbf{a_{30}} & \mathbf{a_{31}} & \mathbf{a_{32}} & \mathbf{a_{33}} \end{matrix}$$

Figura 2.9.- Bloque de entrada al algoritmo **AEC-2001** en una matriz de 4X4.

La clave usada en el algoritmo se despliega como un arreglo rectangular de octetos con 4 filas, el número de columnas de la clave

está determinado por N_K y es igual al tamaño de la clave dividido por 32. Algunas veces la clave es presentada como un arreglo lineal de palabras de 4 octetos. En este caso, las palabras están distribuidas dentro de cada columna. El número de vueltas está descrito por N_r y depende de los valores de N_b y $N_{K.}$

Proceso de Cifrado.

El **AEC-2001** tiene estas etapas:

- Una transformación inicial de **Adición de clave.**
- *Nr -1* rondas.
- Una ronda final.

Cada ronda consta de 4 transformaciones básicas, las cuales. como ya se dijo, son: **Sustitución de octetos, Corrimiento circular de renglones, Mezclado columna por columna,** y **Adición de clave propia de la ronda** La vuelta final del algoritmo es ligeramente diferente, en ésta, la transformación **Mezclado de columna por columna** ha sido eliminada.

El bloque de entrada de 128 bits con la matriz 4x4 de entradas de octetos, está asociado de la siguiente forma:

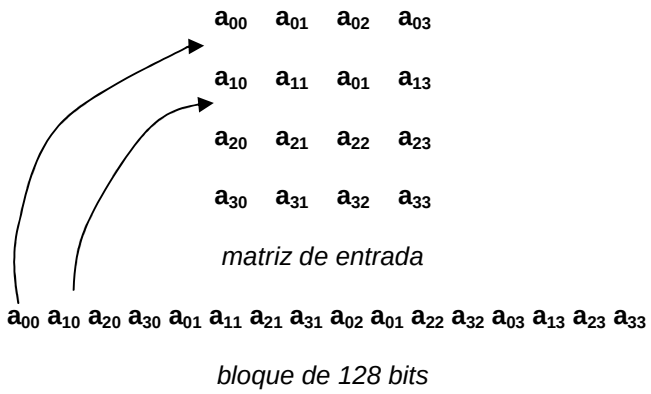


Figura 2.10.- Correspondencia del bloque de 128 bits con la matriz de entrada. Tomada de [17].

Los cuatro primeros octetos se toman para formar la primera columna, los segundos cuatro octetos la segunda columna, y así sucesivamente. La matriz es la entrada del algoritmo **AEC-2001**, y va cambiando en cada una de las rondas, esta matriz será $[a_{ij}]$ y se denomina **matriz estado.** A continuación se describe cada una de las transformaciones. Cada ronda tiene su propia clave, o «subclave», derivadas todas de la primera clave. En esta transformación, la matriz de estado (en la ronda inicial es la matriz de estado inicial con los datos a cifrar) se modifica por la combinación de ella misma con la matriz formada con la subclave de la ronda correspondiente a través de una operación *XOR*.

Adición de clave propia de la ronda.

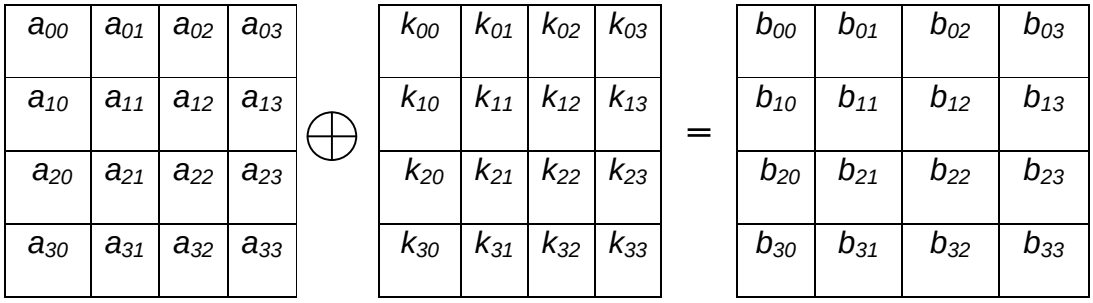


Figura 2.11.- Transformación adición de clave (AddRoundKey). Tomada de [9]

En la figura 2.12 observamos que la transformación toma los elementos $[a_{ij}]$ y $[k_{ij}]$ para realizar una operación *XOR,* de la matriz del bloque con la matriz de la subclave correspondiente.

Sustitución de octetos:

A cada elemento de la matriz estado (octeto) se le substituye por otro octeto, el está en función del octeto a reemplazar.

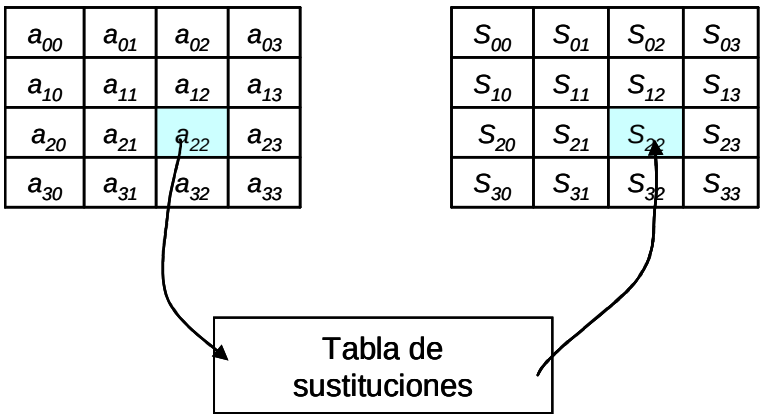


Figura 2.12.- Transformación Sustitución de octetos. Tomada de [9].

Los octetos que sustituyen a los elementos de la matriz de estado se encuentran ya calculados en una tabla llamada **Tabla de sustituciones** (*S-box*) [9].

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	63	7C	77	7B	F2	6B	6F	C5	30	01	67	2B	FE	D7	AB	76
1	CA	82	C9	7D	FA	59	47	F0	AD	D4	A2	AF	9C	A4	72	C0
2	B7	FD	93	26	36	3F	F7	CC	34	A5	E5	F1	71	D8	31	15
3	04	C7	23	C3	18	96	05	9A	07	12	80	E2	EB	27	B2	75
4	09	83	2C	1A	1C	6E	5A	50	52	3B	D6	B3	29	E3	2F	84
5	53	D0	11	ED	20	FC	B1	5B	6A	CB	BE	39	4C	4C	58	CF
6	D0	EF	AA	FB	43	4D	33	85	45	F9	02	7F	50	3C	9F	A8
7	51	A3	40	8F	92	9d	38	F5	BC	B6	DA	21	10	FF	F3	D2
8	CD	OC	13	EC	5F	97	44	17	C4	A7	7E	3D	64	5D	19	73
9	60	81	4F	DC	22	2A	90	88	46	EE	B8	14	DE	5E	0B	DB
A	E0	38	3ª	0A	49	06	24	5C	C2	D3	AC	62	91	95	E4	79
B	E7	C8	37	6D	8D	D5	4E	A9	6C	56	F4	EA	65	7A	AE	08
C	BA	78	25	2E	1C	A6	B4	C6	E8	DD	74	1F	4B	BD	8B	8A
D	70	3E	B5	66	48	03	F6	0E	61	35	57	B9	86	C1	1D	9E
E	E1	F8	98	11	69	D9	8E	94	9B	1E	87	E9	CA	55	28	DF
F	8C	A1	89	0D	BF	E6	42	68	41	99	2D	0F	B0	54	BB	16

Figura 2.13 Tabla de sustitución **AEC-2001**. Tomada de [9].

Cada elemento de la tabla S, [S_{ij}], es resultado de aplicar dos funciones a cada elemento de la matriz de estado [a_{ij}] : 1) su inverso multiplicativo $a_{ij} \rightarrow a_{ij}^{-1}$ ambos en $GF(2^8)$, y posteriormente 2) una transformación lineal en $GF(2^8) \rightarrow GF(2^8)$ [9].

Desde un punto de vista práctico, adecuado para su programación, con los cuatro primeros bits del octeto sustituido se localiza el renglón de la tabla de sustitución, y con los últimos cuatro bits se ubica la columna donde se encuentra el elemento sustituto.

Corrimiento circular de renglones:

La transformación **Corrimiento circular de renglones** se aplica a la matriz estado, a través de corrimientos circulares hacia la izquierda de los elementos de sus renglones. Los elementos que salen del arreglo por su lado izquierdo, van ocupando los huecos dejados en el lado derecho del renglón, En los renglones de la matriz $[a_{ij}]$ se realizan corrimientos izquierdos circulares de octetos de la siguiente manera: recorre cero octetos el primer renglón, un octeto al segundo renglón, dos octetos al tercer renglón, y tres octetos recorridos al cuarto renglón.

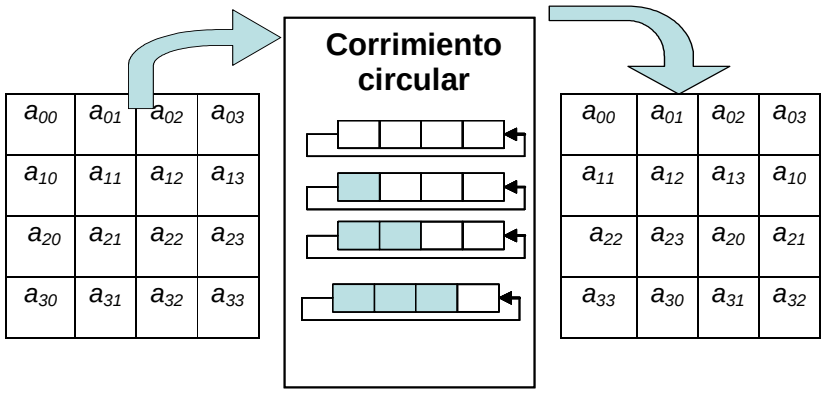


Figura 2.14. Transformación corrimiento circular de renglones. Tomada de [9].

Batido Columna por Columna (*MixColumns*)

Esta transformación se aplica sobre los octetos de cada columna de la matriz de estado $[a_{ij}]$, considerando estas columnas como polinomios cuyos coeficientes pertenecen al **Campo de Galois *GF***(2^8), es decir, octetos de bits.

El «batido columna por columna» consiste en multiplicar, módulo $x^4 + 1$, cada columna $[\ a_{ij}\ ]$ de octetos por un polinomio constante *c(x)* en el campo ***GF***(2^8):

$$c(x) = 03\ x^3 + 01\ x^2 + 01\ x + 02 \text{ ----------(5)}$$

Esta transformación se puede representar como:

$$\hat{\underline{a}}_j(x) = [c(x) \times \underline{a}_j(x)] \mod (x^4 + 1) \text{ ----------(6)}$$

donde $\underline{a}_j(x) = a_{0,j} + a_{1,j}\ x + a_{2,j}\ x^2 + a_{3,j}\ x^3$ es un polinomio que representa una columna de la matriz de estado (primera columna) a la cual se le aplica la transformación, y $c(x) = c_0 + c_1x + c_2x^2 + c_3x^3$ es el polinomio constante; al resultado obtenido de (6) se le aplica módulo $x^4 + 1$, y se puede representar en forma matricial como:

donde $\underline{a}_j(x) = a_{0,j} + a_{1,j}\ x + a_{2,j}\ x^2 + a_{3,j}\ x^3$ es un polinomio que representa a la columna “j” de la matriz de estado, a la cual se le

aplica la transformación, y $c(x) = c_0 + c_1x + c_2x^2 + c_3x^3$ es el polinomio constante; al resultado obtenido de (6) se reduce módulo $x^4 + 1$. La expresión (6) se puede representar en forma matricial como:

$$\begin{pmatrix} a'_{00} \\ a'_{01} \\ a'_{02} \\ a'_{03} \end{pmatrix} = \begin{pmatrix} c_0 & c_3 & c_2 & c_1 \\ c_1 & c_0 & c_3 & c_2 \\ c_2 & c_1 & c_0 & c_3 \\ c_3 & c_2 & c_1 & c_0 \end{pmatrix} \begin{pmatrix} a_{00} \\ a_{01} \\ a_{02} \\ a_{03} \end{pmatrix} \quad \text{----------(7)}$$

Para programar fácilmente la multiplicación de los polinomios, conviene representarlos en su «formato binario desglosado», como se ilustra en el siguiente ejemplo. Sea el polinomio $x^4 + x + 1$, que en forma su forma binaria equivale a *10011 (*en hexadecimal *13)*, se puede descomponer de la siguiente forma (notación en hexadecimal) empleando la operación ***XOR***:

$$01 \oplus 02 \oplus 10 \text{----------(8)}$$

La representación en (8) es muy importante ya que permite programar fácilmente la multiplicación de los polinomios, razón por la cual el polinomio constante *c(x)* se representa en forma hexadecimal.

Cómo el polinomio *c(x)* está en un campo (de Galois), tiene la propiedad de que es invertible, es decir, tiene un inverso multiplicativo, el cual es necesario para realizar el proceso de descifrado.

Por lo anterior, el «batido de columnas» se representa en forma matricial como sigue:

$$\begin{pmatrix} a'_{00} \\ a'_{10} \\ a'_{20} \\ a'_{30} \end{pmatrix} = \begin{pmatrix} 0x02 & 0x03 & 0x01 & 0x01 \\ 0x01 & 0x02 & 0x03 & 0x01 \\ 0x01 & 0x01 & 0x02 & 0x03 \\ 0x03 & 0x01 & 0x01 & 0x02 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

La siguiente figura muestra en forma gráfica la transformación «batido de columnas».

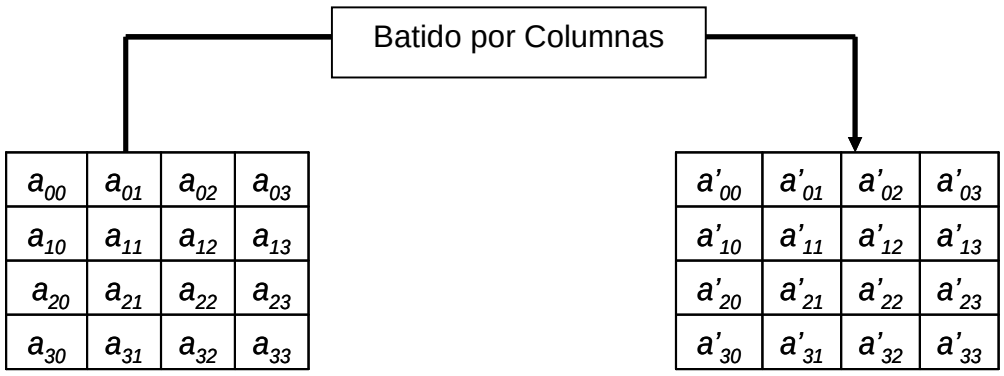


Figura. 2.15.- Transformación batido por columnas. Tomada de [9].

2.3. COMPRESIÓN DE LA VOZ.

La finalidad de la codificación o compresión de voz, es transformar la representación de la señal de voz para su almacenamiento o transmisión, de la manera más eficiente posible sin perder considerablemente su calidad.

En telecomunicaciones, un factor importante que debe tomarse en cuenta es la capacidad de canal disponible para la transmisión de información, especialmente las señales de voz, donde es importante mantener una comunicación en tiempo real y con una calidad aceptable. En el diseño de los sistemas de telecomunicaciones se considera el número de usuarios que pueden ser acomodados en un mismo medio de transmisión. Ya sea que se emplee multiplexión por división de tiempo o multiplexión por división de frecuencia, el número de usuarios está limitado por la capacidad de un enlace para un canal de voz. Esto es muy importante ya que en sistemas como la radiocomunicación o el satelital, la demanda de canales libres para voz, implica el empleo de codificadores o compresores de voz para que los servicios tengan un costo accesible.

En las comunicaciones digitales, la señal de voz es transformada en una secuencia de bits por medio de un codificador, se transmite a través de un canal y nuevamente se transforma en una señal audible por un decodificador.

Existen varios tipos de codificadores de voz, lo cuales pueden ser clasificados como siguen [10]:

- Codificadores de forma de onda (*Waveform*).
- Vocoders.
- Codificadores híbridos.

Los codificadores de forma de onda trabajan muestra por muestra de voz, el objetivo en estos codificadores es obtener a la salida del decodificador una forma de onda lo más parecida a la señal de voz original. Este tipo de codificadores proporcionan una alta calidad de voz a velocidades medias, del orden de 32 kb/s. Sin embargo, no son útiles cuando se quiere codificar a bajas velocidades.

La figura 2.16 muestra el procesamiento básico general de estos codificadores de onda:

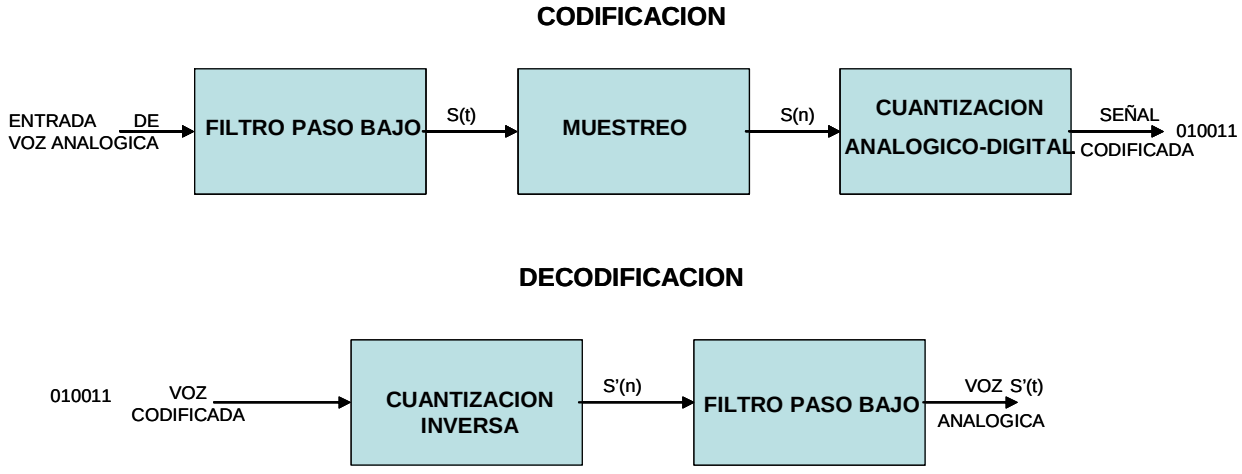


Figura 2.16.- Codificación de forma de onda. Tomada de [6].

Para este tipo de señales, cerca del 1% de la energía se encuentra arriba de los 4KHz. y una parte insignificante arriba de los 7 KHz. En el sistema telefónico usualmente emplean anchos de banda de 0.3 a 3.4 KHz.,

proporcionando una señal de voz con un mínimo de degradación y una calidad perceptible para los usuarios [10].

La señal de voz limitada en banda es sometida a un proceso de muestreo empleando una frecuencia de muestreo según el teorema de Nyquist, produciéndose muestras discretas en el tiempo. En telefonía la frecuencia empleada es de 8 KHz.

La cuantización es el proceso fundamental en la codificación de la voz, que a diferencia del muestreo, este proceso introduce error en la señal decodificada. En esta etapa se representan las muestras de la señal en valores discretos, de tal forma que el número de niveles que se utilicen, determinará el error de cuantización introducido.

La forma más simple de codificación de voz es utilizando una cuantización uniforme, utilizando generalmente 16 bits por muestra, de tal manera que si la señal de voz es muestreada a 8000 Hz. se obtiene una velocidad de codificación de 128 kbps (en modo Stereo). Sin embargo debido a que en una señal de voz la distribución temporal no es uniforme, si se utiliza una cuantización uniforme no se aprovecha adecuadamente el número de bits utilizados para representar la señal, incrementándose la relación *S/N*. Para este problema se emplea la cuantización logarítmica, proporcionando un error de cuantización más uniforme.

Algunos codificadores de onda son por ejemplo: **Codificación para Impulsos Modulados CIM (*PCM:* *Pulse code modulation*)** que codifica una tasa de bits de 64 kbps o **Código de Impulsos para la Diferencia entre Niveles Adaptables CIDNA** (***ADPCM:*** *Adaptive Differential Pulse Code Modulation*) que codifica a 16, 24, 32 y 40 kbps.

Los codificadores de fuente (*vocoders*) tratan de describir una señal de voz en términos de los parámetros de un modelo de producción y recepción de voz a través de señales de excitación y filtros. Estas tecnologías han logrado disminuir la tasa de transmisión considerablemente por debajo de los 2.5. kbps.

En el tercer grupo, los codificadores híbridos, son aquellos que combinando técnicas de los codificadores de fuente con filtros y señales de excitación y de los codificadores de forma de onda aumentan las ventajas de ambos, permitiendo una alta calidad de voz a bajas velocidades. Estos codificadores también son conocidos como **codificadores de análisis por síntesis (*AbS:* *Analysis by Synthesis*)**

Los codificadores de predicción lineal como el de **predicción lineal con excitación de señales combinadas** (***MELP:*** *Mixed-Excitation Linear Predictive*) alcanzan velocidades de hasta 2.4 kbps. Dentro de los codificadores híbridos se encuentran el codificador de **predicción lineal con excitación de señal residual** (***RELP*** : *Residual Excited Linear Prediction*) y el

Codificador de Predicción Lineal con Señal de Excitación del Filtro Seleccionada Según un Código, (*CELP: Code Excited Linear Prediction*).

2.3.1. Modelo Básico de Codificadores de Voz mediante Filtros y señales de Excitación.

La voz es producida por un conjunto de órganos compuesto por los pulmones, la epiglotis, las cuerdas vocales, la cavidad bucal y nasal. La figura 2.17 muestra una sección del cuerpo humano con los órganos involucrados. Los sonidos que conforman la voz se pueden clasificar en entonados (originados en las cuerdas vocales) y sin tono (originados por una fricción en el tracto vocálico), en la práctica la voz está formada por una mezcla de ambos. La señal de voz entonada contiene alto contenido de energía y es casi periódica y el sonido sin tono luce más como ruido aleatorio sin periodicidad.

Para la producción de una señal entonada, los pulmones expulsan aire por la epiglotis haciendo vibrar las cuerdas vocales. Ellas interrumpen el flujo de aire produciendo una onda de presión casi periódica. Los impulsos que son producidos corresponden a una señal que se denomina **tono** o **frecuencia fundamental (*pitch*)** y tiene por función determinar la melodía de la voz de las personas. Al flujo de aire modulado por el **tono** se le denomina señal de excitación,

la cual estimula el aire en la cavidad bucal y nasal resonando y emitiendo una señal de voz.

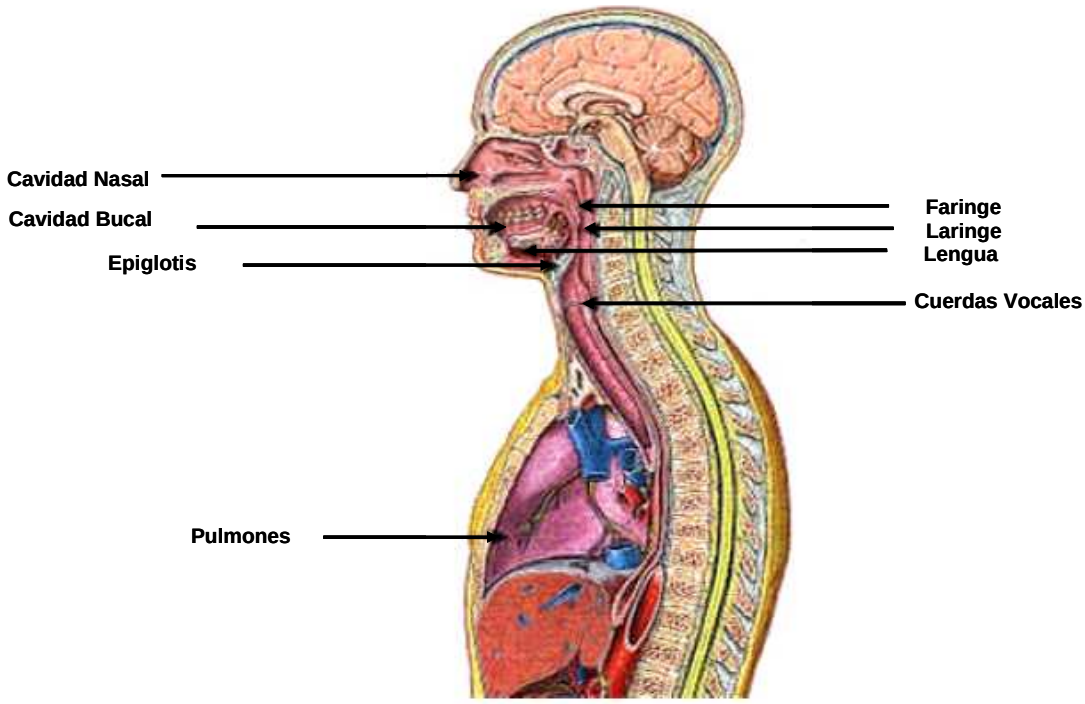


Figura 2.17.- Tracto Vocal Humano. Tomada de [11].

Gracias a la lengua, la cavidad de resonancia puede adoptar diferentes posiciones permitiendo al ser humano pronunciar diferentes sonidos. Para la producción de una señal sin tono, la excitación del tracto vocal es semejante al ruido. Durante el proceso de generación de sonidos no vocalizados, las cuerdas vocales están completamente abiertas, posibilitando la circulación del aire por el tracto vocálico, la que se ve ligeramente obstaculizada por el roce con las paredes del tracto, produciendo un ruido fricativo.

Además del movimiento de las cuerdas vocales y del tracto vocálico, para modelar el proceso de generación de voz se debe considerar también los movimientos de la boca, la lengua, los labios y vibraciones nasales. Por tanto, un modelo básico de este proceso debe considerar lo siguiente:

- La voz es una señal que emerge de una fuente definida: los pulmones actúan como emisores de aire y la señal se produce por la vibración de las cuerdas vocales y la posterior resonancia con las paredes del tracto vocálico.
- La voz está formada por la mezcla de señales de excitación periódica y ruido.
- La variación temporal de la señal en el tracto vocálico produce el timbre característico que diferencia los fonemas, ciertos fonemas son articulados sin la presencia de las cuerdas vocales.
- La fuente emisora posee dos estados: generación de sonidos vocalizados (entonados) y no vocalizados (sin tono).
- Si se toman intervalos de tiempos pequeños se puede modelar el órgano generador de la voz mediante varias funciones de transferencia, las cuales definen a los filtros

que representan la relación entre la entrada (excitación glótica) y la salida (voz generada).

Un diagrama esquemático que representa los órganos involucrados en la producción del habla se muestra en la figura 2.18.

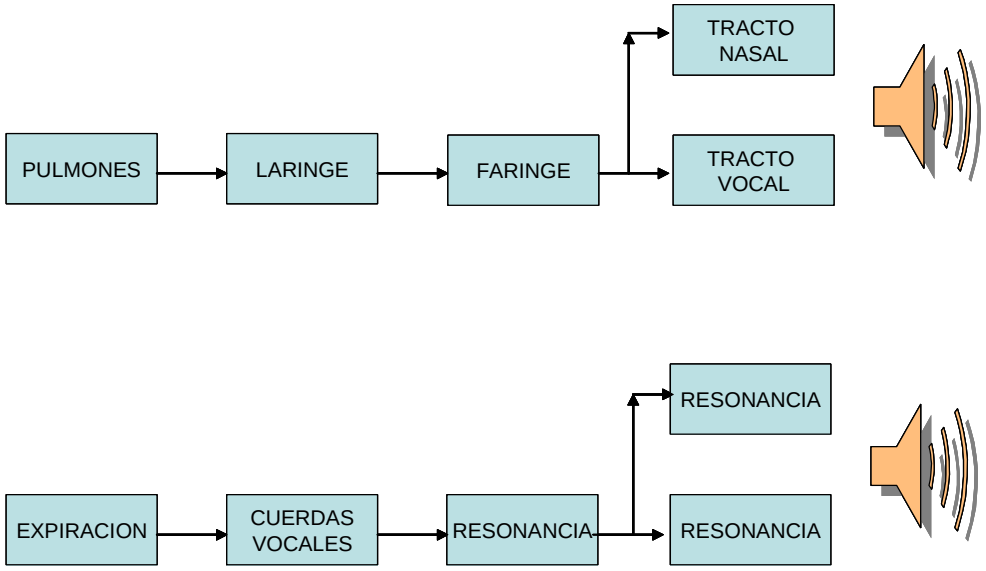


Figura 2.18 Diagrama Funcional de los órganos involucrados en la producción del habla. Tomada de [11].

El modelo de producción del habla se representa en el esquema de la figura 2.19, en el cual las cuerdas vocales son reemplazadas por un generador de impulso, la cavidad resonante por un sistema de filtro lineal, y la excitación de una señal sorda por un generador de ruido. En la práctica todos los sonidos tienen una excitación mixta con porciones entonadas y sin tono, sin embargo, en una primera aproximación es posible utilizar un conmutador simple

que permita seleccionar la excitación entre entonada o sin tono dependiendo del sonido que ha sido generado.

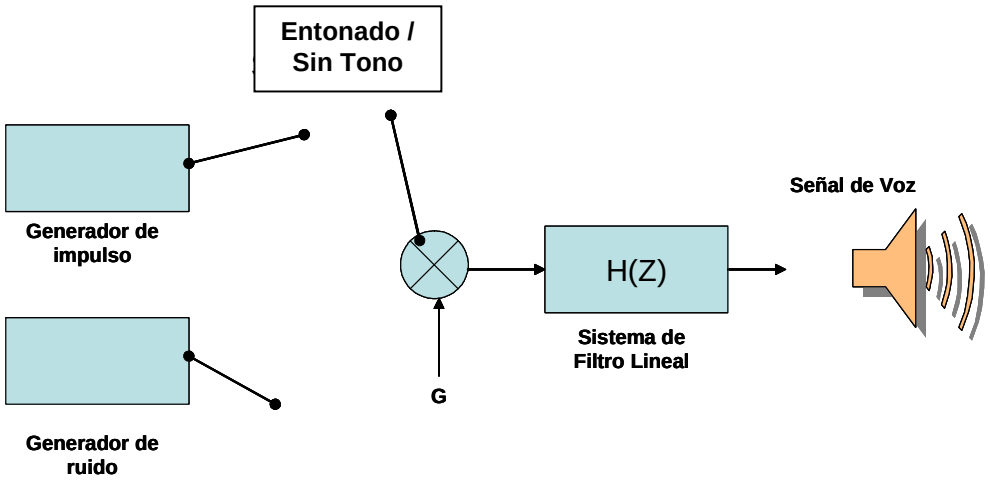


Figura 2.19. Modelo de Producción del Habla. Tomada de [11].

La cavidad resonante es modelada por un filtro lineal, ecuación (7), cuyos coeficientes son calculados para que la respuesta al impulso del filtro corresponda a la envolvente del espectro del fragmento de la señal que está siendo procesada. Es importante que la excitación no distorsione el espectro de la señal original, por lo que se utiliza un generador de ruido blanco Gaussiano para representar los segmentos sin tono de la señal. De este modo, la excitación tiene espectro plano para señal entonada y un tren de impulso para señal sin tono (el espectro de un tren de impulsos en el tiempo es un tren de impulsos en la frecuencia). La señal de salida tendrá entonces la envolvente de la señal procesada, en el caso de una señal sorda, y la

envolvente modulada por el tren de impulsos, en caso de señal sonora.

$$H(z) = \frac{1}{1 - \sum_{i=1}^{p} a_i z^{-i}} \quad \text{----------(7)}$$

Los coeficientes a_i de la ecuación (7) son calculados utilizando la técnica de **Predicción Lineal** (***LPC:*** *Linear Predictive Coding*). De este modo, un *vocoder* elemental calcula los coeficientes ***LPC*** del filtro, calcula el **tono** y si la señal es sonora o sorda. Los parámetros son cuantizados y transmitidos al decodificador, el cual reconstruye la señal de excitación, para finalmente filtrarla y obtener la señal de salida.

2.3.2. Codificador de predicción lineal con señal de excitación del filtro seleccionada según un código, (*CELP: Code Excited Linear Prediction*)

En 1985 se publicó un codificador de voz denominado **Codificador de predicción lineal con señal de excitación del filtro seleccionada según un código** en el cual se utilizó, por primera vez, la técnica de análisis por síntesis. En esta tesis, por simplificación del nombre, el codificador ***CELP*** lo llamaremos Codificador de Predicción Lineal con Excitación **(CPLE)**

La figura 2.20 muestra un esquema básico de este sistema.

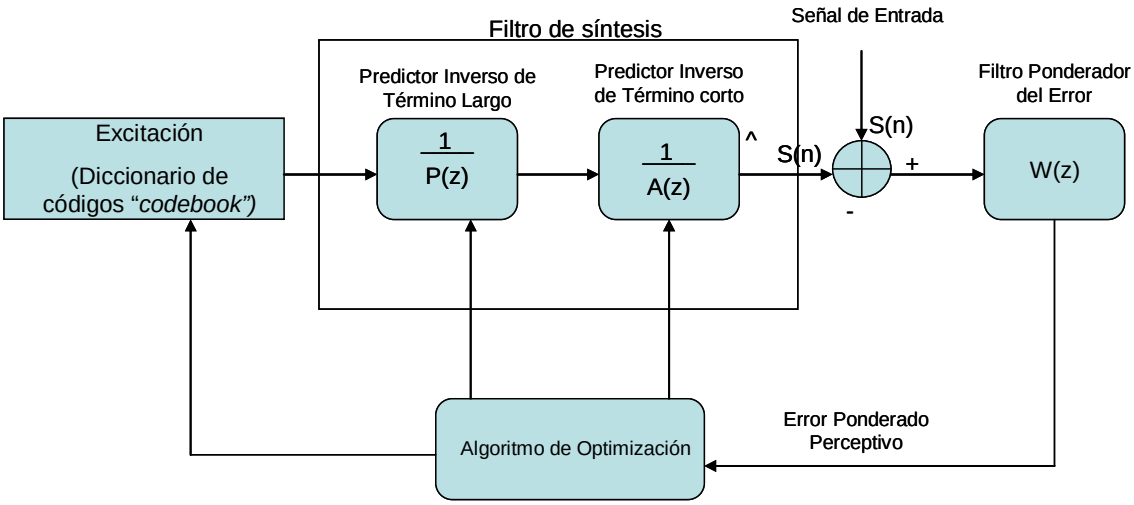


Figura 2.20.- Esquema básico de codificadores CPLE [11].

El principio de este método es que el codificador de voz codifica la señal tomando en cuenta la señal decodificada. Utilizando

los parámetros que posteriormente serán trasmitidos, el codificador reconstruye la señal de modo de minimizar alguna medida de error. Más aun, un filtro ponderador del error es agregado al esquema de análisis por síntesis para explotar la propiedad de enmascaramiento, aprovechando las limitaciones del sistema auditivo humano.

Las tecnologías **CPLE** descomponen la señal de voz, tomando en cuenta características de periodicidad, hasta dejar una señal plana semejante a un ruido aleatorio. Esta señal resultante o residuo es modelado por medio de sistemas de cuantización. El procesamiento es realizado sobre segmentos de voz (*frames*), y no muestra a muestra como en la mayoría de los codificadores de forma de onda. El número de muestras por segmento varía generalmente entre 80 y 280, que muestreadas a 8 kHz corresponden a 10 y 35 milisegundos respectivamente.

El **tono** incorpora una componente casi periódica en la señal de voz. En el dominio del tiempo, esta naturaleza periódica, que normalmente varía entre 2.5 y 20 milisegundos, es llamada correlación de largo plazo (*long term correlation*), la cual puede ser removida utilizando un filtro de predicción. Este filtro, *1/P(z),* corresponde a un filtro con solo polos (finitos) y tiene por función predecir el **tono**. La ecuación (8) entrega la forma general de este filtro, donde corresponde a los coeficientes del predictor, que son

calculados utilizando técnicas de correlación, y *D* representa al período del **tono**.

$$\frac{1}{P(z)} = \frac{1}{1 - z^{-D} \sum_{i=-q}^{r} b_i z^{-i}} \qquad \text{----------(8)}$$

Adicionalmente, la señal de voz presenta correlación muestra a muestra tanto a nivel temporal como espectral. En el dominio del tiempo, este efecto es denominado correlación de corto plazo (short term correlation), el cual es posible reducir pasando la señal de salida del filtro *1/P(z),* por un filtro *1/A(z).* La ecuación (8) muestra la forma general de este filtro, el cual es idéntico a H(z) en la ecuación (7). Los coeficientes son calculados utilizando predictores lineales de orden *p*, el cual usualmente toma el valor de 10 para voces muestreadas a 8 kHz.

$$\frac{1}{A(z)} = \frac{1}{1 - \sum_{k=1}^{p} a_i z^{-k}} \qquad \text{----------(9)}$$

Luego que la señal de voz es procesada por los filtros de predicción de error de término corto y largo, la señal resultante carece de todo tipo de periodicidad. Dado que este residuo presenta características de ruido aleatorio, el codificador **CPLE** lo modela usando segmentos o vectores de ruido blanco Gaussiano. Existe un número grande de estos vectores de representación, los cuales en conjunto forman un **diccionario de códigos** *(codebook)*, el cual está presente tanto en el codificador como en el decodificador. El residuo es comparado usando distancia euclidiana ponderada con cada uno de los vectores que forman el **diccionario de códigos**, para luego seleccionar el más parecido al residuo, o sea el que tenga la distancia mínima. La ventaja es que en vez de enviar al decodificador el vector de residuo, sólo se envía el índice del **diccionario de códigos** correspondiente al segmento seleccionado. Representar una secuencia de vectores en un **diccionario de códigos** por su índice es denominado cuantización vectorial y es parte fundamental en la disminución de la tasa de transmisión de los codificadores tipo *vocoders*. Lo que se manda al decodificador, corresponde al índice de la secuencia del **diccionario de códigos** óptimo, los coeficientes de los predictores de término corto y largo, y las ganancias asociadas a los filtros. Con estos valores se reconstruye la señal de excitación para finalmente filtrarla por el filtro *1/A(z).* La figura 2.21 muestra el diagrama del decodificador.

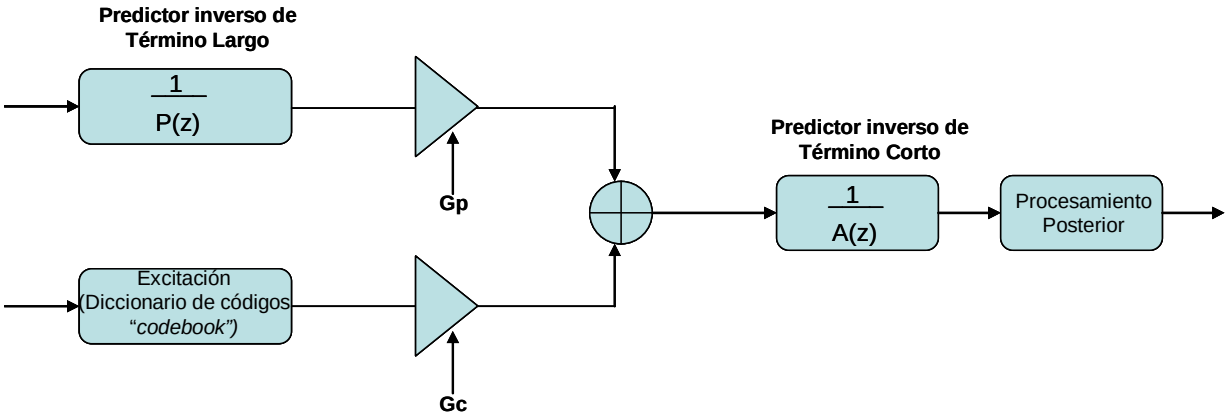


Fig. 2.21.- Esquema básico de un decodifcador CPLE. [11]

2.3.3. El Codificador CPLE *FS-1016* 4.8 kbps.

El codificador **Estándar Federal 1016** (***FS-1016*** *Federal Standard 1016*) es un estándar adoptado por el Departamento de Defensa de los EE.UU y fue desarrollado en conjunto con los laboratorios BELL AT&T. Utiliza una estructura **CPLE** con un **diccionario de códigos** fijo y otro adaptable para producir la excitación filtro todo polo. Opera con segmentos de 30 milisegundos y cada segmento es dividido en 4 subsegmentos de 7.5 milisegundos. Los coeficientes del filtro de orden 10 todo polo son calculados a partir de cada segmento utilizando predicción lineal y cuantizados con 34 bits. La excitación para este filtro es por cada segmento un **diccionario de códigos** fijo y un **diccionario de códigos** adaptable, ambos **diccionarios de códigos** son explorados por el codificador en lazo cerrado para minimizar el error entre la señal original y la señal reconstruida, y las ganancias son cuantizadas con 5

bits. En los subsegmentos impares el índice del **diccionario de códigos** adaptable es codificado con 8 bits, y en los subsegmentos pares este retardo es codificado con 6 bits. Un bit por segmento es utilizado para sincronización, 4 bits por segmento son utilizados para la corrección de errores. Finalmente, 1 bit por segmento es empleado para futuras expansiones del codificador. Los bits resultado de la codificación y que son transmitidos a través del canal de comunicación, se presentan en la figura 2.22.

En el decodificador, los bits son recibidos para proporcionar los coeficientes al filtro de síntesis, los **diccionarios de códigos** fijo y adaptable para excitar este filtro, y posteriormente producir la señal de voz reconstruida.

PARÁMETRO	BITS POR CADA SUBSEGMENTO	BITS POR CADA SEGMENTO
Coeficientes LPC	-	34
Retardo del «**diccionario de códigos**» adaptable.	8 o 6	28
Índice del «**diccionario de códigos**» fijo	9	36
Ganancia del «**diccionario de códigos**» adaptable.	5	20
Ganancia del «**diccionario de códigos**» fijo.	5	20
Corrección de error	-	4
Sincronización.	-	1
Bit de expansión.	-	1
TOTAL	**-**	**144**

Figura 2.22.- Colocación de los bits en el codificador *FS-1016*.

Capítulo

3

DESARROLLO.

3.1. DESARROLLO DEL SISTEMA DE CIFRADO.

3.1.1. Consideraciones generales.

En el presente trabajo, se describe el uso del cifrado con clave simétrica para proporcionar seguridad a la voz que se transmite a través de la **RTB**.

Para aplicar el cifrado en la voz para su transmisión por la **RTB**, se considera lo siguiente:

- Convertir la voz en una forma digital antes de cifrarla.
- Para la transmisión de datos a través del lazo local, se requiere de módems telefónicos. Estos transmiten a una velocidad máxima de 33.6 kbps.
- La voz cuantizada con **CIM** alcanza tasas de 64 kbps, que es mayor a la velocidad máxima con que se comunican

los módems. Si se requiere transmitir voz en tiempo real a través de la **RTB**, es necesario comprimir la voz para evitar pérdida de muestras por sobrecarga de las **regiones de memoria o contenedores para espera de salida** (*output buffers)*.

3.1.2. Propuesta de desarrollo del sistema de cifrado.

La propuesta consiste en implantar un sistema en los aparatos telefónicos que digitalice, comprima, cifre y transmita voz en tiempo real, utilizando la infraestructura de la **RTB**, como se muestra en la figura 3.1.

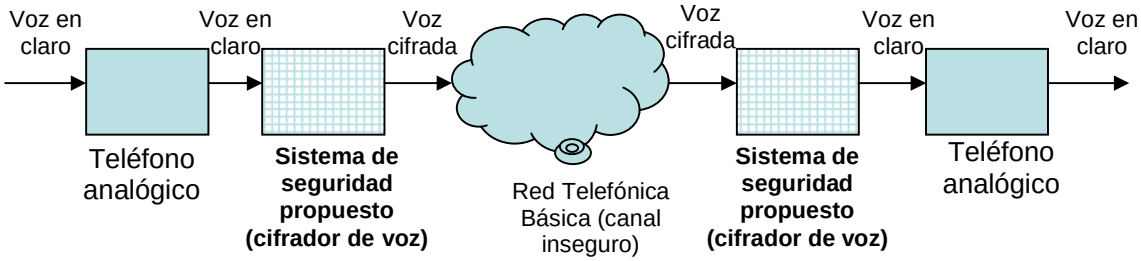


Figura 3.1.- Esquema general del sistema de cifrado de voz propuesto.

El sistema tiene como entrada la voz analógica proveniente del teléfono emisor, realiza la conversión a una forma digital y la comprime hasta una tasa de 4.8 kbps. Después, realiza el cifrado de la información a través del algoritmo de cifrado simétrico con una clave de cifrado de 128 bits. Una vez

cifrada la información, se transmite por la **RTB** con un módem telefónico previo conectado al lazo local. Los datos modulados que viajan a través de la **RTB** están protegidos por el algoritmo de cifrado, y sólo el teléfono receptor que cuente con el mismo sistema y con la clave de cifrado correcta, podrá realizar, además de la demodulación y descompresión, el descifrado correcto de la información. El sistema de cifrado que se propone se puede observar en la figura 3.2.

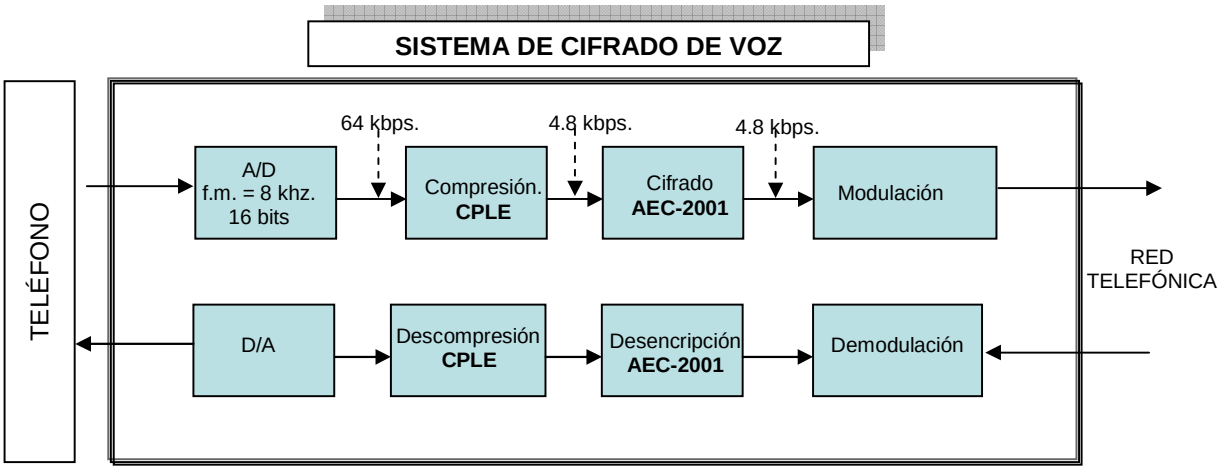


Figura 3.2.- Diagrama a bloques del sistema de cifrado incrustado en las terminales telefónicas.

Las etapas más importantes que se necesitan en el sistema propuesto son:

- **Conversión analógica-digital y digital analógica**, con una frecuencia de muestreo de 8 Khz, con 16 bits por muestra y una velocidad de 64 kbps.
- **Compresión-descompresión** de la voz a una velocidad de 4.8 kbps
- **Cifrado-descifrado** por bloques con clave simétrica conservando la velocidad de 4.8 kbps.

- **Modulación-Demodulación** digital para la transmisión de datos, con velocidades de transmisión de 9.8 kbps a 33.6 kbps.

3.2. HARDWARE DEL SISTEMA DE CIFRADO.

Para la conversión de la voz de una forma analógica a digital, se emplea un **circuito cuantizador σ-δ,** del cual se usa aquí uno de los campos de audio de 16 bits en formato complemento dos, bit más significativo primero; el cuantizador se configuró, mediante un banco de interruptores, para trabajar con una frecuencia de muestreo de 8000 Hz. En la compresión se utiliza el estándar de compresión **CPLE,** y en la etapa de cifrado se implanta el algoritmo **AEC-2001**. Para la transmisión de la voz cifrada se requiere el uso de un módem con velocidad máxima de transmisión de datos de 33.6 kbps. El desarrollo de la propuesta se basa en el esquema mostrado en la figura 3.3 y consta de los siguientes componentes:

- Dos tarjetas de desarrollo *Freescale DSP56858EVM* (1 para cada teléfono) que incluye un **PSD**, convertidores analógico-digital/digital-analógico, **puerto o Interfaz de comunicación serial** (***SCI***: *Serial* Communications *Interface*) y puerto o **Interfaz Serial Síncrona Mejorada (***ESSI: Enhanced Synchronous Serial Interface***)**

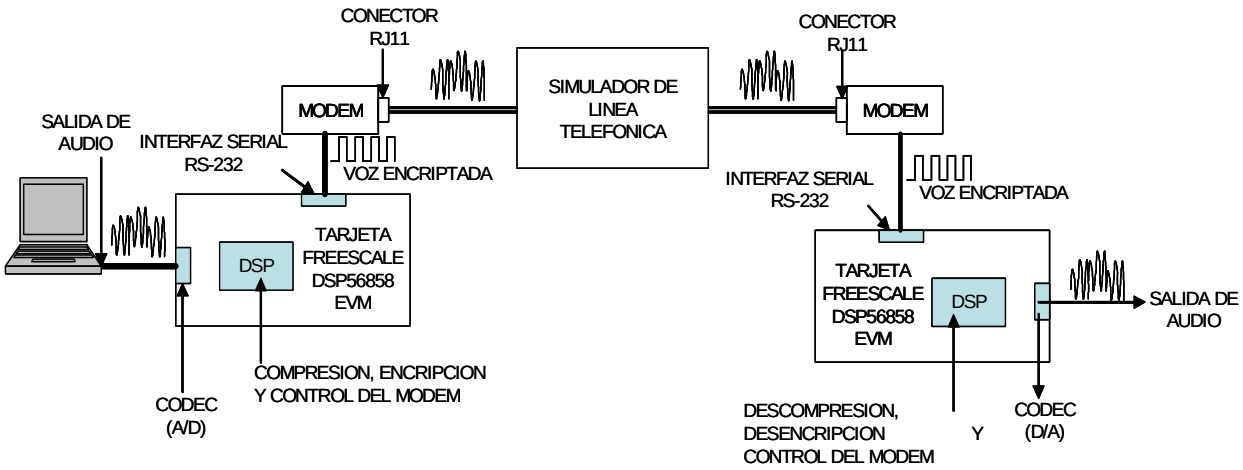


Figura 3.3. Esquema de las diferentes partes empleadas en el desarrollo del sistema de cifrado de voz.

- Dos módems telefónicos externos (1 para cada tarjeta) con interfaz RS-232.
- Un equipo simulador de línea[1] (propiedad de la **Secretaría de la Defensa Nacional**), con la finalidad de sustituir una línea telefónica de la **RTB** y facilitar el desarrollo de pruebas.

En la tarjeta de desarrollo se implantan las etapas de compresión y cifrado, incluyendo los procesos de control necesarios para el funcionamiento del módem externo en la transmisión de los datos.

La voz analógica se digitaliza con el circuito ***CS4218*** incluido en la tarjeta, obteniéndose una velocidad de transmisión de 64 kbps (un solo canal

[1] El simulador de línea empleado es el *BK PRECISION 1050,* permite proporcionar las características de una línea telefónica tales como la tensión en la línea, tono de llamada, timbrado, impedancia entre otras. El empleo de este equipo facilitó de manera importante realizar las pruebas, incluso poder experimentar en distintos lugares debido a su portabilidad, lo cual seria imposible con las líneas telefónicas reales debido a que estas son fijas.

de audio). Posteriormente, la voz se comprime en el **PSD** con el algoritmo **CPLE** hasta una velocidad de transmisión de 4.8 kbps. Una vez comprimida la voz, se cifra en el mismo **PSD** mediante el algoritmo de cifrado **AEC-2001** con una clave de 128 bits. La voz cifrada se envía al módem telefónico a través del puerto serial de acuerdo al estándar de comunicación serial RS-232.

Con el módem telefónico se transmite la voz comprimida y cifrada a través de una línea telefónica proporcionada por el equipo simulador de línea telefónica. Los módems emplean una **modulación por desplazamiento de fase** (***QPSK:*** *Quadrature Phase-Shift Keying*), un código detector y corrector de errores denominado **protocolo de acceso al enlace para módem** (***LAP-M:*** *Link Access Protocol - Modem*) y se conectan a una velocidad de transmisión que varía de 9.8 kbps a 33.6 kbps, dependiendo de las condiciones de la línea telefónica. Para el control del módem, se utilizan los comandos AT (también llamados Hayes), enviados por el **PSD** a través del puerto ***ESSI*.**

En el lado del receptor, un módem recibe la voz cifrada para su demodulación y es transmitida a la tarjeta de desarrollo para su descifrado, descompresión y por último para su conversión digital- analógica.

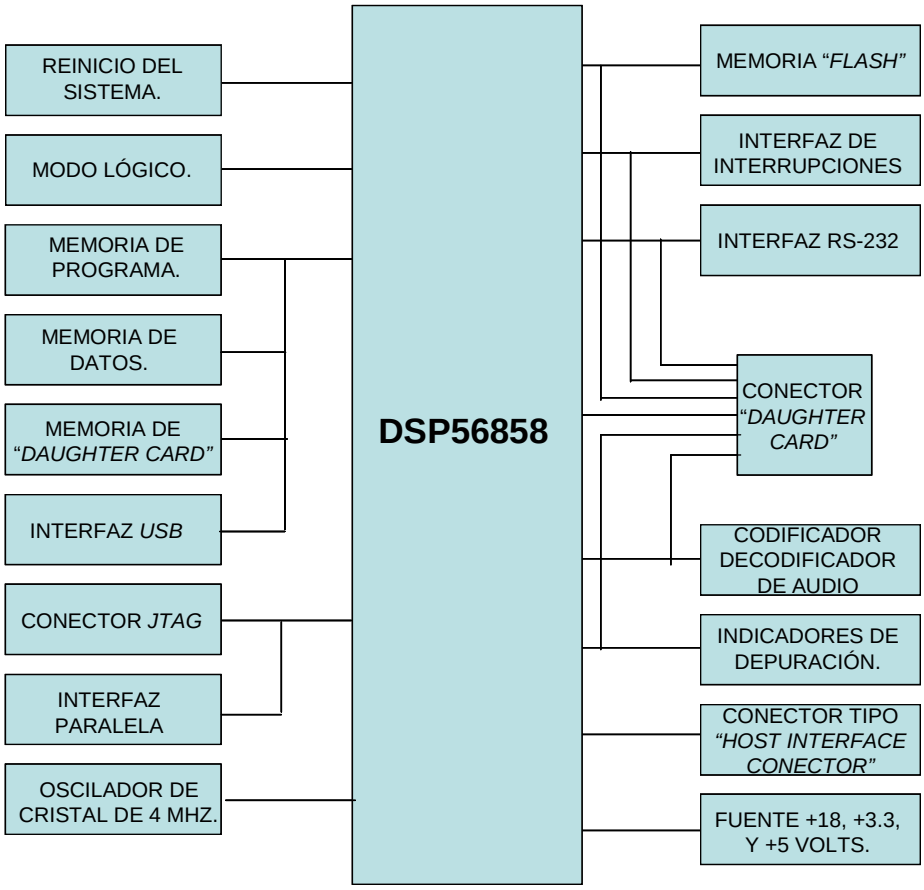
3.3. DESCRIPCIÓN DE LA TARJETA DE DESARROLLO *DSP56858EVM.*

La tarjeta *DSP56858EVM* es un módulo de evaluación para el **PSD** de *Freescale DSP56858* y proporciona el hardware necesario para el desarrollo de aplicaciones sobre el *DSP56858* (figura 3.4). Esta tarjeta tiene las siguientes características generales:

- Procesador de Señales Digitales *DSP56858* de 16 bits, operando con voltajes +1.8V/+3.3 volts a una velocidad de 120MHz.
- Memoria RAM externa y Memoria EEPROM para datos de 1M-bit.
- Oscilador de cristal a 4.00MHz.
- Interfaz paralela.
- Puerto serial RS-232.
- Codificador-decodificador de voz de 16 bits.

3.3.1. *DSP56858* de 16 bits.

En este **PSD** se implantaron los algoritmos de compresión **CPLE** y de cifrado **AEC-2001**. Debido al adecuado diseño del módulo de evaluación del **PSD** empleado, fue relativamente simple realizar el control de módem telefónico mediante los comandos ***AT***.



Memoria *"Daughter Card"*	Memoria para ranura con el estándar de tarjetas auxiliares ("hijas").
Memoria *"Flash".*	Memoria rápida no volátil de borrado en forma eléctrica.
Connector Daughter Card.	Conector de tipo ranura con el estándar de tarjetas auxiliares ("hijas").
Host Interface Connector	Conector de tipo interfaz paralela para procesadores.
***USB** (Universal Bus Serial)*	bus universal tipo serie
***JTAG** (Joint Test Action*	Interfaz de depuración.

Figura 3.4. Diagrama de bloques de la tarjeta *DSP56858EVM.* Tomada de [12].

El *DSP56858* (figura 3.4), es un procesador híbrido que combina la eficiencia de un procesador digital de señales y la

funcionalidad de un microcontrolador con una variedad de periféricos integrados dentro de la pastilla.

Algunas de sus características son:

- 120 millones de instrucciones por segundo a 120MHz.
- Memoria de programa SRAM de 40K x 16 bits, y memoria de datos SRAM de 24K x 16 bits.
- 6 canales independientes de **acceso directo a memoria** (***DMA***: *Direct Memory Access*).
- 2 Interfaces Seriales Síncronas. (***ESSI***).
- 2 Interfaces de Comunicación Serial (***SCI***).
- Interfaz Puerto Serial.
- Interfaz de 8 bits en paralelo.
- Hasta 47 puertos de **Propósito General de Entrada y/o Salida. (*GPIO**: General Purpose Input Output*).**

El *DSP56858* está fabricado con tecnología ***CMOS*** con «tolerancia de conexión a niveles de 5 volts». El término «tolerancia de conexión a niveles de 5 volts» se refiere a la capacidad de un pin de entrada/salida, que diseñado para ser compatible con tecnologías de 3.3 volts, puede soportar un voltaje de hasta 5.5 volts sin dañar el dispositivo. Esto permite la integración de sistemas de 3.3 y 5 volts [13].

3.3.2. Interfaces paralela y de comunicación serial (*SCI: Serial Comunications Interface)* del *DSP56858*.

La interfaz paralela permite al **PSD** con una computadora a través del puerto de impresora. Con este conector se pueden cargar programas y trabajar con los registros del **PSD**.

La interfaz ***SCI*** del **PSD** permite comunicaciones seriales y asíncronas con dispositivos periféricos y otros **PSD**s. Con esta interfaz se implanta una comunicación serial asíncrona a través de un conector RS-232, utilizado para conectar la tarjeta *DSP56858EVM* a un módem telefónico externo con interfaz de comunicación serial. Algunas de las características importantes de este periférico ***SCI*** son:

- Transmisión bidireccional simultánea o simple.
- Formato: Sin retorno a cero (***NRZ****:No return Zero*).
- Selección de velocidad de baudios de 13 bit.
- Formato de datos programable de 8- o 9 bits.
- Habilitación en forma separada del transmisor y receptor.
- Petición separada de interrupciones a la unidad central de procesamiento (***CPU****: Central Processing Unit*) del receptor y transmisor.
- Detección de error de trama en el receptor.

La lectura, escritura y programación de este puerto, es por medio de 5 registros de intercomunicación, mapeados a la memoria del **PSD**.

- Registro de velocidad en baudios (*Baud Rate Register*).- en este registro se configura la velocidad de transmisión.
- Registro de control (*Control Register*).- empleado para el control de la comunicación.
- Registro de control en modo ***DMA*** (*Control Register DMA*).- lo mismo que el anterior, pero aplicado a ***DMA***.
- Registro de estado (*Status Register*).- permite monitorear el estado de la comunicación.
- Registro de datos (*Data Register*).- en este registro se leen los datos recibidos o se colocan los datos a transmitir.

La interfaz ***SCI*** permite una comunicación bidireccional simultánea, asíncrona, sin retorno a cero, entre el **PSD** y un dispositivo remoto. El transmisor y receptor ***SCI*** opera en forma independiente. El **PSD** monitorea el estado del ***SCI***, escribe los datos a ser transmitidos y procesa los datos recibidos.

3.3.3. Puerto RS-232 para comunicación serial de la tarjeta *DSP56858EVM*.

La tarjeta *DSP56858EVM* proporciona un puerto RS-232 (figura 3.5), utilizando un convertidor de nivel *MAX3245EEAI*. Este convertidor de nivel realiza la transición de los niveles proporcionados por la interfaz ***SCI*** del **PSD** a niveles compatibles con el estándar RS-232. Es importante señalar que en esta interfaz no están incluidas las demás señales de control del estándar, por lo que para el desarrollo de esta tesis, el control se implanta utilizando señales de los puertos de propósito general de entrada y salida.

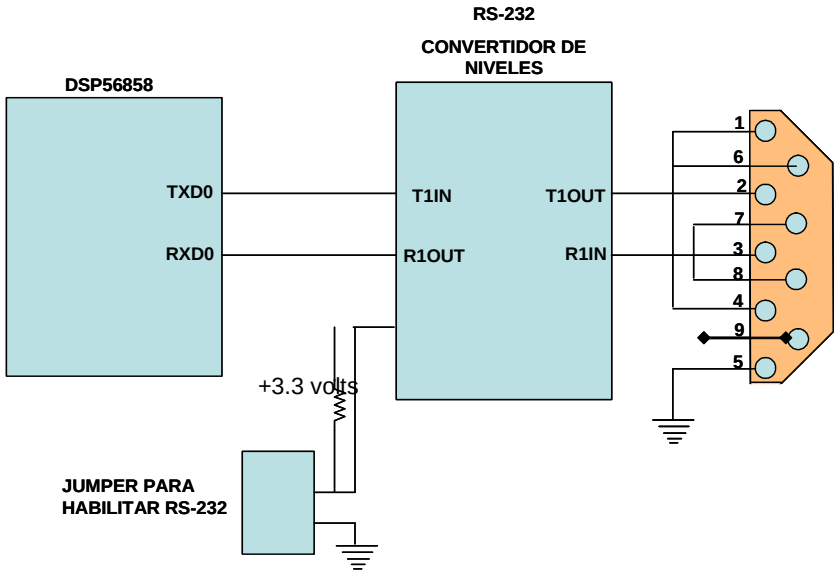


Figura 3.5. Diagrama Esquemático de la interfaz RS-232. Tomada de [12].

La interfaz paralela permite al **PSD** comunicarse con una computadora a través del puerto de impresora. Con este conector se pueden cargar programas y trabajar con los registros del **PSD**.

3.3.4. El codificador-decodificador de voz *CS4218*.

El codificador-decodificador de voz de 16-bit de ***Crystal Semiconductor*** **CS4218**, consta de convertidores A/D y D/A de audio y voz. Está conectado al **PSD** a través de la interfaz ***ESSI*** del ***DSP56858*** y opera a frecuencias de muestreo entre 8KHz y 48KHz.

La interfaz ***ESSI*** transfiere datos de audio digital del codificador-decodificador de voz hacia el **PSD**, así como el audio y voz procesados del **PSD** al codificador-decodificador de voz. La interfaz ***ESSI*** consiste de secciones de transmisión y recepción independientes y establecen una comunicación serial síncrona con el codificador-decodificador.

3.3.5. Puertos de Propósito General de Entrada y/o Salida (*GPIO: General Purpose Input Output*).

Los pines de propósito general están diseñados para compartirse con otros módulos periféricos del **PSD** (Interfaz ***ESSI***, interfaz ***SCI***, interfaz paralela). Es decir, los pines de las diferentes interfaces del **PSD** pueden funcionar en forma independiente también como puertos ***GPIO***.

El módulo ***GPIO*** tiene las siguientes características.

- Control individual de cada pin ya sea en modo normal (interfaz) o modo ***GPIO***.
- Control individual de la dirección cada pin ya sea en modo normal o modo ***GPIO***.

3.4. EL MÓDEM MULTITECH MT5600.

Una vez que la voz digital se comprime y cifra en el **PSD**, se transmite a la **RTB** con un módem telefónico externo. El módem utilizado es el MultiMódem II de Multitech y proporciona comunicaciones V.92 a 56K para datos y 14.4K-bps para fax. Algunas de sus características son:

- Velocidades de transferencia de descarga de V.92/56K y de subida de datos de 48K.
- Transmisión de Fax Clase 1 y Clase 2 a 14.4K
- Configuración por medio de los comandos AT.
- Interfaz serial RS-232.
- Conectores: DB-25 a DB-9
- 115 V/240 V CA, 50/60 Hz.

3.4.1. Conexión del módem a la tarjeta *DSP56858EVM*.

El módem se comunica con la tarjeta *DSP56858EVM* a través del puerto serial con el estándar RS-232 (comunicación asíncrona). El estándar RS-232 consiste en un conector DB-9 de 9 pines y se manejan niveles de voltajes digitales de +12V (0 lógico) y -12V (1 lógico), para la entrada y salida de datos. Cada pin puede ser de entrada o de salida, teniendo cada uno de ellos una función específica.

Este estándar se usa comúnmente para comunicación de datos entre computadoras utilizando módems externos. Para esta tesis, el estándar se utiliza para transmisión de datos entre el **PSD** y el módem externo. Existen señales de entrada y salida, tanto para la transferencia de datos como para el control de flujo de datos y control del módem.

El conector DB-9 consta de 9 pines:

- ***TD*** *(Transmited Data)*.- Transmisión de datos.
- ***RD*** *(Recive Data).*- Recepción de datos.
- ***RTS*** *(Request To Send).*- Petición de envío.
- ***DTR*** *(Data Terminal Ready)* .- Terminal de datos preparado.
- ***CTS*** *(Clear To Send).*- Preparado para transmitir.

- ***DSR*** *(Data Set Ready)****.-*** Módem preparado.
- ***CD*** *(Carrier Detec)****.-*** Detección de portadora.
- ***RI*** *(Ring Indicator).-* Indicador de llamada.

DTR *Y* ***RTS*** son señales enviadas por la tarjeta al módem para indicarle que esta lista para transmitir y recibir los datos. ***DSR*** *y* ***CTS*** son señales que envía el módem a la tarjeta para indicarle que esta listo para transmitir y recibir datos. ***DCD*** es una señal que indica que los módems ya han establecido conexión y que hay una portadora. ***RI*** se activa cuando el módem recibe el tono de llamada. ***RXD*** es la línea por donde el puerto de la tarjeta recibe los datos en forma serial. Y ***TXD*** es la línea por donde el puerto serie de la tarjeta transmite los datos hacia el módem.

En la tarjeta de desarrollo, la comunicación serial se implanta con la interfaz ***SCI***, un convertidor de niveles *MAX3245EEAI* y un conector DB-9 hembra. Sin embargo, únicamente están activadas las líneas ***RXD*** *y* ***TXD***. Las demás líneas del estándar no están consideradas. Para que exista transferencia de datos entre la tarjeta y el módem, es necesario que las demás líneas estén activadas, para lo cual se implantan las líneas de control del estándar RS-232 utilizando otros puertos libres de la tarjeta. Las conexiones se muestran en la figura 3.4.

Las señales ***DTR, DSR, CTS, RTS RI*** **y** ***CD*** se implantan con pines del puerto paralelo**,** los cuales se configuran como puertos ***GPIO*** de entrada o salida, según corresponda. Es importante que entre los pines ***GPIO*** y los pines del conector DB-9 exista un convertidor de niveles, ya que los voltajes que maneja el puerto ***GPIO*** son del orden de 3.3 volts, y los voltajes del módem son de 5 volts. El circuito integrado *MAX3245EEAI* permite convertir niveles de 5 a 12 volts a niveles de 3.3 volts (figura 3.6).

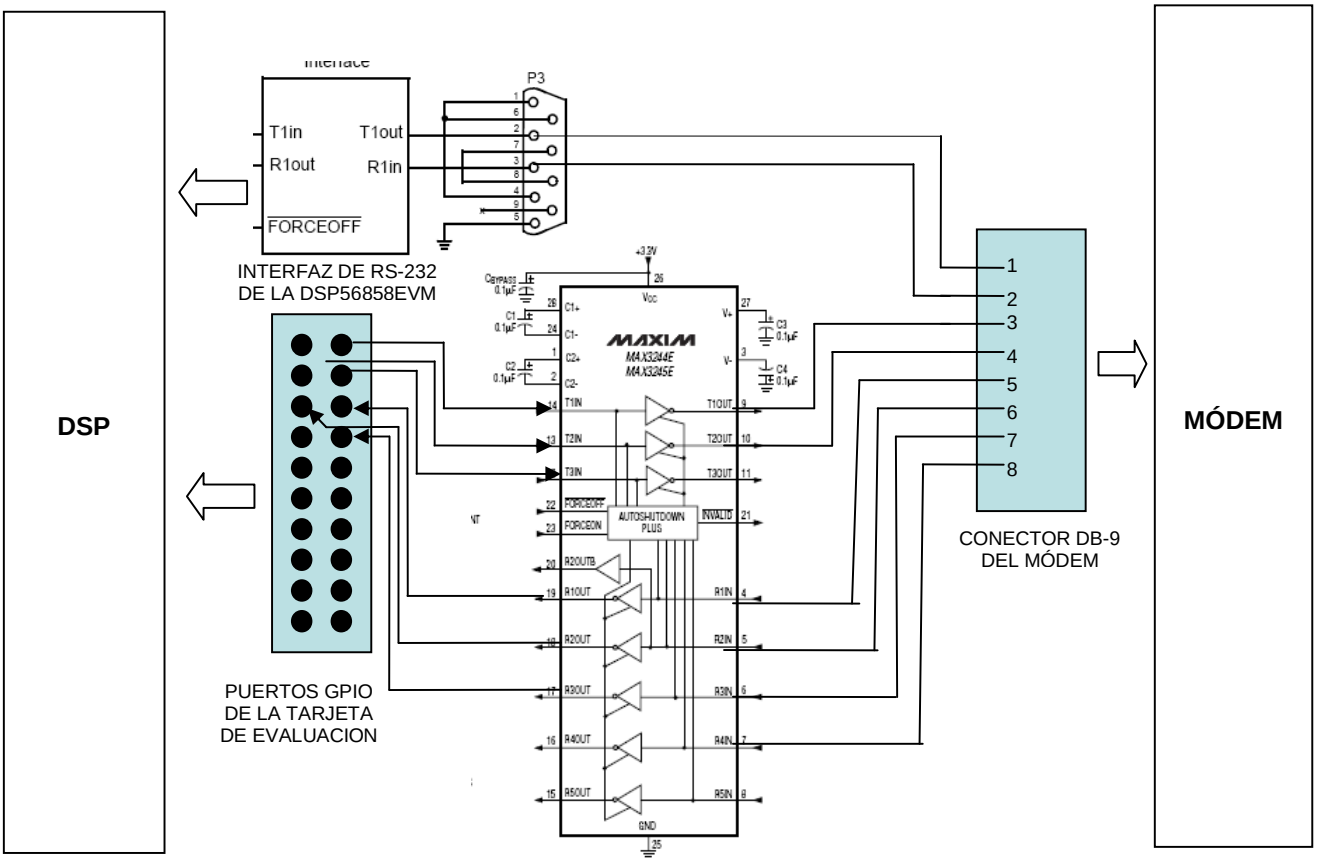


Figura 3.6. Diagrama esquemático de conexión entre la tarjeta DSP56858EVM y el MÓDEM. Tomada de [14].

3.5. EL CODIFICADOR-DECODIFICADOR *CS4218* Y SU CONEXIÓN AL DSP56858.

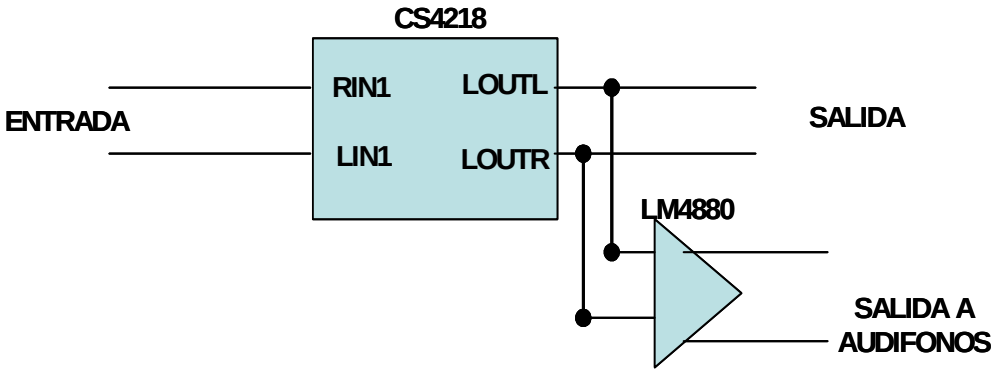
3.5.1. Características.

El codificador-decodificador de voz instalado en la tarjeta es el circuito integrado ***CS4218*** de 16 bits que consiste de 2 convertidores A/D delta-sigma, 2 convertidores D/A delta-sigma, filtros de entrada anti-aliasing, filtros de salida de banda de paso suave, ganancia programable a la entrada y atenuadores programables a la salida. Soporta niveles de +3.3V y permite enviar datos al **PSD**, procesarlos, y transmitirlos inclusive hacia el mismo codificador-decodificador de voz, utilizando para ello el puerto serial síncrono ***ESSI***. Emplea un oscilador de 12.2888 MHZ que le permite operar con frecuencias de muestreo que van de 8KHz a 48KHz y que se configuran manualmente a través de un conmutador de posiciones también integrado en la tarjeta de desarrollo.

En la interfaz analógica del codificador-decodificador de voz, se utilizan conectores para entradas y salidas de audio analógico del codificador-decodificador de voz, además de una salida para audífonos a través de un amplificador operacional *LM4880* como se muestra en la figura 3.7.

En la entrada se introduce la voz analógica, la cual en el desarrollo del presente trabajo, proviene de la salida de audio de una computadora, y la salida de audio del codificador-decodificador de voz se envía a la entrada de las bocinas.

La interfaz analógica del codificador-decodificador de voz está conectado al *DSP56858* a través del puerto ***ESSI***.



RIN1	Entrada de Audio Canal Derecho.
LIN1	Entrada de Audio Canal Izquierdo.
LOUTL	Salida de Audio Canal Izquierdo.
LOUTR	Salida de Audio Canal Derecho.

Figura 3.7 Conexiones del codificador CS4218 (señales analógicas) Tomada de [12].

3.5.2. La interfaz Serial Síncrona Mejorada *(ESSI: Enhanced Synchronous Serial Interface)* del *DSP56858*.

El puerto ***ESSI*** del **PSD** se utiliza para la transferencia de datos (voz digitalizada) con el codificador-decodificador de voz. Consiste de 6 pines, los cuales funcionan también como puertos de entrada y

salida de propósito general, según su configuración. Este puerto es para comunicaciones síncronas cuando se programa como interfaz ***ESSI*** y la función de cada pin se muestra en la Figura 3.8.

NOMBRE DEL PIN	FUNCIÓN DEL PIN
Control en serie 0 Serial Control 0 (SC0/PC0)	Puede ser usado como transmisor o receptor de la señal de reloj, o transmisor de datos.
Control en serie 0 Serial Control 1 (SC0/PC0)	Tiene diferentes funciones según su configuración, ya sea como una bandera de control, receptor o transmisor de la señal de sincronía de trama.
Señal de Reloj en seria Serial Clock (SCK/PC4)	Este pin transmite o recibe la señal de reloj.
Recepción de datos en serie Serial Receive Data (SRD/PC4)	Recepción de datos en forma serial
Transmisión de datos en serie Serial Transmit Data (STD/PC5)	Transmisión de datos en forma serial

Figura. 3.8 Señales de la interfaz ESSI. Tomada de [15]

Cuando estos pines no se utilizan en modo ***ESSI***, funcionan como pines del puerto de propósito general (***GPIO***) Puerto C.

El puerto C del **PSD** puede funcionar como puerto ***GPIO*** o en modo ***ESSI*** configurando los registros de control asociados a este puerto. Cuando trabaja en modo ***ESSI*** se emplean 12 registros de control para determinar su funcionamiento (Figura 3.9).

NOMBRE DEL REGISTRO	FUNCIÓN
Control Register A (CRA)	Controlan los modos de operación de la interfaz ESSI
Control Register B (CRB)	
Status Register (SSISR)	Proporciona el estado de la interfaz ESSI
Transmit Slot Mask Register A (TSMA)	Cuando en la interfaz ESSI se conectan en red varios codec, estos registros determinan el tamaño de datos que a cada codificador-decodificador transmitirán y en que instante.
Transmit Slot Mask Register B (TSMB)	
Receive Slot Mask Register A (RSMA)	Cuando en la interfaz ESSI se conectan en red varios codec, estos registros determinan el tamaño de datos que cada codificador-decodificador transmite al **PSD** y en que instante se reciben.
Receive Slot Mask Register B (RSMB)	
Time Slot Register (TSR)	Este registro se utiliza en combinación con TSMA, TSMB, RSMA, RSMB para determinar el tamaño de datos de cada trama.
Receive Data Register (RX)	Registro de recepción de datos.
Transmit Data Register 0 (TX0)	Registros de transmisión de datos del módulo transmisor de la interfaz ESSI
Transmit Data Register 1 (TX1)	
Transmit Data Register 2 (TX2)	

Figura 3.9 Registros de la Interfaz ***ESSI.*** Tomada de [15].

3.5.3. Conexión del codificador-decodificador de voz al puerto *ESSI* del *DSP56858*.

En la figura 3.10 se muestra la conexión de codificador-decodificador de voz a través del puerto ***ESSI*** del DSP56858.

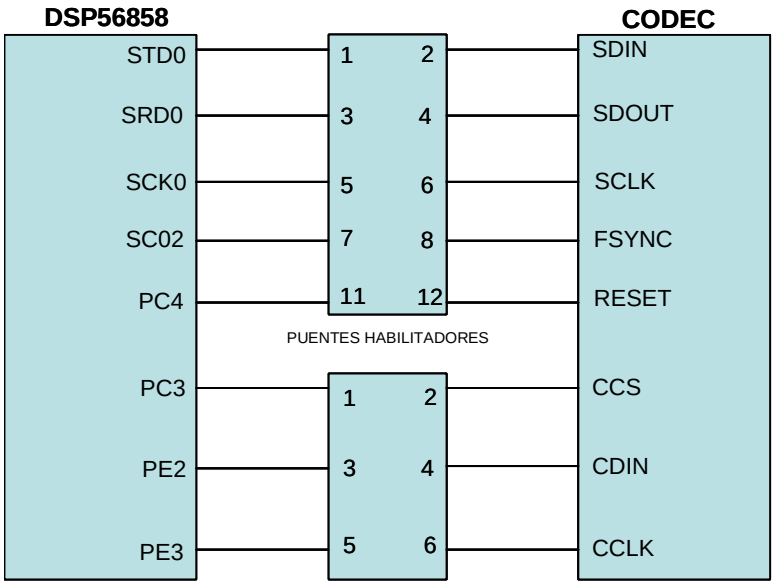


Figura 3.10.- Conexión del codificador-decodificador CS4218 al DSP56858 a través de la interfaz **ESSI**. *Tomada de* [12].

Los pines *STD0* y *SRD0* son entradas y salidas de datos (voz digital) del **PSD** (desde/hacia el codificador-decodificador de voz). El pin *SCK0* recibe la señal de reloj del codificador-decodificador de voz (pin *SCLK* del codificador-decodificador) para la interfaz ***ESSI***. El pin *SC02* del la interfaz ***ESSI*** del **PSD** recibe del codificador-decodificador de voz la señal de sincronía de trama con el fin de sincronizar las tramas de datos. El pin *PC4* del puerto C ***GPIO*** se utiliza para señal de reinicio (*RESET*) al codificador-decodificador de voz. *PC3, PE2 y PE3* del **PSD** se utilizan para fijar ciertos parámetros de la voz digital como atenuación y ganancia.

Para la transmisión y recepción de datos, la interfaz digital del codificador-decodificador de voz tiene cuatros modos de configuración, dependiendo del dispositivo al cual se conecta y el formato de datos. El codificador-decodificador de voz en la tarjeta de desarrollo está integrado utilizando el modo *SM4* el cual corresponde a las conexiones de los pines del codificador-decodificador de voz que se muestran en la figura 3.10. Asimismo, en este modo, el codificador-decodificador de voz esta configurado en modo maestro y transmite y recibe tramas de audio de 32 bits, envía la señal de reloj y de trama de sincronización para el inicio y paro de una trama de datos.

En el modo *SM4*, el codificador-decodificador de voz separa los datos de información (voz) de los datos de control. Los datos de control son parámetros que se envían al codificador-decodificador de voz para fijar parámetros de la voz como la ganancia, y atenuación, y estos datos, a diferencia de otros modos de operación del codificador-decodificador de voz, se envían a través de pines distintos a los utilizados por los datos de información. En la tarjeta de desarrollo *DSP56858EVM* el codificador-decodificador de voz está conectado al **PSD** de forma tal, que la información de control al codificador-decodificador de voz permanece constante y se envía a través de los pines de los puertos de propósito general*, PORT E, y PORT C*.

En la figura 3.11 se muestra la conexión entre pines del codificador-decodificador de voz y la interfaz ***ESSI*** (no se incluyen los pines de la información de control):

Pines de la interfaz ESSI0	**Pines del codificador-decodificador CS4218**	**Descripción.**
STD0	SDIN	Transferencia de datos del puerto ESSI al codificador-decodificador.
SRD0	SCOUT	Transferencia de datos del codificador-decodificador de voz al puerto ESSI.
SCK0	SCLK	Señal de reloj enviado por el codificador-decodificador de voz.
SC00	RESET	Señal de reset al codificador-decodificador de voz enviado por la interfaz ESSI
SC02	SSYNC	Señal de sincronía de trama enviado por el codificador-decodificador de voz.
SC10	CCS	Señal de control de información. Esta señal se mantiene constante (1 lógico).

Figura 3.11 Conexión y descripción de los pines de la interfaz ***ESSI*** del *DSP56858* del codificador-decodificador *CS4218*. *Tomada de* [15]

Los puertos *C y E* se configuran como de propósito general, los registros asociados a estos puertos que determinan su funcionamiento son:

Registros asociados al puerto GPIO y puertos C y E.	**Función.**
Registro de Dirección Puerto C (PRRC)	Controla la dirección del flujo de datos para los puertos ESSI en modo GPIO.
Registro de Dirección Puerto E (PRRE)	
Registro de Datos Puerto C (PDRC)	Almacena los datos recibidos o transmitidos para los puertos ESSI en modo GPIO.
Registro de Datos Puerto E (PDRE)	

3.6. PROGRAMACIÓN PARA LA COMUNICACIÓN DEL *DSP56858* CON EL CODIFICADOR-DECODIFICADOR DE VOZ Y EL MÓDEM.

3.6.1. Comunicación de datos entre el codificador-decodificador de voz y el PSD.

Anteriormente, se explicó la conexión física del codificador-decodificador de voz al **PSD**. Ahora, se describe la programación del **PSD** y codificador-decodificador de voz para la transmisión de datos.

Lo primero que se realiza es una configuración inicial en el **PSD** que consiste en:

- Configuración del módulo del reloj a 120 MHZ.
- Configuración de los puertos ***ESSI*** y ***GPIO***.
- Configuración para transmisión y recepción de datos en modo ***DMA***.

Configuración *del módulo del reloj a 120 MHZ.-* El DSP56858 cuenta con un Módulo Generador de la Señal de Reloj (***CGM****: Clock Generation Module)* que permite el uso de un cristal de 4MHz o una fuente de reloj externa para que el **PSD** funcione a frecuencias desde 0 a 120 MHZ.

Para configurar el **PSD** a una velocidad de 120 MHZ, se utilizan dos registros de 16 bits del ***CGM***: registro de control

(***CGMCR***) y registro divisor (***CGMDB***). En el registro ***CGMCR*** se escribe el valor 0x0800 para activar un oscilador enganchado en fase (PLL "PHASE LOOP LOCKED**")**, el cual permitirá generar una alta frecuencia (120 MHZ) a partir de una baja frecuencia (4 MHZ).

En el registro ***CGMDB*** se escribe el valor de 0x003b para que un oscilador controlado por voltaje (***VCO:*** *Voltage-controlled oscillator*) produzca una frecuencia de 240 MHz. En el *DSP56858*, la fórmula para calcular el valor que se escribe en el registro es:

$$Fvco\_out = Fref \times (PLLDB + 1)$$

donde Fref= 4 MHZ y FVCO_OUT=240MHZ y PLLDB es el valor que se escribe en el registro.

Configuración de los puertos ESSI y GPIO.- El codificador-decodificador de voz está conectado al **PSD** utilizando el puerto *ESSI0*, el cual cada uno de sus pines pueden funcionar también como puerto ***GPIO*** (Puerto C) . Los pines del puerto C son *GPIOC0, GPIOC1, GPIOC2, GPIOC3, GPIOC4 Y GPIOC5.*

Los pines *SDIN, SDOUT, SCLK, FSYNC* del codificador-decodificador de voz se conectan al **PSD** utilizando los pines *GPIO0, GPIOC0, GPIOC1, Y GPIOC5,* los cuales deben configurarse en el modo de *interfaz* ***ESSI***. Asimismo, los pines *CCS, CDIN, y CCLK* del

codificador-decodificador de voz, se conectan a los pines PC3, PC4 (puerto C), PE2 y PE3 (puerto E) del **PSD** y se configura en el modo ***GPIO***.

La configuración en modo ***ESSI***, es a través de registros correspondientes al puerto C. En el registro *GPIO_C_PER* se escribe el valor 0x0018 para que los pines PC3 y PC4 (puerto C) funcionen como ***GPIO*** y los pines PC0, PC1, PC2 Y PC5 se configuran como pines de la interfaz ***ESSI*** (*SDIN, SDOUT, SCLK, FSYNC*). Una vez que se definieron los pines de la interfaz ***ESSI***, se realizará su configuración a través registros de 16 bits asociados a esta interfaz para determinar su funcionamiento.

En el registro *ESSI0_STXCR2* se habilita la interrupción cuando se transmiten o reciben datos , se habilita el transmisor y el receptor de la interfaz ***ESSI***, se habilita la interfaz y se configura una trama de datos con una longitud de 16 bits, esto se debe a que el codificador-decodificador de voz transmitirá tramas con 2 muestras de voz de 16 bits.

Configuración de la interfaz ESSI en modo de acceso directo a memoria (DMA: *Direct Memory Access****).-*** En el registro *ESSI0_SCR3* se configura la trama de datos, es decir, el orden de los bits de cada trama. También, en este registro se establece una

transmisión de datos entre el **PSD** y el codificador-decodificador de voz en modo ***DMA***.

3.6.1.1. Transmisión de datos del codificador-decodificador de voz al PSD (modo *DMA*).

Establecidas las configuraciones iniciales, se inicia la transmisión de datos del codificador-decodificador de voz al **PSD**.

En este proceso de transferencia de datos, se utilizaron 2 contenedores de entrada con un tamaño de 240 registros de 16 bits, que corresponden a 240 muestras de voz. El estándar de compresión **CPLE** utiliza tramas de voz 22.5 milisegundos correspondientes a 240 muestras de 16 bits, y para la implementación de este estándar se requieren 240 muestras para obtener 144 bits de información comprimida y colocarlos en bloques de 128 bits para su cifrado. Estos contenedores hacen posible una transmisión de datos en tiempo real. Mientras un contenedor se “llena”, el otro contenedor se “vacía” hacia el proceso de compresión y cifrado.

Para tener una transmisión en modo ***DMA***, se configura el controlador ***DMA*** correspondiente al canal *DMA0* del **PSD** a través de registros de 16 bits. Primeramente en los registros *DMA_0_SAL* y *DMA _0_DAL* se establece la dirección del registro de recepción de la interfaz ***ESSI***, y en los registros *DMA_0_DAL* y *DMA_0_DAH* la dirección de memoria donde inicia el contenedor de recepción de las tramas de voz. En el registro *DMA_0_TCNT* se determina el número de tramas que se trasferirán desde la interfaz ***ESSI*** hacia el contenedor del **PSD**, siendo 240 tramas de voz para el estándar de compresión **CPLE**. Finalmente se configura el registro de control del canal *DMA0 DMA_0_TCTRL* en donde se habilita una interrupción que ocurre cuando el controlador haya transferido 240 tramas de 16 bits (número de tramas determinadas en el registro *DMA_0_TCNT*). También en este registro *DMA_0_TCTRL* se habilita la transferencia de datos en modo ***DMA***, y se determina el periférico que hará la transmisión ***DMA*** (registro de recepción interfaz ***ESSI***).

Configurado lo anterior, en el mismo registro de 16 bits se “enciende” un bit del registro para que inicie la

transferencia de datos. Una vez que el contenedor se "llena", el controlador ***DMA*** activa una interrupción que permite ejecutar una rutina de cambio de contenedor (en el registro DMA_0_TCTRL se escribe la nueva dirección).

El proceso de compresión de las 240 muestras de voz se inicia con el contenedor "lleno" mientras se almacenan nuevas muestras en el otro contenedor.

3.6.1.2. Transmisión de datos del PSD al codificador-decodificador de voz (modo *DMA*).

Al igual que el proceso anterior, la transmisión de datos es en modo ***DMA***, y se utilizan 2 contenedores para una reproducción de voz en tiempo real. En este caso, se transmiten datos de un contenedor a la interfaz ***ESSI***.

En el registro *DMA_1_SAL* se escribe la dirección donde inicia el contenedor que contiene los datos (voz descomprimida y descifrada), en los registros *DMA_1_DAL* y *DMA_1_DAH* se escribe la dirección de memoria correspondiente al registro de transmisión de la Interfaz ***ESSI***, en el registro *DMA_1_TCNT* se determina el numero de tramas de 16 bits que se transmitirá que para este caso es de 240. Una vez que se configuran los registros

anteriores, en el registro de control del canal *DMA1 DMA_1_TCTRL* se habilita la transmisión en modo ***DMA*** y una interrupción cuando el controlador transfiera 240 tramas de 16 bits. También en este registro se establece el tamaño de cada trama (16 bits).

Una vez que se "vacía" el contenedor que contiene los datos de voz descomprimida, el controlador ***DMA*** detiene la transferencia y se activa una interrupción, lo cual inicia una rutina que realiza la configuración de los registros *DMA_1_DAL y DMA_1_DAH* para establecer la dirección del otro contenedor, y nuevamente "enciende" el bit de activación de la transferencia de datos en modo ***DMA*** en el registro DMA _1_TCTRL.

3.6.2. Comunicación de datos entre el PSD y el módem.

En este trabajo se propone utilizar dos módems, uno en la parte transmisora y otro en la receptora. Para la transmisión de voz comprimida y cifrada a través del módem, se tiene el siguiente protocolo:

- El teléfono de la parte transmisora inicia la llamada a través del módem, el cual marca el número del módem del teléfono receptor.
- El módem receptor contesta y se inicia la negociación entre los 2 módems.
- Una vez que los módem se han establecido el enlace, el teléfono transmisor envía voz al **PSD** para su compresión y cifrado, y este a su vez al módem para su transmisión a la parte receptora, la cual realiza el proceso inverso.
- Se considera que tanto la terminal receptora como la transmisora tienen la misma clave de descifrado. En este trabajo no considera el intercambio de claves.

3.6.2.1. Transmisión.

Antes de iniciar la transmisión, se configura el módem transmisor con una cadena de inicialización, la cual consta de *comandos AT*[2] para establecer la forma en

[2] Los comandos AT conocidos también como comando Hayes (desarrollados en 1977 por Dennos Hayes), son instrucciones codificadas que funcionan como una interfaz de comunicación con un módem para configurarlo y proporcionarle instructores como marcar un numero, "colgar", velocidad en baudios entre muchos más. Se denominan AT por la abreviatura de *attention*.

que trabajará el módem. El algoritmo de transmisión envía la cadena de inicialización del **PSD** al módem por el puerto serial y posteriormente el numero telefónico del módem receptor para iniciar la marcación y el establecimiento del enlace entre los módems.

En una comunicación serial, primeramente se activan las señales de control ***RTS*** y ***DSR*** para indicarle al módem que el **PSD** esta listo para transmitir, por lo que se configuran los pines del puerto B en modo ***GPIO*** mediante los registros *GPIO_B_PER, GPIO_B_DDR, GPIO_B_DR.*

Posteriormente, se realiza la configuración de los parámetros del puerto serial como son: la velocidad de transmisión, 8 bits de datos, sin bit de paridad, bit de inicio y bit de paro, utilizando el registro de 16 bits *SCI_0_CR* y *SCI_0_BR.* También a través del registro *SCI_0_CR2* se activa el módulo de transmisión serial.

La interfaz de comunicación serial cuenta con un registro de corrimiento *SCI_0_DR* donde se colocan los datos que serán transmitidos, y es en este registro donde se envían los comandos AT para control del módem. La información a trasmitir se coloca en este registro sólo

cuando se encuentra vacío, verificando esta condición mediante el registro de estado *SCI_0_SR*.

La cadena de inicio que se envía al módem esta compuesta por los siguientes comandos AT:

AT$SB115200&e0&e3\r

Esta cadena configura el puerto serial del módem a una velocidad máxima de 115200 bps, sin compresión de datos ni corrección de errores. Esto se debe a que existe ya una compresión de voz y la corrección de errores implica retransmisiones, provocando pérdidas de muestras de voz. Es importante considerar que el codificador de voz **CPLE** incluye un código corrector de errores. Además, en las pruebas realizadas en la línea telefónica, la supresión de la corrección de errores del módem no afecta la transmisión de datos. Una vez que el módem fue configurado con parámetros iniciales, se procede a enviar el comando AT al módem para marcar el número correspondiente:

ATDT56587404;

Así como se configura el módem transmisor, el módem receptor se configura para una contestación automática, por lo que el módem receptor al recibir el timbrado descolgará e iniciará la negociación entre los módems receptor y transmisor, si la negociación fue exitosa, se establecerá el enlace y ambos módems pasarán al modo en línea, es decir, que la información que reciban del **PSD** no las considerará como comandos AT, sino como la información que debe modular o demodular.

La transmisión de la voz cifrada se realizará en modo ***DMA*** a través del canal *DMA_1*, y se procede a configurar los registros asociados a éste de manera similar en la transmisión de datos entre el **PSD** y el codificador-decodificador de voz, solo que aquí la comunicación será entre el **PSD** y el módem a través del puerto serial.

Se crearon 2 contenedores de 8 localidades de memoria de 16 bits, para almacenar 128 bits de voz comprimida y cifrada, correspondiente a 240 muestras de voz de 16 bits (30 milisegundos). Es importante recordar

al descifrar los 128 bits, 20 bits son únicamente de relleno.

De estos 2 contenedores, el módem transmitirá la información en modo ***DMA***, siendo necesario configurar el controlador ***DMA*** con los registros asociados a este controlador. En los registros *DMA_1_SAL DMA_1_SAH* se escribe la dirección donde inicia el contenedor que contiene los datos, en los registros *DMA_1_DAL* y *DMA_1_DAH* se escribe la dirección de memoria correspondiente al registro de transmisión de la Interfaz ***ESSI***, en el registro *DMA_1_TCNT* se determina el número bloques de 16 bits que se transmitirán del contenedor al módem por el puerto serial que para este caso es de 8 bloques de 16 bits.

Una vez que se configuran los registros anteriores, se inicia la transmisión de datos del **PSD** al módem, y se configura el registro de control del canal *DMA_1_TCTRL* habilitando la transmisión en modo ***DMA*** y una interrupción para que el controlador ***DMA*** termine el flujo de datos cuando se hayan transmitido 8 bloques de 16 bits. También en este registro se establece el tamaño de cada trama (16 bits).

Una vez que el contenedor que contiene los datos de voz comprimida y cifrada, se "vacía", el controlador ***DMA*** detiene la transferencia y se activa una interrupción, lo cual inicia una rutina de servicio que realiza la configuración de los registros *DMA_1_DAL* y *DMA_1_DAH* para establecer la dirección del otro contenedor, y nuevamente "enciende" el bit de activación de la transferencia de datos en modo ***DMA*** en el registro *DMA_1_TCTRL*.

Los datos enviados al módem a través del puerto serial, son modulados y transmitidos por la línea telefónica a la velocidad que hayan establecido ambos módems. Esta velocidad puede variar dependiendo el ruido de la línea.

3.6.2.2. Recepción.

Al igual que en el proceso de transmisión, mediante registros de 16 bits asociados a la interfaz serial ***SCI*** se activan las señales ***RTS*** y ***DSR*** para indicarle al módem que el **PSD** está listo para enviar recibir datos y posteriormente el módem se configura con una cadena de

inicialización utilizando comandos AT. La cadena de inicio que se envía al módem es la siguiente:

AT$SB115200&e0&e3\r

La cadena configura el puerto serial del módem a una velocidad máxima de 115200 bps, sin compresión de datos ni corrección de errores así como para contestar automáticamente la llamada del módem transmisor.

Establecido entre ambos módems el enlace, el receptor comienza a recibir los datos para demodularlos y transmitirlos al **PSD** a través del puerto serial.

La recepción de la voz cifrada se realizará en modo ***DMA*** a través del canal *DMA0*, y se configuran los registros asociados a este canal. Se crearon 2 contenedores de 8 localidades de memoria de 16 bits para almacenar 128 bits de voz comprimida y cifrada correspondientes a 240 muestras de voz de 16 bits (30 milisegundos) y que son enviados por el módem transmisor.

De estos 2 contenedores, el módem transmitirá los datos recibidos en modo ***DMA*** hacia el **PSD** la información

cifrada, siendo necesario configurar el controlador ***DMA*** con los registros asociados a este controlador. En los registros *DMA_0_SAL DMA_0_SAH* se escribe la dirección correspondiente al registro de corrimiento del puerto serial y en los registros *DMA_1_DAL* y DMA_1_DAH se escribe la dirección de memoria correspondiente al inicio del contenedor donde se almacenaran los datos enviados por el módem. La capacidad de cada contenedor de recepción es de 128 localidades de memoria de 16 bits.

Una vez que se configuraron los registros anteriores, se inicia la transmisión de datos del módem receptor al **PSD**, y se configura el registro de control del Canal DMA0 DMA_0_TCTRL habilitando la transmisión en modo ***DMA*** y una interrupción para que el controlador ***DMA*** termine el flujo de datos cuando se hayan transmitido 8 bloques de 16 bits. También en este registro se establece el tamaño de cada bloque (16 bits).

Cuando se llena uno de los contenedores con voz comprimida y cifrada, el controlador ***DMA*** detiene la transferencia y se activa una interrupción, lo cual inicia una rutina de servicio que realiza la configuración de los

registros *DMA_0_DAL y DMA_0_DAH* para establecer la dirección del otro contenedor, y nuevamente "enciende" el bit de activación de la transferencia de datos en modo ***DMA*** en el registro *DMA_0_TCTRL*.

Simultáneamente, otra rutina del programa se encarga de leer los contenedores "llenos" y pasarlos a las siguientes etapas de descompresión y descifrado de la voz.

3.7. ALGORITMOS DE COMPRESIÓN Y CIFRADO.

3.7.1. Compresión-cifrado.

En la conversión de la voz analógica-digital, se emplean convertidores de 16 bits mono estereo, obteniéndose voz digital a una velocidad de 128 kbps. Para la compresión de la voz se implanta el algoritmo **CPLE** a una tasa de compresión de 4.8 kbps. el cual requiere a la entrada 240 muestras de 16 bits (30 milisegundos) y se obtienen a la salida 144 bits. En el cifrado se utiliza el **AEC-2001**, que cifra la información por bloques de 128 bits con una clave de 128 bits, obteniéndose a la salida el mismo numero de bits.

Podemos observar que el número de bits a la salida del algoritmo de comprensión es diferente al número de bits por bloque que requiere el **AEC-2001**. Por lo que se toman 128 del grupo de 144 bits a la salida del **CPLE** (que corresponden a 240 muestras o 30 milisegundos de voz) para formar un bloque, y los 16 bits restantes se colocan en otro bloque. Esto significa que el segundo bloque de 16 bits corresponde al **CPLE**, y los 112 "espacios" del bloque se "llenan" con ceros.

COMPRESION-ENCRIPCION

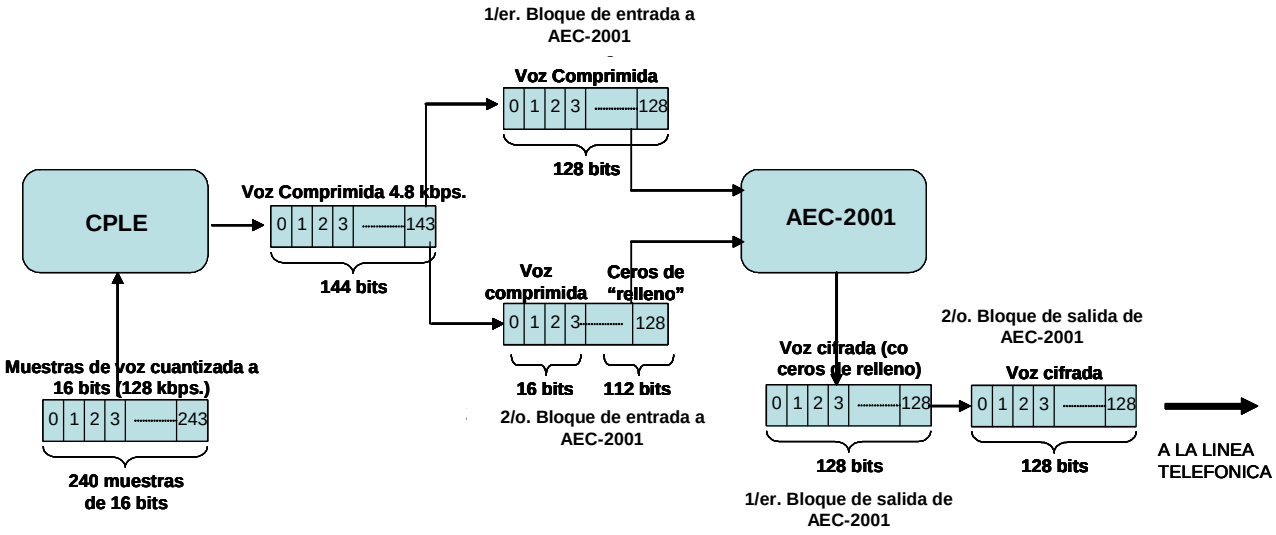


Figura 3.12. Diagrama esquemático del proceso de compresión y cifrado de la voz.

3.7.1.1. Compresión CPLE.

Este algoritmo comienza con una inicialización de parámetros, posteriormente la fase de análisis para

determinar los parámetros del filtro, pitch (ritmo), ganancias, magnitudes de Fourier. Por último, el algoritmo coloca la voz comprimida en 2 contenedores de salida de un tamaño de 9 registros por 16 bits.

3.7.1.2. Cifrado.

El algoritmo consta de las siguientes etapas:

- **Colocación de voz comprimida en un arreglo de 128 bits.** A la salida del proceso de compresión de la voz, se tienen 128 bits por cada 3840 bits a comprimir (240 muestras de voz de 16 bits). Debido a que el estándar **AEC-2001** admite bloques de 128 bits, se toman 128 de los 144 de voz comprimida (30 milisegundos) para colocarlos en una matriz de 4X4 llamada matriz de estado 16 bytes (128 bits).
- **Función de Adición de la clave propia de la ronda.-**Esta función toma una matriz de estado y simplemente hace una operacion *XOR* octeto a octeto con la correspondiente matriz de subclaves

dependiendo de la ronda. Como resultado se obtiene una nueva matriz de estado que se utiliza en las siguientes funciones.

- **Función Sustitución de octetos.-** En esta función cada elemento se sustituye por un elemento de matriz la cual se define al inicio del programa (matriz S[***256***]). Esta matriz, como se explico anteriormente, esta definida por el estándar y se denomina Tabla de Sustituciones **AEC-2001**.
- **Función Corrimiento circular a la izquierda.-** esta transformación se aplica a la matriz estado, aplicando corrimientos circulares de bytes a sus columnas.
- **Función Batido columna por Columna.-** esta transformación se realiza al multiplicar la matriz de estado por una columna constante que es resultado de un polinomio constante *c(x) = 03 x +01 x +01 x + 02*, el resultado es otra matriz de estado.
- **Colocación de la matriz de estado en el contenedor de salida.-** la matriz de estado final contiene la información cifrada y es colocada en el

contenedor de salida. Cada matriz de estado cifrada es colocada en el contenedor de salida de 8 localidades de memoria de 16 bits. El contenedor de salida contiene un parte de voz comprimida de las 240 muestras de voz de 16 bits (30 milisegundos).

- **Colocación de los bits restantes de los 144 bits a la salida del CPLE.-** Inicialmente se forma un bloque de 128 bits de los 144 bits a la salida del **CPLE** requeridos por el algoritmo **AEC-2001**. Este bloque entra al algoritmo de cifrado y se obtiene un bloque cifrado de 128 bits. Ahora, se toman los 16 de los 144 bits restantes a la salida del algoritmo de compresión **CPLE** para formar otro bloque de 128, "llenando" con ceros los 112 "espacios" del bloque.
- **Función de expansión de la clave**.- Esta función extiende la clave de cifrado a 16 subclaves de 16 bits que también son colocadas en una matriz de 4 x 4 , las cuales se emplearan en cada una de las funciones de transformación como son ***Sustitución de Octetos,*** **Corrimiento Circular a**

la Izquierda, *Batido por Columnas y* Adición de la clave propia de la Ronda en las 10 rondas que se ejecutaran estas funciones.

3.7.2. Descifrado-descompresión.

La voz comprimida y cifrada se transmite a una velocidad de 4.8 kbps. a través de la línea telefónica y esta protegida de usuarios no autorizados para poder escuchar la información. El receptor autorizado debe contar con el decodificador de voz **CPLE** y lo más importante la clave de descifrado de 128 bits, de lo contrario sin esta clave le será imposible descifrar la voz.

En la parte receptora, el primer proceso que se debe realizar es descifrar la información que se recibe a través de la línea telefónica. En el transmisor la información se transmite en bloques de 128 bits, por que en la descifrado-descompresión es en bloques de 128 bits.

Recordemos que de acuerdo al algoritmo de compresión **CPLE** y al **AEC-2001** se toman 240 muestras de voz cuantizada a 16 bits para su compresión y cifrado. Los 144 bits se cifran en dos bloques de 128 bits (un bloque con 128 bits y otro con 16 bits de voz comprimida rellenando el ultimo bloque con 112 “ceros”). Por lo que

esto debe ser considerado en el proceso de descifrado y descompresión (figura 3.13).

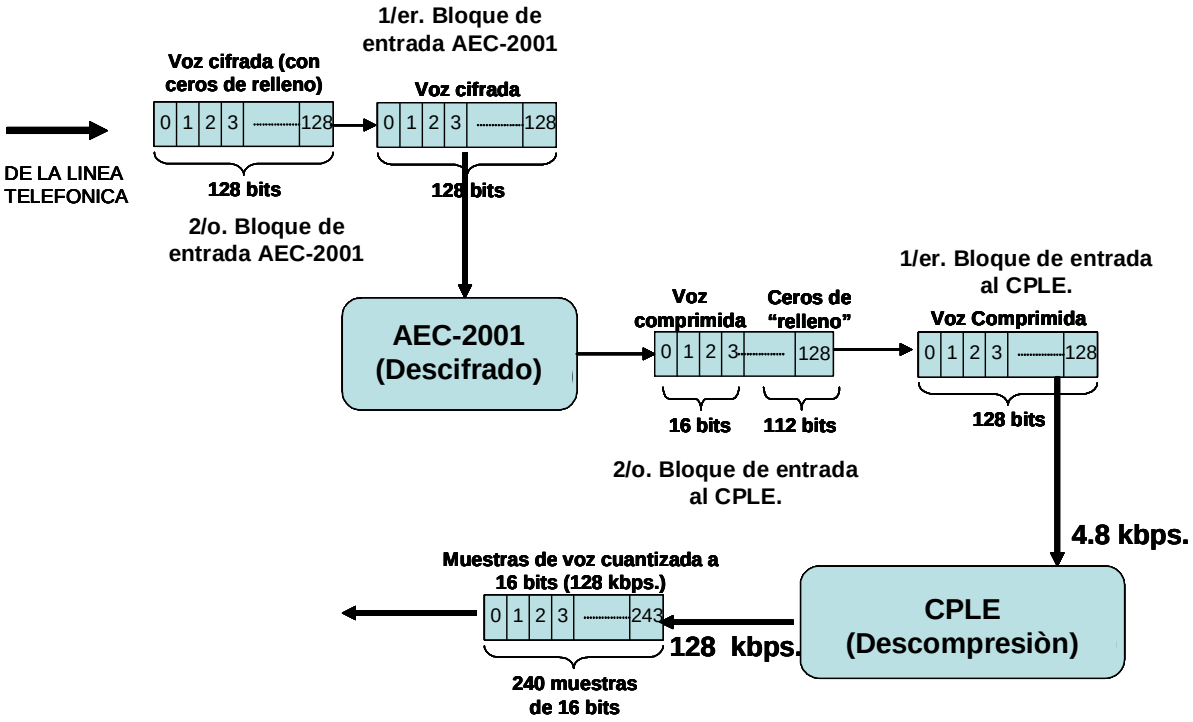


Figura 3.13. Diagrama esquemático del proceso de descifrado y descompresión de la voz.

3.7.2.1. Descifrado AEC-2001.

El algoritmo consta de las siguientes etapas:

- **Colocación de bloques bloque de voz cifrada y comprimida en arreglos de 128 bits.** en el proceso de descifrado, se 2 toman bloques de 128 bits cifrados. El primer bloque cuenta con 128 bits de información cifrada y comprimida, y el segundo

bloque cuenta con 16 bits de voz cifrada y comprimida combinada con “ceros” de “relleno”. Los datos de cada bloque se colocan en matrices de 4X4 bytes llamadas matriz de estado de 16 bytes (128 bits).

Al terminar el descifrado, la información descifrada esta disponible en la matriz de estado de 4 X 4 bytes (128 bits).

La primer matriz de estado descifrada correspondiente al primer bloque de 128 bits, contiene la voz descifrada y comprimida (128 bits). La segunda matriz de estado descifrada se obtiene al introducir al algoritmo el segundo bloque de 128 bits, obteniendo como resultado 16 bits de voz comprimida y 112 bits de “ceros” de relleno.

- **Colocación de la matriz de estado en el contenedor de entrada para el algoritmo de descompresión.-** cuando a la salida del algoritmo se obtiene la matriz de estado final con 128 bits de información descifrada, estos son colocados en los primeros 128 espacios de un contenedor de 144

elementos del tipo entero, y al obtener la segunda matriz solo se toman los primeros 16 bits y se colocan en los 16 espacios restantes del contenedor de 144 elementos. Una vez lleno el contenedor se inicia el proceso de descompresión.

3.7.2.2. Descompresión.

El algoritmo de descompresión **CPLE** toma los datos descifrados del contenedor de salida del **AEC-2001**, el cual contiene 144 bits de voz comprimida.

La rutina de decodificación inicializa los parámetros de decodificación, realiza la síntesis de la voz utilizando los parámetros recibidos, y posteriormente el resultado de la decodificación los coloca en 2 contenedores de salida de 240 registros por 16 bits.

Capítulo

4

RESULTADOS.

En el desarrollo del presente trabajo, se efectuaron distintas pruebas para determinar la mejor opción de compresión y cifrado de la voz. En cada prueba, inicialmente se realizaron simulaciones utilizando el lenguaje C, *Matlab* y un software para edición de voz y audio (*Cool Edit*). Una vez concluidas las simulaciones, se procedió a implantar los distintos algoritmos en el procesador *DSP56858*.

Para transmitir una voz cifrada a través del lazo local, ésta requiere de una digitalización, una compresión y una modulación digital con velocidad de transmisión menor que la permitida por el lazo local. Otra posibilidad para la transmisión de la voz cifrada que evita el proceso de compresión, cuyo tiempo de procesamiento es tardado y complejo, es transmitir la señal de voz cifrada en forma analógica sin el uso de módems telefónicos. Los resultados que se exponen corresponden a pruebas realizadas con voz sin comprimir, cifrada por flujo y transmitida con convertidores D/A

y A/D, así como pruebas de voz comprimida, cifrada por bloques y transmitida por medio de módems.

Voz cifrada con cifradores de flujo y transmitida con convertidores A/D y D/A.

El cifrado por flujo se emplea principalmente cuando se tiene limitación en el tiempo de procesamiento y el canal de comunicación se ve afectado considerablemente por el ruido. En el presente trabajo, se emplean estos cifradores para evitar la compresión de la voz y transmitirla como una señal analógica. Sin embargo, este tipo de señal a diferencia de una digital, se ve más afectada por las características propias del lazo local, tales como el ruido, limitación del ancho de banda y la distorsión.

Cuando un bloque de información cifrada es alterado en tan solo un bit, el algoritmo de descifrado obtiene erróneos todo los bits del bloque. En cambio, cuando la información se cifra bit a bit, la alteración de uno de ellos no repercute en el descifrado de los demás bits. Factores como el ruido afectan más a una señal cifrada por bloques que a una cifrada por flujo, razón por la cual se realizaron pruebas de cifrado de la voz bit por bit transmitiendo la señal cifrada como una señal analógica.

La figura 4.1 muestra un esquema de transmisión de voz cifrada y transmitida de una manera analógica, y que se toma como base para experimentar y determinar la factibilidad de emplear esta forma de cifrado en el desarrollo del sistema de cifrado de voz.

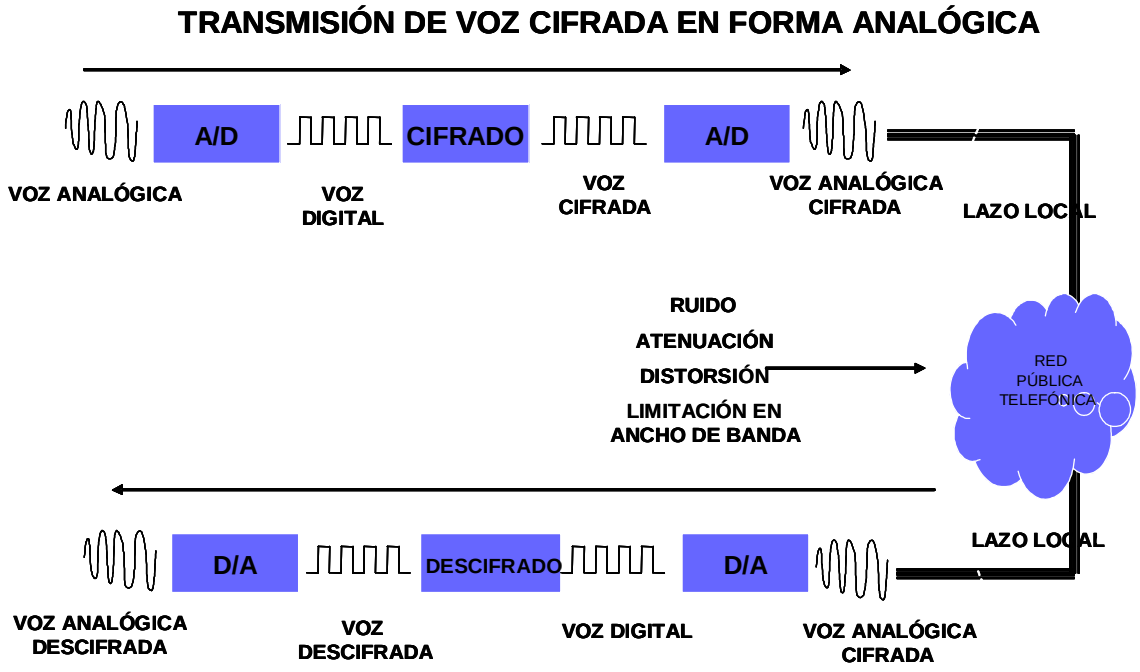


Fig. 4.1.- Esquema de transmisión de voz cifrada en forma analógica.

Con *Cool Edit* se obtiene un archivo de voz con una frecuencia de muestreo de 8000 Hz cuantizada con 16 bits en modo monoestereo; este archivo de voz se cifra con el algoritmo de flujo *SEA* (en lenguaje C) y se obtiene un archivo con voz cifrada. Posteriormente con *Cool Edit* se le aplica un filtro con las características similares a las de una línea telefónica.

La figura 4.2 (a) representa en el tiempo, una señal de voz de 4000 muestras con una frecuencia de muestreo de 8000 Hz y la figura 4.2 (b y c) la señal cifrada con su espectro de frecuencia.

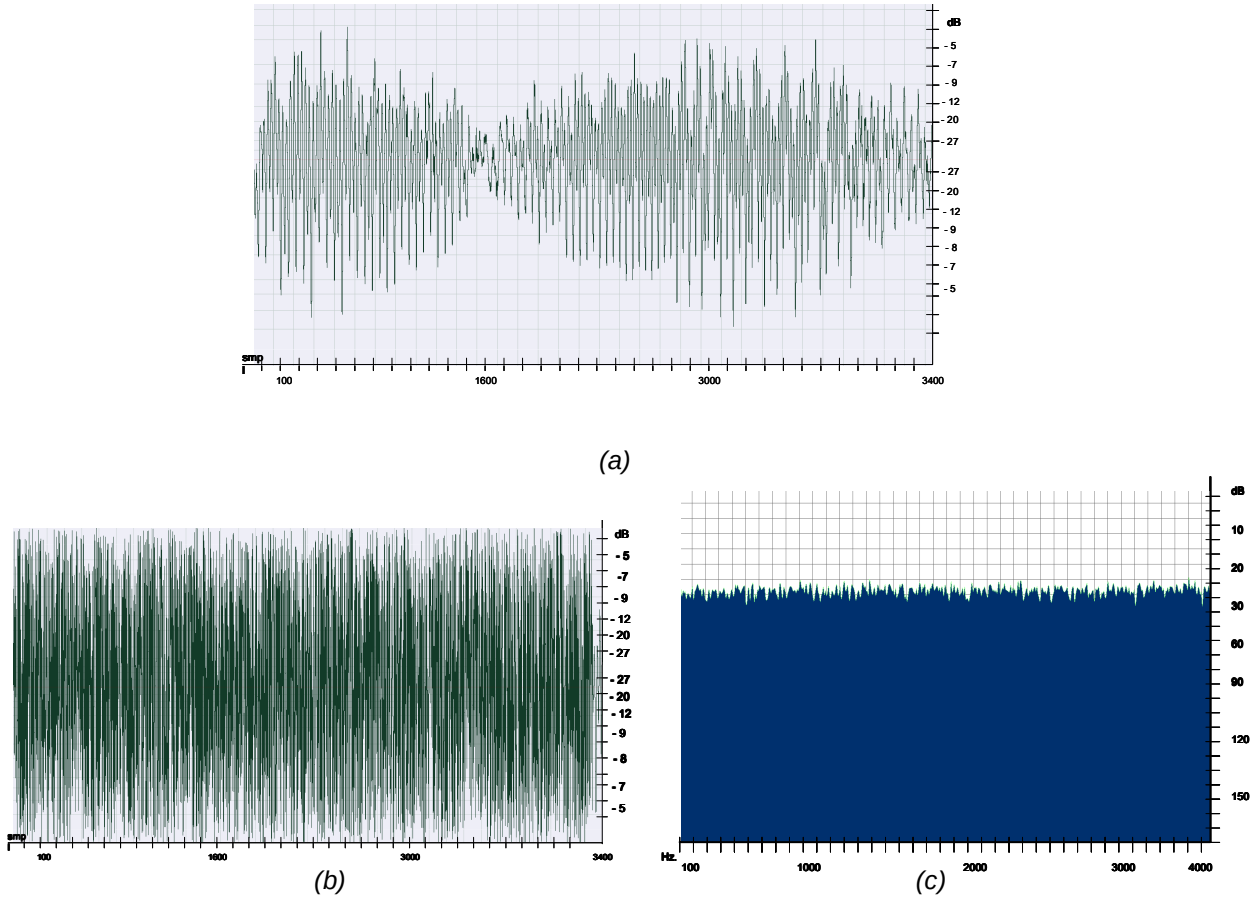
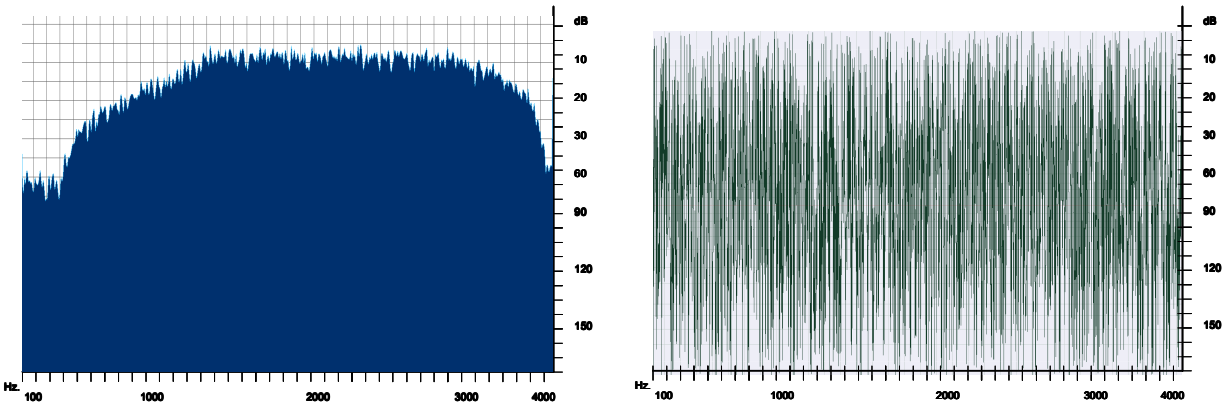


Figura 4.2.- (a) Voz en claro. (b) señal cifrada, y (c) espectro en frecuencia de la señal cifrada

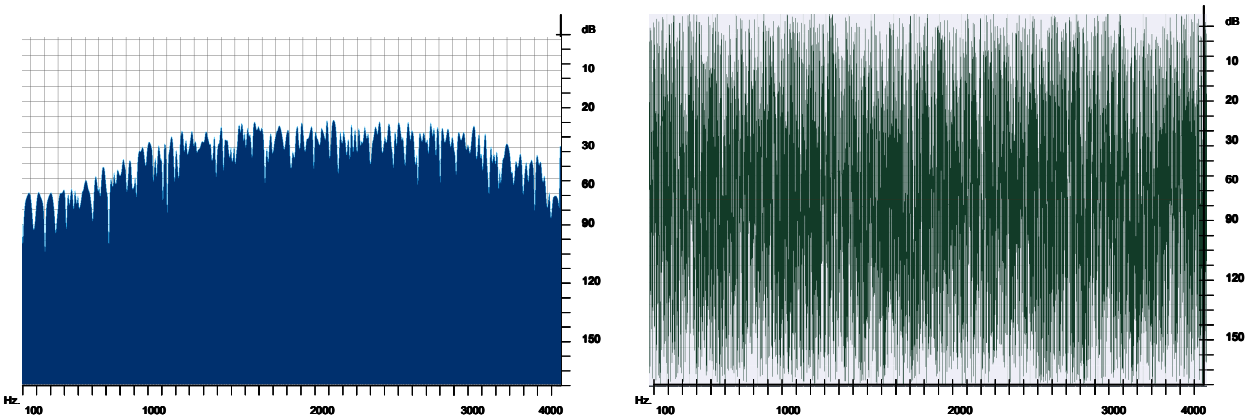
Con el software *Cool Edit* la voz cifrada es limitada en banda con un filtro de características similares al lazo local. En la figura 4.3 (a) se observa el resultado del filtrado de la señal, y por lo tanto, el descifrado no permite recuperar correctamente la señal de voz tal como se observa en la figura 4.3 (b), escuchándose con demasiado ruido.



(a) *(b)*

Figura 4.3.- (a) Espectro de frecuenta de la señal cifrada limitada en ancho de banda, (b) señal de voz descifrada.

Para minimizar los errores, se aplica a la señal cifrada (antes de la conversión digital-analógica y la transmisión por el lazo local) un código corrector de errores BCH (38,48,5). La figura 4.4 muestra la señal cifrada limitada en banda y la señal de voz descifrada. Se observa que el ruido continúa presente en la señal.



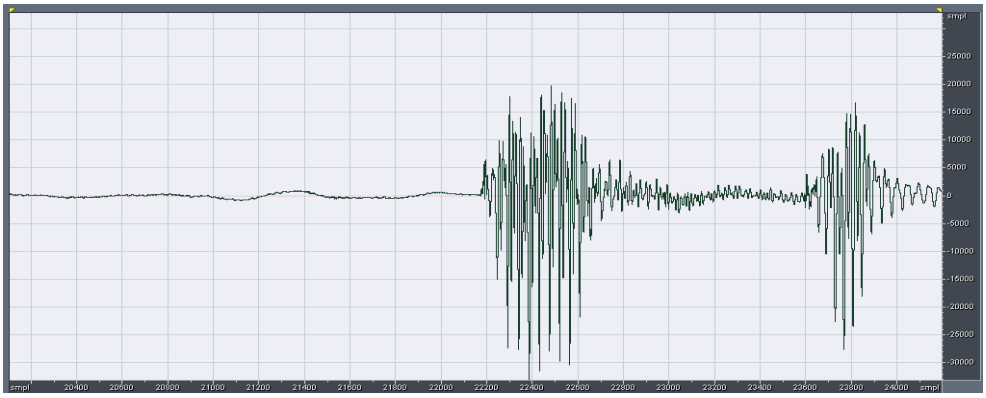
(a) (b)

Figura 4.4.- (a) señal cifrada con código corrector limitada en banda. (b) voz descifrada.

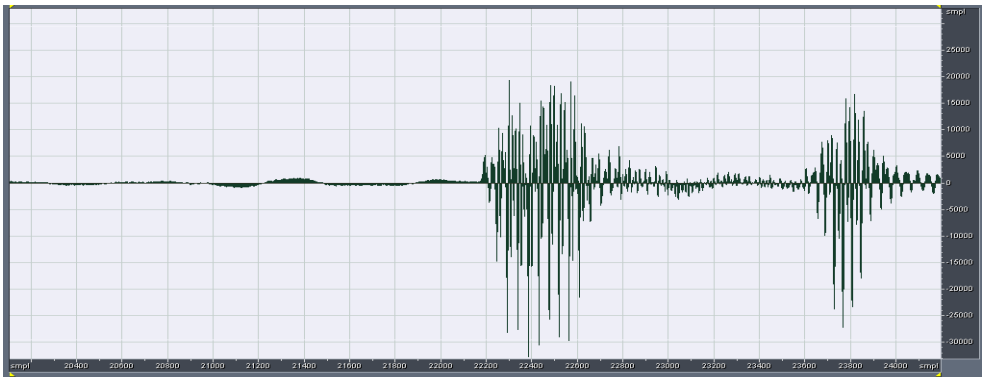
Voz cifrada con cifradores de bloque y transmitida con módems.

El lazo local limita la velocidad de transmisión de datos a 33.6 kbps. en una dirección y se reduce a la mitad cuando es bidireccional, si se sobrepasa esta velocidad se tienen problemas de pérdidas de muestras de voz . Para evitar estas pérdidas se emplea la compresión, sin embargo existe un compromiso entre velocidad y tiempo de procesamiento en el **PSD**. Otro factor importante a considerar es la codificación de canal para protección contra errores incluida en los módems telefónicos, que permiten al usuario elegir la codificación más conveniente; en las pruebas realizadas se seleccionó una codificación de canal sin retransmisión en caso de detección de errores.

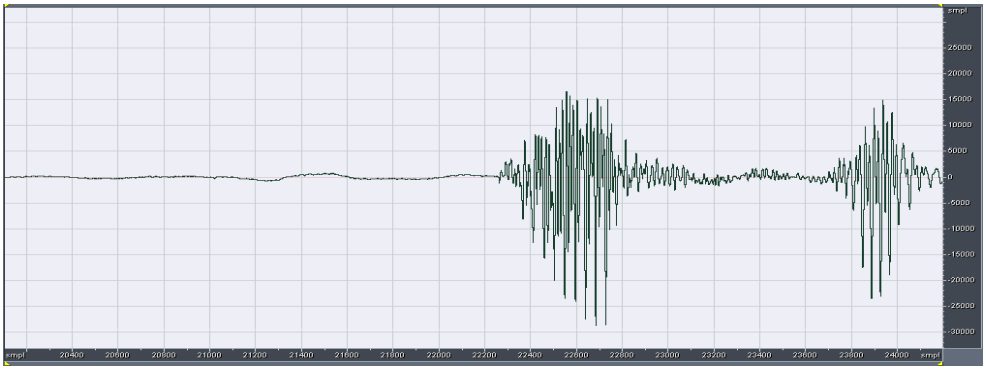
ADPCM es un codificador de forma de onda que permite obtener velocidades de hasta 16 kbps. Para velocidades menores son empleados *vocoders* como el estándar ***FS-1016*** con una compresión a 4.8 kbps. La figura 4.5 muestra una comparación de estos dos tipos de codificadores que comprimen una señal de voz muestreada con una frecuencia de muestreo de 8000 Hz. cuantizada con 16 bits.



(a)



(b)



(c)

Figura. 4.5 (a) Señal original y comparación de la compresión ***ADPCM*** (b) y compresión **FS-1016** (c)

En la siguiente tabla se muestran resultados de los tiempos de procesamiento de los algoritmos ***FS1016, ADPCM y*** **AEC-2001** implementados en el *DSP56858.*

ALGORITMO	MUESTRAS	TIEMPO DE PROCESAMIENTO
ADPCM	240 muestras de 16 bits (30 milisegundos)	0.928 milisegundos
FS1016	240 muestras de 16 bits (30 milisegundos)	39.024 milisegundos
AES (cifrado-descifrado)	240 muestras de 16 bits comprimidas a 144 bits. (FS1016)	5.024 milisegundos
AES (cifrado-descifrado)	240 muestras de 16 bits comprimidas a 480 bits. (ADPCM)	10.048 milisegundos

La implantación de los algoritmos de compresión y cifrado en el *DSP56858* y el empleo de módems telefónicos para la transmisión de la voz cifrada requiere considerar lo siguiente:

1. En el módem telefónico se desactiva la retransmisión en la detección errores, con el fin de no incrementar el tiempo de transmisión y evitar pérdidas de muestras de voz.
2. Si la velocidad de transmisión a la salida del puerto serial del **PSD** es menor que la velocidad de conexión entre ambos módems (33.6 kbps.), no será necesario aplicar un control de flujo entre el **PSD** y el módem
3. El puerto ***SCI*** transmite en forma serial tramas de 10 bits, de los cuales 8 son de datos, un bit de inicio y un bit de parada, de tal forma que sólo el 80 % de los datos que se transmiten corresponden a la información y el 20% a los bits de control.

4. Cuando se transmite la voz comprimida cifrada con **AEC-2001**, se realiza en bloques de 128 bits, siendo necesario una sincronía por bloques a través de bits de sincronía de trama.
5. El algoritmo de compresión codifica la voz en 144 bits, los cuales se colocan en dos bloques de 128 bits para su posterior cifrado con **AEC-2001**. El segundo bloque solo contiene 16 bits de voz comprimida y los espacios libres de este bloque se "rellenan "con ceros (112 bits). Los bits de "relleno", los bits de sincronía de carácter y los bits de sincronía de trama incrementan la cantidad de bits a transmitir respecto a la voz comprimida.

En las pruebas realizadas con el *DSP56858* se observó que para evitar pérdidas de muestras de voz, la velocidad programada en el puerto ***SCI*** debe ser mayor que la velocidad de compresión y menor que la velocidad de transmisión máxima permitida en el lazo local. Estos resultados se muestran en la siguiente tabla.

	Bits de información	**Bits de sincronía por bloque, sincronía de carácter y de "relleno"**	**Bits totales transmitidos por bloque.**	**Velocidad de transmisión mínima en el puerto *SCI*.**	**% de diferencia entre la velocidad de compresión y la velocidad en el puerto *SCI***
ADPCM (32 kbps.)	128	42	170	42.5 kbps.	32.8 %
ADPCM (16 kbps.)	128	42	170	21.250 kbps.	32.8 %
FS1016 (4.8 kbps.)	144	186	330	11.9 kbps	148 %

Costos.

La tarjeta de desarrollo empleada tiene la ventaja de integrar la mayoría de los componentes que se requieren en este trabajo como son un **PSD** para los procesos de compresión y cifrado, un convertidor analógico-digital y digital-analógico, además de los puertos serial asíncrono y síncrono para comunicación con el módem y el convertidor. Además, cuenta con un compilador de C con una interfaz grafica que permite una programación rápida y fácil. El costo de esta tarjeta es de $ 500.00 USD.

El módem utilizado es externo y de fácil programación con un costo de $110.00 USD.

También, es necesaria una licencia adicional cuando se tienen códigos mayores a 8 kbytes de programa, que para el caso del estándar ***FS-1016*** se requieren aproximadamente 80 kbytes (no así para ***ADPCM*** con un código de 5 kbytes). Esta licencia tiene un costo de $ 2500 USD. (Esta es una de las razones por la que no se continuó con la optimización del código). El costo total de este prototipo es el siguiente:

	Cantidad	**Costo**
Tarjeta DSP56858EVM	2	$1000.00 USD.
MODEM Externo	2	$220.00 USD.
TOTAL		$1220.00 USD.

El costo anterior no incluye la licencia.

Capítulo

5

CONCLUSIONES Y RECOMENDACIONES.

4.1. CONCLUSIONES.

En este trabajo se analizaron dos formas de proporcionar seguridad a la voz transmitida por la **RTB**: voz sin comprimir, cifrada por flujo y transmitida con convertidores D/A y A/D, y voz comprimida, cifrada por bloques y transmitida por medio de módems.

La voz cifrada transmitida con convertidores D/A y A/D (como señal analógica) permite evitar la compresión y utilizar un cifrado de flujo para realizar un procesamiento rápido. En la gráfica 4.2 (capítulo 4) se observa que la voz está limitada en banda (8 Khz) y su espectro de frecuencia se expande después del cifrado, de tal forma que cuando se transmite por el lazo local su espectro se limita a 8 khz, se altera su forma de onda en el tiempo y en consecuencia, la voz recuperada se escucha con demasiado ruido.

Los algoritmos de compresión **código de impulsos para la diferencia entre niveles adaptables** (***ADPCM:*** *Adaptive Differential Pulse Code*

Modulation) y cifrado **AEC-2001** consumen un tiempo de procesamiento de 10.976 milisegundos con 30 milisegundos de voz (utilizando el procesador *DSP56858*). Sin embargo, al momento de su transmisión, la velocidad de compresión de 16 kbps se incrementa en un 38% aproximadamente debido a los bits de sincronización a nivel carácter y a nivel bloque. Esta velocidad sigue siendo menor que la permitida por la línea (en una comunicación unidireccional), pero en una comunicación bidireccional simultánea, los módems transmiten y reciben a la mitad de la velocidad de conexión (aproximadamente a 16 kbps). Es importante señalar que en algunos casos, dependiendo del ruido existente en el lazo local, la máxima velocidad de conexión es a 19.2 kbps (unidireccional).

La cifrado y compresión de 30 milisegundos de voz con **AEC-2001** *y* ***FS-1016***, consumen un tiempo de 44.048 milisegundos. La velocidad de compresión es de 4.8 kbps, pero el incremento de información a transmitir debido a los bits de sincronización a nivel carácter y a nivel bloque, así como los bits de "relleno", requiere que la velocidad mínima que se programe en el puerto ***SCI*** sea de 11.9 kbps.

Esta velocidad permite una comunicación en tiempo real al ser menor que la permitida por el lazo local, sin embargo, existen pérdidas de muestras de voz debido al tiempo de procesamiento del algoritmo de compresión. Para evitar este problema se debe optimizar el algoritmo de compresión o utilizar un **PSD** más rápido. La opción de cifrado con **AEC-2001** y compresión con

ADPCM solo permite una comunicación unidireccional, alto nivel de seguridad, pero con una calidad de voz menor que la que proporciona el estándar ***FS-1016***.

La transmisión de voz cifrada por flujo en forma analógica con los convertidores A/D y D/A, es menos compleja que voz comprimida, cifrada y transmitida medio de módems. La desventaja, es que al transmitir la señal cifrada por el lazo local, es afectada por el ruido y la limitación del ancho de banda del lazo local, de tal forma que la voz no se recupera correctamente escuchándose con demasiado ruido. Asimismo, la sincronización durante la transmisión es más compleja.

Se puede concluir que la mejor opción para proporcionar seguridad a la voz que se transmite a través de la red pública telefónica, es por medio de un cifrado por bloques, con la posibilidad de utilizar como compresión el estándar ***FS-1016*** basado en los algoritmos tipo **CPLE** y algoritmo de cifrado **AEC-2001**, y transmitir la información por medio de módems, obteniendo alta seguridad en la voz y una comunicación bidireccional simultánea en tiempo real.

4.2. RECOMENDACIONES.

Por los resultados obtenidos, se recomienda optimizar los algoritmos de compresión y cifrado para su implementación en el *DSP56858* o utilizar un **PSD** de mayor velocidad de procesamiento, asimismo, se debe tener especial atención a la velocidad de compresión, velocidad de transmisión del puerto ***SCI*** y la velocidad de conexión entre los módems en una línea telefónica. En caso de continuar el presente trabajo utilizando el *DSP56858* también se requiere una licencia que permita compilar un código mayor a 8 kbytes.

También se recomienda estudiar la forma de transmitir en forma analógica (con convertidores D/A) la voz cifrada que fue primero adquirida con convertidores A/D. Aquí se usarían cifradores de flujo, y se necesitan soluciones para reducir la afectación por la limitación en ancho de la banda del lazo telefónico.

Otro tema de estudio interesante, es la forma de intercambiar las claves de cifrado a través de la misma red pública telefónica, ya que en el presente trabajo se parte de la suposición que tanto el receptor como el transmisor cuentan con la misma clave que recibieron por algún canal seguro.

En las comunicaciones asíncronas se tiene el inconveniente de que de la totalidad de los datos transmitidos, solo un 80% corresponde a la información y el 20% a los bits de sincronía. Por lo tanto, el empleo del puerto serial para transmitir los datos entre el ***PSD*** y módem requiere de una

velocidad mayor que la de compresión, sin embargo pueden existir problemas de pérdidas de muestras de voz cuando la velocidad del puerto serial sobrepasa la velocidad de conexión entre los módems. Es por eso que es recomendable utilizar una comunicación síncrona entre el **PSD** y el módem, la cual permite transmitir mayor cantidad de datos información que de sincronía.

GLOSARIO.

Acceso directo a memoria
En ingles:: *Direct memory access* ***(DMA)***

Es una característica de los microprocesadores o microcontroladores que permite a ciertos periféricos tener acceso al sistema de memoria para leer o escribir datos sin que intervenga la unidad central de proceso (***CPU:*** *Central Processing Unit*).

Algoritmo estándar de cifrado del 2001 (AEC-2001).
En inglés: *Advanced Encryption Standard* ***(AES)***

También conocido como *Rijndael*, es un esquema de cifrado por bloques adoptado como un estándar de cifrado por el gobierno de los Estados Unidos. El cifrador fue desarrollado por dos criptólogos belgas, Joan Daemen y Vincent Rijmen.

Aparato Telefónico Analógico
En inglés: *Analog Telephone Station*

Es el aparato telefónico que **no** cuenta con convertidores ***AD/DA*** para la transmisión del habla.

Aparato Telefónico Digital
En inglés: *Digital Telephone Station*

Es el aparato telefónico que **sí** cuenta con convertidores ***AD/DA*** para la transmisión del habla.

Aparato Telefónico Inteligente
En inglés: *Intelligent Telephone Station*

Se dice que un teléfono, ya sea analógico o digital, es inteligente, cuando cuenta con un microcontrolador incrustado (*"embedded microcontroler"*), que le permite proporcionar al usuario funciones superiores de control.

Batido columna por columna.
En inglés: MixColumns

Operación que es parte de las iteraciones aplicadas en el cifrado **AEC-2001**; cada columna se expresa como un polinomio de un campo de Galois, y se multiplica cada una por un mismo polinomio preestablecido, según las reglas de dichos campos.

BORSCHT

Designa a un grupo de funciones proporcionadas por los circuitos de cada **«interfaz de línea - analógica- de abonado»** de una **central telefónica local.**
BORSCHT es acrónimo de los nombres en inglés de las funciones batería, protección de sobrevoltaje, timbrado, supervisión, codificación, hibridación de direcciones de propagación, y prueba de funcionamiento: *(**B**attery, **O**vervoltge protection, **R**inging, **S**upervision, **C**oding, **H**ybrid y*

***T**est*). A estos circuitos de interfaz se les conoce en inglés de manera estandarizada como *"Subscriber loop interface circuit (**SLIC)**"*.

Calidad telefónica
En inglés:*telephone voice grade*.

Es la fidelidad e inteligibilidad propia de una señal de audio que se obtiene cuando se transmite por un enlace telefónico típico de la Red Telefónica Básica.

Central Telefónica
En inglés: *Central Office **(CO)***

Es el sitio de una compañía telefónica donde confluyen las líneas de suscriptor y donde los equipos de conmutación realizan la interconexión de los subscriptores (suscriptores transmisor y receptor).

Codificación para impulsos modulados: CIM
En inglés: *Pulse code modulation* **(*PCM)***

Es la codificación en una «palabra de impulsos», del valor cuantizado de cada uno de los impulsos, de un tren de impulsos modulados en amplitud-tope-plano *("Flat-Top-**PAM")**,* según una señal moduladora. Esta modulación en amplitud-tope-plano se conoce como muestreo.
Dentro de un microprocesador, la representación no es mediante «palabras de impulsos», sino mediante niveles lógicos, en Volts, correspondientes a palabras de dígitos binarios.

Codificador de Predicción Lineal con Señal de Excitación del Filtro Seleccionada Según un Código (CPLE).
En inglés: C*ode Excited Linear Prediction* **(*CELP*)**

Es un algoritmo de compresión del habla con el que se alcanza calidad suficiente con envíos menores a 4.8 kbps. Comprende: 1) filtros que modelen la producción biológica del habla, 2) predicción lineal, 3) diccionarios de códigos fijos y adaptables, 4) conjunto finito de señales de excitación y 5) cuantización vectorial.

Código de Impulsos para la Diferencia entre Niveles Adaptables.
En inglés: *Adaptive Differential Pulse Code Modulation* **(*ADPCM*)**.

Estándar de codificación de voz mediante el uso combinado de cuantización adaptable y predicción adaptable.

Compañía Portadora
En inglés: *Carrier company* ó simplemente *Carrier*

Compañía que posee y presta los circuitos de comunicación a través de los cuales se conducen los mensajes entre orígenes y destinos, o entre llamadores y llamados, es decir, entre los entes originarios y entes destinatarios.

Contenedor. En inglés: *Buffer*	Es un segmento de memoria para almacenar datos temporalmente, con el fin de apropiar la comunicación entre dos operaciones o procesos que se ejecutan a diferente velocidad. Zona (de memoria) para sincronizar", o zona de contención.
Corrimiento circular (izquierdo) de los renglones. En inglés*: Shift Columns*	Operación que es parte de las iteraciones aplicadas en el cifrado **AEC-2001;** cada renglón de la **matriz de estado** sufre un desplazamiento circular hacia la izquierda en diferente cantidad de posiciones, según el índice del renglón.
El habla. En inglés*: Speech*	Comunicación oral, discurso, lenguaje hablado. *No todo el habla está constituido por sonidos de voz; por eso, en sentido estricto, voz y habla NO SON SINÓNIMOS.*
Instituto Nacional de Estándares y Tecnología de los Estados Unidos. En inglés: *National Institute of Standards and Technology* ***(NIST)***.	Es una agencia de la Administración de Tecnología del Departamento de Comercio de los Estados Unidos de América. La misión de este instituto es promover la innovación y la competencia industrial en Estados Unidos mediante avances en metrología, normas y tecnología de forma que mejoren la estabilidad económica y la calidad de vida.
Interfaz de comunicación serial Mejorada. En inglés: *Enhanced Synchronous Serial Interface* **(*ESSI*)**	Es un interfaz propia de los procesadores de señales que se utiliza para comunicación serial bidireccional simultánea con otros dispositivos, y que consta registros transmisores y receptores independientes.
Línea analógica	Este término es una contracción de la frase nominativa: «Línea que transporta señales analógicas», donde por señal analógica se entiende aquella de tiempo continuo - amplitud contínua.
Línea de abonado, Línea telefónica, (Cada vez con mayor frecuencia se le dice simplemente **Línea**) En inglés: *Local loop*	Conexión física desde la propiedad del suscriptor hasta el **punto de presencia** (***POP* "*Point of Presence"***) del prestador del servicio telefónico.

Línea digital

Este término es una contracción de la frase nominativa: *«Línea que transporta señales digitales»*. Cabe aclarar que, sobre un par de cables de cobre, la señal digital, en el sentido de "secuencia de números", no existe. Lo que viaja aquí es una señal de voltaje/corriente, continua en el tiempo, y que puede o no tener valores discretos -separados- de amplitud. (*Véase* Señal digital). Lo que sí es una secuencia de números es la información transportada por estas señales.

Órganos del Habla:
En inglés: *Speech organ*

Se agrupan de la manera siguiente:

1. Órganos de respiración (Cavidades infraglóticas (debajo de la lengua): pulmones, bronquios y tráquea).
2. Órganos de fonación (Cavidades glóticas: laringe, cuerdas vocales y resonadores nasal, bucal y faríngeo).
3. Órganos de articulación (cavidad supraglóticas: paladar, lengua, dientes, labios y glotis).

Procesador de señales digitales (PSD):
En inglés: *Digital Signal Processor (**DSP**)*

Es un microprocesador especialmente diseñado para realizar tareas relacionadas con procesamiento de las señales digitales. Un PSD se diferencia de un microprocesador principalmente en que está diseñado para realizar en forma óptima cálculos matemáticos (principalmente multiplicación y la adición), a diferencia de un microprocesador de propósito general, el cual provee únicamente operaciones de manipulación de datos (movimiento de datos u operaciones de decisión)

Procesamiento del habla.
En inglés*: Speech processing:*

El Procesamiento del habla fue originalmente sólo el tratamiento por medio de dispositivos electrónicos *de la información contenida en las señales acústicas del discurso. Este tratamiento requiere del modelado matemático de las señales y de las operaciones a efectuar sobre ellas* [N. del Tut.]

Los objetivos que persigue pueden ser: la transmisión o la grabación de la señal, su excelente reproducción, su síntesis (texto a

discurso), reconocimiento (discurso a texto). o bien, la identificación del sujeto parlante o de su estado, características y atributos (por ejemplo, diagnóstico médico). Limitado originalmente al tratamiento analítico y estadístico de la señales acústicas, atrae ahora a disciplinas muy diversas, desde la fonética y la lingüística, hasta la inteligencia artificial para el desarrollo de sistemas expertos. [Adaptado de: R. Boite y M. Kunt, *Traitment de la parole*, Presses Polytechniques Romandes, 1987]

Pronunciar
En inglés: *Utter.*

Pronunciar, proferir, articular.

Puerto de Propósito General de Entrada y/o Salida.
En inglés: *General Purpose Input Output* **(*GPIO*)**

Tipo de puerto el cual funciona ya sea como una interfaz para comunicarse con otros periféricos, o de manera tal que las líneas que lo componen funcionan de manera independiente.

Red Telefónica Conmutada: RTC
En inglés: *Voice-oriented public telephone network*

Red Telefónica Conmutada es el nombre utilizado por Telefónica de España, *presumiblemente* de acuerdo con las normas de la Unión Internacional de Telecomunicaciones ***ITU-T***. Se define la Red Telefónica Conmutada como aquella constituida por todos los medios de transmisión y conmutación necesarios que permiten enlazar a voluntad dos equipos terminales (aparatos telefónicos, de facsímile, módems, etc.) mediante un circuito físico que se establece especialmente para la comunicación y que desaparece cuando ésta termina. Se trata pues de una red de telecomunicaciones que opera mediante *conmutación de circuitos* (físicos, no virtuales), diseñada primordialmente para la transmisión de voz, aunque pueda también transportar datos.

Las Características *esenciales* de la **RTC** son:

1. Consta de Centrales de conmutación, Troncales entre centrales, y **líneas de acceso** diversas, entre las que *necesariamente* se incluyen las *líneas de abonado tradicionales*, que son pares de cables de cobre**.** Los pares de cobre de un grupo de domicilios vecinos se agrupan

progresivamente en cables que tienen cada ves mayor calibre hasta llegar a la central local, y según la distancia, habrá *«repetidores pasivos»* en el trayecto.

2. Proporciona, además de otros Servicios de Telecomunicaciones posibles, el **Servicio de Telefonía Básica**. Para este fin, como *línea de acceso* instala hasta el domicilio de cada *abonado* (subscriptor) una o varias *líneas de abonado tradicionales* cada una de las cuales entrega al usuario un medio de transmisión para señales eléctricas colocadas en una banda base de 4KHz, y esto es lo que dispone cada abonado, tanto para recepción como para transmisión, cuando se establece una conversación telefónica entre dos domicilios. Esta banda incluye espacio para bandas de guarda y eliminación de interferencias provenientes de las líneas de «Distribución domiciliar de potencia eléctrica».
3. El costo de la ocupación del circuito depende de la distancia entre los extremos y la duración de la conexión.
4. Capilaridad: Los pares de cobre de un grupo de domicilios vecinos se agrupan progresivamente en cables que tienen cada ves mayor calibre hasta llegar a la central local, y según la distancia, habrá *«repetidores pasivos»* en el trayecto.
5. Normalización para su interconexión con otras **RTC**s.
6. Cobertura nacional e interconexión con las redes móviles.

Algunos países tienen sólo una compañía operadora de su **RTC**, como es la situación actual en México, mientras que en otros hay varias, como es en Estados Unidos.

Red Telefónica Conmutada Mundial (RTCM)
En inglés: *Public Switched Telephone Network* (***PSTN)***

Es la red formada por la interconexión de las **Redes telefónicas conmutadas** de la mayoría de los países, la cual da el servicio telefónico local, de larga distancia e internacional, y que usa el grueso de la población de manera cotidiana.

Ritmo o tono.
En inglés: *Pitch*

El tono o ritmo representa la frecuencia fundamental de un sonido.

Señal digital
En inglés: *Digital sigñal*

Es la información formada por una secuencia de números de precisión finita, tal y como puede aproximarse una señal analógica dentro de una memoria de computadora.
Por extensión, también se llama señal digital a cualquier señal eléctrica que represente con exactitud a una secuencia de números de precisión finita.

Servicio de Telefonía Básica (STB) En inglés: *Plain Old Telephone Service* ***(POTS)***

El «Servicio de Telefonía» al que puede accederse a través de los pares de cobre domiciliares conectados a teléfonos analógicos y que transmiten/reciben en la banda base de 4Khz

Sonidos (del habla) sin la Voz
En inglés: *Unvoiced (Speech) Sound*.

Sonidos del discurso que no son VOZ; seudoruidos, y/o semiruidos, del habla.

Sonido de la VOZ
En inglés: *Voiced Sound*.

Sonidos del discurso que sí son VOZ y no seudoruidos ni semiruidos; VOZ.

Subllave de cada ronda
En inglés: *Round Key.*

Cada iteración aplicada en el cifrado **AEC**-2001, incluye una etapa en la que se suma, módulo 2, una llave, relativamente larga, subordinada a una llave primaria más corta. De aquí el nombre de subllave. La subllave empleada en cada iteración es diferente de las otras.

Sustitución de octetos
En inglés: *Byte substitution*

Operación que es parte de las iteraciones aplicadas en el cifrado **AEC**-2001; cada octeto de dígitos binarios (bits) que es un elemento de la **matriz de estado,** se reemplaza con otro octeto mediante un mapeo definido por la **Tabla de substituciones.**

Tabla de substituciones
En inglés: *S-Box*

Tabla, o matriz. preestablecida donde se encuentran los elementos nuevos para la operación **Sustitución de octetos.** La ubicación

	del elemento sustituto dentro de la tabla se efectúa partiendo el octeto original en dos cuartetos, uno de los cuales indica el renglón y el otro la columna.
Transmisor-Receptor Asíncrono Universal (**TRAU).** En inglés: *Universal Asynchronous Receiver Transmitter* (***UART***)	Es un integrado para transmitir y recibir datos en forma asíncrona. Transforma datos en paralelo a una forma serial.
Troncal	Una línea de comunicación entre 2 equipos y/o de conmutación. Actualmente, muchos de éstos son enlaces de transmisión digital, y con frecuencia implantados en fibra óptica.
Voz (Significado coloquial):	Sonido producido por los «órganos del habla *(anatomía)*».
VOZ (Significado en la Fonética y en Ingeniería Electrónica):	Sonido pronunciado *("uttered")* con resonancia de las cuerdas vocales.

REFERENCIAS.

[1] Newton, Harry (2001). Newton`s Telecom Dictionary. New York, EUA: CMP Book.

[2] Tanenbaum, Andrew S. (2003). Redes de Computadoras. México: Pearson Educación. Pp 118-121.

[3] Tomasi (2000). Sistemas Electrónicos. México: Prentice Hall. Pag. 119.

[4] Mahmoud, Harb (1989). Modern Telephony. New Jersey, EUA: Prentice Hall. pp. 18-19.

[5] Bigelow, J. Stephen (1991). Understanding Telephone Electronics.. Indianapolis E.U.: SAMS.pp 8-284.

[6] Haykin Simon (2002). Sistemas de Comunicación. (Traducido por Gabriel Nagore Cazares). México: Limusa Wiley. (Original publicado en 2001.)

[7] A. Menezes, P. van Oorschot, y S. Vanstone, (1996). Handbook of Applied Cryptography. : CRC Press, 1996.

[8] Haykin Simon (2002). Sistemas de Comunicación. (Traducido por Gabriel Nagore Cazares). México: Limusa Wiley. (Original publicado en 2001.) p. 144

[9] Daemen Joan, Rijmen Vincent (2002). The Designed Rijndael. Berlin: Springer-Verlag. Pp.41, 32-34, 210.

[10] Hanzo, L., Somerville F.C. (2001). Voice Compression and Comunications.. New York: Wiley. Pp. 6-11, 207-208

[11] Chu, Wai C. (2003). Speech Coding Algorithms. San Jose California, E.U.: Wiley.Intercience..pp 299-308.

[12] Freescale Semiconductor, Inc. (2004). DSP56858 Evaluation Module User's Manual. Denver, Colorado: p. (1-2).

[13] Freescale Semiconductor, Inc. (2004). DSP56858 16-bit Digital Signal Processor. Denver, Colorado: pp. 1-3.

[14] Dallas Semiconductor (2004). ±15kV ESD-Protected, 1µA, 1Mbps, 3.0V to 5.5V,. Sunnyvale, CA: Dallas Semiconductor Maxim. p. 17

[15] Freescale Semiconductor, Inc. (2004).DSP5685X User Manual Rev 3.0 Denver, Colorado: p. (12-4).

[16] C. E. Shannon. Communication Theory of Secrecy Systems.

[17] José de Jesús Angel Angel. AES - Advanced Encryption Standard. 2005.

[18] Joan Daemen, Vincent Rijmen. AES Proposal: Rijndael. Document version 2 1999.

[19] León-García Alberto, Fundamental Concepts and Key Architectures. Edit. McGraw-Hill Professional, 2004.