

EDITORIAL

Dicen que el 13 es mal número si no crece, pero nosotros hemos superado esa supersticiosa barrera, llevando a vuestras manos éste que es nuestro décimocuarto número y que esperamos sea de vuestro agrado.

Esperamos que este número suponga una nueva etapa de Buran, por dos razones. En primer lugar, tras la celebración del 1er. Congreso Estatal de Ramas del IEEE, que celebramos el pasado mes de Abril en Barcelona, Buran se refuerza en su papel como nexo de unión entre las Ramas del IEEE, y, a buen seguro, veremos incrementada la variedad en la procedencia de los artículos publicados cumpliendo mejor de este modo los objetivos de nuestra revista, formar y comunicar.

Todos aquellos que, de un modo u otro, colaboramos en Buran tenemos un motivo añadido de alegría y es la realización de un viejo sueño: que nuestra revista cruce fronteras y llegue a las universidades de Latinoamérica. Desde estas líneas damos la bienvenida a los nuevos lectores y, cómo no, esperamos recibir pronto colaboraciones del otro lado del Atlántico para la única publicación del IEEE que utiliza el castellano como lengua principal.

La presencia de estos nuevos lectores y colaboradores, debe constituir también un acicate para aquellos que buscan un eco en sus investigaciones porque, no podemos olvidarlo, la comunidad científica ya era global antes de que naciera Internet y hemos de aprovechar las grandes ventajas que nos aporta la «aldea global», para que el intercambio científico, tecnológico y cultural entre nuestros países sea lo más fructífero posible.

Por último, aprovechamos esta oportunidad para animaros a que participéis activamente en las actividades de las asociaciones estudiantiles, que, como esta revista, se realizan con la ilusión y esfuerzo de unos cuantos «afortunados». Por nuestra parte, en las Ramas de Estudiantes del IEEE estamos abiertos a vuestras sugerencias y proyectos, y estamos convencidos de que cuantos más y más activos seamos los estudiantes en este sentido, más grandes cosas se pueden conseguir. Sirvan de ejemplo nuestros compañeros de Madrid, galardonados con el premio AT&T por su proyecto (os referimos a su artículo en estas páginas). Enhora-buena desde aquí por el éxito.

Sed fieles.

*Miguel Ángel Sastre, Presidente.
Daniel Prado, Tesorero.*

EDICIÓN

BARCELONA

José A. López Salcedo
Xavier Palau Marqués
Daniel Prado Rodríguez
Carles Ruiz Floriach
Jorge Sáiz Fernández
Jose Castor Vallés Martínez

REVISIÓN

José A. López Salcedo
Xavier Palau Marqués
Daniel Prado Rodríguez
Carles Ruiz Floriach
Jorge Sáiz Fernández
Jose Castor Vallés Martínez

COORDINACIÓN

BARCELONA

Jorge Sáiz Fernández

AGRADECIMIENTOS

II. Dir. Antoni Elias Fusté, Elisa Pla,
Ángel Cardama, Pere Camps, AESE Estudiantes
y a los puntos de distribución en la UPC:
Abacus, CPET, CPDA, Kiosk Campus Nord
y Reprografía Sant Just.

IMPRESIÓN

RET, s.a.l.

FOTOMECAÑICA

Sistemes d'Edició

DEPÓSITO LEGAL

B-19.950-96

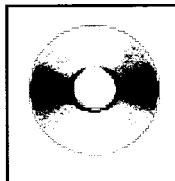
SIP, 1999 (594)

En la portada, una imagen del tornado Floyd acercándose a la costa de Florida. Fotografía obtenida en la web de la NASA:

<http://rsd.gsfc.nasa.gov/rsd/images/Floyd>

La organización se reserva el derecho de publicar los artículos. La opinión expresada en los artículos no tiene por qué coincidir con la de la organización.

Agradecemos las colaboraciones hechas desinteresadamente, y a causa de la falta de espacio, pedimos disculpas a todas aquellas personas a las cuales no se les ha publicado su colaboración. Esperamos que en un próximo número tengan cabida.



MP3

SONIDO DIGITAL AL ALCANCE DE TODOS

David Rincón Rivera

drincon@mat.upc.es

Departament de Matemàtica Aplicada i Telemàtica

Universitat Politècnica de Catalunya (UPC)

INTRODUCCIÓN

El formato de compresión de sonido *MPEG-1 Layer 3* (también conocido como MP3) está adquiriendo una gran notoriedad, debido al revuelo que está provocando en el campo de la grabación y distribución de audio digital. En los últimos meses se han producido repetidos intentos de las compañías discográficas para limitar sus posibilidades de grabación y reproducción, ya que Internet se ha convertido en una vía de distribución paralela a las habituales (y en muchas ocasiones, ilegal).

En éste artículo se va a describir el formato MP3 desde dos puntos de vista: el del técnico y el del usuario. Desde el punto de vista técnico haremos hincapié en los algoritmos de compresión de sonido utilizados por el estándar, así como algunos comentarios sobre una de las aplicaciones más prometedoras de MP3, que es su transmisión a través de redes de conmutación de paquetes. La segunda parte del artículo estará dedicada a proporcionar información y herramientas para experimentar en nuestro ordenador la calidad de éste estándar.

INTRODUCCIÓN AL AUDIO DIGITAL

Como ya sabemos, para digitalizar una señal lo único que necesitamos es disponer de un convertidor analógico/digital (A/D), que se compone de un módulo de muestreo y un codificador.

El papel del primer bloque, también conocido como *sample & hold*, es discretizar la señal en el tiempo. Para asegurar una correcta reconstrucción de la señal, el teorema de Nyquist nos obliga a que la frecuencia de muestreo sea mayor o igual al doble del ancho de banda de la señal original.

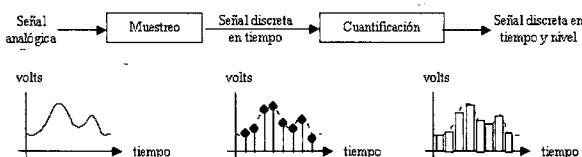


Figura 1. Proceso de muestreo y cuantificación.

El segundo bloque cuantifica los valores obtenidos por el primero, con un cierto número de bits por muestra.

El sistema de audio digital más sencillo es el PCM (*Pulse Code Modulation*), que se limita a cuantificar de manera uniforme la señal muestreada. La cantidad de bits (n) empleados en la cuantificación determinan la relación señal a ruido (SNR, *signal to noise ratio*) del proceso de digitalización, de manera que cada bit adicional añade 6 dB [1].

$$\text{SNR}_q (\text{dB}) = \text{constante} + 6n \text{ dB}$$

Vamos a comentar dos ejemplos de audio digital PCM: telefonía digital y el Compact Disc. En el primero, digitalizamos una señal de 3.1 Khz a un ritmo de 8000 muestras por segundo, con 8 bits/muestra (SNR » 50 dB), lo que genera una tasa de 64 Kbit/s. El CD muestrea una señal de alta fidelidad (20 Khz) a 44100 muestras/segundo y 16 bits/muestra (SNR » 100 dB), en dos canales (stereo), generando una tasa total de 1.4 Mbit/s. Como vemos, esta tasa es elevadísima, y requiere un gran ancho de banda para ser transmitida. Además, el PCM suele ser una codificación muy ineficiente, ya que cada muestra es muy parecida a la anterior, con lo que tenemos una redundancia muy alta.

Por ello se diseñaron algoritmos de compresión basados en la predicción temporal de las muestras, como el DPCM o la modulación Delta, que se basan en la codificación de la diferencia entre la muestra real y la predicha por el sistema en base a las muestras anteriores. La predicción se hace a partir de unos coeficientes que pueden ser fijos o variables, de manera que se adapten a los cambios de la señal de entrada, haciendo que la señal reconstruida sea más fiel a la inicial (un ejemplo de estos sistemas es el ADPCM).

Sin embargo, aunque estos sistemas explotan con éxito la eliminación de la redundancia, sólo son capaces de reducir la tasa en un factor de entre 2 y 4 (se considera que un ADPCM a 32 Kbit/s ofrece una calidad ligeramente superior al PCM de 64 Kbit/s [1]). Por ello se hizo patente la necesidad de crear nuevos esquemas de compresión que explotaran otro tipo de propiedades. Así en el campo de la telefonía aparecieron los detectores de silencio, que eliminan la transmisión cuando la señal es tan baja que no va a ser captada por el oído humano, o los sistemas llamados *vocoder*, que intentan reproducir las características del tracto vocal humano (cuerdas vocales, boca, lengua...) para analizar y sintetizar digitalmente las for-

mas de onda que salen de nuestra garganta. Pero estos sistemas son óptimos cuando son utilizados para codificar voz humana y no música, que es mucho más rica en matices y que contiene más información.

Por ello se inició otra línea de investigación basada en el otro extremo de la comunicación, en el receptor: el oído humano. A estas técnicas se les llama «psicoacústicas» o de «codificación perceptual», porque se basan en las propiedades de nuestro sistema auditivo para comprimir la información acústica a tasas inimaginables hasta el momento. Hay sonidos que no podemos oír, así que podemos eliminarlos y ahorrar una gran cantidad de información. Vamos a describir con más detalle cuáles son las características del sistema auditivo humano, y cómo podemos aprovecharlas en el proceso de compresión.

¿QUÉ OÍMOS Y QUÉ NO PODEMOS OÍR? EFECTOS PSICOACÚSTICOS

El rango frecuencial en el que el oído humano es capaz de detectar sonido está comprendido entre los 20 Hz y los 20 Khz, con una zona especialmente sensible entre los 2 y los 4 Khz, muy cercana al espectro de la voz, situada entre los 500 hz y los 4 Khz [2]. Por tanto, dos tonos de potencia similar situados en los 3 y los 15 Khz serán percibidos de manera muy diferente (el de 15 Khz pasará mucho más desapercibido, pudiendo ser incluso inaudible). Es decir, nuestro oído no ofrece una respuesta plana con la frecuencia, sino que premia a unas bandas y penaliza otras, llegando al extremo de pasar desapercibidas. Por ello se define el **umbral absoluto de audición** como la frontera entre los sonidos que son perceptibles y los inaudibles. El concepto fue acuñado por Fletcher en 1940, durante una serie de experimentos donde se obtuvieron gráficas como la presentada en la figura 2, que presenta el umbral de audición en función de la frecuencia. El gráfico se obtuvo por métodos empíricos, efectuando un muestreo estadístico entre la población. Los valores se establecen respecto a un tono puro de 1 Khz con una potencia tal que se encuentra en el límite de audición.

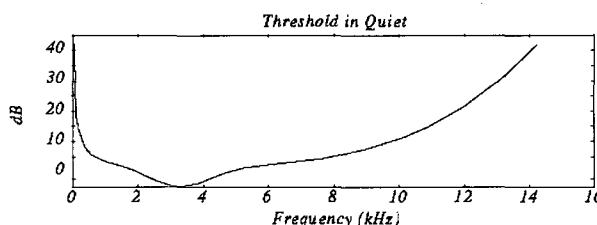


Figura 2. Umbral de audición del oído humano [3].

Según esta gráfica, toda señal que quede por debajo de la curva será inaudible para prácticamente todas las personas, así que no tiene sentido codificarla y puede ser eliminada.

Sin embargo, este umbral de audición no es único. La aparición de señales adicionales puede modificar nuestra percepción de un cierto tono, llegando incluso a producir **enmascarados frecuenciales** (*frequency masking*). En la figura 3 se presenta un ejemplo. Supongamos que disponemos de un tono de 1 Khz a un nivel fijo (60 dB por encima del umbral de audición), que llamaremos «tono enmascarador». Generamos otro tono de 1.1 Khz y medimos el nivel de potencia al que se hace indistinguible. Si repetimos el proceso para toda la banda, obtendremos una segunda curva de umbral, esta vez generada por el tono enmascarador. Como en el caso anterior, toda señal que quede por debajo de esta curva será inaudible (se dice que ha sido enmascarada por el tono dominante). Como es lógico, cuanto más cercano esté la señal al tono enmascarador y menor sea su potencia, más posibilidades de que sea enmascarado.

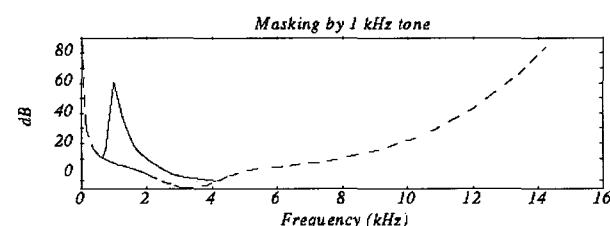


Figura 3. Enmascaramiento causado por un tono de 1 Khz [3].

En la figura 4 se muestra la forma de la campana de enmascaramiento para diversas frecuencias. La característica más destacable es que a medida que crece la frecuencia, más ancha se hace la campana.

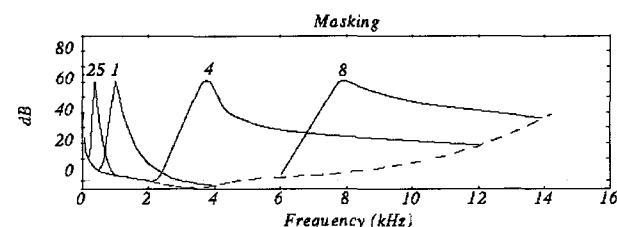


Figura 4. Enmascaramiento frecuencial [3].

Podemos observar que el ancho de la campana de enmascaramiento sigue una regla aproximadamente logarítmica. Por ello parece natural crear el concepto de **banda crítica** (*critical band*), que se define como cada una de las porciones del espectro en la que el oído percibe una señal uniforme [3]. Para medir estas porciones se crea la unidad llamada *bark* (en honor a Barkhausen) para denominar al ancho de banda correspondiente a una banda crítica. A partir del estudio empírico de las bandas críticas, se llegó a la conclusión de que se podían calcular de la siguiente manera:

- Para frecuencias inferiores a 500 Hz, 1 bark » $f/100$
- Para frecuencias superiores a 500 Hz, 1 bark » $9 + 4\log(f/1000)$



Con este convenio se consiguen gráficas como la presentada en la figura 5. Aquí podemos observar cómo la introducción de unidades logarítmicas como los *bark* permiten subdividir el espectro en bandas de tamaño uniforme. Como veremos más adelante, este detalle es importante cuando hay que efectuar un análisis subbanda.

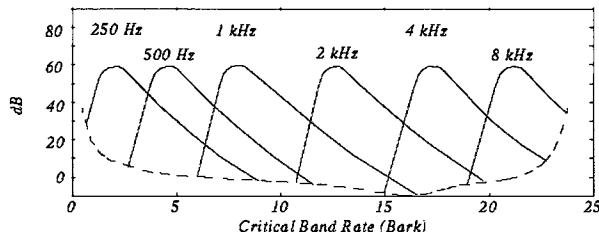


Figura 5. Bandas críticas, representadas en función de los barks [3].

Hasta ahora nos hemos limitado a medir los fenómenos de enmascarado en el dominio frecuencial, pero también se producen **enmascarados temporales**. Un tono muy potente enmascarará durante un cierto intervalo temporal cualquier otro tono de frecuencia parecida y que sea temporalmente cercano. En la figura 6 se presenta el caso que exponemos a continuación. Tenemos un tono enmascarador de 1 KHz y 60 dB, y otro tono de 1.1 KHz y 40 dB, que está enmascarado. En $t = 0$, desactivamos el tono enmascarador y medimos cuánto tarda el oído en percibir el segundo tono (se puede hacer desconectándolo en $t = Dt$, y disminuir Dt hasta que se deja de percibir). Si repetimos el experimento para diferentes potencias, obtenemos una respuesta como la de la figura 6. Cuanto menos potente sea el tono enmascarado, más tarda el oído humano en recuperarse de la saturación que le ha provocado el tono enmascarador. Este es un efecto que cualquiera de nosotros ha experimentado: después de escuchar un sonido fuerte, como por ejemplo una explosión, nos quedamos momentáneamente sordos y necesitamos un poco de tiempo para recuperar la agudeza auditiva habitual.

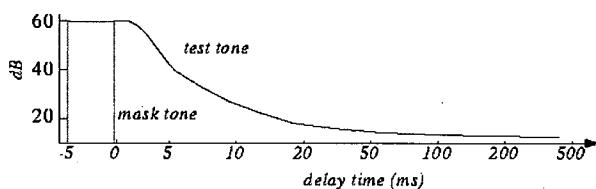


Figura 6. Enmascaramiento temporal [3].

Si repetimos este experimento para diferentes potencias y frecuencias, obtendremos una gráfica como la de la figura 7, donde podemos observar el efecto combinado de los enmascaramientos frecuenciales y temporales. Curiosamente se puede ver que existe el fenómeno del «enmascaramiento previo»: hay sonidos que son enmascarados antes de que se genere el tono enmascarador. Esto no es un error; parece ser que nuestro oído necesita un cierto tiempo antes

de poder identificar un tono. Si en este tiempo se produce el tono enmascarador, el tono enmascarado no será percibido en absoluto.

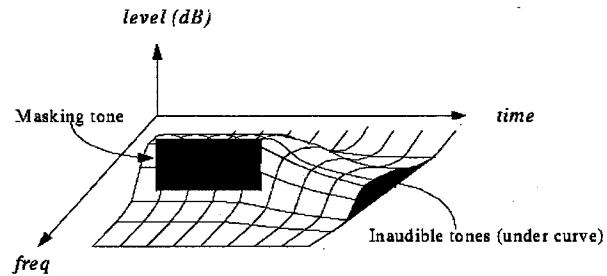


Figura 7. Efecto combinado del enmascaramiento [3].

¿CÓMO PODEMOS APROVECHAR LOS EFECTOS PSICOACÚSTICOS?

Como ya hemos comentado, la gran ventaja de los modelos psicoacústicos es que toda señal que quede por debajo del umbral total de enmascarado (la curva combinada del umbral absoluto, del enmascarado frecuencial y del temporal) es inaudible y, por tanto, no se codifica. También se ha visto que las bandas críticas son las unidades naturales en las que podemos dividir, de manera uniforme, la influencia del enmascarado dentro del espectro. Por tanto parece natural que el proceso de compresión utilice codificación sub-banda (*subband coding*), consistente en separar la señal de entrada en un cierto número de bandas y hacer un análisis independiente de cada una de ellas. Esto se puede conseguir mediante un banco de filtros.

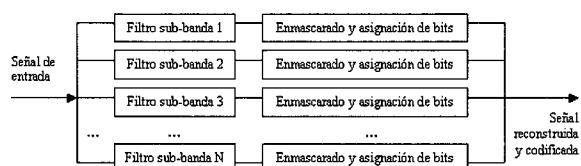


Figura 8. Esquema del análisis sub-banda.

A continuación, a cada sub-banda se le aplica un análisis psicoacústico o perceptual que determina las curvas de enmascaramado, y determina cuáles son los tonos que se encuentran por encima de dichas curvas (los componentes que están por debajo son simplemente descartados). Así se genera la relación señal a máscara (SMR, *signal-to-mask ratio*). A los tonos supervivientes se les asigna una cierta cantidad de bits en función del ruido que podemos introducir, calculado a partir del enmascarado, la SMR y las necesidades de tasa del sistema global. El criterio para asignar bits a la señal es sencillo: la potencia del ruido debe quedar por debajo del umbral de audición. Veamos un ejemplo: supongamos que en una cierta sub-banda con un nivel de enmascarado de 26 dB existe una señal superviviente con una potencia de 40 dB. Dicha señal debería ser codificada con un mínimo de 7 bits (7 bits

$x 6 \text{ dB/bit} = 42 \text{ dB} > 40 \text{ dB}$). Sin embargo podemos ahorrar 4 bits, ya que el ruido los enmascara ($4 \text{ bits} \times 6 \text{ dB/bit} = 24 \text{ dB} < 26 \text{ dB}$). Por tanto, nos basta con 3 bits para codificar la señal superviviente.

Mediante este método combinado de eliminación de señales enmascaradas y reducción de la información correspondiente a las supervivientes se pueden conseguir factores de compresión muy elevados, sin comprometer la calidad (aparente) del sonido. En muchos casos se eliminan sub-bandas enteras que han sido enmascaradas por tonos situados en la banda vecina, haciendo que no se utilice ni un solo bit para codificar la sub-banda en cuestión.

LOS ESTÁNDARES MPEG

MP3 es el nombre con el que se conoce a una de las partes del estándar de codificación de vídeo MPEG (Moving Picture Experts Group) [4] de la Organización Internacional de Estándares (ISO) [5]. Concretamente, MP3 hace referencia a la capa 3 del codificador de audio de MPEG-1, así que lo mejor será empezar describiendo el estándar MPEG.

MPEG-1 fue el primer estándar internacional de codificación de vídeo creado por la ISO que aplicó técnicas de compresión basadas en el enmascaramiento de la información visual y acústica; es decir, lo que el usuario no va a ver ni oír, no se codifica. Esta técnica ya fue aplicada con gran éxito por el comité JPEG (Joint Photograph Experts Group), que dio origen al formato de compresión de imágenes del mismo nombre, y que se ha convertido en el estándar de facto en Internet. Fruto de este éxito, la ISO formó el grupo MPEG a finales de los 80 para crear diversos estándares de vídeo digital de alta calidad. El plan inicial era crear cuatro versiones diferentes, cada una de ellas destinadas a un segmento específico de usuarios y aplicaciones [6]:

- MPEG-1:** Codificador de vídeo a 1.5 Mbit/s con calidad de videoconferencia mejorada, de resolución 352x288 pixels (CIF) o superior. Apareció como la evolución natural del estándar de videoconferencia ITU-T H.261, con mejoras relacionadas con la compensación de movimiento y la predicción temporal.
- MPEG-2:** Codificador de vídeo a tasa de 4-10 Mbit/s con calidad de emisión de TV («broadcast») comparable a los sistemas analógicos PAL, SECAM y NTSC. Destinado a ser el estándar de emisión de TV digital de consumo masivo.
- MPEG-3:** Codificador de vídeo a tasas superiores a los 10 Mbit/s, con calidad de TV de alta definición (HDTV), destinado a ser usado en centros de producción y en redes de transmisión.

MPEG-4: Codificador de videoconferencia a tasa muy bajas (64 – 256 Kbit/s) para ser usado sobre redes de banda estrecha, especialmente de telefonía móvil.

De estos cuatro estándares, sólo 3 han visto la luz (MPEG-1, 2 y 4). MPEG-3 se quedó por el camino, ya que los algoritmos desarrollados para MPEG-2 son tan potentes y flexibles que permiten abarcar tanto la calidad *broadcast* como la de producción, simplemente variando la tasa a la que funciona el codificador.

Aunque el esfuerzo más grande de los ingenieros se dedicó a los algoritmos de codificación de vídeo, no se descuidó el sonido que tenía que acompañar a las imágenes. Uno de los handicaps con los que tuvieron que luchar los diseñadores de MPEG fue el requisito de escalabilidad, que consiste en que se debe permitir que equipos de gamas diferentes puedan reproducir el mismo flujo de información, aunque sea a calidades diferentes. Para ello se definió una arquitectura de tres capas, en la que cada «layer» se basa en un codificador más sofisticado que el de la capa anterior. Así, tenemos MPEG-1 Layer 1, Layer 2 y Layer 3, siendo esta última la más complicada y la más eficiente desde el punto de vista de compresión. Los reproductores de capa 3 son capaces de reproducir flujos codificados con cualquiera de las tres técnicas, mientras que los de capa 1 sólo pueden reproducir información

	Tasa objetivo	Factor de compresión	Calidad a 64 Kbit/s	Calidad a 128 Kbit/s	Retardo teórico de compresión
Capa 1	192 Kbit/s	4:1			19 ms
Capa 2	128 Kbit/s	6:1	2.1 a 2.6	> 4	35 ms
Capa 3	64 Kbit/s	12:1	3.6 a 3.8	> 4	59 ms

Tabla 1. Características de las capas de audio de MPEG-1.

codificada según la capa 1. Las características de cada capa son las siguientes [3]:

La tasa objetivo es el *bitrate* para el que se diseñó cada una de las capas. El factor de compresión nos da la relación entre la tasa generada por el codificador MPEG y la que se necesitaría en PCM para conseguir una calidad equivalente. Vemos que se consiguen factores de hasta 12, lo que nos da una idea de la potencia del algoritmo.

Los otros dos apartados interesantes de la tabla 1 son la valoración de la calidad subjetiva a las tasas de 64 y 128 Kbit/s. Esta medida de calidad se realiza basándose en el criterio MOS (*Mean Opinion Square*), definido por la ITU (Unión Internacional de Telecomunicaciones). Consiste en hacer un análisis estadístico de la calidad percibida por grupos de personas escogidas al azar en diferentes países, que hacen una valoración subjetiva de la calidad de los tests y pruebas presentadas. La escala MOS tiene un rango comprendido entre 1 (ininteligible) y 5



(perfecto). Vemos que incluso a tasas muy bajas, las capas 2 y 3 obtienen valoraciones muy buenas. Una anécdota surgida durante el período de pruebas: parece que uno de los pocos casos en los que el algoritmo no funciona con la calidad adecuada es en la codificación de voz masculina alemana, aunque esto se puede solucionar elevando la tasa del flujo [6].

Existen cuatro modos de funcionamiento para cada una de las capas: canal único (una sola señal de audio en el flujo), canal doble (dos canales separados e independientes), stereo (igual al anterior pero con dos señales pertenecientes a los canales derecho e izquierdo de una señal stereo original), y joint stereo (parecido al anterior pero explota la redundancia entre los dos canales para reducir aún más la tasa).

CODIFICACIÓN Y DECODIFICACIÓN DE AUDIO MPEG

Los estándares MPEG son del tipo denominado **asimétrico**, en el que los codificadores son mucho más complejos que los decodificadores. Esto es así para permitir la comercialización de reproductores baratos, destinados al mercado de electrónica de consumo. Veremos que los codificadores soportan una carga computacional muy superior a la de los decodificadores.

Otro detalle importante desde el punto de vista del implementador es que **no se especifica un estándar de codificación**. Lo que sí existe es una especificación de qué tipo de flujos de bits es capaz de reproducir un cierto «decodificador modelo», y una serie de recomendaciones sobre cómo puede construirse un codificador. Esto permite fomentar el desarrollo de algoritmos de codificación diferentes para cada fabricante (que puede así diferenciarse de sus competidores, promover la investigación y preservar sus patentes), manteniendo al mismo tiempo la compatibilidad (ya que todos los codificadores han de ser compatibles con cualquier reproductor que siga el modelo especificado).

La fuente sobre la cual se aplican los algoritmos debe ser siempre una señal PCM a las frecuencias de muestreo de 32 KHz, 44.1 KHz (propia del Compact Disc) y 48 KHz (propia del sistema DAT), con 16 bits por muestra (unos 100 dB de relación señal a ruido de cuantificación).

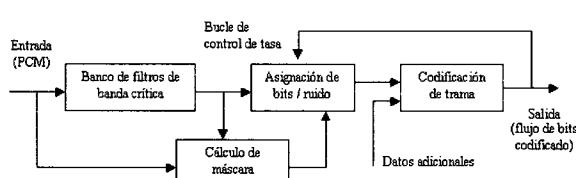


Figura 9. Esquema general del proceso de codificación de audio MPEG-1.

En un apartado anterior hemos comentado cómo se podían aprovechar los efectos psicoacústicos para comprimir (con pérdidas) la información correspondiente a una fuente de sonido. El esquema general utilizado por las tres capas MPEG es el presentado en la figura 9.

El proceso de codificación se realiza sobre el flujo continuo de bits de entrada. Sin embargo, para facilitar la compresión y permitir su segmentación temporal, se define la **trama (frame)** como el bloque unitario mínimo que puede ser decodificado completamente. Las tramas se componen de una cierta cantidad de muestras de entrada, que varía según la capa utilizada.

Podemos observar que la funcionalidad básica es la ya mostrada en la figura 8: el banco de filtros que separa la señal en subbandas críticas, el módulo que evalúa los efectos psicoacústicos y elimina las señales que quedan por debajo de la curva, y el bloque que asigna bits a las señales supervivientes en función del nivel de ruido enmascarado y de las necesidades de tasa instantánea (puede observarse la realimentación desde la salida, para el caso en que se exija una tasa constante). Finalmente encontramos un bloque que se dedica a formatear el flujo de bits de salida, con funciones como controlar la tasa (constante o variable), comprimir aún más los datos mediante algoritmos como Huffman o Ziv-Lempel, segmentar el flujo de salida en tramas, añadir marcas temporales para su correcta reproducción en el decodificador, introducir un canal de datos adicionales con información sobre el autor de la música, códigos de acceso, etc.

Como ya comentamos anteriormente, el esquema del decodificador es mucho más sencillo que el codificador. Se limita a extraer la información formateada de las subbandas, reconstruirlas por separado (con las curvas de enmascaramiento) y finalmente, a unir todas las bandas para formar la señal reconstruida.

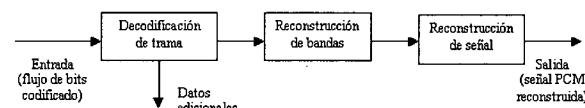


Figura 10. Esquema general del proceso de decodificación de audio MPEG.

Vamos a comentar brevemente las características de cada una de las capas [2, 6, 7, 8]:

MPEG-1 Audio layer 1:

- Segmentación de la señal en 32 subbandas a partir de un filtro polifase de baja complejidad.
- Análisis frecuencial mediante FFT de 512 puntos.
- Cálculo de la SMR a partir del tono dominante en cada subbanda.

- Modelo psicoacústico que sólo tiene en cuenta el enmascarado frecuencial.
- Cuantización fija del máximo valor absoluto de todas las subbandas con 6 bits. Este valor actúa como escala para el resto de bandas.
- Formato de trama: 384 bits, correspondientes a 12 muestras de 32 subbandas (a 48 KHz, equivalentes a 8 ms de sonido).
- Sincronización y ordenación: 16 bits
- Corrección de errores (CRC): 16 bits



- Para cada banda: 6 bits de factor de escala + 4 bits de cuantización del valor dominante.
- Información adicional: 32 bits

MPEG-1 Audio layer 2:

Esta capa es muy similar a la anterior, pero consigue tasas de compresión mayores (a costa de incrementar el retardo). Las diferencias principales respecto a la capa 1 son las siguientes:

- FFT de 1024 puntos (mayor resoluciónpectral).
- 3 valores de escala para cada subbanda en cada trama. Se pueden usar unos, dos o los tres.
- Cuantización variable de cada subbanda, en función del factor de escala.
- El modelo psicoacústico incluye la influencia del enmascarado temporal (no completamente)
- Aumento de la resolución de los cuantizadores a 16 bits.
- Tramas de 1152 bits (equivalentes a 3 de las tramas de capa 1), correspondientes a 36 muestras de 32 subbandas (a 48 KHz, equivalentes a 24 ms de sonido).

MPEG-1 Audio layer 3:

La capa 3 es, con diferencia, la más complicada y sofisticada de toda la familia MPEG. Introduce mejoras sustanciales en casi todos los pasos de la codificación, como:

- Uso de un banco de filtros híbridos conmutados (DCT/polifase) que proporciona resolución variable, ajustándose mejor a las bandas críticas.
- El modelo psicoacústico explota en su totalidad el enmascarado temporal, e incluye una predicción polinómica.
- Aprovecha la redundancia inherente en las señales estereofónicas.

Calidad de sonido	Ancho de banda	Modo	Tasa	Factor de compresión
Sonido telefónico	2.5 KHz	Mono	8 Kbit/s	96:1
≈ Radio AM	4.5 KHz	Mono	16 Kbit/s	48:1
Mejor que AM	7.5 KHz	Mono	32 Kbit/s	24:1
≈ Radio FM	11 KHz	Stereo	56-64 Kbit/s	26-24:1
Cercana a CD	15 KHz	Stereo	96 Kbit/s	16:1
Calidad CD	22.1 KHz	Stereo	112-128 Kbit/s	14-12:1

Tabla 2. Comparación de las diversas calidades de la capa 3 de MPEG-1.

- Cuantización variable con control dinámico del ruido.
- Codificación de trama mediante códigos entrópicos de Huffman y RLE (Run-length encoding).
- Formato de trama: 1152 bits. La trama ya no es la unidad de decodificación, sino que necesita de información codificada (almacenada o transmitida) previamente. Orientado a la transmisión en flujo.

La calidad ofrecida por la capa 3 se compara con la de otros sistemas en la Tabla 2 [9].

MP3 Y EL MUNDO DE LA MÚSICA

Hasta aquí hemos descrito el estándar MP3 desde el punto de vista técnico. ¿Cuál ha sido la razón de su éxito? Lo que en principio era solamente una curiosidad técnica se ha ido convirtiendo en una ola que hace que las discográficas empiecen a preocuparse. ¿Por qué?

La razón es que se ha dado la convergencia de tres factores: la popularización de los ordenadores con capacidad multimedia (especialmente las grabadoras de CD-ROM), el fenómeno del software de libre distribución, y el *boom* de Internet.

La gran novedad del MP3 es que es el primer estándar con el que se puede adquirir y, sobretodo, copiar audio de alta calidad, usando un simple PC. ¿Y cómo se puede hacer eso? Con uno de los cientos de programas de software gratuito que se pueden encontrar tanto en Internet como en forma de regalo por parte de revistas de informática. Básicamente hay tres tipos de programas: los *rippers* (programas que leen las pistas del CD y las convierten a ficheros WAV codificados en PCM a 44.1 KHz),



codificadores (compresores que a partir del WAV aplican los algoritmos y generan un fichero MP3) y decodificadores (reproductores de MP3). Se pueden encontrar versiones comerciales mucho mejor acabadas y con gran cantidad de prestaciones, pero los más populares son programas muy sencillos creados por programadores sin afán comercial, que ceden su producto a toda la comunidad de usuarios (*freeware*).

Un usuario con un equipo medio puede comprimir canciones desde el CD sin ningún tipo de problemas: le basta con poner en marcha el *ripper* y a continuación utilizar el codificador. Normalmente el proceso de compresión no se puede hacer a tiempo real; una canción de tres minutos puede tardar entre diez y quince en ser comprimida, con un Pentium de primera generación. Es necesario un equipo bastante más potente (un Pentium II) para poder generar MP3 en tiempo real. Y para reproducir basta con un equipo de gama baja con una tarjeta de sonido. Esto es posible debido a la característica asimétrica del estándar: el decodificador es mucho más sencillo que el codificador. Así, un simple Pentium a 133 Mhz es capaz de reproducir sin problemas un fichero o un flujo MP3 (siempre que sea el único proceso que corre).

Tal como se comentó en un apartado anterior, se puede conseguir una calidad propia del CD con un factor de compresión que se aproxima a un valor de 10-12. Es el caso de la capa 3 a 44.1 Khz, en modo «*joint stereo*» y una tasa de 128 Kbit/s. Con estos parámetros se consigue que una canción de 4 minutos ocupe menos de 4 Mbytes, lo cual posibilita que en un CD-ROM de 600 Mbytes quepan más de 12 horas de música en MP3, o que sea factible enviar canciones por correo electrónico, o bien capturarlas desde Internet.

Y esto es lo que preocupa a las discográficas, la posibilidad de transportar fácilmente la información, ya sea en CD o a través de la red. Por un lado se está produciendo un fenómeno de piratería (ya es posible tener en un solo CD toda la discografía de un artista); por otro lado, están perdiendo el mercado de la distribución (hay artistas noveles que editan sus trabajos en la red, e incluso algún cantante consagrado como David Bowie que distribuye canciones en su servidor web [10]). Están apareciendo multitud de portales y buscadores especializados en música MP3 [11] que son una auténtica mina para los «piratas musicales», ya que la estructura de Internet hace posible establecer servidores en países donde las discográficas no pueden actuar legalmente en su contra. Es por ello que estas compañías están promoviendo diferentes estándares de compresión que incorporan protección contra copias, pero todavía no está claro que consigan imponerlos en el mercado y, sobre todo, en la red.

La batalla continúa. Hay quien cree que es una lucha entre piratas y empresas, y hay quien piensa que es una pugna por la democratización del acceso a la música, acorde con la filosofía original de Internet (acceso univer-

sal a la información). Como siempre, depende del lado desde el que se mire...

PARA EMPEZAR A JUGAR

En este apartado vamos a ofrecer información práctica sobre programas y productos que nos permitirán experimentar con el sonido MP3. Comentaremos cuáles son los más populares y dónde encontrarlos.



Figura 11. Reproductor WinAmp, con la piel original.

El primer contacto con el mundo MP3 suele ser a través de un reproductor. El más popular es el WinAmp (<http://www.winamp.com>) de la compañía NullSoft Inc. Es un programa actualmente *freeware* que decodifica tanto ficheros como flujos HTTP y RTP (con un *plug-in*) de los formatos MP3, CD-Audio, WMA, MOD y WAV, entre otros. Una de las características más curiosas es su capacidad de cambiar de piel (*skin*). Existen versiones con motivos de Star Trek, el OS de Apple, o el interfaz X-Window de Unix. Los fans más incondicionales pueden diseñar su propia versión y donarla al resto de usuarios.



Figura 12. Reproductor FreeAmp, en su versión Windows.

Otro reproductor, algo menos popular, es el FreeAmp (<http://www.freeamp.org>), del cual está disponible el ejecutable y el código, compilable tanto para Windows (Visual C++) como para Linux. Este decodificador permite recibir flujos HTTP/RTP/Multicast sin ningún tipo de *plug-in* adicional. El código es ideal para estudiar a fondo el funcionamiento del MP3, al menos desde el lado del reproductor.

Si queremos generar nuestros propios ficheros MP3, necesitaremos *rippers* (para fuentes de CD-Audio) y codificadores. Aunque se pueden encontrar por separado, lo habitual es que las dos funciones estén juntas en el mismo programa. Es el caso de AudioGrabber (<http://www.audiograbber.com-us.net>) y AudioCatalyst (<http://audiocatalyst.com>). Los dos permiten seleccionar opciones de adquisición (velocidad del CD, protección contra errores, drivers, etc.) y compresión (pre-normalización de la señal, modo stereo/mono, frecuencia de muestreo, tasa

de bits, preénfasis, etc.). De hecho, estos programas suelen ser simples *front-ends*, interfaces que controlan el motor de compresión. Los motores más habituales son el L3encoder, del instituto Fraunhofer [9], LAME y XingMP3 Encoder (<http://www.xingtech.com/mp3/encoder/>).



Figura 13. Reproductor Rio.

Pero no todo es software: empiezan a aparecer los reproductores basados en hardware. Son portátiles y muy parecidos a los *walkman*. El primero que apareció, y el más popular, es el Rio de Diamond (<http://www.diamondmm.com>). El último modelo, el Rio 500, dispone de una memoria de 64 Mb, ampliable mediante tarjetas flash. La carga de los ficheros de música se realiza mediante las tarjetas o bien mediante un puerto USB con el que se conecta a un ordenador de sobremesa. El peso es de apenas 100 gramos, lo que hace que sea ideal como reproductor portátil.

TRANSMISIÓN DE AUDIO POR RED: RADIO MP3

Para finalizar este artículo, comentaremos una de las aplicaciones más innovadoras del MP3: la posibilidad de crear emisoras de radio a través de Internet, con calidad suficiente como para ser comerciales y con un ancho de banda suficientemente bajo como para poder ser recibidas a través de una conexión de baja velocidad (un módem de 33/56 Kbit/s o un canal RDSI de 64 Kbit/s).

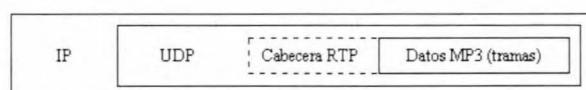


Figura 14. Protocolos involucrados en la transmisión de MP3 en red.

Esto se consigue segmentando los ficheros MP3 en trozos pequeños, de manera que en cada uno de ellos existe un número entero (y no muy elevado) de tramas. Como hemos visto en la descripción del estándar, las tramas nos permiten decodificar por entero un cierto número de muestras PCM. Así generamos un flujo continuo (*stream*) de trozos que se encapsulan en paquetes RTP (Real Time Protocol) [12], que inserta marcas temporales que permiten una reconstrucción fiel del flujo original. Estos paquetes, a su vez, son transportados por los protocolos UDP, o TCP (que a su vez descansan sobre el IP).

El uso de UDP, que se define como un protocolo «ligero» (es decir, de funcionamiento sencillo y con muy poca carga de cabeceras) permite que el transporte de tramas MP3 sea muy eficiente. Sin embargo, el TCP proporciona mucha más protección frente a pérdidas, a costa de retardos y saltos en la reproducción.

Estas dos opciones de transporte han hecho aparecer dos tipos de emisoras de MP3. Por un lado tenemos las basadas en TCP, cuyo mejor exponente es ShoutCast (<http://www.shoutcast.com>). Esta página web, mantenida por la misma empresa del WinAmp, es en realidad un portal que da entrada a varios centenares de emisoras distribuidas por todo el mundo, conectadas a la sede central de Nullsoft a través de Internet. El transporte de la información se realiza sobre HTTP y TCP, y se abre una conexión por cada nuevo usuario, lo que limita la cantidad de oyentes (ya que en caso contrario se desbordaría la capacidad del ordenador emisor y de los enlaces que lo conectan a Internet).

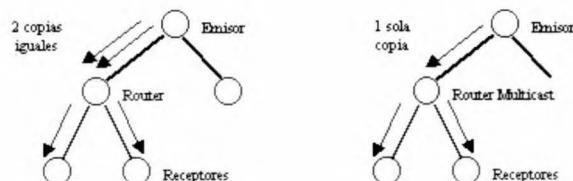


Figura 15. Comparación entre el modo unicast (izquierda) y el multicast (derecha).

Este problema de escalabilidad, que es un obstáculo para una emisión masiva, es resuelto por la otra gran familia de emisoras. Éstas se basan en el uso de UDP y del protocolo IP con extensiones Multicast [12], que permite que sólo se emita una copia de la información desde el emisor, independientemente del número de receptores que estén activos. Cuando la información llega al último *router* del árbol de distribución, se efectúa una copia por cada uno de los usuarios que quieren recibir el flujo. Así se minimiza la carga en los tramos superiores del árbol, donde solo circula una copia de los paquetes. Esta técnica requiere

de routers modernos, con capacidad de encaminar los paquetes multicast, por lo que su uso está restringido por ahora a entornos experimentales y universitarios, pero sin duda se extenderá en el futuro hacia los usuarios comerciales de Internet. Una aplicación muy sencilla y potente para la transmisión de flujos MP3 multicast es liveCaster (<http://www.livegate.com>).

FUTURO

A la vista de la velocidad a la que cambia el mundo de la tecnología, se hace difícil decir cuál va a ser el futuro de los codificadores MPEG y qué impacto van a tener en nuestra vida cotidiana. Lo que sí parece claro es que hay una tendencia hacia la creación de productos de hardware especializado en compresión y reproducción MP3, lo cual podría hacer que se convirtiera en un estándar «de facto» en el mundo de la electrónica de consumo, desplazando a productos como el MiniDisc y compitiendo con el CD Audio.

En el campo de la transmisión de audio con calidad *broadcast*, ya sea asociado a una señal de vídeo o por sí mismo, se continúa el desarrollo de algoritmos basados en efectos psicoacústicos. El último de ellos es el AAC (*Advanced Audio Coding*) de Dolby y NBC, aprobado como estándar para la banda de sonido de MPEG-2. Este sistema se basa en el Dolby Surround de 5+1 canales, y es capaz de multiplexar hasta 48 canales de audio, 15 canales de mejora de baja frecuencia, y 15 canales de datos. Según sus diseñadores, un flujo AAC stereo a 96 Kbit/s ofrece una calidad superior a la de MPEG-1 capa 3 a 128 Kbit/s o MPEG-capas 2 a 192 Kbit/s. Para conseguirlo, suma técnicas predictivas a las psicoacústicas y utiliza un banco de filtros de alta resolución. Pero ya se están anunciando algoritmos de compresión superiores en prestaciones...

Como puede verse, el mundo del audio digital está en plena ebullición. Os recomiendo que os mantengáis al corriente de las últimas novedades visitando los *links* que se proporcionan al final del artículo, y que experimentéis con los programas; es la mejor manera de aprender y disfrutar de la tecnología.

PARA MÁS INFORMACIÓN...

... sobre los estándares MPEG de video y audio:

<http://www.mpeg.org>

... sobre las cuestiones técnicas relacionadas con el formato de audio MPEG:

<http://www.mp3tech.com>

... sobre el Instituto Fraunhofer, creador del estándar MP3:

<http://www.iis.fhg.de>

... sobre música en formato MP3:

<http://mp3.lycos.com>

... sobre productos software y hardware MP3:

<http://www.mp3.com>

... sobre código C de codificadores y decodificadores:

<http://mp3tech.free.fr/programmers/programmers.html>

REFERENCIAS

- [1] B. Sklar, «Digital communications fundamentals and applications», Prentice-Hall International, 1988.
- [2] Marcos Faúndez Zanuy, «Estándares de codificación de audio MPEG», Mundo Electrónico, Septiembre 1999.
- [3] Z. Nian-Li, Audio Compression course notes, http://www.cs.sfu.ca/CC/365/li/material/notes/Chap4/Chap4.4/Chap4.4_prev.html
- [4] Official MPEG Website, <http://drogo.cselt.stet.it/mpeg>
- [5] ISO – International Standards Organisation, <http://www.iso.ch>
- [6] Introducción a la compresión de audio: MPEG 1 Layer 3, <http://www.geocities.com/SiliconValley/Vista/5390/index.html>,
- [7] Davis Pan, «A Tutorial on MPEG Audio compression», IEEE Multimedia, pp 60-74, 1995
- [8] J.L. Mitchell, W.B. Pennebaker, C.E. Fogg, D.J. LeGall, MPEG Video Compression Standard, Chapman and Hall - International Thomson Publishing, 1997.
- [9] Fraunhofer IIS. <http://www.iis.fhg.de/amm/techinf/layer3/index.html>
- [10] David Bowie website, <http://www.davidbowie.com>
- [11] Un buscador de ficheros MP3 se puede encontrar en <http://mp3.lycos.com>
- [12] S.A. Thomas, «IPng and the TCP/IP Protocols», Wiley Computer Publishing, 1996.
- [13] Kosiur, D.R., «IP multicasting the complete guide to interactive corporate networks», John Wiley & Sons, 1998.



SENSORES INTELIGENTES: UNA HISTORIA CON FUTURO



UNIVERSIDAD POLITECNICA DE CATALUNYA, BARCELONA-SPAIN
UNIVERSIDAD NACIONAL EXPERIMENTAL POLITECNICA «ANTONIO JOSE DE SUCRE», PUERTO ORDAZ - VENEZUELA

MsC. Ing. Angel Custodio Ruiz

El Ing. Angel Custodio es Profesor Asistente en el Dpto. de Ingeniería Electrónica de la UNEXPO, Vicerrectorado Puerto Ordaz. Actualmente realiza el Doctorado en el Departamento de Ingeniería Electrónica de la Universidad Politécnica de Catalunya, en Barcelona, España.

*(Dirección: UPC Campus Nord, C/Jordi Girona 1-3, Edificio C4, Dpto. de Ingeniería Electrónica, C.P. 08034)
Telf: 34-934016840, Fax: 34-934016756, e-mail: custodio@eel.upc.es.*

PhD Ramon Bragós Bardía

El Dr. Ramon Bragós Bardía es Profesor Titular en el Dpto. de Ingeniería Electrónica de la Universidad Politécnica de Catalunya en Barcelona, España.

*(Dirección: UPC Campus Nord, C/Jordi Girona 1-3, Edificio C4-311, Dpto. de Ingeniería Electrónica, C.P. 08034)
Telf: , Fax: 34-934016756, e-mail: rbb@eel.upc.es.*

PhD Ramon Pallàs Areny

El Dr. Ramon Pallàs Areny es Catedrático de Universidad en el Dpto. de Ingeniería Electrónica de la Universidad Politécnica de Catalunya en Barcelona, España. Actualmente dirige la División de Instrumentación y Bioingeniería del Dpto de Ingeniería Electrónica de la Universidad Politécnica de Catalunya, en Barcelona, España.

*(Dirección: UPC Campus Nord, C/Jordi Girona 1-3, Edificio C4-305, Dpto. de Electrónica, C.P. 08034)
Telf: , Fax: 34-934016756, e-mail: elera@eel.upc.es.*

INTRODUCCIÓN

La historia de los sensores inteligentes (o Smart Sensors) ha sido un tema apasionante desde el primer y muy sencillo sensor desarrollado por Honeywell en el año 1969, hasta los dispositivos complejos actuales de alta tecnología. Estos últimos integran muchas funciones automáticas: Identificación, calibración, comprobación, etc., que permiten obtener sensores que no sólo entregan una señal digital (o casi digital), sino además, linealizada, calibrada, robusta y compatible con otros dispositivos.

El desarrollo de estos revolucionarios componentes ha permitido aumentar la eficiencia, calidad y velocidad de los procesos industriales, la investigación y el desarrollo científico.

digitales derivados de la conversión de la señal analógica adigital, como la comunicación remota y la direccionalidad. Esta última consiste en la posibilidad de conocer en qué punto se encuentra el sensor (ya sea en una red de multiplexado, red de sensores, o bus de campo). El desa-



Figura 1. Medidor de temperatura industrial, basado en sensores inteligentes. Cortesía de Moore Industries.

LA MOTIVACIÓN

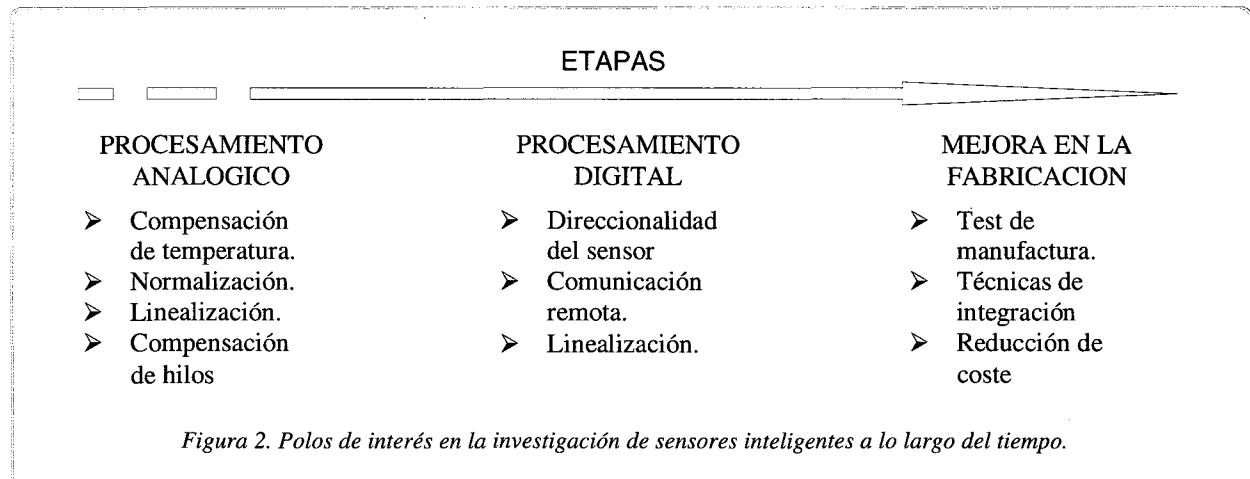
El foco de interés en el desarrollo de los sensores inteligentes ha ido cambiando a lo largo de su vida [1]. Inicialmente la atención se concentró en el procesamiento de la señal generada por el sensor para mejorar la compensación de la temperatura y lograr una señal normalizada. Más tarde la atención se centró en mejorar los sistemas

rrollo más reciente está orientado a los tests de manufactura e integración para mejorar la fabricación de los sensores, con el fin de reducir los costos y mejorar la



relación precio/prestaciones. Esto incluye el diseño de nuevos dispositivos con arquitecturas optimizadas: lo más simple que se puedan realizar, pero conservando los beneficios alcanzados a lo largo de su historia.

sensores. Lo desarrolló Honeywell en los años 60, para el sistema de aire en los aviones DC-9 [2]. Estaba formado por dos piezoresistores que medían la presión, y dos capacitores para crear un desplazamiento de fase. Estos



Estas arquitecturas actuales, aunque tiene el mismo objetivo, tienen forma de implementarse muy diferentes: recurrir a la microelectrónica para integrar componentes ya existentes (convertidores A/D, microprocesador, sensores de silicio, etc.); utilizar sistemas híbridos electromecánicos que utilizan el principio de los servomecanismos; plantear nuevas técnicas de conversión, etc.

En este artículo haremos hincapié en aquella arquitectura que busca simplificar el diseño mediante la técnica de conversión directa de la señal del proceso o sistema a digital. Entiéndase, sin utilizar convertidores A/D convencionales ni amplificadores operacionales. Esto permitirá conocer un mundo interesante donde la lucha entre bajo coste y eficiencia es mucho más fuerte que en otras tendencias.

EL NACIMIENTO

El primer sensor inteligente nació como una solución al problema de compensación de temperatura en los

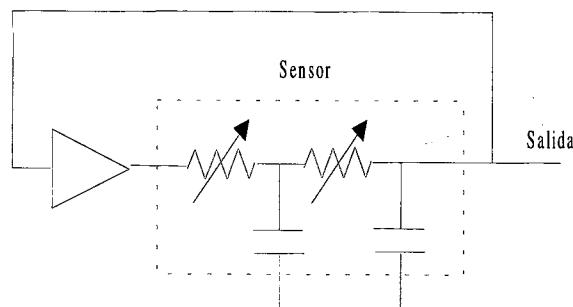


Figura 3. El primer sensor inteligente usaba sensores piezorresistivos y capacitivos integrados, conectados a un inversor para crear una salida en frecuencia proporcional a la presión.

elementos estaban realimentados y conectados a un inversor para crear un oscilador. La frecuencia de salida era proporcional a la constante de tiempo RC, y por ende a la presión. Tiempo después Toyota Research presentó otro sensor de presión similar [3]. Ambas empresas, sin imaginárselo, estaban empezando una revolución sin par en las tecnologías de sensado, que aun hoy continúa.

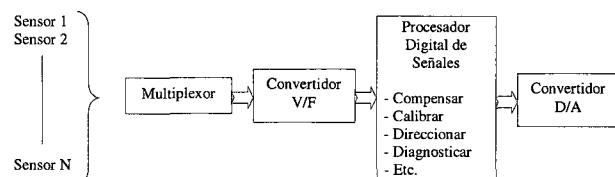


Figura 4. Sensor inteligente de la generación de los ochenta.

La siguiente generación la introdujo nuevamente Honeywell en los años 80. La presentó en dos aplicaciones: Una para el control de procesos (ST3000) [4], y otra para aplicaciones aerospaciales [5]. Ambos estaban formados por un grupo de sensores multiplexados, conectados a un convertidor V/F (tensión/frecuencia). La frecuencia obtenida la procesaba un microprocesador tipo DSP, y la salida se llevaba a un convertidor D/A. La salida era acorde al estándar analógico de 4 mA a 20 mA. El software tenía compensación de la presión estática, calibración remota del rango, direccionalidad y diagnóstico. Los sensores eran de presión diferencial estática y temperatura en el caso del control de procesos, y presión absoluta y temperatura en el caso aerospacial.

En la presente década se ha multiplicado el desarrollo de sensores integrados aplicables en el ámbito industrial, desarrollados por investigadores y académicos [6-9], aunque su precio elevado ha llevado a centrarse en

arquitecturas simples y mejorar los procesos de integración.



Figura 5. Los sensores inteligentes se utilizan en todo tipo de ámbito

LA ACTUALIDAD

Actualmente hay cuatro grandes tendencias en el diseño de sensores inteligentes:

1. Desarrollo de sensores inteligentes integrados. Consiste básicamente en integrar todo el dispositivo junto al sensor. Algunos ejemplos recientes son:

- 1.1. Un oscilador controlado por tensión basado en un simple flip-flop tipo D presentado por Xi. [10]
- 1.2. La integración de un sensor capacitivo junto a un convertidor Σ/Δ propuesto por Llamada y Watanabe [11]. Se consigue linealizar la señal, ajustar el cero y la ganancia.
- 1.3. Un sensor para medición angular o lineal (tipo resistivo sin contacto), con conversión a frecuencia mediante el uso de un oscilador modificado de Martín, desarrollado por Li y Meijer [12].

2. Usar sistemas de conversión integrados y dejar todo el procesamiento y parte del acondicionamiento en una computadora conectada a la red. Scheriber propone que el sistema de interfaz realice sólo la captura y conversión, y dejar todo el procesamiento para corregir los errores a un PC [13]. De esta manera se podría proponer un sistema avanzado de eliminación de errores, tal como los de lógica inteligente.

También se podría usar un microcontrolador muy pequeño (como la serie PIC2XXX de Microchip Technology), y construir un sistema de adquisición de señales (multicanal) muy económico. Pailoor propone un circuito de estas características muy simple y económico [14].

3. Desarrollo de sistemas de medida integrados como el AD654 (convertidor V/F) de Analog Devices el cual permite acoplar termopares y galgas directamente.

Otra idea simple es propuesta por Atmanand para la medida de sensores L, C, y R [15]. Para ello ha desarrollado un circuito de interfaz en el cual el sensor (uno de los tres tipos) se coloca en serie con una resistencia formando la mitad de un puente. La otra mitad está formada por un convertidor DAC multiplicador. Un detector de fase compara ambas señales analógicas de los puentes y su salida es transformada a frecuencia a través de un comparador. El error en la lectura es de 0,7 %.

La familia AD771X de Analog Devices está orientada al desarrollo de sistemas de adquisición integrados (señales simples, diferenciales o pseudodiferenciales, en sensores tipo puente o medida de tensiones). Las señales de entrada son multiplexadas a un convertidor A/D (tipo Σ/Δ), y la señal digital de salida tiene formato serie. El AD280 (del mismo fabricante) tiene 4 canales y está orientado a señales de tensión y termistores, sensores integrados y RTD (permitiendo conexión de 2, 3 o 4 hilos). Debe notarse que en el caso de medida de señales diferenciales, lo que hace el dispositivo es multiplexar ambas entradas. El componente más avanzado de este fabricante es el ADuC812, el cual integra un microcontrolador (8052) y tiene un precio de apenas 20 \$.

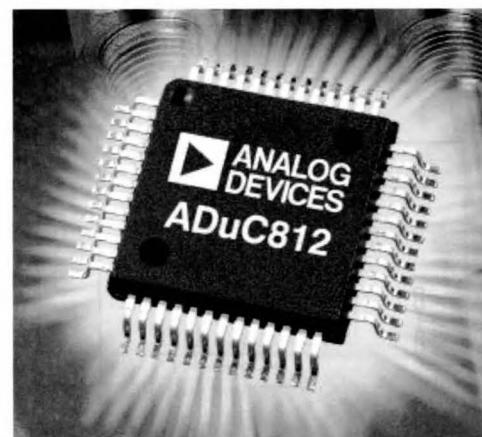
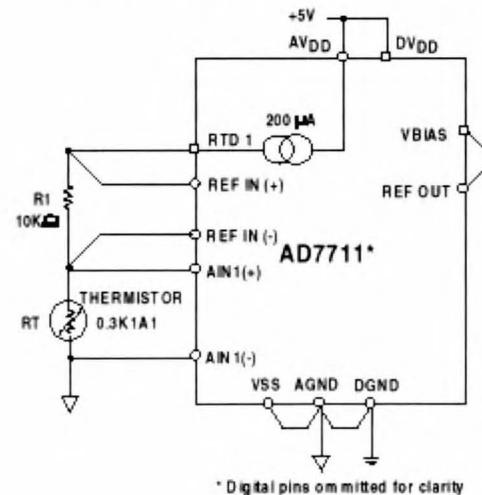


Figura 6. Ejemplos representativos de circuitos de interfaz de sensores inteligentes de la empresa Analog Device.

Crystal tiene el CS550X para conexión a circuitos diferenciales, pero necesita una electrónica compleja de interfaz. También tiene el CS554X, que es multicanal, con autocalibración de cero y ganancia.

Las ideas más revolucionarias pueden encontrarse en los dos trabajos siguientes:

3.1. El USIC (Universal sensor interface chip) desarrollado por ERA Technology Ltd [16]. Este chip posee comparadores, demoduladores Σ/Δ , filtros digitales, interfaz serie y paralelo, convertidor DAC, multiplexor, memoria RAM y un microprocesador RISC. Los convertidores son de 20 bits. Los multiplexores pueden seleccionar tres fuentes de entrada para cada convertidor, haciendo un direccionamiento total de 6 sensores. La salida digital puede ir por RS482 / RS232, o paralelo. Este sistema sin embargo fue retirado del mercado por ser demasiado costoso.

3.2. El UTI (Universal Transducer Interface), creado por Van der Goes y Meijer [17], que se puede utilizar para sensores capacitivos, RTD, termistores, puentes de resistencia y potenciómetros. Se basa en una red de interruptores utilizados para seleccionar el tipo de sensor a utilizar. La salida de la red se conecta a un convertidor de carga a periodo (oscilador de relajación), y éste genera la salida en frecuencia del sistema. El circuito tiene una resolución de 16 bits en un rango de medida de 1 ms a 100 ms. Es muy económico por estar desarrollado en tecnología CMOS. Opera con una fuente simple (3,3 V a 5,5 V), posee auto-calibración de offset y ganancia, medida 2/3/4 hilos, suprime interferencias de 50/60 Hz, y todo en un encapsulado DIP de 16 patillas.

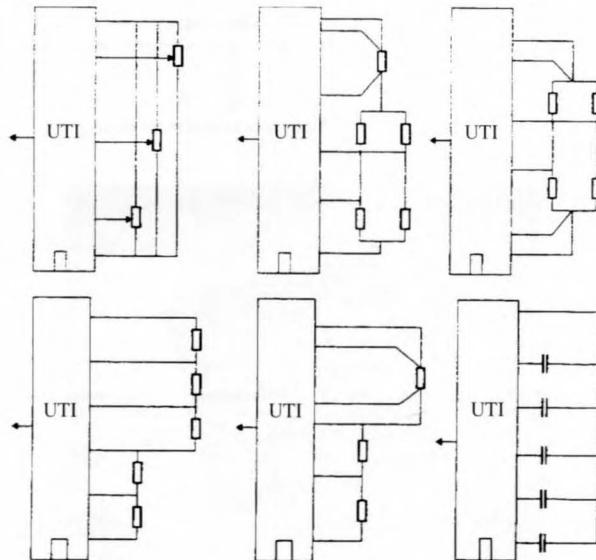


Figura 7. Posibilidades de conexión del UTI en la medida de sensores diferentes.

4. Desarrollo de sensores inteligentes de bajo coste no integrados. En el ámbito discreto se han planteado algunas técnicas de conversión a frecuencia más económicas que las anteriores. Un circuito muy simple formado por dos inversores y un comparador [18], puede alcanzar los 9 bits de resolución y, al proveer una salida que es la relación entre un voltaje de referencia y el voltaje de interés, permite reducir algunas fuentes de error.

Mochizuke y Watanabe proponen un circuito para el procesamiento de señal de alta precisión en sensores capacitivos [19]. Está basado en un oscilador de relajación. Es muy simple pues está formado por 4 operacionales. Su señal de salida es en frecuencia, siendo ésta una relación de valores entre sensores capacitivos (medida ratiométrica). Se puede detectar cambios pequeños de capacitancia de 0,1 % en un tiempo total de 10 μ s.

Ferrari propone el diseño de una interfaz que permite suministrar mediante una sola señal dos informaciones del proceso [20]. La frecuencia la controla un sensor tipo puente, mientras que el ciclo de trabajo lo controla otro sensor. Weinberg propone otro circuito que aprovecha estas dos características para obtener dos informaciones en una misma señal [21].

Otro circuito simple de Mochizuki y Watanabe, pero para sensores resistivos en puente [22], utiliza un oscilador de relajación, cuya salida realimenta a la alimentación del puente. Se obtiene una resolución de 0,05 %, y una excelente linealidad.

En todos estos trabajos se obtiene la relación de la señal medida respecto a una señal de referencia. De esta manera se reduce el efecto de elementos indeseables como son la temperatura y las variaciones en la fuente de alimentación, el efecto de elementos secundarios en la medida, y la corrección de offset y errores de sensibilidad, entre otras cosas.



Figura 8. Para el desarrollo de un sensor inteligente se necesita una fuerte investigación y el apoyo de recursos actualizados.

EL PRESENTE Y EL FUTURO

Los sensores inteligentes diseñados eran incompatibles hasta que se decidió definir una norma que permitiera su interconexión: La norma IEEE-P1451 (Transductor to Microprocessor Interface), la cual fija las directivas de los sensores inteligentes, así como la comunicación con el bus digital [23].

A continuación se exponen unas ideas generales sobre la posible repercusión futura de esta norma.

La norma generaliza el concepto de sensor inteligente al de Transductor Inteligente, todos montados sobre una red común y con la información necesaria para saber en cualquier momento quien es sensor y quien es actuador, y que propiedades tiene cada uno.

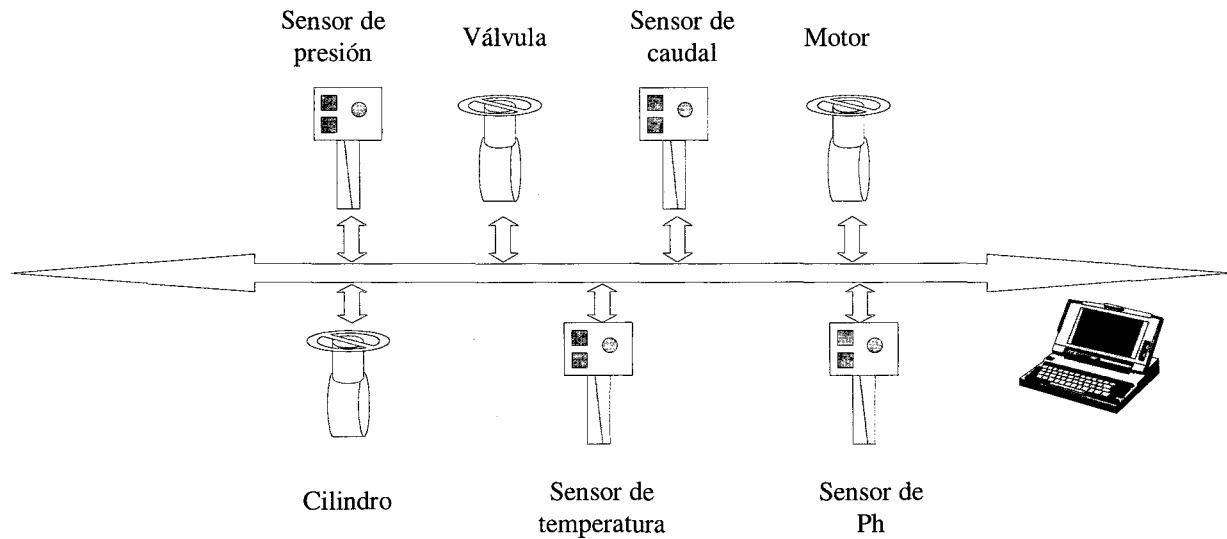


Figura 9. Transductores interconectados en un mismo bus.

La figura 9 recoge la idea central de la norma. Hay varios dispositivos, sensores y actuadores, de diversos fabricantes, con diversas tecnologías. Algunos pueden ser más complejos que otros, pero comparten los requerimientos mínimos de identificación, direccionalidad y comunicación. De esta manera el sensor, no sólo podrá «conocerse a sí mismo», sino que también podrá «conocer a sus compañeros». Esta comunicación interactiva permitirá saber qué tipo de transductor es el dispositivo (sensor o actuador), cuál es su función (sensor de presión, caudal, etc.), posibilidades, ubicación y fabricante. Así podrán realizarse enlaces online de identificación, corrección de parámetros, optimización del proceso, y un amplio etcétera de posibilidades.

Desde el punto de vista del sensor, éste es un circuito (integrado o no), que tiene una o más de las funciones de sensado (uno o más sensores), interfaz (acondicionamiento de la señal, conversión entre dominios, estandarización de la salida), calibración (cero, linealidad,

sensibilidad, corrección de temperatura), e inteligencia (auto-comprobación, auto-calibración y auto-identificación).

La arquitectura en diagramas de bloques se muestra en la figura 10.

El NCAP es un microprocesador encargado de administrar la comunicación con la red donde se instale el dispositivo. El TEDS es una ROM donde se guarda la información de identificación del componente (tipo, fabricante, funciones, etc.). El funcionamiento es muy simple: Los sensores, vía un multiplexor, entregan la información a un acondicionador de señal, el cual se encarga de corregir los errores fundamentales de la señal (linealidad, offset, derivas, etc.). Luego esta señal se digitaliza a digital; esta información, junto con la del TEDS, se envía

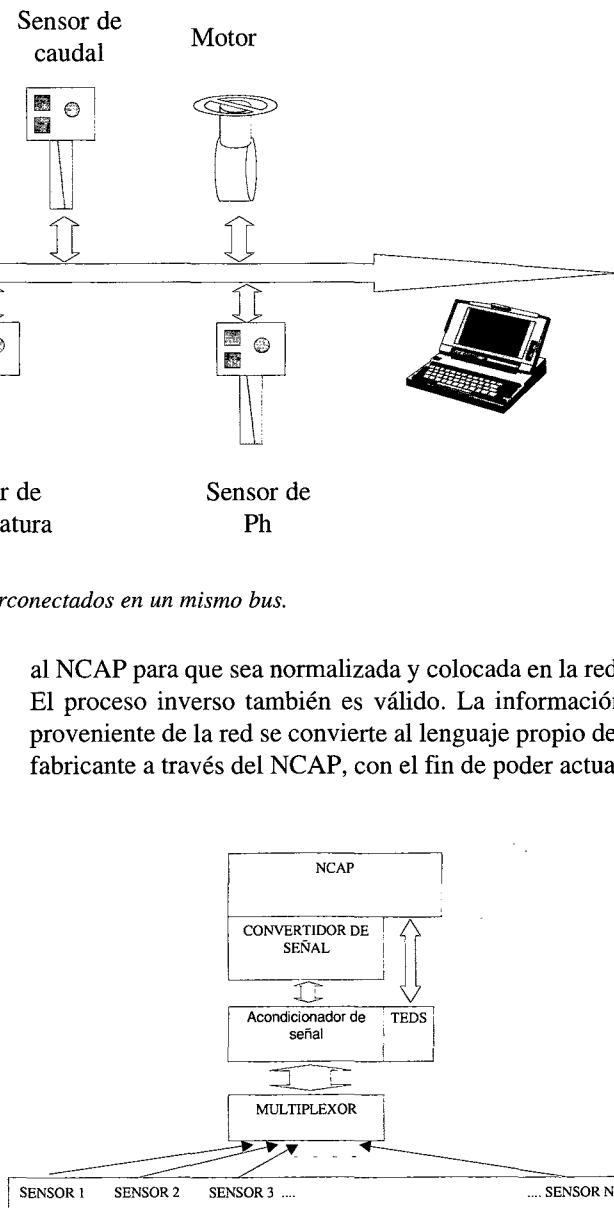


Figura 10. Arquitectura de un "Sensor Inteligente" formado por varios sensores.



sobre el sensor, ya sea para corregir algún parámetro o simplemente para identificarlo.

El futuro de la instrumentación pasa por hacer realidad esta norma, desde el punto de vista de la compatibilidad, ya que fabricantes como Motorola o Siemens, han implementando los conceptos, pero con sus equipos y sus redes.

Este mundo de la instrumentación física, termina por conectarse y complementarse con el de la instrumentación virtual. Esta última busca proyectar en un computador personal toda la información necesaria para supervisar y controlar los componentes presentes en una red, de tal forma que se integre la instrumentación y se transforme al computador en un instrumento.

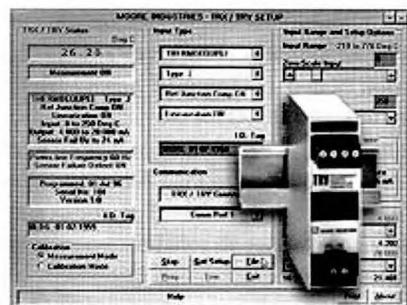


Figura 11. La integración software-hardware forma parte del futuro de la instrumentación

Claro que más allá del PC, la instrumentación virtual mediante el uso de la red de redes, Internet, transportará todas estas funciones a cualquier dispositivo: Televisores domésticos, tercera generación de teléfonos móviles, etc.

A pesar del gran esfuerzo realizado por algunas instituciones como la IEEE y su norma, la fundación Fieldbus, IVI, etc., el reto del futuro próximo seguirá siendo la interconectividad entre sensores.



Figura 12. Sistema Scada donde se interconectan sensores y actuadores en un proceso.

REFERENCIAS

- [1] J Bryzek. (1993, October). «The evolution of smart sensor and transducer design,» Sensors, Vol. 10, N° 10: 12-23.
- [2] E. Peake. (1969). «Solid state digital pressure transducer,» IEEE Trans. on Electron Devices, ED-16(10): 870-876.
- [3] S. Sugiyama, M. Takigawa, I. Igarashi. (1983). «Integrated piezoresistive pressure sensor with both voltage and frequency output,» Sensors and Actuators, 4: 113-120.
- [4] E. Kompass. (1983, October). «‘Smart’ Transmitter stores calibration data,» Control Engineering: 100-110.
- [5] P. DuPuis. (1994, May). «A novel primary air data quality pressure transducer,» Proc. IEEE National Aerospace and Electronics Conference (Dayton, OH.).
- [6] U. Schoneberg et al. (1991, June). «CMOS integrated capacitive pressure transducer with on-chip electronics and digital calibration capability,» Digest of Technical Papers, Transducer '91 (San Francisco, CA.).
- [7] W. Gruber. (1993, March). «Computer literate pressure sensors,» Proc. Sensors Expo West (San Jose, CA.).
- [8] N. Najafi. (1993, March). «A multi-element gas analyzer utilized in a smart sensing system,» Proc. Sensors Expo West (San Jose, CA.).
- [9] W. Yun and R. Howe. (1993, March). «Interfacing force balance accelerometers,» Proc. Sensors Expo West (San Jose, CA.).
- [10] Y. Xi. (1992, Sep.). «Dual D Flip-Flop forms VCO,» Electronics Design: 89.
- [11] M. Llamada, K. Watanabe. (1997). «A capacitive pressure sensor interface using oversampling delta sigma demodulation techniques,» IEEE Trans. Instrum. Meas., 46(1): 3-7.
- [12] X. Li, G. Meijer. (1998). «A novel low-cost noncontact resistive potentiometric sensor for the measurement of low speed,» IEEE Trans. Instrum. Meas., 47(3): 776-781.
- [13] R. Schreiber. (1997). «Air flow control using fuzzy logic,» Application Note AN600, Microchip Technology.
- [14] R. Paloor. (1997). «Smart sensor,» Sensor Interface, Microchip Technology Inc.
- [15] M. Atmanand, V. Kumar, V. Murti. (1996). «A microcontroller-based quasi-balanced bridge for the measurement of L, C and R,» IEEE Trans. Instrum. Meas., 45(3): 757-761.
- [16] P. Wilson, S. Hopkins, R. Spraggs, I. Lewis, V. Skarda, J. Goodey. (1995). «Application of a universal sensor interface chip (USIC) for intelligent sensor application,» Advances in Sensors, IEE Colloquium on: 3/1-3/6.
- [17] F. Vand der Goes, G. Meijer. (1997). «A universal transducer interface for capacitive and resistive sensor elements,» Analog Integrated Circuits and Signal Processing, 14: 249-260.
- [18] M. Walne. (1994, March). «Simple ADC is surprisingly accurate,» EDN: 154.
- [19] K. Mochizuke, K. Watanabe, T. Masuda. (1998). «A high-accuracy high-speed signal processing circuits of differential-capacitance transducers,» IEEE Trans. Instrum. and Meas., 47(5): 1244-1247.
- [20] V. Ferrari, C. Ghidini, D. Marioli, A. Taroni. (1997, May). «Oscillator-based signal conditioning for resistive sensors,» IEEE Instrum. Meas. Technology Conference (Ottawa): 1490-1494.
- [21] H. Weinberg. (1998, Nov). «Accelerometer output gives temperature info,» EDN: 130.
- [22] K. Mochizuki, K. Watanabe. (1996). «A high-resolution, linear resistance-to-frequency converter,» IEEE Trans. Instrum. Meas., 45(3): 761-764.
- [23] S. Woods. (1996, June). «The IEEE-P1451 Transducer to microprocessor interface,» Sensors and Systems: 43-48.



REDES DE ÁREA LOCAL INALÁMBRICAS SEGÚN EL ESTÁNDAR IEEE 802.11

Miquel Oliver, Ana Escudero

Grupo de Comunicaciones Móviles y de Banda Ancha
Departamento de Matemática Aplicada i Telemática (DMAT)
Universidad Politécnica de Catalunya.

I. INTRODUCCIÓN

En los últimos años se ha producido un crecimiento espectacular en lo referente al desarrollo y aceptación de las comunicaciones móviles y en concreto de las redes de área local (Wireless LANs). La función principal de este tipo de redes es la proporcionar conectividad y acceso a las tradicionales redes cableadas (Ethernet, Token Ring...), como si de una extensión de éstas últimas se tratara, pero con la flexibilidad y movilidad que ofrecen las comunicaciones inalámbricas. El momento decisivo para la consolidación de estos sistemas fue la conclusión del estándar IEEE 802.11 el pasado mes de junio de 1997. En este estándar se encuentran las especificaciones tanto físicas como a nivel MAC que hay que tener en cuenta a la hora de implementar una red de área local inalámbrica. Otro de los estándares definidos y que trabajan en este mismo sentido es el ETSI HIPERLAN. En este trabajo nos centraremos en el estudio del primero, el estándar IEEE 802.11.

Los sistemas WLAN no pretenden sustituir a las tradicionales redes cableadas, sino más bien complementarlas.

La idea que queremos resaltar es que los sistemas WLAN no pretenden sustituir a las tradicionales redes cableadas, sino más bien complementarlas. En este sentido el objetivo fundamental de las redes WLAN es el de proporcionar las facilidades no disponibles en los sistemas cableados y formar una red total donde coexisten los dos tipos de sistemas. El presente trabajo está estructurado tal y como se indica a continuación: en la sección II se presentan los sistemas WLAN mediante algunas definiciones y una breve reseña histórica. A continuación, en la sección III, hacemos una descripción de las tecnologías que actualmente se utilizan en la fabricación de estos productos. La sección IV la dedicamos a las configuraciones de redes inalámbricas que podemos encontrar habitualmente. En la sección V nos centramos en el nivel MAC del estándar IEEE 802.11 y hacemos una breve descripción del algoritmo de acceso que se utiliza en este

nivel: el algoritmo CSMA/CA. Finalmente, en la sección VI, resumimos brevemente los puntos tratados a lo largo del artículo.

II. GENERALIDADES SOBRE REDES DE ÁREA LOCAL INALÁMBRICAS

A. Definición de Red de Área Local Inalámbrica

Una red de área local inalámbrica puede definirse como a una red de alcance local que tiene como medio de transmisión el aire.

Por red de área local entendemos una red que cubre un entorno geográfico limitado, con una velocidad de transferencia de datos relativamente alta (mayor o igual a 1 Mbps tal y como especifica el IEEE), con baja tasa de errores y administrada de forma privada.

Por red inalámbrica entendemos una red que utiliza ondas electromagnéticas como medio de transmisión de la información que viaja a través del canal inalámbrico enlazando los diferentes equipos o terminales móviles asociados a la red. Estos enlaces se implementan básicamente a través de tecnologías de microondas y de infrarrojos. En las redes tradicionales cableadas esta información viaja a través de cables coaxiales, pares trenzados o fibra óptica.

Una red de área local inalámbrica, también llamada wireless LAN (WLAN), es un sistema flexible de comunicaciones que puede implementarse como una extensión o directamente como una alternativa a una red cableada.

Este tipo de redes utiliza tecnología de radiofrecuencia minimizando así la necesidad de conexiones cableadas. Este hecho proporciona al usuario una gran movilidad sin perder conectividad. El atractivo fundamental de este tipo de redes es la facilidad de instalación y el ahorro que supone la supresión del medio de transmisión cableado. Aún así, debido a que sus prestaciones son menores en lo referente a la velocidad de transmisión que se sitúa entre los 2 y los 10 Mbps frente a los 10 y hasta los 100 Mbps ofrecidos por una red convencional, las redes inalámbricas son la alternativa ideal para hacer llegar una red tradicional a lugares donde



el cableado no lo permite, y en general las WLAN se utilizarán como un complemento de las redes fijas.

B. Aplicaciones de los sistemas WLAN

Las aplicaciones más típicas de las redes de área local que podemos encontrar actualmente son las siguientes:

- Implementación de redes de área local en edificios históricos, de difícil acceso y en general en entornos donde la solución cableada es inviable.
- Posibilidad de reconfiguración de la topología de la red sin añadir costes adicionales. Esta solución es muy típica en entornos cambiantes que necesitan una estructura de red flexible que se adapte a estos cambios.
- Redes locales para situaciones de emergencia o congestión de la red cableada.
- Estas redes permiten el acceso a la información mientras el usuario se encuentra en movimiento. Habitualmente esta solución es requerida en hospitales, fábricas, almacenes...
- Generación de grupos de trabajo eventuales y reuniones ad-hoc. En estos casos no valdría la pena instalar una red cableada. Con la solución inalámbrica es viable implementar una red de área local aunque sea para un plazo corto de tiempo.
- En ambientes industriales con severas condiciones ambientales este tipo de redes sirve para interconectar diferentes dispositivos y máquinas.
- Interconexión de redes de área local que se encuentran en lugares físicos distintos. Por ejemplo, se puede utilizar una red de área local inalámbrica para interconectar dos o más redes de área local cableadas situadas en dos edificios distintos.

C. Breve reseña histórica

Las redes de área local inalámbricas funcionan desde hace más de quince años en entornos industriales y de investigación.

Este tipo de redes se implementó por primera vez en el año 1979. La casa IBM Suiza, utilizó enlaces infrarrojos creando una red de área local en una fábrica. Posteriormente se utilizaron implementaciones basadas en tecnologías de microondas según los esquemas de transmisión de espectro ensanchado.

En marzo de 1985 la Comisión Federal de Comunicaciones, FCC, organismo encargado de la regulación de las telecomunicaciones en Estados Unidos, asignó a los sistemas WLAN las bandas frecuenciales 902-928 MHz., 2.400-2.4835 GHz. y 5.725-5.850 GHz también conocidas como ISM (Industrial, Científica y Médica) y que pueden utilizarse bajo licencia administrativa. Esta asignación de una localización frecuencial fija propició una mayor actividad industrial. En este punto las redes de área local inalámbrica dejaron de ser meramente experimentales para empezar a introducirse en el mercado. Entre los años 1985 y 1990 se trabajó en el desarrollo de productos WLAN y finalmente, en mayo de 1991, se publicaron algunos trabajos que hablaban sobre redes inalámbricas que superaban la velocidad de transferencia de 1 Mbps velocidad mínima a partir de la cual el comité IEEE considera que una red es de área local.

Hasta ese momento las WLAN habían tenido una aceptación marginal en el mercado por dos razones fundamentales: falta de un estándar y precios elevados de la solución inalámbrica.

En estos últimos años se ha producido un crecimiento en el mercado de hasta un 100 % anual. Este hecho es atribuible a dos razones principales:

- El desarrollo del mercado de los equipos portátiles y de las comunicaciones móviles que han producido que los usuarios puedan estar en continuo movimiento manteniendo comunicación constante con otros terminales y elementos de la red. En este sentido, las comunicaciones inalámbricas ofrecen una prestación no disponible en las redes cableadas: movilidad y acceso simultáneo a los recursos de la red.
- La conclusión de la definición de la norma IEEE 802.11 para redes de área local inalámbricas el pasado junio de 1997 que ha establecido un punto de referencia y ha mejorado muchos de los aspectos de estas redes.

III. TECNOLOGÍAS UTILIZADAS EN LAS REDES INALÁMBRICAS

A. Tecnologías de espectro ensanchado

La tecnología de espectro ensanchado consiste en difundir la señal de información a lo largo del ancho de banda disponible, es decir, en vez de concentrar la energía de las señales alrededor de una portadora concreta lo que se hace es repartirla por toda la banda disponible. Este ancho de banda total se comparte con el resto de usuarios que trabajan en la misma banda frecuencial.

Existen dos tipos de tecnologías de espectro ensanchado:

- Espectro Ensanchado por Secuencia Directa (DSSS)
- Espectro Ensanchado por Salto en Frecuencia (FHSS)

A.1 Tecnología de espectro ensanchado por secuencia directa (DSSS)

Esta técnica consiste en la generación de un patrón de bits redundante llamado *señal de chip* para cada uno de los bits que componen la señal de información y la posterior modulación de la señal resultante mediante una portadora de RF. En recepción es necesario realizar el proceso inverso para obtener la señal de información original.

La secuencia de bits utilizada para modular cada uno de los bits de información es la llamada secuencia de Barker y tiene la siguiente forma:

+1, -1, +1, +1, -1, +1, +1, +1, -1, -1, -1

En la Figura 1 mostramos el aspecto de una señal de dos bits a la cual le hemos aplicado la secuencia de Barker.

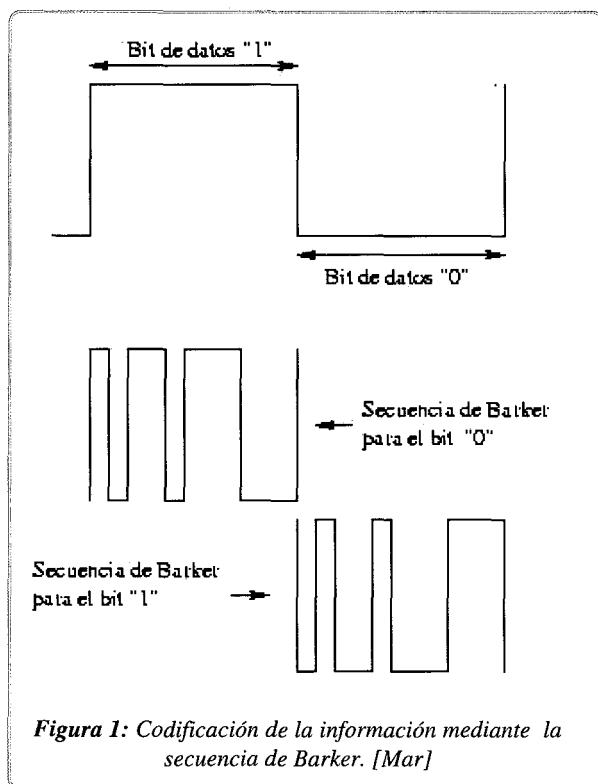


Figura 1: Codificación de la información mediante la secuencia de Barker. [Mar]

DSSS tiene definidos dos tipos de modulaciones a aplicar a la señal de información una vez se sobrepone la señal de *chip* tal y como especifica el estándar IEEE 802.11: la modulación DBPSK, Differential Binary Phase Shift Keying y la modulación DQPSK, Differential Quadrature Phase Shift Keying proporcionando unas velocidades de transferencia de 1 y 2 Mbps respectivamente.

En el caso de Estados Unidos y de Europa la tecnología de espectro ensanchado por secuencia directa, DSSS, opera en el rango que va desde los 2.4

GHz hasta los 2.4835 GHz, es decir, con un ancho de banda total disponible de 83.5 MHz. Este ancho de banda total se divide en un total de 14 canales con un ancho de banda por canal de 5 MHz de los cuales cada país utiliza un subconjunto de los mismos según las normas reguladoras para cada caso particular. En el caso de España se utilizan los canales 10 y 11 ubicados en una frecuencia central de 2.457 GHz y 2.462 GHz respectivamente.

En topologías de red que contengan varias celadas, ya sean solapadas o adyacentes, los canales pueden operar simultáneamente sin apreciarse interferencias en el sistema si la separación entre las frecuencias centrales es como mínimo de 30 MHz. Esto significa que de los 83.5 MHz de ancho de banda total disponible podemos obtener un total de 3 canales independientes que pueden operar simultáneamente en una determinada zona geográfica sin que aparezcan interferencias en un canal procedentes de los otros dos canales. Esta independencia entre canales nos permite aumentar la capacidad del sistema de forma lineal con el número de puntos de acceso operando en un canal que no se esté utilizando y hasta un máximo de tres canales. En el caso de España esta extensión de capacidad no es posible debido a que no existe el ancho de banda mínimo requerido (la información sobre la distribución de las frecuencias en distintas regiones del mundo se encuentra disponible en el estándar IEEE 802.11).

A.2 Tecnología de espectro ensanchado per salto en frecuencia (FHSS)

La tecnología de espectro ensanchado por salto en frecuencia consiste en transmitir una parte de la información en una determinada frecuencia durante un intervalo de tiempo llamada *dwell time* y inferior a 400 ms. Pasado este tiempo se cambia la frecuencia de emisión y se sigue transmitiendo a otra frecuencia. De esta manera cada tramo de información se va transmitiendo en una frecuencia distinta durante un intervalo muy corto de tiempo.

Cada una de las transmisiones a una frecuencia concreta se realiza utilizando una portadora de banda estrecha que va cambiando (saltando) a lo largo del tiempo. Este procedimiento equivale a realizar una partición de la información en el dominio temporal.

El orden en los saltos en frecuencia que el emisor debe realizar viene determinado según una secuencia pseudoaleatoria que se encuentra definida en unas tablas que tanto el emisor como el receptor deben conocer. La ventaja de estos sistemas frente a los sistemas DSSS es que con esta tecnología podemos tener más de un punto de acceso en la misma zona geográfica sin que existan interferencias si se cumple



que dos comunicaciones distintas no utilizan la misma frecuencia portadora en un mismo instante de tiempo.

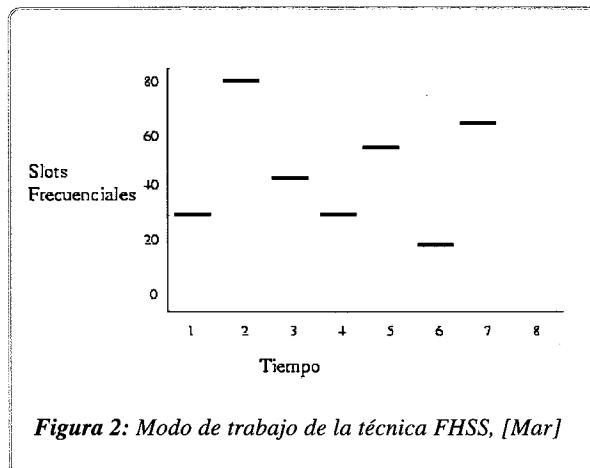


Figura 2: Modo de trabajo de la técnica FHSS, [Mar]

Si se mantiene una correcta sincronización de estos saltos entre los dos extremos de la comunicación el efecto global es que aunque vamos cambiando de canal físico con el tiempo se mantiene un único canal lógico a través del cual se desarrolla la comunicación.

Para un usuario externo a la comunicación la recepción de una señal FHSS equivale a la recepción de ruido impulsivo de corta duración.

El estándar IEEE 802.11 describe esta tecnología mediante la modulación en frecuencia FSK, Frequency Shift Keying, y con una velocidad de transferencia de 1 Mbps ampliable a 2Mbps bajo condiciones de operación óptimas también especificadas en la rma.

B. Tecnología de infrarrojos

Una tercera tecnología, de momento no demasiado utilizada a nivel comercial para implementar WLANs, es la de infrarrojos. Los sistemas de infrarrojos se sitúan en altas frecuencias, justo por debajo del rango de frecuencias de la luz visible. Las propiedades de los infrarrojos son, por tanto, las mismas que tiene la luz visible. De esta forma los infrarrojos no pueden pasar a través de objetos opacos pero se pueden reflejar en determinadas superficies.

Las longitudes de onda de operación se sitúan alrededor de los 850-950 nm, es decir, a unas frecuencias de emisión que se sitúan entre los $3,15 \cdot 10^{14}$ Hz y los $3,52 \cdot 10^{14}$ Hz.

Los sistemas que funcionan mediante infrarrojos se clasifican según el ángulo de apertura con el que se emite la información en el emisor en:

- Sistemas de corta apertura, de haz dirigido o de visibilidad directa que funcionan de manera similar a los mandos a distancia de los aparatos de televisión.

Esto supone que el emisor y el receptor tienen que estar orientados adecuadamente antes de empezar a transmitir información.

- Sistemas de gran apertura, reflejados o de difusión que radian tal y como lo haría una bombilla, permitiendo el intercambio de información en un rango más amplio.

La norma IEEE 802.11 especifica dos modulaciones para esta tecnología: la modulación 16 ppm y la modulación 4 ppm proporcionando unas velocidades de transmisión de 1 y 2 Mbps respectivamente.

Esta tecnología se aplica típicamente en entornos de interior para implementar enlaces punto a punto de corto alcance o redes locales en entornos muy localizados como puede ser una aula concreta o un laboratorio.

IV. CONFIGURACIONES WLAN

El grado de complejidad de una red de área local inalámbrica es variable, dependiendo de las necesidades a cubrir y en función de los requerimientos del sistema que queramos implementar podemos utilizar diversas configuraciones de red.

A. Peer to peer o redes ad-hoc

La configuración más básica es la llamada *de igual a igual* o *ad-hoc* y consiste en una red de dos terminales móviles equipados con la correspondiente tarjeta adaptadora para comunicaciones inalámbricas. En la figura 3 mostramos un ejemplo. Para que la comunicación entre estas dos estaciones sea posible hace falta que se vean mutuamente de manera directa, es decir, que cada una de ellas esté en el rango de cobertura radioeléctrica de la otra. Las redes de tipo *ad-hoc* son muy sencillas de implementar y no requieren ningún tipo de gestión administrativa.

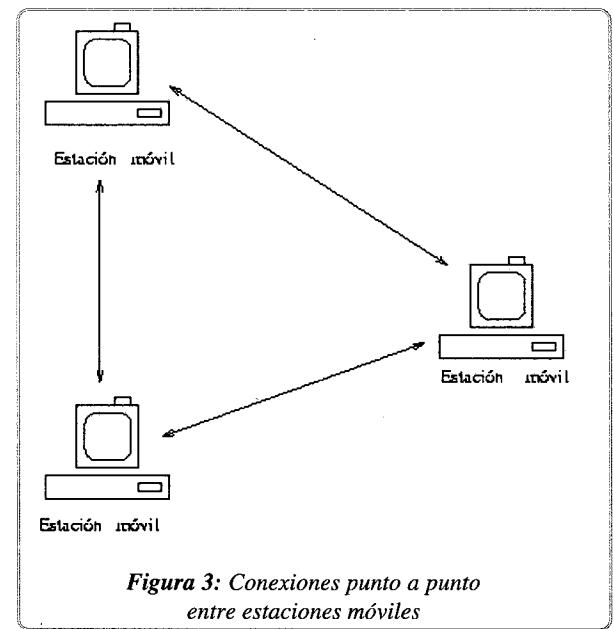


Figura 3: Conexiones punto a punto entre estaciones móviles

B. Extensión de las celdas básicas

Para aumentar el alcance de una red del tipo anterior hace falta la instalación de un *punto de acceso*. Con este nuevo elemento doblamos el alcance de la red inalámbrica (ahora la distancia máxima permitida no es entre estaciones, sino entre cada estación y el punto de acceso). En la figura 4 mostramos un ejemplo. Además, los *puntos de acceso* se pueden conectar a otras redes, y en particular a una red fija, con lo cual un usuario puede tener acceso desde su terminal móvil a otros recursos.

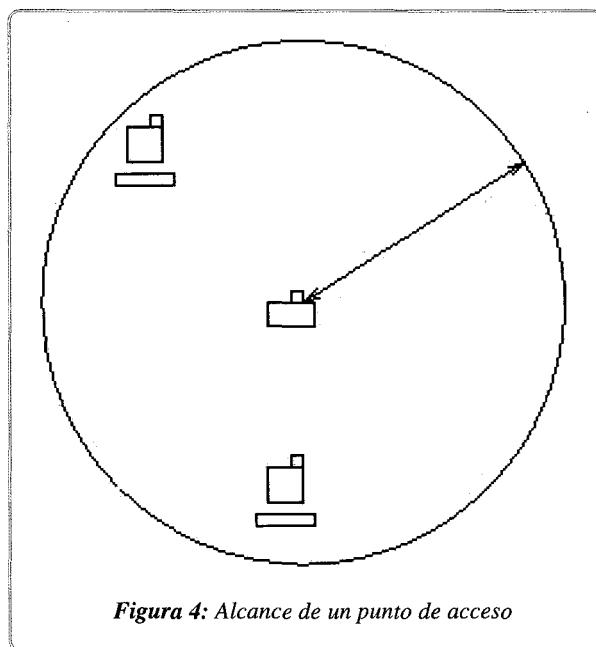


Figura 4: Alcance de un punto de acceso

Para dar cobertura en una zona determinada habrá que instalar varios puntos de acceso de tal manera que podamos cubrir la superficie necesaria con las celdas de cobertura que proporciona cada punto de acceso y ligeramente solapadas para permitir el paso de una celda a otra sin perder la comunicación.

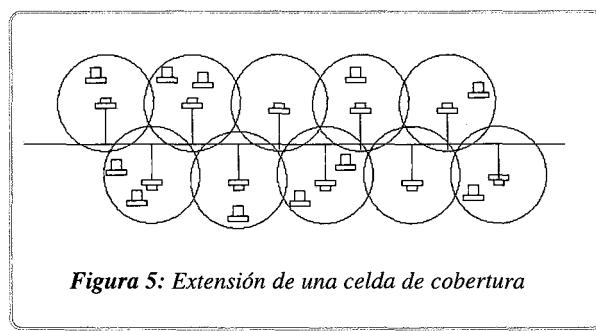


Figura 5: Extensión de una celda de cobertura

C. Enlace entre varias LAN

Para finalizar, otra de las configuraciones de red posibles es la que incluye el uso de antenas direccionales. El objetivo de estas antenas direccionales es el de enlazar redes que se encuentran situadas geográficamente en sitios distintos tal y como se muestra en la figura 6. Un

ejemplo de esta configuración lo tenemos en el caso en que tengamos una red local en un edificio y la queramos extender a otro edificio. Una posible solución a este problema consiste en instalar una antena direccional en cada edificio apuntándose mutuamente. A la vez, cada una de estas antenas está conectada a la red local de su edificio mediante un punto de acceso. De esta manera podemos interconectar las dos redes locales.

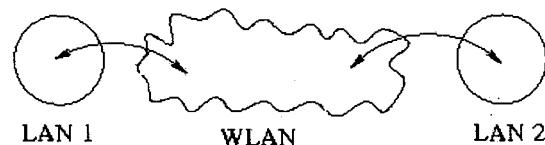


Figura 6: Enlace entre LANs

V. ALGORITMO DE ACCESO MAC

El comité IEEE ha definido un conjunto de estándares para el acceso a las redes de área local. Los diferentes métodos de acceso de la familia IEEE 802 están diseñados según el modelo de referencia OSI y se encuentran ubicados en el nivel físico y en la parte inferior del nivel de enlace o subnivel MAC.

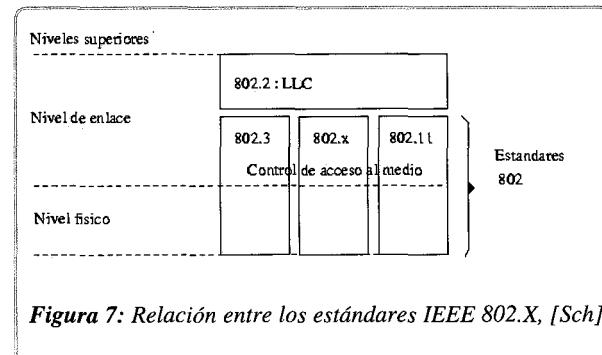


Figura 7: Relación entre los estándares IEEE 802.X, [Sch]

A. Arquitectura del subnivel MAC

La arquitectura MAC del estándar 802.11 se compone de dos funcionalidades básicas: la Función de Coordinación Puntual (PCF) y la Función de Coordinación Distribuida (DCF).

Definimos *función de coordinación* como la funcionalidad que determina, dentro de un conjunto básico de servicios (BSS), cuándo una estación puede transmitir y/o recibir unidades de datos de protocolo a nivel MAC a través del medio inalámbrico.

En el nivel inferior del subnivel MAC se encuentra la función de coordinación distribuida y su funcionamiento se basa en técnicas de acceso aleatorias de contienda por el medio. El tráfico que se transmite bajo esta funcionalidad es de carácter asíncrono ya que estas técnicas de contienda introducen retardos aleatorios y no predecibles no tolerados por los servicios síncronos. El algoritmo básico de acceso a este nivel es muy similar al



implementado en el estándar IEEE 802.3 y es el llamado CSMA/CA (Carrier Sense Multiple Access / Collision Avoidance). Este algoritmo funciona tal y como describimos a continuación:

- 1.- Antes de transmitir información una estación debe testear el medio, o canal inalámbrico, para determinar su estado (libre / ocupado).
- 2.- Si el medio no está ocupado por ninguna otra trama la estación ejecuta una espera adicional llamada *espaciado entre tramas* (IFS).
- 3.- Si durante este intervalo temporal, o bien ya desde el principio, el medio se determina ocupado, entonces la estación debe esperar hasta el final de la transacción actual antes de realizar cualquier acción.
- 4.- Una vez finaliza esta espera debida a la ocupación del medio la estación ejecuta el llamado algoritmo de Backoff, según el cual se determina una espera adicional y aleatoria escogida uniformemente en un intervalo llamado *ventana de contienda* (CW). El algoritmo de Backoff nos da un número aleatorio y entero de ranuras temporales (slot time) y su función es la de reducir la probabilidad de colisión que es máxima cuando varias estaciones están esperando a que el medio quede libre para transmitir.
- 5.- Mientras se ejecuta la espera marcada por el algoritmo de Backoff se continua escuchando el medio de tal manera que si el medio se determina libre durante un tiempo de al menos IFS esta espera va avanzando temporalmente hasta que la estación consume todas las ranuras temporales asignadas. En cambio, si el medio no permanece libre durante un tiempo igual o superior a IFS el algoritmo de Backoff queda suspendido hasta que se cumpla esta condición.

Por encima de la funcionalidad DCF se sitúa la función de coordinación puntual, PCF, asociada a las transmisiones libres de contienda que utilizan técnicas de acceso deterministas. El estándar IEEE 802.11, en concreto, define una técnica de interrogación circular desde el punto de acceso para este nivel. Esta funcionalidad está pensada para servicios de tipo síncrono que no toleran

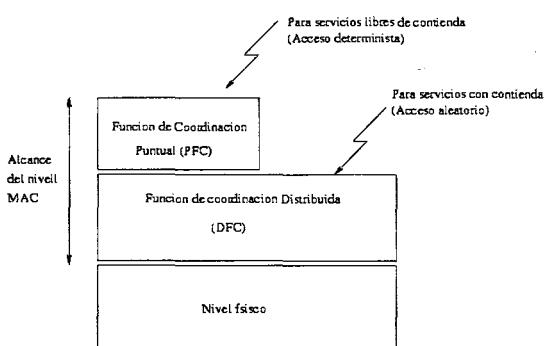


Figura 8: Arquitectura IEEE 802.11 de niveles 1 y 2 [Sta]

retardos aleatorios en el acceso al medio. En la figura 8 mostramos la relación entre estos dos modos de operación.

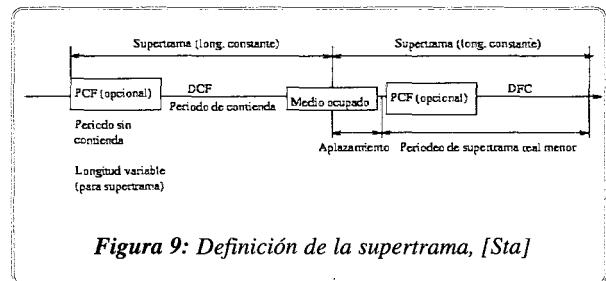


Figura 9: Definición de la supertrama, [Sta]

Estos dos métodos de acceso pueden operar conjuntamente dentro de una misma celda o conjunto básico de servicios dentro de una estructura llamada *supertrama*. Un parte de esta *supertrama* se asigna al periodo de contienda permitiendo al subconjunto de estaciones que lo requieran transmitir bajo mecanismos aleatorios. Una vez finaliza este periodo el punto de acceso toma el medio y se inicia un periodo libre de contienda en el que pueden transmitir el resto de estaciones de la celda que utilizan técnicas deterministas.

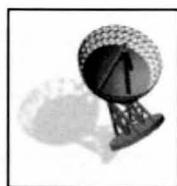
VI. RESUMEN

En este artículo hemos tratado los siguientes puntos:

- Aspectos generales sobre los sistemas WLAN. Hemos dado una definición de lo que entendemos por red de área local inalámbrica, hemos resaltado las aplicaciones típicas y hemos dado una breve reseña histórica.
- Tecnologías utilizadas actualmente en la implementación de las WLAN. En concreto nos hemos centrado en las tecnologías que se especifican en el estándar IEEE 802.11 y que son: las tecnologías DSSS y FHSS de espectro ensanchado y la tecnología de infrarrojos.
- Otro de los puntos que hemos tratado ha sido la arquitectura de los sistemas WLAN y las configuraciones que podemos encontrar en una red de este tipo: redes de tipo ad-hoc, extensión de celdas básicas y enlace entre varias LANs.
- Finalmente hemos descrito el subnivel MAC del nivel de enlace, i en concreto el algoritmo de acceso al medio CSMA/CA.

VII. REFERENCIAS

- [I11] IEEE 802.11, Wireless LAN MAC and Physical Layer Specifications. Editors of IEEE, juny 1997.
- [Mar] Michael M. Martin, Wireless Local Area Networks, <http://www.csm.ohio.edu/comt391w/mmartin/final.html>
- [Sch] Mischa Schwartz, Redes de Telecomunicaciones, Ed. Addison-Wesley Iberoamericana, México 1994
- [Sta] William Stallings, Comunicaciones y redes de computadoras, Prentice Hall, 1997
- [WLA] Wireless LAN Alliance, <http://www.wlana.com>



COSPAS-SARSAT

UN SISTEMA PARA LA LOCALIZACIÓN DE ACCIDENTES VÍA SATÉLITE

Humberto J. Ayala Cabrera.

Chairman de la Rama de estudiantes del IEEE de Canarias.

hjac@cic.teleco.ulpgc.es, hjac@ieee.org

1.- ANTECEDENTES

El 1º de Julio de 1988 se firmó en París el ACUERDO INTERNACIONAL DEL PROGRAMA COSPAS-SARSAT.

las señales de socorro emitidas por radiobalizas en situaciones de emergencia.

A través de este Acuerdo se invitó a participar a cuantos Países desearan cooperar con aportación de Segmento Terrestre o bien como Estado Usuario del Sistema.



Figura1. Logotipos de la Cospas-Sarsat y de la estación de Maspalomas

Este acuerdo fué suscrito por los Gobiernos de Canadá, Francia, URSS y USA, comprometiéndose por quince años y períodos sucesivos de 5 años de prolongación, a mantener una constelación de satélites en órbita polar baja para la retransmisión a Estaciones Terrenas de

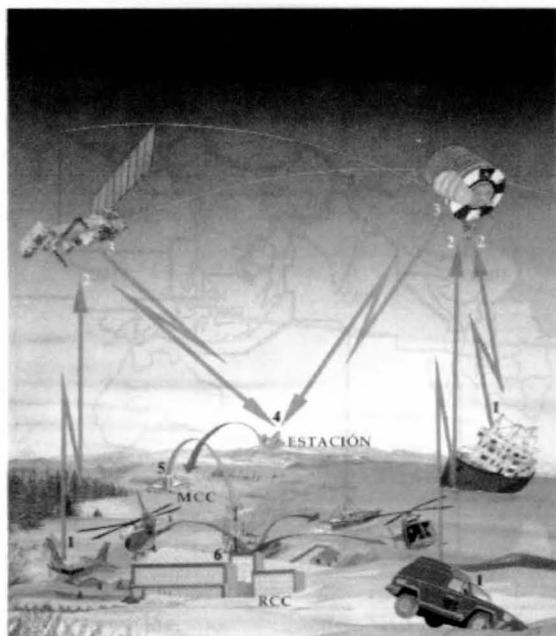


Figura2. Ejemplo de localización de accidentes



Figura3. Satélites en órbita alrededor de la Tierra

Con fecha del 31 de Mayo de 1991 el Consejo del Ministerio del Gobierno Español aprueba la incorporación de España al Programa COSPAS-SARSAT en calidad de Proveedor de Segmento Terreno, nombrándose al INTA (Instituto Nacional de Técnica Aeroespacial) como el organismo responsable de cumplir los compromisos adquiridos. El 1º de Enero de 1993 la Estación Espacial de Maspalomas del INTA inició oficialmente su actividad en el Programa Internacional COSPAS-SARSAT: un programa en el que actualmente, colaboran un total de 32 países de todo el mundo.



2.-DESCRIPCIÓN DEL PROGRAMA

Los primeros cuatro países mantienen actualmente seis satélites de órbita polar, dos rusos (COSPAS) y cuatro norteamericanos (SARSAT), girando alrededor de la Tierra y retransmitiendo las señales de socorro que puedan ser emitidas por buques, aviones o expediciones cualesquiera en situación de emergencia en una región aislada.

Estas señales se transmiten en la frecuencia de 406 MHz desde radiobalizas especialmente diseñadas para esta función, indicando en su mensaje, no solo la propia

nal), su País de origen y la propia identificación con que ha sido registrada en dicho País.

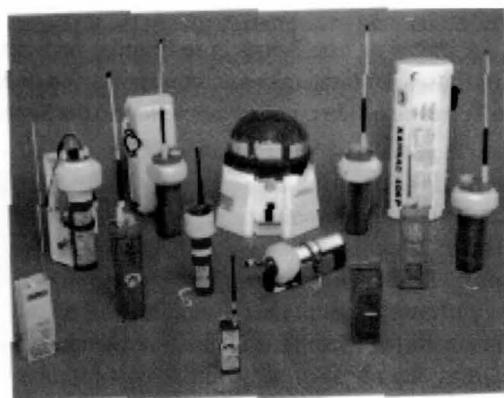


Figura 4. Balizas obsoletas

identificación, sino otro tipo de información que puede ser crítica a la hora de realizar las labores de rescate.

Todas ellas transmiten un mensaje de alerta que indica el tipo de usuario (aeronáutica, marítima o perso-

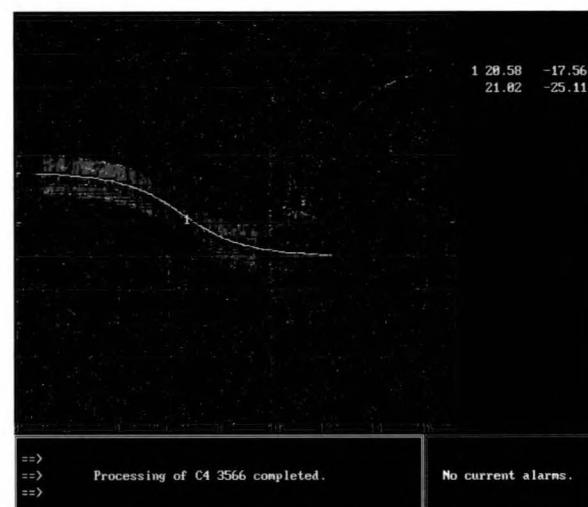


Figura 6. Detalle del proceso de detección de alarmas

Otras balizas ya obsoletas emiten en 121.5 MHz, no dan ningún tipo de mensaje, dan gran número de falsas alertas y la precisión en su localización no alcanza los 25 Km, por lo que no son recomendables y el Sistema Cospas-Sarsat dejará de darles servicio a partir del año 2008.

La Estación de Maspalomas recibe esta señal en su estación receptora (LUT), calcula la posición del siniestro mediante el uso de técnicas doppler y el conocimiento preciso de la posición instantánea del satélite y desde su MCC (Centro de Control de la Misión) da la alerta a los

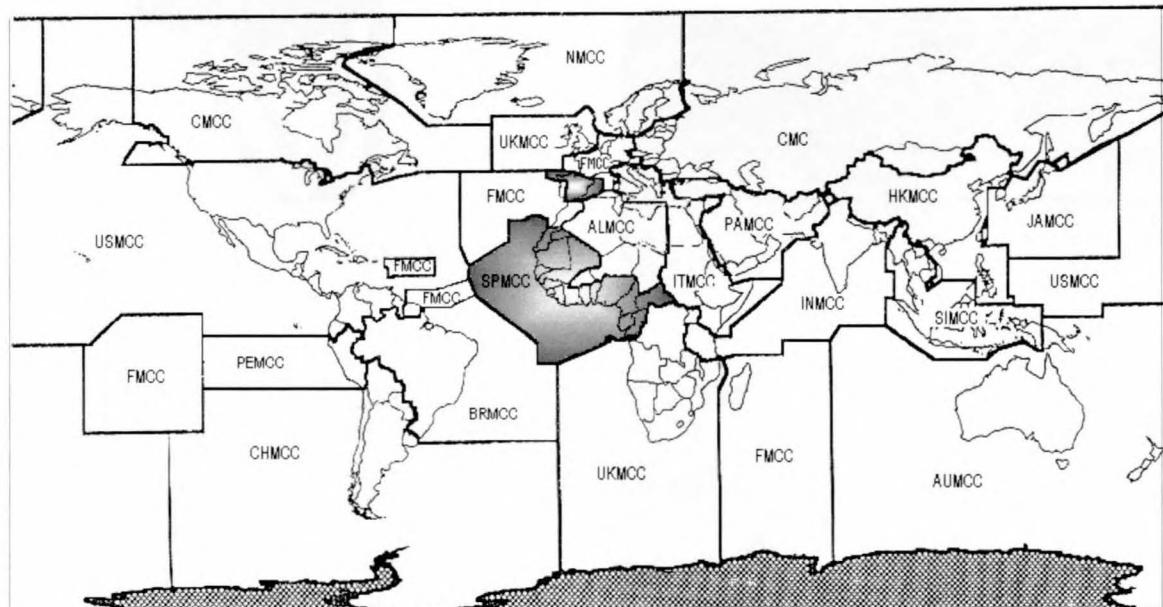


Figura 5. Radio de acción del SPMCC

servicios de rescate dentro de los cinco minutos siguientes al pase del satélite.

Esta operación es igualmente llevada a cabo desde 19 Centros Coordinadores de la Misión COSPAS-SARSAT pertenecientes a otros tantos Países dentro de su



Figura 7. Estación de Maspalomas

área de visibilidad. La distribución de alertas entre MCCs se hace de forma coordinada para lo cual el mundo se ha dividido en 4 regiones, existiendo un MCC nodal en cada

rescate Españoles y a los de otros 21 países africanos como se observa en el mapa (Figura1).

La Estación de Maspalomas, como sede del LUT (Local User Terminal) y del SPMCC (Spanish Mission Control Center), cuenta con un equipo de siete ingenieros expertos para efectuar los mantenimientos y las operaciones necesarias de seguimiento, recepción de las señales de emergencia, procesado de las mismas, localización de la posición del accidente, y envío del mensaje de alerta a los servicios de rescate del área en la que ha sido localizado el siniestro. Este servicio está cubierto durante las 24 horas de todos los días del año. Además, este mismo equipo humano, bajo la Dirección del Jefe del SPMCC, realiza todas las tareas de desarrollo de los planes de test, adaptación y actualización, análisis del sistema, etc.

Dentro del mismo programa, se cuenta de forma ya operativa con un innovador sistema de recepción de señales de emergencia de radiobalizas vía satélites geostacionarios (GOES, INSAT), que se prevé como sistema futuro sustitutivo o complementario de los satélites de órbita polar cuando la mayoría de radiobalizas de 406 MHz queden dotadas de receptores GPS integrados, de manera que la posición quede automáticamente insertada en el propio mensaje de socorro transmitido, sin necesidad de tener que calcularla por efecto Doppler en la estación receptora. En este campo España, desde Maspalomas, ejerce una función de liderazgo en el Siste-

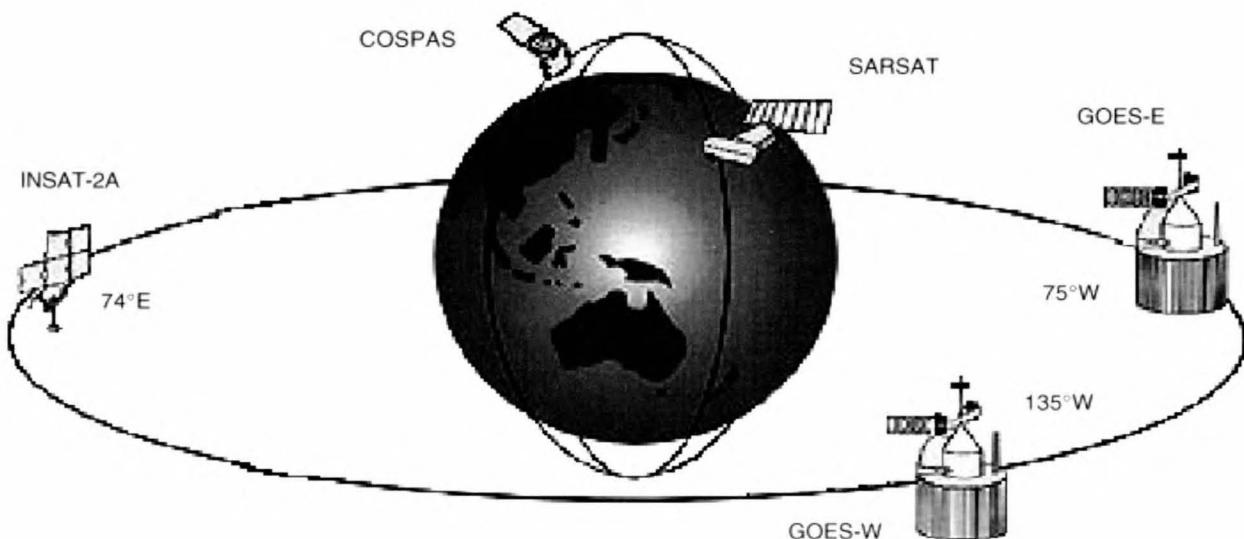


Figura 8. Satélites geostacionarios (GOES y INSAT)

una a través del cual se rutan los mensajes entre regiones. Europa y África conforman la Región Central de la cual Francia asume la responsabilidad de MCC nodal y España (SPMCC) le sustituye en caso de anomalía.

El SPMCC es también responsable directo de la transmisión de los mensajes de alerta a los servicios de

ma GEOSAR, pues únicamente disponen de él otros 4 países.

Con carácter de exclusividad (primer aviso de socorro a los servicios de salvamento), este Programa ha contribuido desde sus inicios (1983) al salvamento de más

de 9.000 vidas humanas que no tuvieron otro medio para alertar de su siniestro que el Sistema COSPAS-SARSAT.

internacionales que se alcancen en los distintos Organos Decisorios, participación en reuniones técnicas, operativas, etc.

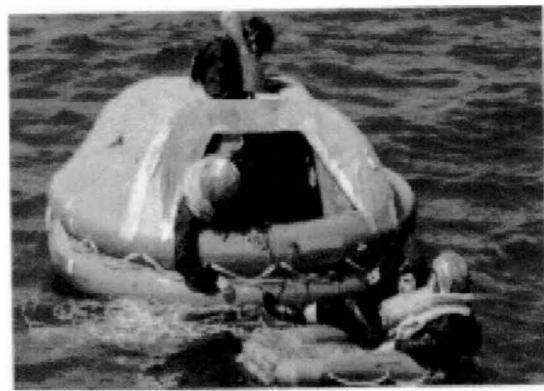


Figura 9. Primer aviso de socorro a los servicios de salvamento

Maspalomas da un soporte crucial importantísimo por cubrir un área hasta ahora desprotegida en la zona del Noroeste Africano y Atlántico Central. Su cobertura, para los satélites de órbita polar, alcanza 3.500 km de radio: del Norte de Irlanda hasta pasado el Ecuador y desde mitad del mar Mediterráneo hasta cerca de las costas del Brasil.

La incorporación de España a este Programa ha sido llevada con ejemplar eficacia desde Canarias, correspondiendo al INTA la responsabilidad de que nuestro País cumpla todos los compromisos que su adhesión al Programa exige: como es la Regulación del Uso de las radiobalizas, Registro de radiobalizas, la eliminación de interferencias, dar el servicio de vigilancia satelitalia de 24 horas, alertar a los Servicios de Salvamento de cualquier país donde se produzca el siniestro, cumplimentación de los acuerdos

Desde Enero de 1993 hasta Abril de 1999, y bajo la consideración de primer ó único mensaje de alerta, el Sistema Español tiene en su haber la contribución al salvamento de 137 vidas humanas en zona de rescate de responsabilidad española, por su intervención en 31 accidentes dando la alarma a los Servicios del SAR del Ejército del Aire o al Centro Nacional de Coordinación de Salvamento de la Marina Mercante Española. Otras 314 personas fueron rescatadas en esta misma área, en 86 emergencias en las que la Estación de Maspalomas prestó su apoyo. Y su balance concluye a través de su participación complementaria en un total de 297 siniestros aéreos y marítimos en otras partes del Mundo hasta un total de casi 2000 vidas humanas rescatadas.

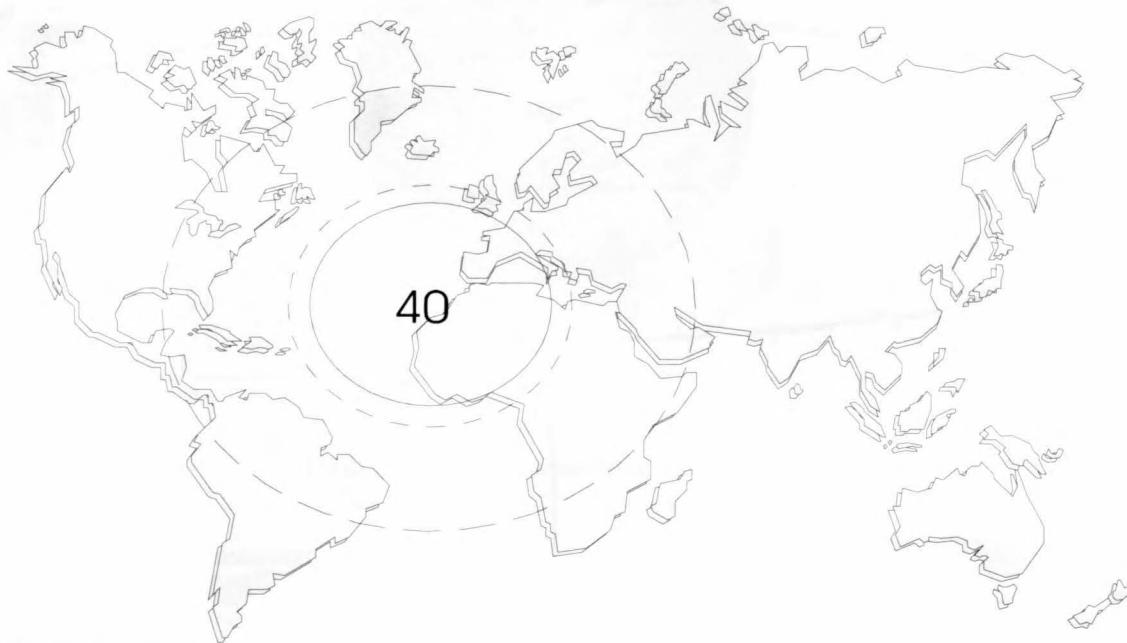


Figura 10. Radio de acción de Maspalomas



1er CONGRESO ESTATAL DE RAMAS DE ESTUDIANTES DEL IEEE

José Antonio López Salcedo

Estudiante de la ETSETB y Miembro de la Rama de Estudiantes del IEEE de Barcelona
j.lopez@ieee.org

I. LA HISTORIA

Todos nosotros, pero fundamentalmente los que editamos esta revista, recordamos de manera especial el momento en que el último número de Buran (el número 13) veía la luz allá por el mes de abril.

Sin duda, la aparición de cada nuevo número de Buran supone una gran satisfacción personal, al ver finalmente reflejado el producto de varios meses de trabajo (no sin antes haber superado innumerables problemas y obstáculos) pero qué duda cabe que mayor satisfacción aún nos produjo el hecho de ser partícipes, en primera persona, de un hecho sin precedentes en la historia de las Ramas de Estudiantes del IEEE en España: el 1er Congreso Estatal de Ramas de Estudiantes del IEEE.

La idea en sí había comenzado ya a gestarse durante la asistencia el verano de 1998 al Student Branch Congress en Estambul, al que tuve el placer de asistir como Secretario de la Rama de Barcelona, acompañando a nuestro Presidente, Miguel Ángel Sastre. Aquel congreso supuso a su vez, el primer encuentro entre Ramas de Estudiantes del IEEE de la Región 8 (Europa, África y Medio Oriente) y allí, de entre la comitiva española representada por miembros de las Ramas de Barcelona, Canarias, Madrid y Valencia, surgió el comentario sobre la posibilidad de organizar un encuentro entre representantes de estudiantes del IEEE, al estilo del que estábamos asistiendo, pero enmarcado en España.

En realidad, aquella era también la primera vez que miembros de Ramas del IEEE de España se encontraban, teniendo la oportunidad de conocerse en persona y de poder tratar de manera directa sobre temas comunes relacionados con todo lo que supone llevar adelante una asociación de estudiantes como una Rama de Estudiantes del IEEE. Fue por ello que, junto con el hecho de que nuestra Rama de Barcelona celebraría al año siguiente (el ahora actual 1999) su 20 aniversario, se propuso celebrar un congreso similar en Barcelona para mediados del mes de febrero.

El congreso en Turquía finalizó, nos despedimos todos, pero volvimos a casa con la convicción de que algún día nos volveríamos a ver.

II. PREPARATIVOS

A decir verdad, no fue en este mes de febrero cuando se organizó el esperado congreso de Ramas españolas, sino que fue entonces cuando, y producto de los problemas habituales que solemos tener aquellos que compaginamos nuestra actividad de estudio de una carrera técnica con la colaboración en una asociación de estudiantes, que se comenzó a organizar.

Lo que había comenzado meses atrás como un simple comentario, se había convertido ya a esas alturas en un reto que había que conseguir, con el condicionante de que nadie antes lo había hecho en España.

El primer paso fue contactar con las diferentes Ramas del IEEE en España, rescatar las direcciones de contacto que teníamos, y ponernos en contacto con los interesados para hacerles saber que el "Congreso" ya estaba en marcha.



Logos de la Rama de Estudiantes del IEEE y de la Universidad de Las Palmas de Gran Canaria.



Logo de la Rama de Estudiantes del IEEE de Madrid.

Por otro lado, las fechas definitivas de su realización, que serían el 29, 30 de abril y 1 de mayo de 1999, se eligieron de manera que el Congreso coincidiera con la fiesta más multitudinaria de nuestra Escola Tècnica Superior d'Enginyeria de Telecommunicació de Barcelona, la *Telekogresca*, de manera que los asistentes pudieran no sólo asistir a un congreso sino también pasar una buena noche de fiesta.



III. EL CONGRESO

Justo el día anterior acabamos de recibir las últimas confirmaciones por parte de algunos asistentes, y por aquellos momentos nos empezamos a dar cuenta de lo que realmente teníamos entre manos.

Los representantes de Las Palmas de Gran Canaria fueron los primeros en llegar, y luego les fueron siguiendo los de Madrid, Málaga y finalmente Valencia.

Tras el control de asistencia y la entrega de acreditaciones, aquella misma tarde del jueves 29 de abril de 1999 quedó inaugurado el 1er Congreso Estatal de Ramas de Estudiantes del IEEE, en un acto que tuvo lugar en el Aula Master del Campus Nord de la UPC, y al cual asistieron como invitados, el director de nuestra escuela, Sr. Antoni Elias, y el *counselor* de la Rama de Estudiantes del IEEE de Barcelona, Sr. Ángel Cardama.

Ellos fueron los encargados de dar oficialmente la bienvenida a Barcelona, y a este Congreso, a los representantes de las Ramas de Estudiantes del IEEE de las principales Escuelas Técnicas Superiores de Ingeniería de Telecomunicación de España, que aquí se encontraban reunidos.

El día siguiente, día central del Congreso, tuvieron lugar las presentaciones de las Ramas asistentes, tratando cada una sobre su historia (algunas como la Rama de Málaga, recién creadas), las actividades que organizan, sus logros, sus proyectos y también sus problemas, los que mediante este Congreso se pretendía dar a conocer para intentar buscar entre todos vías de solución.



Instantes previos a la presentación de la Rama de Estudiantes del IEEE de Las Palmas de Gran Canaria.

Uno de los puntos de más interés de estas presentaciones fue las actividades que realizan las diferentes Ramas, ya que uno de los objetivos de este Congreso era precisamente dar a conocer estas actividades, de manera que así, cursos y seminarios (por ejemplo) impartidos en una cierta Rama, pudieran ser también

organizados en otras, existiendo una colaboración mutua entre ambos organizadores.

La tarde sin embargo, dio paso a las tertulias y mesas redondas. Uno de los temas principales de aquellas reuniones fue sin duda Buran, ya que, a pesar de que esta revista fue creada en 1993 por la Rama de Barcelona, y desde entonces se ha venido editando y publicando aquí (con algunas colaboraciones puntuales de las Ramas de Valencia y Canarias), pretende ser a partir de ahora un medio en el que todas las Ramas del IEEE de España puedan tener su propia voz. Claro está que Buran nació con la vocación de ser una revista hecha por y para estudiantes, y de esta manera es lógico que tengan cabida no sólo los estudiantes del Campus Nord, o de los diferentes puntos de distribución de la Rama de Barcelona en Cataluña, sinó los de todo el resto de Escuelas Técnicas de Ingeniería de Telecomunicación de España, y especialmente, aquellas en donde está presente una Rama de Estudiantes del IEEE.

Es por ello que, con un poco de suerte, esperamos que en el próximo número de Buran se pueda ya apreciar este cambio, y sean muchas más las colaboraciones que se reciban de alumnos de otras Escuelas Técnicas, a través de colaboraciones con sus Ramas de Estudiantes del IEEE.

Esa misma tarde además, pudimos asistir a la Exposición Virtual que sobre la Agencia Espacial Europea, habían organizado nuestros compañeros de AEES Estudiantes. Allí pudimos contemplar algunos de los últimos proyectos de la ESA, un simulador de vuelo y varias maquetas de los satélites y lanzaderas utilizadas.

Finalmente, ya caída la tarde, se dieron por concluidas las reuniones, y nos emplazamos para la *Telekogresca*, que tuvo lugar aquella misma tarde y dio por clausurado el Congreso.

Con la promesa de un futuro reencuentro en la que sería la segunda edición de este Congreso, posiblemente organizado por la Rama de Valencia, nos despedimos



Visita a la Exposición sobre la Agencia Espacial Europea (ESA), organizada por nuestros compañeros del Capítulo de AEES Estudiants.

todos para, al día siguiente, tomar cada uno su camino de vuelta a casa.

A nosotros, y en nombre de la Rama de Estudiantes del IEÉE de Barcelona, tan sólo nos queda la satisfacción de haber hecho el trabajo lo mejor que supimos, y la ilusión de que todo este esfuerzo que un día hicimos para llevar este Congreso adelante, no fuera en vano sino que sirviera para que en un futuro, hubiera continuidad. Ésa sería la mejor muestra de que el Congreso fue realmente un éxito.

IV. AGRADECIMIENTOS

Aprovechamos desde aquí para agradecerle una vez más su presencia al director de nuestra escuela, y especialmente a nuestro *counselor*, Sr. Ángel Cardama, el cual tras el acto de inauguración, tuvo la amabilidad de guiarnos a través de una visita por las principales instalaciones técnicas y de investigación del Departamento de Teoría de la Señal y Comunicaciones (TSC), y del Departamento de Arquitectura de Computadores (DAC).

Por lo que respecta al departamento de TSC, pudimos contemplar la Sala Anecoica, el Estudio de Grabación y laboratorios donde se realizan ensayos con laser y pruebas holográficas, mientras que por lo que respecta al DAC, descendimos a las salas donde se encuentran algunos equipos multiprocesador (de hasta 128 procesadores) y sistemas de última generación.



Entrada a la Sala Anecoica de la ETSETB.

No quisiéramos tampoco olvidar al Sr. Manel Domínguez, el cual, y de manera desinteresada, nos ofreció una visita a la "Sala Blanca" del Departamento de Electrónica.

Y por supuesto, agradecer a todos aquellos que dieron una parte de su tiempo para hacer que este Congreso fuera posible.



De izquierda a derecha y de arriba a abajo, asistentes y Rama a la que pertenecen: Luis Miguel Hernanz (Madrid), José Cástor Vallés (Barcelona), José Antonio López Salcedo (Barcelona), Rafael Gallego (Barcelona), María Dolores Carreño (Madrid), Juan Antonio Martínez (Valencia), Ángel León (Las Palmas), Miguel Ángel Sastre (Barcelona), Javier Muñoz (Málaga), Daniel Prado (Barcelona), Marta Castillo (Barcelona), Víctor Canero (Las Palmas) y Pablo García (Málaga).





PREMIO AT&T PARA LA RAMA DE ESTUDIANTES DEL IEEE DE MADRID

Maria Dolores Carreño Ruiz y Luis Miguel Hernanz Iglesias

Estudiantes de la ETSIT de la UPM.

dolores@ieeesb.etsit.upm.es luish@ieeesb.etsit.upm.es

«La primera sorpresa que te da la electrónica cuando la practicas es la carestía de los componentes y las dificultades que tienes que afrontar cuando diseñas e implementas tu primer prototipo. Los estudiantes no tienen tiempo ni dinero para superar estos obstáculos».

Así comienza la propuesta titulada «**Versatile and autonomous programmable logic training board**» que los autores de este artículo presentamos al concurso «AT&T Labs Student Enterprise Award» y que resultó ser agraciada con uno de los premios.

EL PREMIO

El «AT&T Labs Student Enterprise Award» es un concurso que anualmente organiza la compañía de telecomunicaciones AT&T en colaboración con el IEEE. En él tienen cabida todas las ramas de estudiantes del mundo. Su objetivo es dar a los estudiantes del IEEE la oportunidad de llevar a cabo sus proyectos aportando la financiación necesaria para los mismos, así como motivar y/o fortalecer las actividades de las ramas, fomentando la elaboración de sus propuestas. Los proyectos pueden ser técnicos o no, con temáticas que varían desde la investigación en la tecnología más avanzada al estudio de servicios para la comunidad.

A la hora de valorar los proyectos presentados, no sólo se tiene en cuenta la dificultad o la originalidad del proyecto en sí, sino que también son muy importantes los «aspectos sociales» del mismo: cuánta gente se verá involucrada en su desarrollo (y cuánto aprenderán con ello) y a cuánta gente beneficiará una vez que el proyecto se lleve a cabo.

Cinco ramas son las ganadoras anualmente, y como antes se ha mencionado, reciben como premio una determinada cantidad de dinero y el reconocimiento por parte de la comunidad. En la memoria que los participantes envían de su proyecto, ha de constar un presupuesto estimado del gasto total que supondrá la realización de su idea, hasta un tope de \$1000. Luego, la comisión del jurado será la que adjudique la suma que considere oportuna a los distintos ganadores.

EL PROYECTO

En la última edición de estos premios (98-99), la rama de estudiantes de la Escuela Técnica Superior de

Ingenieros de Telecomunicación de la Universidad Politécnica de Madrid fue una de las galardonadas. Tres estudiantes de cuarto curso unimos nuestra creatividad y nuestras ganas de trabajo para desarrollar una «placa entrenadora autónoma y versátil basada en lógica programable». Nuestros nombres son Mª Dolores Carreño Ruiz, Luis Miguel Hernanz Iglesias y Álvaro Barge Cebrián.



Estos son los protagonistas, de izquierda a derecha: María Dolores Carreño Ruiz, Álvaro Barge Cebrián, Javier Macías Guarasa (counselor de la Rama de Estudiantes del IEEE de Madrid), y Luis Miguel Hernanz Iglesias.

¿QUÉ NOS IMPULSÓ?

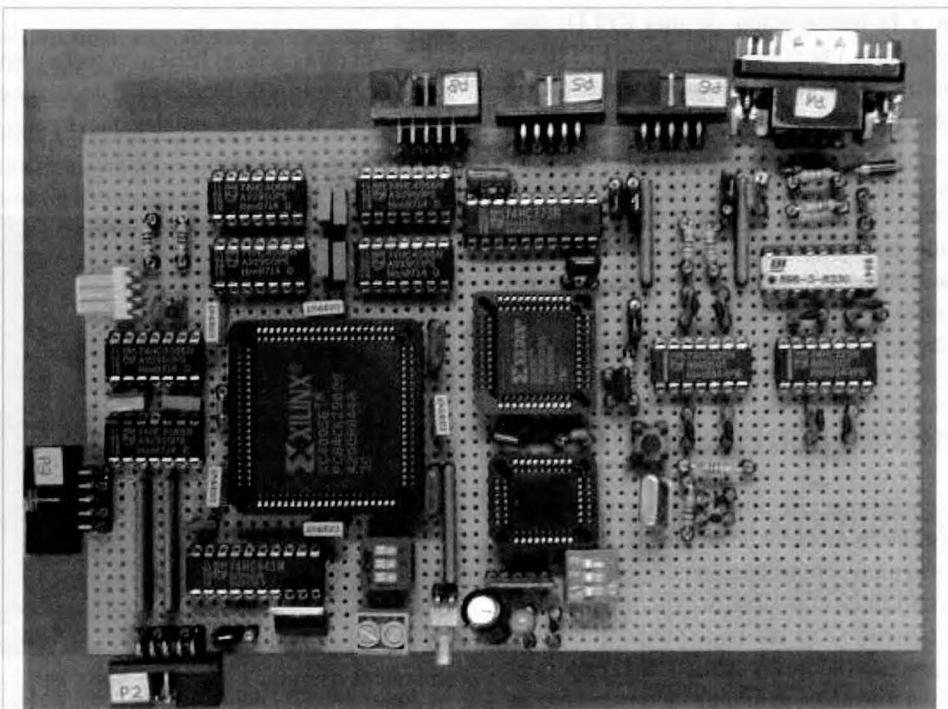
En los laboratorios de electrónica de la escuela los estudiantes se encuentran con que han de desembolsar una cantidad importante de dinero comprando componentes que sólo van a servirles durante el laboratorio. Además, pasan gran parte del tiempo revisando los cientos de conexiones que implican los diseños a realizar, en lugar de abordar diseños más avanzados. La falta de medios de los laboratorios aleja a los estudiantes de las técnicas de diseño electrónico actual (montaje superficial, por ejemplo) y obligan a los alumnos a trabajar con chips ya anticuados que sólo permiten crear circuitos de moderada complejidad.

Para lograr el doble objetivo de usar técnicas modernas y mantener un bajo coste, se nos ocurrió que era posible recurrir a la lógica programable. Esta tecnología, que tiene muy pocos años en el mercado, está basada en chips que internamente constan de miles de puertas lógicas, cuyas interconexiones se programan por el usuario con la ayuda de un programa específico. Por tanto, la funcionalidad del chip no está prefijada desde la fábrica, sino que cada usuario la programa todas las veces que sean necesarias. Esto permite realizar los diseños de los circuitos en un ordenador, utilizando todas las técnicas de diseño existentes (diagramas de estados, esquemáticos, lenguajes de descripción hardware), y luego mediante un sencillo proceso de programación lograr que el chip implemente el circuito.

El problema fundamental si se quiere que esta tecnología fuese fácilmente accesible por los alumnos del laboratorio, era necesario disponer de una placa que ofreciese facilidades para programar el chip e incluir otra funcionalidad (por ejemplo que la placa pudiese guardar varios circuitos), así como para proteger a los propios

LA PLACA ENTRENADORA

Consiste fundamentalmente en una FPGA (Field Programmable Gate Array) que es el corazón de todo el sistema. Internamente contiene un elevado número de puertas lógicas, multiplexores y buffers agrupados en bloques lógicos que se interconectan convenientemente cuando es configurada. Nosotros hemos trabajado con productos de la compañía Xilinx, pionera en muchos campos de la lógica programable. El modelo de FPGA que nosotros empleamos permite realizar circuitos de hasta 5000 puertas lógicas, pero puede ser sustituido por otro de mayor capacidad para albergar diseños más grandes. Este tipo de dispositivos permiten ser reconfigurados muchas veces con potentes diseños digitales. Estos diseños se realizan con un software específico ligado a la marca de la FPGA. En el PC se puede validar el funcionamiento del diseño, y posteriormente descargarlo en la FPGA mediante un cable basado en el estándar JTAG. En nuestro caso lo hemos variado ligeramente para incrementar su funcionalidad. Como la placa está pensada para usarse en los más variopintos diseños, se ha cuidado especialmente el



En esta imagen se puede ver la placa entrenadora autónoma y versátil basada en lógica programable, con un corazón FPGA que permite realizar circuitos de hasta 5000 puertas lógicas.

chips de las manos de los estudiantes (¿quién no ha quemado alguna vez un chip porque conectó mal algún cable?).

Con esta idea en mente, presentamos una práctica especial para uno de los laboratorios de nuestra escuela, y como el proyecto se fue liando, haciéndose cada vez más ambicioso, decidimos presentarlo al concurso de AT&T. El resto es historia...

que ninguna de las posibilidades del chip quedase recortada por la estructura propia de la placa entrenadora. Por tanto, todos los pines de la FGPA son accesibles desde el exterior a través de unos puertos de 10 contactos muy cómodos de usar.

El segundo elemento importante de la placa es una memoria FLASH que permite, entre otras cosas, que la placa pueda funcionar independientemente de un ordenador. Esto es así porque esta memoria la destinamos para

guardar archivos de configuración de la FPGA. Nuestra FPGA está basada en memoria RAM, por lo que pierde su configuración cuando se apaga la alimentación. El contar con la memoria, permite que la FPGA vuelva a reconfigurarse leyendo la información necesaria de la misma. Además el tamaño de la memoria nos permite tener varios programas almacenados, hasta ocho circuitos, en una misma placa y cambiar de uno a otro con ayuda de unos pequeños switches. Incluso también es posible que la propia placa se reinicie de forma automática y decida qué programa quiere cargar, sin necesidad de ninguna ayuda externa. Esto permite una suerte de reconfiguración dinámica que hace posible que la placa elija el circuito más adecuado para la situación actual.

También hemos incluido en la placa la lógica necesaria para interactuar con un microcontrolador Motorola 68HC11, muy extendido en nuestra escuela. Es posible que el micro acceda a la memoria e intercambie datos con la FPGA de forma sencilla. La cooperación entre un microcontrolador y una FPGA, puede permitir la realización de funcionalidad muy complicada, pues nos ofrece la versatilidad del software y la fuerza bruta de velocidad del hardware. Esta interesante funcionalidad es posible gracias a la intervención de una CPLD, otro dispositivo programable, que ponestar basado en EEPROM mantiene su configuración al retirarle la alimentación a la placa. Concretamente se trata de la XC9536 de Xilinx.

Para hacer más fácil la interacción entre la placa y el ordenador, hemos creado nuestros propios programas, que permiten una gestión total de la programación de la placa, así como de los contenidos almacenados en la memoria de la misma. Con ellos, se puede borrar, grabar o recuperar datos de la memoria, eligiendo qué programas de la FPGA hay en cada sector de la misma.

EL FUTURO DEL PROYECTO

Cuando se publique este artículo esperamos haber recibido ya los primeros modelos de las placas producidos en serie. Tras hacer unas comprobaciones, este mismo año se empezarán a introducir con carácter experimental en el laboratorio, permitiendo a un reducido número de alumnos desarrollar prácticas sobre las mismas. Si la experiencia resulta positiva y no aparece ningún inconveniente, se fabricarán más placas para permitir que cada vez más alumnos puedan acceder a esta tecnología.

Esperamos que con la ayuda de los futuros alumnos del laboratorio, se vayan corrigiendo los errores que pueda haber en el diseño actual o en los programas que la acompañan, y que vayan añadiendo con su trabajo más posibilidades a la misma, bien construyendo otras placas accesorias, bien diseñando circuitos que añadan funcionalidad a la FPGA.

En un futuro lejano, puede que alguna otra universidad se interese por contar con estas placas para sus laboratorios de electrónica.

En otro orden de cosas, esperamos que este premio sirva como guía al resto de las ramas de estudiantes del IEEE de España (y también a la nuestra) para favorecer la aparición de numerosos proyectos que se presenten en sucesivos años al concurso. Existen numerosas oportunidades de que se reconozca nuestro trabajo y muchas veces las dejamos pasar por desconocimiento o pereza. Es muy gratificante ver como tu esfuerzo es reconocido y nos consta que en nuestro país hay muy buena «materia prima». Tan sólo hay que animarse y ponerse manos a la obra.

RECONOCIMIENTO

Cuando se supo la buena nueva de la concesión del galardón, fueron varios los medios de comunicación locales y nacionales que se interesaron en conocer más detalles del premio, lo cual no dejó de sorprendernos gratamente (no en vano hemos sido los primeros europeos que han recibido este premio).

Nuevamente desde aquí animamos a nuestros compañeros a que con ilusión y muchas ganas, se pongan manos a la obra; seguro que no somos los últimos.

Por otra parte, el IEEE nos recompensó ofreciendo a uno de nosotros la oportunidad de asistir como invitado a una de las reuniones anuales que organiza.



Javier Macías Guarasa, junto con Magdalena Salazar, recibiendo el premio de manos de Levent Onural

AGRADECIMIENTO

Deseamos agradecer a nuestro counselor Javier Macías Guarasa su inestimable colaboración y el incansable apoyo que nos ha prestado durante todos estos meses.

Fue él precisamente el que recogió de manos de Levent Onural (IEEE Student Activities Chair) el diploma que AT&T e IEEE concedieron a la rama de Madrid, y que le fue entregado durante el pasado «R8 Committee Meeting» celebrado en Berlín, en Mayo de 1999.



PRESENTACIÓN DE LA RAMA DE ESTUDIANTES DEL I.E.E.E. DE MÁLAGA

Pablo García Vázquez

Presidente de la Rama de Estudiantes del IEEE de Málaga

Nos es grato dirigirnos a los lectores de «Buran» con el fin de presentar la Rama de Estudiantes del I.E.E.E. de Málaga en nombre de cada uno de sus integrantes.

Todo comenzó en Marzo de 1998, cuando dos estudiantes de la E.T.S. de Ingenieros de Telecomunicación de la Universidad de Málaga creyeron que sería una gran idea el establecimiento de una asociación de estudiantes que, bajo la ayuda del I.E.E.E., desempeñara localmente actividades de difusión científica y tecnológica.

Queremos reseñar y agradecer la valiosa ayuda que nos brindó la Rama Universitaria del I.E.E.E. de Madrid en todo lo referente a los procesos previos que había que seguir para que, al fin, el ocho de Marzo de 1999, el I.E.E.E. reconociera a un grupo de 54 estudiantes miembros como Rama de Estudiantes. De igual modo, también resultó fundamental la colaboración prestada por los Departamentos de Tecnología Electrónica y de Ingeniería de Comunicaciones de la Universidad de Málaga, y por la Dirección de nuestra Escuela.

Una vez superados los obstáculos iniciales, el número de miembros se ha incrementado hasta alcanzar 67 estudiantes en la actualidad, cantidad que se estima que va a aumentar en breve considerablemente, debido a la gran aceptación que este proyecto ha tenido tanto entre alumnos como profesores.

En pasadas fechas tuvimos la oportunidad de presentar nuestros planes en el Primer Congreso Estatal de Ramas de Estudiantes del I.E.E.E., celebrado en la E.T.S.E.T.B. de la UPC a finales de Abril y principios de Mayo, y también en la Reunión de la Sección Española que tuvo lugar en Madrid el pasado dos de Julio.

Nuestra premisa de mirar hacia delante ha supuesto la puesta en marcha de actividades que enriquezcan la formación recibida mientras dure nuestro estudio, saliendo un poco de la rutina que puedan suponer las clases, y recibiendo una formación adicional en aquellos temas científicos y tecnológicos que resulten de mayor interés. De esta forma se han constituido grupos de interés específico:

- 1.- Astronomía
- 2.- Audio
- 3.- Bioingeniería
- 4.- Electrónica
- 5.- Ingeniería Software
- 6.- Robótica

Entre las actividades que en breve desarrollaremos se encuentran:

- Ciclo de Conferencias sobre «Inventos y Descubrimientos del siglo XX».
- Seminario de Programación en C++.
- Curso de Audio Digital.
- Conferencia sobre Astronomía impartida por Juan Pérez Mercader, astrofísico de gran prestigio internacional, y que en la actualidad ocupa entre otros cargos el de director del I.N.T.A.

Mediante la creación de nuestra Asociación queremos manifestar la idea de progreso y de desarrollo que día a día tratamos de reflejar en nuestra actitud y pensamiento.

Emplazamos a todos los interesados en conocer cuanto aquí desarrollamos a que nos visiten. Estáis todos invitados.

LA JUNTA DIRECTIVA

Counselor: Francisco Sandoval Hernández
sandoval@dte.uma.es

Presidente: Pablo García Vázquez
pablo_garcia@ieee.org

Vicepresidente: Javier Muñoz Kirschberg
jmunozk@clientes.unicaja.es

Secretario: José Luis Vázquez Fernández-Baca
joselvazquez@yahoo.com

Tesorero: Jerónimo Vázquez García de la Vega
jvazquez@ieee.org

Rama de Estudiantes del I.E.E.E.

E.T.S. de Ingenieros de Telecomunicación
Universidad de Málaga.
Campus de Teatinos s/n
29071 Málaga.
Tlf: 95 2133301
Fax: 95 2132416





EL TEOREMA DE GÖDEL Y LA DISCUSIÓN SOBRE EL CARÁCTER COMPUTACIONAL DE LA MENTE HUMANA

Carles Ruiz Floriach

Estudiante de la ETSETB y miembro de la Rama de Estudiantes del IEEE de Barcelona
carles25@casal.upc.es

1- GÖDEL Y SU TIEMPO

1.1 Introducción

El Teorema de Incompletitud de Gödel se considera hoy en día uno de los más importantes que ha dado el siglo XX y se lo compara, en cuanto a profundidad y consecuencias, con la relatividad de Einstein o el principio de incertidumbre de Heisenberg. Como veremos más adelante es un teorema que conlleva implicaciones filosóficas de todo tipo, desde la discusión sobre la naturaleza de la verdad hasta aspectos relacionados con la mente y el pensamiento. Este artículo pretende ser una breve exposición de algunas ideas interesantes entorno al teorema, y en concreto de las ideas referidas a las implicaciones del teorema en las reflexiones sobre el pensamiento humano. Para ello, se explica brevemente quién fue Kurt Gödel, se exponen las ideas básicas subyacentes al teorema, y finalmente se discuten las consecuencias en lo referente a la filosofía de la mente.

1.2 Breve biografía

Kurt Gödel nació en 1906 en Brno, ciudad que actualmente pertenece a la República Checa. Allí estudió en el «Realgymnasium» (escuela técnica) sacando siempre las mejores notas, especialmente en matemáticas y lengua. Su hermano Rudolf explicó en cierta ocasión que se rumoreaba que Kurt no había cometido nunca ningún error gramatical en su época de escolar.



Gödel y Einstein en 1953

El joven Gödel era un chico introvertido pero muy despierto. En 1923 se trasladó a Viena para estudiar física en la universidad, aunque al poco tiempo optó por estudiar matemáticas. Se doctoró en 1929 y pasó a ser miembro de la universidad. En 1931 publicó la demostración de su famoso Teorema de Incompletitud. En 1933 Hitler llegó al poder. Al principio esto no tuvo mucha influencia en su vida, pero ocurrió que Schlick, que fue quien introdujo Gödel a lógica, fue asesinado por un estudiante nacional socialista. Este hecho marcaría profundamente la vida emocional de Gödel. En 1940 emigró a los Estados Unidos para trabajar en el Institute for Advanced Study en Princeton y en 1946 se nacionalizó americano. A pesar de sus logros no ocupó una cátedra hasta 1953, en parte debido a los rumores sobre su estabilidad mental. Publicó su último trabajo en 1958. Después de esto, fue encerrándose progresivamente en sí mismo y volviéndose paranoico. El temor a ser envenenado hizo que muriera de hambre en 1978.

1.3 Contexto histórico y noción de sistema formal

En el siglo XIX la matemática hizo grandes progresos. El descubrimiento de las geometrías no Euclídeas revolucionó la concepción que los matemáticos tenían acerca de la relación entre la matemática y el mundo real. Hilbert trabajó en la fundamentación de la geometría y Cantor creó la teoría de conjuntos que arrojaba un poco de luz sobre el esquivo concepto de infinito. Aún así, seguían habiendo paradojas a las que nadie sabía dar respuesta. El ejemplo paradigmático es la paradoja de Bertrand Russell que aparece cuando nos preguntamos acerca de la naturaleza del conjunto de todos los conjuntos. Así pues, la matemática se enfrentaba a situaciones que hacían tambalear sus fundamentos.

En este contexto, uno de los principales objetivos de los matemáticos en el siglo XIX y principios del XX fue establecer una base sólida sobre la cual edificar las matemáticas. Russell y Whitehead hicieron un intento descomunal de fundar las matemáticas en base a la lógica, eliminando la posibilidad de la aparición de paradojas. Su obra, los Principia Mathematica, efectivamente estaba libre de contradicciones. De todas formas quedaba por saber si el sistema definido por los Principia era completo, es decir, si podía expresar todas las verdades de las matemáticas. Ante esta cuestión, Hilbert propuso un programa que tuvo mucha influencia en la lógica del primer



cuarto del siglo XX. A este programa se le llamó formalismo.

El concepto clave de las ideas de Hilbert era el de sistema formal. Un sistema formal no es nada más que un conjunto de secuencias de símbolos que se construyen en base a unas reglas determinadas. Estas secuencias representan realidades de las matemáticas. Hay ciertas secuencias iniciales (axiomas) que se aceptan como verdades, y en base a estas secuencias se deducen todas las otras. Cualquier secuencia que se deduzca de los axiomas es un teorema, y cualquier secuencia que no lo haga es una proposición no demostrable (y es falsa si el sistema formal representa la realidad).



David Hilbert

Así pues, Hilbert pretendía hallar un sistema formal axiomático que permitiese fundamentar las matemáticas liberándolas de contradicciones.

Pero sucedió que en 1931 Gödel publicó su Teorema de Incompletitud, que demostraba que este objetivo era inalcanzable. El Teorema de Gödel decía que no puede formularse un sistema formal suficientemente potente que sea a su vez consistente y completo. "Suficientemente potente" significa que el sistema contenga en sus reglas y axiomas la aritmética de números naturales. Gödel demostró que en un sistema de este tipo siempre existirá una proposición sobre los números naturales que es verdadera pero que no puede ser demostrada dentro del sistema. En otras palabras, en cualquier sistema formal que contenga la aritmética de números naturales, siempre hay una proposición indecidible.



Kurt Gödel

El Teorema de Gödel señalaba límites inherentes a la matemática y eliminaba la pretensión de adquirir un conocimiento completo y cierto. Entonces, nuestra concepción del mundo tuvo forzosamente que cambiar.

Poco después, paralelamente a la aparición de los primeros computadores, aparecieron los primeros intentos de trasladar los resultados del Teorema de Gödel a otras disciplinas como las ciencias de la mente y las ciencias de la computación.

Parecía que los ordenadores, sujetos a reglas como los sistemas formales, deberían tener ciertas limitaciones que los humanos no tendrían necesariamente que tener. Pero esto lo dejamos para la sección 3. De momento intentaremos trazar las líneas básicas del razonamiento del teorema. A esto se dedicará la sección 2.

2- EL TEOREMA DE GÖDEL

2.1 Procedimiento del Teorema de Gödel

Asumimos un sistema axiomático suficientemente poderoso como para incluir la aritmética de los números naturales. Esto quiere decir que cualquier proposición sobre la teoría de los números se puede escribir en el lenguaje del sistema formal, y que además, si es verdadera, se deduce de los axiomas mediante las reglas de inferencia del sistema. Llamaremos a este sistema S , y como acabamos de decir, S es un sistema consistente y completo.

El punto clave de la demostración es la llamada proposición de Gödel, que llamaremos G y que escribimos a continuación:

$G: \text{ Esta proposición no es un teorema de } S$

Supongamos que esta proposición se puede escribir en el lenguaje de S . Entonces, siguiendo el significado de G , vemos que si G es falsa tenemos un teorema falso de S y por lo tanto el sistema S no es consistente. Por otro lado, si G es verdadera tenemos una proposición verdadera que no es un teorema de S , es decir, un teorema cierto que no puede ser probado dentro del sistema y por lo tanto S es incompleto. Resumiendo:

$G \text{ falsa} \Rightarrow \text{Tenemos un teorema falso del sistema} \Rightarrow \text{El sistema falla (no es consistente)}$

$G \text{ verdadera} \Rightarrow \text{Tenemos un teorema cierto que no es del sistema} \Rightarrow \text{El sistema es incompleto}$

Pero tomemos distancia: ¿Podemos en realidad escribir G como una proposición de S , como hemos supuesto anteriormente? Hemos asegurado que S contiene todos los teoremas posibles sobre la teoría de números naturales, pero la proposición G no habla de números, sino que habla de si misma. Si fuésemos capaces de hacer que el lenguaje de S hablase de si mismo, podríamos escribir la proposición G en el lenguaje de S . Entonces tendríamos efectivamente una proposición verdadera que no es un teorema de S , y habríamos demostrado que el sistema S es incompleto. Gödel demostró que esto es posible en cualquier sistema que contenga la aritmética de números



MEDIR LAS DISTANCIAS

Miguel Mañas

*Estudiante de la ETSETB y miembro de AEES Estudiants
miguel27@casal.upc.es*

«No puede haber un lenguaje más universal y más simple, más libre de errores y de oscuridades [...] más digno de expresar las relaciones invariables de las cosas naturales [que las matemáticas]. Interpreta [todos los fenómenos] con el mismo lenguaje, como si quisiera atestiguar la unidad y simplicidad del plan del universo, y hacer aún más evidente este orden inalterable que preside todas las causas naturales».

Joseph Fourier, *Théorie analytique de la chaleur.*
Discurso preliminar, 1822.

El pasado mes de Octubre, y para gran alegría de los aficionados a la astronomía, AEES Estudiants organizó un ciclo de conferencias de cosmología en el Campus Nord de la UPC. Este ciclo constó de tres conferencias cuya pretensión era ilustrar, de forma introductoria e intuitiva, tres importantes temas de la cosmología.

- ***El Big-Bang y el satélite Plank de la ESA***
Por Enrique Gaztañaga.

- ***Fundamentos de relatividad para la astrofísica***
Por Alberto Lobo.

- ***Nacimiento, evolución y muerte de las estrellas***
Por el Dr. Jordi Isern.

Pese a la gran aceptación de las conferencias, éstas dejaron a los asistentes con la sensación de que todavía quedan muchos temas de los que les gustaría saber más. Y efectivamente, así es. La cantidad de temas relacionados con la astronomía, la astrofísica y la cosmología es tan grande, que intentarlos resumir en tres conferencias es francamente imposible. Es por ello que se ha escrito este artículo, esperando sirva de ayuda para algunos de vosotros.

Uno de los temas que parecía resultar más atractivo, de entre todos los que no fueron abordados en las conferencias, era el de la forma en que los astrónomos median las distancias a que se encuentran las estrellas y las galaxias. Sus inquietudes son razonables, puesto que si observamos el estado de las teorías cosmológicas en la actualidad, nos encontramos regularmente con argumentos de tamaño y velocidad de expansión del universo. Deben existir métodos para medir distancias suficientemente potentes como para que estos argumentos tengan algún sentido. Evidentemente estos

métodos existen, y si bien la mayoría de los asistentes a las conferencias tenían una noción de las bases en que se apoyan, siempre quedaba la duda de como es en realidad este procedimiento.

Pues bien, veamos si somos entre todos capaces de esclarecer el misterio de una forma entendedora. Para ello, me veo en la obligación de tratar un par o tres de temas que luego van a hacerme falta. Antes que nada:

1.- EL PARALAJE

El paralaje no es más que un método que permite calcular la distancia del observador a un objeto concreto. Para ilustrar de forma extremadamente simple en que consiste esta técnica, realicemos ahora un conocido experimento. Coloquemos el dedo pulgar delante de los ojos (contra un fondo no uniforme) a una distancia de unos 10 cm. Manteniendo inmóviles la cabeza y el dedo, miremos primero con un ojo y luego con el otro. Observamos que la posición del dedo con respecto al fondo varía. Estamos modificando nuestro punto de vista al mirar con uno u otro ojo. Si repetimos el experimento pero con el dedo más alejado de los ojos (un palmo por ejemplo) observamos el mismo efecto, pero esta vez, la variación no será tan grande.

Este efecto existe ya que los ojos están separados entre ellos varios centímetros, de modo que la línea imaginaria que une el dedo con uno de los ojos forma un ángulo apreciable con respecto a la línea imaginaria que une el dedo con el otro ojo. Sin embargo, esto deja de ser apreciable para objetos colocados más allá de unos 15 m de nuestros ojos, con lo cual no es en absoluto útil para el caso de la medición de las distan-



cias de las estrellas, que de hecho se encuentran mucho más lejos. Sin embargo, si pudiésemos agrandar la distancia entre los ojos, podríamos también agrandar el rango en que el paralaje es apreciable. Esto no es posible, naturalmente, pero sí que podemos utilizar dos observatorios separados por distancias de cientos de kilómetros, y considerar que cada uno de ellos es uno de los "ojos" del sistema. La base analítica sobre la que descansa el problema es la que sigue.

Supongamos, por ejemplo, que deseamos medir la distancia que hay entre la Tierra y la Luna utilizando el paralaje. Lo primero que tenemos que hacer es suponer que disponemos de un fondo fijo a suficiente distancia del objeto a medir (la Luna). En este caso podemos considerar como fondo fijo las estrellas, ya que se encuentran a distancias suficientemente grandes como para ello. Llamaremos a la línea que une

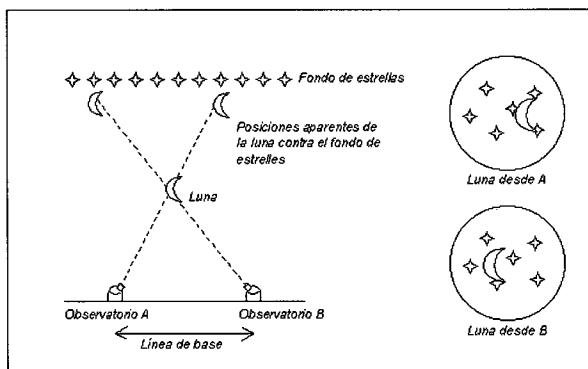


Figura 1. Esquema de paralaje. Dos observatorios y la Luna. El fondo son las estrellas fijas.

nuestros dos observatorios línea de base. El esquema de la figura 1 resultará aclaratorio.

Podemos representar de forma esquemática los dos triángulos de la **figura 1** y dar nombre a ciertos parámetros relevantes. Observacionalmente se puede encontrar el valor de los ángulos a y b , y la distancia d ,

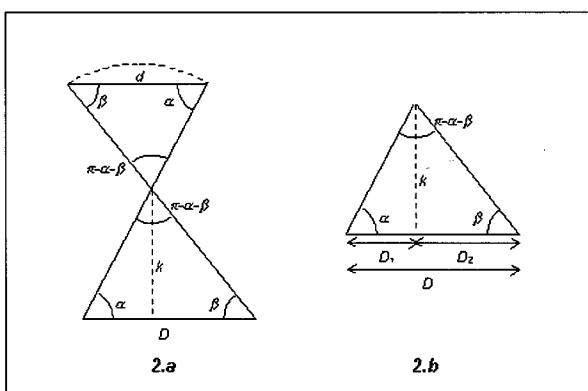


Figura 2

que (medida en grados) es el ángulo de paralaje. De este esquema (**figura 2.a.**), nos bastará el triángulo de la **figura 2.b** para calcular la distancia k , del observador al objeto observado.

Como puede verse, la realización de ciertas aproximaciones es indispensable para poder utilizar este modelo; sin embargo, no nos detendremos a estudiarlas, ya que tan sólo aportarían confusión al concepto de paralaje.

Utilizando un poco de geometría y trigonometría básicas podemos plantear sistema de ecuaciones siguiente:

$$\operatorname{tg} b \cdot (D - D_2) = k$$

$$\operatorname{tg} a \cdot (D_2) = k$$

Cuya solución es:

$$k = \operatorname{tg} \alpha \cdot (h \cdot D) / (h + 1)$$

donde

$$h = \operatorname{tg} \beta / \operatorname{tg} \alpha$$

que sería la solución únicamente para el punto $P = D_1 - D_2$. Sin embargo, el error de considerar ésta

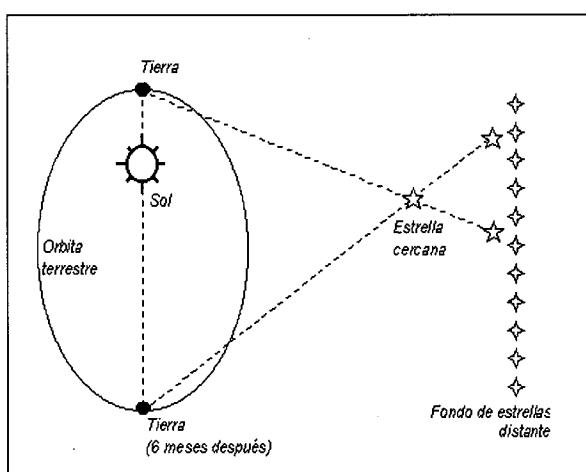


Figura 3. Paralaje de línea de base igual al diámetro de la órbita terrestre.

como la distancia entre el objeto y cualquiera de los observatorios no es muy grande y puede utilizarse de esta forma sin preocupaciones.

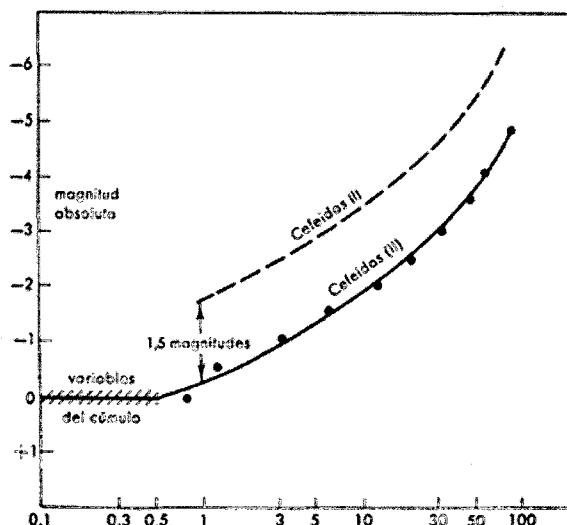
Utilizando el paralaje podemos llegar a medir distancias muy grandes, aunque todo tiene su límite. Pese a que la invención del telescopio ha resultado de gran ayuda en la medición de ángulos de paralaje muy pequeños, la línea de base continúa siendo un gran limitador del método. Cuanto más grande es ésta, más lejano puede estar el objeto y sea detectable su paralaje.

En la **figura 3** se puede ver un método que hace que la línea de base sea suficientemente grande como para observar el paralaje de las estrellas más cercanas.

Sin embargo, aún con los mejores instrumentos, no es posible detectar paralajes de estrellas a más de unos 100 años luz de nosotros. Y es cierto que más allá de estas distancias existen todavía muchas estrellas. Para poder medir esas distancias se hace necesario emplear otro método. Pero antes, quizás sea conveniente repasar las variables cefeidas.

2. LAS VARIABLES CEFEDAS

Las estrellas Cefeidas entran dentro del grupo de estrellas variables, o sea, que cambian el brillo cada



Gráfica 1. Relación entre el periodo y la luminosidad de las Cefeidas

cierto periodo de tiempo. Las razones de estos cambios de brillo no deben preocuparnos por ahora, pero debemos quedarnos con la idea de que la variabilidad de algunas de estas estrellas está motivadas por causas externas a las propias estrellas (eclipses, por ejemplo). Uno de los casos en que la causa de la variabilidad es debida a la propia estrella corresponde a la estrella

Delta Cephei, cuarta estrella en brillo de la constelación de Cefeo. Su periodo es muy corto y extremadamente regular. A las estrellas del tipo de Delta Cephei se les llama variables Cefeidas, o simplemente Cefeidas.

Las Cefeidas tienen una característica muy peculiar, y es que su periodo y su luminosidad están relacionados. Cuanto más luminosa es una Cefeida, mayor es su periodo. Esto puede apreciarse en la **gráfica 1**. Como se puede observar, existen dos tipos de variables Cefeidas, pero para el caso no nos preocuperán sus particularidades. Bastará con conocer la existencia de la relación.

Sin embargo, nosotros percibimos de una estrella su brillo, y no su luminosidad. El brillo de una estrella depende de la luminosidad de la misma, pero

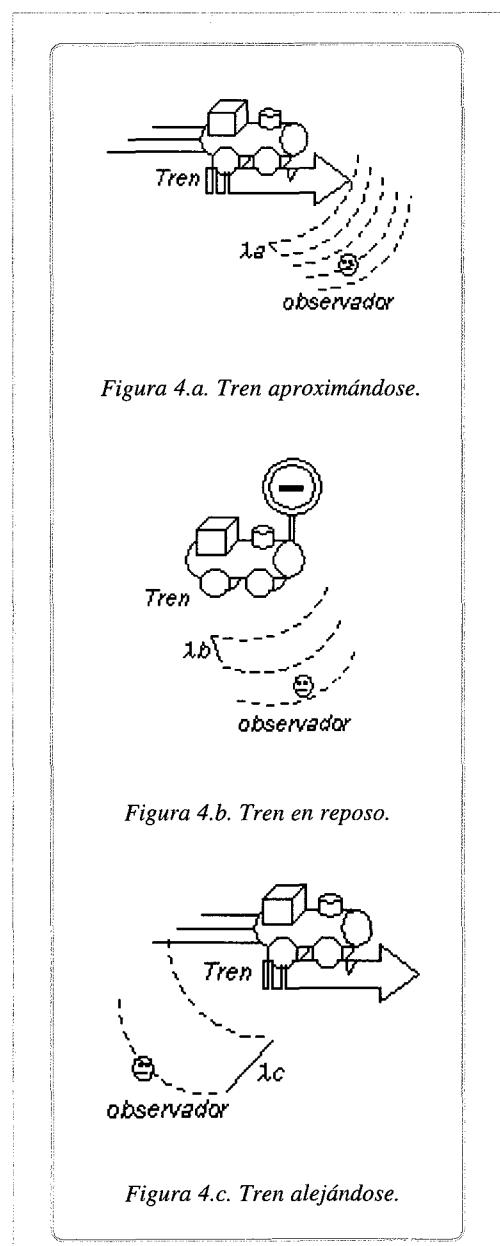


Figura 4.a. Tren aproximándose.

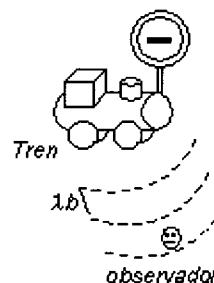


Figura 4.b. Tren en reposo.

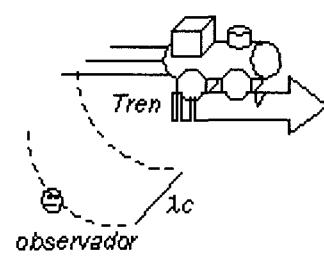


Figura 4.c. Tren alejándose.

Figura 4. Efecto Doppler.



también de la distancia que hay entre ella y nosotros sus observadores. De esta forma, si supiésemos la distancia a una sola Cefeida podríamos calcular (observando su período), la distancia de cualquier otra Cefeida de la misma población. El problema es que no sabemos la distancia a ninguna.

3.- EL EFECTO DOPPLER

No es mi intención aquí explicar el efecto Doppler tal y como lo haría un libro de física, sino dar una idea intuitiva para que cuando lo utilicemos posteriormente pueda entenderse sin dificultad. Por ello, al igual que con el paralaje, veamos un ejemplo fácil a partir del cual podamos extrapolar a situaciones más generales.

Consideremos la siguiente situación: nos encontramos en el andén de una estación ferroviaria cuando se acerca un tren que emite un silbido. Mientras el tren se acerca oímos el silbido más agudo de lo que oiríamos si el tren estuviese respecto a nosotros en reposo. Cuando el tren pasa justo por delante nuestro, notaremos un cambio brusco en el sonido, y mientras el tren se aleja, lo que oiremos será más grave que si el tren estuviese en reposo con respecto nosotros.

La explicación de esto es bastante sencilla y se puede entender sin entrar en un análisis profundo. Lo único que hace falta saber es que el sonido se propaga en forma de ondas de presión por el aire (o cualquier otro medio material), y que la distancia entre un frente de onda y el siguiente se llama *longitud de onda* (λ). La inversa de la longitud de onda por la velocidad de propagación es la *frecuencia* (f). Los sonidos agudos son aquellos que tienen frecuencias altas (longitudes de onda pequeñas), y los sonidos graves son aquellos que tienen frecuencias bajas (longitudes de onda grandes). En la **figura 4** podemos observar por qué el movimiento del tren respecto al observador hace que el sonido sea diferente según se acerque o aleje la fuente emisora (silbato del tren).

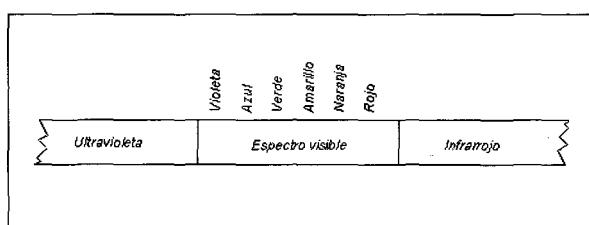


Figura 5. El espectro.

En la **figura 4.b.** se observa la longitud de onda del silbato del tren, estando este en reposo con respecto a nosotros.

Cuando el tren se está acercando lo que ocurre es la siguiente: El primer frente de la onda es emitido con normalidad, y se propaga de la forma habitual (velocidad del sonido). Sin embargo, cuando el segundo frente de onda sea emitido, el tren se habrá desplazado hacia el observador, con lo cual la distancia entre los dos frentes será menor que el caso de reposo. Con los siguientes frentes ocurrirá lo mismo, con lo cual, percibiremos una frecuencia mayor, y por lo tanto un sonido más agudo. Puede apreciarse en la **figura 4.a.**

La situación contraria es semejante: si el tren se aleja, la emisión del segundo frente de onda se producirá cuando el tren se haya movido y esté más lejos, cosa que quiere decir que la longitud de onda será mayor que en los casos anteriores, y por lo tanto la frecuencia será menor. El sonido que oiremos será más grave (**figura 4.c.**)

Utilizando el efecto Doppler podemos calcular la velocidad a la que se acerca o aleja un objeto con respecto a un observador.

El efecto Doppler puede aplicarse también a la luz, ya que ésta no es otra cosa que ondas electromagnéticas, y como tales, son válidas sobre ellas todas las definiciones anteriores.

4.- EL ESPECTRO

Los experimentos de Newton demostraron que la luz blanca no es otra cosa que la suma de luz de distintos colores. De hecho, se sabe que cada color corresponde a una longitud de onda diferente. También es conocido que la luz visible tan sólo corresponde a una pequeña parte de las longitudes de onda emitidas por un cuerpo. De hecho existen longitudes de onda por encima y por debajo de la pequeña franja que corresponde a la luz que somos capaces de ver. En conjunto, todas estas longitudes de onda, de forma ordenada, corresponden a lo que se llama el espectro electromagnético (en el arco iris vemos el trozo de espectro correspondiente a la luz visible).

Podemos ver en la **figura 5** el espectro de una forma muy simplificada.

Para poder detectar los movimientos relativos de una estrella con respecto a nosotros, deberíamos fijarnos en su espectro, de forma que si una estrella se aproximase a nosotros veríamos que la luz que emite cambiaría a tonos más azulados (corrimiento hacia el violeta); y al revés, una estrella que se acercara a nosotros debería parecer más rojiza de lo que sería si no existiese tal movimiento (corrimiento hacia el rojo).

Sin embargo, nos encontramos en este razonamiento con un par de puntos que hacen que eso no sea así:

- Existen otros motivos que hacen que la luz emitida por una estrella sea de diferentes colores. (temperatura, composición química...)
- Un corrimiento hacia el rojo o el azul no sería apreciado (excepto a muy grandes velocidades), ya que el espectro se desplaza en bloque, y lo que sale de la franja visible por un lado, entra por el otro.

Llegamos a la conclusión de que debemos buscar algún método para identificar un posible corrimiento, que no sea basándose en el color aparente. Este método existe, y consiste en localizar ciertas marcas en el espectro que puedan servir de referencia. Estas marcas son llamadas líneas espectrales, y no son otra cosa que longitudes de onda que no están presentes en la luz recibida. La razón de la ausencia de estas longitudes de onda es una característica de absorción energía de cada elemento, de lo que no es momento de hablar.

Así pues, sabemos que el espectro de una estrella presenta discontinuidades, y si somos capaces de determinar las longitudes de onda a que corresponden, tendremos una marca de la estrella. Si tenemos un modelo adecuado del espectro de una estrella en reposo, tan sólo nos hará falta medir las diferencias entre las líneas espectrales del modelo y la estrella observada, y podremos calcular la velocidad con que se mueve con respecto a nosotros.

5.- LA DISTANCIA DE UNA CEFEDA

Podemos utilizar el efecto Doppler para medir la distancia que nos separa de una Cefeida y poder así tener una escala fácil para calcular distancias mucho más grandes que las permitidas por la técnica de paralaje. (Efectivamente, el efecto Doppler te permite medir velocidades, ¡pero no distancias! Sin embargo, veremos una forma mediante la cual podemos establecer la distancia a una Cefeida. Paciencia).

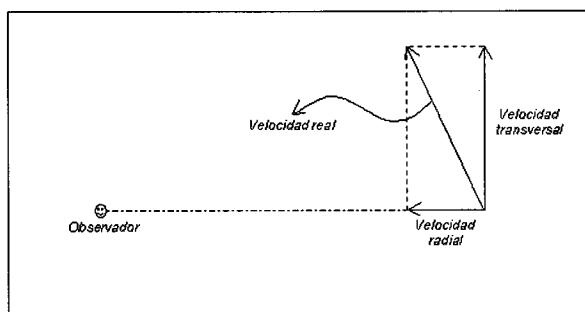


Figura 6. Componentes del movimiento.

Deberemos tener en cuenta que los movimientos son tridimensionales, y se hará necesario utilizar algo de probabilidad de forma intuitiva y geometría vectorial básica.

La situación habitual con la que nos encontraremos es la que se muestra en la **figura 6**. Donde considerado tan sólo el plano en que se mueve el objeto, y hemos descompuesto la velocidad en sus componentes transversal y radial.

La velocidad radial es la velocidad que nosotros medimos con el efecto Doppler-Fizeau; la velocidad transversal (medida en ángulos de arco) es la velocidad que se observa cuando un cuerpo se desplaza por el firmamento. Como ya sabemos, el cálculo de la velocidad radial no depende de la distancia a que se encuentre el objeto. Sin embargo, el cálculo de la velocidad transversal sí que depende de esa distancia (No es lo mismo que un objeto a tres metros se mueva 1 segundo de arco, que lo haga un objeto a tres kilómetros). La necesidad de conocer la distancia para poder calcular la velocidad transversal puede interpretarse a la inversa: si conociésemos la relación entre las componentes radial y transversal (el ángulo con respecto a la velocidad real), seríamos capaces de determinar la distancia.

El problema reside en que tampoco conocemos esta relación. Es ahora donde entra en juego el cálculo estadístico. El truco de la cuestión está en saber que por regla general, en unos casos la velocidad tangencial será mayor que la radial, y en otros a la inversa; pero en media, estas diferencias son canceladas y podemos considerar ambas componentes iguales.

Podemos calcular la distancia a una Cefeida de forma bastante fácil entonces, lo que sucede es que el error que se puede producir es enorme. De hecho, lo más probable es que el resultado sea absolutamente disparatado. Sin embargo, si realizamos el cálculo para un gran número de Cefedas, todas ellas del mismo cúmulo y de igual período, entonces, la distancia media que calcularemos será una muy buena aproximación a la distancia real.

6.- MIDIENDO DISTANCIAS

Una vez hemos determinado la distancia de una Cefeida, tal y como ya habíamos visto anteriormente, podemos determinar la distancia de cualquier otra del mismo tipo con tan sólo conocer su período. La razón de utilizar como referencia una variable Cefeida para medir las distancias viene dada por la inexactitud de los cálculos por el método de las velocidades. Una vez determinada con gran exactitud la escala Cefeida y la distancia de algunas de ellas, resulta mucho más preciso utilizar este método que no el otro. Debe tenerse en cuenta que las variables Cefedas son relativamente abundantes en el Universo, y por lo tanto, son fáciles de localizar; con lo que el método es bastante útil.



ADSL: ACCESO INDIRECTO AL BUCLE DE ABONADO

Xavier Lago Santín

Estudiante de 5º curso de Telecomunicación en la ETSETB,
Proyectista de Ingeniería Electrónica en el Dpto. Matemática Aplicada i Telemática.
jlagosan@retevision.es

En estos últimos años hemos podido oír que el mundo se encuentra en una era donde la información es uno de los bienes máspreciados por el ser humano. El acceso a este recurso tan valioso se ha convertido en algo primordial, tanto que se ha bautizado a esta época como la “era de la información”. Para que esto sea posible han de desarrollarse una serie de tecnologías que permitan que tanpreciado bien llegue a todos sus posibles consumidores, que cada vez exigen una mayor cantidad de servicios y una mayor calidad. Hemos pasado de la telefonía, la radio y la televisión analógica, como necesidades básicas, al acceso a las redes de datos (INTERNET es el mayor exponente de este hecho), la televisión digital y el video bajo demanda, la videoconferencia, el comercio electrónico y otros servicios, que serán considerados tan básicos como eran los anteriores, en un futuro no demasiado lejano. Pero para que todo esto sea posible, sobre todo el acceso a INTERNET y los servicios multimedia, necesitamos de una tecnología que permita transmitir información a altas velocidades, mucho más de lo que hasta ahora estamos habituados.

Con la tecnología ADSL se aprovecharía la infraestructura existente y se llegaría prácticamente a toda la población con menor coste de implantación

El principal problema que nos encontramos a la hora de llevar una red de Banda Ancha hasta el usuario es que esto suele tener un coste muy elevado. Una alternativa que se impone con mucha fuerza son las redes de cable, en un principio pensadas únicamente para permitir aplicaciones de televisión digital y video bajo demanda, y que ahora se ven también como una forma de permitir el acceso a todos los otros servicios que hemos nombrado. El gran inconveniente es que hasta el momento no existe una infraestructura de cable que llegue a todos los usuarios potenciales, y que es complicado que esto ocurra en un plazo de tiempo corto, dada la terrible inversión y la gran dificultad material que la instalación del cable comporta. Otra posibilidad sería el uso del acceso por satélite, tal como ahora se usa para la difusión de televisión, pero ésta

tecnología también se encuentra con graves problemas dado que el coste del equipo transmisor que necesita el usuario para enviar información hacia la red es bastante alto. Por otra parte surge también la siguiente pregunta: ¿qué vamos a hacer con la red de acceso existente de telefonía básica? Esta red de acceso, de poder aprovecharse, tiene unas ventajas considerables. Por un lado se aprovecharía la infraestructura existente, se llegaría prácticamente a toda la población, y el coste de implantación (tanto para el usuario como para el proveedor de servicio) serían previsiblemente menores que en los otros casos. Con esta idea nacieron la familia de tecnologías xDSL, y en concreto la conocida como ADSL, que será el centro de este artículo.

En un caso como el español la red de acceso de telefonía básica, lo que conocemos normalmente como bucle de abonado, es propiedad de una operadora “dominante”, que es Telefónica S.A.. El ministerio de fomento, en unas ordenes del 26 de Marzo de este mismo año (O.M. 8181 y 8182, publicadas en el BOE del 10 de Abril de 1999), obliga a esta operadora a ceder sus líneas permitiendo que otras operadoras den servicio de acceso al abonado utilizando su bucle local, pero de forma indirecta. Esto, que a primera vista puede parecer un poco enrevesado, no lo es tanto si lo miramos detenidamente. La operadora Telefónica está obligada a ir adecuando sus centrales locales paulatinamente, de forma que en el extremo del bucle de abonado que se encuentra en ellas se instale la infraestructura necesaria para permitir el servicio ADSL. A partir de aquí, cuando un usuario contrate el servicio a una operadora, ésta instalará en su domicilio el resto de infraestructura. El tráfico que genere el usuario irá

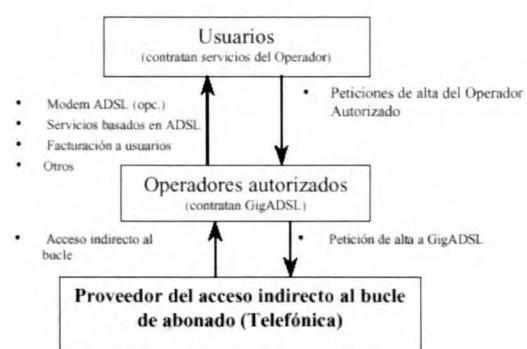


Figura 1: Estructura del acceso ADSL y sus partes

hacia la central local, donde se separará en dos partes: telefonía básica (que se dirigirá por la RTB) y datos, los cuales se conducirán por la red de datos de Telefónica (red ATM) hasta un punto donde se interconectarán con la red ATM de la operadora que proporciona el servicio al abonado. Esto se verá más claramente en el siguiente punto, cuando presentemos el esquema global de la tecnología.

ESQUEMA GLOBAL DE LA TECNOLOGÍA ADSL

Lo primero que hemos de ver es que es exactamente la tecnología ADSL. Su nombre mismo ya nos da una indicación de esto. ADSL es el acrónimo de *Asymmetric Digital Subscriber Line*. ADSL es un método de transmisión digital que consigue velocidades muy altas en un sentido de la comunicación mientras que en el otro son relativamente bajas (tecnología de transmisión asimétrica), aprovechando para ello el bucle local de abonado. En el mejor de los casos podemos conseguir velocidades de hasta 8 Mb/s en un sentido y del orden de 640 kb/s en el otro. Este tipo de transmisión de datos asimétrico se justifica en tanto en cuanto vamos a asociarlo a servicios para los cuales el usuario transmite muy poca información (comandos, secuencias de control) pero recibe un gran volumen de datos (imágenes, ficheros de datos, etc.). Esta tecnología pretende aprovechar el par de cobre del bucle de abonado, y transmitir simultáneamente por éste la señal de voz (POTS = Plain Old Telephone Service) junto con una conexión de datos. Esto es lo que está recogido en los estándares, pero lo que queda es ver como vamos a conseguirlo.

En primer lugar tomamos una división frecuencial del espectro por el que vamos a transmitir. Aprovecharemos frecuencias de hasta 1.1 MHz, en las cuales la atenuación que introduce el cable todavía permite transmitir de forma fiable por líneas de hasta unos 5 o 6 km, pudiendo dar servicio a gran parte de los abonados (aproximadamente un 90% se encuentran a esta distancia o inferior de la central local). Distancias mayores requieren de otro tipo de tecnologías de acceso. Este ancho de banda del que partimos vamos a dividirlo en tres fragmentos:

1. En el caso más usual (el que no usa cancelación de ecos en los equipos receptores) el espectro queda separado de la siguiente forma:

- Un canal hasta los 4 KHz reservado para los servicios de voz (POTS).
- Un canal de los 25 KHz a los 150 kHz reservado para tráfico de subida.
- Un canal de 150 kHz a 1.1 MHz para tráfico de bajada.

2. En el caso de estar usando técnicas de cancelación de ecos la parte dedicada a tráfico de bajada puede extenderse desde los 25 kHz hasta 1.1MHz.

En primer lugar vamos a considerar tres zonas a la hora de analizar el acceso al usuario mediante la tecnología ADSL. Por un lado vamos a considerar la parte de usuario, por otro lado la parte que atañe a la operadora dominante (poseedora de la infraestructura de cobre) y finalmente a la operadora que va a proveer servicios mediante esta tecnología. Vamos a entrar un poco más en detalle en ellas.

En la parte del usuario vamos a tener dos posibles configuraciones que permiten utilizar esta tecnología. Por un lado tenemos la versión de ADSL cono-

Este tipo de transmisión de datos asimétrica se justifica en tanto en cuanto vamos a asociarlo a servicios para los cuales el usuario transmite muy poca información (comandos, secuencias de control) pero recibe un gran volumen de datos (imágenes, ficheros de datos, etc)

cida como “full-rate”. En este caso podemos obtener toda la velocidad que la línea sea capaz de dar, teniendo en cuenta las apropiadas consideraciones en cuanto a distancia, ruido, interferencias, etc., hasta el máximo que permite esta tecnología. En este caso lo primero que vamos a encontrar en casa del abonado es un “POTS splitter”, un dispositivo que separa el canal telefónico de voz de los dos canales de datos. A este dispositivo tendremos directamente conectados el teléfono y un módem ADSL, que a su vez irá conectado a una tarjeta de interfaz de red (NIC) en un PC, a un router, si estamos interconectando toda una red de área local (LAN) o a un “bridge”. Esta configuración es cara y costosa de instalar. La solución a estos dos inconvenientes los dio la norma G-lite, que consiste en una versión de ADSL que no usa “splitter”. A cambio de esto, la velocidad máxima que podemos alcanzar se ve limitada a 1.5 Mb/s de flujo de bajada y 512 kb/s de subida. En este caso sólo necesitamos el módem, que puede ser instalado fácilmente por el propio usuario. Tendremos problemas con ruidos espúreos, pero po-

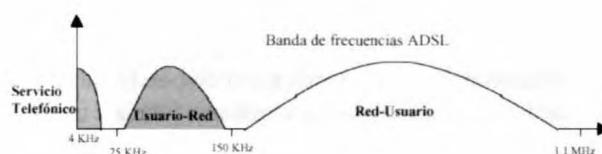


Figura 2: Diagrama espectral típico de una línea ADSL



demos solucionarlo fácilmente mediante un filtro pasivo paso bajo conectado a la clavija del aparato telefónico.

Otra parte importante del esquema de conexión va a ser la central local que está interconectada al usuario mediante el par de cobre. En esta central vamos a encontrarnos también con un “splitter”, con la misma funcionalidad que habíamos visto en el extremo cliente. En este caso la presencia de este dispositivo es obligatoria. Por un lado vamos a conectar el “splitter” a la RTBC (Red Telefónica Básica Conmutada) por donde se va a dar el servicio de telefonía. Por otro lado, los canales de datos van a ir a otro módem ADSL, una tarjeta situada en el dispositivo conocido como DSLAM (Digital Suscriber Line Access Multiplexer) que va a recibir los paquetes de datos de cada usuario, los va a multiplexar y los va a llevar mediante la red de transporte, que usualmente va a ser ATM, hacia lo que conoceremos como PAI, o punto de acceso indirecto. Cada una de estas tarjetas típicamente contiene 4 o 8 puertos, pudiendo soportar un rack de multiplexores generalmente entre 500 y 1000 líneas ADSL.

El punto de acceso indirecto es la frontera con la tercera parte implicada en la estructura que hemos planteado. En este caso se plantea este punto como la interconexión de la red de transporte de la operadora que posee el par de cobre con el proveedor de servicios, sea otra operadora telefónica, un proveedor de acceso a INTERNET, de vídeo bajo demanda, etc. Esta interconexión se realiza entre las redes de transporte ATM de las partes implicadas.

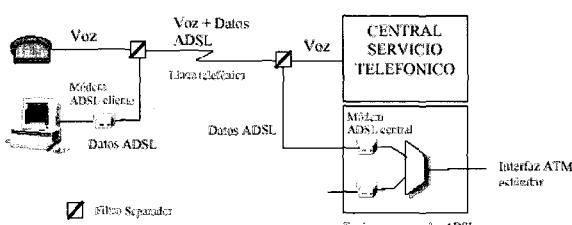


Figura 3: Esquema típico de una conexión ADSL

Hemos visto el esquema global de conexión y hemos nombrado los diferentes dispositivos que hemos de tener para poder dar servicio ADSL. Vamos ahora a hacer algunos comentarios interesantes al respecto. Los módems ADSL trabajan fundamentalmente con dos modulaciones distintas, CAP y DMT, la segunda de las cuales veremos de una forma más amplia en el siguiente apartado. Hemos de tener en cuenta que puede haber problemas de compatibilidad y habremos de asegurarnos cual es la modulación que se va a usar en el extremo del DSLAM. Por otro lado la separación de los canales de datos del canal vocal mediante filtrado permite que tengamos un servicio telefónico

como el que estamos acostumbrados, mientras que tendremos una conexión de datos que estará permanentemente en servicio, de forma que no tendremos que llevar a cabo fases de establecimiento de llamada ni otros procesos a los que estábamos acostumbrados. De la misma forma estamos descongestionando la red conmutada, separando los dos tráficos y enrutándolos hacia redes distintas.

Debido a los efectos capacitivos y resistivos distribuidos, los cables telefónicos atenuan y distorsionan las señales de voz, con efectos que se incrementan con la frecuencia y la longitud de la línea

Hasta aquí hemos hablado de la topología de la conexión y de algunas características básicas del ADSL. Vamos ahora a ver algunos de sus principales problemas. Uno de los peores problemas que tenemos son las bobinas de carga. Debido a los efectos capacitivos y resistivos distribuidos, los cables telefónicos atenuan y distorsionan las señales de voz, con efectos que se incrementan con la frecuencia y la longitud de la línea. Para ecualizar las líneas en la banda vocal las compañías instalan bobinas de carga en el extremo de la línea que está en la central local. Estos inductores mejoran la respuesta en frecuencia para la banda vocal, pero incrementan la atenuación de las señales por encima de esa banda. En consecuencia van a ser perjudiciales para la transmisión de datos a alta velocidad y deberían ser eliminadas. Otro gran problema que nos vamos a encontrar van a ser los efectos de dispersión. Las diferencias en la propagación de las distintas frecuencias van a provocar la aparición de ISI, de forma que vamos a tener una velocidad de transmisión limitada por este efecto. Como en el caso de la atenuación este efecto empeora con la frecuencia y la distancia (en ambos casos al aumentar la frecuencia y la distancia el efecto es más grave). Otros problemas graves son las líneas en circuito abierto (que provocan reflexiones por desequilibrio de impedancias a altas frecuencias) y el efecto de “cross-talk” en sus dos posibilidades. El NEXT (near end cross talk) es la interferencia provocada por otro usuario situado en el mismo extremo del sistema que yo. Puede ser un efecto muy nocivo, dado que la potencia del emisor es alta mientras que la señal que yo recibo viene atenuada. El FEXT (far end cross talk) viene provocado por transmisores en el otro extremo de la línea. Estos problemas hacen que nos veamos obligados a tener muy en cuenta la potencia que hemos de enviar así como los filtros que hemos de poner para eliminar estos ruidos espúreos.

LAS MODULACIONES EN ADSL

Fundamentalmente vamos a ver que se usan dos modulaciones en esta tecnología. Por un lado está la CAP, que es el acrónimo de Carrierless Amplitude-Phase, es similar a la QAM en tanto en cuanto es un esquema de transmisión en dos dimensiones multinivel. La diferencia fundamental estriba en que no usa portadoras ortogonales generadas por un mezclador en seno-coseno, sinó que la modulación de los dos flujos de bits se lleva a cabo digitalmente por dos filtros transversales paso banda. Estos filtros tienen igual característica de amplitud pero una diferencia en fase de $\pi/2$ (lo que se conoce como un par de Hilbert).

No hablaremos más sobre esta modulación ya que los estándares contemplan a la DMT como modulación a usar.

DMT son las siglas de “Discrete Multi-Tone”, una tecnología de modulación multiportadora. La banda de interés se divide en 256 subcanales cada uno con un ancho de banda de aproximadamente 4 KHz. Cada subcanal puede llevar un cierto número de bits. Estos bits son asignados a una subportadora FDM (un tono frecuencial), cada una operando a una frecuencia diferente. Estas subportadoras han de ser ortogonales entre sí, siendo conveniente para esto la FFT. Esta multiplexación en frecuencia utilizando diferentes tonos es lo que conocemos como multiportadora.. Estas subportadoras utilizan una QAM para transmitir la información cada una de ellas.

La comunicación de datos de subida (“upstream”) suele ocupar 25 subcanales, mientras que la de bajada (“downstream”) ocupa entre 225 y 250 subcanales dependiendo de si se usan o no técnicas de cancelación de ecos.

El receptor se autoajusta para poder trabajar con un número óptimo de bits. Un subcanal puede llegar a ser eliminado si en su subbanda hay un nivel de ruido demasiado grande, por ejemplo.

La DMT es flexible, tiene una gran inmunidad al ruido y optimiza la velocidad de transmisión dependiendo de las condiciones de la línea. Si todo el ancho de banda del canal opera de forma consistente en todas las frecuencias podremos enviar el mismo número de bits en cada uno de los subcanales. Pero esto no será así, de forma que los canales que trabajen en dominios frecuenciales de alta calidad llevarán más bits que los que sean de baja calidad.

DMT es pues actualmente un conjunto de sistemas QAM trabajando en paralelo con cada subcanal correspondiendo a un tono frecuencial de una FDM. Esta modulación no sólo se ajusta a las condiciones espectrales del canal (forma en que el canal atenúa las

distintas frecuencias), sino que también se puede ajustar para evitar problemas causados por el ruido impulsivo, las interferencias por radiofrecuencia y el “crosstalk”.

Las normas de los diferentes organismos internacionales que han emitido documentos estandarizando la tecnología ADSL (ANSI, ETSI, ITU) han reconocido la modulación DMT como estándar para esta tecnología.

Diversos tests hechos comparando equipos que usan ambas tecnologías han puesto de manifiesto varias cosas:

- CAP necesita más potencia que DMT ya que es menos eficiente.

- CAP no es tan compatible espectralmente como DMT y puede causar interferencias en otros sistemas.

- CAP tiene peores resultados para velocidades y líneas equivalentes.

- Los equipos construidos utilizando la modulación CAP son muy poco compatibles, no así los que utilizan DMT.

También hay que decir que la técnica DMT no es recomendable si tenemos estrictos requerimientos de tiempo, si no requerimos versatilidad, adaptación y flexibilidad (puntos fuertes de esta técnica) o si hemos de ser muy eficientes en cuanto a potencia. De la misma forma CAP tuvo una clara ventaja: no tuvo que pasar por los organismos de estandarización, lo cual permitió un desarrollo mucho más rápido.

CONCLUSIONES

Como conclusiones podemos decir que la tecnología ADSL posiblemente sea la opción más barata y fácilmente implementable para suministrar un acceso de Banda Ancha a las redes de Telecomunicaciones al usuario. Esta tecnología se puede ver como un directo competidor del cable, aunque las operadoras de cable lo puedan usar para conseguir dar servicio a usuarios aislados, a los cuales todavía no haya llegado la infraestructura de la red de cable. Es una tecnología probada y regulada por estándares internacionales y existe la suficiente apertura en el mercado de servicios de telecomunicaciones como para que haya una seria competencia entre operadoras, hecho que provocará una bajada en el precio de las tarifas, posibilitando un acceso a estos servicios que, si no masivo, sea al menos bastante amplio.

REFERENCIAS

[1] <http://www.adsl.com>

[2] [http://www.telefonica.es:80/index/novedades/
datamegaviaadsl.html](http://www.telefonica.es:80/index/novedades/datamegaviaadsl.html)

[3] <http://www.3com.com>





SEGURIDAD DEL CLIENTE EN EL SERVICIO WORLD WIDE WEB

Felipe Moreno Strauch

*Estudiante de la ETSETB, UPC y socio colaborador de BJT.
felipe@bjt.upc.es*

INTRODUCCIÓN

El World Wide Web empezó como una forma sencilla de distribuir información, pero ahora se ha transformado en algo que va mucho mas allá. A medida que el Web se ha ido desarrollando las herramientas involucradas en el proceso de creación de las páginas también lo han hecho. Al lenguaje HTML se han añadido nuevas tecnologías como el JavaScript, el Java o el Active-X y los navegadores han aumentado en tamaño para ofrecer más funcionalidad.

Los navegadores ahora son programas complejos, con muchas líneas de código y esto tiene dos consecuencias en cuanto a la seguridad: a medida que aumentan sus prestaciones se vuelven más difíciles de configurar y a medida que su código se vuelve más complejo, son más susceptibles a tener fallos. Por ello, es muy importante dedicar algún tiempo a estudiar la configuración de estos programas y mantenerse informado sobre las posibles vulnerabilidades que pueden tener.

En este artículo, aclararemos algunos conceptos relativos su configuración que están relacionados con la seguridad, veremos algunos de los fallos conocidos más importantes que poseen los dos navegadores de uso mas extendido y propondremos algunas soluciones para navegar de forma más segura.

Cuando hablamos de seguridad en el navegador hablamos básicamente de dos puntos: la privacidad del usuario que navega y la integridad de su máquina. En cuanto a privacidad, nos referimos a navegar de forma anónima, divulgando la mínima información posible a la red y en cuanto a integridad, al acceso que los programas o scripts contenidos en las páginas Web pueden tener a los recursos de la máquina.

1. Helper Applications

Para ampliar la funcionalidad del navegador podemos configurar aplicaciones externas como

«helper applications» y asociarlas a tipos MIME que el navegador no sea capaz de tratar por sí mismo. De esta forma al recibir un documento de tipo desconocido, este consulta la lista de «helper applications» y si encuentra alguna aplicación asociada al tipo recibido la ejecuta y le pasa el documento.

En cuanto a seguridad es importante controlar que aplicaciones están configuradas como «helper applications» ya que esto puede suponer una posible vulnerabilidad. Por ejemplo si se configurará una shell «command.com» como «helper application» de un tipo MIME application/x-bat sería posible llegar a ejecutar archivos por lotes («.bat») enviados desde un servidor Web. Lo que debemos hacer es controlar a aquellas aplicaciones que son capaces de ejecutar macros o scripts y que están configuradas como «helper applications» (por ejemplo el Microsoft Word o el Microsoft Excel).

En general basta con configurar el navegador de forma que NUNCA se ejecuten aplicaciones externas de forma automática sino que se pregunte siempre antes (la típica ventana de «ejecutar» o «grabar en disco») para que el usuario siempre tenga la ultima palabra.

2. Plug-ins

Los plug-ins, como las «helper applications» sirven para ampliar la funcionalidad del navegador permitiéndole que interprete documentos de tipos que antes desconocía. La diferencia es que en lugar de aplicaciones externas, son módulos que se instalan sobre el propio navegador. El funcionamiento es muy parecido: al recibir un documento de tipo desconocido se consulta la lista de plug-ins instalados y se ejecuta el adecuado.

En cuanto a seguridad es importante tener en cuenta que no se debe instalar un plug-in si no se sabe exactamente para que sirve y que puede hacer. Además es recomendable instalar solamente plug-ins desarrollados por empresas conocidas como: Macromedia (Flash o Shockwave), RealNetworks

(Real Player), etc. Ya que en este caso la marca supone cierta garantía.

3. Cookies

Los «cookies» son pequeños trozos de información enviados por un servidor de páginas Web y almacenados por el navegador en el disco duro. El objetivo es subsanar la falta de estado del protocolo HTTP y permitir crear la ilusión de una sesión a lo largo de varias páginas visitadas (por ejemplo, para que el usuario solamente tenga que autenticarse una vez). El navegador controla el envío de los cookies hacia los servidores en función del atributo «dominio». De esta forma, el cookie solo puede ser enviado hacia un servidor en el mismo dominio del que lo ha generado (para evitar que un determinado dominio tenga acceso a cookies creados por servidores de otros dominios). Un mecanismo de control impide que se establezcan cookies con atributo «dominio» con dominios de alto nivel («top level domains») como el «.com».

Los cookies están muy relacionados con la privacidad, ya que muchos sitios los utilizan para hacer un seguimiento de las acciones del usuario. Por ejemplo, los buscadores como el Yahoo o el Altavista guardan todas las búsquedas realizadas por un usuario en una base de datos que utilizan para realizar marketing personalizado (la publicidad aparece en función de los intereses del usuario). Para poder realizar este seguimiento establecen un cookie en cada cliente donde almacenan un número identificador que sirve de índice en esta base de datos.

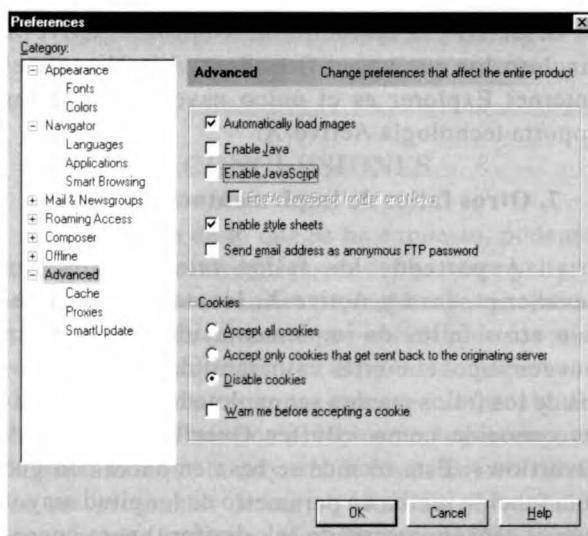


Figura1. Pantalla de configuración del Netscape Communicator (File-Preferences).

Para evitar este tipo de acciones lo más eficaz es desactivar la aceptación de cookies o al menos configurar el navegador para pedir una confirmación al usuario antes de aceptar uno. Por defecto, los navegadores tienen activada la opción de aceptar los cookies activada.

4. JavaScript

El JavaScript es un lenguaje de scripts desarrollado por Netscape para añadir más interactividad a las páginas Web. A principio se incorporaba solamente a sus navegadores pero a medida que aumentó su popularidad, Microsoft lo incorporó también al Internet Explorer (en la versión 3.0).

La historia de JavaScript está llena de ejemplos de agujeros de seguridad, pero dada la propia característica de ese lenguaje (no se puede cambiar la información almacenada en el disco) los fallos repercuten casi siempre en pérdida de privacidad.

Algunas de las vulnerabilidades relacionadas con JavaScript permiten obtener información sobre las páginas visitadas como por ejemplo el «JavaScript Cache Browsing Bug» (versiones Windows de Netscape 3.04, 4.07 y 4.5) que permite que un script sea capaz de leer la información almacenada en la memoria cache; o obtener información sobre la configuración del navegador como la «Preferences Bug» (versiones 4.0 a 4.04 de Netscape en todas las plataformas) que permite acceder al fichero «prefs.js» y obtener la dirección de email, servidores de correos y noticias, contraseña de correo y de FTP, etc.

Otros fallos de seguridad más importantes permiten que un script pueda acceder (lectura solamente) a los ficheros del disco duro de la máquina cliente. El «Injection Bug» (versiones 4.0 a 4.07 y 4.5PR de Netscape en todas las plataformas) permite acceder a la información almacenada en los cookies, determinar otras páginas visitadas y obtener listados de ficheros y directorios. Otros fallos permiten el envío de ficheros al servidor como el «The Cuartango Hole» o el «The Son of Cuartango Hole» que afectan a las versiones de Netscape 4.0 y 4.01 así como la versión preview del Internet Explorer 5. Otros fallos parecidos afectan también a la versión 4.0 del Internet Explorer. Estos tipos de fallos suelen utilizar formularios en frames ocultos para enviar los ficheros a un CGI mediante elementos de tipo «file» (INPUT TYPE=FILE). Para evitar la restricción de acceso que tienen los scripts a estos elementos se suele utilizar el método de «copiar-y-pegar».

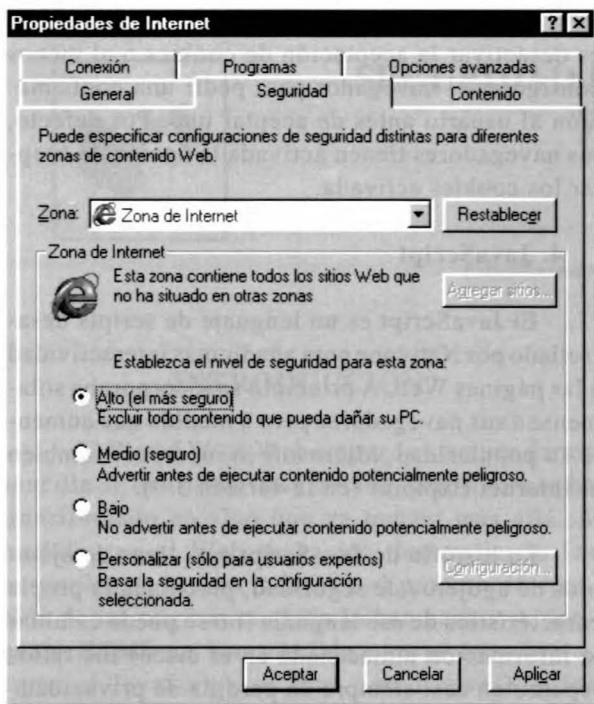


Fig.2 - Opciones de seguridad en «Panel de Control» - «Configuración de Internet».

5. Java

Aunque Java es un lenguaje de programación en todos sus aspectos, los programas hechos para su utilización en las páginas Web, los applets, tienen fuertes restricciones. Estas restricciones las controla un objeto «security manager» creado por el navegador que impide al applet ejecutar ciertos comandos de sistema, cargar ciertas librerías, acceder a ciertos recursos como el sistema de ficheros o establecer conexiones con otros servidores en la red a parte de su servidor de origen.

A pesar de ser seguro en la teoría, los fallos de implementación hacen con que en la práctica un applet pueda generar efectos desastrosos. Casi todas las vulnerabilidades relacionadas con Java permiten burlar al «security manager» y obtener acceso total a la máquina como el «Java Security Vulnerability» (Netscape 4.0 a 4.5 en todas las plataformas) del 29/03/99; el «Classloader Java Vulnerability» (Netscape 4.0 a 4.05 en todas las plataformas) del 14/08/98 o el «Virtual Machine SandBox Vulnerability» (Explorer 4.0 y 5.0) del 25/08/99. Este ultimo, al afectar la implementación de la máquina virtual de Java (JVM) puede ser explotado desde todos los programas de Microsoft que la utilizan (como el Outlook por ejemplo).

En caso de utilizar Internet Explorer es muy recomendable obtener la versión corregida de la JVM de Microsoft, que se puede bajar de Internet.

6. Controles Active-X

El Active-X es una tecnología desarrollada por Microsoft para distribuir software por Internet. De la misma forma que un applet de Java, un control Active-X puede ser incluido en un página Web. Los controles son distribuidos en código binario y por tanto han de ser compilados para cada posible plataforma cliente a diferencia del código precompilado de Java.

El modelo de seguridad utilizado por Microsoft para el Active-X es bastante diferente del modelo de Java. No hay ninguna restricción sobre que puede hacer un control sino que se permite que cada control sea «firmado» digitalmente utilizando un sistema llamado «Authenticode». De esta forma se puede comprobar el origen del control comprobando la firma y la autenticidad del certificado utilizando mediante una Autoridad de Certificación (los applets de Java también pueden «firmarse» mediante un mecanismo parecido).

El «Scriptlet.typeLib/Eyedog Vulnerability» del 31/08/99 que afecta las versiones 4.0 y 5.0 de Internet Explorer permite explotar el fallo de dos controles Active-X (el «scriptlet.typeLib» y el «eyedog») para crear y modificar ficheros, acceder al registro de Windows e incluso permitir la ejecución de código (realizando un ataque de «Buffer Overflow» sobre los métodos del control «eyedog»). Para corregir este problema se puede descargar un parche desde Microsoft, pero lo más recomendable es desactivar la ejecución de controles Active-X (incluso los que vayan firmados). Actualmente el Internet Explorer es el único navegador que los soporta tecnología Active-X.

7. Otros fallos de implementación.

A parte de los fallos relacionados con JavaScript, Java y Active-X, los navegadores tienen otros fallos de implementación que también pueden suponer ciertas vulnerabilidades. La mayoría de los fallos pueden ser explotados con la técnica conocida como «Buffer Overflow» o «Stack Overflow». Esta técnica se basa en hacer con que una función reciba un parámetro de longitud mayor que el espacio reservado (el «buffer») para conse-

uir sobre escribir la dirección de retorno almacenada en la pila («stack»). Así se puede sustituir esta dirección por la de otro punto de la memoria (en general un punto dentro del propio buffer) donde se pone el código a ejecutar.

El Internet Explorer 4.0 y 4.01 contiene varios de estos fallos relacionados con URLs o elementos HTML de longitud muy larga y el Netscape 4.03 y 4.04 presenta problemas al almacenar páginas con títulos muy largos en los bookmarks. El «Long Filename Mail Vulnerability» (Netscape 4.0 a 4.05 y PR1) es otro caso de «Buffer Overflow» y se puede explotar si un mensaje de email o news contiene como attachment un fichero con un nombre muy largo.

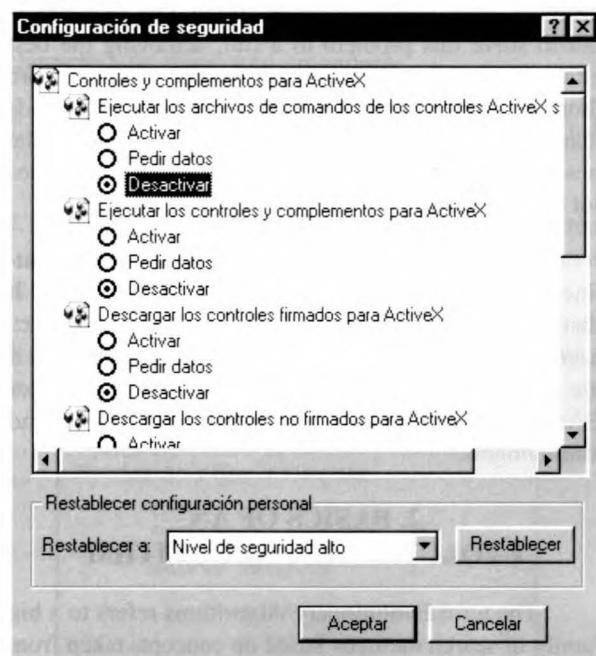


Fig.3 - Configuración de la opción Personalizar donde se puede cambiar los parámetros de Java, JavaScript y Active-X.

CONCLUSIONES

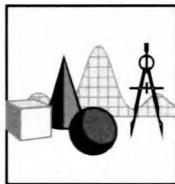
En vista de lo que se ha expuesto, podemos comprobar que navegar puede ser bastante peligroso. No debemos sentirnos protegidos por el aparente anonimato de ser uno entre millones de usuarios ya que estos ataques no van dirigidos a una persona, sino que son como trampas esperando a que alguien las active. La situación es incluso más preocupante si pensamos que la mayoría de los usuarios utilizan en sus programas de correo electrónico la opción que permite formatear mensajes con HTML ya que entonces las páginas «trampa» se pueden enviar por email y para activarlas basta con abrir el mensaje.

El resultado de conseguir ejecutar código en la máquina cliente puede tener serias consecuencias, no solo para la máquina local (borrar el disco duro, cambiar la configuración del sistema operativo, obtener información sobre el usuario), sino incluso para la propia red donde esta está conectada (obtener ficheros de password, contaminarla con virus, o instalar programas de control remoto o «sniffers»). Como la mayoría de los sistemas de firewall habilitan el tráfico HTTP, se podría atacar a una máquina utilizando uno de los fallos de los navegadores para instalar en ella un programa. Este programa podría funcionar como un cliente HTTP «especial» que se utilizaría para intercambiar datos con el exterior anulando el firewall. Sería una forma de salida para la información obtenida dentro de la red (ficheros, contraseñas, nombres de servidores) y además se podría utilizar como forma de entrada para realizar ataques dentro de la red (podríamos por ejemplo llegar a montar una sesión de Telnet por encima del HTTP).

Para evitar estos problemas, es necesario tener siempre las últimas versiones de los navegadores (4.7 para el Netscape Communicator y 5.0 para el Internet Explorer). De esta forma aunque no se podrá asegurar invulnerabilidad (incluso las últimas versiones tienen fallos), podemos garantizar que estamos protegidos de los fallos conocidos y difundidos por Internet, que son los más peligrosos. Además es necesario configurar adecuadamente el navegador y si se navega por páginas «potencialmente peligrosas», desactivar la ejecución de JavaScript, Java, Active-X y deshabilitar los cookies. Por último, también es importante, estar siempre informado de los fallos encontrados y de sus posibles soluciones o parches, lo que se puede hacer consultando los boletines de seguridad a través de Internet en las páginas de los propios navegadores.

PARA MÁS INFORMACIÓN

- [1] Página de seguridad de Netscape:
<http://home.netscape.com/security/notes/index.html>
- [2] Página de seguridad de Internet Explorer:
<http://www.microsoft.com/windows/ie/security/default.asp>
- [3] FAQ sobre seguridad del World Wide Web Consortium:
<http://www.w3.org/Security/faq/www-security-faq.html>
- [4] Secure Internet Programming Laboratory - Princeton University:
<http://www.cs.princeton.edu/sip/>



AN APPLICATION EXAMPLE OF THE BREEDER GENETIC ALGORITHM TO FUNCTION OPTIMIZATION

Lluís A. Belanche

Secció d'Intel·ligència Artificial

Dept. de Llenguatges i Sistemes Informàtics

Universitat Politècnica de Catalunya

belanche@lsi.upc.es

ABSTRACT

Evolutionary Algorithms (EA) have demonstrated their ability to solve optimization tasks in a wide range of applications. In this paper, after outlining the basics of such algorithms, the possibilities of one of the latest to emerge, the Breeder Genetic Algorithm (BGA) are exemplified by addressing a classical numerical optimization problem: the Fletcher-Powell pseudo-random function.

1. INTRODUCTION

Since the early days of computer science, function optimization has been one of the topics of more active research, because of the enormous potential of application to solve engineering or mathematical problems. For discrete solution spaces, traditional Artificial Intelligence (AI) techniques for solving constrained combinatorial optimization tasks, such as the A* or IDA* algorithms and derivations [1] have proven to be useful for certain applications. However, these methods rely on heuristics to guide the search process that are not always easy to find or to express. In addition, if the search space is big, the computational cost of the algorithm can become unacceptably high.

An alternative to these methods are Evolutionary Algorithms (EAs) [2], among which Genetic Algorithms (GAs) [3] have been extensively used for discrete as well as continuous optimization tasks, although they were not conceived for this [4]. Their main advantages over other methods are the coarsegrained *global search* mechanism, the neat tradeoff between *exploration* of the search space for new solutions and *exploitation* of the promising ones, the lack of an explicit heuristics, and the easiness of parallelization. In addition, no knowledge of the problem has to be coded apart from telling the algorithm how good (how *fit*, in EA jargon) a potential solution is.

Among the latest algorithms we find the Breeder Genetic Algorithm (BGA) [5]. It is characterized by a truncation selection procedure and direct representation of continuous variables. Truncation selection is a simple use of rank-based selection, proven to be very useful in traditional GAs [6]. The direct representation of variables eliminates the need for a coding scheme -that usually changes the search space- and permits to develop new, continuous genetic operators. The BGA has been shown

to be superior to traditional GAs for classical continuous optimization test problems [7].

We show in this paper how a relatively simple algorithm like the BGA is able to cope with the Fletcher-Powell function, an extremely difficult optimization problem, often used for benchmarking. Our intention is not to solve this problem to a full, achieving the best attainable solution, because this would require a more thorough study and much higher computational demands than those used here. Rather, we use it to illustrate the possibilities of EAs as black-box and robust techniques for optimization.

In the next two sections, the basics of EAs are outlined, and in particular the BGA is briefly introduced. It then follows a description of the selected problem. Next, a small experimental setup is devised and the results on it are commented on, together with a comparison to other EAs. The paper ends with some conclusions and afterthoughts.

2. BASICS OF AN EVOLUTIONARY ALGORITHM

The term Evolutionary Algorithms refers to a big family of search methods based on concepts taken from Darwinian evolution of species and natural selection of the fittest. Some concepts from genetics are also present. Given a problem to be solved EAs maintain a population of *individuals* that represent potential solutions to it. Each individual in the population is represented by a *chromosome* consisting of a string of atomic elements called *genes*. Each gene contains (represents) a variable, either for the problem or for the algorithm itself. The possible values of a gene are called *alleles* and the gene's position in the chromosome is called *locus* (pl. *loci*). There is also a distinction between the *genotype*, the genetic material of an individual, and the *phenotype*, the individual result of genotype development (that is, the born living thing). In EAs the genotype coincides with the chromosome, and the phenotype is simulated via a *fitness function*, a scalar value -similar to a reinforcement- expressing how well an individual has come out of a given genotype.

An EA can be formally described by the conceptual algorithm in g. 1, parameterized by a tuple:

$$< \text{EA-Setup} > = < \Pi_0, (\mu, \lambda), \Upsilon, \Omega, \Psi, \Theta, \Phi, \Xi >$$

where $\Pi_t = (i'_1, i'_2, \dots, i'_{\mu})$ is the population at time t and thus Π_0 is the, usually random, initial population, μ the population size, λ the offspring size (out of μ), Υ the selection operator, Ω the recombination operator, Ψ the mutation operator, Θ the termination criterion, Ξ the replacement criterion and Φ the fitness function. In this algorithm, operator sequencing on the population is as follows: Π_t represents the population at time (i.e., generation) t , Π'_t the population after selection, Π''_t after recombination and Π'''_t after mutation, to end in a new population Π_{t+1} .

The search process usually starts with a randomly generated population Π_0 and evolves over time in a quest for better and better individuals where, from generation to generation, new populations are formed by application of three fundamental kinds of operators to the individuals of a population, forming a characteristic three-step procedure:

1. *Selection* of the fittest individuals, yielding the socalled *gene pool*;
2. *Recombination* of (some of) the previously selected individuals forming the gene pool, giving rise to an offspring of new individuals;
3. *Mutation* of (some of) the newly created individuals.

```

Procedure Evolutionary-Algorithm
{
    t:=0;
    create  $\Pi_t$ ;
    evaluate  $\Phi(i)$ ,  $\forall i \in \Pi_t$ ;
    while not ( $\Theta(\Pi_t)$ ) do
    {
        /* Create the gene pool  $\Pi'_t$  */
        select:  $\Pi'_t := \Upsilon(\Pi_t)$ ;

        /* Apply genetic operators */
        recombine:  $\Pi_t := \Omega(\Pi'_t)$ ;
        mutate:  $\Pi''_t := \Psi(\Pi_t)$ ;

        /* Evaluate their effect */
        evaluate  $\Phi(i)$ ,  $\forall i \in \Pi''_t$ ;

        /* Form the new generation */
        replace:  $\Pi_{t+1} := \Xi(\Pi_t \cup \Pi''_t)$ ;
        t := t+1
    }
}

```

Figure 1: Evolutionary Algorithm.

By iterating this three-step mechanism, it is hoped that increasingly better individuals will be found (that is, will appear in the population). This reasoning is based on the following ideas:

1. The selection of the fittest individuals ensures that only the best ones¹ will be allowed to have offspring, driving the search towards good solutions, mimicking the natural process of selection, in which only the more adapted species are to survive.

2. By recombining the genetic material of these selected individuals, the possibility of obtaining an offspring where *at least* one child is better than any of its parents is high.
3. Mutation is meant to introduce new traits, not present in any of the parents. It is usually performed on freshly obtained individuals by slightly altering some of their genetic material.

Hence, an EA may be seen as a non-empty sequence of ordered operator applications: fitness evaluation, selection, recombination, mutation and replacement.

The entire process iterates until one of the following criteria is fullfiled:

1. *Convergence*: it happens because the individuals are too similar. Fresh and new ideas are needed, but recombination is incapable of providing them because the individuals are very close to one another, and mutation alone is not powerful enough to introduce the desired variability. Convergence can be monitored by on-line (average of the best individuals) and off-line (average of average individuals) throughout the generations;
2. *Problem solved*: the global optimum is found up to a satisfactory accuracy (if optimum known);
3. *End of resources*: the maximum number of function evaluations has been reached.

Evolutionary Algorithms are effective mainly because their search mechanism keeps a well-balanced tradeoff between *exploration* (trying to always drive the search to the discovery of new, more useful, genetic material) and *exploitation* (trying to fine-tune good already-found solutions). Exploration is mainly dealt with by the mutation operator. Exploitation is carried out by the selection process and the use of recombination operators, although mutation may also play a role in the netuning of solutions. The fitness function is built out of the function to be optimized (called the *objective function*). All EAs represent the decision variables in the chromosome in one way or another, either directly as real values or resorting to a discrete coding, usually binary (like most GAs). The particular coding scheme is the classical knowledge representation problem in AI, and completely conditions the results. In addition, some algorithms (like the Evolution Strategies, ES) append their own variables to the representation in the form of auxiliary information that evolves with time like the other variables.

An excellent state-of-the-art and review of EAs, and a useful departure point because of its rich set of references

¹Or the luckiest in some EA instances, like most GAs.



is [8]. There is also a very complete FAQ with pointers to papers, books, software and the main groups working on EAs all over the world [9].

3. THE BREEDER GENETIC ALGORITHM

In traditional GAs, selection is stochastic and meant to mimic -to some degree- Darwinian evolution; instead, BGA selection (named *truncation* selection) is a deterministic and artificial procedure driven by a *breeding* mechanism (as used in livestock), where only the best individuals-usually a fixed percentage of τ total population size μ - are selected and enter the gene pool to be recombined and mutated, as the basis to form a new generation. Genetic (recombination and mutation) operators are applied by randomly and uniformly selecting two parents until the number of offspring equals $\mu \cdot q$. Then, the former q best elements are re-inserted into the population, forming a new generation of μ individuals that replaces the previous one. This guaranteed survival of some of the best individuals is called *q-elitism*. For the BGA, the typical value is $q = 1$. The BGA selection mechanism is then deterministic (there are no probabilities), extictive (the best elements are guaranteed to be selected and the worst are guaranteed *not* to be selected) and 1-elitist (the best element is always to survive from generation to generation).

The BGA chromosomes are potential solution vectors \vec{x} of n components, where n is the problem size, the number of free variables of the function to be optimized. The common aspect of BGAs with ordinary GAs is the fact that both are mainly driven by recombination, with mutation regarded as an important but background operator, in the double role of solution fine-tuner (for very small mutations) and as the main discovery force (for moderate ones). We will now briefly describe the different possibilities for the genetic operators. The reader is referred to [10] for a detailed description.

3.1 Recombination

Any operator Ω combining the genetic material of the parents is called a recombination operator. In BGAs, recombination is applied unconditionally, $Pr(\Omega) = 1$. Let $\vec{x} = (x_1, \dots, x_n)$, $\vec{y} = (y_1, \dots, y_n)$ be two selected gene-pool individuals \vec{x}, \vec{y} such that $\vec{x} \neq \vec{y}$. Let $\vec{z} = (z_1, \dots, z_n)$ be the result of recombination and $1 \leq i \leq n$. The following are some of the more common possibilities to obtain an offspring \vec{z} :

1. Discrete Recombination (DR).

$$z_i \in \{x_i, y_i\} \quad (\text{chosen with equal probability})$$

2. Line Recombination (LR).

$$z_i = x_i + \alpha(y_i - x_i)$$

with a fixed $\alpha \in [0, 1]$. Typically, $\alpha=0.5$.

3. Extended Intermediate Recombination (EIR).

$$z_i = x_i + \alpha_i(y_i - x_i)$$

with $\alpha_i \in [-\delta, 1+\delta]$ chosen with uniform probability. The parameter δ expresses to what degree an offspring can be generated out of the parents's scope, the imaginary line that joins them in \mathcal{R} . More precisely, it works by controlling the maximum fraction $a = \delta[y_i - x_i]/(r_i^+ - r_i^-)$ of the distance between parents where the offspring can be placed, either left to the leftmost parent or right to the rightmost parent. A typical value for $\delta=0.25$, although any non-negative real number not exceeding 0.5 is a potential value. The bigger the δ , the more the effect of the parents is diminished in creating offspring. A method for dynamically setting its value called range, was introduced in [10] and shown to have a remarkable effect in performance. It works as follows:

$$z_i = y_i + \alpha_i(x_i - y_i), \quad \text{with } x_i \geq y_i$$

such that $\alpha_i \in [-\delta_i^-, 1 + \delta_i^+]$ with uniform probability and,

$$\delta_i^- = \frac{y_i - r_i^-}{r_i^+ - r_i^-}$$

$$\delta_i^+ = \frac{r_i^+ - x_i}{r_i^+ - r_i^-}$$

This procedure assigns different values for the left (δ_i^-) and right (δ_i^+) limits of the interval from which α_i is to be selected, and does never generate a value outside the range $[r_i^-, r_i^+]$ for the variable i , an aspect not fulfilled by the other methods that otherwise has to be dealt with a posteriori.

4. Fuzzy Recombination (FR). This operator, introduced in [11], basically replaces the uniform *pdf* (probability distribution function) by a bimodal one, where the two modes are located at x_i and y_i . The label "fuzzy" comes from the fact that the two parts $Pr_{x_i}(t); Pr_{y_i}(t)$ of the probability distribution resemble triangular fuzzy numbers.

3.2 Mutation

A mutation operator Ψ is applied to each gene with some probability $Pr(\Psi)=1/n$ so that, on average, one gene is mutated for each individual. Let $\vec{z} = (z_1, \dots, z_n)$ denote the result of mutation of an individual \vec{x} . The elements of \vec{z} are formed as follows:

1. Discrete Mutation (DM).

$$z_i = x_i + \text{sign}_i \cdot \text{range}_i \cdot \delta$$

with $\text{sign}_i \in \{-1, +1\}$ chosen with equal probability, $\text{range}_i = p(r_i^+ - r_i^-)$, $p \in [0.1, 0.5]$ and

$$\delta = \sum_{i=0}^{k-1} \varphi_i 2^{-i}$$

where $\varphi_i \in \{0,1\}$ from a Bernouilli probability distribution where $\Pr(\varphi_i = 1) = 1/k$. In this setting $k \in \mathbb{N}^+$ is a parameter originally related to the *precision* with which the optimum was to be located, a machine-dependent constant. In practice, however, the value of k is related to the *expected* value of mutation steps: the higher k is, the more fine-grained is the resultant mutation operator. The factor ρ is the *range ratio*, related to the *maximum* step that mutation is allowed to produce as a ratio of variable range.

2. Continuous Mutation (CM). Same as DM but with

$$\delta = 2^{-k\beta}$$

where $\beta \in [0, 1]$ with uniform probability.

4 THE FLETCHER-POWELL FUNCTION

This highly multimodal function was introduced in 1963 as a regression problem [12], where a collection of parameters have to be estimated such that a quadratic error term, depending on a non-linear expression, has to be minimized. Let us denote the n parameters as a vector \vec{x} . Let $A = (a_{ij})$, $B = (b_{ij})$ be two $n \times n$ real matrices and $\vec{\alpha}$ a fixed real vector of dimension n . Define the two terms A_i and B_i as:

$$A_i(\vec{\alpha}) = \sum_{j=1}^n (a_{ij} \sin \alpha_j + b_{ij} \cos \alpha_j)$$

$$B_i(\vec{x}) = \sum_{j=1}^n (a_{ij} \sin x_j + b_{ij} \cos x_j)$$

Then, the function is defined as:

$$FP(\vec{x}) = \sum_{j=1}^n (A_i(\vec{\alpha}) - B_i(\vec{x}))^2$$

The solution is $FP(\vec{x}^*) = 0$, where $\vec{x}^* = \vec{\alpha}$. The function is clearly non-symmetric (which would make it easier) and non-separable (it cannot be optimized on each x_j separately). The sinusoidal factors lead to a high number of hills and valleys, making a gradient-based method unfeasible. There are up to 2^n extrema located in the subspace $|x_j| \leq \pi$, the one usually selected to constrain the solution space. The vector $\vec{\alpha}$ (the position of the absolute minimum) is chosen at random from this subspace $\alpha_j \in [-\pi, \pi]$. Moreover, the matrices A , B are also chosen at random with $a_{ij}, b_{ij} \in [-100, 100]$. This makes the

suboptimal extrema to be randomly distributed over the search space. In our case, a value of $n = 30$ is chosen, so that the function includes a total of 1,830 random numbers.

5. EXPERIMENTAL RESULTS

An execution with the BGA involves the choice of a mutation operator Ψ and its parameters ρ and k and of a recombination operator Ω and its parameter δ (only for EIR). The truncation threshold τ and population size μ are also to be set. These last two values are of great importance because they are strongly interrelated. For the sake of clarity and simplicity we do not discuss them here. Instead, an educated guess is made just in order to obtain reasonable results within low computing resources. To this end, a value of $\tau = 15$ is set, to ensure a high selective pressure so as to enforce the quick discovery of fairly good solutions. The low value of τ entails a moderate-to-high population size in correspondence, to account for enough diversity. We set then $\mu = 100$.

The choice of genetic operators is more elaborated. To begin with, ρ is set to 0.5, a value that stands for high average mutation steps, needed to broad-tune rather than to refine solutions. In addition, a low value of k (lower than 16) is in favour of this scheme, and known to have a much deeper impact on performance [10] than , so we carried out some preliminary runs with $k \in \{6, 8, 10, 12, 16\}$ that lead to the selection of $k = 10$. In addition, the continuous operator showed to be better.

All this tuning runs take only a few seconds (always less than a minute) in a shared SUN™ Ultra-60 System. The stopping criterion is based on the number of function evaluations permitted (given by the variable `FFEvals`). In particular, given a finite number of `FFEvals`, the algorithm will stop each run whenever `[FFEvals / μ]` generations are reached. This stopping criterion allows to compare different general settings in a fair way, since, for example, a smaller population would be allotted more generations, but always keeping the number of evaluations in similar values. For each configuration, a number of independent runs are performed - denoted by `NRuns` - keeping track of the mean and best solutions found. In these initial experiments, `FFEvals=50,000` and `NRuns=5`.

Regarding the recombination operator, ten possibilities are tested from within the set {DR, LR with $\alpha = 0.5$, EIR with $\delta = 0$ to 0.30, EIR with $range_\delta$ and FR}. Among them, the EIR operator quickly stands out over the rest, with varying performance that depends on its parameter δ . The results for this last operator are



presented in a single table for ease of reading -Table 1- summarizing the information as a function of the sample values for tested. For each configuration, average and best solutions found are shown.

EIR (0)	Average	Best
0.05	51,541	18,296
0.10	28,837	23,151
0.15	25,368	2,265
0.20	32,459	11,798
0.25	41,262	14,017
0.30	47,356	20,561
range _d	25,953	9,030

Table 1: Results for the EIR (δ) recombination operator. Each entry shows the average and best results across five runs.

As it can be seen, EIR (0.15) shows to be the best setting, both on average and in terms of the best solution found. Note that the curve of performance across is concave, with EIR (0.15) at the bottom. The operator with the modification range shows to have a performance almost equal than the best one achieved with a fixed δ . This is remarkable because it means that a comparable performance can be obtained without the need of a search along . Both settings ($\delta = 0.15$ and $range_d$) are selected for further experiments. It has to be said that the already obtained results are quite good. The initial function evaluations are in the order of millions (between one and five, depending on the run).

Once the algorithm has been roughly tuned to the problem landscape, a series of runs are performed to assess its potential to a deeper degree. Operators are then set to continuous mutation with $\rho = 0.5$, $k = 10$ for mutation, and the two choices EIR ($\delta = 0.15$), EIR ($range_d$) for recombination, using the knowledge gained so far; this time, however, $NRuns = 100$ runs of $FFEvals = 100,000$ are carried out. The results are presented collectively in Table 2, along with some other results present in the literature.

Bäck [2] (p. 157), for example, reports results on this function with a standard genetic algorithm (SGA), an evolutionary programming algorithm (EP) and the powerful evolution strategies (ES), reproduced in the table for

Algorithm	Average	Best
SGA	$4.581 \cdot 10^4$	$1.032 \cdot 10^4$
EP	$1.107 \cdot 10^5$	$2.997 \cdot 10^4$
BGA-1	$1.987 \cdot 10^4$	$8.165 \cdot 10^2$
BGA-2	$9.238 \cdot 10^3$	$8.649 \cdot 10^2$
ES	$1.749 \cdot 10^3$	$3.190 \cdot 10^1$

Table 2: Comparative results found by some evolutionary algorithms. BGA-1 is with EIR (0.15), and BGA-2 with EIR ($range_d$)

convenience. The first notable point is the inferior performance of the SGA and EP compared to the other algorithms. Eiben and van Kemenade [14] report also some results for a GA, product of their study on diagonal and n-point crossover² operators. They show these operators to be generally superior to traditional 2-parent, 1- or 2-point crossover. Although they do not provide numerical results, this improvement seems to make the GA vary between 30,000 and 10,000 on average, occasionally going below this mark. Altogether then, the binary GA is not likely to go beyond this order of magnitude, unless heavily modified or tailored to the problem.

With respect to the BGAs, both have improved their average performance, due to the double number of evaluations they have been allowed. Also noteworthy is the superior average performance of EIR ($range_d$) over EIR ($\delta = 0.15$). We believe that, in the long run, the first method is likely to widen the gap because of its adaptive nature.

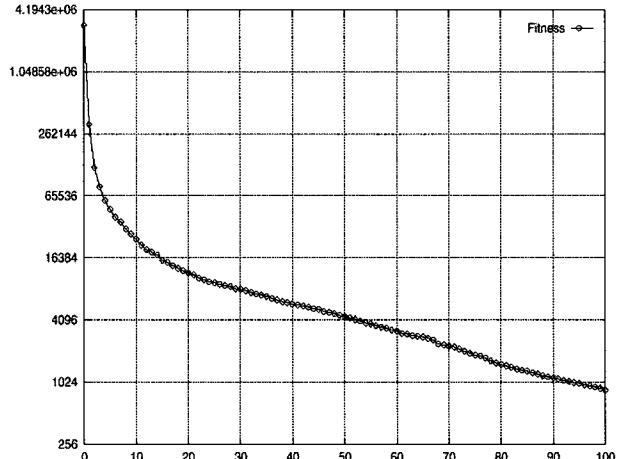


Figure 2: Performance curve for BGA-2 (EIR with range). The x-axis shows the number of generations (in multiples of ten). The y-axis shows the fitness in a log2 scale. Note the quasilinear slope and the fact that further progress was still being achieved. The error decrease from the initial random guess to the final solution is in four orders of magnitude.

Note that, very roughly, both departing from similar average values, doubling the number of function evaluations has reflected in an average result for EIR ($range_d$) that is twice as good as that of EIR ($\delta = 0.15$). However, the best absolute result is very similar for both, possibly indicating a limit in what can be achieved with the allowed resources and the selected μ, τ setting. A plot of the run finding the best solution for BGA-2 is shown in Figure 2.

The results reported by Bäck for the ES are by far the best, both on average and, especially, in the best solution found, which is excellent. This can be explained in several ways. First, ES [13] is a far more complicated algorithm that takes into account the interactions between variables more explicitly by selfoptimizing the amount of mutation

²Crossover is a form of recombination used in discrete GAs.

necessary. It is nonetheless very sensitive to its parameters so that tuning it is a more difficult task. However, once these are correctly set, ES performance is often superior to other algorithms. Second, it is interesting to note that the data for Bäck's experiments were obtained in similar conditions ($\mu = 100$, $\tau = 15$ and 100 runs) but resorting to 200,000 evaluations. The random initialization for the function is also a different one, which could well result in an overall easier or more complicated landscape. Bäck also reports the great variability found: of the 100 runs, only 44 were "successful runs", defined as those reaching a final value under 400.0 (compare it with the best result obtained: 0.319). In our case, for example, the standard deviation was computed to be 16792.8 for BGA-1, and 6563.7 for BGA-2. To see whether the BGA would be able to better its performance, we carried out a last experiment, a single run of BGA-2 with 200,000 evaluations. The result was 358.379, a value qualifying as successful.

6 CONCLUDING REMARKS

The quest for general-purpose search mechanisms is still an open and very active field. Since the 70's, new and powerful heuristic methods have emerged that are particularly well suited for function optimization -although this was not exactly their original purpose- mainly because of their generality, robustness, and conceptual (though not necessarily analytical) simplicity. Three of these methods are Simulated Annealing (SA), Tabu Search (TS) and Evolutionary Algorithms (EA). In the last decade, these properties and their remarkable successes have boosted their widespread use. The EA family is the biggest and subject of continuous improvement. In addition, many classical and modern problems are being reinstated and explored under the light of these methods. Artificial neural networks are a good representative of this. In particular, the training process of a supervised feed-forward network can be easily cast as a function optimization problem.

In this paper we have given an impression of how different EAs can cope with a well-known task, one for which methods of non-linear optimization are prone to end up in local minima of the function. Among the former, the Breeder Genetic Algorithm (BGA) is relatively simple and yields reasonably good solutions in a very limited time. Note that although it has not been our intention to solve the task at hand, and only a small experiment setup has been devised for it, the quality of the solutions and the promise of better ones show the BGA (and EAs in general) as feasible alternatives for a great variety of tasks, ranging from the mentioned neural network training problem [15] to aerofoil design in Aerodynamics [16]. Work is in progress toward a thorough application of these methods in a principled way.

REFERENCES

- [1] Pearl, J. Heuristics. Intelligent search strategies for computer problem solving. Addison-Wesley, 1984.
- [2] Bäck, Th. Evolutionary Algorithms in Theory and Practice. Oxford University Press, New York, 1996.
- [3] Goldberg, D.E. Genetic Algorithms for Search, Optimization & Machine Learning. Addison-Wesley, 1989.
- [4] Holland, J.H. Adaptation in natural and artificial systems. The University of Michigan Press. Ann Arbor, MI, 1975.
- [5] Mühlenbein, H., Schlierkamp-Voosen, D. Predictive Models for the Breeder Genetic Algorithm. *Evolutionary Computation*, 1 (1): 25-49, 1993.
- [6] Whitley, D., The GENITOR Algorithm and Selection Pressure: Why Rank-Based Allocation of Reproductive Trials is Best. In Procs. of the 3rd Intl. Conf. on Genetic Algorithms, (ed.) Schaefer, J.D., Morgan Kaufmann: San Mateo, 116-121, 1989.
- [7] De Falco, I., Del Balio, R., Della Cioppa, A., Tarantino, E. A Comparative Analysis of Evolutionary Algorithms for Function Optimisation. In Procs. of the Second Workshop on Evolutionary Computing (WEC2), Nagoya, JAPAN.
- [8] Bäck, Th., Schwefel, H.P. An Overview of Evolutionary Algorithms for Parameter Optimization. *Evolutionary Computation*, 1 (1): 1-23, 1993.
- [9] Heitkötter, J., Beasley, D. (Eds.). The Hitch-Hiker's Guide to Evolutionary Computation: A list of Frequently Asked Questions (FAQ). USENET: comp.ai.genetic. Available from anonymous FTP at rtfm.mit.edu, on /pub/usenet/news.answers/ai-faq/genetic.
- [10] Belanche, Ll. A Study in Function Optimization with the Breeder Genetic Algorithm. LSI Research Report LSI-99-36-R. Dept. de Lenguajes i Sistemes Informatics. Univ. Politecnica de Catalunya, 1999.
- [11] Voigt, H.M., Mühlenbein, H., Cvetkovic, D. Fuzzy recombination for the continuous Breeder Genetic Algorithm. In Procs. of the 6th Intl. Conf. on Genetic Algorithms, (ed.) L. Eshelman, Morgan Kaufmann: San Mateo, 104-113, 1995.
- [12] Fletcher, R., Powell, M.J.D. A rapidly convergent descent method for minimization. *Computer Journal*, 6:163-168, 1963.
- [13] Bäck, Th. Evolution Strategies: An alternative Evolutionary Algorithm. Technical Report of the Informatik Centrum, Dortmund.
- [14] Eiben, A.E., van Kemenade, C.H.M. Diagonal Crossover in Genetic Algorithms for Numerical Optimization. *Journal of Control and Cybernetics*, 26(3):447-465, 1997.
- [15] Yao, X. A Review of Evolutionary Artificial Networks. *Intl. J. of Intelligent Systems*, 8(4): 539-567, 1993.
- [16] De Falco, I. An introduction to Evolutionary Algorithms and their application to the Aerofoil Design Problem - Part I: the Algorithms. Part II: the Results. Invited papers at the von Karman Lecture Series 1997, Bruxelles. Both papers can be retrieved from the IRSIP Web Pages <http://www.irsip.na.cnr.it/~hotg>.





EL ABANDONO DE “LIS ESTELES”

Miguel Escudero

Profesor del Departamento de Matemática Aplicada y Telemática de la UPC
escudero@mat.upc.es

Las posibilidades de comunicación dictarán la sentencia. El próximo siglo -sería solemne en exceso decir milenio- asistirá a una convergencia de las culturas oriental y occidental. De estas imbricaciones lo que interesa es la fecundidad, esto es, la potencia que depare a la vida personal. Si cultura es, como entendía Ortega, ese lugar donde podemos trasladar nuestras entrañas, entonces es claro que la primera misión de cada ser humano en ese camino será *identificarlas*, hacernos a la idea de que *nosotros somos alguien y, por eso, ellos*, los otros humanos, también.

El monje católico norteamericano Thomas Merton (1915-1968) escribió a propósito del sabio chino Chuan-Tzú que la vida en sociedad trae como secuela confundir nuestra existencia, pues nos lleva a obsesionarnos con lo que no somos y a olvidar *quienes somos* realmente. Todo podría quedar como un mero juego de palabras si no especificásemos que “yo soy yo y mi circunstancia, y si no la salvo a ella no me salvo yo”. El “yo”, individual o social, no se salva si no se atiende a lo que nos rodea, siempre cambiante. El hombre no tiene naturaleza, sino que tiene historia; es persona, es el animal que tiene vida humana, a pesar de que no la aproveche o de que sea tratado sin compasión. Los pueblos no tienen identidad, sino herencias con las que construir el futuro. *Nada humano me es ajeno* es el lema de todo pueblo que sabe prolongarse.

Con un punto de sosiego, y curándonos de la ansiedad que nos abruma, podríamos advertir que “todo el cosmos es un grano de arroz, y la punta de un cabello es grande como una montaña”. Una sentencia con *sabor oriental* pero compatible con el estilo occidental, así nuestro poeta Maragall decía creer que “el hombre más sabio del mundo es aquel cuyos ojos, habiendo visto mucho, conservan del todo la visión del niño de antes de ir a la escuela”. De este modo, acabamos por cruzarnos con la música del lenguaje y la mística.

El propio Maragall escribió a principios del siglo XX un artículo de prensa profundamente personal y que merece ser leído íntegramente, “*Las lenguas francesas*”, en donde busca *la niñez fermentada de las lenguas*, esto es, en clave orteguiana, su poesía. Habla Juan Maragall (así se firmaba cuando escribía en castellano) de su enamoramiento de las variantes múlti-

ples y encantadoras de un mismo lenguaje áspero y dulce, *libremente matizado*. Camino del pueblo gascón de Gavarnie en la oscuridad de la noche, nos cuenta que acercándose a Gedre, “pueblocillo que lleva un nombre bonito” surgió “como una pequeña hada, una niña de cinco o seis años, pidiendo limosna -¿podíamos negársela?-”. La hicimos hablar por el gusto de oírla. -¿Cómo llamas tú a las estrellas? -le preguntamos (a aquellas alturas sólo se ocurre nombrar cosas grandes y maravillosas). -*lis esteles*- contestó con su vocecilla de hada en el infinito silencio. *¡lis esteles!*

Unamuno, entrañable amigo del escritor catalán (Juan Maragall), pedía oír en el silencio los ecos dulces de la niñez lejana como rumor de aguas frescas y vivas, esto es, con afán de recuperar el tiempo pasado.

Alzamos los ojos al cielo y las estrellas nos parecieron brillar con nueva luz del inmortal misterio”. Y explica que la dulce libertad del verbo pirenaico le penetró deliciosamente en el hondón del alma. Concluye exclamando: “¡Cómo quisiera meter este sentimiento en las entrañas de nuestros hermanos pirenaicos que desprecian su *patois* (*patois*, pero suyo) por la hermosa lengua de Racine (hermosa, pero de Racine)!”. Es el orgullo de una lengua *secundaria* que se defiende de toda ideología por *propia* que sea; con presión oficial y chantaje afectivo hay en la España de hoy casos de domesticación y sometimiento mediante la lengua, de este modo se estrangula el pensamiento, la expresión libremente matizada.

Unamuno, entrañable amigo del escritor catalán, pedía oír en el silencio los ecos dulces de la niñez lejana como rumor de aguas frescas y vivas, esto es, con afán de recuperar el tiempo pasado. Pero “¡Santa sencillez! Una vez perdida no se recobra”, confiesa en su *Diario Intimo* (publicado por primera vez en 1970). No obstante, creyendo que hay que vivir con toda el alma, se propone vivir en adelante “obsesionado en salvar mi alma”.

Demasiado anunciado, ¿a quién le seduce el plan del medieval místico alemán que recomendaba: «Acuéstate a la noche como si fuesen a enterrarte a la mañana, y levántate por la mañana como si hubieran de enterrarte a la noche»? ¿Quién lo podría soportar siguiendo el consejo de forma literal? Otra cosa es sentenciar: «Cuanto más vivas en Dios más en tí mismo vivirás, más dentro de tí mismo, y serás más tú». Hay en las páginas de esos cuadernos de Miguel de Unamuno dos frases, en particular, que resultan hoy día insólitas por su franqueza y decencia. En la primera, el escritor vasco, despierto de su vanidad, afirma: «He vivido soñando en dejar un nombre». Claro está que esto no es negativo en sí mismo, salvo que nos reduzca la autenticidad personal. Por su parte, la segunda frase es un diagnóstico singular, de valor imprecededor: «Estoy muy enfermo, y enfermo de yoísmo».

La cultura no sería entonces el lugar o la posada donde podemos trasladar nuestras entrañas, sino el zurrón que llevamos en nuestro itinerario íntimo: una bolsa con poco peso al costado de quien está dispuesto a quemar las naves fantasmas de la vanidad y llegar a puerto para bañarse como naufrago de la vida

Para sobreponerse de esa caída hay que seguir una ruta de soledad, donde lo importante no es llegar primero sino saber llegar para *volver a vivir*. La cultura no sería entonces el lugar o la posada donde podemos trasladar nuestras entrañas, sino el zurrón que llevamos en nuestro itinerario íntimo: una bolsa con poco peso al costado de quien está dispuesto a quemar las naves fantasmas de la vanidad y llegar a puerto para bañarse como naufrago de la vida. Algebra de estas metáforas es Compostela, hacia ella conduce por los campos, peregrinando, un recorrido mágico y magnético. El camino alfombrado de estrellas se *inauguró* con las debidas licencias eclesiásticas en el siglo IX, pocos años después de que un ermitaño observara una lluvia de estrellas y la asociara con el anuncio divino del oculto sepulcro del apóstol Santiago. Hoy día, el camino de los peregrinos, nombre específico de los que se dirigen a la localidad gallega para acabar dando un abrazo al busto del hijo del Zebedeo, está de moda. Las razones de tal peregrinaje han evolucionado con los siglos, pero ahora ¿qué se busca con esa marcha?

«Deportivamente» se aproxima a las tan concurridas carreras pedestres por las urbes, pues de lo que se trata es de participar y llegar a la meta en ambiente masivo de fiesta. Pero en este caso aparece un componente religioso intermitente. Se sea creyente en Cristo -con mayor o menor fervor, con mayor o menor coherencia- o no se sea,

se acaba el *juego* llegando con gozo al adorado templo. Muchos evadidos de las ciudades, acaso con más de un desespero a cuestas, buscan renovarse, sentirse capaces de alguna gesta reconocida, huyen de la anomia de las ciudades, cuyas luces deslumbrantes impiden ver. Sin embargo volverán a casa con indefinible saudade al finalizar su visita por esos terrenos, semejantes a los ocupados hasta no hace mucho por sus antepasados rurales, y que han hecho suyos con su andar. La presencia patente y benigna de la historia fortalece el deseo.

Es evidente que sólo el tiempo libre nos ha permitido seguir el *camino*, un camino también ruta de comunicación social y económica. Y aquí nos encontramos con las entreveradas razones líricas del ser humano. Lo más hondamente interesante que hace un hombre con su vida no es para hacer negocio. Por eso conviene reivindicar el ocio como tiempo personal, reconquistarlo de su imagen más engañosa, el cebo de un negocio organizado con ávidos buscadores de dinero y placer. Maragall se refirió lúcidamente al ocio como «la condición indispensable de ciertos descubrimientos científicos y de la producción poética, artística. La condenación del ocio conduciría a resultados desastrosos. Porque ocio se llama, por ejemplo, la conversación y la lectura; y, sin embargo, una y otra son las que forman principalmente la opinión y las costumbres». El ocio nos permite así el tránsito de la holganza a la holgura en la vida. Con sosiego siempre podemos reconsiderar las cosas y las personas, el quehacer vital.

Nunca acaban de desaparecer las estrellas, tarde o temprano reaparecen. Pero no podemos quedar prendidos de por vida a una hoja, analógica o digital, examinando el anuncio de la hora afortunada, los horóscopos quedan la nota de los vividores. Ansiosos de forma crónica por el espectáculo universal, hemos celebrado hace pocos días que el Sol abandonase la escena y desertase de su papel. Esto significa el término de origen griego *eclipse*: deserción, abandono, desaparición. La mayoría ha asistido al evento con atropellado y frustrante jolgorio; otros han confiado en el desastre final. Entre aquellos ha sido inevitable para algunos sentir la «injusticia cósmica» de la distribución del oscurecimiento solar; en España, la capital menos *afortunada* fue, según leo, Las Palmas de GranCanaria, con un 23 por ciento, mientras que San Sebastián y Santander encabezaron la lista con un 78 por ciento. Ha habido asimismo pueblos enteros de la Tierra que, dominados por la ignorancia e inconscientes del saber científico, han tomado al pie de la letra sus mitologías. Aunque estemos a años luz de la comprensión de la mecánica celeste y de la ciencia astronómica, ¿cómo podemos suplantarla? En su camino por la vida el hombre ha de saber arrastrar su *sombra lunar*, sus carencias y deficiencias. Desde ese *saber no saber*, no abandonará su guía, a veces invisible, de «los esteles».

(Publicado en la revista de pensamiento «Cuenta y Razón»).

