

# EDITORIAL

Después de un gran esfuerzo, por fin os presentamos una nueva entrega de Buran. Como podréis comprobar, este año se ha realizado un cambio de junta en la Rama de estudiantes del IEEE. Para todos nosotros es nuestra primera Buran, aunque para nuestros lectores más fieles ya sea su vigésimo segunda entrega. Es por ello, que hemos tenido un pequeño retraso sobre el plazo previsto de salida. A favor nuestro tenemos que decir, que tenemos muchas ideas nuevas para mejorar Buran. Esperamos que nuestras propuestas presentes y futuras sean de vuestro agrado.

Antes de continuar, debemos agradecer la ayuda prestada por los antiguos miembros de la rama. Su ayuda y su paciencia han hecho posible que esta edición salga a la luz. Agradecemos también la ayuda prestada por Juan Carlos Pimentel de la Rama de estudiantes del IEEE de la universidad Ricardo Palma de Lima (Perú).

En la penúltima página de la revista, se incluyen con detalle las novedades introducidas en esta edición. Como adelanto, os mencionamos que a partir de ahora Buran incluye una sección llamada: Hombres escondidos en fórmulas, donde gracias a la ayuda de Miguel Escudero, nos adentraremos en los aspectos históricos más interesantes de los científicos y matemáticos más relevantes de los últimos tiempos. Además a partir de esta Buran ya incluimos junto al número de depósito legal, el número de ISSN. Lo cual reportará un aliciente añadido para todos aquellos que quieran publicar en nuestra revista. Damos las gracias a Josep Pegueroles por aconsejarnos sobre el tema.

En fin, como habréis visto, esta edición viene cargada de novedades. Esperamos que obtengan la acogida esperada, y seguir trabajando para mejorar Buran día a día.

Raúl Cortés Delgado  
Vicepresidente de la Rama  
Coordinador de Buran

## COMITÉ EJECUTIVO DE LA SECCIÓN ESPAÑOLA DEL IEEE

**Presidente:** Prof. José Antonio Delgado-Penín  
**Vicepresidente:** Prof. Manuel Sierra-Pérez  
**Secretario:** Prof. Eduardo Bertrán-Albertí  
**Tesorero:** Prof. Andrés de Santos-y-Lleó  
**Responsable de la actividad de "Membership development":** Prof. Magdalena Salazar-Palma  
**Responsable de relaciones con las Ramas de estudiantes:** Ing. Juan-Jesús Rodríguez-Yubero

## COORDINACIÓN BURAN

Raúl Cortés Delgado

## RAMA DE ESTUDIANTES DE BARCELONA

**Presidente:** Jordi Núñez Vera  
**Vicepresidente:** Raúl Cortés Delgado  
**Secretario:** Guillem Fernández Solá  
**Tesorero:** Alfredo C. López Salcedo  
**Colaboradores:** Ricardo Blasco, Daniel Navarro, Carlos Hernández, Xavier Bielsa, José Luis Hernández, Felipe Calderero, Eduard Calvo, Juan Carlos Pimentel  
**Diseño Portada:** Jordi Núñez Vera

## REVISIÓN

Carles Gómez Montenegro, José A. López Salcedo, Miguel Angel Sastre, Josep Paradells

## AGRADECIMIENTOS

II. Dir. Juan A. Fernández Rubio, Ángel Cardama, Jose A. Delgado-Penín, Jorge Luis Sánchez-Ponz y a los puntos de distribución en la UPC: Abacus, CPET, y Kiosk Campus Nord.

We would also like to thank Ms. Laura Durrett (IEEE Student Services Manager), and IEEE International for their helpful support, encouragement and financial funding for distributing Buran across south american Region 9 IEEE Student Branches.

## IMPRESIÓN

RET, s.a.l.

## FOTOMECAÑICA

Sistemes d'Edició

## DEPÓSITO LEGAL

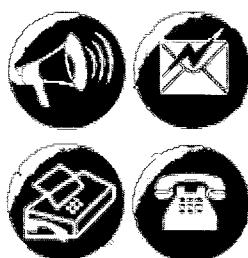
B-19.950-96

## ISSN

1698-7047

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.



# PETRA: ADVANCED ORAL INTERFACES FOR UNIFIED MESSAGING APPLICATIONS

David Hernando, Javier Hernando

and Xavier Anguera

{aldhd,javier,xanguera}@gps.tsc.upc.es

Department of Signal Theory and Communications

Universitat Politècnica de Catalunya (UPC)

## ABSTRACT

A new unified messaging system which integrates voice messages, fax and e-mail in a common mailbox has been developed. The combination of speech and linguistic engineering advances allows a natural oral man-machine interaction with the user when accessing the messages by phone. Advanced features are supplied such as texts summarization, messages classification and notification through the phone of new messages received.

## 1. INTRODUCTION

Nowadays, the use of e-mail, voice mail and fax has become indispensable in any working environment. In all cases, a dependency exists between each message type and its associated electronic device. The user must use a computer to access the e-mail, a phone to listen the voice mail and a fax machine to manage the faxes.

We present Petra\*, which solves this dependency, providing a common mailbox and two access methods easily available, webmail and telephone.

The whole work included three work lines:

1. Integration of phone, internet and fax services.
2. Development of an advanced dialogue system which brings the user a friendlier access than the provided by a system working only with the DTMF tones.

\*The project has been funded by the Spanish Government (CICYT TIC-2000-0335) and is related to the European project Majordome (E!-2340).

3. Intelligent information management for text classification and summarization.

The convergence of speech processing and natural language processing technologies is a crucial factor in the development of such a system, which requires the concourse of a wide range of knowledge and experience in linguistic engineering.

The paper is organized as follows. Firstly we present the main aspects of the unified messaging platform developed. In section 3, the phone access system to the messages is discussed. Evaluation results are showed in section 4. We finish with some conclusions.

## 2. THE UNIFIED MESSAGING SYSTEM PETRA

The new system presented gives the user two alternatives to access his messages: a webmail and a telephone. Features such as mail filtering and classification, notifications of new mail received, a dialogue driven by a natural man language or the text summarization, added to the usual functionality of a mail agent, convert Petra in an advanced unified messaging system.

The architecture of the demonstrator platform installed at our lab is shown on figure 1. This demo system is accessible by registered users from the public telephone network and from Internet. The external mail accounts from the users can be also accessed from Petra and are treated as extra folders.

The filtering module includes an automatic classifier that, for demo purposes, has been trained with a Spam messages corpus to detect this class of

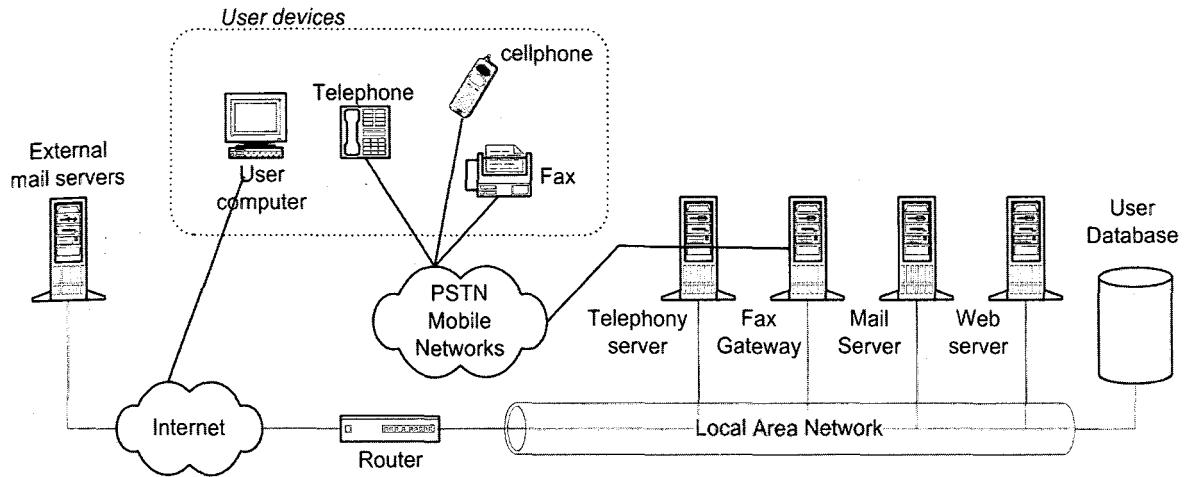


figure 1: System architecture

mail. Alerts, when a new message complies the filter's rules, are received by the users via a short message in the cellphone or via a phone call.

### 3. PHONE ACCESS SYSTEM

The phone access system is based on six main modules as seen in figure 2. A CTI board from Intel Dialogic is used as communication interface. Verbio (Verbio, 2004) libraries provide speech recognition and text-to-speech functionality. A speaker verification part has been developed and integrated for security authentication purposes and a new dialogue manager was built. Finally, extern data modules are needed, mainly mail servers and user database.

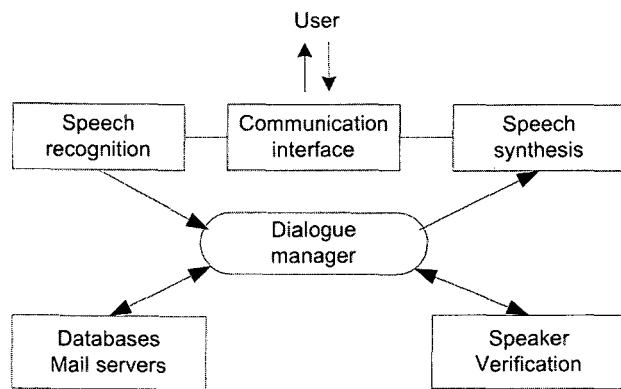


Figure 2: Phone access modules

#### 3.1 Dialogue Manager

This module offers the user an easy and efficient way to listen and send messages using a natural dialogue with the system. The design and implementation of the new dialogue manager is based on the one used in Attemps (Padrell and Hernando., 2002), a previous project giving real-time meteorological information by phone.

Recognition of user utterances is based on ABNF grammars. In each stage of the dialogue, the group of active grammars define the user's possible answers. Some grammars, which contain user specific data, as his folders or contact addresses, are dynamically generated after the authentication process. Some commands are enabled all the time to offer help, add recognition error correction or to give access to the main menu.

An experienced user is able to interrupt the system to answer a question before it is finished, completing tasks in a minor time (although barge-in can degrade recognition performance in noisy environments).

The confidence level of an user utterance returned by the recognizer is used to decide between two confirmation policies. For levels higher than a threshold, last answer is confirmed implicitly, asking also for next data. If lower, or in critical situations, as in a message deletion, an explicit confirmation policy is used.



Oral response generation is based on sentence selection from patterns, which are combined and completed with particular session information to give the desired meaning to the sentence. In confirming the answers, the system creates sentences with the same form as the user when referring to dates or numbers (for example, "the 13th of May" or "next Thursday"), avoiding confusing the user as in:

U: send the message next Thursday?  
 S: Do you want to send the message on May 15th?

For most common responses, the dialogue manager answers with one of multiple equivalent sentences, selected randomly each time it is required, increasing this way the sensation of a natural dialogue.

Besides the help provided when an user asks for it, the system follows an automatic helping strategy. In a first turn, the manager faces the user with a short and open sentence. If the user answer is wrong, the system asks for a second time with a more explicit question, showing the possibilities and, in further turns, giving the exact words to say or the key tones that user can press instead of his voice.

### 3.2 User authentication

In order to ensure confidentiality of the information, the user needs to sign in using a speaker verification system. The system we present can be trained and configured by the user independently.

Upon dialing the system phone number and receiving the greeting, the user is prompted to say the user login and the password. If the speaker verification system is activated, the user's identity is verified and entrance to the main menu granted or rejected.

As we see in figure 3, the user needs to provide a valid login and password. Which are first assessed using digits recognition and, if the speaker verification system is enabled, independent likelihood values are returned.

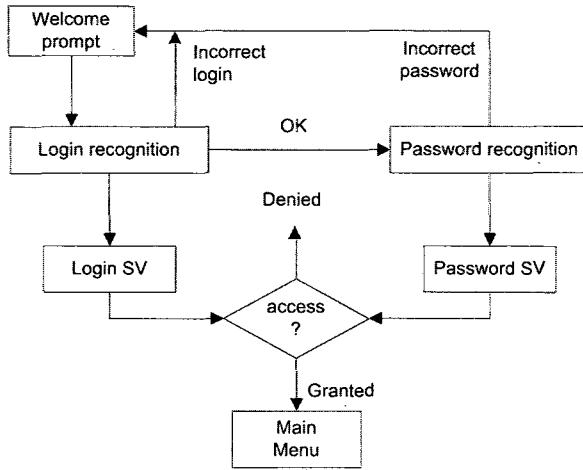


Figure 3: Speaker authentication system diagram.

Such likelihoods are compared with the speaker threshold values and the speaker is granted entrance only if both are positive.

The speaker verification system is based on Gaussian Mixture Models (GMM). Both login and password consist of 4 digits. Two different approaches can be used when training the GMM and decoding: a) In independent digits mode the models for digits (0-9) are independently trained. Decoding is done using the corresponding models only. b) In sentence mode both login and password are treated as a whole sentence, a model for each is trained.

The user can train and modify his personal verification models through a voice menu accessible from the main menu. Before using the speaker verification capability he will have to select the kind of verification to use and train the models. Once the system is working he can change some of the system variables by voice through the same menu.

### 3.3 Message adaptation to speech synthesizer

For multimodal e-mail messages received, only text and audio parts are selected to be reproduced. The system, will inform about the existence of other parts, as images or binary documents.

The appropriate speech synthesizer is activated detecting each paragraph language following an

stochastic method. A count of words in Spanish, Catalan and English is done, selecting the language that has more words. The efficiency of the method increases with the number of words.

In order to synthesize correctly the text, some structures as dates, electronic addresses, abbreviations and initials are rewritten. Any format of date is written in an extended version. Electronic addresses are spelled and for abbreviations and initials an equivalence is searched in a dictionary of exceptions.

### 3.4 E-mail summarization

Petra gives, upon request, a message summary provided by Carpanta [reference] module. The goal is to create summaries that are indicative of the number of topics discussed, in contrast with informative summaries, which try to synthesize most of the relevant information. Carpanta does not build new texts, but selects most representative sentences in the text. Its modular architecture permits to difference between language-independent modules, which are the core of the system, and language-dependent modules, which guarantees the portability to other languages different from Spanish.

## 4. EVALUATION

To evaluate the phone access system, the participation of 30 people was requested. Three calls were required for each participant. The first one, common for all people, offered a first contact with the application and its use. For second and third calls, different simulated scenarios were prepared, each one with a goal to be achieved during the call.

Results were obtained from the analysis of 72 registered calls. We took the following measures: 1) Understanding rate, the system understood the user command and acted properly. 2) Wrong user entry, tries to identify dialogue turns where the users had some difficulties. 3) Recognition failure, the system was not capable to understand the user, although his

command was right. It could be due to a vocabulary lack or due to a noisy environment.

<i>Measures (%)</i>	<i>w/o barge-in</i>	<i>w barge-in</i>
Understanding rate	76.6	77.1
Wrong user entry	2.9	1.7
Recognition failure	20.4	21.2

Table 1: - Evaluation results

To see the effects of barge-in when it's used, those turns were measured separately. Table 1 shows these results.

Barge-in did not show a clear degradation of the recognition rates. This is probably due to the users being more confident in their answers when they interrupt the system prompt. We arrived to such conclusion by listening to the recorded calls.

Finally, users were asked to fill on a website a survey measuring user satisfaction. Table 2 shows the results, where punctuation goes from 1 (worst) to 5 (best).

Table 2: - User satisfaction results

<i>Statements</i>	<i>Mark</i>
I understood what the system said	4,23
The system understood me	3,33
I knew what I could say to the system	3,50
The system helped me when I needed	4,23
The system behaviors as I could expect	3,61
The dialogue is natural	3,39
The system voice results friendly	2,72
The goal of call 2 was easy to achieve	3,72
The goal of call 3 was easy to achieve	3,06

## 5. CONCLUSION

We have presented the new unified messaging system developed using recent advances in linguistic engineering. E-mail, voice mail and fax messages can be accessed from a common mailbox. The main features of the system have been explained and phone access to the system has been discussed in more detail. The modular design used for implementing the dialogue system will permit to



reuse the new tools developed in different domain applications, getting a minor time of implementation.

The phone access behavior and usability has been evaluated, getting satisfactory results. Useful conclusions were extracted from the campaign which will permit the improvement of those stages of dialogue where the users had difficulties.

A computer with internet connection and a telephone with hands-free capabilities will be needed for make the presentation in the meeting.

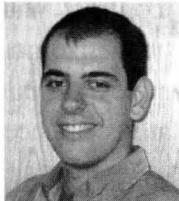
## REFERENCES

- Alonso, L., Casas, B., Castellón, I., Climent, S. and Padró, Ll. 2003. Carpanta eats words you don't need from e-mail. XIX Congreso anual de la Sociedad Española para el procesamiento del lenguaje natural.
- Padrell J. and Hernando J. 2002. Access to Meteorological Information by Telephone. ICSLP'02. Denver. 2173-2176.
- Reynolds, D.A. 2001. Speaker Verification: From Research to Reality. ICASSP Tutorial. Salt Lake City
- Verbio from Atlas-CTI. [www.atlas-cti.com](http://www.atlas-cti.com)



**Francisco Javier Hernando Pericás.**

Ingeniero de Telecomunicación en 1988 y Doctor Ingeniero de Telecomunicación en 1993 con Premio Extraordinario de Doctorado por la Universidad Politécnica de Cataluña. Profesor Titular de Universidad del Departamento de Teoría de la Señal y Comunicaciones de la Universidad Politécnica de Cataluña. Sus trabajos de investigación se centran en técnicas de procesado digital de la señal orientadas a la representación robusta de la voz para reconocimiento del habla y del locutor. Es autor de publicaciones en revistas y de comunicaciones en conferencias y es miembro de asociaciones relevantes en el área tanto en el ámbito nacional como internacional. Trabaja en proyectos subvencionados institucionalmente.



**Xavier Anguera Miró.** Ingeniero de

Telecomunicación en el 2001, máster Europeo en lenguaje y habla en el 2001 y estudiante de doctorado en Telecomunicaciones en la actualidad. Especializado en procesado digital de la señal, especialmente del habla. Ha trabajado en sistemas de síntesis de voz, de diálogo entre hombre y máquina y de verificación de locutor. Ha realizado su actividad profesional mayoritariamente en los EEUU, donde en la actualidad se encuentra en una estancia en el International Computer Science Institute (Berkeley, California).

## AUTORES



**David Hernando Davalillo.** Ingeniero de Telecomunicación en febrero de 2004. Su proyecto final de carrera lo realizó en el Grupo de procesado del habla del TSC participando en diferentes tareas del proyecto Petra aquí presentado, principalmente en el sistema de diálogo oral. En la actualidad, trabaja en una Biomeric Technologies, empresa de nuevas tecnologías dedicada a la identificación de personas a través de su voz y otras biometrías.

# ¿ES POSIBLE LEGISLAR SOBRE CUESTIONES ÉTICAS?



Victor Federico Dios Otin

fede@tsc.upc.edu

Profesor Titular del Dept. de Teoría de la Señal y Comunicaciones  
Universitat Politècnica de Catalunya (UPC)

## RESUMEN

La legitimidad del Estado para legislar sobre conductas consideradas no-éticas (sea desde el punto de vista de la razonabilidad práctica o, al menos, desde la óptica de la mayoría de los ciudadanos) cuando se circunscriben al ámbito privado, es un problema recurrente dentro de la teoría política. En cierta medida podríamos afirmar que está en el núcleo de todos los debates parlamentarios y sociales importantes que surgen periódicamente en el seno de las sociedades occidentales. En este artículo, sin ánimo de tratar todas las derivaciones posibles, que son muy numerosas, se expone parte del recorrido histórico de esta cuestión y algunas de sus implicaciones actuales.

## ¿ES POSIBLE LEGISLAR SOBRE CUESTIONES ÉTICAS?

Si nos referimos exclusivamente al *harmless wrongdoing* - que ha sido tratado extensivamente en la literatura política anglosajona - esto es, a aquellos actos que no conllevan un daño aparente para los demás ciudadanos, el liberalismo moderno más ortodoxo afirma que no: tal legislación estaría fuera de lugar. Uno de los actuales máximos representantes de la doctrina liberal, John Rawls, en la clásica obra '*A theory of justice*' [1] opina que la legislación ha de contener una mínima base común de principios éticos, un *overlapping consensus*, capaz de satisfacer a todos los grupos sociales, políticos o religiosos con, quizás, serias divergencias en sus respectivas filosofías de vida. El modo en que Rawls construye su paradigma de sociedad es el siguiente: imagina que en un estadio original, presocial, se reúnen todos los partícipes de la futura sociedad, pero detrás de lo que él denomina

*el velo de ignorancia*, esto es: sin saber a priori cuáles serán las posturas que cada uno defenderá en la sociedad ya formada. En esa situación, opina Rawls, *sería forzoso llegar a un acuerdo de mínimos, de modo que nadie pudiera después sentirse incómodo en su papel definitivo*. La posición de Rawls ha sido criticada desde diferentes presupuestos, pero tiene el mérito de haber sido el primero en tratar de fundar racionalmente un orden ético social desde el liberalismo moderno. A la hora de la práctica, sin embargo, se presentan escollos de difícil solución, y surgen discusiones sobre determinados puntos que para unos son claros - deberían formar parte de aquel *overlapping consensus* - y para otros no. Tal es el caso, por ejemplo, del aborto. A Rawls le parece evidente que nadie debería tener dificultades en admitir la legalidad del aborto dentro de los tres primeros meses de vida del feto, por lo que sobre ese particular debería haber consenso. Sabido es, sin embargo, que tal consenso no existe.

En las mismas filas liberales han surgido corrientes críticas con enfoque rawlsiano, y opinan que el legislador no podrástraerse a la influencia moralizante o 'amoralizante' de las leyes, por lo que ha de procurar en todos los casos crear un ambiente propicio para 'hacer mejores' a los ciudadanos. Tal es la postura, por ejemplo, de Joseph Raz, uno de los más firmes exponentes de esta corriente [2]. No sería prudente, en su opinión, conformarse con una moral de mínimos, sino que la sociedad en su conjunto y el poder político muy especialmente, deben aspirar positivamente a un enriquecimiento ético progresivo de los miembros de la sociedad.

La postura de Raz, y de otros pensadores actuales, enlaza con el pensamiento clásico, si bien



existen siempre matices diferenciados propios de cada escuela. Hagamos un pequeño repaso histórico de esta cuestión, que desde siempre ha estado en el núcleo de la teoría política. No pretendo que sea un examen riguroso, porque se necesitarían muchísimas páginas para hacerlo. Nos conformaremos con recordar las posiciones más representativas que se han sucedido a lo largo de la historia.

La postura de Aristóteles, expresada en su *Ética a Nicómaco* y en su *Política*, debe entenderse desde la peculiar concepción de la sociedad o, mejor, de la *polis*, propia de la Grecia clásica. Siendo el hombre un animal social por antonomasia, éste no puede esperar alcanzar su perfección como persona sino a través de la convivencia organizada propia de una comunidad. Para el pensador griego, por otra parte, la familia es una célula demasiado pequeña, y un país, o lo que sería equivalente a una confederación de ciudades, resulta ya demasiado grande para que los ciudadanos puedan aspirar a ejercer con sabiduría, prudencia y conocimiento de causa, un papel dirigente, quizá limitado, pero decisivo en su particular ámbito. Sin la posibilidad de jugar tal papel, el hombre no podrá desarrollar sus potencialidades. Desde esta perspectiva se adivina que Aristóteles ve la organización política como la cima de la ética, de aquel conocimiento y actitud moral que llevará al hombre a una vida plena o, en sus propias palabras, a una vida lograda. Como conclusión "es evidente que ha de preocuparse por la virtud la que de verdad se llama (a sí misma) ciudad... Pues, en otro caso, ... la ley se convierte en un pacto que es garante de los derechos mutuos, pero que es incapaz de hacer buenos y justos a los ciudadanos". Mas adelante insiste: "... la comunidad política tiene por objeto las buenas acciones y no únicamente la vida en común" [3].

En el pensamiento posterior de la tradición cristiana no se espera, sin embargo, que sea la comunidad política la que haya de contribuir decisivamente al perfeccionamiento de los ciudadanos. La persona humana tiene una naturaleza social, pero la adquisición de las virtudes es un empeño personal más que comunitario y, en cualquier caso, las iniciativas formativas no han de esperarse del poder político. Tomás de Aquino afirma en esa línea

que las leyes humanas no pretenden prohibir todas las acciones contrarias a la virtud "*sino tan sólo las más graves, aquellas que pueden ser evitadas por la mayoría de las personas y, sobre todo, las que van en perjuicio de los demás, sin cuya prohibición la propia sociedad humana no podría subsistir*" [4]. Es cierto, sin embargo, que Tomás de Aquino asigna un papel importante a las leyes en cuanto a la formación ética de los ciudadanos. Las leyes no pueden evitar todos los comportamientos no virtuosos, pero tampoco pueden alestarlos.

En la línea del empirismo y del racionalismo, presentes en el pensamiento occidental a partir del siglo XVI, la postura de Thomas Hobbes es indudablemente más dura. Hobbes no se plantea ni siquiera la cuestión de cómo legislar sobre las cuestiones éticas de mayor dificultad, puesto que su idea de la naturaleza humana es bastante más negativa y no le permite remontarse a esas sutilezas. Para él, las leyes tienen suficiente trabajo con impedir, en cualquier ámbito de la convivencia social, la prevalencia del más fuerte. No en vano Hobbes no ve tampoco con demasiado optimismo la asamblea democrática, demasiado fácil de manipular en su provecho por un grupo de hombres decididos y sin escrúpulos.

Los padres del liberalismo clásico, como John Stuart Mill o Benjamin Constant, tienen en su punto de mira la defensa de los derechos del individuo frente a las pretensiones omnímodas de un poder estatal de signo absolutista, que tiende a absorber todos los focos de decisión. Sin duda, como muestran en sus numerosos escritos, están preocupados por el bien común y el imperio de la virtud, pero es mayor su temor a las injerencias ilegítimas del Estado. Afirma Mill: "*La única finalidad por la cual el poder puede, con pleno derecho, ser ejercido contra un miembro de una comunidad civilizada contra su voluntad, es evitar que perjudique a los demás. Su propio bien, físico o moral, no es justificación suficiente. Nadie puede ser obligado justificadamente a realizar o no realizar determinados actos, porque eso fuera mejor para él, porque le haría más feliz, o porque, en opinión de los demás, hacerlo sería más acertado o más justo*" [5].

A propósito del Informe Wolfenden, remitido en 1957 al Parlamento británico, se desató un largo debate acerca del legítimo alcance de la ley sobre las acciones consideradas inmorales cuando éstas se circunscriben a un ámbito no público. Patrick Devlin opinó que no habría de cerrarse por principio la posibilidad de actuación legal en esos casos, siquiera sea para evitar un peligro de pérdida de la cohesión social, en cuyo fundamento juegan un papel muy importante los sentimientos éticos de la mayoría de la población [6]. Alrededor de esta opinión se sucedieron durante años críticas y matizaciones, que sería largo detallar aquí.

En la actualidad es claro que el debate sobre si puede o no legislarse sobre cuestiones éticas, o hasta qué punto puede hacerse legítimamente, continua abierto.

Por mi parte, he de confesarlo, estoy más cerca de la posición de John Neuhaus, de tendencia conservadora. Neuhaus sostiene que es una ingenuidad pretender mantener la acción política apartada de implicaciones éticas, ya que, al final, es sobre cuestiones éticas de fondo sobre lo que se está decidiendo cuando se promulgan leyes, en uno o en otro sentido. El problema, según Neuhaus, no es si se debe legislar o no sobre cuestiones éticas, sino cómo debe legislarse.

Es claro, y en esto coinciden prácticamente todas las filosofías políticas, que deben separarse cuidadosamente las cuestiones que tienen, otendrán a la larga, una inevitable repercusión pública de aquellas que quedan en la esfera privada de los ciudadanos. El legislador debería preocuparse por las posibles actitudes amorales o inmorales de los ciudadanos cuando éstas conllevan siempre, o con alta probabilidad, consecuencias negativas para otros miembros de la sociedad. Por poner algunos ejemplos: el que un hombre tenga la costumbre de emborracharse los fines de semana en su casa, o en la de algún amigo o en el bar, puede tener consecuencias o no. Es sabido que, a la larga, de esa costumbre se derivarán efectos negativos en el que la practica, como que se convierta en un alcohólico o que desarrolle alguna grave enfermedad, hepática o psíquica. Sin embargo no es prudente, ni posible en la práctica, que el Estado

se inmiscuya en ese tipo de conductas privadas. La situación, sin embargo, adquiere otra dimensión cuando se comprueba que el consumo inmoderado de alcohol es causa de un elevado tanto por ciento de los accidentes de tráfico. Ciertamente el Estado sigue sin poder, razonablemente, inmiscuirse en la conducta privada, pero prohibirá la conducción de vehículos en determinadas condiciones, por el riesgo que supone para los demás y para uno mismo. En esta misma línea, el Estado no entrará en el asunto de las borracheras 'tranquilas' mientras quienes se emborrachen sean personas adultas, pero si observa que esas desafortunadas costumbres se introducen de forma generalizada entre la gente joven, entonces se ve obligado a ejercer un papel 'paternalista', a fin de evitar que numerosos jóvenes ciudadanos adquieran el vicio del excesivo consumo de alcohol desde edades tempranas. Es indudable que el Estado actuará con buen criterio cuando ponga trabas legales para dificultar esas costumbres, de acuerdo, sin duda, a criterios éticos bien fundados.

La existencia de principios éticos básicos, subyacentes a toda acción política, es una evidencia. Ocurre que nos hemos acostumbrado de tal manera a respetar determinados principios que nos parecen simplemente de sentido común. Por otra parte, la función formativa de las leyes es también clara. Veamos algunos ejemplos, aunque podrían aducirse muchos más. La política fiscal tiene como primera misión, no alimentar las arcas del Estado, sino la de contribuir a limar diferencias de renta entre ciudadanos y entre regiones geográficas, porque se entiende que unas excesivas diferencias de disponibilidad económica no son acordes con la igual dignidad de todas las personas. Cuando los hombres nacen y se forman en una sociedad en la que se aplican de forma natural esos principios de solidaridad, acaban formando parte de la propia manera de pensar. En la misma línea del ejemplo anterior: a la propiedad privada no se le concede en la sociedad occidental, un valor inatacable de pleno derecho, sino que se considera que toda propiedad privada tiene una función social. Actualmente se estudia en nuestro país la posibilidad, de grabar con fuertes impuestos a quienes tienen pisos desocupados, para que los pongan a disposición de los demás ciudadanos, sea en forma de venta o de alquiler. Esa política en particular se aplica ya en



otros países desarrollados. Los clásicos derechos de paso o el derecho de usucapión - el hacerse con la propiedad pacífica de un bien no reclamado y que se lleva años utilizando - que proceden del derecho romano, abogan también por un empleo racional y justo de los bienes.

Hoy en día el asunto que tenemos planteado sobre la mesa es el futuro de la familia, a raíz de los profundos cambios que desde algunos sectores pretenden introducirse. Parece evidente que no es un tema menor. En este particular sigue siendo razonable mantener que no es el Estado quien ha de dictar a cada cuál cómo ha de enfocar su vida o sus relaciones personales, pero también es cierto que la defensa del bien común le obligara a matizar muy bien las disposiciones legales, ya que no le es posible encontrar una posición de hipotética neutralidad. El aspecto que el legislador debería tener muy en cuenta es el mensaje que desde el corpus legislativo se está mandando a los ciudadanos. Si las leyes protegen a la familia, en su acepción clásica, los ciudadanos seguirán comprendiendo mayoritariamente su importancia y la necesidad que existe de mantener sus estatus diferencial, frente a otros tipos de uniones estables que no cumplen la misma función social. Por el contrario, si las leyes tratan por igual todas las uniones estables, los ciudadanos, dentro de quizás no muchos años, acabarán por considerar que la familia hoy día considerada normal, no tiene mayor relevancia que aquéllas otras uniones.

En los países occidentales tenemos ya conocimiento de algunas de las consecuencias que se derivan de la disolución del núcleo familiar, debido al hasta ahora creciente número de divorcios. La Unión Europea ha hecho públicos, por ejemplo, los datos relativos al índice de pobreza, mostrando que tienen más posibilidad de caer bajo el umbral de pobreza los hogares monoparentales [7]. Más preocupantes son los datos referentes al crecimiento de la violencia juvenil: es un dato estadístico que el índice de violencia juvenil entre los chicos aumenta cuando les falta un adecuado referente masculino [8].

En este tema, como en tantos otros, no es posible sustraerse a consideraciones éticas. Sin duda el legislador que pretende equiparar legalmente

cualquier tipo de unión estable está ejerciendo una opción libre de elección entre dos concepciones éticas divergentes.

Si investigamos más hacia el fondo de las diferentes actitudes se advierte la ya clásica pugna entre la postura ultraliberal y las posiciones de la tradición ética de raíz judeo-cristiana. Para la primera el bien supremo es la libertad de elección, para la segunda lo prioritario son los bienes genuinos del hombre, razonablemente considerados como tales. Para los liberales el hecho de poder elegir es un bien en sí mismo, para los segundos se trata no tanto de elegir sino de acertar en esa elección. Pero estas consideraciones nos llevarían ya muy lejos.

## REFERENCIAS

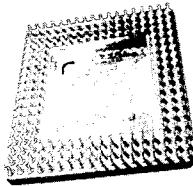
- [1] John Rawls, *A Theory of Justice*, Harvard University Press, 1971.
- [2] Joseph Raz, *A Morality of Freedom*, Clarendon Press, 1986.
- [3] Aristóteles, *Política*, libro III, cap. 9.
- [4] Tomás de Aquino, *Summa Theologica I-II*, q. 96, a. 2
- [5] John Stuart Mill, *On Liberty* (1859), Penguin Books, 1958.
- [6] Patrick Devlin, *The Enforcement of Morals*, Oxford University Press, 1965.
- [7] El 'riesgo persistente de pobreza' es de un 15% para la Europa de los quince. A la cola, por países, están Irlanda, Grecia y Portugal, con un riesgo que se sitúa entre el 20% y el 21%. Por tipos de familia, las mujeres u hombres solos, con uno o más hijos a su cargo, alcanzan un índice del 35%, según Eurostat.
- [8] María Victoria Barrio y Kathleen Heide, en el debate *Violencia y juventud*, VIII Reunión Internacional sobre Biología y Sociología de la violencia. Valencia, 2004.

## AUTOR

**Federico Díos Otín** es profesor titular del departamento de Teoría de la Señal y Comunicaciones. Imparte las asignaturas de Campos Electromagnéticos y de Ingeniería de Sistemas Ópticos. Sus campos de interés científicos son la influencia de la turbulencia atmosférica en la propagación de ondas luminosas y los sistemas de medida mediante láser. En su faceta como pensador ha publicado recientemente el libro "Mi querido agnóstico", Desclée de Brouwer, 2004.



# AN INTRODUCTION TO MICRO ELECTRO MECHANICAL SYSTEMS (MEMS)



Juan Pablo Sáenz

jpsaenz@ieee.org

## ABSTRACT

The remarkable progress of MEMS (Micro Electro Mechanical Systems) research has reached the stage where practical applications are intensely pursued. Based on high-performance systems commercial products have made remarkable progress in many fields. This paper discusses the underlying principles for MEMS, going over the materials and processes used for their fabrication. It also presents two state of the art examples, under development in the industry. A brief discussion of terminology is also included.

## 1. INTRODUCTION

Micro Electro Mechanical Systems (MEMS), also known as Micromechanics, Microsystems technology (MTS) or nanotechnology, is an interdisciplinary field of study committed to the physical integration of micromechanical systems with microelectronics, resulting in miniature embedded system that involve micromachined components and structures.

MEMS have been studied since the 1960's. The first devices appeared in the 70's and the public interest gave MEMS a boost in the 80's. During the last decade, the research has grown through funding from governmental agencies, involving researchers from universities and industries all over the world (MEMS Exchange, 2003). Most of this research and development has been directed toward the replacement of conventional technologies to increase functionality, reduce cost, and improve reliability.

The strong demand of these systems in different fields of work makes the study and general understanding of MEMS required for any engineer. This paper introduces the general concepts of this technology, by summarizing the materials and processes involved in fabrication, and provides a closer look at two applications where MEMS have been successfully implemented.

## 2. MEMS

Since their first appearance, Micro Electro Mechanical Systems have been studied for research groups around the globe. Their numerous applications have made their market grow exponentially.

### 2.1. Microengineering

The design and development of MEMS requires knowledge not only of one single field, but of different disciplines of science. This knowledge includes quantum mechanics, molecular theory of matter, mechanics thermo fluids and chemistry, just to mention a few. Figure 1 illustrates the relationship between some of these fields. The technology and techniques involved in MEMS are grouped in what is called Microengineering referring to the technologies and practice of making three dimensional structures and devices with dimensions in the order of micrometers.

### 2.2. State-of-the-art

The application of MEMS can be grouped in three major categories: passive structures, sensors and actuators. Currently, they are being used for many industries such automobile,



aerospace, telecommunications and health care. The dominant applications are pressure and inertial sensors, and inject print heads. Work is also being done on high resolution displays and high-density storage devices. Figure 1 shows some of these applications.

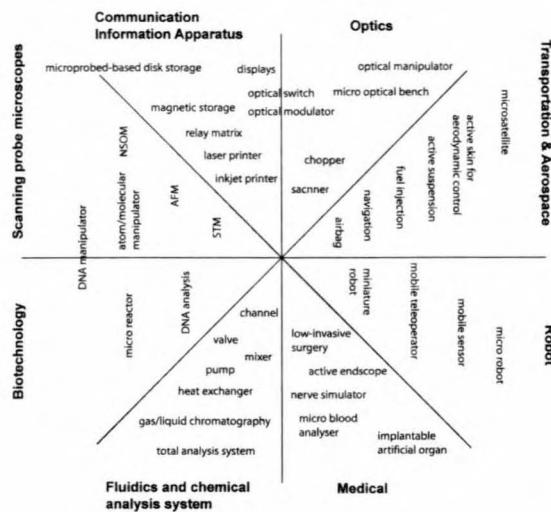


Figure 1. Applications of MEMS

Recent efforts have been directed towards biomedical and genetic engineering, developing new medical diagnostic tools and drug delivery systems. The market for MEMS has grown exponentially during recent years reaching 82 billion dollars in 2000.

### 3.MEMS FABRICATION

The origin of MEMS fabrication takes place in the integrated circuit (IC) industry. Most of the processes used to produce IC are used to develop MEMS. Some recent devices require certain specific processes or a combination of them for their elaboration. All these processes are referred to as Micromachining. Micromachining adds layers of material over a

silicon wafer and etches precise patterns in these layers or in the underlying substrate.

### 3.1.Materials

The variety of materials used in MEMS is very broad. These materials are used as substrate material, where Silicon is by far the most commonly used. Exploration and development of new techniques drive the research in the use of new compounds, and thin films of silicon nitrides, oxides, glass, organic polymers and some shape-memory alloys.

The control of the mechanical properties and its extensive use, make silicon and polysilicon the materials of preference in industry (Petersen 1982). Figure 2 shows some silicon ingots and the crystal structure of this element.

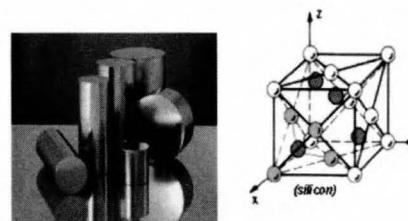


Figure 2. Silicon ingots and crystalline structure of silicon.

The study of crystallography and solid state physics are strongly encouraged for people interested in working with MEMS. The properties of a selected group of materials used on MEMS can be found in table 1.

### 3.2.Processes

Micromachining processes are divided into basic tools (well-established methods available

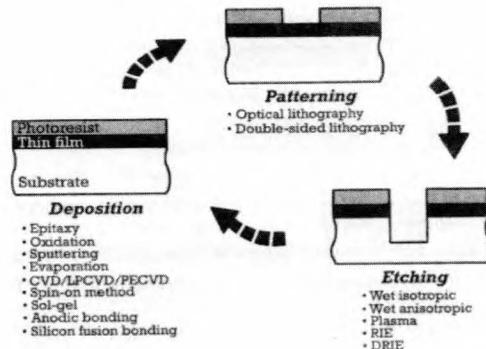
Table 1. Properties of selected materials

Property	Si	SiO <sub>2</sub>	Si <sub>3</sub> N <sub>4</sub>	Quartz	SiC	AlN
Density (g/cm <sup>3</sup> )	2.4	2.3	3.1	2.65	3.2	3.26
Specific Heat (J/g*K)	0.7	1.0	0.7	0.787	0.8	0.71
Hole mobility (cm <sup>2</sup> /V*s)	400	-	-	-	40	-
Electron mobility	1500	-	-	-	1000	
Relative permittivity ( $\epsilon_0$ )	11.8	3.8	4	3.75	9.7	8.5
Dielectric strength (V/cm x 10 <sup>6</sup> )	3	3-5	5-10	25-40	4	13
Melting temperature C	1415	1700	1800	1610	2830	2470

in most foundries) and advanced tools (unique in their nature and limited). Figure 3 show the basic process flow in micromachining. As previously mentioned, micromachining layers are deposited; photoresist is lithographically patterned, and then used as a mask to etch the underlying material.

The patterning is made by means of photolithography with three sequential steps: application of photoresist (photosensitive emulsion), its optical exposure to print an image of a mask (made using CAD tools) onto the resist, and the immersion in an aqueous developer to dissolve the exposed resist and render the image visible. More advanced techniques entail double-side lithography that requires a high degree of accuracy.

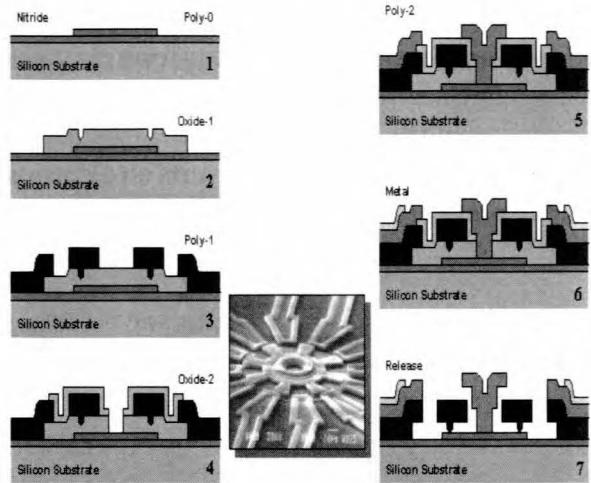
*Figure 3. Basic process flow in micromachining*



One of the most important processes in micromachining is etching, which involves the removing of materials in desired areas by chemical and physical methods. Chemical etching uses diluted chemicals like HF or Si<sub>3</sub>N<sub>4</sub> to dissolve substrates. Two kinds of etching are available, isotropic (uniform in all directions) and anisotropic (just in preferred directions). Physical etching or plasma etching uses a stream of positive-charged-carrying ions of a substance with a large number of electrons. These ions bombard the surface of the target, gnawing out the substrate material from its surface (Hsu, 2002).

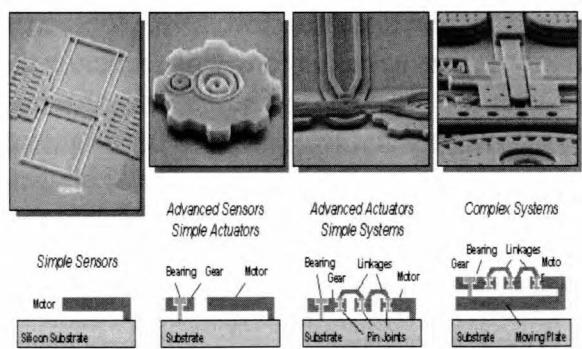
Figure 4 shows the use of these basic processes for the fabrication of a micromotor. Layers are deposited; photoresist is

lithographically patterned and then used as a mask to etch the underlying material.



*Figure 4. Micromotor design and processes steps.*

More advanced techniques are used for the manufacturing of different devices. This group includes Lithography Electroforming and Molding (LIGA), Silicon Fusion Bonding and Reactive Ion Etching (SFB-DRIE) and Single Crystal Reactive Etching And Metallization (SCREAM). Given that the objective of this paper is to give an introduction, a more detailed explanation of every single one of all processes can be found in the references for this paper. It is important to note that the potential complexity of the systems increases with the number of unique process features and independent structural layers, as shown in figure 4.



*Figure 5. Device complexity by structural layers.  
From left to right 2, 3 , 4 and 5 levels of complexity.*

#### 4.MEMS STRUCTURES AND SYSTEMS

To have a better understanding of MEMS, it is important to go over the commonly used principles of applications: sensing and actuation.



#### 4.1.Sensing and actuation

Sensors and actuators are special types of transducers. A sensor is a device that converts one physical or chemical quantity to an electrical one; sometimes an intermediate step takes place, for processing by the microsystem. Similarly, an actuator is a device which converts an electrical quantity into a physical or chemical one. (Banks, 1999)

Some of the sensing techniques rely on the properties of the materials used. One commonly used principle is the dependency of most materials' properties on temperature; piezoresistivity and piezoelectricity effects. Some other applications have their origin on capacitive sensing, relying on external parameters, changing the spacing or the relative dielectric constant between two plates of a capacitor (acceleration and relative humidity sensors). Moreover, another approach is its use of electromagnetic signals to detect and measure physical parameters (e.g. Magnetoresistive sensor in the read heads of high-density computer storage systems). Table 2 shows the characteristics of these methods.

Actuation methods can be grouped in five primary categories: electrostatic actuation, the attractive force between two plate elements carrying opposite charges (opposite in sign) in such way that an applied voltage results in an attractive electrostatic force; piezoelectric actuation; thermal actuation, which relies on the difference of the coefficients of thermal expansion between two joined layers, causing the structure to bend with temperature changes; magnetic actuation, based on the principle that

an electrical current in a conductive element that is located within a magnetic field reacts to an electromagnetic force (Lorentz force) perpendicular to the current and magnetic field; and lastly actuation by Shape-memory alloys. Table 3 compares these actuation techniques on basis of the maximum energy output density.

Actuation	Max energy density	Physical & material parameters
Electrostatic	$\frac{1}{2}\epsilon_0 E^2$	E = Electric field $\epsilon_0$ = dielectric permittivity
Thermal	$\frac{1}{2}Y(\alpha\Delta T)^2$	$\alpha$ = coefficient of expansion $\Delta T$ = temperature rise Y = Young's modulus
Magnetic	$\frac{1}{2}B^2/\mu_0$	B = magnetic field $\mu_0$ = magnetic permeability
Piezoelectric	$\frac{1}{2}Y(d_{33}E)^2$	E = Electric field Y = Young's modulus $d_{33}$ = piezoelectric constant
Shape-memory alloys	-	Critical temperature

Table 3. Actuation techniques based on maximum energy density.

#### 5.APPLICATIONS

To illustrate the applicability of this technology this paper presents two of the most prominent applications, Radiation Sensors and Digital Micromirror Devices.

**Radiator sensor** for infrared imaging (Sensing). The basic approach is to measure the change in temperature of a suspended sense-resistor. This temperature change is due to incident infrared radiation. The intensity level is then proportional to the change in resistance of the device. Figure 6 shows a single element in the infrared imaging array from Honeywell. This model achieves high sensitivity to radiation by providing extreme thermal isolation. (Cole, 1998).

Piezoresistive	Capacitive	Electromagnetic
Simple fabrication	Simple mechanical structure	Structural complexity varies
Low cost	Low Cost	Complex Packaging
Voltage or current drive	Voltage drive	Current drive
No need for circuits	Requires circuitry	Simple control circuits
High temperature dependence	Low temperature dependence	Low temperature dependence
Small sensitivity	Large dynamic range	Sensitivity a magnetic fields
Medium power consumption	Low power consumption	Medium power consumption

Table 2. Characteristics of Piezoresistive, Capacitive and Electromagnetic sensing methods.

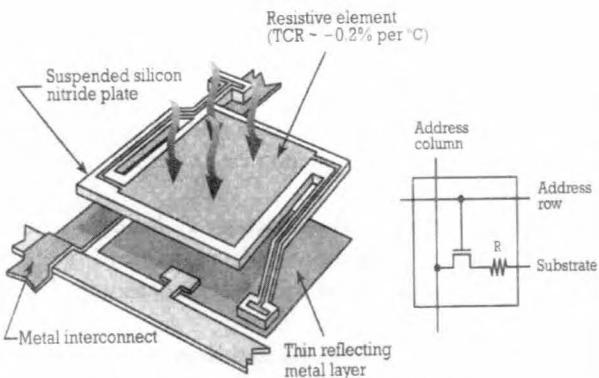


Figure 6. Single radiation sensor for infrared imager. (HoneyWell)

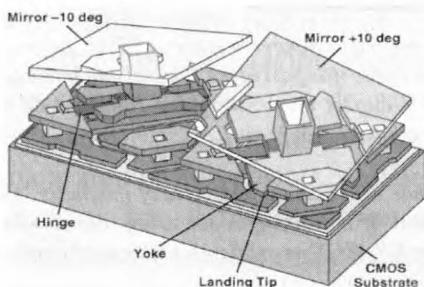
The square of silicon nitride is used as sense-resistor,  $50\mu\text{m}$  by  $50\mu\text{m}$  and  $0.5\mu\text{m}$  thick. Its fabrication process relies on a surface micromachining approach, involving an organic layer, polyimide, rather than polysilicon. The fabrication is made in arrays of 240 by 336 pixels; the electronics offers multiplexing, scan and read-out operations, in exchange for the difficulties of having a single lead to each pixel.

#### Digital Micromirror Device (Actuation).

This device consist in a two dimensional array of optical switching elements (pixels) on a silicon substrate. Each pixel consists of one mirror supported from a central post, mounted in a lower platform, suspended by two other compliant torsional hinges anchored directly to the platform.

Two electrodes positioned underneath the platform provide electrostatic actuation. This device DMD is a trademark of Texas Instruments (TI), U.S. patent #4615595. The angle of tilt is geometrically limited to  $10^\circ$ . Its control (state ON,  $+10^\circ$ , or OFF,  $-10^\circ$ , and duration) is made by a CMOS static-random-access-memory (SRAM) cell underneath the device. Every pixel measures  $17\mu\text{m} \times 17\mu\text{m}$ .

Figure 7. A pair of DMD pixels of Texas instrument



The fabrication used by TI uses surface micromachining on wafers which incorporate de CMOS electronics. The release of the mechanical structures is done by etching. A typical DMD includes 1.3 Million micromirrors and is used for digital projection TV and cinema. (Texas Instruments, 2003)

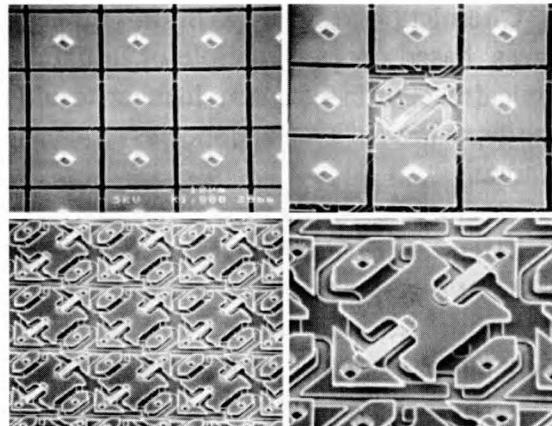


Figure 8. Details of the complete array.

## 6.FUTURE PERSPECTIVE AND SUMMARY

After a brief overview of MEMS' technology, it is obvious that this extremely diverse technology could significantly affect every category of commercial products. The nature of this technology and its diversity of useful applications make it potentially more pervasive than even integrated circuit microchips.

Also, MEMS blurs the distinction between complex mechanical systems and integrated circuit electronics. Costly and unreliable parts of a macroscale electronics system, as sensors and actuators, increase reliability and decrease their cost using batch fabrication techniques in micro electro mechanical systems. The performance of MEMS is expected to be superior to macroscale systems, and the price is predicted to be much lower.

MEMS are already used for tasks ranging from in-dwelling blood pressure monitoring to active suspension systems for automobiles. Their great promise lies in their potential to enable a new range of applications, a major growth industry. Some of these applications under development are: miniature biomechanical

chambers for DNA synthesis and addressing, optical communication switches, thermomagnetical data storage systems, as well as new and improved micropumps and mechanical resonators.

## GLOSSARY

**Chemical Vapor Deposition (CVD):** A process, based on the principle of initiating a chemical reaction in a vacuum chamber, resulting in the deposition of a reacted species on a heated substrate.

**Micro Electro Mechanical Systems (MEMS):** A generic descriptive term, common in the United States, for a broad technology having the objective of miniaturizing complex systems by integrating a diverse set of functions in a small package.

**Micromachining:** A term describing the set of design and fabrication tools for the machining of microstructures and very small mechanical features on a substrate.

**Piezoelectricity:** The property exhibited by a class of materials to develop voltage in response to applied mechanical stress or pressure. Conversely, an external applied electrical voltage strains and deforms this material.

**Piezoresistivity:** The property of certain type of materials, including impurity doped silicon, to change their electrical resistivity in response to mechanical stress.

**Shape memory alloy:** Group of materials that have the ability of returning to a predetermined shape when heated above critical "transition temperature." These materials "remember" their original shape after being strained and deformed.

**Transducer:** A device that converts from one physical quantity to another. One example is the deformation of a piezoelectric crystal under an applied electric field.

## REFERENCES

- [1] Banks, David (1999). "Introduction to Microengineering." MEMS Micromachines and MST. <<http://www.dbanks.demon.co.uk/ueng/>> (Nov 3, 2003)
- [2] Bustillo, James M., Howe, Roger T., Muler, Richard S. (1998). "Surface Micromachining for Microelectromechanical Systems". Proceedings IEEE 86, 8, 1552-1574.
- [3] Cole, B. E., R. E. Higashi, and R. A. Wood. (1998). "Monolithic two-dimentional Arrays of Micromachined Microstructures for Infrared Applications". Proceedings IEEE 86, 8, 1679-1686.
- [4] Colorado University (2003). Colorado University MEMS web page. <<http://mems.colorado.edu/>> (Nov 3, 2003)
- [5] Hsu, Tai-Rai (2002). MEMS & Micosystems. Design and Manufacture. Mac Graw Hill Higher Education. Boca Raton, FL.
- [6] Maluf, Nadim (2000). An Introduction to Micromechanical System Engineering. MEMS Series, Artech House. Norwood, MA.
- [7] MEMS exchange (2003). About MEMS technology.<http://www.mems-exchange.org/MEMS/> (Nov 3, 2003)
- [8] Nayoga, Japan (1997). 10th International Workshop on MEMS. <http://www.cs.arizona.edu/japan/www/atip/public/atip.reports.97/atip97.039r.html> (Nov 3, 2003)
- [9] Petersen, Kurt E. (1982). "Silicon as a Mechanical material". Proceedings IEEE 70, 5, 420-457.
- [10] Texas Instrument (2001). White papers on DLM and MEMS. <http://www.dlp.com> (Nov 3, 2003)

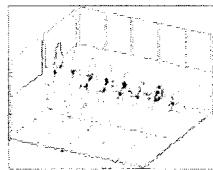
## AUTORES

*Student member of the IEEE since 1998, Juan Pablo Saenz joined the Cryogenic Electronics Group at San Francisco State University in fall 2002, where he pursued a MS in Engineering. As Graduate Research Assistant he worked on microfabrication processes set up, superconductor materials characterization and device Layout, for MEMS integration with Cryogenic devices.*



*His MS thesis "Study of processes and materials for MEMS integration with cryogenic devices," recently won the 18th Annual research competition in Engineering and Computer Science at SFSU. Before joined the CEG Juan Pablo was part of the Control and intelligent System Research group at the Universidad Nacional de Colombia, where he got his BS in Electrical Engineering.*

*He is author with Camilo A. Cortes of the thesis "Design and construction of a walking robot using nitinol wire and the calculator HP 48GX as artificial intelligence platform."*



# SOFTWARE BASADO EN EL MÉTODO DE ELEMENTOS FINITOS PARA LA ENSEÑANZA DE ELECTROMAGNETISMO

Diego Sánchez Repila, Luis E. García

Departamento de Teoría de la Señal y Comunicaciones,  
Universidad de Alcalá Escuela Politécnica (Madrid).  
Email: luise.garcia@uah.es

## ABSTRACT

Se presenta una herramienta software para la enseñanza de electromagnetismo. Ésta, permite analizar una gran variedad de fenómenos electromagnéticos en las bandas de microondas y ondas milimétricas. Dos aspectos muy cuidados en el desarrollo han sido la portabilidad y la difusión, estando disponible para sistemas operativos **Linux** y **Windows** y en diferentes idiomas. El interfaz gráfico, está basado en un **pre-post procesador de propósito general** para análisis computacional. Por otra parte, el análisis electromagnético está asentado en programas autónomos desarrollados con propósitos de investigación, que hacen uso del Método de los Elementos Finitos. La flexibilidad en el diseño de la arquitectura del software desarrollado, permite una fácil introducción de nuevos núcleos de cálculo. De este modo, aumenta periódicamente el número de problemas electromagnéticos que pueden ser analizados.

## 1. INTRODUCCIÓN

El objetivo de este trabajo, es proveer a los alumnos de una herramienta software, que sea capaz de analizar diferentes problemas electromagnéticos [1]. De los posibles resultados que se originan tras el análisis, se incluyen la visualización del campo eléctrico y/o magnético, en el interior y alrededor de las estructuras analizadas, y el cálculo de distintos parámetros de interés, (tales como la constante de propagación, modos dentro de una guíaonda, frecuencias de resonancia existentes en una cavidad de microondas...). De esta forma, al alumno le es posible entender, de un mejor modo, los diferentes fenómenos electromagnéticos. Como ejemplos de algunos de estos fenómenos, cuya resolución puede relacionarse

con este trabajo, citaremos: resonancia en cavidades 3D, análisis de líneas de transmisión multiconductoras, radiación y dispersión de ondas electromagnéticas...

La obtención de este objetivo, hace inevitable la elección de un método numérico, para el análisis de las estructuras restadas a estudio. El Método de los Elementos Finitos (MEF), fue la mejor opción de entre las múltiples posibilidades barajadas. Sus cualidades más sobresalientes son la precisión y capacidad para manipular geometrías y configuraciones arbitrarias con el mismo código.

Cada uno de estos fenómenos electromagnéticos, está ligado a un código de resolución. Estos códigos cubren una amplia variedad de problemas electromagnéticos, incluyendo: líneas de transmisión y, en general, estructuras de guíaondas, cavidades 3D, discontinuidades en guíaondas y problemas de scattering y radiación en 3D. De entre sus características fundamentales, podemos reseñar que hacen uso de elementos de orden superior, del tipo curl-conforming o que utilizan novedosos procedimientos MEF de truncación de malla [2]. Para su creación se ha utilizado FORTRAN. Estos programas se basan en funciones procedentes de la librería MODULEF [3], de fuentes disponibles, acceso gratuito y de características muy versátiles para crear cualquier programa de MEF, dado su carácter modular. Algunos ejemplos representativos de la versatilidad de MODULEF son: la posibilidad de definición de elementos finitos, de virtualmente cualquier tipo, o el método de asignación de condiciones de contorno.

Estos códigos presentan dos virtudes fundamentales: potencia y flexibilidad. Sin embargo, adolecen de sencillez y atractivo, sobre todo para los estudiantes. Más concretamente, los inconvenientes aparecen en la entrada y salida de



los datos, y también en la presentación de los resultados. Con el objeto de afrontar estas dificultades, se recurrió a GID, un prepostprocesador, que proporciona, gracias a su interfaz potente, intuitiva y bien documentada [4], una rápida y fácil introducción de los módulos de cálculo, dependiendo del problema tipo a implementar. Para la etapa de preprocesso, cuenta con un generador de mallados de tipo estructurado y desestructurado. A destacar en la parte de postproceso, está su muy buena y variada visualización de resultados. Otros criterios para su elección, han sido su gran portabilidad (Linux, Windows y Mac), y la existencia de una licencia académica (gratuita, de 700 elementos en 2D y 3000 en 3D), que cubre, adecuadamente las necesidades de los alumnos.

Por tanto, lo que se busca es utilizar la potencialidad de cálculo de los códigos, y visualizar los resultados en GID.

En este momento, el software que se presenta en esta comunicación, permite el análisis en onda completa de estructuras de guiado arbitrarias, (incluyendo la caracterización de líneas multiconductoras, mediante sus parámetros primarios o secundarios, dependientes de la frecuencia, y el empleo de algoritmos de mallado auto-adaptativo) y los modos y frecuencias de resonancia de cavidades 3D de microondas. En ellos, se hace uso de tetraedros curl-conformes de primer y segundo grado, para los análisis de cavidades y de triángulos no estándar del tipo híbrido curl-conforme/Lagrange, tanto rectos como curvos, en el análisis de estructuras de guiado [2]. En breve se incluirán los códigos desarrollados, correspondientes a elementos curl-conforme de tercer grado y al análisis de problemas abiertos (dispersión y radiación).

## 2. FORMULACIÓN EMPLEADA

La formulación 3D, utilizada en el cálculo de los modos resonantes de las cavidades de microondas, es la correspondiente a la ecuación de onda del doble rotacional:

$$\nabla \times (\llbracket f \rrbracket^{-1} \nabla \times \overset{\mu}{V}) = \omega^2 \cdot [g] \cdot \overset{\mu}{V} \quad (1)$$

siendo  $\overset{\mu}{V}$  el campo y  $\llbracket f \rrbracket$  y  $[g]$  los tensores que vienen dados en la tabla I.

A partir de la ecuación 1, con unas condiciones de contorno de pared eléctrica y/o magnética perfectas, y la normalización de campo de la forma siguiente:

$$\overset{\mu}{V}_n = \sqrt{g_o} \cdot \overset{\mu}{V} \quad (2)$$

teniendo en cuenta que:

$$\llbracket f \rrbracket = f_o \cdot \llbracket f_r \rrbracket \quad [g] = g_o \cdot [g_r] \quad (3)$$

obtenemos el siguiente funcional:

$$F = \int_{\Omega} \left\{ \nabla \times \overset{\mu}{V}_n \cdot \llbracket f_r \rrbracket^{-1} (\nabla \times \overset{\mu}{V}_n) - k_o^2 (\overset{\mu}{V}_n \cdot [g_r] \overset{\mu}{V}_n) \right\} d\Omega \quad (4)$$

Por otro lado, refiriéndonos a la formulación 2D usada en el análisis en onda completa de estructuras de guiado arbitrarias, realizaremos la descomposición del campo en componentes transversal y longitudinal y con simetría de traslación respecto al eje z:

$$\overset{\mu}{V}_n(x, y, z) = (\overset{\mu}{V}_{tn}(x, y) + j\overset{\mu}{V}_{zn}(x, y)) \hat{a}_z e^{-j\beta z} \quad (5)$$

el funcional obtenido es:

$$F_n = A_n - k_o^2 B_n \quad (6)$$

TABLA I  
CORRESPONDENCIAS PARA LA FORMULACIÓN DEL DOBLE ROTACIONAL

	$\overset{\mu}{V}$	$\llbracket f \rrbracket$	$[g]$
Form. H	$\overset{\mu}{H}$	$[\varepsilon]$	$[\mu]$
Form. E	$\overset{\mu}{E}$	$[\mu]$	$[\varepsilon]$

donde  $k_o^2$ ,  $A_n$  y  $B_n$  son:

$$k_o^2 = \frac{\omega^2}{c_0^2}$$

$$A_n = \int_{\Omega} \left\{ (\nabla_t \times V_{zn} \hat{a}_z) \cdot [f_{tr}]^1 \cdot (\nabla_t \times V_{zn} \hat{a}_z) + \right. \\ + (\hat{a}_z \times V_m^P) \cdot [f_{tr}]^1 \cdot (\hat{a}_z \times V_m^P) - \\ - 2 \cdot (\nabla_t \times V_{zn} \hat{a}_z) \cdot [f_{tr}]^1 \cdot (\hat{a}_z \times V_m^P) + \\ \left. + \frac{1}{\beta^2} \frac{1}{f_{zz}} \cdot (\nabla_t \times V_m^P) \cdot (\nabla_t \times V_m^P) \right\} d\Omega$$

$$B_n = \int_{\Omega} \left\{ \frac{1}{\beta^2} \cdot V_m^P \cdot [g_{tr}] \cdot V_m^P + g_{zz} V_{zn} V_{zn} \right\} d\Omega$$

donde  $f_o, g_o, [f_{tr}]$  y  $[g_{tr}]$ , se corresponden con  $\epsilon_o, \mu_o, [\epsilon_{tr}]$  y  $[\mu_{tr}]$  de manera análoga a lo indicado en la Tabla I:

$$\epsilon = \epsilon_o \begin{bmatrix} \epsilon_{xxr} & \epsilon_{xyr} & 0 \\ \epsilon_{xyr} & \epsilon_{yyr} & 0 \\ 0 & 0 & \epsilon_{zzr} \end{bmatrix}$$

$$\mu = \mu_o \begin{bmatrix} \mu_{xxr} & \mu_{xyr} & 0 \\ \mu_{xyr} & \mu_{yyr} & 0 \\ 0 & 0 & \mu_{zzr} \end{bmatrix}$$

Notese que  $[\epsilon]$  y  $[\mu]$  se han supuesto, tanto en 2D como en 3D, simétricos.

Las discretizaciones del campo eléctrico o magnético de los funcionales anteriores (ecuaciones: 4 y 6), se hacen en base a los elementos tipo *curl-conforming* (triángulos en 2D y tetraedros en 3D) [2]. En el caso particular del análisis en onda completa de estructuras de guiado arbitrarias en 2D, se ha implementado un *postproceso* para la caracterización de líneas *multiconductoras*, (la impedancia característica de una línea de transmisión, parámetros primarios y secundarios) [6].

En el análisis en 2D, el software utiliza malla adaptativo, para realizar de un modo más eficiente los cálculos [5]. Entre las ventajas de utilizar este método, pueden encontrarse entre otras, que este algoritmo genera buenas mallas de cálculo, con

errores locales del mismo orden de magnitud, a partir de una malla inicial tosca; o que proporciona una tasa de convergencia independiente del grado de regularidad de la solución. Otra ventaja muy importante, es que los resultados son producidos en base a un nivel de precisión previamente establecido.

### 3. DETALLES DE USO E IMPLEMENTACIÓN

Para clarificar los pasos seguidos, ilustraremos la explicación con un diagrama (Figura 1). En él se pueden observar los pasos seguidos en el análisis del primer modo de una cavidad *ridge*. El estudio de cavidades 3D incluye el cálculo de los autovalores y la visualización de la distribución de campo eléctrico o magnético, según haya elegido el usuario.

En primer lugar, nos encontraremos en la etapa de preproceso, ubicada en el entorno GID. En ella se creará, o bien se importará (DXF, IGES...), la geometría a analizar, la cual, será mejor comprendida, gracias a las posibles rotaciones que se pueden realizar en la interfaz gráfica, programado en *Open GL*. A la geometría se le dotará de condiciones de contorno y se le asignarán los materiales de los que estará compuesto (Figura 1, esquina inferior izquierda). Las condiciones de contorno dependerán de la elección del *problema tipo* a analizar. A continuación, se introducirán los datos generales propios a cada *problema tipo*. Por ejemplo, en el estudio de cavidades 3D tenemos, el tipo de formulación, el intervalo de búsqueda de autovalores... El interfaz donde se incluirán estos datos, está programado en TCL/TK, un lenguaje interpretado que funciona, sin necesidad de compilar, en muchas plataformas y que suele ser utilizado para realizar aplicaciones gráficas [7].

El preproceso proseguirá con la creación de la malla, que en estructuras de seis caras (3D) o de cuatro líneas (2D), puede ser estructurada (cuadrángulos y hexaedros) o, genéricamente desestructurada (triángulos y tetraedros).

Uno de los propósitos conseguidos, es que el usuario realice todas las tareas propias del análisis, en el entorno del *pre-postprocesador*. Para ello se ha realizado una personalización (*customization*)



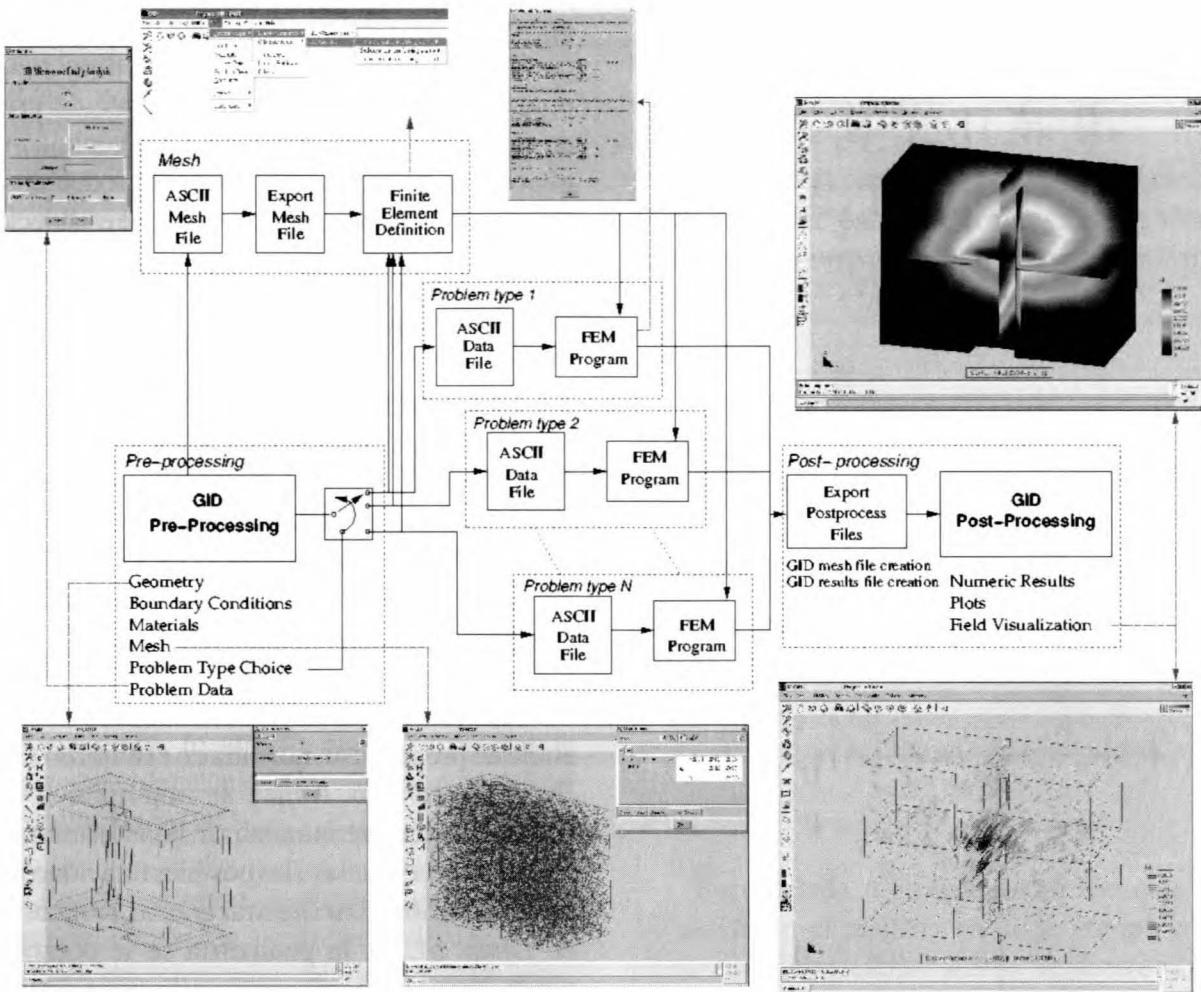


Figura 1. Diagrama del desarrollo software.

del interfaz de GID, de modo que el alumno perciba una integración total del módulo de cálculo con el *pre-postprocesador*.

Una parte de esta personalización (*customization*), como por ejemplo: las posibles condiciones de contorno, o la base de datos de materiales, se lleva a cabo con la ayuda de ficheros ASCII [4]. El resto, como la creación de la barra de herramientas, propia del *problema tipo*, o la creación de nuevos menús, está programado en TCL/TK [7]. El usuario también dispone, en todo momento, de una completa ayuda referente a los detalles de funcionamiento del *problema tipo*, integrada en GID y presentada en formato HTML (Figura 2).

Una vez finalizado el preproceso con la creación de la malla, tenemos que calcular la solución al problema. Para ello, se emplea un código basado en elementos finitos, e implementado en FORTRAN, que tiene como entrada dos ficheros. Por un lado, un fichero ASCII, conteniendo los datos generales

del problema y, por otro, un archivo binario, definido por MODULEF, que incluye la información geométrica de la malla [3].

Los dos ficheros de entrada al código de cálculo, se generan desde el *preprocesador*, cuando se pulse la opción *Calculate*. En este momento,

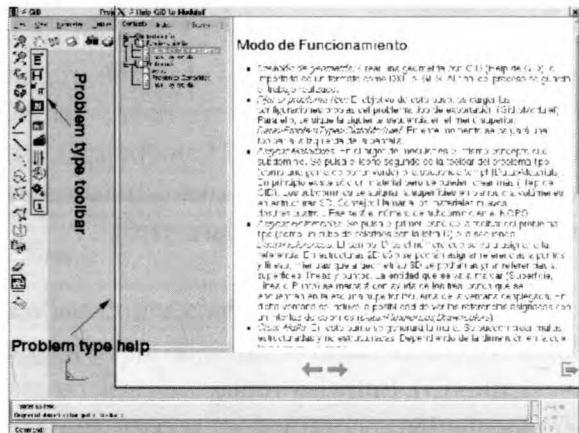


Figura 2. Detalle de la toolbar y de la ayuda de un problema tipo.

dos módulos son ejecutados por parte de GID. En primer lugar, se crean archivos ASCII, en base a códigos programados en un pseudo lenguaje propio de GID, que contienen los datos del preprocesado (información geométrica de la malla, tensores de los materiales...). En segundo lugar, se ejecutará una *shell script*, que irá llamando a sucesivas aplicaciones creadas por el programador del *problema tipo*. Entre estas aplicaciones se encuentran: un conversor *GIDtoMODULEF*, implementado en C (por la rapidez de manipulación con grandes volúmenes de datos); el código para resolver y un conversor *MODULEFtoGID*, tanto para los resultados como para la malla.

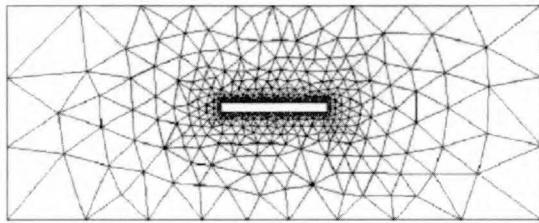


Figura 3. Mallado de la línea stripline.

En las aplicaciones electromagnéticas basadas en FEM, para conseguir una buena discretización de campos eléctricos o/y magnéticos, se hace necesaria la existencia de mallas compuestas por elementos *curl* o *div conforming*. Resaltaremos, que las características de la malla proporcionada por GID, son insuficientes para los cálculos, que potencialmente se podrían realizar, ya que sólo nos ofrece hasta elementos de *Lagrange* de segundo orden. Para salvaresta deficiencia, hemos recurrido

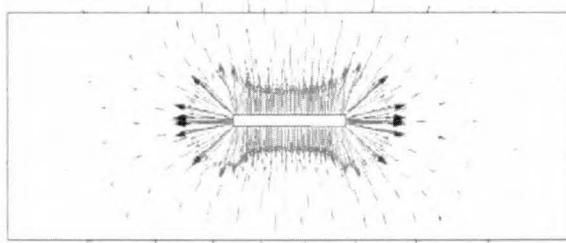


Figura 4. Distribución de módulo de campo por medio de vectores.

a aplicaciones sobre MODULEF, que facilitan la creación de mallas con elementos de ubicación de nodos arbitraria, a partir de mallas con nodos en los

vértices de cada elemento. Así mismo, con estas aplicaciones también dividimos cuadrángulos y cuadriláteros en triángulos y tetraedros, según sea en 2D o 3D, puesto que los códigos trabajan con estas geometrías de elementos. En referencia a esto último, es interesante reseñar, que la herramienta es capaz de crear malla regular. Es decir, la formada por cuadrángulos, en superficies formadas por cuatro líneas, o por hexaedros, en volúmenes formados por seis superficies.

De entre las múltiples posibilidades de la herramienta, y persiguiendo el objetivo de integración total en el interfaz del *pre-postprocesador*, se incluye la posibilidad de poder seguir el progreso del módulo de cálculo a través de una tubería (*pipe*).

Una vez ejecutado el código de elementos finitos correspondiente, resta ordenar los resultados en formato legible para GID. Como se ha dicho anteriormente, también se tuvo que implementar un conversor de malla *MODULEFtoGID*, ya que además de existir casos, en los que la malla inicial difiere de la final estructuralmente (de hexaedros a tetraedros, de cuadrángulos a triángulos o en malla adaptativa), existe una reenumeración de los elementos y nodos cuando el tipo de interpolación cambia.

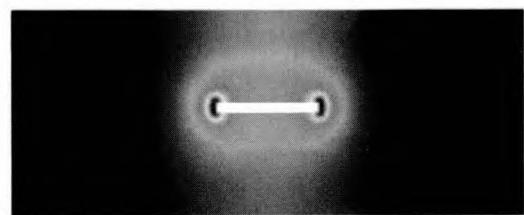


Figura 5. Distribución de módulo de campo por medio de Contour Fil.1

Otra característica digna de mención, es el reducido tiempo empleado por cada *problema tipo*, tanto en convertir la malla a formato NOPO, como en el cálculo. Como muestra, baste indicar que en el ejemplo de la estructura *ridge*, descrito con 3000 elementos de malla y en un Pentium III 1.3 *Celeron* de 256M de RAM, el análisis dura alrededor de cuarenta segundos, para el cálculo de los cinco primeros modos.

#### 4. EJEMPLOS

Para ilustrar, de una forma gráfica, el tipo de resultados proporcionados por la herramienta que se presenta, se ha recurrido a dos ejemplos. Como primer ejemplo se ha elegido el análisis de una *stripline* (anchura de tira unidad, altura del substrato 0.1, características del substrato  $\epsilon_r = 3$ ,  $\mu_r = 1$  y

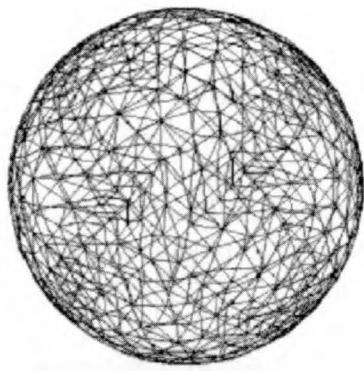


Figura 6. Mallado de una cavidad esférica. Contour Fil.l

y dimensiones normalizadas de la caja que lo aloja 5 por 2. El mallado se presenta en la Figura 3. A modo de ejemplo elegimos la resolución del primer modo TEM, el más intuitivo. Una vez en el postproceso de GID, los resultados pueden ser plasmados, tanto en concentración de campo (*Contour Fill*) (Figura 5), como en distribución de módulo por medio de vectores (Figura 4). Resaltar, que el módulo de cálculo tiene la capacidad de calcular diferentes parámetros, primarios y secundarios, como las posibles variaciones de

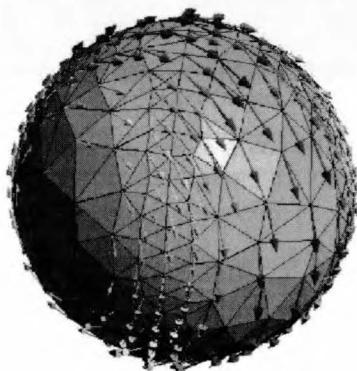


Figura 7. Distribución de módulo de campo por medio de vectores.

impedancia característica, en función de la frecuencia de líneas *microtiras*.

En el segundo ejemplo, se realiza el estudio de los modos de resonancia de una cavidad esférica vacía de radio unidad. El análisis se realiza sin considerar pérdidas, es decir, con condición de pared eléctrica perfecta en la superficie metálica que delimita la esfera. El mallado empleado en el análisis se muestra en la Figura 6. Se puede observar (Figuras 7 y 8), que los resultados proporcionados por el módulo de cálculo, son suficientes para la correcta asimilación de la distribución de campo magnético en este primer modo. El autovalor

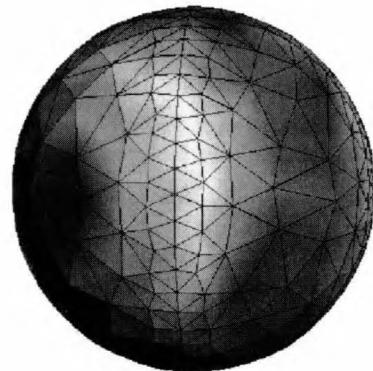


Figura 8. Distribución de módulo de campo por medio de Contour Fil.l

calculado en este ejemplo, es de 7.634382, con un error sobre el analítico de 1.3976 %.

#### 5. CONCLUSIONES

Se ha presentado una herramienta de fácil utilización, rápida, versátil y potente, para resolver varios problemas electromagnéticos, apoyándose en un *pre-postprocesador* de carácter general. Se ha empleado el mismo estándar de datos, referentes a la información geométrica de la malla, para los diferentes tipos de problemas. Ello y una modularización de todas las tareas en lo posible, facilitará en gran medida, la creación de nuevos *kernels* que solucionen los posibles problemas electromagnéticos que vayan surgiendo. A fin de conseguir la mayor difusión posible, la herramienta está portada a *Windows* y *Linux*, además de estar disponible en varios idiomas (Inglés, Francés, Alemán, Japonés, Coreano y Castellano).

## AGRADECIMIENTOS

El autor quiere expresar su agradecimiento a:

*Prof. M. Salazar Palma* por la posibilidad de utilizar su código de mallado adaptativo en estructuras 2D.

*Ministerio de Ciencia y Tecnología* por financiar el proyecto al cual este trabajo pertenece (TIC 2001-1019).

## AUTORES



*Diego Sánchez Repila* nació en Salamanca, en 1980. Actualmente está realizando el proyecto sobre el trabajo presentado, que le permitirá obtener el título de Ingeniero de Telecomunicación. Desde enero del presente año, colabora en el Departamento de Teoría de la Señal de la Universidad de Alcalá (Madrid) merced a una beca concedida por el Ministerio de Educación y Cultura. Es autor de 3 artículos en conferencias internacionales colaborando desde 2002 en un proyecto financiado por el Ministerio de Ciencia y Tecnología, enmarcado en el plan de investigación científica, desarrollo e innovación tecnológica.

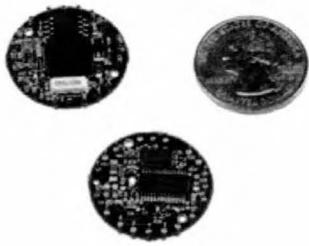
## AGRADECIMIENTOS

- [1] D. Sánchez-Repila, L. E. García-Castillo, «A Software Tool based on the Finite Element Method for Electromagnetic Education», IEEE Antennas and Propagation Society International Symposium (AP-S). Columbus, Ohio, USA, Jun 2003.
- [2] M. Salazar-Palma, T. K. Sarkar, L. E. García-Castillo, T. Roy, y A. R. Djordjevic, *Iterative and Self-Adaptive Finite-Elements in Electromagnetic Modelings*, Artech House Publishers, Inc., Norwood, MA, 1998.
- [3] «*The MODULEF Library*», Institut National de Recherche en Informatique et en Automatique (INRIA), URL: <http://www-rocq.inria.fr/modulef/>.
- [4] «*GiD: The Personal Pre and Postprocessor*», International Center for Numerical Methods in Engineering (CIMNE), URL: <http://gid.cimne.upc.es/>.
- [5] M. Salazar-Palma, L. E. García-Castillo, *Finite Element Software for Microwave Engineering*, capítulo 16, pp. 401-432, Wiley Series in Microwave and Optical Engineering. John Wiley & Sons, 1996.
- [6] M. C. Jiménez-González, L. E. García-Castillo, M. Salazar-Palma, «Computation of Characteristic Impedance and Losses of Multiconductor Anisotropic Transmission Lines Using the Finite Element Method», IEEE Antennas and Propagation Society International Symposium (AP-S). Salt Lake City, USA, Jul 2000.
- [7] John Ousterhout, *Tcl and the Tk Toolkit*, Addison-Wesley.



*Luis E. García Castillo* nació en Madrid, en 1967. Obtuvo los títulos de Ingeniero de Telecomunicación y Doctor Ingeniero de Telecomunicación por la Universidad Politécnica de Madrid en 1992 y 1998, respectivamente. Su tesis recibió dos premios por parte del Colegio Oficial de Ingenieros de Telecomunicación y de la Universidad Politécnica de Madrid. De 1997 a 2000 fue Profesor Titular de Escuela Universitaria de la Universidad Politécnica de Madrid. En el 2000 ingresó en la Universidad de Alcalá (Madrid) como Profesor Titular de Universidad. Su actividad investigadora se centra en la aplicación de métodos numéricos y procesado de señal a problemas electromagnéticos. Es autor de 1 libro, 5 contribuciones de capítulos y artículos en libros, 10 artículos en revistas internacionales, y 51 artículos en conferencias internacionales. Actualmente, es investigador responsable de un proyecto de investigación científica, desarrollo e innovación tecnológica financiado por el Ministerio de Ciencia y Tecnología. También ha participado en 7 proyectos y contratos financiados por instituciones internacionales, europeas y nacionales y/o compañías.





# TÉCNICAS DE ENCAMINAMIENTO EN REDES AD HOC BASADAS EN CRITERIOS DE POTENCIA Y BATERÍA

*Jesus Martín Hermosín, Antonio Pascual Iserte*

*{almartin,tonip}@gps.tsc.upc.es*

*Departamento de Teoría de la Señal y Comunicaciones. Universitat Politècnica de Catalunya*

## RESUMEN

Las redes ad hoc están formadas únicamente por terminales móviles que funcionan a partir de baterías, es decir, no hay estaciones base. El consumo de energía es, por lo tanto, un punto muy importante a la hora de diseñar la red. Para maximizar el tiempo de vida de una red ad hoc, el consumo de energía tiene que ser distribuido de forma equitativa entre el conjunto de nodos que la forman, y a la vez ha de minimizarse el consumo total de energía para cada transmisión. Los algoritmos de encaminamiento clásicos para ad hoc no tienen en cuenta la energía y por tanto no consiguen maximizar el tiempo de vida de la red. En este artículo se hace un estudio del estado del arte en métricas de encaminamiento que tengan en cuenta la batería al seleccionar las rutas y se proponen nuevas métricas al respecto. Dichas métricas se comparan mediante simulación intentando llegar a un compromiso entre parámetros como el tiempo de vida de la red y la calidad de servicio, entre otros.

## 1. INTRODUCCIÓN

Uno de los campos de investigación más activos actualmente es el de las redes inalámbricas ad hoc. De origen militar, durante más de 30 años se ha venido trabajando en este tipo de redes, pero ha sido en los últimos años cuando la proliferación de la tecnología inalámbrica y su bajo coste ha provocado el mayor avance en el desarrollo de la investigación en este campo. Una red inalámbrica ad hoc es una colección de nodos (posiblemente móviles), que se comunican entre sí, sin disponer de ninguna infraestructura fija en la que apoyarse ni estructura predeterminada de enlaces disponibles

[1]. Dichos nodos pueden comunicarse con otros nodos que pueden o no estar dentro de su radio de alcance. En este último caso uno o varios nodos intermedios transportan la información desde el origen hasta el destino, lo que forma una red ad hoc de múltiples saltos.

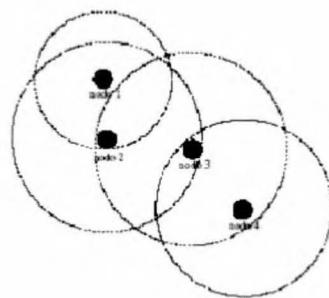
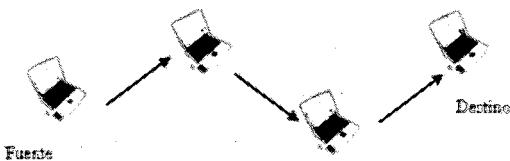


Figura 1.1. Ejemplo de red ad hoc.

Los nodos han de ser capaces de discernir con qué otros nodos pueden comunicarse directamente y con cuáles no, y de detectar la presencia o la desaparición de nodos vecinos. Una característica relativamente rápida debido a la movilidad de los nodos.

Una red ad hoc es auto-organizativa y adaptativa, esto es, que se puede formar o destruir en cualquier momento y lugar sin necesidad de ningún sistema de administración. "Adhoc" tiende a implicar que "puede tomar diferentes formas" y que "puede tener movilidad".

Cada nodo, por lo tanto, hace a la vez las funciones de terminal y router, ya que todos deberán ser capaces de retransmitir la información de la que no sean destinatarios hacia otros nodos hasta que se alcance el destino, como se observa en la figura 1.2.



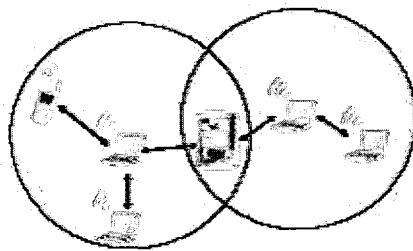
*Figura 1.2. Ejemplo de funcionamiento ad hoc.*

Se puede dar el caso de que uno o más nodos de los que forman la red puedan agotar su batería en tareas de retransmisión y de mantenimiento de la red, sin haber llegado a ser fuente o destino de una llamada o transmisión. Este hecho muestra claramente la necesidad de administrar de manera justa la energía de la red, es decir, de las baterías de los nodos que la componen.

## 2. SERVICIOS

La actual y previsible futura demanda de tecnología ad hoc hace que constantemente salgan a la luz nuevas aplicaciones que en la mayoría de los casos no hacen más que responder a dicha demanda.

Los servicios van principalmente destinados a cubrir las necesidades de nuevas redes autónomas que puedan conectarse a accesos fijos siguiendo la tecnología IP.



*Figura 2.1. Servicios de valor añadido.*

A continuación se hace un listado de posibles aplicaciones y servicios de las redes ad hoc:

" El origen militar de las redes ad hoc hace que en este campo continúen teniendo un papel muy importante tanto en servicio como en desarrollo de la tecnología. La industria militar demanda sistemas totalmente autónomos y muy dinámicos.

" Aplicaciones para rescates, emergencias, y en general en zonas donde no existan infraestructuras disponibles y donde se haga necesaria la formación, en ocasiones muy rápida, de una red de comunicaciones. Adecuadamente combinada con información proveniente de los satélites de comunicación, las redes ad hoc proporcionan un método muy rápido y flexible de establecer comunicaciones.

" Cada vez más son una alternativa en comunicaciones marítimas e incluso se han realizado pruebas de comunicación entre aviones y entre aviones y barcos.

" La creación de WLAN con todas las aplicaciones de las LAN y la aparición de nuevas aplicaciones sacan provecho de la gran ventaja del rápido despliegue de las redes wireless ante las redes cableadas.

" Creación de redes provisionales en todo tipo de eventos o lugares donde sea necesario, tales como congresos, eventos deportivos, etc.

" Tienen un papel importante en las redes de 4G y en el desarrollo de las redes de área personal (PAN), donde se forman redes ad hoc heterogéneas que conectan por ejemplo el ordenador personal con el teléfono móvil o la agenda.

" Un campo al que recientemente se ha prestado bastante atención es el de las redes de sensores [2]. Los sensores son de tamaño muy reducido y poseen capacidades de medición, comunicación vía radio y almacenamiento. Se usan principalmente para realizar mediciones en tareas militares, pero también tienen otras muchas aplicaciones. En la industria médica, los sensores permiten monitorizar información crítica para la vida. En la industria alimenticia se aplican al control de calidad de los productos. El control medioambiental y el uso en la agricultura son otras de sus muchas aplicaciones.

### **3.PROBLEMAS Y RETOS: ENERGÍA**

A pesar de su relativamente larga historia, las redes ad hoc presentan todavía algunos retos por resolver. Principalmente van ligados a los problemas de escalabilidad, seguridad, calidad de servicio o scheduling. En varios trabajos se profundiza en estos aspectos que han sido tratados a lo largo de la historia de las redes ad hoc [3,4,5]. En este artículo nos centraremos en un problema del que no se ha tratado hasta hace bien poco, el problema de la gestión de la energía de la red.

La vida de la red ad hoc depende de la vida de sus terminales. A su vez los terminales móviles dependen de sus baterías como única fuente de energía. Esto hace de la gestión eficiente de la energía de los nodos algo vital para la red. La capacidad de las baterías es finita y representa uno de los mayores problemas a la hora de diseñar un algoritmo de encaminamiento. Estudios en el avance de la tecnología de las baterías muestran que no se esperan muchos progresos en la capacidad de las mismas [3]. La tecnología de las baterías está muy por detrás de la tecnología de los microprocesadores. La vida de una batería de Litio hoy en día es sólo de 2 ó 3 horas [1]. Las limitaciones en la vida de las baterías y los requerimientos de energía debido al encaminamiento hace de la conservación de energía uno de los principales problemas en ad hoc.

#### **3.1 Estrategias para la conservación de energía**

Se han investigado estrategias para la conservación de la energía en los diferentes niveles de los terminales móviles [5] [6]:

" Capa física. Actualmente se han hecho avances en hardware a nivel de circuito en la optimización de la potencia. Pantallas y CPUs que consumen poca potencia o algoritmos computacionales eficientes en consumo de potencia son algunos ejemplos. Otros trabajos han ido dirigidos a conseguir mejorar la eficacia en los componentes actuales.

" Capa de enlace. En la capa de enlace la conservación de la energía puede conseguirse usando esquemas de retransmisión efectivos y operación en modo sleep.

La capa de enlace detecta errores en la transmisión y retransmite los paquetes necesarios. En las redes ad hoc, debido a la presencia de movilidad o interferencia co-canal, los errores de transmisión pueden ocurrir con bastante frecuencia, lo que lleva a frecuentes peticiones de retransmisión. Debido a que dichas peticiones provocan un alto consumo de energía y de interferencia, un nuevo sistema de petición de retransmisiones es necesario. Alguna propuesta al respecto se presenta en [5].

" Capa de red. En las redes con cables los esfuerzos se han centrado tradicionalmente en aumentar el throughput de la red y minimizar el retardo. Elegir rutas que requieran niveles de consumo de potencia menor puede perjudicar los parámetros tradicionales que se basaban en escoger la ruta con el menor número de saltos.

Transmitir con mucha potencia puede disminuir la probabilidad de pérdida de las transmisiones y por lo tanto incrementar el throughput. Sin embargo, esto puede provocar un aumento de la interferencia, bloquear enlaces, y por lo tanto puede provocar una disminución de la capacidad de la red [7], por lo que transmitir a baja potencia no siempre es perjudicial para la red. Desde el punto de vista de la energía, las redes ad hoc necesitan protocolos que maximicen la vida de todos los nodos y con ello maximicen el tiempo antes de que la red se particione, es decir, se dé el caso en que dos nodos no se puedan comunicar entre sí ya que los nodos intermedios necesarios para establecer la comunicación tengan su batería agotada.

### **4.ENCAMINAMIENTO EN AD HOC**

Se puede definir encaminamiento como un proceso mediante el cual se trata de encontrar una ruta o camino entre dos puntos de la red. Los algoritmos de encaminamiento tienen por fin encontrar la mejor ruta entre pares de nodos.

Si nos referimos a redes de comunicación fijas, típicamente al decir "mejor ruta" se consideran aquéllas que cumplen unas condiciones como presentar el menor retardo medio, ofrecer el menor coste o presentar la mayor cadencia media sin tener en cuenta el retardo. Respecto a la métrica pueden tenerse en cuenta parámetros como el número de saltos (el camino que toma un paquete de información desde un nodo a otro de la red), el tiempo de retardo que sufre la información durante su transmisión o la longitud de la cola, etc. Los algoritmos y protocolos de encaminamiento clásicos para ad hoc han mantenido en gran medida las métricas clásicas de selección de rutas, centrando sus esfuerzos principalmente en realizar la gestión de la información de control de la manera más eficiente para así mejorar la escalabilidad de la red, es decir, la capacidad de mantener una calidad de servicio a medida que la red crece.

En ad hoc la forma de actuar de los algoritmos clásicos no funciona de manera eficiente. Ante este hecho recientemente están apareciendo trabajos destinados a buscar nuevos mecanismos de encaminamiento más apropiados para ad hoc [5] [11].

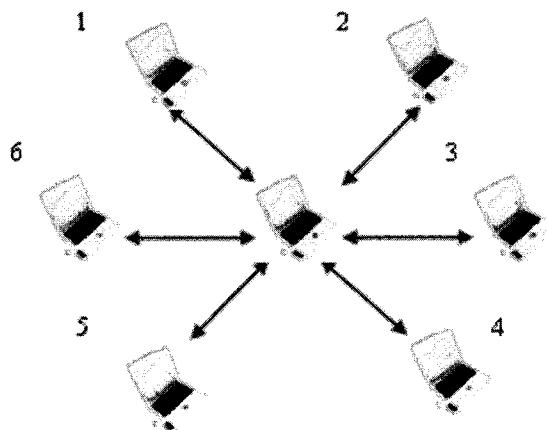


Figura 4.1. Ejemplo de selección de rutas.

En la figura 4.1. se observa que desde el punto de vista del encaminamiento clásico el nodo central agotará su batería en muy poco tiempo. Esto es debido a que forma parte de las rutas con menor número de saltos entre los nodos 1-4, 2-5 y 6-6. Esto provoca que la red se particione rápidamente

ya que el nodo central agotará su batería antes que el resto. Existe la necesidad de la aparición de nuevos esquemas de encaminamiento que gestionen la energía de la red de manera más eficiente, aunque ello suponga aumentar el número de saltos de las rutas.

#### 4.1 Protocolos existentes para ad hoc

Son muchos los protocolos que han aparecido específicamente para redes ad hoc. Se pueden clasificar en proactivos, reactivos e híbridos.

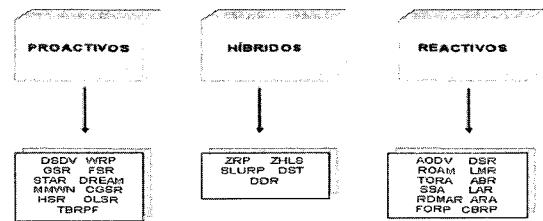


Figura 4.2. Clasificación de los protocolos para ad hoc.

En los protocolos proactivos o globales cada nodo mantiene información de encaminamiento de los demás nodos de la red. La información de encaminamiento se almacena en diferentes tablas. Responden a los cambios de topología propagando actualizaciones para mantener una red estable [8]. La diferencia entre ellos está en la forma en que se propaga esta información de actualización y en el tipo de información guardada en las tablas de encaminamiento.

Los protocolos reactivos se diseñaron para reducir la carga de control que se produce en los protocolos proactivos. Para ello se mantiene la información solamente de los nodos activos (On-demand). Las rutas se determinan y mantienen solamente cuando un nodo en cuestión tiene que enviar información.

Los protocolos híbridos son una nueva generación de protocolos que incluyen características proactivas y reactivas. Se han pensado para mejorar la escalabilidad y reducir la carga de información necesaria para el descubrimiento de rutas. Para ello permiten a los nodos cercanos trabajar como una unidad, es decir, sin necesidad de intercambio periódico de información de control.



## 5. METRICAS DE ENCAMINAMIENTO PARA REDES AD HOC

En este apartado se presentan esquemas de encaminamiento que utilizan métricas que pretenden ser justas a la hora de administrar el consumo de batería de los nodos de una red ad hoc. Este estudio trata de encontrar nuevas métricas que sustituyan a las clásicas aplicadas por la mayoría de los protocolos de encaminamiento.

### 5.1. MTPR (Minimum Total Transmission Power Routing)

Este es uno de los primeros algoritmos que introdujeron el consumo de energía como métrica de selección de rutas. Fue propuesto por K. Scott y N. Bambos [7]. La potencia requerida por cada nodo puede dividirse en potencia de procesado y potencia de transmisión. La potencia de transmisión es la potencia consumida en la comunicación por radio con otros nodos, y es en su estudio en el que se centra el algoritmo.

Para transmisiones correctas, la relación señal a ruido (SNR) recibida en un nodo tiene que ser superior a un umbral. Este umbral está estrechamente relacionado con la probabilidad de error de bit (BER) máxima tolerable en el enlace para la transmisión de la señal recibida. Por lo tanto se observa que para una transmisión satisfactoria desde un nodo  $n_i$  hasta un nodo  $n_j$ , la relación señal a ruido en el nodo  $n_j$  tiene que satisfacer la siguiente ecuación:

$$SNR_j = \frac{P_i^\alpha G_{i,j}}{\sum_{k \neq i} P_k^\alpha G_{k,j} + \eta_j} \geq \psi_j$$

donde  $P_i^\alpha$  es la potencia de transmisión del nodo  $n_i$ ;  $G_{ij}$  es la ganancia en potencia del canal entre los nodos  $n_i$  y  $n_j$  y  $\eta_j$  es la potencia de ruido térmico en el nodo  $n_j$ .

Según este criterio, la potencia mínima de transmisión depende del ruido interferente, de la probabilidad de error deseada y, lo que es más interesante, de la distancia entre nodos.

El algoritmo MTPR tiene como objetivo minimizar la potencia total de transmisión de la ruta.

A continuación se especifica el algoritmo:

" Se define:

$A$ , conjunto de todas las rutas posibles que unen dos nodos.

" La potencia total de transmisión para una ruta  $l$ , se puede obtener de:

$$P_l = \sum_{i=0}^{D-1} P(n_i', n_{i+1}') \text{ para todo nodo } n_i \in l$$

Donde  $n_o$  y  $n_d$  son los nodos fuente y destino respectivamente.

" De lo que se trata es de tomar la ruta deseada,  $l_c$ , como la ruta que minimiza la potencia total de transmisión necesaria:

$$P_{lc} = \min P_l, \text{ para } l \in A$$

" Se puede observar que debido a su naturaleza, este algoritmo puede tomar rutas más largas que otros algoritmos, lo que provoca un mayor retardo. Esto hace que en ocasiones no sea atractivo para una red ad hoc.

### 5.2. MBCR (Minimum Battery Cost Routing)

La métrica del coste total de potencia presenta el inconveniente de que la reducción en el consumo de energía puede no verse reflejado en cada terminal. Por ejemplo si las rutas de menor coste total pasan todas por un nodo en concreto, la batería de este terminal se consumirá en muy poco tiempo.

En [9] se proponen varias métricas que buscan maximizar el tiempo de vida de los terminales que forman una red. Una de ellas es utilizada en [5] como base del algoritmo MBCR. El algoritmo MBCR propone tener en cuenta la batería residual

de los nodos para evitar las rutas que contengan los nodos con menor batería.

" Se define:

$B'_i$ , capacidad residual de batería de un nodo  $n_i$  en un instante de tiempo  $t$ .

$f_i(B'_i)$ , función de coste de batería de un nodo  $n_i$ .

Se propone una función de coste que haga que un nodo sea más reacio, es decir, le "cueste más" encaminar, a medida que su capacidad residual de batería es más baja. Se propone la siguiente:

$$f_i(B'_i) = \frac{1}{B'_i}$$

" Se calcula el coste de batería  $C_l$  para la ruta  $l$  como la suma de los costes individuales de los nodos que la forman:

$$C_l = \sum_{i=0}^{D_l} f_i(B'_i)$$

donde  $D_l$  es el nodo destino de la ruta  $l$ .

" Se toma la ruta que presenta el menor coste de batería de la ruta:

$$C_{lc} = \min \{C_l \mid l \in A\}$$

Debido a que la capacidad de batería disponible se incorpora directamente en el algoritmo de encaminamiento, esta métrica evita que los nodos puedan ser sobrecargados. A pesar de que este esquema prolonga tanto el tiempo de vida de la red como el de los nodos en individual., presenta un problema que queda patente en la figura 5.1.

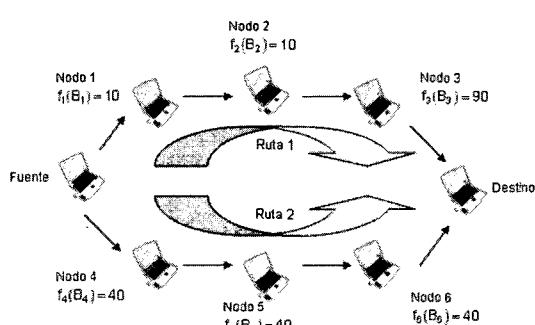


Figura 5.1. Problema de encaminamiento.

En la figura anterior, existen dos posibles rutas entre el nodo fuente y destino. Aunque el nodo 3 tiene mucha menos batería restante que los demás nodos (su función de coste es sensiblemente superior), la ruta 1, la que incluye el nodo 3, presenta un menor coste total de batería. Por lo tanto la ruta 1 es la elegida, reduciendo el tiempo de vida del nodo 3, algo como se ha venido comentando, no es deseable.

### 5.3 MMBCR (Min-Max Battery Cost Routing)

El objetivo del esquema MMBCR es distribuir el consumo de batería entre los diferentes nodos, protegiendo la red del rápido consumo de la batería de unos pocos nodos. Se trata de evitar que los nodos con menor capacidad de batería restante soporten la mayoría del tráfico sólo por el hecho de estar situados en una ruta con menor coste de batería total. Con ello se pretende alargar la vida de los nodos y de la red.

Se propone la modificación del cálculo del coste de batería  $C_l$  para la ruta  $l$  de la siguiente manera:

$$C_l = \max f_i(B'_i) \text{ para } i \in l$$

Es decir, el coste de la ruta viene marcado por el nodo con valor de la función de coste más alto, esto es, con menor capacidad de batería restante.

" La ruta elegida  $lc$  es aquella que tenga el menor coste de todas las posibles:

$$C_{lc} = \min \{C_l \mid l \in A\}$$

En principio al evitar los nodos más débiles parece que la energía de la red se trata de manera más justa. Sin embargo, por un lado no se garantiza que la ruta escogida sea la más corta, con el mayor retardo que ello conlleva. Por otro lado, no hay garantía de que en todos los casos la ruta con menor coste de energía de transmisión total sea la seleccionada. Esto, de hecho, reduce el tiempo de vida de todos los nodos. Todas estas consideraciones quedan patentes en la figura 5.2,



donde se selecciona la ruta 2, con mayor coste total de batería.

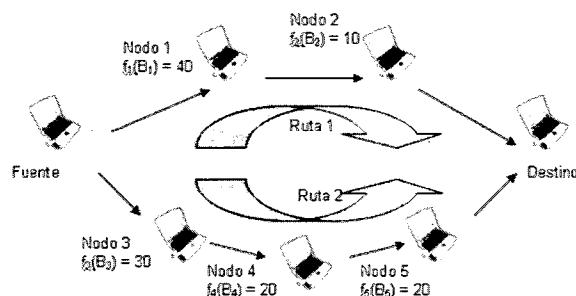


Figura 5.2. Problema de encaminamiento.

#### 5.4CMMBCR (Conditional Max-Min Battery Capacity Routing)

Los esquemas anteriores no consiguen los objetivos de maximizar el tiempo de vida de los nodos y de la red simultáneamente. Para resolver este problema, el esquema CMMBCR [5] introduce un umbral de batería a la hora de seleccionar las rutas. Lo hace de la siguiente manera:

1. Cuando todos los nodos incluidos en las posibles rutas disponen de suficiente capacidad de batería restante, es decir, por encima de un umbral  $\square$ , entonces se escoge la ruta con el mínimo consumo total de potencia transmitida según el esquema MTPR.

2. Cuando todas las rutas posibles contienen algún nodo por debajo del umbral, se evitan las rutas que contengan los nodos más débiles.

"Se define la capacidad residual de batería para una ruta  $l$  en un tiempo  $t$  como:

$$C_l^B = \min B_i^t \text{ para } i \in \text{ruta } l$$

"Se define un vector  $Q$  donde se encuentran todas las rutas con batería residual suficiente:

$$C_l^B \geq \gamma, \forall l \in Q$$

entonces se llega a:

1. Si  $A \cap Q \neq \emptyset$ , lo que implica que existe alguna ruta entre el nodo fuente y destino en la que

todos los nodos disponen de un nivel de batería restante superior al umbral  $\gamma$ , se elige la ruta incluida en  $A \cap Q$  aplicando el esquema MTPR [18].

2. En otro caso, se selecciona la ruta con la máxima batería restante en su nodo más débil usando la ecuación:

$$C_{lc}^B = \max \{C_l^B \mid l \in A\}$$

El factor  $\gamma$  es un margen de protección y presenta un compromiso para el funcionamiento del algoritmo.

#### 5.5 CMBCR (Conditional Minimum Battery Cost Routing)

Con este nuevo esquema se propone utilizar la idea propuesta en el algoritmo CMMBCR pero destinada íntegramente a minimizar el coste de batería a la hora de seleccionar las rutas. Se propone lo siguiente:

1. Por un lado, y al igual en el esquema CMMBCR, se define un umbral  $\square$ . Si todos los nodos que componen la ruta tienen capacidad de batería restante por encima del umbral, se utiliza en esquema de mínimo coste de batería total para así alargar la vida de toda la red.

2. Cuando no exista más remedio que usar una ruta que contenga algún nodo con capacidad de batería restante por debajo del umbral, se tratará de evitar los nodos con menor batería, para así alargarles la vida.

" Se utilizan las funciones de coste y vectores que se han venido utilizando a lo largo del texto.

" Se define el coste total de batería:

$$C_l^1 = \sum_{i=0}^{D_{l,1}} f_i(B_i^t)$$

" Por otro lado se define la capacidad residual de batería de la ruta:

$$C_l^2 = \min B_i^t \text{ para } i \in l$$

" A partir de estos costes se aplica la misma idea que el algoritmo anterior y se evitan los nodos más débiles cuando la batería de todos los que forman la ruta es inferior a un umbral.

Esta técnica controla la gestión de la batería pero no tiene control sobre la potencia de transmisión en la selección de rutas, lo que hace que no en todos los casos se gestione la energía de la red de la manera más justa.

### 5.6 TVNR (Tiempo de Vida de los Nodos en Ruta)

Este esquema trata de hacer hincapié en una cuestión que no han tenido en cuenta los esquemas anteriores, al menos de manera explícita. Por mucha batería restante que disponga un nodo, si éste tiene que soportar el tráfico derivado de los demás nodos, agotará su batería rápidamente. Para evitar este problema se propone tener en cuenta el consumo medio por unidad de tiempo que sufre un nodo mediante la definición de un parámetro  $Bc_i^t$ , que no es más que el consumo de batería medio por un nodo  $n_i$  por unidad de tiempo, en un instante de tiempo  $t$ .

" Se define el tiempo de vida estimado bajo el consumo actual de un nodo según la ecuación:

$$T_i(B_i^t, Bc_i^t) = \frac{B_i^t}{Bc_i^t}$$

" Se selecciona la ruta con el mayor tiempo de vida como:

$$T_{vida} = \max \{T_i \mid i \in A\}$$

Una parte crítica del diseño de este esquema es la selección del parámetro  $Bc_i^t$ , para el cual se tiene en cuenta el valor calculado en diferentes intervalos de tiempo según la forma:

$$Bc_i^t = \alpha Bc_{ant} + (1 - \alpha) Bc_{act}$$

Con el parámetro  $\alpha$  se regula si se quiere dar más valor al consumo medio calculado durante todo el tiempo de vida del nodo o al calculado en el último intervalo.

### 5.7 TPBCR (Transmission Power-Battery Cost Routing)

Los algoritmos anteriormente descritos tienen en cuenta la potencia de transmisión o la batería en sus métricas pero no simultáneamente, aunque en todos ellos ambas tienen un impacto importante en el tiempo de vida de la red. Si se selecciona una ruta atendiendo únicamente a la batería restante de los nodos sin tener en cuenta la potencia que consumirá, puede suceder que un nodo agote rápidamente su batería si éste se encuentra muy alejado del que le sucede en la ruta; es decir, se pierde un término importante a la hora de controlar el consumo de energía, y por lo tanto de controlar el tiempo de vida de la red.

Lo que se pretende con el algoritmo TPBCR es que a la hora de decidir qué nodo sigue en la selección de la ruta se tenga en cuenta la potencia que se va a consumir al transmitir y simultáneamente la capacidad de batería restante del nodo en cuestión.

" Según lo anterior se define una función de coste que tenga en cuenta la potencia de transmisión y la capacidad de batería restante en el nodo:

$$f_i(B_i^t, P_{ij}^t) = \frac{P_{ij}^t}{B_i^t}$$

" De lo que se trata es de tomar la ruta con menor coste total en sus nodos:

$$C_l = \sum_{i=0}^{D_{l-1}} f_i(B_i^t, P_{ij}^t)$$

$$C_{lc} = \min \{C_l \mid l \in A\}$$

Se puede pensar que un nodo con poca batería restante pero que necesita poca potencia de transmisión puede ser elegido en lugar de un nodo con más batería si éste necesita una potencia de transmisión mayor, por lo que en este caso se penalizará al nodo con menor batería. Lo que prima es la relación entre potencia y batería.



## 5.8 CTPBCR (Conditional Transmission Power-Battery Cost Routing)

Para evitar lo comentado anteriormente se ha propuesto aplicar un umbral de batería al esquema TPBCR. Así, de la misma manera que en los demás esquemas condicionales, cuando no existen rutas formadas por nodos con batería residual por encima de un umbral, se evitan los nodos con menor batería restante.

Para comprobar las suposiciones teóricas y determinar qué soluciones dan mejores resultados se hace necesaria la comparación mediante simulación.

## 6. EL SIMULADOR

### 6.1. Estructura del simulador

Se permite simular tanto redes ad hoc de múltiples saltos como de salto único, dentro de un área limitada definida por el usuario. La creación del escenario y de la red se controla a partir de los parámetros de entrada. Así se pueden crear desde una red de dos nodos en un área pequeña hasta una red de cientos de nodos que ocupe una gran área.

El simulador permite la formación de redes estáticas y dinámicas. En estas últimas los nodos toman direcciones aleatorias que varían cuando llegan a los límites del área permitida.

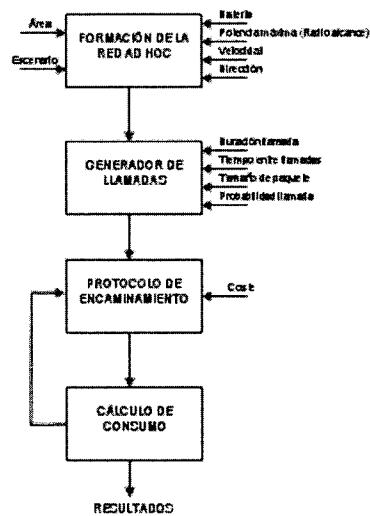


Figura 6.1. Diagrama de bloques del simulador.

## 6.2. Generador de llamadas

A partir de los parámetros de tiempos, se generan las llamadas de forma aleatoria según una variable aleatoria exponencial.

Los nodos generan llamadas con una probabilidad determinada e indicada, y el nodo destino se elige de manera aleatoria.

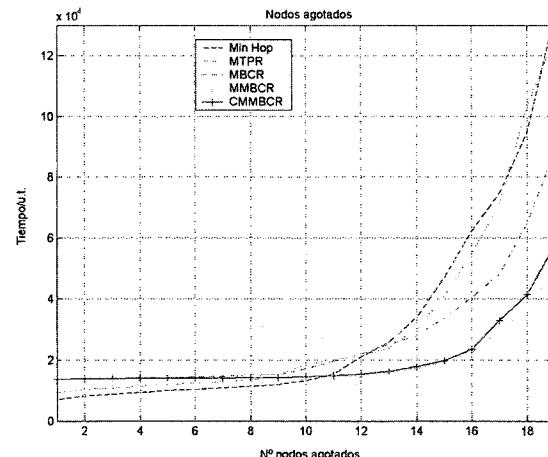
## 6.3. Protocolo de encaminamiento

Para evitar tener que aplicar el protocolo de encaminamiento a nivel de paquete se ha tomado la solución de realizar un protocolo a nivel de llamada. Los nodos generan peticiones de llamada que se envían por la red y ponen en marcha los diferentes algoritmos de encaminamiento.

## 7. RESULTADOS

Se han realizado un gran número de simulaciones y comparaciones en diferentes escenarios con diferentes características cada uno. El objetivo principal ha sido comparar el comportamiento de las métricas en diferentes situaciones y desde diferentes puntos de vista (tiempo de vida de la red, número medio de saltos por ruta, tráfico cursado...). Además de todas las métricas expuestas anteriormente se ha introducido la métrica del mínimo número de saltos (MH), típica de los protocolos clásicos.

Aquí se presenta alguna muestra de los resultados obtenidos.



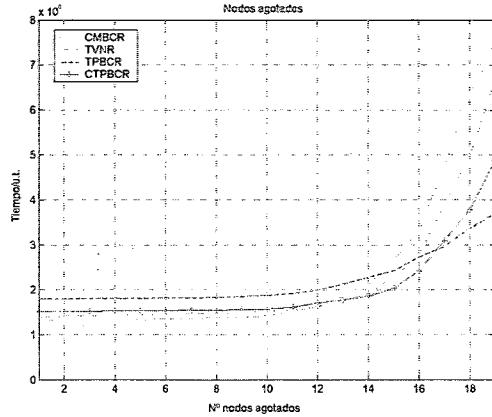


Figura 6.2. Secuencia de expiración de los nodos

Se observa que el esquema que agota sus nodos más rápidamente es el MH. Al no tener control sobre la energía, selecciona rutas que perjudican a unos nodos en concreto, que agotan su batería rápidamente. A medida que los nodos caen se llega al punto de que no presenta el menor número medio saltos (figura 6.3). Esto, aunque puede parecer extraño, es lógico ya que se al no gestionar la batería de los nodos adecuadamente, agota los recursos de la red rápidamente viéndose obligado a tomar rutas largas, ya que son las únicas que permanecen activas. Respecto al tráfico cursado, se observa que presenta nuevamente el peor comportamiento (figura 6.4). Esto se debe nuevamente a la mala gestión de los recursos de la red.

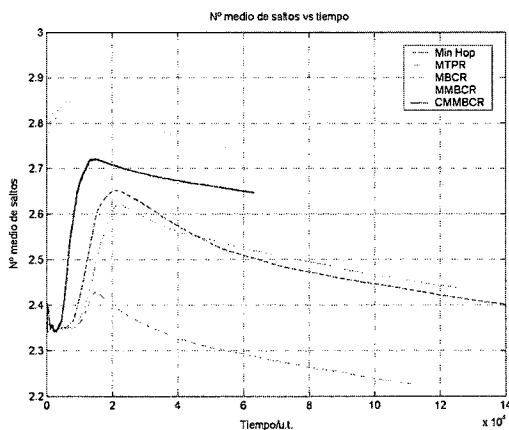


Figura 6.3. Número medio saltos vs tiempo.

Esquemas como el MBCR o CMBCR, que seleccionan sus rutas a partir del coste total de batería, presentan un buen comportamiento respecto

a la gestión del tiempo de vida de la red. El problema que presentan los esquemas que no introducen la potencia de transmisión es que son poco robustos ante los cambios en el radio de cobertura. La potencia de transmisión está muy relacionada con el radio de cobertura de los nodos. Es por esto que para redes con nodos muy alejados se pueden tomar enlaces muy largos que agoten rápidamente la batería de los nodos.

La métrica TPBCR es la que trata de manera más justa la energía de los nodos. Los nodos se agotan en instantes próximos. Esto significa que los recursos de la red son gestionados para sacar el máximo partido de la misma, sin abusar de unos pocos nodos y evitando que la red se particione. Esto da lugar a altos porcentajes de tráfico cursado respecto a los demás esquemas (figura 6.4).

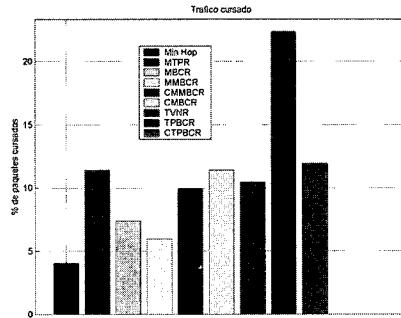


Figura 6.4. Porcentaje de tráfico cursado.

## 8. CONCLUSIONES

En este artículo se ha puesto en evidencia la necesidad de aplicar nuevas técnicas de encaminamiento que tengan en cuenta la energía de la red. Se ha hecho una revisión del estado del arte y se han propuesto nuevas técnicas destinadas a mejorar las ya realizadas.

Se puede concluir que el comportamiento de las diferentes métricas depende de las características de la red. Claramente la técnica comúnmente aplicada, la minimización del número de saltos, no es óptima para redes donde la gestión de la batería de la red es vital para la misma. Las propuestas que solamente tienen en cuenta la batería de los nodos al seleccionar las rutas presentan un buen comportamiento en general para redes con los



nodos relativamente próximos. El mejor comportamiento para todo tipo de redes lo presenta el esquema TPBCR, que incluye la métrica de la potencia de transmisión y el coste de batería simultáneamente.

Quedan muchas cuestiones todavía por resolver sobre las redes ad hoc. Cualquier avance en cualquiera de las capas ayudará a formar redes más eficientes. En concreto se está poniendo énfasis en la relación de la distancian entre los nodos y la potencia de transmisión necesaria [10]. Nuevos avances en este terreno mejorarán los resultados de los esquemas de encaminamiento que introducen la métrica de la potencia de transmisión.

## REFERENCIAS

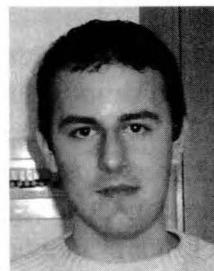
- [1] Toh,C-K. (2002). Ad hoc Mobile Wireless Networks, Procols and systems. Prentice-Hall.
- [2] Xu,Y.; Heidemann,J.; Estrin,D.(2000). Adaptive Energy-Conserving Routing for Multihop Ad Hoc Networks.USC/ISI Technical Report.
- [3] Chlamtac, I; Conti, M.; Liu, N. (2003). Mobile ad hoc networking: imperatives and challenges. Ad hoc networks. 1(1):13-64.
- [4] Internet Engineering Task Force (IETF) internetdraft, <http://www.ietf.org/html.charters/manet-charter.html>, 1998.
- [5] Toh, C-K..(2001). Maximum Battery Life Routing to Support Ubiquitous Mobile Computing in Wireless Ad Hoc Networks. IEEE Communications magazine. (39):138-147.
- [6] Jones,C.; Sivalingam,K.; Agarwal,P.; Chen,J.C. (2001). A survey of energy efficient network protocols for wireless and mobile networks, ACM/Kluwer Wireless Networks 7(4)343-358.
- [7] Scott, K.; Bambos, N. (1996). Routing and channel assignment for low power transmission in PCS. Proceedings of IEEE International Conference on Universal Personal Communications (ICUPC '96), Cambridge.
- [8] Abolhasan, M.; Wysocki, T.; Dutkiewicz,E.(2004). A review of routing protocols for mobile ad hoc networks. Ad hoc networks. 1(2).

[9] Singh, S.; Woo, M.; Raghavendra, C. (1998). Power-aware routing in mobile ad hoc networks. Proceedings of Mobicom'98.181-190.

[10] Stojmenovic,I.; Lin,X. (2000). Power-Aware Localized Routing in Wireless Networks. Proceeding of IEEE International Parallel and Distributed Computing Simposium, 2000.

[11] Martín, J; (2004). Técnicas de encaminamiento en ad hoc basadas en criterios de potencia y batería. PFC en la ETSETB. Director: Pascual, A. Biblioteca Rector Gabriel Ferraté (BRGF).

## AUTORES

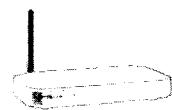


*Jesús Martín, nacido en 1980, comenzó los estudios de ingeniero superior de telecomunicación en 1998 en la Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona (ETSETB) de la Universitat Politècnica de Catalunya (UPC). En septiembre de 2004 presentó su proyecto final de carrera sobre la investigación de nuevas técnicas de encaminamiento en redes ad hoc que tengan en cuenta y gestionen la capacidad de energía de la red en su funcionamiento, realizado en el departamento de Teoría de la Señal y Comunicaciones (TSC).*



*Antonio Pascual nació en 1977 en Barcelona. Cursó los estudios de ingeniero superior de telecomunicación en el período 1995-2000 en la Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona (ETSETB) de la Universitat Politècnica de Catalunya (UPC), obteniendo el premio nacional de fin de carrera correspondiente al curso 2000-2001. Durante el período de septiembre de 1998 hasta junio de 1999 trabajó como becario en el departamento de ingeniería electrónica de la UPC, y desde junio de 1999, y hasta enero de 2001, trabajó, primero como becario y después como ingeniero contratado, en el departamento de I+D de Retevisión, involucrándose en temas relacionados con la implantación de la redes de televisión digital y radio digital terrestre en España. En enero de 2001 se incorporó como estudiante de doctorado y becario del programa de "Formació d'Investigadors" de la Generalitat de Catalunya al departamento de Teoría de la Señal y Comunicaciones (TSC) de la UPC, donde actualmente desarrolla en su última etapa de tesis actividades de investigación en el grupo de procesado de arrays y señales multicanal. Desde septiembre de 2003, es profesor en la Escuela Politécnica Superior de Castelldefels (EPSC) de la UPC.*

# SISTEMA INTEGRAL DE COMUNICACIÓN SEGURA BASADO EN POWERLINE



Fernando Méndez Rebollo

*fernando\_mendez@ieee.org*

Universidad de Alcalá – Escuela Politécnica, Campus Universitario, Ctra. Madrid-Barcelona  
Km. 33.600 28806 Alcalá de Henares (Madrid), España.

## ABSTRACT

En este artículo se presenta un sistema digital de comunicación basado en subsistemas Powerline y Ethernet. El fin perseguido es proveer de un canal de comunicación segura a cualquier aparato conectado a la red eléctrica. Sus principales ventajas son fiabilidad, escalabilidad, facilidad de instalación y bajo coste. El hardware de transmisión sobre red eléctrica ha sido diseñado para cumplir con los estándares CENELEC y debido a su precio y tamaño, puede ser incorporado a casi cualquier dispositivo eléctrico. La seguridad está garantizada mediante el uso de SSL en el enlace Ethernet y SSL o TEA-Block sobre el enlace Powerline. Se ha prestado especial atención a la gestión del sistema, siendo éste capaz de detectar nuevos elementos en la red y configurarlos automáticamente.

## 1. INTRODUCCIÓN

La solución que se presenta a continuación pretende abordar la implementación de un sistema de comunicación segura extendida a través de la red eléctrica doméstica de 125 ó 220V y 50 ó 60Hz. Las principales características que se han impuesto como requisitos del sistema son: facilidad de integración en un entorno doméstico o laboral, escalabilidad, inmunidad ante ataques contra la seguridad del sistema y bajo coste.

Aunque las aplicaciones para este tipo de red de comunicación son innumerables en el campo de la domótica, este artículo se centra en la implementación de un sistema antirrobo para todo tipo de aparatos eléctricos conectados a la red eléctrica. Este concepto es objeto de la patente en trámite número: P200300038, presentada por el autor en Enero de 2003.

El sistema aquí presentado pretende cumplir con las restricciones que una aplicación de este tipo impone al diseño. La idea general para abordar el problema es utilizar la red eléctrica como medio de transmisión, por el que se mantendrá una comunicación constante con los aparatos que pudieran ser sustraídos. Si un aparato intenta ser robado, necesariamente ha de ser desconectado de la red eléctrica, con lo que se pierde la comunicación y se detecta el hurto. Un equipo maestro debería ser el encargado de manejar todas las comunicaciones y de iniciar la alarma en caso de hurto, pero ya que la distancia de transmisión sobre red eléctrica es limitada, serán necesarios varios maestros. Para poder controlar todo el conjunto de forma centralizada, se ha optado por conectar todos los equipos maestros a la red de área local presente en el edificio en el que se quiere instalar el sistema. De esta forma, se puede situar en cualquier parte de la red una consola central de monitorización. Esta consola tendrá un control total y a través de ella se informará de los posibles hurtos.

Se intentará dar solución a todos los problemas que históricamente han tenido este tipo de sistemas distribuidos [1]. Existen varias ventajas en la utilización de la red eléctrica para la comunicación de información en un sistema de seguridad, entre las que destacan:

- 1) La infraestructura de red ya existe.
- 2) La red eléctrica se extiende a todos los puntos de un edificio.
- 3) Los equipos a proteger se encuentran constantemente conectados al medio.
- 4) Es fácil implementar un canal de bajo ancho de banda sobre red eléctrica.
- 5) Es imposible saber a priori si un equipo está o no siendo monitorizado.



Como desventajas principales de este medio de comunicación se tiene:

- 1) El ancho de banda se contrapone a la distancia máxima de comunicación en implementaciones de bajo coste.
- 2) La distancia máxima de transmisión está además limitada por factores como la atenuación del canal y la topología de la red.
- 3) El ruido introducido en la red eléctrica por otros aparatos a ella conectados, puede disminuir las prestaciones del conjunto.

El hecho de tener un ancho de banda limitado, no es crítico en un sistema que sólo pretenda dotar de protección antirrobo a los equipos eléctricos, ya que la información enviada es mínima. El problema más grave es, por tanto, la distancia máxima de transmisión. Este problema es solventado mediante el uso adicional de una red de área local. Para explotar de forma lógica ambas redes, se utilizan dos tipos de dispositivos en concordancia con lo ya expuesto:

1) Esclavos: Dispositivos que serán incorporados a los equipos eléctricos que desean ser protegidos contra robo. Este dispositivo sólo dispone de comunicación a través de la red eléctrica y por tanto no podrá comunicarse a grandes distancias.

2) Maestros: Dispositivos capaces de conectarse tanto a la red eléctrica como a la red de área local y que harán de nexo entre ambas redes. Así, un maestro da la posibilidad de comunicación a través de la red de área local a todos aquellos esclavos que se encuentren a una distancia inferior a 200 metros.

La figura 1 proporciona una visión general de la arquitectura.

En las siguientes secciones, se abordará primeramente el funcionamiento general del sistema, tras lo cual se describirán los distintos medios de transmisión utilizados. A continuación se expondrá el hardware empleado y por último se describirán los protocolos y sistemas de seguridad aplicados.

## 2. FUNCIONAMIENTO DEL SISTEMA

Para cumplir con las especificaciones, se ha optado por incluir en la fuente alimentación de cada equipo eléctrico que quiere ser protegido contra robo, un dispositivo esclavo capaz de comunicarse a través de la conexión a la red eléctrica del propio aparato. Esto posibilita la instalación del sistema sin variación de las infraestructuras del edificio. Así, una vez que los equipos están dotados de esta vía de comunicación, pueden ser interrogados constantemente para garantizar su presencia, de tal forma que un equipo que sea desconectado del enchufe será incapaz de responder a tales requerimientos.

Como ya se ha dicho anteriormente, las comunicaciones a través de la red eléctrica tienen notables limitaciones de distancia, por lo que será necesario que el dispositivo encargado de realizar el control de los esclavos sea un dispositivo con conexión a la red de área local. Por ejemplo, en un edificio de oficinas, donde se requiere tener un control centralizado de todos los ordenadores alojados en las distintas plantas, será necesario disponer de varios maestros.

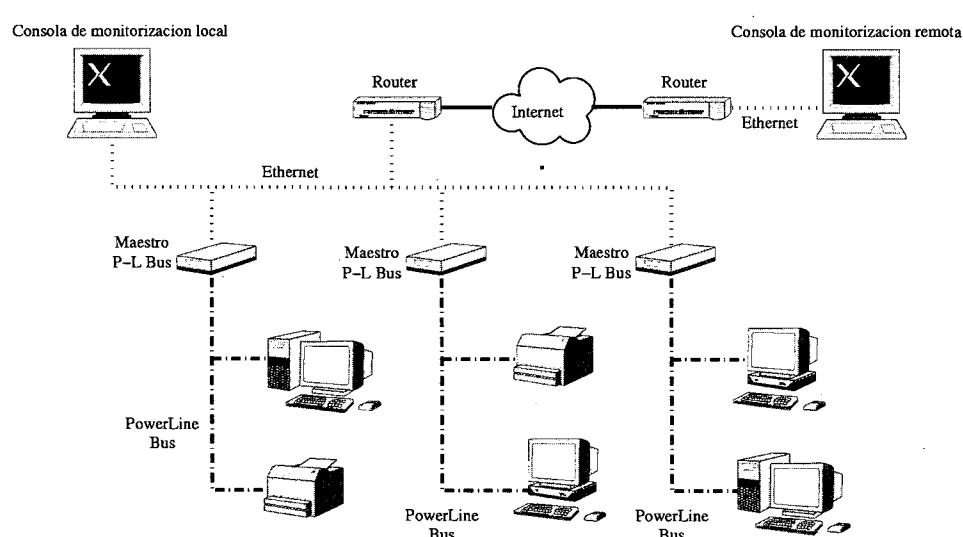


Figura 1. Implementación del sistema

Este tipo de red, representa la realidad de las comunicaciones informáticas de la inmensa mayoría de entornos laborales, por lo cual será accesible desde casi cualquier parte del edificio. Así pues, existirá una consola central encargada de realizar en última instancia la monitorización de los aparatos que se quieren proteger. Para ello, preguntará a los distintos maestros dentro del edificio sobre la presencia o ausencia de los equipos que residen bajo su control. Los maestros a su vez estarán encargados de mantener la comunicación con todos los esclavos bajo su control y poner esta información a disposición de la consola central. De esta forma si un aparato es sustraído, el maestro al que estaba conectado registra su ausencia y comunica tal información a la consola central. De semejante manera, si un maestro es eliminado, la consola central perderá la comunicación con él e informará de tal suceso.

### 3. DESCRIPCIÓN DE LOS MEDIOS DE TRANSMISIÓN

#### 3.1. Red eléctrica doméstica

La red eléctrica doméstica posee unas características bastante inusuales bajo el punto de vista de los medios tradicionales de comunicación [2]. Uno de los puntos más conflictivos a la hora de caracterizar este medio, es el hecho de que no cumple dos de las condiciones necesarias para poder aplicar el teorema de superposición de señales. Así, en muchas ocasiones, la red eléctrica no cumple condiciones de linealidad ni de invarianza en el tiempo. Otro problema que se presenta en la utilización de este medio de transmisión, es el ruido de múltiples tipos que es injectado en la red eléctrica por los aparatos a ella conectados. En la figura 2 se representan algunas de las fuentes de ruido más comunes en entornos domésticos.

El ruido producido por estos aparatos puede clasificarse de la siguiente manera:

1) Ruido impulsivo de baja frecuencia: Tiene frecuencia doble de la corriente alterna de la red. Es generado principalmente por aparatos que utilicen triacs, como por ejemplo reguladores de luz.

2) Ruido tonal: Es generado principalmente por las fuentes de alimentación conmutadas. Su frecuencia fundamental puede estar en el rango desde 20KHz hasta 1MHz.

3) Ruido impulsivo de alta frecuencia: Es producido por motores de corriente alterna conectados a la red. Su frecuencia es de varios kilohercios.

En nuestro análisis, las premisas de utilización pasan por utilizar la red eléctrica para realizar comunicaciones a distancias alrededor de 200 metros con frecuencias en torno a 135KHz. En este caso, la longitud de cable es inferior a 1/8 de la longitud de onda efectiva de la señal. Esto permite asumir que la degradación del canal debida a efectos de onda estacionaria será despreciable. Esta banda se encuentra lo suficientemente alejada de la señal de potencia eléctrica de la señal que porta la información. Dicha banda de frecuencias es objeto del estándar CENELEC EN 50065-1 que regula la utilización de la banda entre 3KHz y 148,5KHz en instalaciones eléctricas.

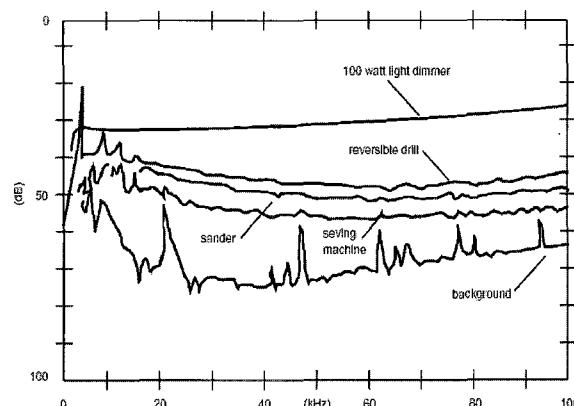


Figura. 2. Diagrama espectral de algunas fuentes comunes de ruido en powerline

Siguiendo con una filosofía de bajo coste se ha optado por la utilización de un modem FSK con 2 portadoras a 133,05KHz y 131,85KHz respectivamente [4]. El medio restringe las posibilidades a un modo half duplex con un régimen binario de 4800bps. A las citadas distancias también son notables los efectos por atenuación. La figura 3 muestra algunas atenuaciones típicas para distintos caminos de propagación encuadrados en un entorno doméstico. Se han podido realizar transmisiones exitosas a distancias de 200 metros, observándose atenuaciones medias en amplitud de 25dB. Es esta limitación la que obligará a asociar a los equipos esclavos en grupos, marcados por su cercanía entre sí.

Éstos a su vez, deben comunicarse con una consola central de supervisión. Para realizar tal comunicación, los maestros utilizan la red de área local. En la implementación actual, los maestros disponen de interfaz para conectarse a redes Ethernet.

Cada uno de estos grupos se caracteriza porque los esclavos que a él pertenecen, comparten el mismo bus de comunicación a través de la red eléctrica. Se denominará a cada uno de estos buses PL-Bus (PowerLine Bus).

Cada PL-Bus será monitorizado por un equipo maestro, que se encargará además de relacionar a todos los equipos esclavos con la consola central de monitorización. Ésta es la razón de la adopción del segundo medio de transmisión, la red de área local Ethernet, que será la encargada de comunicar a todos los maestros de PL-Bus con la consola central de monitorización.

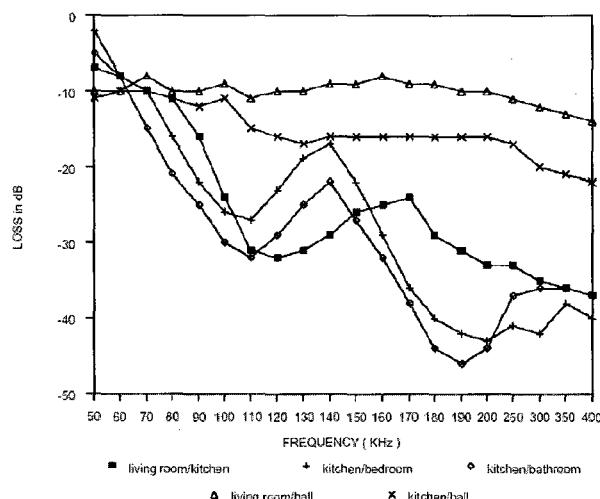


Figura 3. Atenuación de la red eléctrica por distintos caminos  
(Por Daniel Chaffanjon)

### 3.2. Red de área local Ethernet

Este tipo de redes son actualmente una constante en cualquier entorno empresarial. El caso más extendido en entornos laborales es aquel en el que los equipos de dicha red de área local tienen acceso a Internet a través de un gateway.

Este caso proporciona la posibilidad de instalar la consola de monitorización central en un punto remoto, permitiéndose así el control a distancia e incluso la posibilidad de controlar de forma centralizada varios edificios. La estructura de este esquema puede observarse en la figura 1.

Una de las ventajas de las tecnologías de Internet es la madurez de los métodos de protección de datos, que serán discutidos más adelante.

## 4. HARDWARE UTILIZADO

### 4.1. Esclavos

Cada esclavo dispone de una fuente de alimentación conmutada independiente de la alimentación del aparato al que se adhiere. Esto permite que el sistema de protección funcione incluso cuando el aparato no esté en funcionamiento. El sistema de comunicación por el que acceden al PL-Bus consta de una interfaz de aislamiento, que discrimina la banda de comunicaciones, una sección de amplificación para la señal de salida y un modem FSK.

El modem está dotado de un detector de portadora independiente del correlador, que es imprescindible para la detección de errores irrecuperables en el protocolo de comunicaciones. Dispone además de la posibilidad de sensar la señal de entrada tras un pequeño filtrado que facilita la medida. El módulo de control del esclavo es un microcontrolador de 8 bits. Éste es el encargado de implementar la máquina de estados que gobierna el sistema y de realizar la encriptación y desencriptación de la información. Este microcontrolador, gracias a un ADC es capaz de realizar medidas de amplitud de la señal de entrada al modem. Esta característica permite la ejecución de un algoritmo de autogestión, por el cual, esclavos y maestros son capaces de agruparse automáticamente basando esta asociación en el sensado de la amplitud de las portadoras de cada equipo. Así cada esclavo escoge al maestro al que recibe con mayor amplitud. Este algoritmo será discutido más adelante. Las medidas de amplitud, son además útiles para caracterizar el medio y poder planificar la inserción de maestros de PL-Bus en zonas donde la señal está muy debilitada.

### 4.2. Maestros de PL-Bus

Aunque el modelo en el que se trabaja actualmente implementa los equipos maestros mediante el mismo hardware que los esclavos y mediante una conexión RS-232 a un PC, se prevé el diseño de hardware específico para éstos.

La propuesta más interesante la compone el uso un microprocesador embebido, lo que conferiría a los maestros las altas prestaciones de cálculo necesarias para resolver los esquemas de autenticación necesarios en las comunicaciones sobre Internet.

La utilización de un microprocesador embebido, permite además la inclusión de un sistema operativo, lo que facilita notablemente la tarea de desarrollar software para los dispositivos maestros. Actualmente existen en el mercado varias alternativas, siendo las más estudiadas aquellas que ofrecen una solución basada en un único chip que implementa procesador, sistema de almacenamiento e interfaz Ethernet. En cuanto al sistema operativo, se ha escogido GNU/Linux como plataforma de desarrollo. Éste ofrece implementación completa de todos los protocolos de red, librerías de autenticación y aplicaciones en código abierto, permitiendo la adaptación a las necesidades específicas que se requieran. GNU/Linux dispone además de herramientas de desarrollo de gran calidad, lo que facilitará la generación del software necesario.

#### 4.3. Consola central de monitorización

Esta consola es propiamente un ordenador personal. Esto mejora la capacidad para interactuar con el usuario y permite implementar de forma cómoda todos los algoritmos de encriptación necesarios y la representación gráfica. Esta consola podrá ser local o remota, ya que la señalización de control puede enviarse de forma segura a través de Internet gracias a la utilización de SSL en las comunicaciones con los maestros.

### 5. PROTOCOLOS DE COMUNICACIÓN

#### 5.1. Comunicaciones en PL-Bus

Las comunicaciones utilizadas sobre el PL-Bus se basan en una trama de comunicaciones que engloba el nivel de enlace y red en una única trama semejante a HDLC (High Level Data Link Control). En la figura 4 puede observarse la estructura de campos de dicha trama.

La trama comienza con una secuencia de entrenamiento, que produce la puesta en marcha de los correladores en los receptores así como de los detectores de portadora.

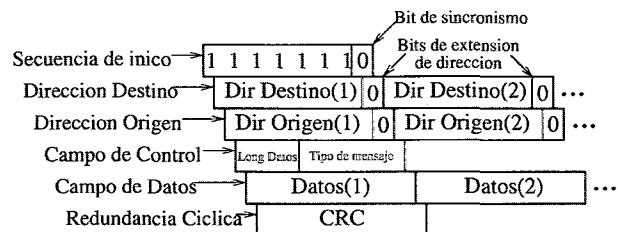


Figura. 4. Trama de PL-Bus.

Esta secuencia termina con un bit a nivel bajo que permite la sincronización de la trama en el flaco de bajada. Obsérvese que el detector de portadora es imprescindible para discernir las auténticas secuencias de entrenamiento, ya que éstas sólo aparecen cuando el detector de portadora indica ausencia de portadora.

La dirección es ampliable y crece en bloques de 7 bits, lo que proporciona una escalabilidad sólo limitada por los retardos de transmisión. El campo de control reserva 3 bits para indicar la longitud del campo de datos, de tal forma que se pueden seleccionar tamaños de campo de datos de entre 0 y 64 bytes. Así, la relación que indica el tamaño del campo de datos es:

$$\text{datos} = 2^{\text{long}-1}$$

salvo el caso del cero, que se computa directamente como la anulación del campo de datos. Los otros 5 bits del campo de control, indican el tipo de mensaje.

Entre estos tipos de mensajes se establecen primitivas para ceder el bus a una entidad superior con capacidad de autenticación por medio de la autoridad certificadora implementada en la consola central. Otras primitivas implementadas son el ping simple, o las primitivas que permiten la aceptación en el bus de un nuevo esclavo que será verificado por la autoridad certificadora.

El proceso de aceptación de un nuevo esclavo es completamente automático, no requiriéndose ningún tipo de conocimiento sobre la organización zonal de los PL-Bus. La última implementación realizada del sistema pretende además el autonegociado de agrupación de maestros y esclavos de forma automática basándose en la amplitud de las portadoras recibidas, para lo que se ha desarrollado un protocolo más. En la implementación actual, los datos contenidos en el paquete se transmiten encriptados.



El algoritmo de encriptación utilizado es TEA-Block [5], ya que la potencia de cálculo requerida por éste es asumible para un microcontrolador de 8 bits.

Todos los esclavos y maestros poseen una clave preprogramada en memoria, lo que evita tener que distribuir la primera clave para iniciar la comunicación.

Así, como se aprecia en la figura 5, esta clave es adecuada para cifrar un primer mensaje de intercambio de clave de sesión. Esta clave de sesión, distribuida en tiempo de ejecución es diferente para cada esclavo y tiene un tiempo de vida que depende del número de esclavos en el bus. La figura 5 muestra el proceso de intercambio de una nueva clave de sesión.

Se realizará una futura mejora de este sistema de encriptación, haciendo uso de una implementación de Serpent [6], una propuesta para AES (Advanced Encryption Standard). Serpent admite implementaciones que no requieren de elevadas capacidades de procesado, lo que hace factible su utilización con microcontroladores de 8 bits, a la vez que supera la protección contra ataques de fuerza bruta que otros métodos como triple-DES ofrecen. Otra posibilidad dependiente del desarrollo de precios de los ya citados procesadores embebidos, es la utilización de éstos tanto en dispositivos maestros como en esclavos, lo que mejoraría la seguridad y facilitaría el desarrollo del sistema debido a la unificación de las arquitecturas hardware. Una implementación de este tipo permitiría la utilización de SSL también en las comunicaciones a través de Powerline.

Para minimizar el ancho de banda consumido por las peticiones de respuesta del maestro, éstas son eliminadas en su mayoría. En lugar de solicitar a cada esclavo una respuesta cifrada con la clave de sesión, se realiza una asignación de turno por token.

El maestro indicará un número aleatorio que cada esclavo deberá responder encriptando con la clave de sesión en su turno de token. El maestro indica a cada esclavo cual es su turno enviándole la dirección del esclavo inmediatamente anterior a él en el bus.

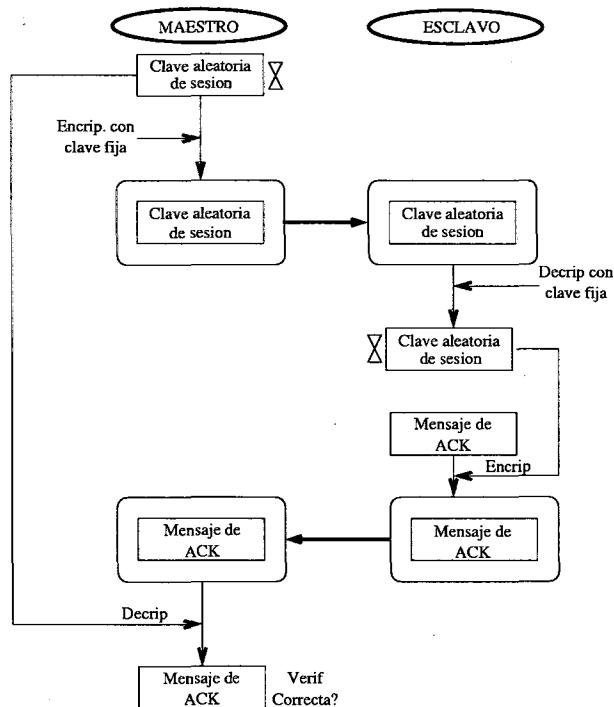


Figura. 5. Intercambio de clave de sesión en PL-Bus.

Así cada esclavo responde en su turno de token sin intervención del maestro, que sólo tiene como misión la verificación de los mensajes. Sólo se producen comunicaciones directas entre un esclavo y su maestro en situaciones de cambio de posición en el bus, intercambio de nueva clave de sesión o errores de autenticación. El maestro puede cambiar el orden de los esclavos después de un tiempo aleatoriamente generado. Todos estos procesos, dificultan enormemente la suplantación de un esclavo, elevando así el nivel de seguridad.

Como sistema de detección de errores, el sistema utiliza un campo de CRC de 1 byte. Para la generación de esta redundancia se utiliza el polinomio  $x^8 + x^2 + x + 1$

Esta redundancia permite detectar cualquier error en 1 bit, cualquier número impar de bits erróneos y cualquier salva de error de menos de 8 bits [7].

## 5.2. Autonegociado y autoconfiguración del sistema

La última implementación del sistema pretende dotar a éste de funcionalidad adicional que permita una administración del sistema más sencilla. Uno de los problemas más graves en PL-Bus es la situación dentro de la red eléctrica de maestros y esclavos y qué maestro es asignado a cada esclavo.

El pin marcado como RAI indica la entrada de la señal de la red eléctrica tras pasar la etapa de adaptación al medio.

Obsérvese que en el punto marcado como RxFO se obtiene la señal filtrada antes de ser amplificada. La señal extraída en este punto será sensada, de tal forma que se puede obtener la amplitud.

Este valor será indicativo de la potencia con la que se está recibiendo una fuente. Este proceso de sensado se realizará para cada una de las portadoras recibidas de los distintos maestros cercanos.

El conjunto se comporta de la siguiente manera:

Cada 20 segundos, los maestros inicián un periodo de aceptación de nuevos esclavos. Estos periodos están sincronizados entre los maestros a través de la Ethernet, de tal forma que no choquen en sus emisiones por la red eléctrica. El proceso

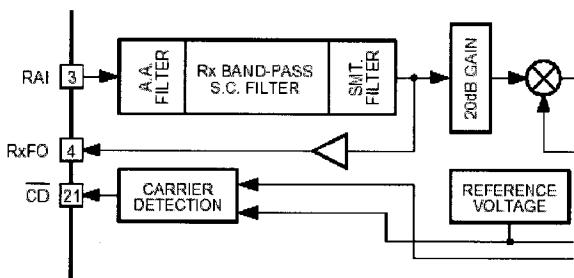


Figura. 6. Detalle de la estructura interna del modem utilizado.

consta de la emisión consecutiva de secuencias de entrenamiento por parte de cada maestro.

El esclavo interesado en registrarse en el sistema, recibirá todas las secuencias y decidirá cuál tiene mejor nivel de señal. Tras esto, existe un periodo de contienda, en el que todos los esclavos interesados en registrarse intentarán hacerlo.

Si existe choque, los maestros lo sabrán, pues detectarán las portadoras, pero la trama recibida estará corrupta (CRC no válido). En tal caso, los

maestros no responden y los esclavos darán por hecho que han chocado, esperando un tiempo aleatorio antes de reintentarlo.

El choque sólo sucede si la emisión de al menos 2 esclavos comienza en el mismo instante, puesto que si uno de ellos empieza a transmitir, el resto de esclavos detectan su portadora y no intentan emitir.

Las claves de todos los procesos concernientes a la detección de choques, se basan en la utilización del detector de portadora y la verificación del CRC de las tramas.

### 5.3. Comunicaciones en Ethernet

Ya que los mensajes se transmitirán por la misma red corporativa, éste es el punto más evidente para un ataque de seguridad por suplantación. Además, si la consola de monitorización es remota, los mensajes entre maestros y consola serán enrutados vía Internet, lo que provoca un nuevo problema de seguridad.

Para evitar las agresiones contra esta parte del sistema de seguridad, se ha optado por la implantación de SSL (Secure Socket Layer) para proporcionar el nivel de seguridad necesario en las transacciones.

SSL puede ser utilizado sobre cualquier protocolo de transporte, en nuestro caso TCP/IP. SSL confiere al canal de comunicaciones privacidad, autenticidad e integridad.

SSL utiliza certificados X.509, un esquema de clave pública/clave privada y una suite de cifrado que cuenta entre otros algoritmos con Triple-DES, MD5, RSA y RC4 [8].

SSL se alza como una de las soluciones integrales más extendidas en Internet, ya que autentica al cliente y al servidor, cifra el canal y asegura la integridad de los datos.

La implementación completa de SSL está disponible como código abierto bajo licencia GNU, lo que acorta enormemente el tiempo de desarrollo. Ya que actualmente SSL es el estándar de facto para seguridad en Internet, no se entra en más detalles sobre su implementación.

## 6.CONCLUSIONES

Se ha presentado un sistema de comunicación segura adaptable a prácticamente cualquier equipo conectado a la red eléctrica.

Una aplicación interesante es la utilización de este desarrollo en un sistema de control antirrobo. El sistema cumple con las premisas de fiabilidad, solidez ante ataques, bajo coste y fácil integración y administración.

La seguridad ha sido una prioridad y por ello se han implementado métodos de autenticación que permiten administrar el sistema desde Internet sin temer por la seguridad del sistema.

El futuro inmediato del desarrollo pasa por la mejora de la encriptación sobre powerline, la implementación de un hardware definitivo para los equipos maestros y la verificación del funcionamiento del algoritmo de autonegociado de maestro.

## REFERENCIAS

[1] RJ Anderson, SJ Bezuidenhoudt, "On the Reliability of Electronic Payment Systems", IEEE Transactions on Software Engineering v 22 no 5, May 1996.

[2] Phil Sutterlin, "A PowerLine Communication Tutorial - Challenges and Technologies", Proceedings of the 1998 International Symposium on Power-line

Communications and its Applications Soka University, Tokyo, March 1998.

[3] Roger M. VINES, Jel TRUSSEL, Louis J. GALE, "Noise on Residential power distribution circuits", IEEE Transactions on Electromagnetic Compatibility, Vol EMC-26, N°24, pp 161-168, November 1984.

[4] S.G. Wilson, «*Digital Modulation and Coding*», Prentice-Hall, 1996.

[5] David Wheeler and Roger Needham, "TEA, a Tiny Encryption Algorithm", Proceedings of the K. U. Leuven Workshop on Cryptographic Algorithms. Springer-Verlag, 1995.

[6] Anderson, Ross and Eli Biham, "Serpent: A Proposal for the Advanced Encryption Standard", NIST AES Proposal, June 1998.

[7] S.B. Wicker, "Error Control Systems for Digital Communication and Storage", Prentice-Hall, 1995.

[8] John Ousterhout, "Timing attacks on implementation of Diffie-Hellman, RSA, DSS, and other systems", Lecture Notes in Computer Science, vol.1109, Springer-Verlag, 1996.

## AUTOR



Fernando Méndez Rebollo nació en León en 1980. Actualmente es becario de investigación del Departamento de Electrónica de la Universidad de Alcalá (Madrid), donde realiza el proyecto fin de carrera de la titulación de Ingeniería de Telecomunicación en el campo de la visión artificial mediante arquitecturas hardware basadas en FPGA's.

Otro campo de interés es la robótica, habiendo formado parte del primer equipo español competidor en el campeonato europeo de robótica Eurobot2002.

En el campo de la domótica, colabora en un proyecto del Departamento de Electrónica de la Universidad de Alcalá para el desarrollo de una plataforma domótica integral para la interacción de un robot móvil de ayuda a la discapacidad con un entorno doméstico automatizado, utilizando tecnología Home-Plug.



# DYNAMIC TIME WARPING APPLIED TO DETECTION OF CONFUSABLE WORD PAIRS IN AUTOMATIC SPEECH RECOGNITION

*Jan Anguita Ortega, Francisco Javier Hernando Pericás*

{jan,javier}@talg.upc.es (TALP Research Center)  
Universitat Politècnica de Catalunya, Barcelona, Spain

## ABSTRACT

In this paper we present a method to predict if two words are likely to be confused by an Automatic Speech Recognition (ASR) system. This method is based on the classical Dynamic Time Warping (DTW) technique. This technique, which is usually used in ASR to measure the distance between two speech signals, is used here to calculate the distance between two words. With this distance the words are classified as confusable or not confusable using a threshold. We have tested the method in a classical false acceptance/false rejection framework and the Equal Error Rate (EER) was measured to be less than 3%.

## 1. INTRODUCTION

Using speech to communicate with the machines is a great improvement since it has a lot of advantages: speech is the natural way of communication for humans, speaking is faster than typing, while speaking hands and eyes are free for other tasks, some channels (Phone) are made for speech, etc. Unfortunately, although Automatic Speech Recognition (ASR) technology is already mature enough for some consumer products, in order to obtain acceptable performances the vocabulary and the structure of the sentences that the system is able to recognize must be limited. Even so, the systems make errors. These errors are sometimes caused by the words of the vocabulary that are phonetically similar. Therefore, the error rate can be reduced by designing the vocabulary with words as less similar as possible. The aim of this study is to design a method to predict if two words are likely to be confused by an ASR system. A tool

like this can help to design the vocabulary of a speech recognition system since it can warn the designer if two words are too similar, and sometimes one of them can be changed for another one with the same meaning but less similar.

For example, suppose that we want a speech recognition system for a mail application, where the user will be able to control all the options by his voice. First of all, we have to define the vocabulary that the system will be able to recognize. For example, suppose that we chose the following words:

- Supprimer: to delete a message.
- Imprimer: to print a message.
- Envoyer: to send a message.
- Lire: to read a message.
- Ecrire: to write a message.
- Suivant: to go to the next message.
- Précédent: to go to the previous message.

We have chosen French words because in this project we have worked in French. Once the vocabulary is chosen we have to define the syntax, i.e. the structure of the sentences the system will recognize. In this case isolated words is enough. This means that the user can only say one word each time, preceded and followed by a silence. He cannot say supprimer suivant for example. This application may seem very simple but, as we have already said, we must do these simplifications because, nowadays, the ASR technology is not good enough to let the user say what he wants and how he wants. Even if this system is very simple it will make errors. Sometimes the user will say one word and the system will recognize another one. This is very dangerous because, imagine that the



users says lire and the system understands supprimer. The message is lost forever. Therefore, these errors must be reduced as much as possible. Imagine that, when the system is already in use, we realize that it often confuses two words, for example supprimer and imprimer. If we would have known this when we were designing the vocabulary of the system, we could have changed one word by a synonym, for example supprimer by effacer. In this case the application would have been exactly the same and we would have avoided the confusions between supprimer and imprimer.

In this project we have developed a method to predict if two words are likely to be confused by an ASR system if they are both in its vocabulary. We will use the terms confusable and not confusable to refer to the pairs of words that are often confused by an ASR system and the ones that are not confused respectively. The developed method is based on a technique called Dynamic Time Warping (DTW) [1], which is usually used in ASR to measure the distance between two speech signals. Here, we use it to calculate a measure of distance between the phonetic transcriptions of the words to compare [2,3] and, after, we classify the pair of words as confusable or not confusable using a threshold. This method can help to design the vocabulary of an ASR system, because it will warn the designer if he chose confusable words so, he can change them if it is possible. The principle of this measure is to do an alignment between the phonetic transcriptions of the two words and calculate the distance as the sum of the distance between the phones that are in correspondence according to the alignment. Although the developed method can be used in any language, the used language in this work is French.

The organization of this paper is as follows. In section 2 we explain the DTW distance. The first step of this technique is to align the phonetic transcriptions to compare. Therefore, first of all we explain the notation used to describe an alignment. After, the formulation of DTW and its algorithm are presented. In section 3 we present the distance measure between phones that we have used to calculate the distance between phonetic transcriptions. In section 4 we describe the experiments performed to test the method and the

obtained results. Finally, in section 5 we present the conclusions of this work.

## 2. DYNAMIC TIME WARPING

### 2.1. Alignment between Phonetic Transcriptions

Let  $W_1 = \{p_{1i}\}$  and  $W_2 = \{p_{2j}\}$ , with  $i=1, \dots, I$  and  $j=1, \dots, J$ , be the phonetic transcriptions of the two words to compare. The values  $I$  and  $J$  are the lengths of the phonetic transcriptions and  $p_{1i}$  and  $p_{2j}$  are their phones. Let us consider an  $i$ - $j$  grid, shown in Fig. 1, where  $W_1$  and  $W_2$  are developed along the  $i$ -axis and the  $j$ -axis respectively. A path through the grid is written as  $F = \{c(1), c(2), \dots, c(K)\}$ , and it represents an alignment between the two transcriptions. The generalised element of the path is  $c(k)$  and it consists of a pair of coordinates in the  $i$  and  $j$  directions. The  $i$  and  $j$  coordinates of the  $k$ th path element are  $i(k)$  and  $j(k)$  respectively.

$$c(k) = (i(k), j(k)) \quad (1)$$

The path  $F$  fulfills the following conditions [1]:

1) Monotonic conditions:

$$i(k-1) \leq i(k) \text{ and } j(k-1) \leq j(k) \quad (2)$$

2) Continuity conditions:

$$|i(k) - i(k-1)| \leq 1 \text{ and } |j(k) - j(k-1)| \leq 1 \quad (3)$$

3) Boundary conditions:

$$i(K) = I \text{ and } j(K) = J \quad (4)$$

The alignment is defined by the path  $F$  as follows:

- if  $i(k) = i(k-1) + 1$  and  $j(k) = j(k-1) + 1$  then  $p_{1i(k)}$  and  $p_{2j(k)}$  are aligned.

- if  $i(k) = i(k-1) + 1$  and  $j(k) = j(k-1)$  then  $p_{1i(k)}$  is aligned with the null character (symbol of an insertion or an omission)

- if  $i(k) = i(k-1)$  and  $j(k) = j(k-1) + 1$  then  $p_{2j(k)}$  is aligned with the null character.

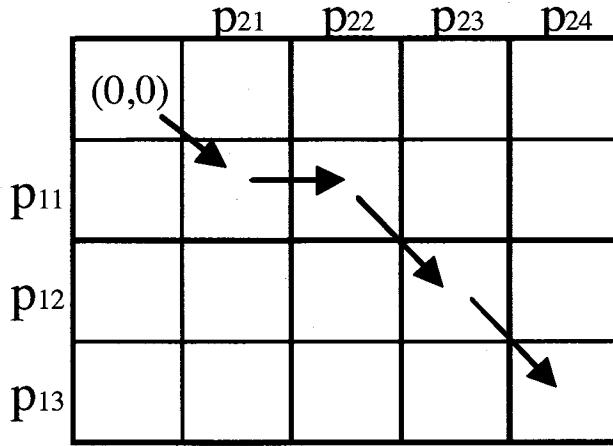


Fig. 1. Example of a path F in the grid, and the steps c(k).

For example, the alignment associated to the path of Fig. 1 is the following one:

$$\begin{array}{cccc} p_{21} & p_{22} & p_{23} & p_{24} \\ p_{11} & - & p_{12} & p_{13} \end{array}$$

## 2.2. DISTANCE BETWEEN TWO PHONETIC TRANSCRIPTIONS

The proposed application of this work is to predict if two words are likely to be confused by an ASR system, i.e., if they are confusable or not. In order to do this, a distance is calculated between the two words and, if the distance is lower than a threshold, the word pair is considered confusable:

$$\begin{cases} \text{if } D_{DTW}(W_1, W_2) \leq \text{Threshold} \Rightarrow \text{Confusable} \\ \text{if } D_{DTW}(W_1, W_2) > \text{Threshold} \Rightarrow \text{Not Confusable} \end{cases}$$

where  $D_{DTW}(W_1, W_2)$  is a distance between the phonetic transcriptions of the words  $W_1$  and  $W_2$ . Dynamic Time Warping (DTW) [1] is a technique that was used in speech recognition to calculate a distance measure between two speech signals. In this work we apply this technique to calculate the distance between the phonetic transcriptions of two words:

(5)

$$D_{DTW}(W_1, W_2) = \min_F \left[ \frac{\sum_{k=1}^K d(c(k))w(k)}{\sum_{k=1}^K w(k)} \right]$$

$$\begin{aligned} c(1) &= (1,1) \\ c(2) &= (1,2) \\ c(3) &= (2,3) \\ c(4) &= (3,4) \end{aligned}$$

where  $w(k)$  is a weighting function introduced to normalise by the path length and  $d(c(k))$  is a distance measure between the elements that are aligned according to  $c(k)$ . For example  $d(c(1))$  in Fig. 1 is the distance between the phones  $p_{21}$  and  $p_{11}$ . How to obtain the distance  $d(c(k))$  is explained in the following section. In this work we have used the following weighting function [1]:

$$w(k) = i(k) - i(k-1) + j(k) - j(k-1) \quad (6)$$

This implies that:

$$\sum_{k=1}^K w(k) = I + J$$

Then, the denominator of (5) is constant and, therefore, independent of the path F. The DTW distance is the minimum weighted summation of the distances between the aligned phones, for all the possible alignments between the phonetic transcriptions of the words. Since the denominator  $N(w) = I + J$  is a constant, in order to solve (5) we only have to minimize the numerator and after, divide by  $I + J$ . Recall that the points that can lead to the point  $(i(k), j(k))$  are  $(i(k)-1, j(k)), (i(k)-1, j(k)-1)$  and  $(i(k), j(k)-1)$  (monotonic and continuity conditions). Therefore, the weights associated to each step are (using (6)):



$$\begin{aligned}
(i(k)-1, j(k)) \rightarrow (i(k), j(k)): w(k) &= i(k)-i(k-1)+j(k)- \\
j(k-1) &= i(k)-(i(k)-1)+j(k)-j(k) = 1 \\
(i(k)-1, j(k)-1) \rightarrow (i(k), j(k)): w(k) &= i(k)-i(k-1)+j(k)- \\
j(k-1) &= i(k)-(i(k)-1)+j(k)-(j(k)-1) = 2 \\
(i(k)-1, j(k)) \rightarrow (i(k), j(k)): w(k) &= i(k)-i(k-1)+j(k)- \\
j(k-1) &= i(k)-i(k)+j(k)-(j(k)-1) = 1
\end{aligned}$$

The solution, i.e DDTW(W<sub>1</sub>,W<sub>2</sub>), can be found using the variable s(i,j) defined as follows:

(8)

$$s(i, j) = \min \begin{cases} s(i-1, j) + d(i, j) \\ s(i-1, j-1) + 2d(i, j) \\ s(i, j-1) + d(i, j) \end{cases}$$

where s(i,j) is the accumulated distance of the optimal path that goes from the point (0,0) to the point (i,j). Therefore,

(9)

$$D_{DTW}(W_1, W_2) = \frac{s(I, J)}{I + J}$$

When all the values s(i,j) have been calculated over all i,j DDTW(W<sub>1</sub>,W<sub>2</sub>) can be calculated using (9). Below we present the algorithm to find the solution[4].

$$s(0,0) = 0$$

for(j=1;j J;j++) s(0,j)=

for(i=1;i I;i++){

$$s(i,0) =$$

for(i=1;i I;i++){

$$s(i, j) = \min \begin{cases} s(i-1, j) + d(i, j) \\ s(i-1, j-1) + 2d(i, j) \\ s(i, j-1) + d(i, j) \end{cases}$$

}

$$D_{DTW}(W_1, W_2) = \frac{s(I, J)}{I + J}$$

### 3.DISTANCE BETWEEN PHONES

In the previous section we have explained the DTW technique, which is used to calculate a distance measure between two phonetic transcriptions. Since this technique depends on a distance between the phones of the phonetic transcriptions, in this section we explain how to obtain this distance. In modern ASR systems the acoustic units are usually modeled

by Hidden Markov Models (HMM) [5]. Therefore, it is possible to obtain a distance measure between two phones by calculating the distance between their HMMs. In this paper we propose the following distance between two HMMs:

(10)

$$d_{HMM}(p_1, p_2) = \begin{cases} \sum_Q P(Q) \left[ \frac{1}{L} \sum_{i=1}^L D_N(N_{q_{1i}}, N_{q_{2i}}) \right] & \text{if } p_1 \neq p_2 \\ 0 & \text{if } p_1 = p_2 \end{cases}$$

where Q is an alignment between the states of the HMMs of the phones p<sub>1</sub> and p<sub>2</sub>, P(Q) is the probability of Q, L is the length of the alignment, q<sub>1i</sub> and q<sub>2i</sub> are states of the models that are aligned according to Q, N<sub>q<sub>1i</sub></sub> and N<sub>q<sub>2i</sub></sub> are the Gaussian distributions associated to the states q<sub>1i</sub> and q<sub>2i</sub>, and D<sub>N</sub>(·) is a measure of distance between the two Gaussian distributions. The numerator is a weighted sum of the average distance between the Gaussians of the aligned states for each alignment Q. In [6], this average distance between Gaussians is calculated for each Q and the minimum one is chosen. On the other hand, we sum all these average Gaussian distances weighted by the probability of the alignment. Since only a subset of the possible alignments is used, the denominator is introduced in order to normalise by the probability of the subset of alignments. In this work, we used the alignments associated to the possible paths in a grid of dimension M<sub>1</sub> x M<sub>2</sub>, where M<sub>1</sub> and M<sub>2</sub> are the number of states of the models. Fig. 2 shows an example with M<sub>1</sub> = M<sub>2</sub> = 3. This subset avoids the alignments where there are loops in states of the two models at the same time.

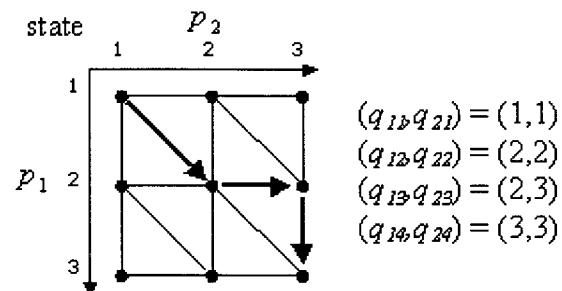


Fig. 2: Subset of alignments used to calculate the inter-HMM distance. The bold line shows one of these alignments. The values of q<sub>1i</sub> and q<sub>2i</sub> are the aligned states according to the path in bold.

The models used to obtain a distance value between the phones with the proposed measure have one Gaussian per state. This does not imply that the real ASR systems must have one Gaussian per state. We considered several monomodal Gaussian distances such as Euclidean, Mahalanobis, Kullback-Leibler, Bhattacharyya and Jeffreys-Matusita [7,8]:

Euclidean distance:

$$D_{EUC}(N_1, N_2) = (\mu_2 - \mu_1)^T (\mu_2 - \mu_1) \quad (11)$$

Bhattacharyya distance:

(12)

$$\begin{aligned} D_{BHA}(N_1, N_2) &= (1/8) (\mu_2 - \mu_1)^T ((\Sigma_1 + \Sigma_2)/2)^{-1} \\ &+ (\mu_2 - \mu_1) + 1/2 \log (I(\Sigma_1 + \Sigma_2)/2I)/(I(\Sigma_1)I(\Sigma_2))^{1/2} \end{aligned}$$

Jeffreys-Matusita distance:

(13)

$$D_{JM}(N_1, N_2) = \sqrt{2(1 - \exp(D_{BHA}(N_1, N_2)))^{1/2}}$$

Kullback-Leibler distance:

(14)

$$\begin{aligned} D_{KL}(N_1, N_2) &= 1/2 (\mu_2 - \mu_1)^T (1/\Sigma_1 + 1/\Sigma_2) (\mu_2 - \mu_1) \\ &+ 1/2 \text{tr}((1/\Sigma_1)\Sigma_2 + (1/\Sigma_2)\Sigma_1 - 2I) \end{aligned}$$

Mahalanobis distance:

(15)

$$D_{MAH}(N_1, N_2) = (\mu_2 - \mu_1)^T (\Sigma_1 \Sigma_2)^{-1} (\mu_2 - \mu_1)$$

where  $\mu_i$  and  $\Sigma_i$  are the mean vector and the covariance matrix of the Gaussian  $N_i$  respectively.

This distance has to be extended to cover pairs consisting of a phone and the null character, which corresponds to the operation of insertion or omission. The extended inter-phone distance, which

is the one used to calculate the DTW distance measure is:

$$(16)$$

$$d(c(k)) = \begin{cases} d_- & \text{if } (i(k) = i(k-1) \text{ or } j(k) = j(k-1)) \\ d_{HMM}(p_{i(k)}, p_{j(k)}) & \text{otherwise} \end{cases}$$

where  $d_-$  is the distance between a phone and the null character. This value was set at the arithmetic mean of the distances between all the phones:

(17)

$$d_- = \frac{1}{P^2} \sum_{i=1}^P \sum_{j=1}^P d_{HMM}(p_i, p_j)$$

where  $P$  is the total number of phones.

## 4. EXPERIMENTS AND RESULTS

### 4.1. Experimental Setup

In order to test our method we need to determine which pairs of words are usually confused by ASR systems to compare them with the prediction of our method. We constructed two kinds of ASR systems: one to detect the confusable word pairs, and the other to detect the not confusable word pairs.

-NCD Systems (No Confusability Detection): 223 systems, each one with only one word in its vocabulary and a garbage model to reject out-of-vocabulary data. Each system was tested with the 223 words.

-CD System (Confusability Detection): One system with 841 words and a garbage model, tested with the 841 words.

If one of the NCD systems, with only the word A in its vocabulary, is tested with another word B and they are never confused, it means that they are very different and, therefore, they are not confusable. On the other hand, if they are sometimes confused, it only means that B is more similar to A than to the garbage model, not necessarily that A and B are similar. Therefore, with this kind of systems we can only determine if two words are not confusable in general.



If we test the CD system with several pronunciations of a word A, and a word B is never recognized, we cannot say that A and B are not confusable, we can only say that A is more similar to some of the other words of the vocabulary than to B. On the other hand, if they are sometimes confused, we can assure that they are quite confusable. Therefore, with this system we can detect confusable word pairs.

The vocabulary of the CD and NCD systems consisted of French isolated words such as numbers, cities, commands, etc. Each word was pronounced by 700 speakers in average. The speech signal was sampled at 8 kHz and parameterized using MFCCs. The feature vectors consisted of 27 coefficients: the frame energy, 8 MFCCs, and the first and second time derivatives. The models of the words were constructed by concatenating context dependent HMMs of the phones with one Gaussian per state. By testing these systems the following three groups of word pairs are obtained:

"Low Probability of Confusion (LPC): 21506 word pairs which were never confused when the NCD systems were tested.

"Medium Probability of Confusion (MPC): 150 word pairs which had a confusion rate lower than 5% and higher than 0% when the CD system was tested.

"High Probability of Confusion (HPC): 189 word pairs which had a confusion rate higher than 5% when the CD system was tested.

We consider a False Rejection to classify as confusable a LPC word pair, and a False Acceptance to classify as not confusable a HPC word pair. The MPC word pairs were not taken into account in the evaluation because we considered that is not a severe error neither to classify them as confusable nor as not confusable.

The HMMs used to calculate the inter-phone distances are not the models used in recognition. In the first case we used models without context with 3 states and 1 Gaussian per state.

#### 4.2. RESULTS

In order to test our method we measure the False Rejection Rate (FRR) and the False

Acceptance Rate (FAR). The FRR is the error rate when classifying the pairs of words that are not confusable, i.e., the percentage of pairs words belonging to the group LPC classified as confusable. The FAR is the error rate when classifying the pairs of words that are confusable. That is to say, the percentage of word pairs belonging to the group HPC classified as not confusable. Our objective is to minimize both FRR and FAR, and both values depend on the chosen threshold. If we decrease the threshold the FRR decreases and the FAR increases and vice versa. In order to compare the different methods with only one value we use the Equal Error Rate (EER). The EER is the False Acceptance Rate and the False Rejection Rate obtained with the threshold that makes them equal as shown in Fig. 3.

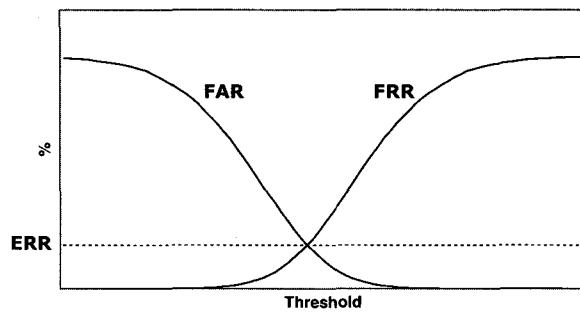


Fig. 3: FRR and FAR in terms of the threshold. The EER is the point where the two lines intersect.

Table 1 shows the EER for each Gaussian distance in (10). We can see that the better results are obtained when using the Mahalanobis Gaussian distance to calculate the distances between the phones. With the Euclidean and the Kullback-Leibler distances also low error rates are obtained. On the other hand, the Battacharyya and the Jeffreys-Matusita distances give high EERs and therefore are not useful to our purpose. We can conclude that the developed method can be used useful to predict if two words are likely to be confused by an ASR system because we have obtained an EER of 2.6%.

#### 5.CONCLUSIONS

In this paper we have presented a method to predict if two words will be confused by an ASR system. This method is based on the classical DTW technique, which is used to calculate a distance

between two phonetic transcriptions. We also have described how to obtain the data to test. We have tested the method in a classical false acceptance/false rejection framework and the EER was measured to be less than 3%.

	<b>DTW</b>
<b>EUC</b>	3,1%
<b>KL</b>	3,2%
<b>MAH</b>	2,6%
<b>JM</b>	7,5%
<b>BHA</b>	8,9%

Table 1: EER obtained with each Gaussian distance in (10)

## 6.ACKNOWLEDGMENTS

This paper reports the work performed during an internship at Telisma R&D, Lannion, France. The authors would like to thank all the members of Telisma specially Stephane Peillon and Alexandre Bramoulle.

## 7.REFERENCES

- [1] Hiroaki Sakoe and Seibi Chiba, "Dynamic Programming algorithm optimization for spoken word recognition". In IEEE Trans. on ASSP, vol. ASSP-26, N°1, 1978.
- [2] Beng T. Tan, Yong Gu, Trevor Thomas, "Word Confusability Measures for Vocabulary Selection in Speech Recognition", Proceedings of the ASRU, Keystone, December 1999.
- [3] Sandrine Pouységur, "Etude du taux de confusion de mots pour la reconnaissance de mots isolés". 4e Rencontres jeunes chercheurs en parole, 2001.
- [4] Calliope, La Parole et son Traitement Automatique, Masson, Paris, 1989.
- [5] Rabiner L. "A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition". Proceedings of the IEEE, 77(2):257-286, Feb. 1989.

[6] Bahlmann C. and Burkhardt H. "Measuring HMM similarity with the Bayes probability of error and its application to online handwriting recognition", Proceedings of the ICDAR, pp 406-411, 2001.

[7] Basseville M. "Distance Measures for Signal Processing and Pattern Recognition", Signal Processing, Vol. 18(4), pp. 349-369, 1989.

[8] Sooful J., Botha E. "An Acoustic Distance Measure for Automatic Cross-language Phoneme Mapping". Proceedings of the PRASA, 2001.

## 8.AUTHORS



*Jan Anguita Ortega received his degree in telecommunication engineering and the European master in speech and language from the Universitat Politècnica de Catalunya (UPC) in 2003. He is currently a PhD student in the department of signal theory and communications at UPC. His interests are speech processing, robust speech and speaker recognition and speech perception.*



*Francisco Javier Hernando Pericás received his degree and PhD degree in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC) in 1988 and 1993 respectively. He is presently an associate professor in the department of signal theory and communications at UPC. His main interests are speech processing and robust speech and speaker recognition. He has published papers in international journals and conferences and he is member of both national and international associations. He works in projects supported by Spanish and European institutions.*



# VIDEO STREAMING WITH QUALITY ADAPTATION USING COLLABORATIVE ACTIVE GRID NETWORKS

Juan Carlos Peláez López,

Jorge Mata Díaz

Departamento de Ingeniería Telemática, Universitat Politècnica de Catalunya

Jorge Girona 1-3, Campus Nord, Mód. C3, UPC. 08034 Barcelona

Teléfono: +34-934016028 Fax: +34-934011058

E-mail: {jpelaez,jmata}@entel.upc.es

## ABSTRACT

Due to the services and demands of the end users, Distributed Computing (Grid Technology, Web Services, and Peer-to-Peer) has been developed rapidly in the last years. The convergence of these architectures has been possible using mechanisms such as Collaborative work and Resources Sharing. Grid computing is a platform to enable flexible, secure, controlled, scalable, ubiquitous and heterogeneous services. On the other hand, Video Streaming applications demand a greater deployment over connected Internet users. The present work uses the Active Grid technology as a fundamental platform to give a solution of multimedia content recovery. This solution takes into account the following key concepts: collaborative work, multi-source recovery and adaptive quality. A new architecture is designed to deliver video content over a Grid Network. The active and passive roles of the nodes are important to guarantee a high quality and efficiency for the video streaming system. The active sender nodes are the content suppliers, while the passive sender nodes will perform the backup functions, based on global resource control policies. The aim of the backup nodes is minimize the time to restore the system in case of failures. In this way, all participant peers work in a collaborative manner following a multi-source recovery scheme.

Furthermore, Video Layered Encoding is used to manage the video data in a high scalable way, dividing the video in multiple layers. This video codification scheme enables the quality adaptation according to the availability of system resources. In addition, a buffer by sender peer and by layer is needed for an effective control of the video retrieve. The QoS will fit considering the state of each buffer and the measurement tools provide by the Active Grid on the network nodes.

**Keywords:** Peer-to-Peer Grid Architecture, Services for Active Grids, Streaming Media, Layered Coding, Quality Adaptation, Collaborative Work.

## 1. INTRODUCTION

The users of Internet demand more and better multimedia services. The number of users who request multimedia information in Internet is increasing. This rising demand involves two challenges: 1) Internet offers only best-effort services, whereby intelligence techniques are required to deal with the user requirements of multimedia applications; 2) Streaming multimedia demands higher and ensured resources availability in the network, in the content supplier and distributor, and moreover higher capabilities in the receiver user system. The challenges mentioned have been reached by the distributed computation evolution and development. This evolution has changed the centralized models for the distributed architectures.

Grid computing [1] and Peer-to-Peer models [2] arise from the last evolution in the development chain of the distributed computation and communication. Grid technology creates secure and efficient Virtual Organizations (VO). Grid is the novel enabling technology to transparently access computing and storage resources anywhere, anytime and with guaranteed QoS. The new application field for the Grid Computing is the multimedia services management. These services are actually the more requested by Internet users. File Sharing P2P applications emerge as the best alternative for new Internet services. File sharing systems, such as Napster [2], is an example of the large current acceptance of P2P architecture. Hence, the contribution of P2P models is the flexible and

scalable support of VOs with a fast discovery of accessible services.

A general classification of multimedia applications is needed to clarify the type of multimedia services addressed in this work. Two groups of multimedia applications are: 1) Real-Time: video-on-live, or videoconference. They receive and visualize the video at the same time the capture is accomplished. 2) Not-real-Time: Video-on-Demand (VoD). Here, a user decides the moment that wishes to receive and visualize the video; moreover, the video is pre-captured and pre-stored. Streaming video is classified within the second group. The current file sharing services allow to retrieve a multimedia content in the download-playback mode. The present work is focused in video streaming services which follows a Playback-during-download mode.

Different architectures have been proposed to support decentralized media streaming services. Caching Server [3] was the first attempt to decentralize the content. It consists in storage the content among points that have already visualized the video. Nevertheless, caching technique transfers the bottleneck to several points but does not eliminate them completely. The Content Distribution Networks (CDN) [4] provide multimedia content recovery with an acceptable degree of scalability. But the high implementation cost of powerful equipment servers deployed in different geographic points limits this option. Multicast technology emerged under the idea to minimize the number of packets transmitted through the network. Real-time multimedia transmissions have been addressed by both IP multicast or Network Level Multicast [6] and Application Level Multicast or also called Overlay Multicast [5]. But multicast schemes are not able to support not-real-time solutions. Nowadays, Grid architectures are the emergent solutions for real-time and non-real-time multimedia services. Grid architectures provide a fast deployment of these services by means of interoperability, the easy management of specific QoS policies, and the organization scalability.

Home entertainment and shopping, digital video library, movie-on-demand, distance learning, medical information services, interactive games,

news/services-on-demand and emergency and disasters management are some of the most outstanding applications of a media streaming service. Push and pull services are very attractive in the current market of Internet users. For instance, a push service (initiated by the server) could be the transmission of a descriptive video of the tourist site to some user with a mobile device triggered by the user location and current time. On the other hand, a pull service (initiated by client) could be the multimedia message transmission from an isolated site where communicating is only possible by wireless networks.

Parallel technologies like third generation mobile communications, wireless Internet, video compression, among others, provide an improved structure of communications for the media streaming services. A controlled Active Grid architecture is the best alternative for supporting VoD applications. Moreover, a global efficiency of the V.O. will be obtained applying collaborative work concepts over a common interaction environment. This efficiency is measured in terms of used bandwidth, experimented latency and losses. The Active Grid QoS framework [7] is based on OGSA (Open Grid Service Architecture), where the Grid Security Infrastructure (GSI) protocol and Global Resource Control techniques are included.

In conclusion, an efficient and scalable VoD Streaming Service using the P2P Grid technology is introduced in this work. The rest of the paper is organized as follows. The problem definition is addressed in Section 2. The designed model and architecture is explained in Section 3. Section 4 concludes the paper and describes future works.

## 2. PROBLEM DEFINITION

There are two ways to recover a video file over Internet. The first one is called ‘Download-and-play’ mode, and the second one is the ‘Play-during-download’ mode. Videostreaming is a service that demands the application of the second style. The users of video streaming expect to watch the video few seconds after their request. Likewise, the users employ different terminals to access to multimedia services networks. Hence, the ubiquity in user terminals must be considered. Also, the

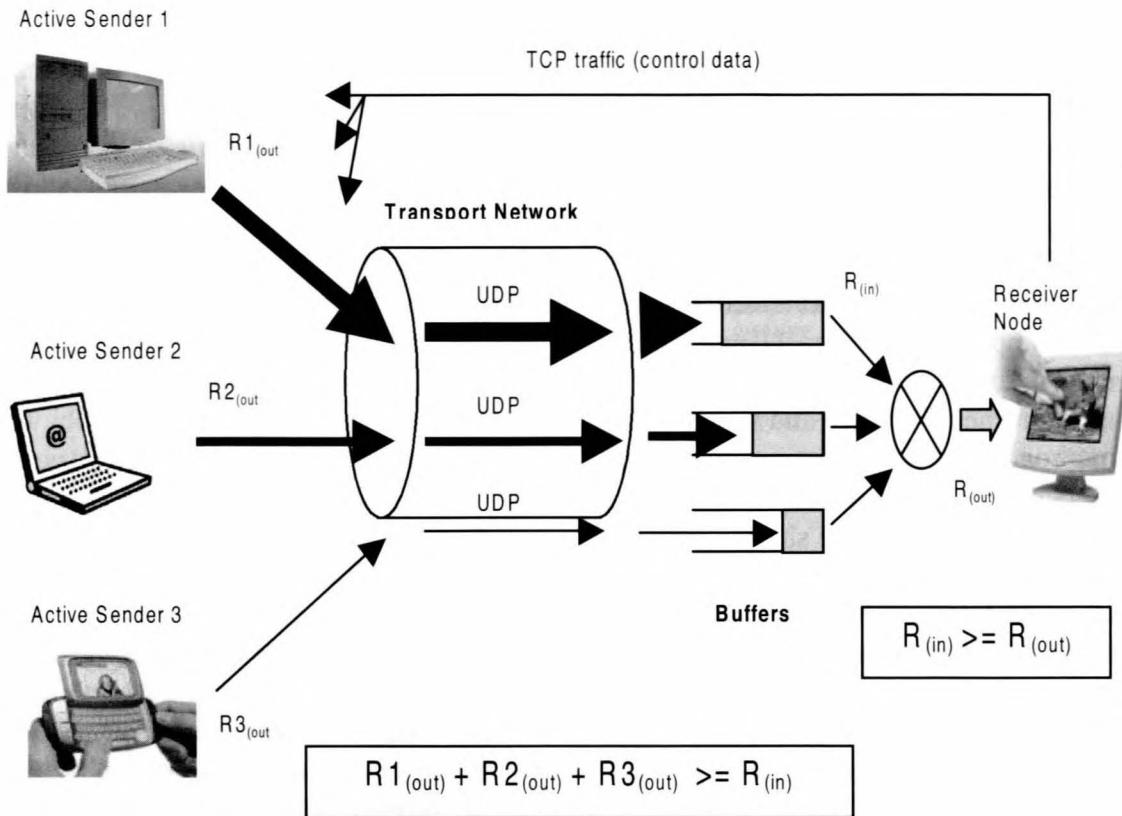


Fig. 1: Multi-source recovery

network conditions to transmit the video over the networks are variable and unpredictable. These conditions should be controlled by the active Grid in order to guarantee the QoS in the audiovisual services.

The video recovery over Grid networks in a play-during-download mode includes the following processes and characteristics: 1) search of nodes with the stored video; 2) selection of the best content suppliers; 3) congestion control mechanism in the active network nodes; 4) buffering techniques in the receiver side; 5) adaptation of the video playback rate in the receiver; 6) retardando nodo emisor, as nodos y la ruta a seguir hasta el nodo receptor, midiendo el ancho de banda disponible y ofrecido) system recovery after a failure; and, 7) security mechanisms within the V.O.

The Active Grid architecture solves the search, selection and classification of nodes that store the video content required. The selection of the best content suppliers will be the results of the analyzing

resources of each candidate node and transport network resources. The delay time and packet lost probability are also considered to select the best peers in the admission control function of the video services. This selection will be upgrade each time the network conditions change. A Grid QoS mechanism is required to evaluate the changes in the network conditions. A congestion situation in the network will imply to fire a reconfiguration of the video streaming recovery system. This mechanism is based on global resource control policies of the Grid Network. Buffering techniques in the receiver side are used to manage the playout video during an irregular system condition. The content in the buffers should be controlled every time, adapting the quality of the video.

Video quality adaptation will be supported by the encoding technique called ‘Video Layered Encoding’. There will be multiples layers, since the base layer to enhanced layers. This scheme of layered encoding will allow adapting the quality video in a gradual manner. A buffer per layer and

per peer is proposed in this work. The model of multiples buffers is more efficient and more reliable when a system failure or a network congestion happen. Quality Adjustment Mechanisms are supported by the Active Grid nodes, which control the dynamic buffering.

In conclusion, the problem to deal is the video streaming retrieval. This recovery must be reliable, efficient, scalable, economic and adaptive. Dynamic and restricted environments like Internet are considered as part of the problem. The system constraints are: 1) dynamic behaviour and cooperative level of the V.O. nodes; 2) heterogeneity of nodes; 3) changes in the network conditions; 4) video application non tolerant to the delay; and 5) reception constrains in the receiver peer side.

### 3. ACTIVE P2P GRID MODEL

#### A. Principles of the proposed model

**1) Collaborative work and Load Balancing:** To apply the foundations of collaborative work means to obtain the results of a process with a smaller cost. Optimizing the use of the available

resources and achieving an efficient distribution of work between the nodes of a network. The Active Grid network facilitates the application of this concept, using its global knowledge of the system. An initial load balance is performed for the active grid service. The resources of each candidate node are considered in the streaming distribution process. A load balancing efficient mechanism will help to reach suitable results derived from a distributive fairness. A initial media content distribution is performed by the Grid system, based on the popularity of the video and the hierarchical organization of the nodes [7]. Likewise, Video sequences are distributed over the grid network taking into account the CPU, Memory, disk I/O workload, and disk capacity of the active grid nodes.

#### 2) Multi-source recovery of Video Streaming:

The global behaviour of the system depends on the individual performance of each one of its components. In the case of video streaming services will be much more efficient in time and cost when a recovery from multiple sources is carried out instead of a single recovery. Efficient exploitation of resources, scalability for media services and

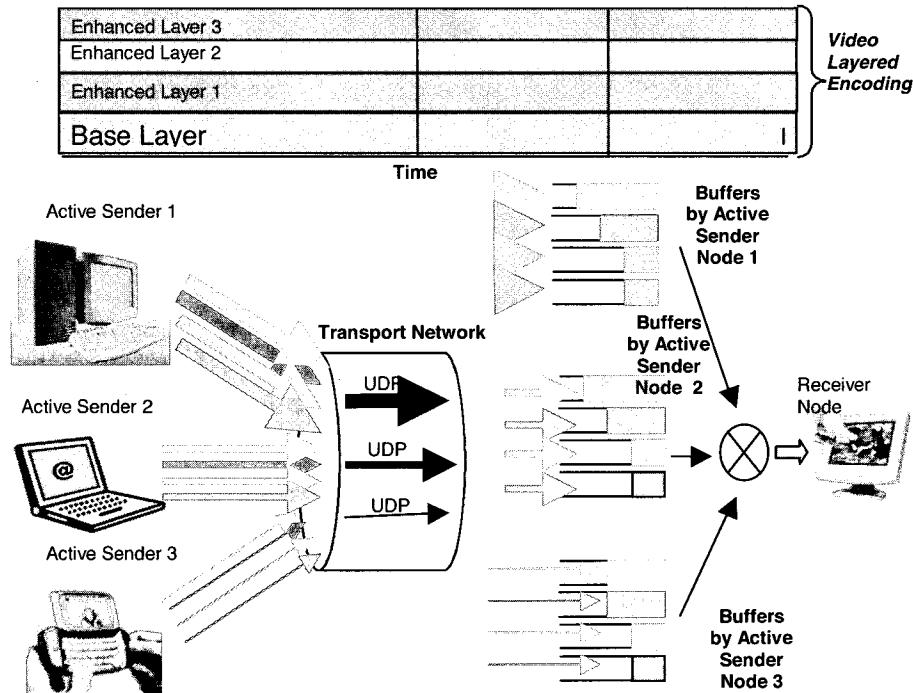


Fig. 2: Buffer Distribution by peer and by layer.



protection to error-prone wireless networks are achieved by this multi-source recovery scheme (see Fig. 1).

**3) Adaptive Quality and Buffering Techniques:** The dynamism of the P2P Grid network is motivated by the variability of the peer connection time. On the other hand, the conditions of the network change too. This variation is produced by the congestion, resources reservation, management policies, etc. Hence, it is required the implementation of a Grid Computing QoS framework, where the video layered encoding joined to TCP-friendly congestion control protocols are used. Moreover, the use of buffers in the receiver side is requested. The buffers smooth the variations of the arrival rate and allow to maintain a constant presentation rate. A control mechanism of the buffers state in the receiver has high-priority in our proposal. Multiples buffers are employed, one buffer per peer and per layer. This scalable model of multiples buffers is more efficient and more reliable during a system failure and network congestion situations.

#### B. Components of the proposed model

**1) Receiver Node:** This node will be the main participant in the video recovery system. It will initiate the request of a multimedia content and interacts with Quality Adjustment Mechanism of the Active Grid in order to redistribute the work between the active nodes, and acts over the active networks nodes to maintain the QoS. This work will be allocated in a fair and weighted sense, considering the constraints of sender nodes. The receiver node will demand as minimum the fulfilment of the service level specification to obtain a playout quality.

**2) Active Sender Node:** It will be one of the content suppliers, where the contribution is according to its transmission capabilities and its processing availability. This node will receive an allocation packet from the receiver node. Furthermore, it will establish a congestion control TCP-friendly mechanism. Also, each Active Sender Node will communicate with Passive Sender Nodes providing control information to all participants.

**3) Passive Sender Node (backup):** This peer will perform backup functions when an active

sender node fails or its behaviour degrades the system quality. This node acts as a Pro-active listener and agent. It attempts to cover immediately any incidence in the subgroup of active sender nodes. The backup node has information of the whole system, since it exchanges control information with receiver node as well as all active sender nodes. Receiver node could pre-assign (using initial distribution algorithms) the backup role with respect to a possible failure of a determined active sender node. This pre-assignment will be based on active test algorithms of possible disconnections or failures. The target is to reduce the reaction time of the backup nodes, and hence, minimize the time to restore the system.

**4) Video Layered Encoding:** Our objective is to implement the video codification by multiple layers. This implementation will improve gradually the playout quality, according to the availability of system resources. There are two mechanisms to transmit the video in a flexible way. The first option is the Video Layered Encoding which provides high scalability. It divides stream into multiple video layers, being the base layer the fundamental one. The other layers, called enhanced layers, are used to improve the image quality. These enhanced layers require the base layer to be decoded in the receiver side. The second option is the Multiple Description Coding (MDC). This technique is a scalable solution, where each layer can be decoded separately. However, a reduction in the efficiency of the video compression is introduced; in consequence the transmission needs more bandwidth. MDC is more suitable in scenarios where the error-prone channels are considered. Additionally, when the video format is not the same in both sender and receiver sides, mechanisms for video format adaptation are required. Transcoding techniques could be used to adapt the video format (i.e. Format change of MPEG to H.263). In this case, a third element within the Active Grid Network is needed to implement the transcoding function. The new transcoding nodes contribute with their CPU and bandwidth resources.

**5) Number and capacity of receiver buffers:** The receiver node has a buffer distribution mechanism by sender node and by layer (see Fig.

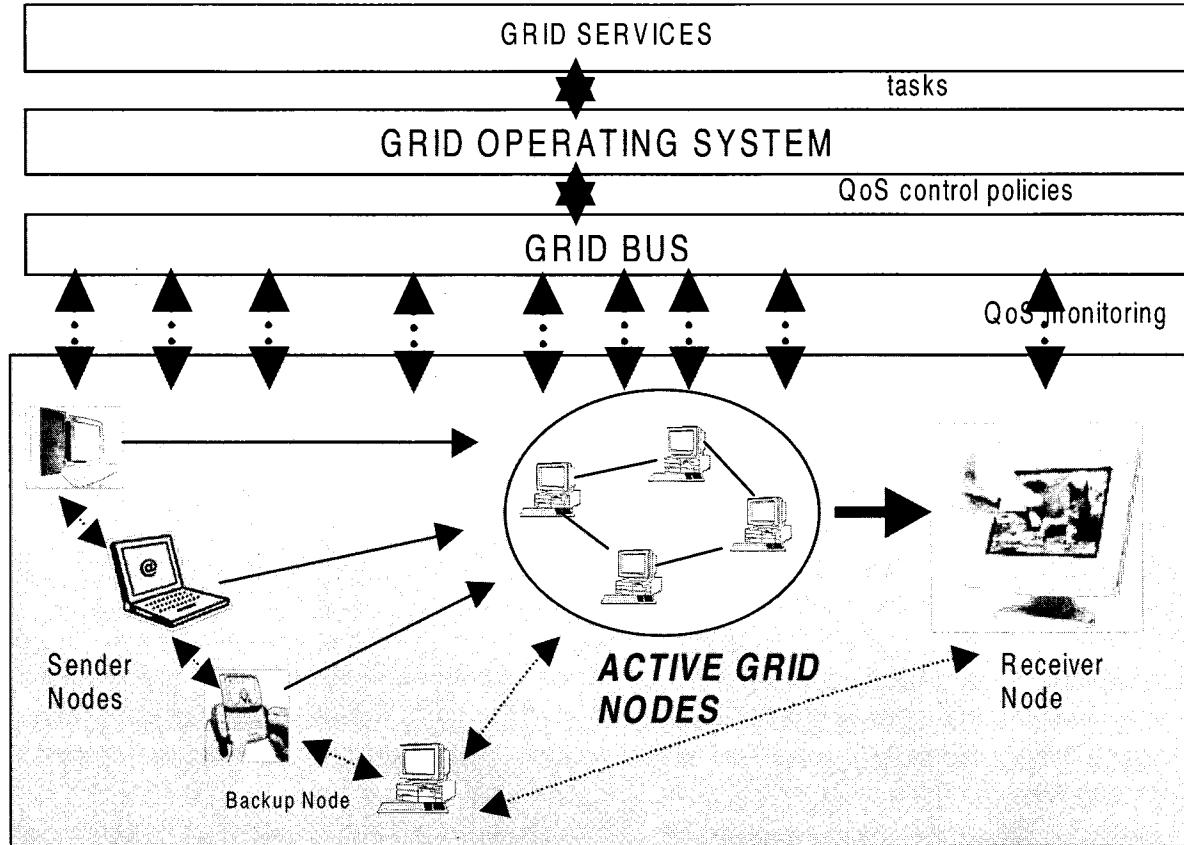


Fig. 3: Active P2P Grid Transmission Architecture

2). The receiver node will request a percentage of base layer and one or more enhanced layers to each active sender nodes. Hence, a control by peer and by layer will be applied. Our model prioritizes the control of buffers corresponding to the base layer. This priority is required to mitigate the effects of congestion, loss or delay in the network; hence, the first buffers that will be withdrawn will be the buffers associated to the enhanced layers. To find a function that helps to define the optimal state of each buffer is one of our challenges. Finally, a control function is implemented to regulate the incoming flow rate of each buffer in order to control their fullness. In an extreme situation, the system must react to fitting the video quality.

**6) Active Grid Computing QoS framework:** The quality adjustment could be seen from two points of views. The first viewpoint is an adjustment of coarse granularity. If a degradation of the system occurs and there is not enough information in buffers, the receiver must eliminate the highest video layer. The second viewpoint is an adjustment of fine granularity, which implies the reduction of

intra-layer transmission rate. When a temporal congestion occurs, the receiver request to active nodes to reduce the transmission rate per layer. Major changes will be managed by the Active Grid. Each active grid node is able to process additional complex functions, and not only routing traffic. These functions could be the video transcoding, transmission forwarding decisions, terminal dependence forwarding, QoS monitoring, among others.

In conclusion, in this section the basis of the proposed model (Fig. 3) have been presented. The model consists in the following relevant principles: collaborative work, load balancing, multi-source recovery, adaptive quality and buffering techniques. All of these principles are implemented by the above mentioned elements.

#### 4. CONCLUSION AND FUTURE WORK

In this work, an exhaustive analysis of the topics that involve the Video-on-Demand streaming services using Active P2P Grid networks has been

addressed. The emergence of these services and the heterogeneity of resources in VOs make this topic an important and vast field of latest research work. This work presents a novel Active Grid Video Streaming Recovery architecture. The concepts of Collaborative Computing, Fairness Distribution and Quality Adaptation are the basis of this proposal.

This architecture has been designed using two roles of sender nodes. These nodes are differentiated by their capabilities and their contribution to the global system performance. The presented model divides the nodes set in two subclasses. The first subclass is the active sender nodes which send the video to the receiver node in a collaborative manner. This collaboration mechanism involves a multi-source retrieval operation. The other subclass is the passive sender nodes, called backup nodes. These backup nodes are prepared to react in whatever irregular situation or whenever the Active Grid operating system decides. The collaboration between active and passive sender nodes improves the behaviour of the system and hence, the video quality perceived by the users.

Our model uses an Active Grid QoS framework to adjust the performance of the system. The Grid system applies QoS policies; meanwhile the active grid nodes will be monitored by processes to accomplish the quality requirements. Video Layered Encoding is used to exploit the implementation of multiple buffers. Buffering techniques are implemented to control the state of each buffer. These techniques ensure available buffers for the base layer mainly, and others buffers for the enhanced buffers. Security Policies are also integrated to guarantee the authenticity and integrity of the received video information. VoD services are provided based on Grid Security Infrastructure (GSI) protocols.

Finally, we are convinced that the present proposal can be extended in several directions, not without before making an analysis process of results and simulations. We are considering to use NS2 and Globus Toolkit. The future work could include the following subjects: Design of a prototype and extend to real applications.

## REFERENCES

- [1] I.Foster, C.Kesselman and S.Tuecke. «The Anatomy of the Grid: Enabling Scalable Virtual Organizations». Int. J. Supercomputer Apps., Vol. 15, No. 3, 2001
- [2] Napster: <http://www.napster.com>
- [3] L.Fan, P.Cao, J.Almeida, Z.Broker, «Summary cache: A Scalable Wide-Area Web Cache Sharing Protocol». IEEE Trans. Networking vol. 8, pp. 281-293, June 2000.
- [4] G.Peng, »Content Distribution Network« [www.ecsl.cs.sunysb.edu/~gpeng/](http://www.ecsl.cs.sunysb.edu/~gpeng/) June 2003.
- [5] Y.Chu, S.G.Rao, H.Zhang, «A Case For End System Multicast», Proc. of ACM SIGMETRICS, Santa Clara, CA, June 2000, pp 1-12.
- [6] S.Banerjee, B.Bhattacharjee, C.Kommareddy, G.Varguese. «Scalable Application Layer Multicast». In Proc. of ACM SIGCOMM'02, 205-220, USA, August 2002
- [7] X.He, X.Tang, J.You. «Supplying Instantaneous Video-on-Demand Services Based on Grid Computing». GCC2003: The Second International Workshop on Grid and Cooperative Computing, Shanghai, China, December 2003.

## AUTHORS



*Jorge Mata Díaz, es Doctor Ingeniero en Telecomunicación y profesor del Departamento de Ingeniería Telemática de la Universidad Politécnica de Cataluña. Su trabajo de investigación se enmarca dentro del grupo «Servicios Telemáticos» del mismo departamento y su actividad docente se desarrolla en la ETS de Ingeniería de Telecomunicación de Barcelona. Entre sus campos de investigación se encuentra el desarrollo y propuesta de servicios telemáticos de redes de comunicación terrestres, sin hilos y VSAT: servicios audiovisuales, sistemas distribuidos multimedia, servicios corporativos en entornos intranet e Internet, y servicios de seguridad y comercio electrónico. Los resultados de sus trabajos se han publicado en numerosas revistas y congresos tanto nacionales como internacionales (<http://sertel.upc.es>)*



*Juan Carlos Peláez López, es estudiante de Doctorado de la Universidad Politécnica de Cataluña, en el Departamento de Ingeniería Telemática, donde se encuentra vinculado al grupo de investigación de Servicios Telemáticos desde 2001. Recibió el grado de Ingeniero en Sistemas Computacionales de la Universidad Católica de Santiago de Guayaquil, Ecuador, en 1998. Entre sus campos de investigación se encuentran el desarrollo y propuesta de servicios distribuidos multimedia, tecnología peer-to-peer, Web Services, Grid Computing, y comercio electrónico. Actualmente se encuentra trabajando en su proyecto de tesis sobre «Video Streaming Peer-to-Peer with QoS for the recovery and delivery».*



## COSMOS PROLONGADO Y RETRAIDO

Miguel Escudero Royo

*escudero@mat.upc.edu*

*Departament de Matemàtica Aplicada, Universitat Politècnica  
de Catalunya*

En los últimos meses y con inusitada frecuencia se están ofreciendo noticias astronómicas que apenas nadie entiende con plenitud pero que la Prensa resalta como relevantes. Habría que preguntarse si sabemos bien lo que ya se sabía con anterioridad. Así, antes de las más recientes informaciones los especialistas enmarcaban la antigüedad del Universo entre diez y quince mil millones de años. Nuestra Tierra, a la que como planeta se le otorga unos 4.500 millones de años de edad, sigue siendo opaca, no irradia luz o energía en cantidad apreciable. Por eso sólo se nos puede ver desde fuera mediante «reflexión», es el destino de la condición humana.

Los objetos más luminosos del Universo son los *quasars* y a la hora de hoy marcan sus confines. Estas «casi estrellas» están a unos 10.000 millones de años luz de nosotros. Traduzcamos para poder comparar: Un solo año luz representa nueve billones de kilómetros (para obtener este resultado basta multiplicar la velocidad de la luz, estipulada en 300.000 kilómetros por segundo, por los 86.400 segundos que tiene cada uno de nuestros días y por los 365 días de rigor anual; siguiendo la notación europea, un billón es igual a 10 elevado a 12 unidades, 9 en versión norteamericana).

Parece ser que existen millones de galaxias semejantes a la Vía Láctea, la cual a su vez tiene del orden de cien mil millones de estrellas. Hay quienes estiman que el Universo contiene más galaxias que estrellas tiene una galaxia. Ante eso no puedo por menos que recordar unos versos de Quevedo: «El mentir de las estrellas / es muy seguro mentir / porque ninguno va a ir / a preguntárselo a ellas» y repensar esta machadiana: «Sólo el poeta puede / mirar lo que está lejos / dentro del alma, en turbio / y mayo sol envuelto». Contemplar con serenidad y cordura lo que está lejos dentro del alma nos acerca a la realidad personal. Los rayos que despiden el «poeta» enriquecen la mirada de un receptor atento, humilde y generoso, ayudándole a destellar con

particular luz y calidez. En nuestras manos está el uso de la técnica del zoom personal.

En los últimos años está de moda una peculiar paternidad a distancia, mediante vínculos a cuenta bancaria. Desde el «primer mundo» se «apadrinan» niños de otros lares (pero no a los destinados al aborto en la propia casa, los hijos no deseados por sus padres biológicos deberían tener oportunidad de vivir con padres adoptivos en lugar de ser condenados al reciclaje). Asimismo los zoos también buscan «padrinos» o sponsors para sus criaturas. Quizá sea curioso saber que si alguien propusiera ahora el apadrinamiento de nuestras estrellas galácticas, a cada hombre (de los, redondeando, siete mil millones de vivos en el planeta azul) le podría corresponder por igual diez estrellas ahijadas, y aún quedarían unas pocas por «colocar».

La distancia que media entre el Sol, la estrella de nuestro sistema, y la Tierra es de unos 150 millones de kilómetros. Y la que separa a nuestro vagabundo (eso significa planeta) de la estrella actualmente más próxima al Sol, la *Próxima Centauri*, es de 4 años luz; por eso la «vemos» ahora como fue hace cuatro años, y si un día desapareciese no nos enteraríamos de ello hasta cuatro años después. Se podrían ofrecer muchos más ejemplos para dar idea del contraste de los límites que marcan los órdenes de magnitud, los hay que nos obligan a estar alerta contra la inercia mental: Mercurio posee el período de revolución más corto de los planetas, 88 días terrestres tarda en dar la vuelta al Sol.

Su velocidad orbital es muy alta, unos 48 kilómetros por segundo, en cambio, su velocidad de rotación es extremadamente baja y tarda en girar sobre sí mismo unos 180 días terrestres. Así pues, en Mercurio dos años (2 vueltas completas al Sol) equivalen a un solo día (una rotación planetaria completa). De cada dos noches mercurianas, una es Nochevieja en la Tierra (cada año nuestro se



corresponde con 1,94 días mercurianos). Por otro lado, si se situase en Venus, un planeta de tamaño similar al que habitamos, un observador tendría el mundo al revés, el Sol sale por el Oeste y se pone por el Este.

El sistema solar está casi vacío, hay mucha más radiación que materia (el volumen ocupado por la materia viene a ser una billonésima parte del volumen total del sistema). Su silencio no me llega a aterrizar, sí, en cambio, el dolor concentrado que almacena. «Cuando recordar no pueda, / ¿dónde mi recuerdo irá? / Una cosa es el recuerdo / y otra cosa recordar», escribió Antonio Machado. El tiempo se acumula en ese silencio de forma misteriosa e inaprensible. Por cierto que los rayos del Sol tardan unos ocho minutos en pasar de la superficie solar a la terrestre. En cambio, sujetos a un proceso aleatorio de absorción y emisión, esos rayos pueden tardar millones de años en ir del núcleo del Sol a su superficie (unos 700.00 kilómetros en línea recta). El misterio es inseparable de la ciencia, un conocimiento irrenunciable para nosotros. Las preguntas «¿Quién soy yo?» y «¿Qué será de mí?» son incombustibles en cualquier lugar del Universo y en cualquier momento de la Historia, son un intento de sobrevivir y siempre van juntas a estas otras dos: «¿Quién eres tú?» y «¿Qué será de ti?». A la ciencia no se le puede pedir lo que no puede dar.

Permítanme volver, impenitente, la mirada de nuevo hacia Antonio Machado: «Ya soy más viejo que eras tú, padre mío, cuando me besabas. / Pero en el recuerdo soy también el niño que tú llevabas de la mano. / ¡Muchos años pasaron sin que yo te recordara, padre mío! / ¿Dónde estabas tú en esos años?». Estos versos filiales transportan a la existencia de una inexistencia. El hombre comienza por ser el que aún no es como tal, afirmó Ortega. Somos también el que dejamos de ser y quien quisimos ser, con mayor o menor intensidad, en cada una de nuestras trayectorias biográficas. Forman parte de nuestra realidad.

Con sensibilidad podemos añorar sentir los corazones conocidos o desconocidos que antaño latieron y que mañana latirán. Sería la conquista de la actualidad del pasado y del futuro. La clave del quehacer humano está en saber si su serie asociada, al modo matemático, converge, diverge u oscila. Si convergiera, esto es, si la suma de sus infinitos términos tuviera límite real, importaría conocer su valor, pues no se trata sólo de la posibilidad de

rebobinar cada vida humana sino de liberarla y poderla llevar más allá y más acá de modo intra y extra-vagante. Acaso la trascendencia sea un nefable zoom vital y amoroso por una Historia con n dimensiones.

*Giordiano Bruno* clasificó magias, entre ellas destacó la «*Mathematica u occulta philosophia*». El profesor rumano *Ioan Petru Culianu* (1950-1991; una muerte temprana) -autor del libro «*Eros y magia en el Renacimiento*», prologado por *Mircea Eliade*- sitúa la función de la magia renacentista no como un factor de desorden, sino como un medio para restablecer una coexistencia pacífica entre el inconsciente y el consciente. Su terreno propio sería aglutinar relaciones intersubjetivas. Para *Culianu*, «la tecnología viene a ser una magia democrática que permite a todo el mundo gozar de las facultades extraordinarias de las que, hasta ahora, sólo podía presumir el mago». Comprendo su consideración de que un libro es una siembra de fantasmas que va dirigida a un segador desconocido. Verán ustedes enseguida adónde vamos a ir a parar. Podríamos divagar por el espacio al compás de la suite orquestal «*Los Planetas*», de *Gustav Holst* (1874-1934), en pos de Júpiter (símbolo de la alegría) y en conjunción con Saturno (símbolo de la vejez) o alcanzar Venus (símbolo de la paz) tras eludir Marte (símbolo de la guerra). Pero recalemos de nuevo en la Tierra. Les propongo aplicar el zoom a una crónica americana no muy presente para casi nadie. Me refiero a la «*Historia natural de la Antigua California*», de *Miguel del Barco* (1706-1790). Hay una clase de libros a los que me asomo con la certeza de tener que acelerar mi lectura ya desde el principio, pero también con casi la seguridad de acabare encontrando algo que me merezca la pena la dedicación. Así me sucedió con las páginas de este jesuita extremeño que durante treinta años vivió en México como misionero. Hube de esperar a una adición final para aplicar el zoom de mi cámara lectora, acercar esos párrafos y clavarles la mirada.

Tratando de asuntos etnológicos y lingüísticos, *Del Barco* habla, dos siglos y medio atrás, de las diversas naciones y lenguas que pueblan la California. Advierte que en América el nombre nación suele tener distinta significación que en Europa, donde a su vez también se toma de diferentes maneras. «En Europa -dice- se da nombre de una nación a los que viven en cierta extensión de terreno, o bajo de cierto dominio, sean o no de un lenguaje». Pero «en la América, por lo regular, no habiendo entre los

indios que ahora se conquistan, ni distinción o límites de provincias ni separación de dominios, cuales se hallaron en los dos imperios de México y de Perú, se reputan por una nación todos los indios que usan un mismo lenguaje, sean pocos o muchos; bien que vivan cerca unos de otros; bien que derramados en distantes rancherías; o que si se diferencian en el idioma, es poco, por ser unas lenguas, dialectos de las otras, de modo que puedan entenderse entre sí mismos». Destaca que hay misioneros que han escrito que las lenguas de esta península son seis, otros que cinco y otros que tres, si bien un examen más profundo revela pocas diferencias entre algunas que «no merecen el nombre de idioma distinto». Del Barco se decanta por el dictamen del padre *Segismundo Taraval* que da tres lenguas: la cochimí, la pericú y la de Loreto. Las naciones principales serían tres: la nación de los pericúes (llamados vulgarmente pericos), la de los monquis o lauretanos y, por último, la nación cochimí. «Para proceder con la claridad posible entre tanta confusión, y no tropezar después», prosigue nuestro autor distinguiendo que al indio que cae al sur o al mediodía de su territorio se le llama edú («gente de otra lengua» en cochimí, que a su vez significa «gente que vive por la parte del norte»), pero esta palabra «comprende varias naciones, de las cuales cada una tiene su particular nombre con que se distingue de las demás, será más acertado no usar este vocablo para evitar equivocación y confusión: ni en la California se usa de él, hablando en castellano», apostilla *Del Barco*.

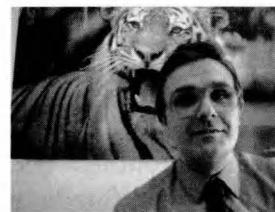
Entre los hábitos de los «californios», el jesuita cuenta que «las suegras no miraban a sus yernos, porque estaban persuadidas a que, si los miraban, enfermarían de los ojos o perderían la vista», lo que califica de bobería y superstición. Siempre simpatizando con los naturales del lugar, *Del Barco* explica que los misioneros optaron «mientras no se hallaba modo mejor para explicarles nuestros misterios» por la palabra huayib para traducir «resucitar» en el Credo. Esa era la primera palabra cochimí en que prorrumpían los indios al ver moverse las moscas, y denotaba también el acto de levantarse de quien estaba acostado o enfermo, pero no sentado. Por cierto que dice asimismo que los cochimíes llamaban al día ibó, nombre que también daban al Sol.

Claro está que la asimilación de estas curiosidades se produce en contextos variados. Yo las reúno con mis particulares conexiones y las enfoco bajo mi propia perspectiva, no puede ser de otro modo. Cavilo acerca de las identidades, problema con el que entraremos en el nuevo milenio, y busco componentes y fuerzas integradoras de la realidad. Me encuentro en el camino con *Julián Marías*, quien desde su perspectiva cristiana nos participa «la esperanza de que la realidad tenga sentido, la creencia de que lo tiene y por eso puede buscarse e indagarse».

Envueltos en incertidumbre y vacunados contra la alienación y la crueldad de la cerrazón, abiertos a las decepciones nos sentimos invitados a pensar a Dios como plenitud de la realidad. ¿Cómo? Desde lo más próximo, su imagen accesible e inteligible de persona.

En ese trayecto cósmico, multidireccional en busca de paz y sosiego, de redención del pasado y afán de porvenir, de luz afectuosa y compasiva, del mejor yo que deseamos ser, nos encontraremos con edúes, gentes de otras lenguas y otras galaxias, siquiera sean mentales, personas como nosotros. Con compostura y desde el punto de adorno que el cosmos oculta y exhibe a la vez, nos disponemos a reabsorber nuestras circunstancias con el destino concreto de pasar de persona a estrella: «Podríamos decir que la infinita distancia no sólo se acorta, sino se anula, sin mengua de su infinitud. Dios se ha hecho hombre, el hombre participa de la vida divina. Pero todo ello procede de Dios», nos apunta Marías.

## AUTOR



*Miguel Escudero* es profesor titular de la Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona, adscrito al Departamento de Matemática Aplicada IV de la Universidad Politécnica de Cataluña. También es Doctor en Filosofía y Letras y ha escrito artículos de opinión en publicaciones como 'Cuenta y Razón', el diario 'La Vanguardia' o 'Buran'.



# HOMBRES ESCONDIDOS EN FÓRMULAS



Carl Linné

## EL PADRE DEL HOMO SAPIENS

Miguel Escudero Royo

escudero@mat.upc.edu

Departament de Matemàtica Aplicada, Universitat Politècnica  
de Catalunya

Finalmente lo he conseguido. Era una cuestión pendiente para mí, puede que hasta exagerando, una «asignatura pendiente», aunque no tenía demasiada prisa por aprobarla. El otro día, inesperadamente, sin buscarlo, topé con el dato. *El homo sapiens* fue definido por un sueco nacido en 1707, el mismo año que nació el gran matemático suizo *Leonhard Euler*. Se trata de *Carl Linné* o *Carolus Linnaeus* (con el nombre latinizado). ¿Quién era este hombre?

De niño fue un gran aficionado a la botánica (una palabra que proviene del griego y que significa hierba). Cuando tenía veinte años empezó a estudiar medicina, pero con veintidós ya firmaba su primer trabajo de botánica. Participó en expediciones científicas como la célebre a Laponia (dentro del Círculo Ártico), donde se midió un grado del meridiano para comprobar el achatamiento polar de la Tierra. Contreinta años fue designado miembro de *l'Academie des Sciens* de Francia, y se estableció en Estocolmo como médico. (En aquella temporada el español *Antonio de Ulloa* descubrió el platino y *Daniel Bernouilli* –el más destacado de una impresionante familia de científicos suizos– planteó la teoría cinética de los gases.)

Poco después se casó, tuvo su primer hijo y fue nombrado profesor de medicina en la Universidad de *Uppsala*. Un año más tarde pasó a ser profesor de botánica, su vocación «desinteresada».

Ya en 1744 murió su compatriota *Anders Celsius* (seis años mayor), el astrónomo que propuso la escala de temperaturas centígrada, y *Pierre Maupertuis* formuló el principio de la mínima acción de la naturaleza. En seguida comenzaría a publicarse *l'Encyclopédie*, editada por *Diderot y D'Alembert*. Poresas fechas, *Carl Linné* detallaría en el libro «*Philosophia Botanica*» la manera de nombrar y clasificar las plantas.

La gran manía de *Linné* era clasificar y ordenar, tanto animales como plantas y minerales. Siguió una nomenclatura binomial: cada especie era designada mediante dos palabras en latín: una genérica (en mayúsculas) y otra específica (en minúsculas). Así llegó al *HOMO sapiens*, unas 8000 especies animales y unas 6000 plantas (estas según sus flores; un criterio artificial pero práctico). Hoy por hoy tengo entendido que hay clasificados un millón y medio de especies animales y un millón de plantas.

Al establecer las categorías taxonómicas, popularizó la ciencia botánica. Parece interesante señalar que la palabra taxonomía proviene de dos vocablos griegos: de *taxis*, derivado de «tasso» y significa ordenación, orientación; y *nomos*, la ley. *Carl Linné* creía que las especies de los seres vivos eran inmutables y escribió: «Tot sunt species quae ab initio creavit infinitudens», es decir, «Hay tantas especies como el ser infinito creó en el inicio».

El rey de Suecia lo nombró caballero y le concedió el derecho a llamarse *von Linné*. Aquél hombre, padre del *homo sapiens*, murió en *Uppsala* en 1778 (el mismo año que murieron *Voltaire* y *Rousseau*) de un ataque al corazón. Unos años atrás había sufrido una angina de pecho y sufrió mal de gota, ciática y fuertes migrañas.

*González Bueno* dice en «El principio de los botánicos» (Nivela) que *Linné* tenía los ojos oscuros, era bajito y atlético. Tenía mala oratoria y mal oído para disfrutar de la música, y además le costaban mucho los idiomas. Pero su memoria visual era extraordinaria. No tengo noticia de cómo eran su olfato y tacto. Sí que fumaba en pipa y le gustaba bastante el café. Su casa la dejó totalmente en manos de su mujer.