

El reto de emparejar contenidos

Investigadores de la Carlos III trabajan para mejorar la equivalencia entre búsqueda de información y resultados en Internet, y con ello, para una mejor gestión de la comunicación en las redes

Todo cuanto viaja por las redes de telecomunicaciones es información. De ella se espera que pueda satisfacer las expectativas de un usuario que a menudo se ve sometido a un bombardeo indiscriminado en forma de textos, de voz, de imágenes fijas o de video que recibe sin haberlo requerido o respondiendo a peticiones de contenido que no siempre se ajustan a lo solicitado inicialmente. Una fórmula para evitar la recepción de información no deseada y al mismo tiempo facilitar su búsqueda podría ser el emparejamiento de contenidos, la asociación entre usuarios e informaciones precisas. En esta línea trabaja el grupo dirigido por Aníbal Figueiras en el Departamento de Teoría de la Señal y Telecomunicaciones de la Universidad Carlos III de Madrid.



Aníbal Figueiras, responsable del Departamento de Teoría de la Señal y Telecomunicaciones y del Grupo de Tratamiento de Información y Datos

Xavier Pujol Gebellí

Uno de los grandes retos de Internet, y por extensión de las distintas redes de telecomunicaciones existentes, es cómo acomodar los flujos de información y de contenidos en tiempo y forma para un usuario determinado. Algo así como lograr que aquello que uno precisa o desea recibir en un momento dado responda exactamente a sus expectativas. La cuestión, aunque pudiera parecer simple, no lo es en absoluto y, en cambio, su importancia es alta: no son pocos los que creen que el éxito o el fracaso del negocio futuro de las telecomunicaciones pasa por ahí. Para ello se precisa de tecnología y desarrollo de producto, pero también de conceptos. La idea del emparejamiento, en la que trabajan algunos de los grupos más punteros del mundo, es una de ellas.

El grupo de Tratamiento de Información y Datos de la Universidad Carlos III de Madrid, dirigido por Aníbal Figueiras, se dedica a esta tarea. En su ánimo está aplicar los "conceptos clásicos" de las telecomunicaciones, según los describe Figueiras, para el tratamiento de la información multimedia que hay en las redes de comunicaciones y en los sistemas informáticos. Esto es, voz, textos, imágenes fijas, video, gráficos o, en definitiva, "todo lo que pueda aparecer en un sistema de comunicaciones o en una máquina de cómputo". Muchas de estas informaciones, puntualiza, son susceptibles de explotación económica.

El objetivo fundamental de la línea iniciada por el grupo es determinar qué parte de la información que genera una fuente o está depositada en una base o un repositorio debe ser transmitida (es lo que se llama codificación de fuente). En paralelo, determinar también cómo hay que disponer esa información para poder transmitirla (codificación de canal). Un tercer objetivo es determinar cómo hay que recuperarla. Y finalmente, definir cómo hay que presentarla. Todo ello culmina en un punto "esencial": establecer cómo hay que gestionar la información en las redes.

"Estamos concibiendo el problema de las telecomunicaciones como radicado en la dificultad de personalizar o hacer corresponder contenidos o informaciones específicas con usuarios". Eso significa, a juicio de Figueiras, que hoy por hoy la

La capacidad de las redes es enorme, lo difícil es gestionar adecuadamente las posibilidades que presenta toda esa capacidad

capacidad de transmisión y la potencia de cómputo, no constituyen ningún cuello de botella, pero sí lo que se ha convenido en llamar personalización de la información. "Cada usuario precisa de una información servida en tiempo y forma", insiste el experto. La complejidad viene dada por la enorme cantidad de información que circula y por los millones de usuarios existentes. Acomodar a cada uno requiere de algo más que ingenio.

¿Cómo hacerlo? "Primero hay que caracterizar la información, después hay que caracterizar al usuario, después hay que determinar procedimientos que sean capaces de crear esas parejas". Es decir: clasificar la información para cada usuario y luego segmentar a los usuarios para las informaciones que puedan desear. Figueiras lo resume como un problema de toma de decisiones: "¿Es una información determinada útil o relevante para una persona, máquina, organización, sitio y momento?"

Muy probablemente, la solución deba pasar por la elaboración de protocolos de decisión, mecanismos que sean capaces de dotar al sistemas de parámetros para discriminar lo que es o no relevante según el tipo de información y el usuario. De algún modo, aunque resulta una expresión "un tanto burda", sería como añadir inteligencia al sistema.

Más allá del buscador

Los robots de búsqueda y los catálogos de índices que actualmente pueblan Internet son lo más parecido al concepto que describe Figueiras, pero se quedan todavía a años luz de lo que debería ser a su juicio la resolución del problema. "Los buscadores son todavía muy elementales", advierte. Y añade: "Parece poco verosímil que quien busque se lea todos los índices".

El equipo de la Carlos III participa en distintos proyectos que tratan de dar solución a esta problemática. En uno de ellos, en colaboración con una agencia de noticias, el objetivo que se busca es servir textos con información precisa a un usuario concreto que se encuentra en un lugar determinado. Por ejemplo, las últimas noticias de índole económica de Brasil.

El problema en este caso es construir una aplicación que analizase los textos, que descubriese si esa información se adapta a la demanda de ese usuario y se la enviase. Si se hace adecuadamente, el usuario "estará contento" porque recibirá lo que desea en el momento oportuno. Si se hace de forma abusiva o escasa, "el usuario se defiende", dejando de consultar o perdiendo interés por el servicio. Por tanto, la aplicación debe afinar en el análisis de ese texto para que la tasa de aciertos sea relevante.

Para que la aplicación resulte exitosa no puede limitarse, como ocurre con muchos de los actuales buscadores, a seleccionar palabras en titulares o en el texto entero. "Hay que diseñar sistemas capaces de efectuar otro tipo de exploraciones, que tengan un cierto nivel de inteligencia o de autonomía y que puedan determinar la relevancia de lo encontrado". El modelo ideal, admite, sería el basado en "etiquetas humanas" que informaran con precisión acerca del contenido, aunque la evolución más lógica probablemente sea el establecimiento de un modelo en el que "la cosa [la información] tiene que facilitar su reconocimiento y el reconocedor [el sistema de



Figueiras, en su despacho.

Se necesitan
sistemas de
búsqueda en
Internet mucho más
independientes y
autónomos, que
funcionen con
"etiquetas humanas"

análisis o el robot de búsqueda] tiene que ser lo suficientemente hábil para encontrarla bien”.

El concepto de negocio

La evolución de Internet y de las redes de telecomunicaciones hace previsible que surjan nuevos formatos o incluso nuevos productos. Pero ello no quita que, de algún modo, deban continuar indexándose los contenidos y facilitando su búsqueda o su recuperación. Y el problema es que la comunicación no se establece de forma bidireccional, como ocurre con el teléfono, ni está localizada en unos pocos nodos. “A medida que aumenta el volumen de información y el número de usuarios, la complejidad se incrementa”. A ello contribuye también el tipo de información (las imágenes presentan dificultades añadidas) u otras cuestiones como la necesidad de preservar la seguridad o la propiedad intelectual.

“Nadie esperaba una explosión de las telecomunicaciones tan brutal como la ocurrida”, continúa Figueiras. “Y ahora están en crisis porque son víctimas de si mismas”. A su juicio, la entrada de agentes en el sector con un “escasísimo nivel de conocimiento”, tanto de los fundamentos técnicos como del negocio de las telecomunicaciones, ha distorsionado un tanto la realidad. Es parte, dice Figueiras, de la explicación de la burbuja de la nueva economía.

El emparejamiento de contenidos y usuarios, en su sentido más amplio, podría utilizarse como justificación para la prestación de servicios de pago. Figueiras, no obstante, plantea una objeción: “Si lo que se me sirve es lo que ya obtengo por menor coste en otro sitio, pagar no tiene sentido”. Lo contrario, en cambio, sí estaría justificado. Representaría ensayos del tipo periódico a medida, formación de comunidades o, en definitiva, adaptación de contenidos a usuarios. Sobre esta base, cree el experto, es probable que se rescriba parte del negocio de las telecomunicaciones.

"Inteligencia" en la red

A los expertos que trabajan en telecomunicaciones y gestión de la información la palabra 'inteligencia' les suena un tanto excesiva. Pero para un profano, emplear algoritmos, fórmulas o sistemas que actúan selectivamente o son capaces de aprender de la experiencia o recibir supervisión, escapa a lo cotidiano. Redes neuronales, lógica difusa, algoritmos para clasificación e identificación, algoritmos genéticos, agentes inteligentes, minería y explotación de datos, junto con métodos estadísticos convencionales, conforman algunas de las herramientas 'inteligentes' empleadas por este y otros grupos del sector.

Con ellas, el grupo ha ensayado el diseño de agentes de comunicaciones que, depositados en un 'ecosistema' determinado, son capaces de competir entre sí de modo que al final del proceso sólo quedan "los mejor preparados" para acceder a una información concreta. El agente debe encontrar la información, valorarla, determinar si es útil para el usuario y efectuar las transacciones necesarias para adquirirla o

transportarla. Es aplicable sobre bases de datos, sobre conjunto de 'sites' o, en general, en repositorios de información.

Una segunda línea de investigación del grupo plantea como objetivo tratar de mejorar las condiciones de transmisión. Esto es, crear dispositivos para mejorar las condiciones del canal. Han combinado varios esquemas de manera que de forma automática el sistema es capaz de adoptar la fórmula más adecuada a cada situación.

En paralelo, el grupo ha iniciado desde hace unos años cursos específicos para fomentar la creatividad, una experiencia que pretende que las nuevas generaciones de técnicos aprendan a "vencer el miedo, la falta de estímulo o el excesivo acomodo a las reglas" que acaban limitando su libre pensamiento. El planteamiento, según Figueiras, es que los estudiantes "aprendan a plantearse las cosas de un modo distinto al tradicional".

FICHA TÉCNICA

Grupo de Tratamiento de Información y Datos

Centro: Departamento de Teoría de la Señal y Telecomunicaciones. Escuela Politécnica Superior. Universidad Carlos III de Madrid.

Coordinador: Aníbal Ramón Figueiras

Dirección: c/ Butarque, 15 - 28911 Leganés (Madrid)

Teléfono: 91 624 99 23

Fax: 91 624 87 49

Email: arfv@ing.uc3m.es arfv@tsc.uc3m.es

Líneas de investigación: Minería de datos y Ayuda en toma de decisiones (Telecomunicaciones, Finanzas, Salud, Defensa); Personalización de servicios y accesos en Telecomunicaciones (Intranet, Internet, GPRS/UMTS); Gestión de la información y el conocimiento (teleeducación, teleocio, trabajo colaborativo).