

INFORMACIÓN EMBARGADA HASTA EL MARTES 22 DE AGOSTO, 12:00 HORAS, hora central europea)

Jon Kleinberg Premio Nevannlinna

El trabajo de Jon Kleinberg ha proporcionado una visión teórica para abordar importantes cuestiones prácticas que se han convertido en esenciales para la comprensión y la gestión de nuestro cada vez más interconectado mundo. Ha trabajado en una amplia variedad de áreas, desde el análisis de redes y el enrutado, a la *minería* de datos, la genómica comparativa o el análisis de la estructura de las proteínas. Además de haber realizado contribuciones fundamentales en investigación, Kleinberg ha reflexionado sobre el impacto de la tecnología en sus esferas social, económica y política.

Uno de los más importantes logros de Kleinberg se centra en la estructura de red de la World Wide Web. Sus aportaciones han tenido una enorme influencia en la forma de operar de los más importantes sistemas de búsqueda en internet. Mientras que las páginas web individuales tienen un grado de estructura impuesto por sus creadores, la estructura de la WWW en su conjunto carece por completo de planificación; emerge y cambia continuamente a medida que se incorporan nuevos enlaces y contenidos. Por ello, es todo un desafío el diseño de formas eficientes de seleccionar las páginas relevantes para ofrecer a los usuarios de los sistemas de búsqueda. Antes del trabajo de Kleinberg, estos sistemas se centraban solo en el contenido de las páginas, no en la estructura de enlaces. En 1996, Kleinberg introdujo las ideas de "autoridades" y "hubs". La primera se refiere a las páginas que contienen información sobre un tema concreto, y la segunda a la que contiene enlaces a muchas "autoridades". Por ejemplo, si se teclea la búsqueda de "cámaras digitales", una web que ofrezca guías, comentarios o un panorama general sobre diferentes cámaras será un "hub", mientras que las páginas de los fabricantes serán "autoridades". Aunque la distinción entre ambos conceptos es intuitivamente sencilla, definirlos matemáticamente es difícil por la circularidad que presenta: un "hub" ofrece conexiones a muchas "autoridades" y cada una de éstas tiene enlaces con muchos "hubs". La contribución clave de Kleinberg fue proporcionar una forma para romper esta circularidad, abriendo el camino para un análisis matemático de la estructura de enlaces de las redes. Kleinberg desarrolló las herramientas algorítmicas para resolver este problema y demostró su efectividad probándolas en la World Wide Web.

Otro campo en el que Kleinberg ha realizado aportaciones fundamentales es en las redes del tipo "el mundo es un pañuelo". Estas redes fueron ideadas por el psicólogo Stanley Milgram en experimentos realizados en los años 60, y se convirtieron en el centro de diferentes modelos matemáticos basados en el trabajo de Duncan Watts y Steve Strogatz en los 90. En los experimentos de Milgram, "A" debía hacer llegar un mensaje a "B", a quien no conocía directamente,

a través de otras personas, en el menor número de pasos posible. Cada persona de la cadena solo podía enviar el mensaje a otra que conociera personalmente. Milgram observó que, de media, el mensaje llegaba a su destinatario en tan solo seis pasos (éste es el origen de la expresión "seis grados de separación"). El número 6 es sorprendentemente pequeño, dado el tamaño de la población global y que cada persona operaba solo con información local de la red. Hasta bien avanzados los 90, había una amplia literatura científica sobre la existencia de atajos para conectar individuos en la red social, pero apenas se había prestado atención a la forma de encontrar estos caminos solo con información local: ¿cómo puede encontrar el camino más corto gente que no tiene una visión global de la red completa y solo conoce a sus amigos inmediatos? Esta es la pregunta que Kleinberg se hizo, y descubrió algunas cosas sorprendentes.

Los nodos de una red social son las personas que la forman. Dos nodos están conectados si las dos personas se conocen mutuamente. Kleinberg observó que en una red de este tipo la probabilidad de que dos nodos se conecten desciende a medida que aumenta la distancia geográfica entre ellos, y mostró que si lo hacía con el cuadrado de la distancia entre ellos, existían algoritmos capaces de encontrar el camino más corto entre ambos. De forma más sorprendente aún, mostró que si decrecía más rápido o más despacio que el cuadrado de la distancia no existía ese algoritmo. Además de resolver aspectos teóricos de redes de este tipo, el trabajo de Kleinberg ha sido útil en el desarrollo de sistemas P2P de intercambio de archivos, en los que la información tiene que localizarse sin que esté centralizada.

Uno de los primeros resultados de Kleinberg también se refiere a algoritmos eficientes, en esta ocasión para el problema de encontrar "vecinos cercanos" en grupos de datos de alta dimensión. El problema se presenta en situaciones como las siguientes. Supongamos que tenemos un grupo de documentos y un diccionario de n palabras. Para cada documento contamos el número de veces que aparece la primera palabra del diccionario, la segunda, la tercera, etc. De esta forma, obtenemos n cuentas, que pueden ser consideradas como un vector de n dimensiones representativo del documento. Para un documento cualquiera, "A", ¿cómo podemos identificar el vecino más próximo, es decir, el documento "B" que es el más parecido a A de todos los documentos del grupo? Pensando en estos documentos como vectores, lo que pretendemos es encontrar B de forma que la distancia entre A y B sea lo más pequeña posible. Una solución, por supuesto, consiste simplemente en medir las distancias entre A y todos los demás vectores y mirar cual es la más corta. Pero si n (el número de palabras en el diccionario) es muy grande, nos enfrentamos con un espacio de muy altas dimensiones, donde esta técnica de fuerza bruta, de comprobar todas las posibilidades se convierte rápidamente en una "carrera de dimensionalidad" y en consecuencia, cara y excesivamente larga. Kleinberg rompió la situación de estancamiento desarrollando una nueva e ingeniosa aproximación al problema del vecino más cercano, mediante la combinación aleatoria de proyecciones unidimensionales de los vectores.

Otro hito en el trabajo de Kleinberg es el desarrollo de un modelo matemático para reconocer *fogonazos* en el interior de flujos de datos. Una forma de analizar la estructura de la información contenida en un texto es localizar las explosiones de actividad que aparecen repentinamente y se sostienen durante un cierto periodo de tiempo. Por ejemplo, aunque usted no hable inglés, si observa la aparición repentina de la palabra "Katrina" y la elevada frecuencia de su uso en los medios de comunicación en agosto de 2005, usted comprenderá que algo significativo ha ocurrido. Kleinberg empezó estudiando estos *fogonazos* por una razón eminentemente práctica: quería encontrar el mejor medio para organizar su propio archivo personal de mensajes electrónicos que se le estaban acumulando. Aunque la idea de estos *fogonazos* de actividad informativa es intuitivamente clara, definir con rigor la noción no es sencillo, debido a la dificultad de distinguir estos fenómenos de las fluctuaciones estadísticas en un mar de información aleatoria. La aportación fundamental de Kleinberg fue proporcionar semejante definición, utilizando el concepto matemático de los modelos de Markov. Comprobó sus ideas con diferentes grupos de datos, como los títulos de los artículos científicos distribuidos muchas décadas antes en un área determinada. En este caso, los fogonazos correspondían a la aparición y atención prestada a nuevos temas en ese área de investigación.

La interacción entre la comprensión teórica de las redes y la observación de redes en funcionamiento ha impregnado la investigación de Kleinberg y proporcionado nuevas perspectivas sobre el papel de la tecnología en la sociedad. "Internet y la web nos obligan a pensar sobre las consecuencias sociales de un mundo en el cual la información es más completa y viaja más ampliamente que nunca antes, y en el que cualquiera tiene la posibilidad, a través de nuevos mecanismos de comunicación y a un precio muy bajo, de convertirse en un autor con una audiencia global", escribió Kleinberg en una entrevista por correo electrónico con la revista digital Noticias de Investigación Tecnológica (Technology Research News). "Pero hay un buen número de retos aquí. Sabemos que el discurso *on-line* puede estar altamente polarizado, ¿significa que las herramientas *on-line* que hemos creado están contribuyendo a incrementar el nivel de polarización en el diálogo social en general? ¿Cómo podemos determinar con certeza este fenómeno y qué debemos pensar sobre el diseño de nuevos instrumentos para hacer el discurso *on-line* más productivo?" El trabajo de Kleinberg se distingue por su riqueza y diversidad, y también por su habilidad para utilizar visiones teóricas y profunda comprensión en su aplicación a problemas prácticos. Esta poderosa combinación asegura a Kleinberg un papel estelar como uno de los pensadores más influyentes en la ciencia de computación teórica de los próximos años.

Nota Biográfica

Jon Kleinberg nació en 1971 en Boston (Massachusetts, EE.UU.). Se doctoró en 1996 en el Massachusetts Institut of Technology (MIT). Es profesor de ciencia de la computación en la Universidad Cornell. Entre

los reconocimientos que ha tenido se encuentran el *Sloan Foundation Fellowship* (1997), *Packard Foundation Fellowship* (1999) y el *Initiatives in Research Award* de la U.S.: *National Academy of Sciences* (2001). En el 2005, Kleinberg recibió la designación de *MacArthur genius Fellowship* de la *John D. and Catherine T. MacArthur Foundation*.

Imágenes:

hubs-authorities.pdf: Los puntos rojos representan 'hubs', y los negros, 'autoridades'. Imagen cortesía de Jon Kleinberg.

networks.gif: Estructuras en red. [Es necesario el permiso de la revista Science para reproducir esta imagen]

Artículo completo en:

<http://www.cs.cornell.edu/home/kleinber/sci01.pdf>

