

Datos de móviles predicen en qué zona de Londres sucederá un crimen

Combinados con información sobre cada barrio, los investigadores pudieron determinar con un 70% de eficacia qué zonas de la ciudad sufrirían actos criminales el mes siguiente

[Miguel Ángel Criado 21 SEP 2014 - 11:49 CEST1](#)

¿Cuánto daría un buen alcalde o su jefe de policía por saber en qué barrio de la ciudad se va a producir un crimen? Pues un grupo de investigadores sólo ha necesitado un perfil sociodemográfico de la zona y la actividad de sus antenas de móviles para predecir, con hasta un 70% de grado de acierto, en qué zona se cometerá un robo o un asesinato en las calles de Londres. Todo un ejemplo de que el [Big Data](#) también sirve para el bien común.

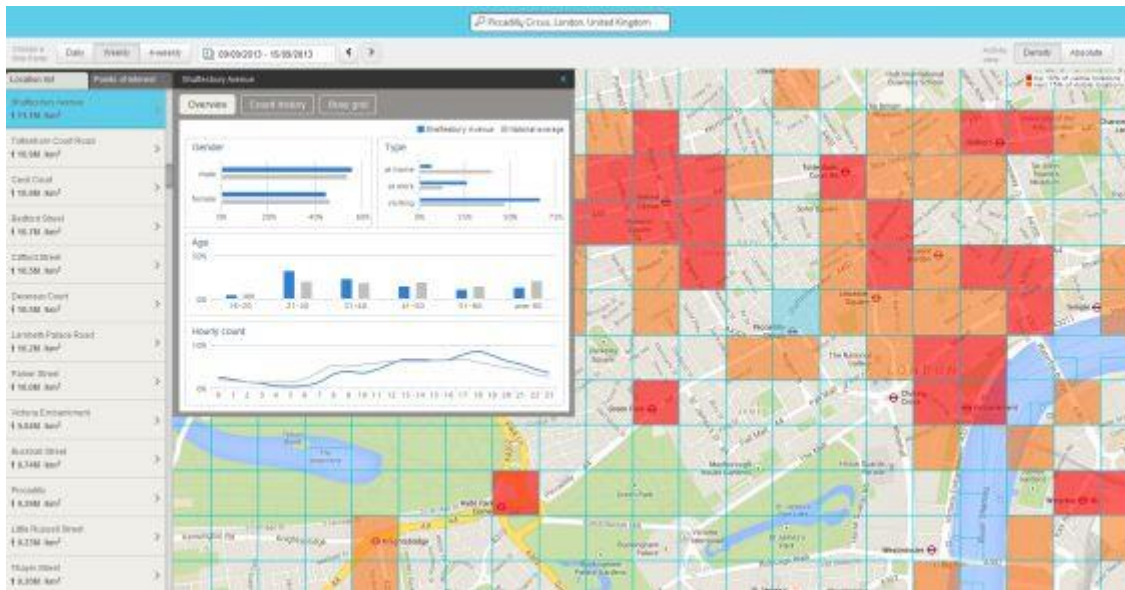
En esta noticia

- [Documento](#)

En los últimos años se creado un gran negocio alrededor de la predicción del crimen. Apoyadas en las estadísticas policiales, empresas como IBM han desarrollado herramientas de análisis predictivo que dan pistas de dónde se va a cometer un asalto o el robo de un coche. Ya son muchas las policías, en especial en Estados Unidos, que usan estos complejos y caros sistemas de minería de datos. Pero, ¿Hay alguna forma más económica, asequible y eficaz de mejorar sus predicciones? Sí, con los datos de los móviles y sin saber nada de la identidad de sus usuarios, sólo con la actividad recogida en las antenas de telefonía móvil.

Durante la [Campus Party](#) de 2013 celebrada en Londres, el Instituto Tecnológico de Massachussets (MIT), el británico Instituto de Datos Abiertos y Telefónica retaron a varios equipos de investigadores a sacarle el mayor jugo posible a grandes cantidades de datos. Una de esas bases de datos es [Smart Steps](#), un producto de la operadora española (por ahora sólo disponible en Reino Unido), que usa los datos móviles de red anónimos y agregados para calcular cuánta gente hay en una zona determinada.

“Usando los datos conocidos del barrio [variables como su demografía, su perfil étnico, clase social, tipo de casas...] se puede predecir en un 62% si habrá un crimen, con Smart Steps, el porcentaje se eleva hasta el 68% y combinado ambos tipos de datos aún más”, asegura la directora científica de Telefónica Research y coautora de un estudio sobre la predicción del crimen en Londres, [Nuria Oliver](#).



Con Smart Steps, de Telefónica, se registra la actividad de los móviles de forma anonimizada y agregada sobre una zona determinada. / Telefónica Research

Los investigadores dividieron Londres en 124.119 cuadrículas. Con los datos del [Open Data Institute](#) pudieron hacer un perfil exacto de cada celda con 68 variables, desde su historial de crímenes hasta el tiempo, pasando por el régimen de propiedad típico de los inmuebles de la zona. Y, sobre esa capa, añadieron los datos del tráfico recogido durante tres semanas por las antenas de telefonía móvil de O2 (la filial de Telefónica en el Reino Unido) previamente anonimizados y agregados. A pesar de esta salvaguarda de la privacidad, los investigadores pudieron hacer una estimación de la edad, sexo y si el tráfico procedía de los móviles de residentes del barrio, trabajadores o personas que iban de paso.

“Para cada localización, computamos una respuesta a la variable alta criminalidad-baja criminalidad basada en la base de datos con los casos criminales geolocalizados facilitada por la policía de Londres, explica el investigador de la Universidad de Trento (Italia) y principal autor de la investigación publicada en Arxiv, Andrey Bogomolov.

Por desgracia y por razones de privacidad, las autoridades habían eliminado de los registros la fecha exacta en que se había producido cada crimen y sólo estaban catalogados por meses. Por eso, su modelo sólo era capaz de señalar qué cuadrícula en la que habían dividido el Gran Londres había caería en la categoría alta criminalidad (o baja) en el siguiente mes. Aún así el mapa generado gracias a sus algoritmos coincidía casi en un 70% con la distribución geográfica de los crímenes en el mes estudiado. “Creo que, con los datos diarios, las métricas de predicción serían mucho mejores”, asegura Bogomolov.

Y dice aún más. Aunque no era el objetivo de su estudio, con los datos inferidos del tráfico de los móviles, hay crímenes que se pueden predecir mejor que otros. “Desafortunadamente, no teníamos todos los datos sobre los crímenes durante el tiempo que duró el reto”, reconoce el investigador. Sin embargo, ya sí los tienen y están trabajando en la publicación de sus resultados, que presentarán en el congreso [ICMI 2014](#).

“La gran ventaja de los datos móviles es que se pueden hacer estimaciones casi en tiempo real”, recuerda la investigadora española. A diferencia de las estadísticas oficiales, de los censos, que se elaboran y publican en lapsos de tiempo muy grandes, la información ofrecida por las redes de [telefonía móvil](#) está disponible al instante.

Además, inferir la conducta humana de este tipo de datos es más barato que hacerlo desde las herramientas que usan el Big Data para hacer análisis predictivo. Productos como el [SPSS de IBM](#) necesitan un mar de datos de todo tipo, un gran centro de datos para almacenarlos y muchos ingenieros para analizarlos. “Somos los primeros que ha resuelto este problema del Big Data con los datos de telecomunicaciones, que son muy accesibles y baratos de recoger. Redes GSM hay en todas la megápolis del mundo”, recuerda Bogomolov.

Más información

- [Documento: 'Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data' ArXiv:1409.2983](#)
- [El gripazo de Google muestra las flaquezas del 'big data'](#)