

We know what you did last Sónar

Real-time user tracking & Interactive data visualization

Project website:

<http://www.bsc.es/viz/whatyoudid/>

About the project:

Nowadays, we are being followed and tracked by multiple sensors and networks. In addition to digitally generated data, physical tracking of people is on the rise, and there is a strong trend to join physical and online tracking. For example, to link customer's purchase records and behavior in a physical store to their online social activity, and gain new insight on their preferences, tastes, interests, and shopping habits. The data we generate is analyzed and mined for value, many times without our consent or, at least without our fully awareness on the information we give away.

At the BSC Visualization Team, with a small budget and using available technology, we developed proprietary software and hardware to track the position of all mobile devices during the festival Sónar by Day 2015. Dedicated teams in larger companies can easily go beyond this point, and many are, in fact, doing so. Personal data mining, collection, and analysis by third parties will certainly become an important issue and a central topic of discussion especially in the absence of clear regulation and legislation.

Our project aims to raise awareness on this issue, and to highlight the affordances in this area made possible by already available technology and knowledge.

The data:

We tracked the mac address of all mobile devices in Sónar by Day 2015 that had wi-fi connection on. We deployed a series of tracking devices in the different spaces of the venue that collected this data during the three days of the festival. Through proprietary software and hardware, we were able to collect the mac address of the devices, and their approximate location at any given time. With this information we can find out the exact time they arrived and left each stage. We also used a structured database of all the events scheduled in the festival to create a personalized recommendation system based on what they did before.

Our tracking devices are passive detectors based on wi-fi signals, so users don't need to register or connect to any site for the system to track the approximate location of their mobile devices. Each device can track signals within a 100 meters radius approximately, depending on

whether they are located on open or confined space, the presence of obstacles, and other similar factors.

This kind of technology is not new. However, our system makes use of very economic and affordable components that may be used potentially to obtain the precise location of mobile devices, only by developing and implementing specific software.

Nevertheless, data detection is just the first step of the process. Highly complex algorithms are needed to build and handle Big Data databases, to “clean up” the incoming data, to remotely communicate the information from the tracking devices to dedicated servers, and to visualize the resulting information.

The tracking devices detect signals continuously and send the collected data to the server every 30 seconds. Once there, proprietary software based on Java Servlets clean up the raw data and store every detected mac address and its associated metadata in a Cassandra database, an open source Big Data managing platform. In computer simulation tests before Sónar, our system was able to process 30.000 mac addresses for each tracking device every 15 seconds, maintaining the same speed for our 10 tracking devices.

The Cassandra database permits to add computing nodes with almost linear scalability. Thus, with relatively small effort our system could scale to track larger amounts of people at even bigger events.

Visualization:

We created three kinds of visualizations to let the users explore different aspects of the data. One kind shows the movement of all Sónar By Day attendants, the second displays the personal unique trajectory of a single person, and the third is a personalized recommendation page based on the person’s activity during the festival.

The visualization “What they do” shows a real-time flow map featuring the number of attendants on each space of the venue and the flows of people between spaces.

This visualization features filters to visualize different information, like the signal strength of the detected devices, the provenance of the signals in each stage, and

Similar to the above, the Sankey diagram “What they did” is a flow-chart that shows the historic movements between stages for all detected devices focusing on time and stage, like a live schedule enriched with the attendants’ movements information. Both visualizations are based on a graph generation algorithm that compares the current and existing mac addresses in the database every 5 minutes.

In contrast, in the page “What I did”, users can register their device mac address to see the unique map of what they did during the whole festival: The stages they visited, the hour they arrived and left each one, and the events they saw at those times.

That information is also the base for the recommendations system featured in the page “What you like”. Based on the historic record of spaces visited by a person, and the events that took place at that moment, the system suggests related artists and events.

Public Installation:

The interactive web app was shown between June 18-20, 2015 during Sónar by Day 2015. Visitors were able to see the construction of the visualizations on real time, and interact with the application at the VizTeam stand of Sónar+D.

Interactive web app:

The web app offering the same interactivity and functions as in Sónar 2015 is available on <http://www.bsc.es/viz/whatyoudid/>

Artists:

The project was created by the Visualization Team of the Barcelona Supercomputing Center. Team members are Fernando Cucchiatti, Guillermo Marin, Luz Calvo, David García, and Artur García.

Contributors:

Antoni Artigues (BSC), David Fusté, and Jorge García Vidal (BSC).

Summary

Nowadays, we are being followed and tracked by multiple sensors and networks. In addition to digitally generated data, physical tracking of people is on the rise, and there is a strong trend to join physical and online tracking. For example, to link customer's purchase records and behavior in a physical store to their online social activity, and gain new insight on their preferences, tastes, interests, and shopping habits. The data we generate is analyzed and mined for value, many times without our consent or, at least without our fully awareness on the information we give away.

At the BSC Visualization Team, with a small budget and using available technology, we developed proprietary software and hardware to track the position of all mobile devices during the festival Sónar by Day 2015. Dedicated teams in larger companies can easily go beyond this point, and many are, in fact, doing so. Personal data mining, collection, and analysis by third parties will certainly become an important issue and a central topic of discussion especially in the absence of clear regulation and legislation.

Our project aims to raise awareness on this issue, and to highlight the affordances in this area made possible by already available technology and knowledge.

We deployed a network of sensors during Sonar by Day for three days to detect and follow wireless communication devices, such as mobile phones, and make a real time analysis of how they move around the space of the festival. Our tracking devices are passive detectors based on wi-fi signals, so users don't need to register or connect to any site for the system to track their approximate location. Though this kind of technology is not new, our system makes use of

very economic and affordable components, potentially capable to obtain the precise location of mobile devices, only by developing and implementing specific software.

Nevertheless, data detection is just the first step of the process. Highly complex algorithms are needed to build and handle Big Data databases, to “clean up” the incoming data, to remotely communicate the information from the tracking devices to dedicated servers, and to visualize the resulting information.

We developed proprietary software based on Java Servlets to clean up the raw data and store every detected mac address and its associated metadata in a Cassandra database, an open source Big Data managing platform. In computer simulation tests before Sónar, our system was able to process 30 000 mac addresses for each tracking device every 15 seconds. As the Cassandra database permits to add computing nodes with almost linear scalability, with relatively small effort our system could scale to track more people at even bigger events.

Because of ethical and legal requirements, we anonymized all the data collected in our project so the audience can't be identified without their consent. Only the people that stop by the BSC space (or our web, after Sónar) and register their devices will access visualizations of their own unique activity, how they fare in comparison to the rest of the crowd, and obtain personalized recommendations based on which events they have attended or liked. In this way we also aimed at showing the useful, “bright” side of machine learning and artificial intelligence.

The results are shown in appealing and information-dense interactive visualizations that let the users explore the movement patterns of thousands of people moving around different spaces, access a personalized recommendation system based on what they did before, and obtain their own personal “footprint” of their activity during the festival.

Short description

Nowadays, we are being tracked by multiple sensors and networks. In addition to our online social activity, physical tracking of people is on the rise, and there is a strong trend to join physical and online tracking. The data we generate is analyzed and mined for value, many times without our consent or without fully awareness on the information we give away. Our project aims to raise awareness on this issue, and to highlight the affordances in this area made possible by already available technology and knowledge.

At the BSC Visualization Team, we developed proprietary software and hardware to track the position of all wireless communication devices during the festival Sónar by Day 2015. We deployed a network of sensors that detect and follow those devices across the venue. For data management and processing we developed proprietary software based on Java Servlets to clean up the raw data and store every detected mac address and its associated metadata in a database specialized for Big Data management. Our system was able to process 30.000 mac addresses for each tracking device every 15 seconds.

We anonymized all the data collected in our project so the audience can't be identified without their consent. Those who register their devices during the festival or via the project's website may access personalized visualizations and recommendations based on the events they have attended or liked.

The results are shown in appealing and information-dense interactive visualizations that let the users explore the movement patterns of thousands of people moving around different spaces, access a personalized recommendation system based on what they did before, and obtain their own personal "footprint" of their activity during the festival.

About BSC

Barcelona Supercomputing Center - Centro Nacional de Supercomputación (BSC-CNS) hosts the supercomputer MareNostrum. It also counts with well-known supercomputing research groups that develop tools for the academia and industry. BSC focuses its research areas in Computer Sciences, Life and Earth Sciences and Computer Applications in Science and Engineering. In the context of this multi-disciplinary approach, BSC counts with more than 350 researchers and experts in HPC (High Performing Computing) and 100 of those are from outside Spain.

BSC-CNS was constituted as a public consortium formed by the current Spanish Ministry of Economy and Competitivity (Ministerio de Economía y Competitividad), the Department of Economy and Knowledge of the Catalan Government and the Technical University of Catalonia. Barcelona Tech (UPC), and is headed by Professor Mateo Valero.

In 2011, the BSC-CNS was recognized as a "Severo Ochoa Centre of Excellence" for its contributions and research agenda in the area of computing and applications. In the first edition of the Severo Ochoa programme, the Ministry of Science and Innovation selected 8 research centres and units in Spain to be among the best in the world in their respective fields.







