

ANALYSING HUMAN MOBILITY PATTERNS

Romuald Mosdorf, Tomasz Rybak

Faculty of Computer Science, Bialystok Technical University

Many ways of calculating positions of people or distance they travel exists. The most known system allowing for calculating position in almost any part of the globe is GPS (Global Positioning System). It uses triangulation to calculate position basing on distance from satellites. Necessity of using many satellites means that GPS does not work inside buildings and has problems working reliably in cities. WiFi or BlueTooth access points' known positions can be used to estimate position in dense-populated areas. However they also pose some problems. One needs to know positions of many access points; the more, the more precisely current position can be estimated. This means that they require maintaining enormous database of access points. As BlueTooth devices are mobile (cellular phones, computers, etc.), their usage as static points of reference may be problematic. WiFi access points can be created and disabled, so one needs many of them to have secure overlay allowing of calculating positions. Above technologies were not intended to get positions — their main reason of existence is to transmit data. Because of that, they use varying power levels of signals, and some of the devices do not even allow reading this strength without much trouble. So one cannot calculate exact position but only roughly estimate it.

Custom system build with intention of calculating positions inside buildings is needed to overcome those problems. Sputnik/OpenBeacon is one of many systems that intend to trace participants of conference. There are even systems that contain infrared cameras and microphones to trace interactions between participants. Others use GSM phones, which are constantly carried by people, to trace their movements. Using custom build system allows for having total control and to research accuracy of different algorithms.

Sputnik system was used during 24th Computer Chaos Conference, where 220 participants were carrying badges transmitting signals. Participants were either on the lectures or at the venue, socialising with other participants. 37 readers were placed inside building. This data set was used to calculate positions of tags by using strength of signal transmitted by tags and known positions of readers. Three different algorithms of calculating positions were used, and then data coming from them used as source for further analysis.

From data set with positions few interesting were chosen to further investigation. The longest sequences were chosen, as amount of data they contained promised more basics for analysis. To analyse data non-linear chaotic functions were used. Among chosen data sequences three different classes were discovered.

Although results were very interesting, noise in data makes them less certain. This means that there is need for more advances in hardware design, firmware, and algorithms for calculating positions, or at least more planning with placements of readers, so one can get better position estimation. Nonetheless informations are very interesting and point for further investigation.

Lack of social context of participants also disallows further analysis, making impossible answering detailed questions why some behaviours occurred.