

Inferring social contacts relevant for disease transmission from mobile phone call data records

Background:

Each time a mobile phone is used to make a call, send a text message or access the Internet, the service provider retains a record of this event for billing purposes. These records contain the time, duration and destination of each event as well as the location of where it was made. In contrast to using these data for commercial purposes, this project would repurpose them for public good. Call Data Records (CDR) offer a wealth of information on the social interactions of subscribers, on a massive scale.

However, these data are not immediately representative of potentially infectious contacts; using the CDR directly for modelling infectious disease spread between individuals would be flawed; when person A calls person B, they are unlikely to be in the same place and therefore unable to transmit disease. Under the assumption that pairs of individuals are more likely to have social contact if they talk on the phone, simple analysis of the call data can already give us some useful insights.

Objectives:

Using a combination of call data and location data, the aim is to develop novel methods to infer real social ties and build a social network with unprecedented detail. Given such social interaction data, we will develop a simple model of disease transmission and examine potential strategies for control of epidemics.

What the student will do:

The student will have access to CDR from a European country, from February 2009 onwards. The project will be centred on the statistical analysis of these data and the development of a transmission model.

Inferring potentially infectious contacts will be done in a number of ways. For example:

- 1) Pairs of individuals that have a history of phone call interaction that are found to be in the same vicinity at the same time, can be assumed to be socially linked. Such pairs will be identified and their behaviour characterised.
- 2) Individuals in the same age group found in the same vicinity may help describe settings such as schools (for younger users) or workplaces.

Once defined, the social network will serve as a substrate for the transmission of infection. The end goal is to have a model of infectious disease spread based on the social interaction data extracted from CDR, aiding our understanding of infectious disease spread in complex social settings.

Skills, techniques and follow on:

This project is ideally suited to a student with strong numerical/computational skills, basic knowledge of statistical analysis and graph theory. The student will have the opportunity to acquire the knowledge necessary to embark upon a PhD project on network inference, infectious disease modelling and understanding collective human behaviour.

References:

1. Eagle et al. Inferring friendship network structure by using mobile phone data. Proc Natl Acad Sci USA (2009) vol. 106 (36) pp. 15274-8
2. González et al. Understanding individual human mobility patterns. Nature (2008) vol. 453 (7196) pp. 779-782