

DOCTORAL CANDIDATE: Pål Sundsøy
DEGREE: Doctor Philosophiae
FACULTY: Faculty of Mathematics and Natural Sciences
DEPARTMENT: Department of Informatics
AREA OF EXPERTISE: Data science / Computational social sciences
DATE OF DISPUTATION: 22th of February 2017

DISSERTATION TITLE: Measuring patterns of human behaviour through large-scale mobile phone data: Big Data for social sciences.

English version :

Analysis of large amounts of data, so called Big Data, is changing the way we think about science and society. One of the most promising rich Big Data sources is mobile phone data, which has the potential to deliver near real-time information of human behaviour on an individual and societal scale. Several challenges in society can be tackled in a more efficient way if such information is applied in a useful manner.

Through seven publications this dissertation shows how anonymized mobile phone data can contribute to the social good and provide insights into human behaviour on a large-scale. The size of the datasets analysed ranges from 500 million to 300 billion phone records, covering millions of people. The key contributions are two-fold:

Big Data for Social Good: Through prediction algorithms the results show how mobile phone data can be useful to predict important socio-economic indicators, such as income, illiteracy and poverty in developing countries. Such knowledge can be used to identify where vulnerable groups in society are, improve allocation of resources for poverty alleviation programs, reduce economic shocks, and is a critical component for monitoring poverty rates over time. Further, the dissertation demonstrates how mobile phone data can be used to better understand human behaviour during large shocks and disasters in society, exemplified by an analysis of data from the terror attack 22nd July 2011 in Norway and a natural disaster on the south-coast in Bangladesh. This work leads to an increased understanding of how information spreads, and how millions of people move around. The intention is to identify displaced people faster, cheaper and more accurately than existing survey-based methods.

Big Data for efficient marketing: Finally, the dissertation offers an insight into how anonymised mobile phone data can be used to map out large social networks, covering millions of people, to understand how products spread inside these networks. Results show that by including social patterns and machine learning techniques in a large-scale marketing experiment in Asia, the adoption rate is increased by 13 times compared to the approach used by experienced marketers. A data-driven and scientific approach to marketing, through more tailored campaigns, contributes to less irrelevant offers for the customers, and better cost efficiency for the companies.

Norsk versjon

Analyse av store datamengder, populært kalt 'Big Data', er med på å forandre måten vi tenker om vitenskap og samfunn. En av de mest lovende kildene til Big Data er mobildata, som har potensial til å levere tilnærmet sanntidsinformasjon om menneskelig atferd, både på individ- og samfunnsnivå. Mange utfordringer i samfunnet kan håndteres mer effektivt dersom informasjonen i store datamengder anvendes på en hensiktsmessig måte. Denne avhandlingen viser gjennom syv arbeider hvordan anonymisert mobildata kan anvendes til samfunnsnyttige formål, og gi innsikt i menneskelig atferd i stor skala. Størrelsen på de analyserte datasettene varierer fra 500 millioner til 300 milliarder obsevasjoner, og dekker millioner av mennesker. Nøkkelbidragene er todelt:

Big Data anvendt til samfunnsnyttige formål: Gjennom prediksjonsalgoritmer viser resultatene at mobildata kan brukes til å måle sosioøkonomiske indikatorer, som for eksempel inntekt, analfabetisme og fattigdom i utviklingsland. Slik kunnskap er nyttig for å identifisere hvilke områder sårbare grupper i samfunnet befinner seg i, forbedre allokering av ressurser til fattigdomsbekjempelse, redusere økonomiske sjokk, samt å monitorere fattigdom over tid. Videre demonstrerer avhandlingen hvordan mobildata kan brukes til å forstå menneskelig atferd i katastrofesituasjoner. Gjennom å analysere data fra terrorangrepet i Oslo 22. Juli 2011, og en naturkatastrofe på sør-kysten av Bangladesh bidrar arbeidene til økt forståelse av hvordan informasjon spres, og hvordan millioner av mennesker forflytter seg. Slik informasjon kan bidra til at man raskt kan prioritere områder der behovene for hjelp er størst – både mer nøyaktig og rimeligere enn eksisterende metoder.

Big Data anvendt til effektiv markedsføring: Til slutt viser avhandlingen hvordan anonymisert mobildata kan brukes til å kartlegge store sosiale nettverk, med millioner av mennesker, for å forstå hvordan produkter spres innenfor disse nettverkene. Resultater fra et stort eksperiment i Asia viser at ved å inkludere sosiale mønstre og teknikker fra kunstig intelligens, kan produktopptak økes 13 ganger sammenlignet med markedsføring utført av erfarne markedsførere. En datadrevet og vitenskaplig tilnærming til markedsføring gjennom mer skreddersydde kampanjer, bidrar til færre irrelevante tilbud for kundene, og bedre kostnadseffektivitet for bedriftene.