

Når vi bruger smartphones og sociale medier som Facebook og Twitter, efterlader vi en masse virtuelle spor og data. Disse er særdeles nyttige for forskere, som ønsker at blive klogere på komplekse netværk. Lektor ved DTU Informatik, Sune Lehmann, forsker i netværksvidenskab og fokuserer på løsninger, som tilgodeser de etiske aspekter ved indsamling af data, der vedrører private oplysninger om brugerne.

# SMARTPHONES BIDRAGER TIL VIDENSKABEN

AF  
SUNE LEHMANN

DTU INFORMATIK,  
DANMARKS TEKNISKE UNIVERSITET

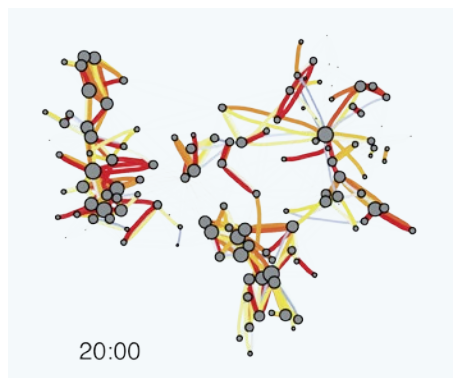
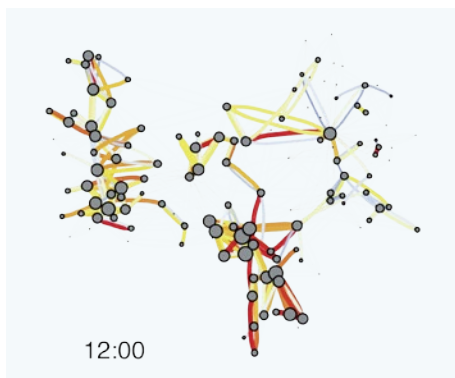
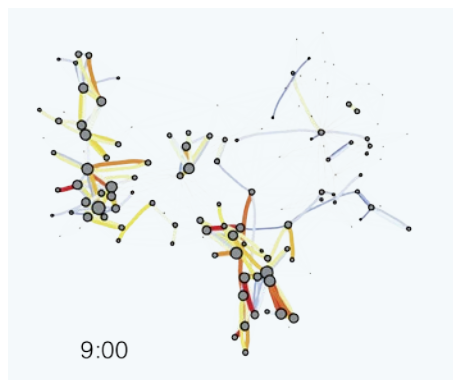
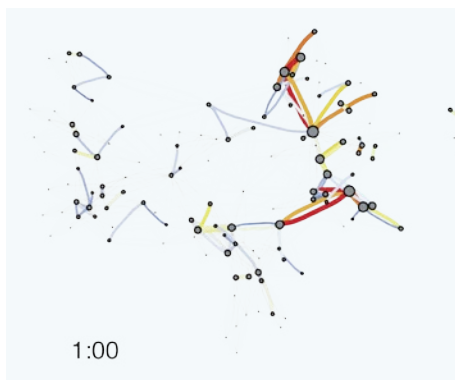
Netværksvidenskab er en interdisciplinær forskningsretning, der drejer sig om at udvikle en kvantitativ forståelse af komplekse netværk. Et komplekst netværk er, kort fortalt, en repræsentation af et naturligt system, hvor alle detaljer er abstraheret bort, og systemets elementer er beskrevet som *nodes* og deres interaktioner som *links*. Et netværk kaldes kun 'komplekst', hvis det er karakteriseret ved en ikke-triviell topologi. Selv om netværksvidenskab som en selvstændig videnskab er en ny ide, er studiet af netværk ikke nyt. Emnet er opstået - og studeres aktivt - inden for mange videnskabelige grene, f.eks. ren matematik, biologi, fysik, sociologi og datalogi. Fysikere - som jeg selv - studerer netværk, netop fordi mange matematiske værktøjer fra statistisk fysik har vist sig at være gode til at beskrive netværk, deres dynamik, og de dynamiske processer, der udfolder sig på netværk.

## BIG DATA OG SOCIALE NETVÆRK

En af grundene til, at netværksvidenskab er vokset dramatisk som felt siden årtusindskif-

tet, er, at adgangen til data har ændret sig fundamentalt. Hvert muse-klik, hvert mobilopkald, hver tur til hæveautomaten og hver statusopdatering på Facebook eller Twitter bliver gemt i en database et sted, og mængden af disse data stiger eksplosivt. Alt imens denne indsamlede 'big data' har givet lovning på en ny forståelse af den menneskelige natur, har resultaterne ladet vente på sig. Med mere end 900 millioner brugere har sociale medier som Facebook en enorm samling af vekselvirkninger mellem mennesker. På trods af den potentielle forklaringskraft i sådanne store datasamlinger, er studier baseret på data fra en enkelt kommunikationskanal stadig problematiske. Interaktioner mellem mennesker udfolder sig på mange forskellige måder. God gammeldags samtale ansigt-til-ansigt er stadig uundværlig, vi taler i telefon, sender SMS'er, skriver emails og bruger instant messaging og sociale netværk på internettet.

Jeg er netop startet på et stort grundvidenskabeligt projekt, der drejer sig om at flytte græn-



Et netværk af kommunikationsmønstre. Et udsnit af et stort mobiltelefon netværk. Hver node repræsenterer en person, og kommunikationsvolumen er repræsenteret som en farve fra blå til rød. Bemærk hvordan forskellige grupper udviser forskellige kommunikationsmønstre. Nogle er aktive indtil sent på aftenen, mens andre er aktive om morgenen (mens den første gruppe sover længe).

serne for en kvantitativt funderet forståelse af sociale systemer. Indtil videre har forskningen i sociale netværk drejet sig om at forstå netværk, som de udspiller sig på en enkelt kommunikationskanal (f.eks. Facebook som i eksemplet ovenfor). Min hypotese er, at for virkelig at forstå sociale systemer, er vi nødt til at indsamle data fra alle vore kommunikationskanaler samtidig - og i en stor population. Derfor har vi udstyret en stor gruppe frivillige med moderne smartphones, der dynamisk indsamler det sociale netværk, som det udspiller sig på alle de kanaler, som er nævnt ovenfor.

#### VIDENSKAB OG ETIK

Projektet er i sin natur tværvideenskabeligt, og mens mit eget fokus er på matematisk-tekniske emner som sampling, community detection, diffusion, spredning af information og prædiktions, håber og tror jeg, at mit datasæt kan være med til at flytte grænser for kvantitativt arbejde inden for samfundsvidenskaberne (f.eks. sociologi, antropologi og økonomi) og folkesundhed. Der er desuden etiske aspekter

ved min forskning, som gør det nødvendigt at konsultere andre fageksperter. Hensynet til eksempelvis privatlivets fred er vigtigt at diskutere med filosoffer og jurister. Detaljerede målinger af menneskers opførsel kan ikke undgå at afdække følsomme detaljer om deltagerne i forsøget, og det er man som forsker nødt til at tage højde for. Lige nu eksisterer der ganske enkelt ikke løsninger til at håndtere privatlivets fred i forsøg som vores. Datasættet er simpelthen for komplekst til, at det er praktisk muligt at anonymisere det med klassiske metoder.

Mens dette på den ene side er et praktisk problem, er det samtidig en mulighed for at gøre en forskningsmæssig indsats på dette område. Jeg har derfor allokeret ressourcer til specifikt at fokusere på dette emne. Vi forsøger altså at udvikle praktiske løsninger, som andre vil kunne benytte, og som vil give os mulighed for at dele datasættet med andre forskergrupper, mens vi samtidig sikrer privatlivets fred for deltagerne i eksperimentet.