Copenhagen Networks Study

Piotr Sapiezynski
email: pisa@dtu.dk
twitter: @sapiezynski

DTU
The problem is to find a walk through the city that will cross each bridge once and only once.
The problem is to find a walk through the city that will cross each bridge once and only once.
The problem is to find a walk through the city that will cross each bridge once and only once.
Hint: a node with an uneven number of links has to be either the starting or the end point.
Hint: a node with an uneven number of links has to be either the starting or the end point.
Networks in sociology

six degrees of separation
Networks in political science
Networks in biology
Computer networks
-map of physical connections between routers in 2005
infrastructure
questions so far?
social science

Information (bits)

big data

subjects
Information (bits)

social science

big data

subjects
First longitudinal dataset (year+) that captures human social life down to smallest temporal (5 mins) scale for approximately 1000 individuals.
proximity
multiple channels of interactions

Face-to-face

Texts

Calls

12 pm - 6 am
6 -12 am
12 am - 6 pm
6 - 12 pm
what did we learn so far?
what did we learn so far?

data bleeds through different channels
Photos/Media/Files
Uses one or more of: files on the device such as images, videos, or audio, the device's external storage

Wi-Fi connection information
Allows the app to view information about Wi-Fi networking, such as whether Wi-Fi is enabled and names of connected Wi-Fi devices

Device ID & call information
Allows the app to determine the phone number and device IDs, whether a call is active, and the remote number connected by a call
"results": [  
  {  
    "ssid": "RoyalPalace_5GHz",  
    "bssid": "44:94:fc:3a:0a:bc",  
    "level": -82,  
    "timestamp": 1398075203,  
  },  
  {  
    "ssid": "RoyalPalace",  
    "bssid": "64:70:02:8f:14:e8",  
    "level": -60,  
    "timestamp": 1398075203,  
  },  
  {  
    "ssid": "HomeBox-4296",  
    "bssid": "7c:03:4c:ca:42:9c",  
    "level": -86,  
    "timestamp": 1398075203,  
  },  
  {  
    "ssid": "litmark",  
    "bssid": "20:e5:2a:f4:2c:fb",  
    "level": -75,  
    "timestamp": 1398075203  
  }  
]
"results": [
    {
        "ssid": "RoyalPalace_5GHz",
        "bssid": "44:94:fc:3a:0a:bc",
        "level": -82,
        "timestamp": 1398075203,
    },
    {
        "ssid": "RoyalPalace",
        "bssid": "64:70:02:8f:14:e8",
        "level": -60,
        "timestamp": 1398075203,
    },
    {
        "ssid": "HomeBox-4296",
        "bssid": "7c:03:4c:ca:42:9c",
        "level": -86,
        "timestamp": 1398075203,
    },
    {
        "ssid": "litmark",
        "bssid": "20:e5:2a:f4:2c:fb",
        "level": -75,
        "timestamp": 1398075203
    }
]
This is how your last 48 hours look like from the WiFi perspective.

Each line shows when you saw one of those routers: RoyalPalace eduroam eduroam eksamen CMTW HomeBox-F8B2
Try zooming in to see more detail.
total time and location of meetings
total time and location of meetings
intimacy of meetings
total time and location of meetings
intimacy of meetings
predictability of meetings
total time and location of meetings
intimacy of meetings
predictability of meetings
similarity of social networks
what did we learn so far?
data bleeds through different channels
what did we learn so far?
data bleeds through different channels
things might be more difficult than they seem
python toolbox for visualizing geographical data and making maps

119 commits  2 branches  0 releases  2 contributors

Branch: master  geoplastlib / +

labels API and layer

andrea-cuttone authored 7 days ago  latest commit a9c2a7cfcd
what did we learn so far?

data bleeds through different channels

things might be more difficult than they seem
what did we learn so far?

data bleeds through different channels

things might be more difficult than they seem

things might be less difficult than they seem
Networks in sociology

six degrees of separation
why doesn’t it work this way?
I have friends from Warsaw, volunteer friends, my sister's friends, and friends I know through my sister.
With high-resolution data, community detection becomes unnecessary.
With high-resolution data, community detection becomes unnecessary.
what did we learn so far?

data bleeds through different channels

things might be more difficult than they seem

things might be less difficult than they seem
what did we learn so far?
data bleeds through different channels
things might be more difficult than they seem
things might be less difficult than they seem
sampling matters
snapshot subsampling

link subsampling
\( \beta = 0.03 \)

- \( T_s = 5 \text{ min} \)
what did we learn so far?
data bleeds through different channels
things might be more difficult than they seem
things might be less difficult than they seem
sampling matters
what did we learn so far?

WiFi = location

finding those locations is not always trivial

no need for community detection with good enough data

sampling matters
Measuring Large-Scale Social Networks with High Resolution
A. Stopczynski, V. Sekara, P. Sapiezynski, A. Cuttone, M. Madsen, JE Larsen, S. Lehmann

Tracking Human Mobility through WiFi Signals
P. Sapiezynski, A. Stopczynski, R. Gatej, S. Lehmann

Inferring Physical Interactions and Social Network Ties Using WiFi Signals
P. Sapiezynski, A. Stopczynski, D.Wind, J.Leskovec, S. Lehmann

Opportunities and Challenges in Crowd-Sourced War-Driving
P. Sapiezynski, R. Gatej, A. Mislove, S. Lehmann

The fundamental structures of dynamic social networks
V. Sekara, A. Stopczynski, S. Lehmann

Temporal Fidelity in Dynamic Social Networks
A. Stopczynski, P. Sapiezynski, A. Pentland, S. Lehmann