

Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 1

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

MINING DATA FROM MOBILE DEVICES

Algorithms: Location & Context

Spiros Papadimitriou, *Tina Eliassi-Rad*
[spiros.papadimitriou, tina.eliasirad@rutgers.edu]

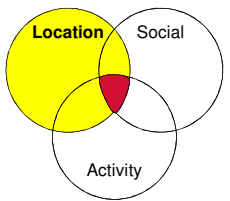


Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 2

RUTGERS

Context includes...

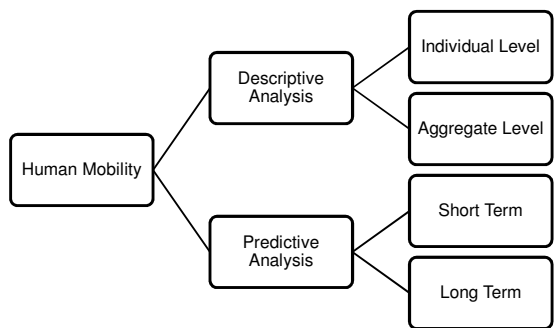
- **Location/Local**
 - What resources are nearby?
 - Where are you?
- **Social**
 - Who are you with?
- **Activity**
 - What are you doing?



Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 3

RUTGERS

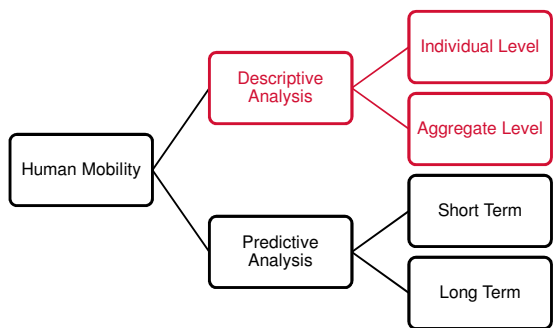
Work in human mobility



Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 4

RUTGERS

Work in human mobility

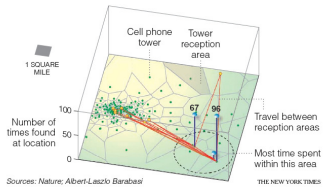


Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 5

RUTGERS

Human mobility at the individual level (descriptive)

- Human trajectories are not random!
- They have high degree of temporal and spatial regularity
- Individual humans follow simple reproducible patterns
 - Exploration + preferential return
- Impact: epidemic prevention, emergency response, urban planning, ...



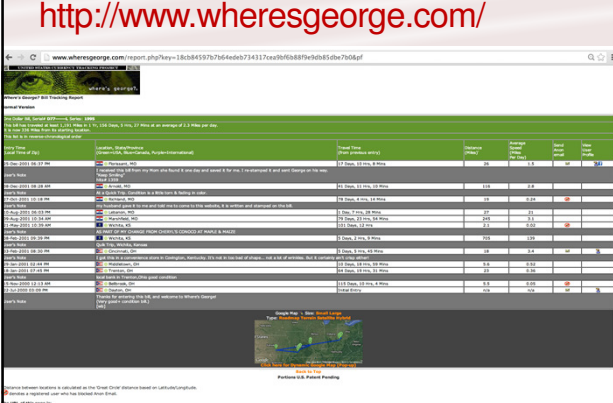
Sources: Nature; Albert-Laszlo Barabasi

Marta Gonzalez, Cesar Hidalgo, Albert-Laszlo Barabasi: Understanding individual human mobility patterns, Nature 453, 779-782, 2008.

Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 6

RUTGERS

<http://www.wheresgeorge.com/>



RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 7

Human mobility at the aggregate level (descriptive)

- Billions of anonymized Call Detail Records (CDRs) from a cellular network
- Characterized *daily travel, carbon emissions, number of workers and event goers, and traffic volumes* of hundreds of thousands of people

Richard A. Becker, Ramon Caceres, Karrie Hanson, Sibren Isaacman, Ji Meng Loh, Margaret Martonosi, James Rowland, Simon Urbanek, Alexander Varshavsky, Chris Volinsky: [Human mobility characterization from cellular network data](#). Commun. ACM 56(1): 74-82 (2013).

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 8

Human mobility (descriptive)

- M. Kim, D. Kotz, S. Kim: [Extracting a mobility model from real user traces](#). In InfoCom 2006:1-13
- K. Lee, S. Hong, S. Kim, I. Rhee, S. Chong: [Slaw: A new mobility model for human walks](#). In InfoCom 2009, 855-863
- Z. Li, B. Ding, J. Han, R. Kays, P. Nye: [Mining periodic behaviors for moving objects](#). In KDD 2010, 1099-1108
- M. Kim, D. Kotz, D: [Identifying unusual days](#). Journal of Computing Science and Engineering 5(1), 2011:71-84
- ...

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 9

Work in human mobility

```

graph LR
    HM[Human Mobility] --> DA[Descriptive Analysis]
    HM --> PA[Predictive Analysis]
    DA --> IL[Individual Level]
    DA --> AL[Aggregate Level]
    PA --> ST[Short Term]
    PA --> LT[Long Term]
  
```

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 10

Human mobility at the individual level (predictive; short-term; GPS)

- "Where are you going to be in the next hour?"
 - Successful techniques: hidden Markov models, random walk based formalisms
 - Performance around 3-5 km off; classification accuracy low 90%
 - Learning from GPS alone
 - D. Ashbrook, T. Starner: [Using GPS to learn significant locations and predict movement across multiple users](#). Personal Ubiquitous Comput. 7, 2003:275-286.
 - L. Liao, D. Fox, H. Kautz: [Location-based activity recognition using relational Markov networks](#). In IJCAI 2005.
 - J. Krumm, E. Horvitz: [Predestination: Inferring destinations from partial trajectories](#). In UbiComp 2006: 243-260.
 - B. Ziebart, A. Maas, A. Dey, J. Bagnell: [Navigate like a cabbie: Probabilistic reasoning from observed context-aware behavior](#). In UbiComp 2008:322-331.
- ...

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 11

Human mobility at the individual level (predictive; long-term; GPS)

- "Where are you going to be 285 days from now at 2PM?"
- FarOut
 - Identifies periodicity via Fourier analysis (mapping time to frequency)
 - Uses PCA for pattern extraction
 - Utilizes PCA-based classification
- Data: 32K days worth of GPS data across 703 subjects (½ people; ½ cars)
- High variance in area across subjects
 - From 30 to more than 10^8 km²
 - Surface area of earth = 5.2×10^8 km²
- Number of contiguous days = 7 to 1247 ($\mu = 45.9; \sigma = 117.8$)
- Captures both continuous (raw GPS) and discretized (triangular cells) data
- Each subject has a matrix D, where each row is a day.
- Performance continuous rep.: 1 km off; baseline 2.5km off
- Performance discrete rep.: 80% accuracy up to 80 weeks into the future; baseline ~60%

Figure 3: Our continuous vector representation of a day *d* consists of the median latitude and longitude for each hour of the day (00:00 through 23:59), binary encoding of the day of week, and a binary feature signifying whether a national holiday falls on *d*.

Figure 4: Our cell-based vector representation of a day *d* encodes the probability distribution over dominant cells conditioned on the time within *d*, and the same day-of-week and holiday information as the continuous representation (last 8 elements).

Adam Sadilek and John Krumm: [Far Out: Predicting Long-Term Human Mobility](#). AAAI 2012.

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 12

Human mobility at the individual level (predictive; long-term but < 24 hours)

- N. Eagle and A. Pentland: [Eigenbehaviors: Identifying structure in routine](#). Behavioral Ecology & Sociobiology 63(7), 2009:1057-1066
- Predictions up to 12 hours into the future
- Class labels: {Home, Elsewhere, Work, No Signal, Off}.
- PCA based classification
- 79% accuracy

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 13

Context includes...

- **Location/Local**
 - What resources are nearby?
 - Where are you?
- **Social**
 - Who are you with?
- **Activity**
 - What are you doing?



RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 14

Location & Social

1. Nathan Eagle and Alex (Sandy) Pentland. 2006. *Reality mining: sensing complex social systems*. Personal Ubiquitous Comput. 10, 4 (March 2006), 255-268.
2. Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 2011. *Friendship and mobility: user movement in location-based social networks*. KDD 2011, 1082-1090.
3. Salvatore Scellato, Anastasios Noulas, Renaud Lambiotte, Cecilia Mascolo. *Socio-spatial properties of online location-based social networks*. ICWSM 2011
4. Salvatore Scellato, Anastasios Noulas, Cecilia Mascolo. *Exploiting place features in link prediction on location-based social networks*. KDD 2011: 1046-1054
5. Yu Zheng: *Location-based social networks - users*. In Computing with Spatial Trajectories, Chapter 8, Eds: Yu Zheng, Xiaofang Zhou, Springer, 2011
6. Yu Zheng, Xing Xie: *Location-based social networks - location*. In Computing with Spatial Trajectories, Chapter 9, Eds: Yu Zheng, Xiaofang Zhou, Springer, 2011
7. Huiji Gao, Jiliang Tang, Huan Liu: *gSCorr: Modeling geo-social correlations for new check-ins on location-based social networks*. CIKM 2012: 1582-1586
8. Huiji Gao, Jiliang Tang, Huan Liu: *Exploring social-historical ties on location-based social networks*. ICWSM 2012
9. Chloé Brown, Vincenzo Nicosia, Salvatore Scellato, Anastasios Noulas, Cecilia Mascolo: *Where online friends meet: Social communities in location-based networks*. ICWSM 2012
10. Miltiadis Allamanis, Salvatore Scellato, Cecilia Mascolo: *Evolution of a location-based online social network: analysis and models*. Internet Measurement Conference 2012: 145-158
11. M. Domenico, A. Lima, and M. Musolesi. *Interdependence and predictability of human mobility and social interactions*. In *Nokia Mobile Data Challenge: http://research.nokia.com/page/12000*, MDC 2012, 2012
12. ...

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 15

Location & Social

1. Nathan Eagle and Alex (Sandy) Pentland. 2006. *Reality mining: sensing complex social systems*. Personal Ubiquitous Comput. 10, 4 (March 2006), 255-268.
2. Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 2011. *Friendship and mobility: user movement in location-based social networks*. KDD 2011, 1082-1090.
3. Salvatore Scellato, Anastasios Noulas, Renaud Lambiotte, Cecilia Mascolo. *Socio-spatial properties of online location-based social networks*. ICWSM 2011
4. Salvatore Scellato, Anastasios Noulas, Cecilia Mascolo. *Exploiting place features in link prediction on location-based social networks*. KDD 2011: 1046-1054
5. Yu Zheng: *Location-based social networks - users*. In Computing with Spatial Trajectories, Chapter 8, Eds: Yu Zheng, Xiaofang Zhou, Springer, 2011
6. Yu Zheng, Xing Xie: *Location-based social networks - location*. In Computing with Spatial Trajectories, Chapter 9, Eds: Yu Zheng, Xiaofang Zhou, Springer, 2011
7. Huiji Gao, Jiliang Tang, Huan Liu: *gSCorr: Modeling geo-social correlations for new check-ins on location-based social networks*. CIKM 2012: 1582-1586
8. Huiji Gao, Jiliang Tang, Huan Liu: *Exploring social-historical ties on location-based social networks*. ICWSM 2012
9. Chloé Brown, Vincenzo Nicosia, Salvatore Scellato, Anastasios Noulas, Cecilia Mascolo: *Where online friends meet: Social communities in location-based networks*. ICWSM 2012
10. Miltiadis Allamanis, Salvatore Scellato, Cecilia Mascolo: *Evolution of a location-based online social network: analysis and models*. Internet Measurement Conference 2012: 145-158
11. M. Domenico, A. Lima, and M. Musolesi. *Interdependence and predictability of human mobility and social interactions*. In *Nokia Mobile Data Challenge: http://research.nokia.com/page/12000*, MDC 2012, 2012
12. ...

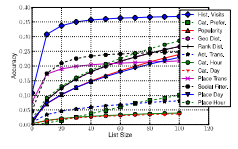
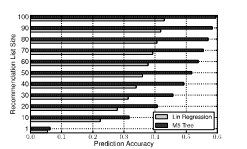
• Long distance movements are influenced by ties in the social network.

• This is not true for short-range movements or temporally periodic movements.

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 16

Predicting the next check-in on a LBSN

- Predicting the next FourSquare check-in
- A supervised approach
 - decision tree and logistic regression
- Tried out a bunch of features
 - User mobility features
 - Global mobility features
 - Temporal features
- How well?
 - Much lower accuracies compared to GPS
 - Much harder task

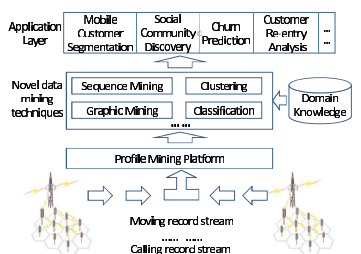



Anastasios Noulas, Salvatore Scellato, Neal Lathia, Cecilia Mascolo: *Mining User Mobility Features for Next Place Prediction in Location-Based Services*. ICDM 2012: 1038-1043.

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 17

Location & Activity

- MobileMiner [Wang et al. SIGMOD 2009]



Tengjiao Wang, Bishan Yang, Jun Gao, Dongqing Yang, Shiwei Tang, Haoyu Wu, Kedong Liu, Jian Pei: *MobileMiner: A real world case study of data mining in mobile communication*. SIGMOD 2009: 1083-1086

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 18

Context includes...

- **Location/Local**
 - What resources are nearby?
 - Where are you?
- **Social**
 - Who are you with?
- **Activity**
 - What are you doing?



RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 19

Location & Social & Activity

- Lots of work analyzing Twitter data in this space
- A recent best paper winner is from U. of Rochester
- *Flap* uses a dynamic Bayesian network per user to predict his/her locations given location of friends, time of day, type of day
- Experiments on over 4M tweets from users in LA and NYC
- It can correctly place a user within a 100m radius with up to 85% accuracy




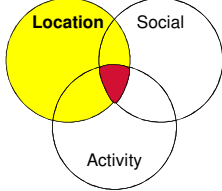
Figure 2: Flap's visualization of a sample of geotagged tweets in NYC. Red links between users represent friendships.

- Adam Sadilek, Henry A. Kautz, Jeffrey P. Bigham: *Finding your friends and following them to where you are*. WSDM 2012:723-732 (**best paper**)
- Adam Sadilek, Henry Kautz, Jeffrey P. Bigham: *Modeling The Interplay of People's Location, Interactions, and Social Ties*. IJCAI 2013.

RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 20

Context includes...

- **Location/Local**
 - What resources are nearby?
 - Where are you?
- **Social**
 - Who are you with?
- **Activity**
 - What are you doing?



RUTGERS THE STATE UNIVERSITY OF NEW JERSEY Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 21

MINING DATA FROM MOBILE DEVICES


Algorithms: Location & Context

Spiros Papadimitriou, *Tina Eliassi-Rad*
spiros.papadimitriou.tina.elias@rutgers.edu



RUTGERS Mining Data from Mobile Devices / Papadimitriou, Eliassi-Rad 22

License



These slides are made available under a Creative Commons Attribution-ShareAlike license (CC BY-SA 3.0):
<http://creativecommons.org/licenses/by-sa/3.0/>

You can share and remix this work, provided that you keep the attribution to the original authors intact, and that, if you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

© 2013 Spiros Papadimitriou, Tina Eliassi-Rad