Mobile Sensing:
Location Sensing

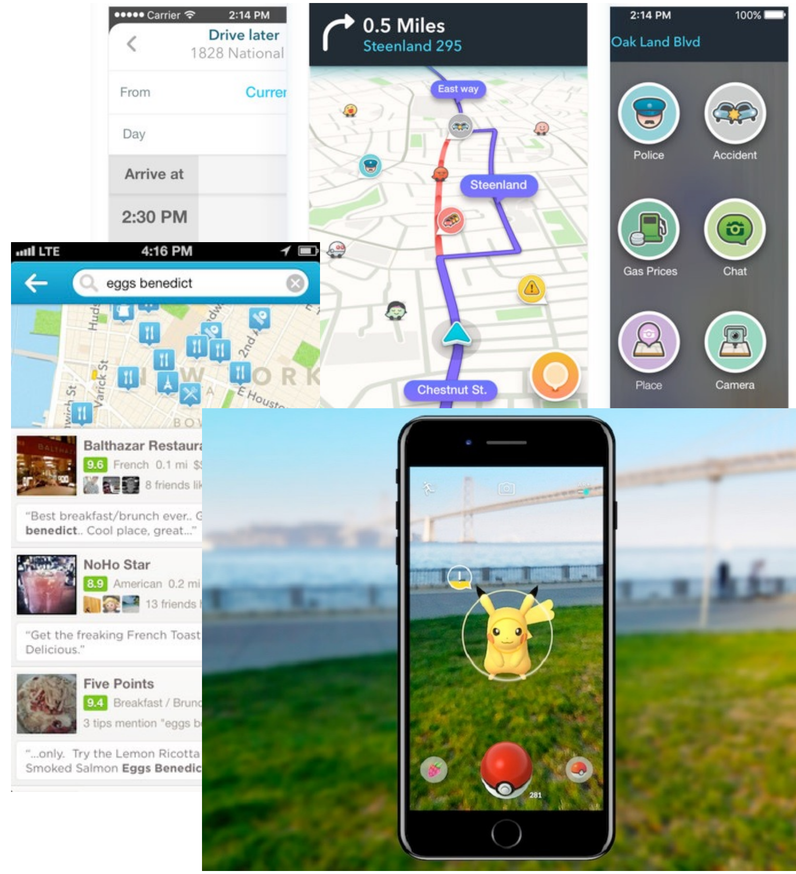
Master studies, Winter 2021/2022

Dr Veljko Pejović
Veljko.Pejovic@fri.uni-lj.si

University of Ljubljana
Faculty of Computer and Information Science

Location-Based Services

- Real-time navigation
- Geo-based alerts and emergency services
- Location-based recommendations
- Location-based online social networks
- Augmented reality and games
- High-level inferences/predictions: habits, health

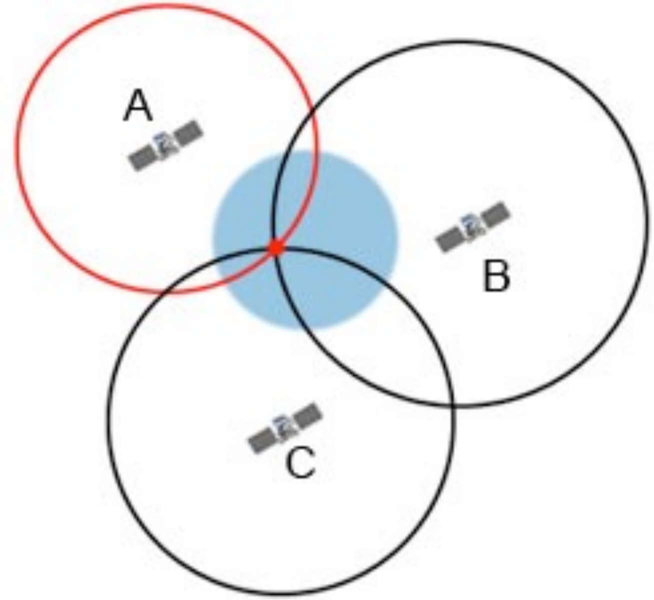


University of Ljubljana
Faculty of Computer and Information Science

Location Determination

- Methods:
  - Global navigation satellite systems (GNSS)
  - Proximity detection
  - RF modelling
  - RF fingerprinting
- Tradeoff among:
  - System complexity
  - Cost
  - Energy consumption
  - Applicability
  - Accuracy

GNSS



- Trilateration:
  - If we know the locations of A, B, and C, and a receiver's distance from each of the points, we can calculate the receiver's absolute position
- Knowing A, B, C's location
  - Satellites have well-specified orbits
- Knowing the distances to the receiver:
  - EM waves travel at $3*10^8$ m/s; measure the time the signal, sent from a satellite, took to reach the receiver

GNSS

- Satellites orbit around the Earth and trasmit signal that contains:
  - The satellite's almanach, from which the satellite's precise location can be found
  - The precise time the signal was sent
- Problem:
  - The receiver is not synchronized with the satellites!
- Solution:
  - The receiver adjusts its clock so that the calculated distances from four satellites intersect at one point

GNSS

- GNSS implementations:
  - GPS (USA)
  - GLONASS (Russia)
  - Galileo (EU)
  - BeiDou (China)
- GNSS limitations:
  - Line of sight needed
  - Realtively high energy requirements
  - Relatively poor accuracy
    4 meter RMSE for GPS

In practice solved with Kalman filtering

Kalman Filter

- A method to calculate more accurate estimates from noisy measurements
- Widely applicable in positioning, navigation, control, computer vision (for tracking), sensor data processing in general
- Very simplified explanation:
  - Prediction step: produce estimates of the current state along with the uncertainty
  - Update step: when the next measurement is observed, the estimate is updated using a weighted average, with weights distributed according to the certainty of the estimate and the measurement

Kalman Filter – More Detailed Explanation

Proximity Detection

- Fixed-position anchors with Rx/Tx; moving targets with Tx/Rx
- The target is assigned the location of the strongest signal anchor
- Active Badge (1992)
  - Infra-red technology
  - Beacon from a mobile Tx every 15 seconds
  - Location at the room level
  - Sunlight impacts performance

Want, R., Hopper, A., Falcao, V., & Gibbons, J. (1992). The active badge location system. *ACM Transactions on Information Systems (TOIS)*, *10*(1), 91-102.

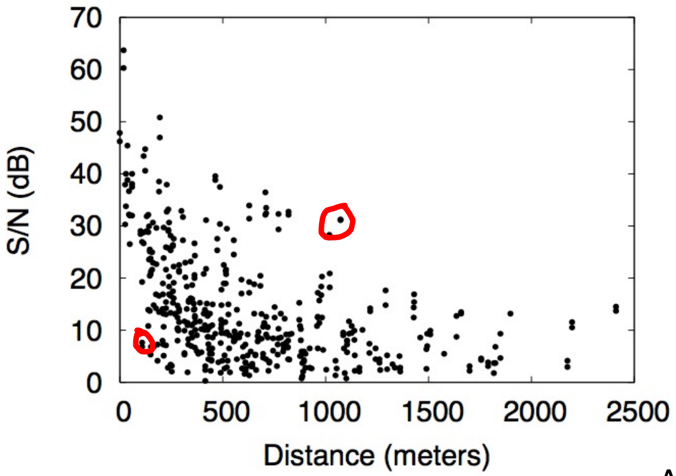
Proximity Detection

- Cricket (2000)
  - Idea: RF signal travels faster than sound
  - Fixed: RF and Ultrasound transmitters
  - Mobile: RF and Ultrasound receivers
  - When a mobile detects RF signal, it turns the sound receiver on, calculates the time difference of arrival
  - The mobile's location is set to the location of the fixed transmitter for which the shortest time difference between the RF and Ultrasound signal is observed
  - Drawback: Fitting the space and the mobile with extra Rx/Tx equipment

Priyantha, N. B., Chakraborty, A., & Balakrishnan, H. The cricket location-support system. ACM MobiCom 2000

RF Modelling

- Wireless signal properties taken into account:
  - RF signal propagation speed
  - Received signal power falls of with distance
    - Free space path loss (FSPL)
  - Signal phase changes with distance
    - With multiple antennae you can infer the angle of arrival
- Models often fail to capture real-world behaviour



Aguayo, Daniel, et al.
"Link-level measurements from an 802.11 b mesh network." *ACM SIGCOMM* 2004.

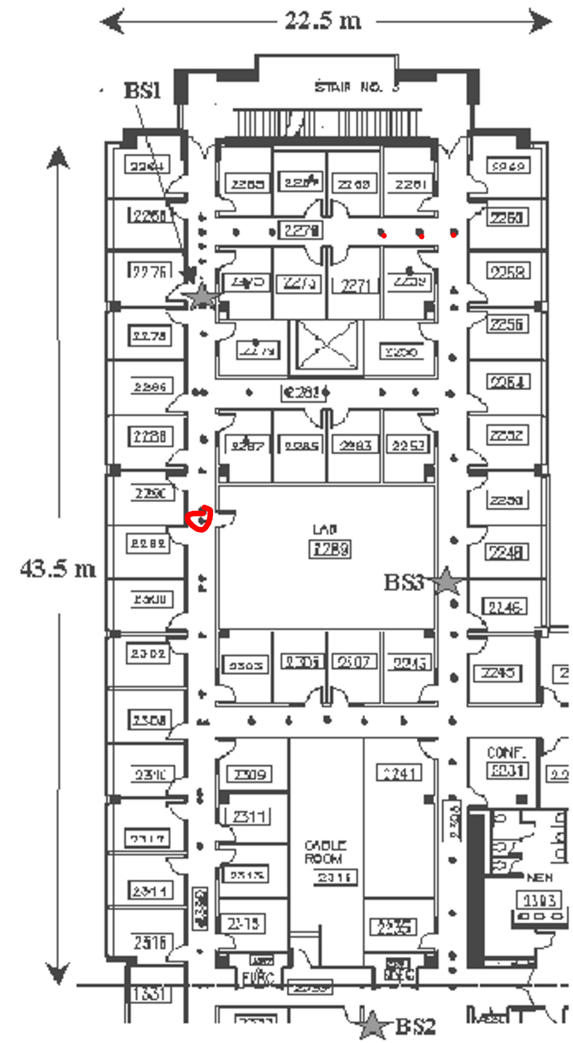
University *of Ljubljana*
Faculty *of Computer and Information Science*

RF Fingerprinting

- RF propagation between a transmitter and a receiver might be difficult to model
- But at least it might not vary drastically in time
- Idea:
  - Record the RF signal properties at different locations
  - Compare the real-time recorded RF signal with the recordings and find the most similar one – this determines the user's location

RF Fingerprinting

- RADAR (2001)
  - Later improved to take the signal propagation models into account
  - Drawbacks:
    - Need to visit each of the possible locations in order to do the fingerprinting
    - Different WiFi card might report different results
    - Change in the infrastructure placement might impact the results



Bahl, V., & Padmanabhan, V.
RADAR: An in-building RF-based user location and tracking system, INFOCOM 2000

University of Ljubljana
Faculty of Computer and Information Science

Location Determination - Conclusions -

- Precise location determination is still an open problem
  - Result accuracy of RADAR: 50% within 2.5 m, 90% within 5.9 m
  - More accurate solutions require more diverse and expensive hardware
- Further opportunities:
  - Increasing density of RF signals, including 60 GHz short-range communication
  - Restrictions of the real world: cars (usually) travel on roads, pedestrians move with a limited speed, etc.

University *of Ljubljana*
Faculty *of Computer and Information Science*

Location Determination
- Conclusions -

- In (current) practice
  - More accurate location determination requires more hardware, calculation, energy
    - Think about the accuracy level that your application needs!
    - In Android:
      - android.permission.ACCESS_COARSE_LOCATION
      - PRIORITY_LOW_POWER for your LocationRequest
  - Indoor/custom localization can be implemented via trilateration
    - Android 9 and 10 support roundtrip time (RTT) querying from a mobile to a supported WiFi AP https://developer.android.com/guide/topics/connectivity/wifi-rtt

Location Prediction
- Applications -

- Cell tower association prediction for handover management in mobile networks
- Significant places recognition
- Smart home adaptation (e.g. MavHome project light, heating)
- Encounter prediction
- Routine deviation

Location Prediction - Approaches-

- What to predict?
  - Absolute location (latitude, longitude)
  - Semantic location (home, work, restaurant)
- How far in future to predict?
  - Next hour
  - Next place that a user will visit
  - Next year
- For whom to predict?
  - Individual
  - Group (e.g. encounter with OSN friends)

Something in between (a particular restaurant, etc.)

University of Ljubljana
Faculty of Computer and Information Science

Next Location Prediction

- What is next location?
- Depends on your data:
  - GPS traces:
    extract significant locations using clustering
    - DBSCAN algorithm
  - Access point (AP) associations:
    AP BSSID as location e.g. “at Huawei-01234”
  - Service-provided locations:
    Foursquare gives you nice semantic location
    e.g. “at Kinodvor theatre”

Next Location Prediction

- Dartmouth University dataset:
  - AP associations for 6,000 users over a few years on a college campus
  - Available (together with many other datasets) on https://crawdad.org
- Prediction problem:
  - Next AP association

Song, L., Kotz, D., Jain, R., & He, X. (2006). Evaluating next-cell predictors with extensive Wi-Fi mobility data. IEEE transactions on mobile computing, 5(12), 1633-1649.

Next Location Prediction

- Prediction method: Markov Chain
  - A stochastic model describing a sequence of possible events in which the probability of an event depends only on the previous event
  - Markov property

$$P(X_{n+1} = a \mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) =$$
$$P(X_{n+1} = a \mid X_n = x_n)$$

Does next location depend only on the previous location?

Next Location Prediction

- Higher order (*k*) Markov Chain
  - The probability of an event depends only on the last *k* previous events

$$P(X_{n+1} = a \mid X_1 = x_1, X_2 = x_2, .., X_n = x_n) =$$
$$P(X_{n+1} = a \mid X_{n-k+1} = x_{n-k+1}, X_{n-k+2} = x_{n-k+2}, .., X_n = x_n)$$

  - Estimate the above probability from the data trace:

$$P(X_{n+1} = a \mid X_{n-k+1} = x_{n-k+1}, X_{n-k+2} = x_{n-k+2}, .., X_n = x_n) = \frac{N(ca, L)}{N(c, L)}$$

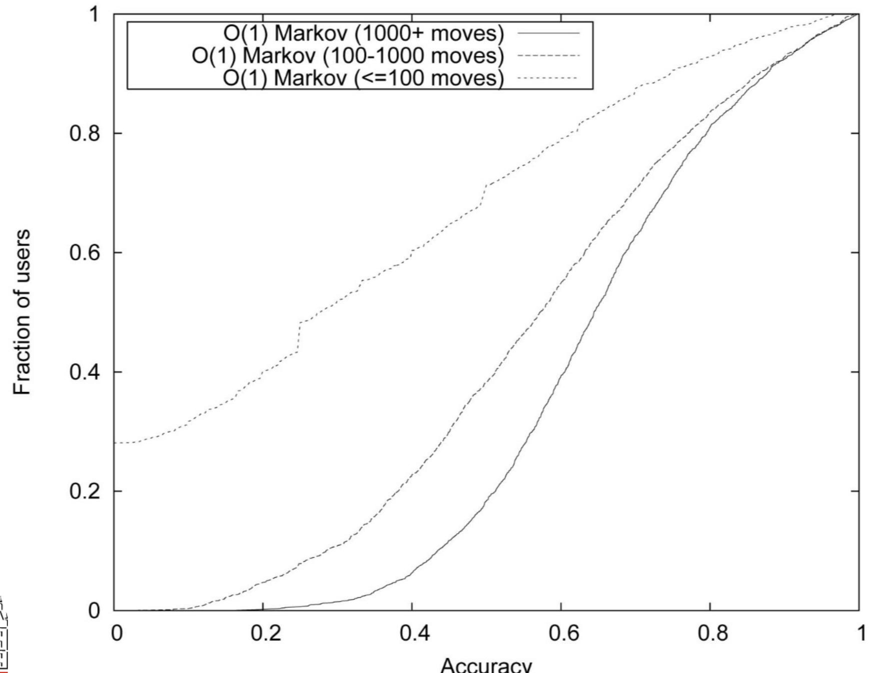
$N(c, L)$ - Num of times a sequence c appears in the data trace L

University of Ljubljana
Faculty of Computer and Information Science

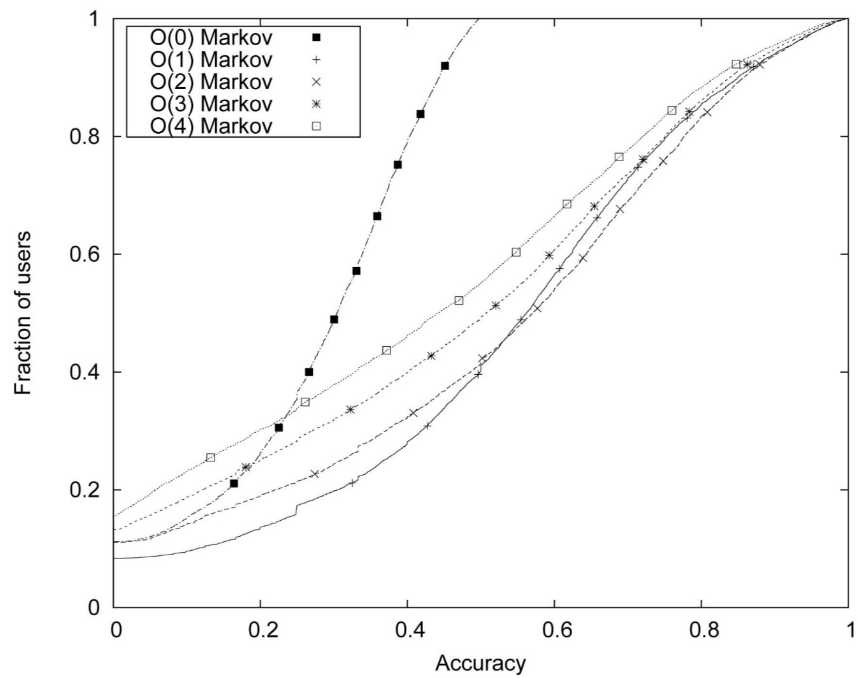
Next Location Prediction

- Evaluation

How much data do we need? A lot!



How far in the past do we need to look? Not far at all!



University of Ljubljana
Faculty of Computer and Information Science

Next Location Prediction

- Foursquare dataset:
  - 1 million users
  - 35 million check-ins
  - 5 million venues
- Prediction problem:
  - Predict a list of N possible venues where the user will check in next
  - Limit the problem to a city (i.e. not trying to predict if a user is going to check in when travelling)

Mining User Mobility Features for Next Place Prediction in Location-based Services
A. Noulas, S. Scellato, N. Lathia, C. Mascolo. In IEEE International Conference on Data Mining. Brussels, Belgium. December 10-13, 2012.

Next Location Prediction

- Prediction method: ranking scores – r
  - *t' – current prediction time*
  - *C_u – user u's check-ins (loc, time)*
  - *k - venue*
  - Historical visits rank:

$$\hat{r}_k(u) = |\{(l,t) \in C_u : t < t' \wedge l = k\}|$$

  - Category preferences:
    z – venue category

$$\hat{r}_k(u) = |\{(l,t) \in C_u : t < t' \wedge z_l = z_k\}|$$

  - Popularity:

$$\hat{r}_k(U) = \sum_{u \in U} |\{(l,t) \in C_u : t < t' \wedge l = k\}|$$

  - Popularity among friends:
    Γ_u – u's friends

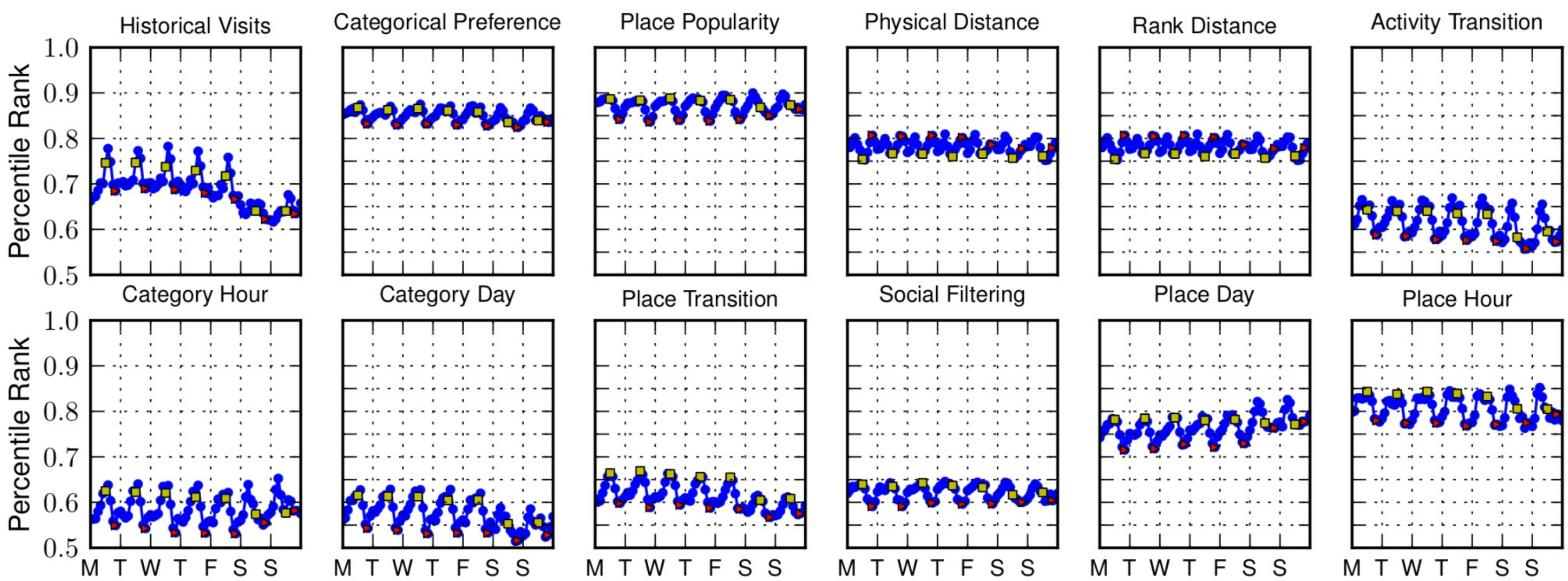
$$\hat{r}_k(u) = \sum_{v \in \Gamma_u} |\{(l,t) \in C_v : t < t' \wedge l = k\}|$$

University of Ljubljana
Faculty of Computer and Information Science

…also distance, temporal rankings, etc.

Next Location Prediction

- Evaluation

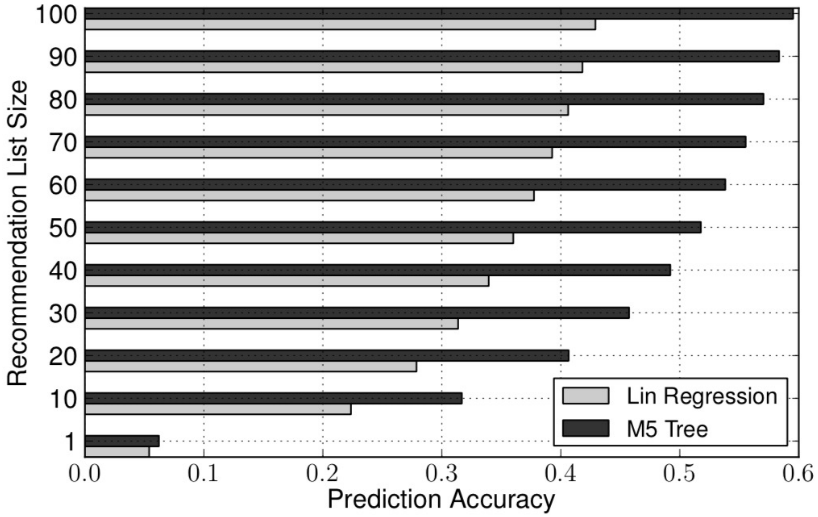


…but, what if multiple factors impact one's decision to visit a venue?

University *of Ljubljana*
Faculty *of Computer and Information Science*

Next Location Prediction

- Prediction method: supervised learning from the venue/visit properties
  - Features:
    - popularity
    - distance
    - temporal activity score
  - Label: positive when visited, randomly sample other locations and assign negative labels to them
  - Classifiers:
    - Tree-based and Linear regression



University of Ljubljana
Faculty of Computer and Information Science

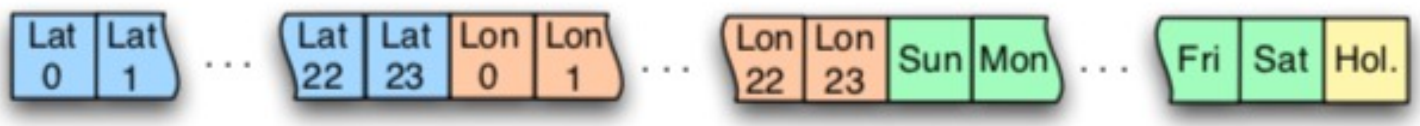
Long-Term Prediction

- Dataset:
  - GPS traces of people and vehicles
  - 703 users
  - From 7 to 1247 days per user
- Prediction method: Motif Identification
  - Identify regularities in user's movements, model the evolution over time, and estimate future locations by projecting the patterns into the future
  - Assume that people's exhibit 24h-periodic behaviour, yet, that this can differ based on the day of the week, holidays

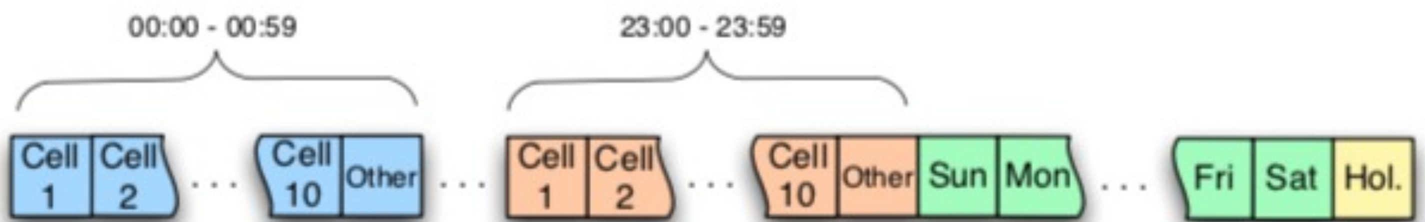
Sadilek, Adam, and John Krumm. "Far out: Predicting long-term human mobility", AAAI 2012

Long-Term Prediction

- Represent data as vectors
  - Each day is a vector
  - lat, lon represent mean values in the given hour



  - Alternatively, divide space into cells; fill with a probability that a user is in the cell at the given hour

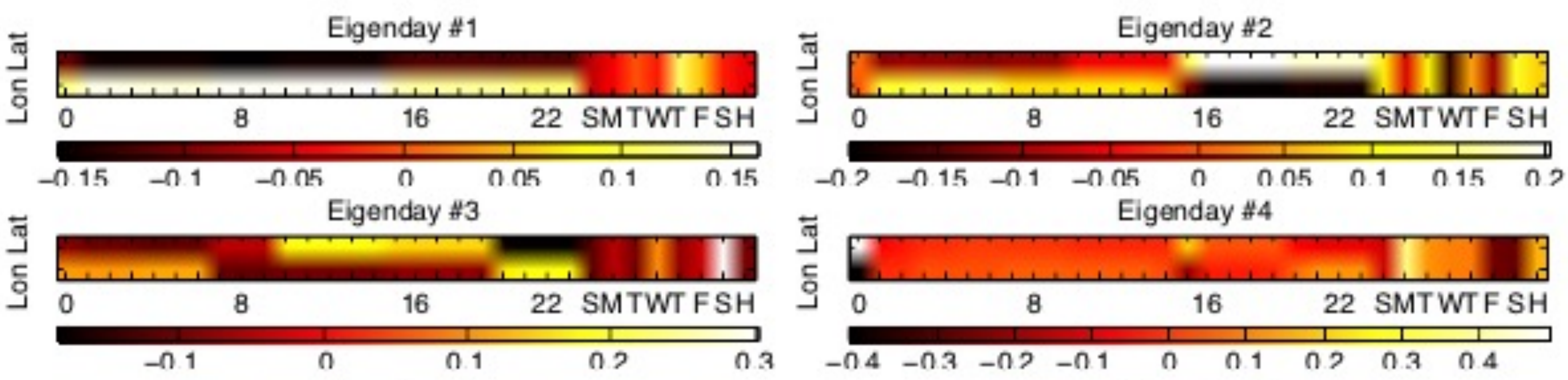


University of Ljubljana
Faculty of Computer and Information Science

Long-Term Prediction

- Now we have a matrix of vectorised days
- Perform Principal Component Analysis (PCA) to identify the essential behaviour
- PCA Eigenvectors represent characteristic days (eigendays)

What's PCA?

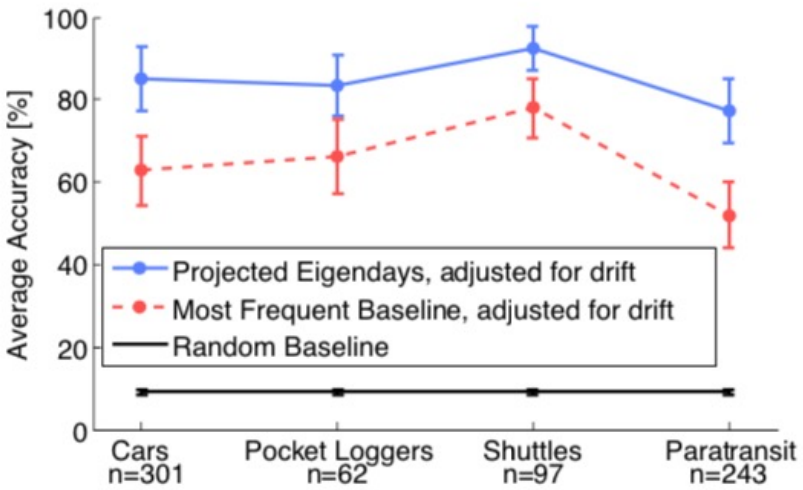


Long-Term Prediction

- Prediction:
  - Extract the observed features of the target, i.e. day of the week, whether it is a holiday or not and pack them into a vector
  - Project the above vector onto the PCA space to get a weight for each eigenday that captures how dominant that eigneday is, given the observed feature values

Long-Term Prediction

- Evaluation
  - Comparison with a method that predicts the most frequent location prev. visited with the given feature values



- Other uses
  - Find deviations from the routine – if a new day deviates from the subspace formed by historical eigendays

Detect if someone stole the user's phone!

University *of Ljubljana*
Faculty *of Computer and Information Science*

Location Prediction
- Conclusions -

- Both short-term and long-term predictions are very powerful enablers of new apps and services
  - Health care, security, smart cities, etc.
- Usually, different methods for long-term and short-term mobility prediction
- For other methods, start with:
  - Pejovic, Veljko, and Mirco Musolesi. "Anticipatory mobile computing: A survey of the state of the art and research challenges." *ACM Computing Surveys (CSUR)* 47.3 (2015): 47.
- Remember: Data is the king!