DES535
Ubiquitous Computing

Dr. Pragma Kar
Assistant Professor
Department of Human-Centered Design
INDRAPRASTHA INSTITUTE *of* INFORMATION TECHNOLOGY **DELHI**

Google Classroom Code : pcwnf5t

Location Sensing

Module IV (Part II)

Queries to the location models

- In order to derive requirements on location models various forms of queries to location models from the perspective of users and applications are required.
- Types of queries:
    - Position queries
    - Nearest neighbor queries
    - Range queries

https://link.springer.com/article/10.1007/s00779-004-0270-2

Position queries

- The determination of the positions of mobile and static objects like users, buildings, bus stops, etc. is a common building block of location-based and context-aware systems. Therefore, all location models contain this information, but they differ in the way it is represented.
- The definition of a position requires some form of coordinates.
- Based on an object's position, actions can be carried out, such as teleporting the user's interface, controlling the input and output of applications to arbitrary spaces in the physical environment via projection techniques, or in industrial settings, such as a smart factory, the positions of resources and tools can be monitored in a production planning system.
- This shows that a general location model has to support different coordinate reference systems, **global and local ones**.
- Beside well-known geometric coordinates, some positioning systems provide symbolic coordinates.

https://link.springer.com/article/10.1007/s00779-004-0270-2

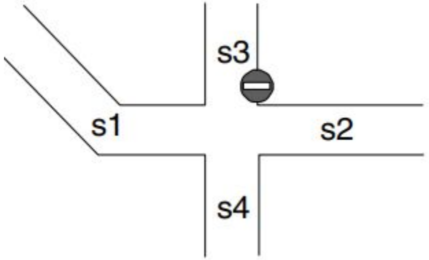
Nearest neighbor queries

- A nearest neighbor query is the search for the n objects closest to a certain position.
- Beside known object positions, the definition of a distance function on the coordinates is required for this type of query.
  - For geometric coordinates, the direct physical distance between two positions can be calculated using well-known formulas like Pythagoras in Cartesian systems.
  - If only symbolic coordinates are modeled, then the model must contain explicit definitions of distances between these coordinates, e.g., to define the distance between room number X and the printers in the rooms number Y and Z, since symbolic coordinates do not contain a natural embedment into a metric space.
- There are other notions of distance that are often more relevant than the direct physical distance.
  - For instance, for a pedestrian, it might be impossible to cross a highway. Therefore, a restaurant across the highway with a direct physical distance of 100 m might be farther away than a restaurant with 200 m direct physical distance not located across this highway. In these cases, additional model information like the road network a user uses to get from location A to B has to be taken into account.

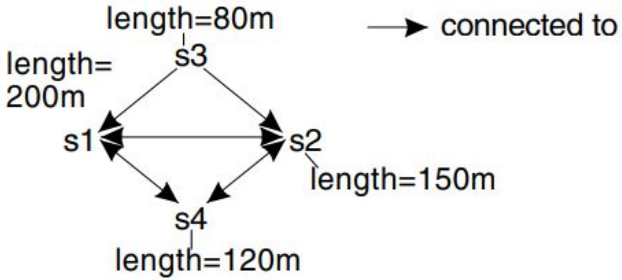
https://link.springer.com/article/10.1007/s00779-004-0270-2

Navigation

- Navigation systems have become standard equipment in many modern cars. Such systems require a location model to find paths between locations. Possible paths are defined by the transportation network (roads, train or bus routes, etc.) and consist of several interconnected locations.
- It does not suffice to know the geometry, e.g., of roads, but it is also important to know how to get from one location to neighboring locations, e.g., from one road segment to another road segment at a junction, and finally, to the destination.



Road geometry



Road topology

https://link.springer.com/article/10.1007/s00779-004-0270-2

Range queries

- A range query returns all objects within a certain geographic area. It can be used, for instance, to query the occupancy of a room as well as to check whether an evacuation plan is processed correctly, i.e., if a room is empty before the fire doors are closed and sealed.
- Also, simple algorithms for new types of communication can be implemented on the basis of range queries, e.g., geocast, i.e., the sending of messages to receivers in a certain geographic area. First, a range query can be used to determine all receivers in the target area of the message. Second, the message is sent to these receivers, e.g., using multiple unicast messages.
- First of all, object positions have to be known to answer a range query. Additionally, the topological relation ‘‘contains’’ has to be modeled, i.e., it has to be defined whether a coordinate lies within a spatial area.
  - For geometric coordinates, this information can be derived from the known geometry. But for symbolic coordinates, this relation has to be defined explicitly.
- For instance, a model can define that the room 2.062 is on (‘‘within’’) the second floor that, in turn, is part of (‘‘within’’) a certain building, etc. Thus, querying for a larger area automatically includes all objects from locations that lie within that area.

https://link.springer.com/article/10.1007/s00779-004-0270-2

Requirements for location models

- **Object positions:** Positions of objects have to be modeled in the form of coordinates. Supported coordinates and reference systems are:
  - • – Geometric and symbolic coordinates
  - • – Multiple local and global coordinate reference systems
- **Distance function:** Distances between spatial objects have to be modeled. This can also be the “size” of a location, e.g., the length of a road segment, which represents the distance one has to travel when crossing this location in order to reach another location.
- **Topological relations:** The following topological relations between spatial objects have to be modeled:
  - • – Spatial containment in order to allow range queries
  - • – Spatially connected to for navigation services
- **Orientation**: In addition to the positions of mobile objects, the orientation in the horizontal and/or vertical dimensions can be supported.

https://link.springer.com/article/10.1007/s00779-004-0270-2

Modeling Effort

- **Accuracy**: The model should describe the real world as accurately as possible, i.e., the stored information should be consistent with the real world. Accuracy is not a question of the model type, but of how the model is created and updated, and of the dynamics of the modeled objects: highly dynamic objects require high update rates, e.g., highly mobile objects will have to update their position frequently to get accurate position information.
- **Level of detail**: The level of detail describes the precision or granularity of the model. Fine-grained models describe locations down to room level or below; coarse-grained models stop at buildings or larger. A flexible model allows both ends of the scale.
- **Scope**: The scope is the area covered by the model. Local models may only describe one single room, whereas global models, at the other end of the scale, describe locations all over the world.

The requirements have to be regarded in conjunction with the requirement of minimal modeling effort.

https://link.springer.com/article/10.1007/s00779-004-0270-2

Dataset

GeoLife (Microsoft Research Asia)

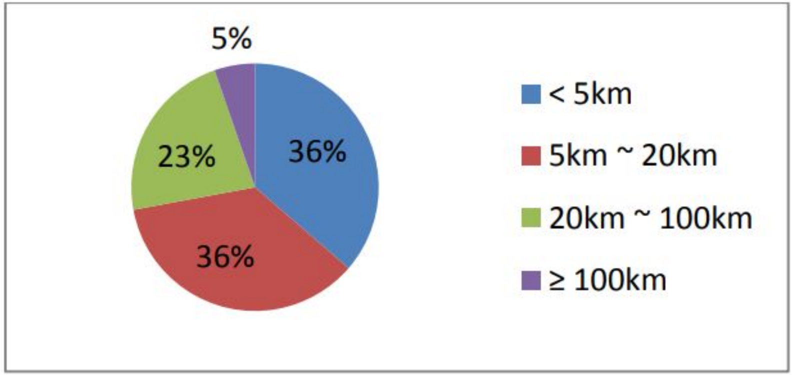
- This GPS trajectory dataset was collected in (Microsoft Research Asia) Geolife project by 178 users in a period of over four years (from April 2007 to October 2011).
- A GPS trajectory of this dataset is represented by a sequence of time-stamped points, each of which contains the information of latitude, longitude and altitude.
- This dataset contains 17,621 trajectories with a total distance of 1,251,654 kilometers and a total duration of 48,203 hours.
- These trajectories were recorded by different GPS loggers and GPSphones, and have a variety of sampling rates. 91 percent of the trajectories are logged in a dense representation, e.g. every 1~5 seconds or every 5~10 meters per point.

https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/

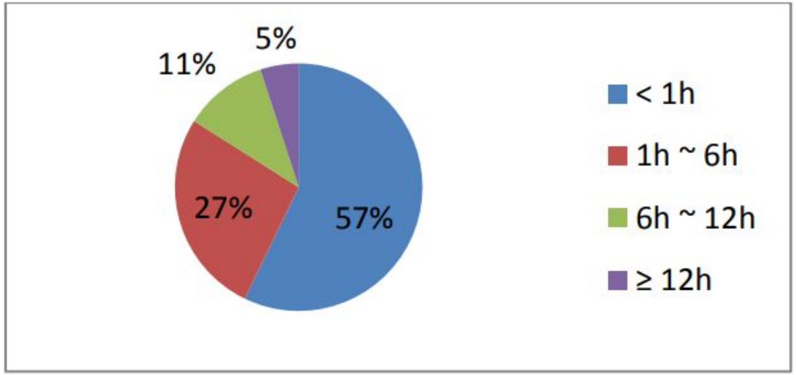
Dataset

GeoLife (Microsoft Research Asia)

- This trajectory dataset can be used in many research fields, such as mobility pattern mining, user activity recognition, location-based social networks, location privacy, and location recommendation.



Distribution of trajectories by distance



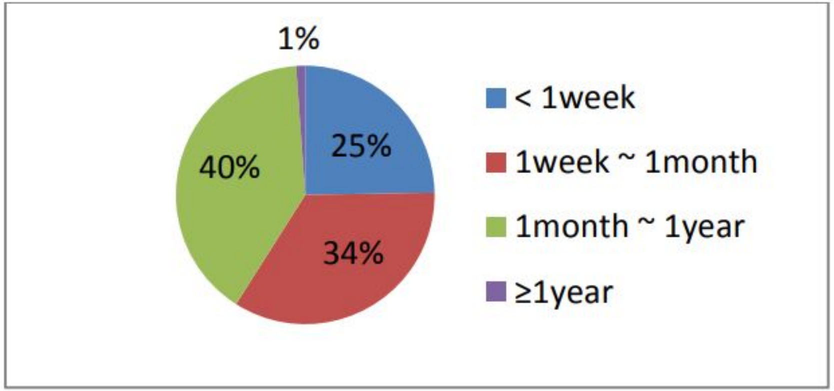
Distribution of trajectories by effective duration

https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/

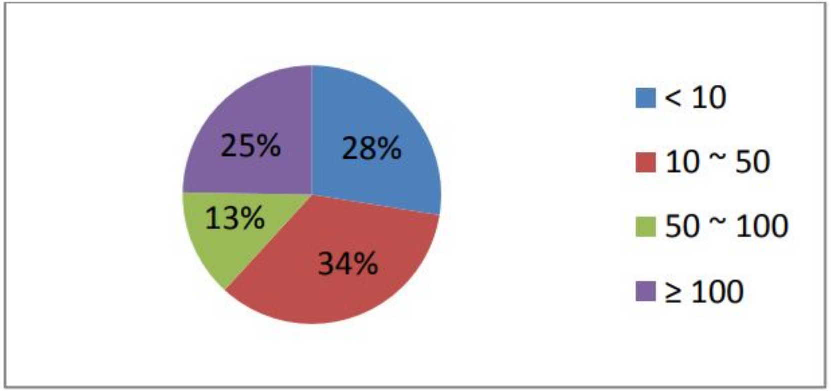
Dataset

GeoLife (Microsoft Research Asia)

- 69 users have labeled their trajectories with transportation mode, such as driving, taking a bus, riding a bike and walking. There is a label file storing the transportation mode labels in each user’s folder



Distribution of users by data collection period



Distribution of users by trajectories

https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/

Dataset

GeoLife (Microsoft Research Asia)

- Line 1…6 are useless in this dataset, and can be ignored. Points are described in following lines, one for each line.
- Field 1: Latitude in decimal degrees.
- Field 2: Longitude in decimal degrees.
- Field 3: All set to 0 for this dataset.
- Field 4: Altitude in feet (-777 if not valid).
- Field 5: Date - number of days (with fractional part) that have passed since 12/30/1899.
- Field 6: Date as a string.
- Field 7: Time as a string.

```
39.999749,116.321866,0,151,39993.3441550926,2009-06-29,08:15:35
39.999952,116.321858,0,145,39993.344212963,2009-06-29,08:15:40
40.000145,116.321818,0,150,39993.3442708333,2009-06-29,08:15:45
40.000375,116.321809,0,134,39993.3443287037,2009-06-29,08:15:50
```


https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/

Dataset

GeoLife (Microsoft Research Asia)

- Line 1…6 are useless in this dataset, and can be ignored. Points are described in following lines, one for each line.
- Field 1: Latitude in decimal degrees.
- Field 2: Longitude in decimal degrees.
- Field 3: All set to 0 for this dataset.
- Field 4: Altitude in feet (-777 if not valid).
- Field 5: Date - number of days (with fractional part) that have passed since 12/30/1899.
- Field 6: Date as a string.
- Field 7: Time as a string.

```
39.999749,116.321866,0,151,39993.3441550926,2009-06-29,08:15:35
39.999952,116.321858,0,145,39993.344212963,2009-06-29,08:15:40
40.000145,116.321818,0,150,39993.3442708333,2009-06-29,08:15:45
40.000375,116.321809,0,134,39993.3443287037,2009-06-29,08:15:50
```

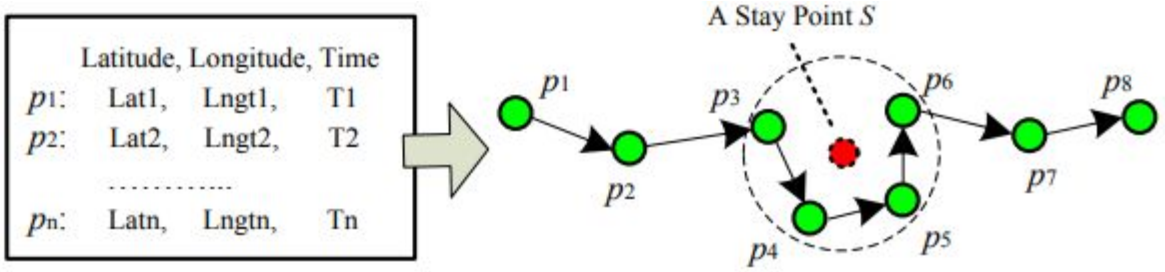

https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/

Usage of GeoLife: Mining Interesting Locations and Travel Sequences from GPS Trajectories

- Typically, people would desire to know which locations are the most **interesting places** in a geospatial region.
- Further, given these interesting locations in a geospatial region like a city, users might also wonder what the most **classical travel sequences** are among them.

https://dl.acm.org/doi/pdf/10.1145/1526709.1526816?casa_token=fGeySXvYYNYAAAAA:pO-lbvLKW-sXezCIVkPFFnywFOstVW09sWiEMtQsoSTFbumd230cXZllKrQ2H38hDS mCwXMpWs8

Usage of GeoLife: Mining Interesting Locations and Travel Sequences from GPS Trajectories

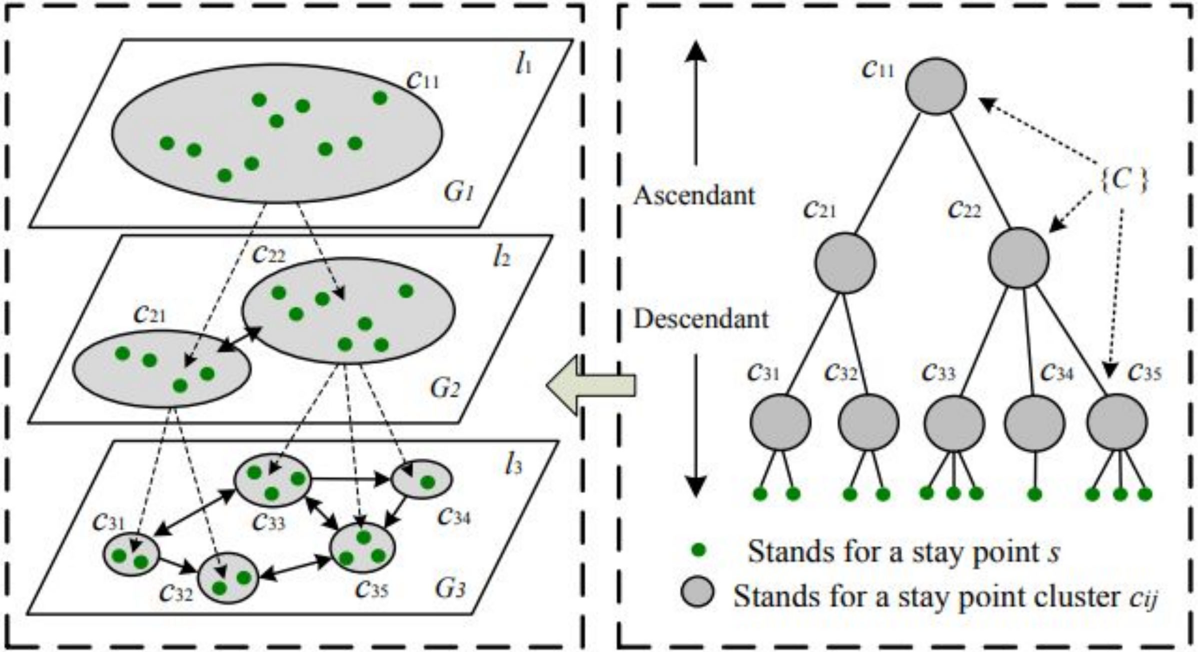


- **GPS log** is a collection of GPS points P={p1 , p2 , … , pn}
- On a two dimensional plane, we can sequentially connect these GPS points into a curve based on their time serials, and split this curve into **GPS trajectories**
- A **stay point** (S) stands for a geographic region where a user stayed over a certain time interval. The extraction of a stay point depends on two scale parameters, 4
    - a time threshold and
    - a distance threshold

https://dl.acm.org/doi/pdf/10.1145/1526709.1526816?casa_token=fGeySXvYYNYAAAAA:pO-lbvLKW-sXezCIVkPFFnywFOstVW09sWiEMtQsoSTFbumd230cXZlIKrQ2H38hDSmCwXMpWs8

Usage of GeoLife: Mining Interesting Locations and Travel Sequences from GPS Trajectories

- A **location history** is a record of locations that an entity visited in geographical spaces.
- A **Tree-based hierarchical graph** is a collection of stay point-based clusters C with a hierarchy structure L



https://dl.acm.org/doi/pdf/10.1145/1526709.1526816?casa_token=fGeySXvYYNYAAAAA:pO-lbvLKW-sXezClVkPFFnywFOstVW09sWiEMtQsoSTFbumd230cXZllKrQ2H38hDSmCwXMpWs8

Further reading

- https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4653466
- https://link.springer.com/content/pdf/10.1007/s40558-017-0092-5.pdf
- https://openaccess.thecvf.com/content_CVPR_2020/papers/Weyand_Google_Landmarks_Dataset_v2_-_A_Large-Scale_Benchmark_for_Instance-Level_CVPR_2020_paper.pdf
- https://dl.acm.org/doi/abs/10.1145/1526709.1526816