Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking

Marco Gruteser, Dirk Grunwald

Derya Erkan

Paper propose

"New technologies can pinpoint your location at any time and place. They promise safety and convenience but threaten privacy and security"

Cover story, IEEE Spectrum, July 2003

Outline

- Accuracy Requirements of Location-Based Telematics Services
  - System Assumptions
  - Scenarios
- Privacy Threats Through Location Information
- Solution
  - k-Anonymous Location Information
- Implementation
- Results
- Analysis
- Conclusion

Accuracy Requirements of Location-Based Telematics Services

- **The paper tries to achieve reach the E-911 requirements which is to be able to estimate the caller's position with an accuracy of 125 m (RMS) in 67 percent of cases.**
- **The system assumes that the clients should be able to provide the location with very high precision.**

Accuracy Requirements Scenarios

- **Driving Conditions Monitoring**
- **Road Hazard Detection**
- **Road Map**

Service	Position Accuracy	Time Accuracy	Frequency of Access
Driving Conditions Monitoring	100 meters	minutes	continuous
Road Hazard Detection	10 meters	> days	sporadic
Road Map	100 meters	sub-second	sporadic

Table 1: Approximate accuracy requirements of telematics services

Privacy Threats Through Location Information

- Restricted Space Identification.

  If A knows that space L exclusively belongs to subject S then A learns that S is in L and S has sent M.

- Observation Identification.

  If A has observed the current location L of subject S and finds a message M from L then A learns that S has sent M.

- Location Tracking.

  Location Tracking. If A has identified subject S at location Li and can link series of location updates L1;L2;...;Li;...;Ln to the subject, then A learns that S visited all locations in the series.

LBS

A: adversary

M: message that contains L

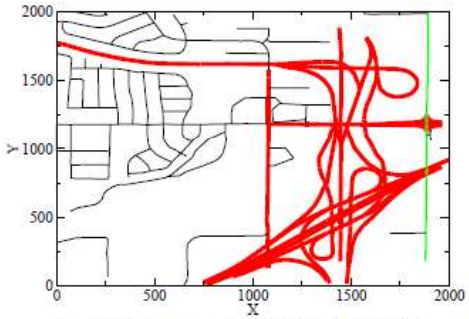
S: subject who located in L

k-Anonymous Location Information

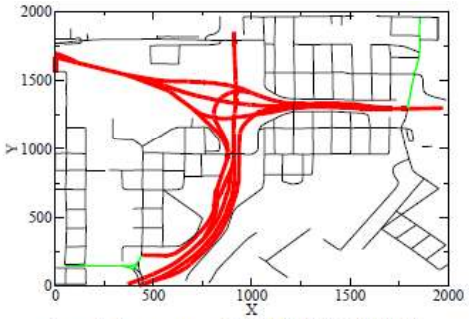
- Anonymity is etymologically defined as being nameless or of unknown authorship, information privacy researchers interpret it in a stronger sense.
- K-Anonymous: We consider a subject as k-anonymous with respect to location information, if and only if the location information presented is indistinguishable from the location information of at least k -1.

Implementation

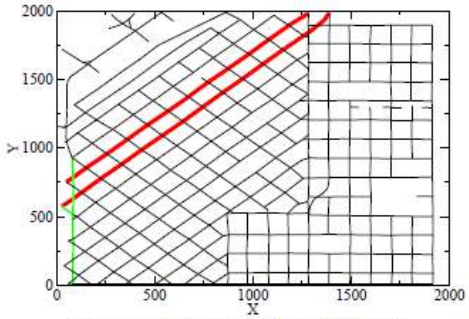
- The paper, uses automotive traffic simulations based on US geological survey (USGS) cartographic material.



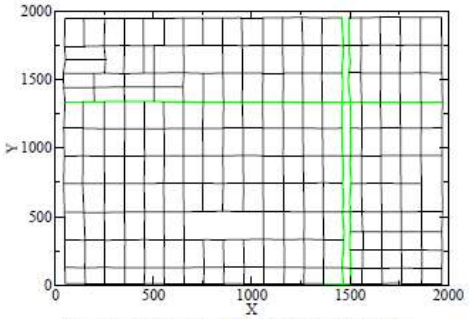
Area 1: Expressways (X:500000, Y:4407000)



Area 2: Expressways (X:500000, Y:4402000)



Area 3: Collectors (X:501000,Y:4400000)



Area 4: Collectors (X:506000,Y:4400000)

Implementation

- A traffic study reports the 24 hour traffic volume at specific points along roads.

USGS Class	Road Type	Traffic Volume
1	Expressway	70000
2	Arterial	22000
3	Collector	6000

Table 3: Mapping of Traffic Study volumes to road classes from USGS data

Implementation

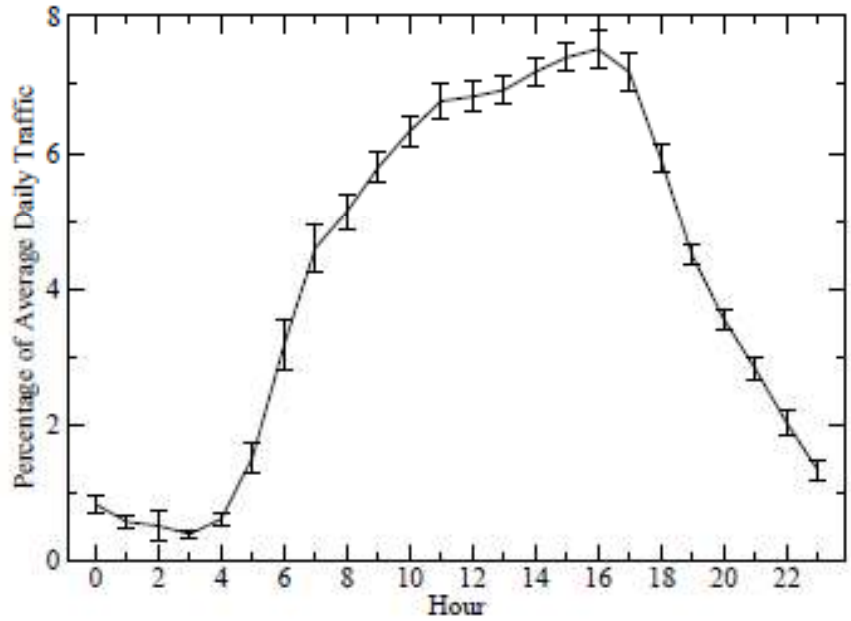


Figure 2: Hourly traffic volume relative to daily traffic volume during a typical August 2002 day

Implementation

- Data to be used is generated from the average traffic count for each and every hour , and for each segment using the traffic information available for the city of Denver, Colorado using an hour adjustment factor.

$$n = \frac{l \times c \times h}{v}$$

with traffic count c, hour adjustment h, vehicle velocity v, length of a road segment l.

Implementation by spatial discretization

An area is chosen around the requester of LBS, such that there are at least k min vehicles in the region. This is achieved using **quadtree** algorithms.

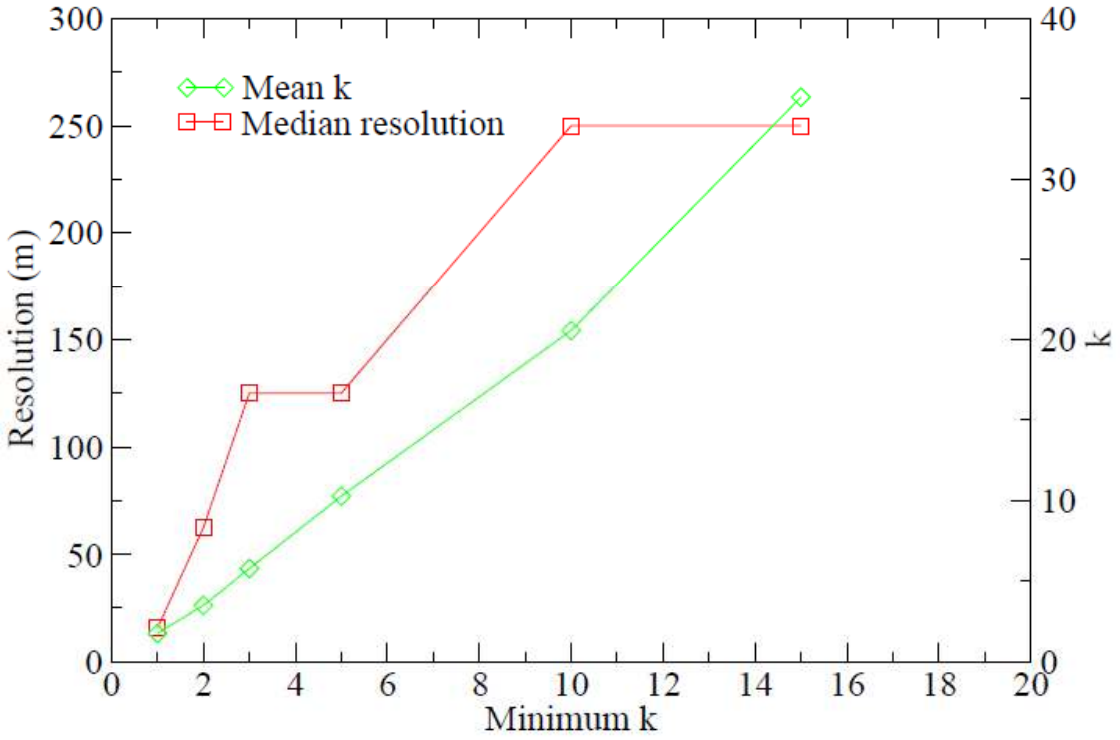


Figure 6: Dependency of spatial resolution and mean anonymity on anonymity constraint. This figure illustrates how spatial resolution (left scale) and mean actual anonymity (right scale) vary with different anonymity constraints (x-axis).

Implementation by temporal discretization

Similar to the previous algorithm but, spatial accuracy is maintained by delaying the service to be processed until, kmin different vehicles have visited the area.

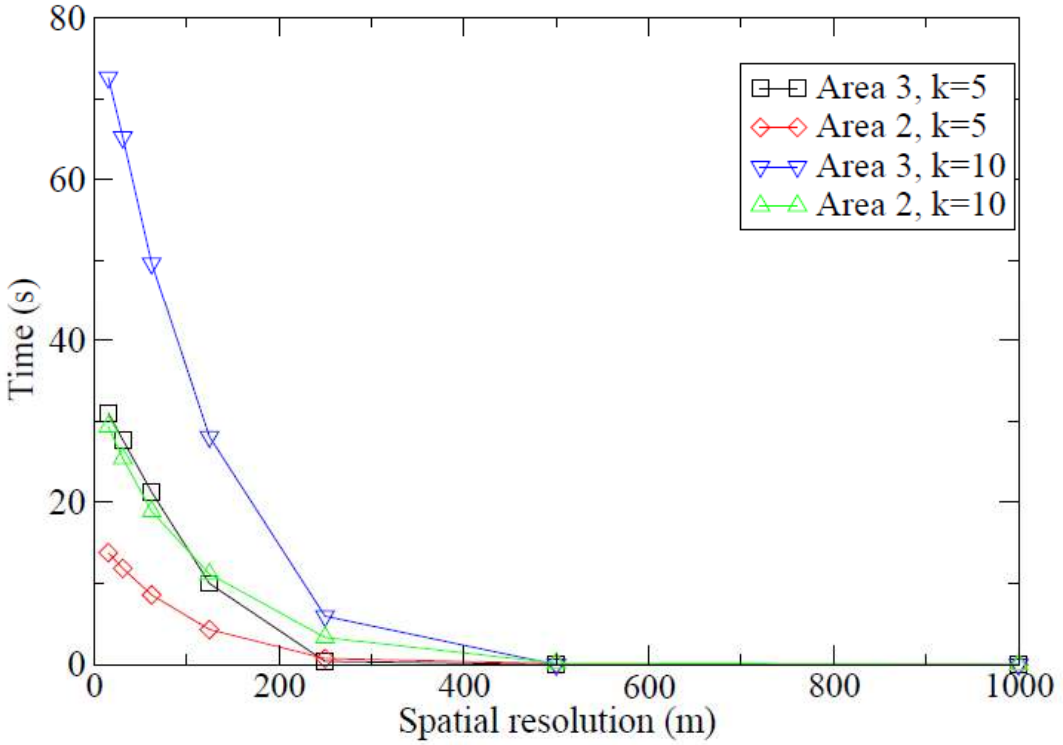


Figure 7: Tradeoff between temporal and spatial resolution The figure shows the mean reduction in temporal resolution necessary to reach a specified spatial resolution. The tradeoff is shown for a highway (2) and a collector (3) area at different anonymity constraints

Results

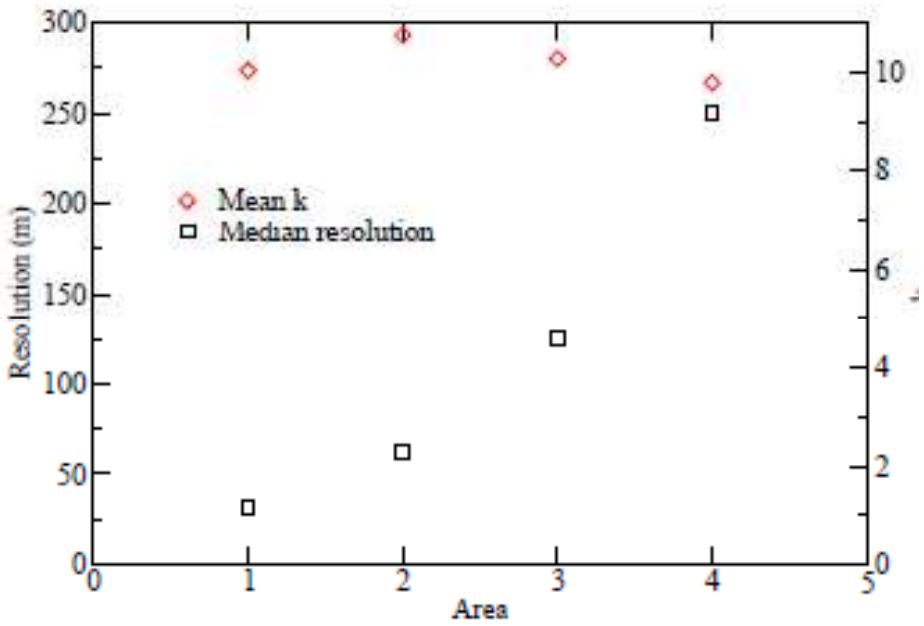


Figure 3: Dependency of spatial resolution and mean anonymity on area characteristics. For each evaluation area, the figure shows the median resolution from a large number of requests (left y-axis scale) and the mean *actual* anonymity—the number of subjects indistinguishable from the requestor (right y-axis scale).

Results

Accuracy: While in the highway area less than 10% of requests reach a resolution poor than 125 meters, it is about 60% for the collector street

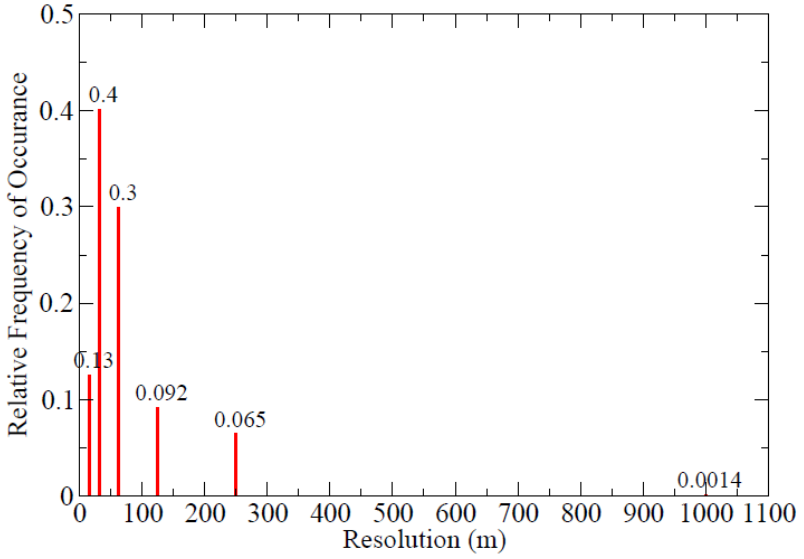


Figure 4: Relative frequency of spatial resolution for highway area (1). This figure illustrates the distribution of resolutions over a large number of simulated requests.

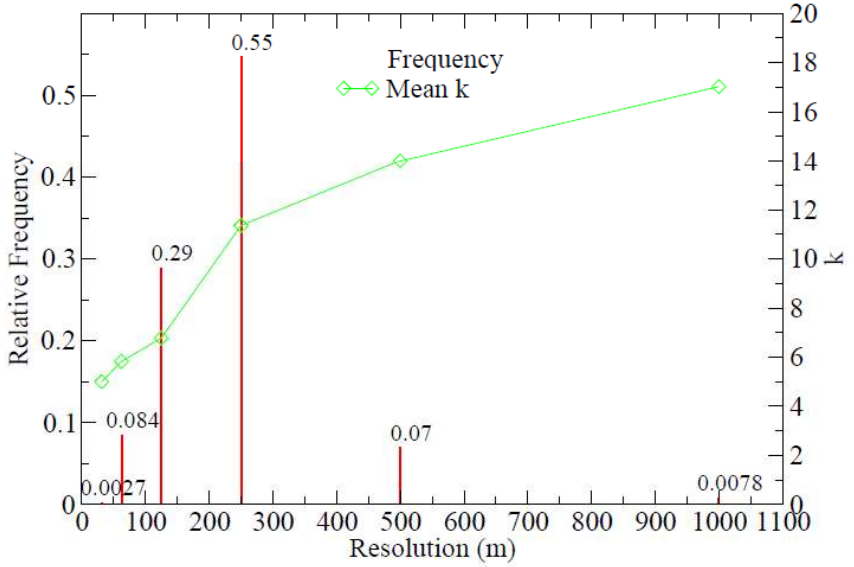


Figure 5: Relative frequency of spatial resolution for collector road area (4). In addition to the distribution of esolutions (left y-axis scale), the figure shows the mean actual anonymity at each resolution (right y-axis scale).

Analysis

1. Accuracy: The accuracy in case of collector street(4) where , 60% of requests reach a resolution poor than 125 meters, does not meet the requirements.

2. The paper suggest to adjust the spatial resolution by adjusting the temporal resolution. This works only in case of situations where services doesn't require quick response.

3. Driving conditions monitoring and road hazard detection are well served by this approach.

Analysis

Security:

1. Data transfer using authentication channels between the anonymity server and location client along with timestamps, avoid eaves-droppers from listening or replaying the coordinate information.

2. Consider the case when adversary tries to spoof virtual vehicles to make the anonymity server believe kmin has been reached. This can be avoided by authenticating the vehicles.

Analysis

Anonymity:

Given no other information the re-identification risk is 1/k. But in the figure 8 on the right, if Kmin is 3 the quadrant returned for 1,2,3 nodes are the single quadrant containing the node. But in case of node 4, the area returned is all four quadrants together. Using this, we can easily figure out that the request is from a node which is in the range(1,1) and (2,2).

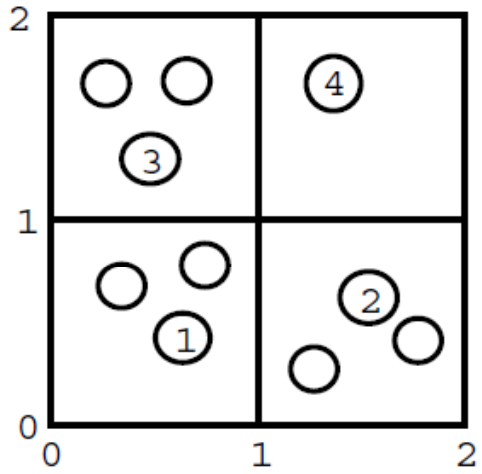


Figure 8: Compromised anonymity through overlapping requests. The circles and squares represent subjects and the quadrants computed by the cloaking algorithm, respectively. If each numbered subject requests a service simultaneously, subject 4 could be identified.

Conclusion

- The quadtree-based algorithm reached accuracy levels comparable to the phase II E-911 requirements, and thus should be suitable for many location-based services.
- In areas with major highways the median accuracy is approximately 30m and increases to 250m for city areas with large block sizes. These results were obtained with an anonymity constraint of 5, yielding a mean anonymity level of approximately 10 people who may have issued a particular request.
- Spatial resolution can be significantly improved through a several seconds reduction in temporal resolution. Because of the imposed delay, this method is most applicable to non-interactive services.

Questions

?