Privacy Implications of Automated GPS Tracking and Profiling

Muhammad Usman Iqbal

Samsung Lim

The University of New South Wales

School of Surveying & Spatial Information Systems
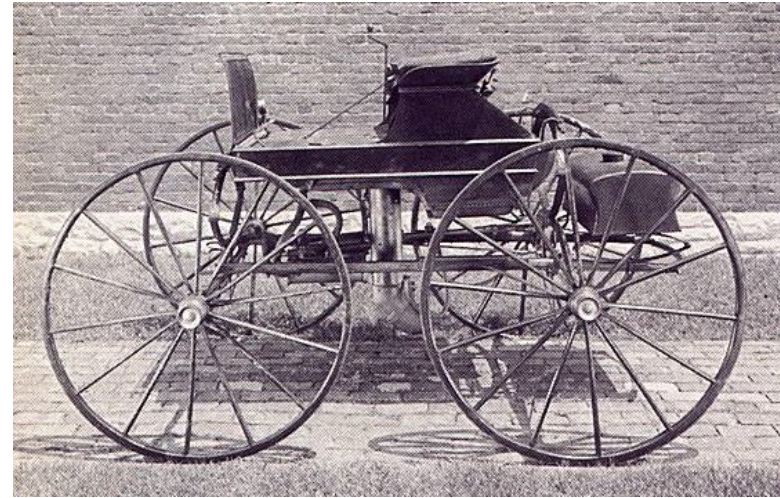The University of New South Wales, Australia

Presentation Outline ..

- GPS and the Smart Car Revolution

- GPS-based Profiling

- The Profiling Experiment

- Related Privacy Issues

- Conclusion

The Automobile Evolution

Source: http://patentpending.blogs.com/patent_pending_blog/automobile/index.html

- ## Past
  - Mechanical & hydraulic Parts
  - Unreliable
  - Uncomfortable

- ## Present
  - Uses computers
  - Telematics and GPS navigation
  - Keyless ignitions, ABS brakes
  - Various sensors

Source: http://autoshow.blogs.cnnmoney.cnn.com/2007/04/

GPS and The Smart Car Revolution

- Cars would be one of the major users of GPS

- Traditionally used for route navigation

- Recent uses include

  – GPS as alibi

  – Covert Surveillance

  – Mobility-based charging

  – And more recently...Congestion Charging

GPS Alibi

- Sets a legal precedent of GPS admissibility in NSW courts

- Lead to charges being dropped and fine overturned

Father and son stick to guns to prove radar wrong

Jerry Simotas and his son Michael, who have both beaten speeding fines in court by claiming police used faulty equipment.
Photo: Ben Rushton

Source: Sydney Morning Herald (12 March 2007)

Covert Surveillance

- Law Enforcement
  - US vs. Garcia
- Employers
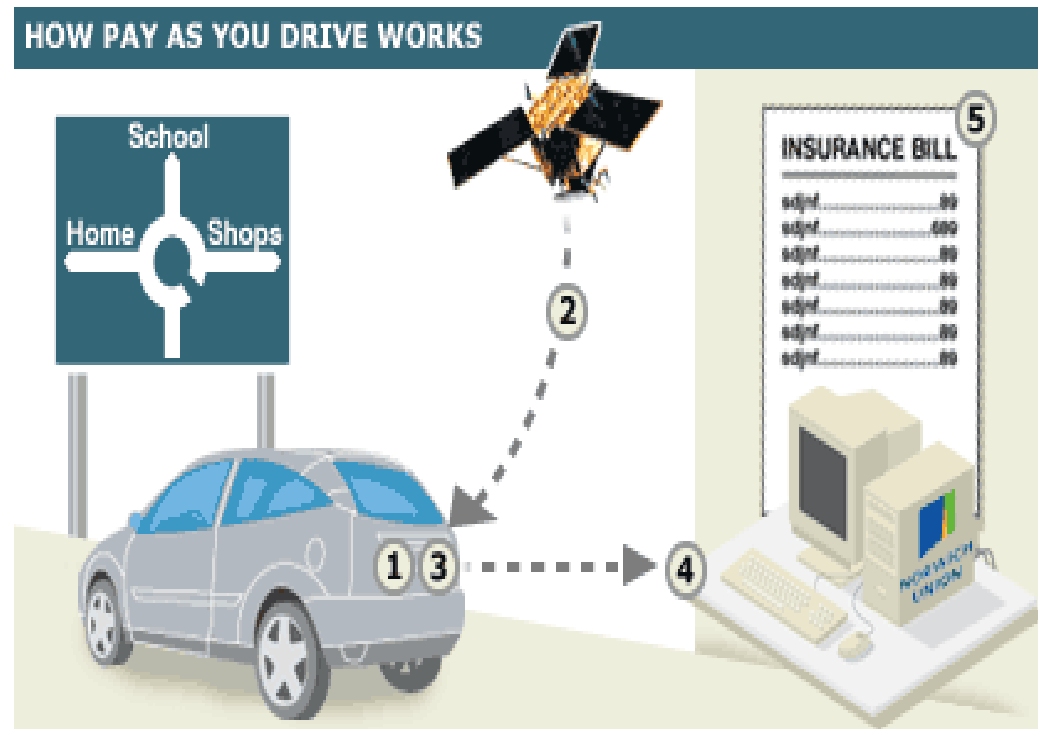- Curious spouse
  - George Ford Case

Source: TrackStick Pro userguide, pp 30

Motorists have no expectation of privacy on public roads

Mobility-based Charging

– Insurance
– Toll Collection
– Road Tax
– Parking

HOW PAY AS YOU DRIVE WORKS

– And Congestion-charging

Data Collection and Retention

Data Collection and Retention

- Sources of GPS data
  - Surveillance
  - Commercial projects (insurance, LBS, etc.)

- Decreasing cost of data storage
  - Data held indefinitely
  - Gives sufficient time to seek alternate use

- Significant privacy threats
  - No information or consent
  - Misreporting
  - Profiling ..

Profiling

".. a technique whereby a set of characteristics of a particular class of person is inferred from past experience, and data-holdings are then searched for individuals with a close fit to that set of characteristics"
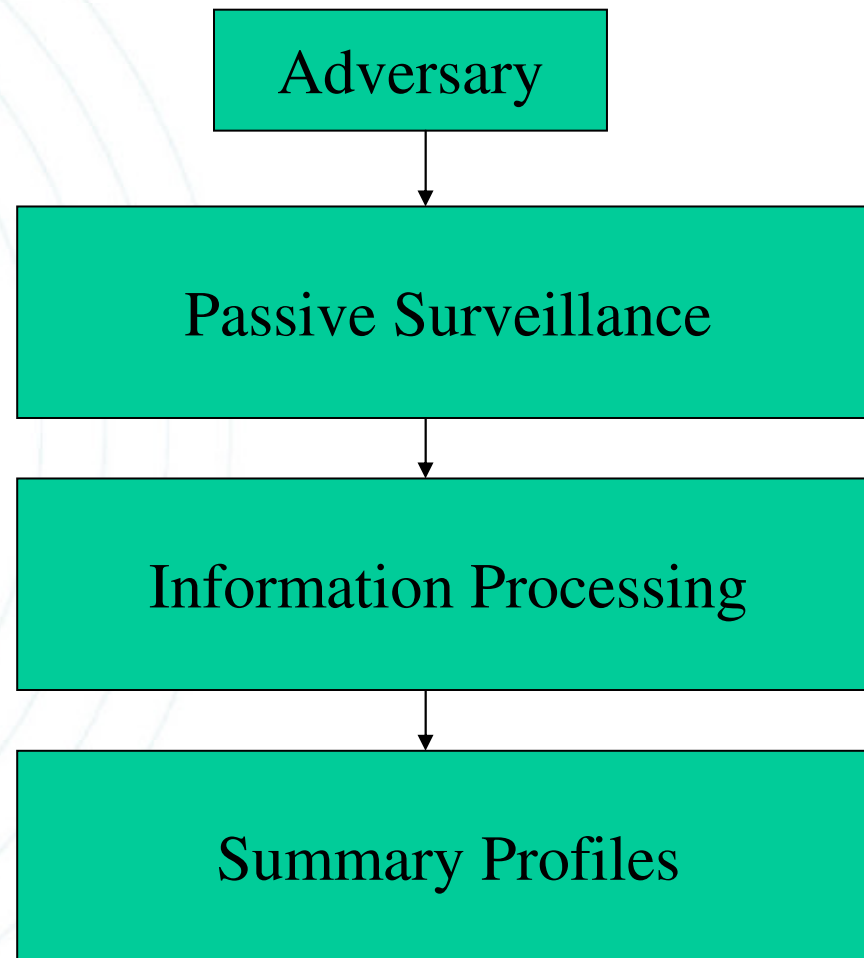
Roger Clarke, 1993.

GPS-based Profiling

Drawing inferences about individuals based on their GPS logs.

- Inferences have been made
  - Driving behaviour and Transport studies

Profiling Paradigm

```
┌─────────────────────────┐
│       Adversary         │
└─────────────────────────┘
             │
             ▼
┌───────────────────────────────────────────┐
│          Passive Surveillance             │
└───────────────────────────────────────────┘
             │
             ▼
┌───────────────────────────────────────────┐
│         Information Processing            │
└───────────────────────────────────────────┘
             │
             ▼
┌───────────────────────────────────────────┐
│            Summary Profiles               │
└───────────────────────────────────────────┘
```

Automated GPS Profiling

- 1020 Google Scholar hits for
  - "GPS + Data mining + Privacy"
  - Very few efforts to critically evaluate the privacy threats of profiling
- To demonstrate
  - Wealth of inferences may be drawn
  - Which may not necessarily be correct
  - Violate user's privacy
  - Automated: future abuses would be automated

The Experiment

- Collect data with a passive GPS device
- Process this data
  - Identify "significant locations"
  - Use data mining techniques, heuristics and rules
  - Correlate information from other databases
  - Infer home, personality, pattern and road behaviour
- Create summary profiles using tables and maps
- Compare with actual by interviewing volunteers

The Tools

- Passive GPS Surveillance device

- Software
  - PSMA's GeoCoded National Address File
  - PSMA's transport Maps
  - ESRI ArcGIS and ArcObjects
  - Visual Basic
  - PostGIS

- Heuristics

Surveillance

- Volunteers included
  - Academic staff
  - Administrative staff
  - Research student
  - Undergrad student
- Data collected for at least one week
- Device requires no input from user
- Plugs into the cigarette lighter jack
- Logs data 4-6 times per minute

Information Processing

- Import "anonymous" GPS data
- Cleanse data
- Identify significant locations

Rec #	Date	Time	Latitude	Longitude	Altitude	Status	Course	GPS Fix
134	03/17/2007	11:15	-33.9120°	151.1194°	24.7 m	31 kph	NE	Y
135	03/17/2007	11:15	-33.9116°	151.1199°	25.4 m	31 kph	NE	Y
135	03/17/2007	11:15	-----	-----	-----	Power Off	-----	-----
140	03/17/2007	11:22	-----	-----	-----	Power On	-----	-----
141	03/17/2007	11:22	-33.9054°	151.1275°	0.0 m	0 kph	N	Y

Black dots - significant locations

Information Processing
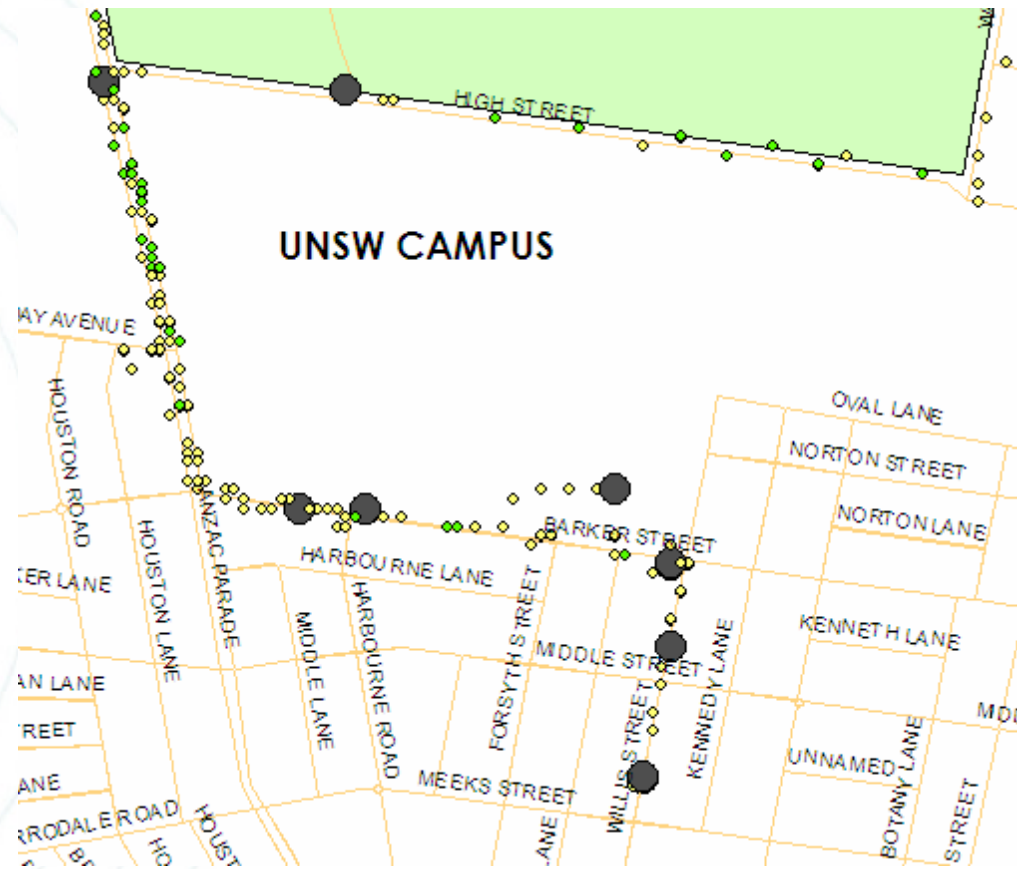
● Vehicle parked – Significant location

Home Determination Algorithm

- Exclude insignificant locations using temporal heuristics to identify home and work locations
  - Morning destinations
  - Afternoon/Evening destinations
- A set of locations short-listed as potential candidates
- Use **PSMA's GNAF** data address file for NSW
  - Apply a nearness query for a residential address to these points using PostGIS
  - The statistical mode considered as the most likely address
  - Classic example of correlating databases
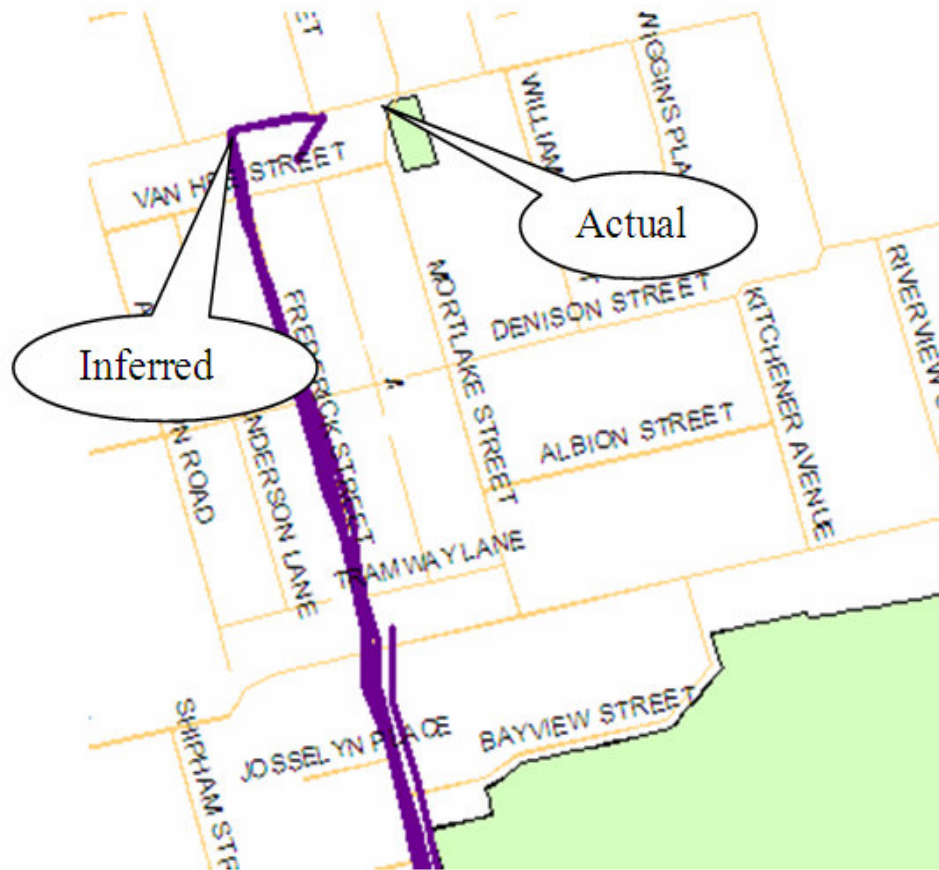
Home Determination Algorithm

Reported Results

Home location	Volunteer C	Volunteer B	Volunteer Y	Volunteer J	Volunteer U
Street number(s) inferred	7	39	24, 25	44	53
Actual street number	11	39	22	Different street	51

Assumption: 2 address buffer

Inferring Accuracy

Street address for volunteer incorrectly guessed by the home determination algorithm

Summary Profile

| Work and commute profile	Volunteer C	Volunteer B	Volunteer Y	Volunteer J	Volunteer U
Total GPS records	5240	1997	2330	4812	2147
Total Distance	301 km	174.59 km	172 km	284.9 km	149.72
Average distance	27.38 km	34.59 km	31.2 km	40.7 km	37.43 km
Total travel time	12 hr 45 m	4 hr 25m	5 hr 1 m	11 hr 44 m	4 hr 51 m
Average travel time	1hr 10 m	52 m	54 m	1 hr 40 m	1 hr 12 m
Max Speed	101 kph	83 kph	86 kph	98 kph	91 kph
Average Speed	32 kph	45 kph	39 kph	33 kph	39 kph
Average time leaves home	7:33 am	8:21 am	9: 10 am	07:46 am	9:54 am
Average time leaves work	3:30 pm	5:09 pm	4:54 pm	08:58 pm	5:07 pm
Average time arrives at work	8:03 am	8:55 am	9:32 am	08:40 am	10:15 am
Average time at work	7 hr 58 min	8 hr 10 min	7 hr 25 min	12 hr 18 m	6 hr
Parks car in	University parking lot	University parking lot	University parking lot	University parking lot	Around university
Type of person	Academic Or Support	Academic Or Support	Academic Or Support	Research Student	Undergrad Student

Visual Profiles
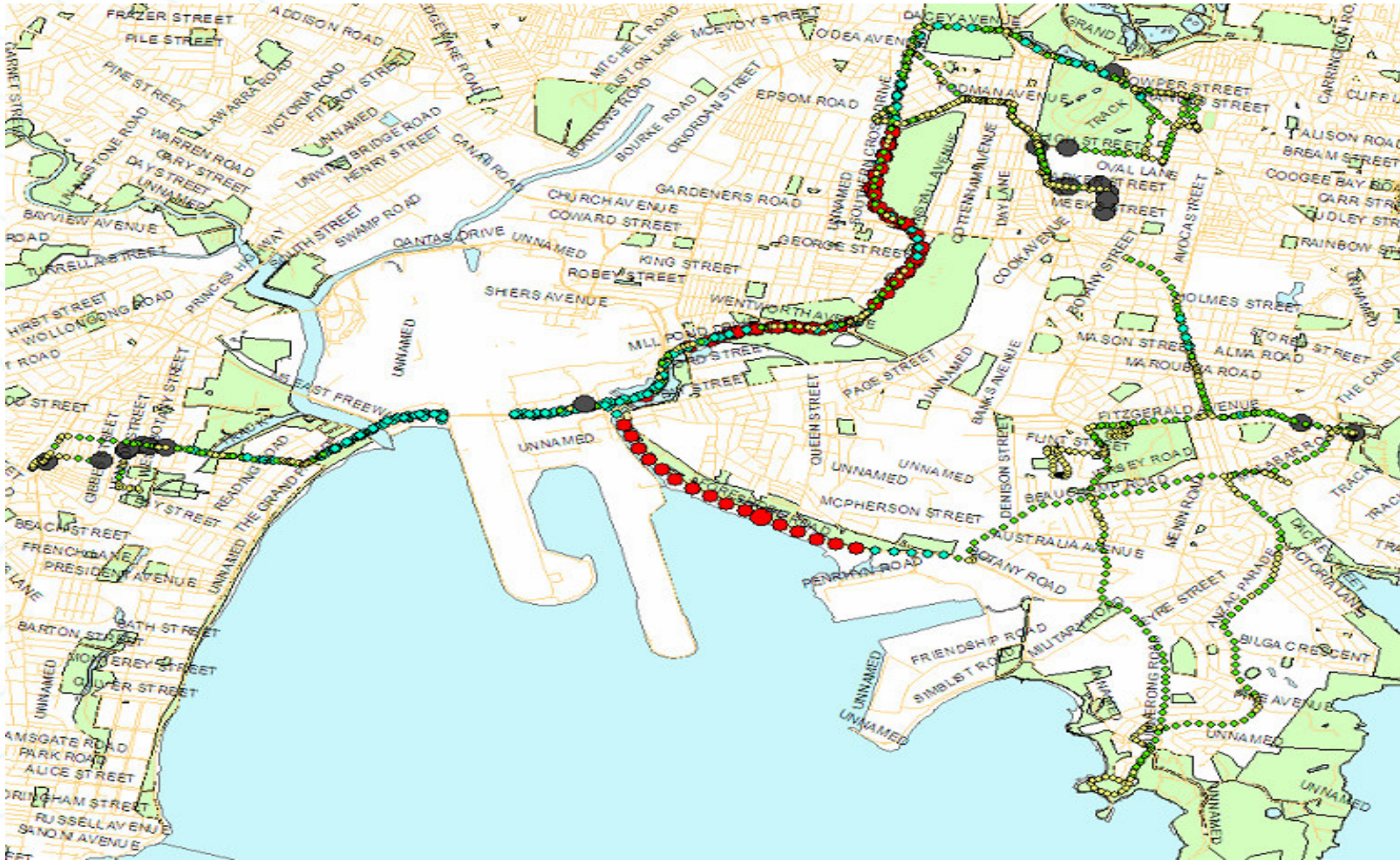
● Vehicle stopped – Significant location

● Vehicle speed > 80km/hr

Speed and Acceleration Monitoring
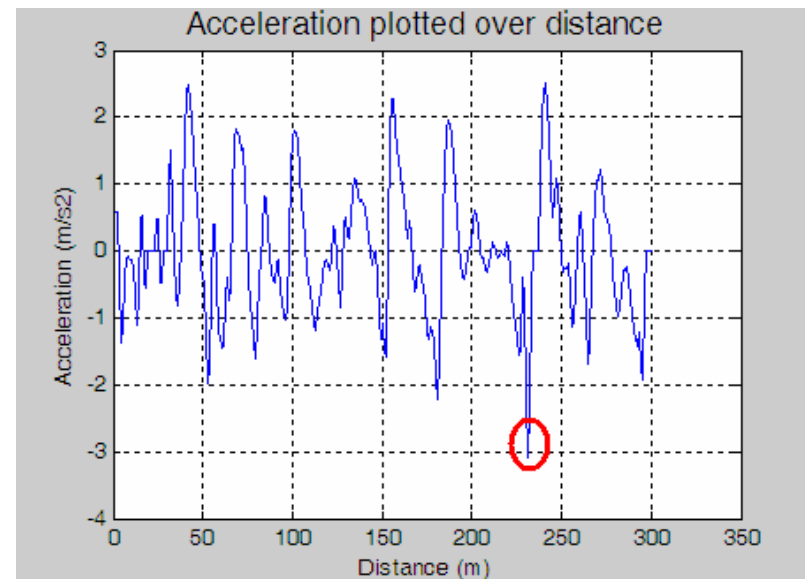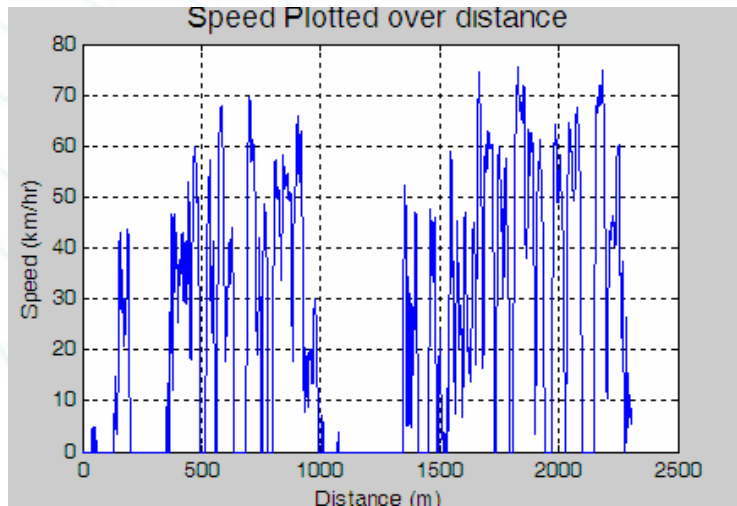
Red spot indicates deceleration of 3 m/s2---- (hard braking)

Why is Profiling Privacy Invasive?

- Experiment demonstrates risks of profiling
- Secondary use of data
- No consent sought
- Can embarrass or cause harm to a person
- Possibility of Mis-Profiling
- Profiling and stereotyping
  - Risky driver for insurance

Concluding Remarks

- Data gathered only for 1 week
  - Insufficient for highly accurate profile generation
- Predicted the personnel type and residence
- Profile may be circumstantial
  - An academic preparing lecture notes
  - Being on campus doesn't necessarily mean working
  - High speed, brakings due to emergency
- Mobility-pricing and complete GPS log disclosure raises these threats
  - Important to negotiate the use and retention
  - Use aggregated data reporting

Thank You for your time!

- Questions?