

# Towards Intra-Vehicular Sensor Data Fusion

Paulo H. L. Rettore, Bruno P. Santos, André B. Campolina,  
Leandro A. Villas, and Antonio A. F. Loureiro

# Schedule

- Introduction
- Background
- Vehicular Data
- Case Study
- Conclusion

# Introduction

- The world's population has increased
- Over 50% live in huge cities
- Issues related to transportation and traffic begin to grow:
  - Safety and quick mobility
  - Injuries and accidents causes:
    - Expensive cost medical
    - Decrease productivity



**Brazil**



**China**



**USA**



**India**

# Introduction

- Possible solutions?
  - Governments strategies
    - Rotating vehicles
    - Traffic restriction on selected regions

Monday	Last # 1 and 2
Tuesday	Last # 3 and 4
Wednesday	Last # 5 and 6
Thursday	Last # 7 and 8
Friday	Last # 9 and 0

**Not scalable**

# Introduction

- Possible solutions?
  - Governments strategies
    - Rotating vehicles
    - Traffic restriction on selected regions
  - Additional capacity



# Introduction

- Possible solutions?
  - Governments strategies
    - Rotating vehicles
    - Traffic restriction on selected regions
  - Additional capacity



**Cost prohibitive**

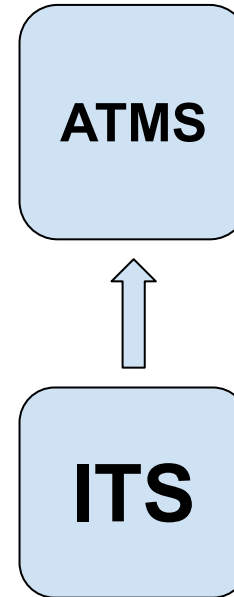
# Introduction

- Possible solutions?
  - Governments strategies
    - Rotating vehicles
    - Traffic restriction on selected regions
  - Additional capacity
  - Intelligent Transportation System (ITS)
    - Can be a feasible way
    - But ITS dependent of **data** and communication
- **We are interested in data for ITS.**
  - **Especially in car data**



# Background: ITS and Subsystems

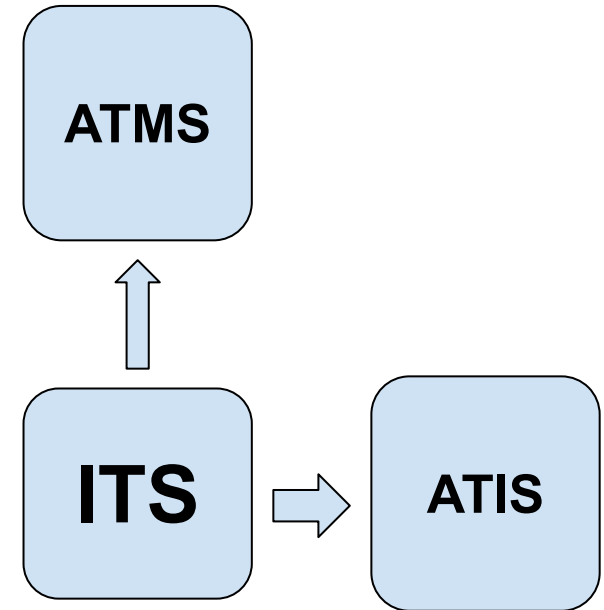
- Advanced Transportation/Traffic Management Systems
  - To control and manage traffic devices (signals, monitoring and safety devices etc...)





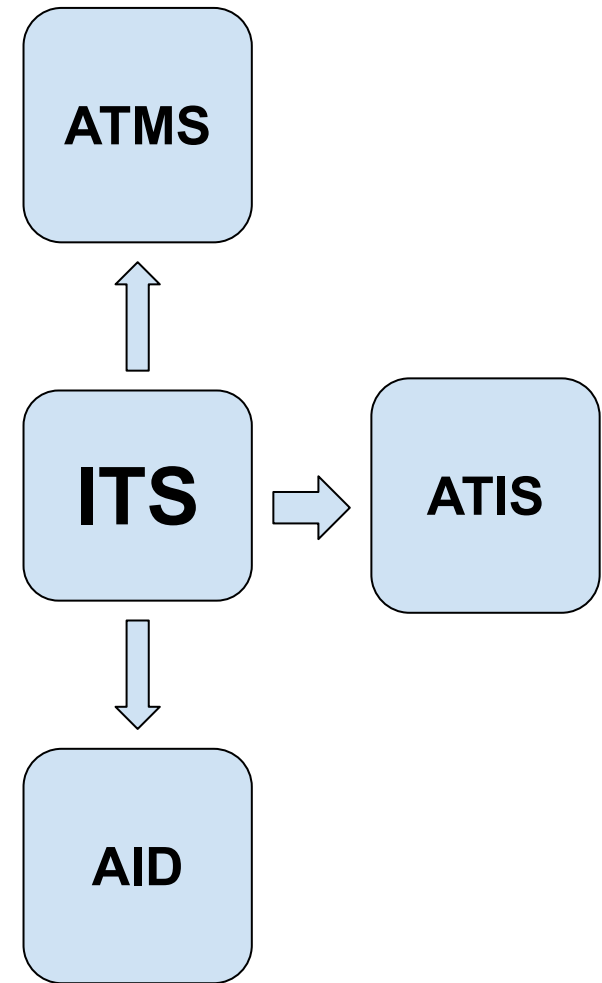
# Background: ITS and Subsystems

- Advanced Traveler Information Systems
  - To collect data and process it to improve understanding of traffic conditions



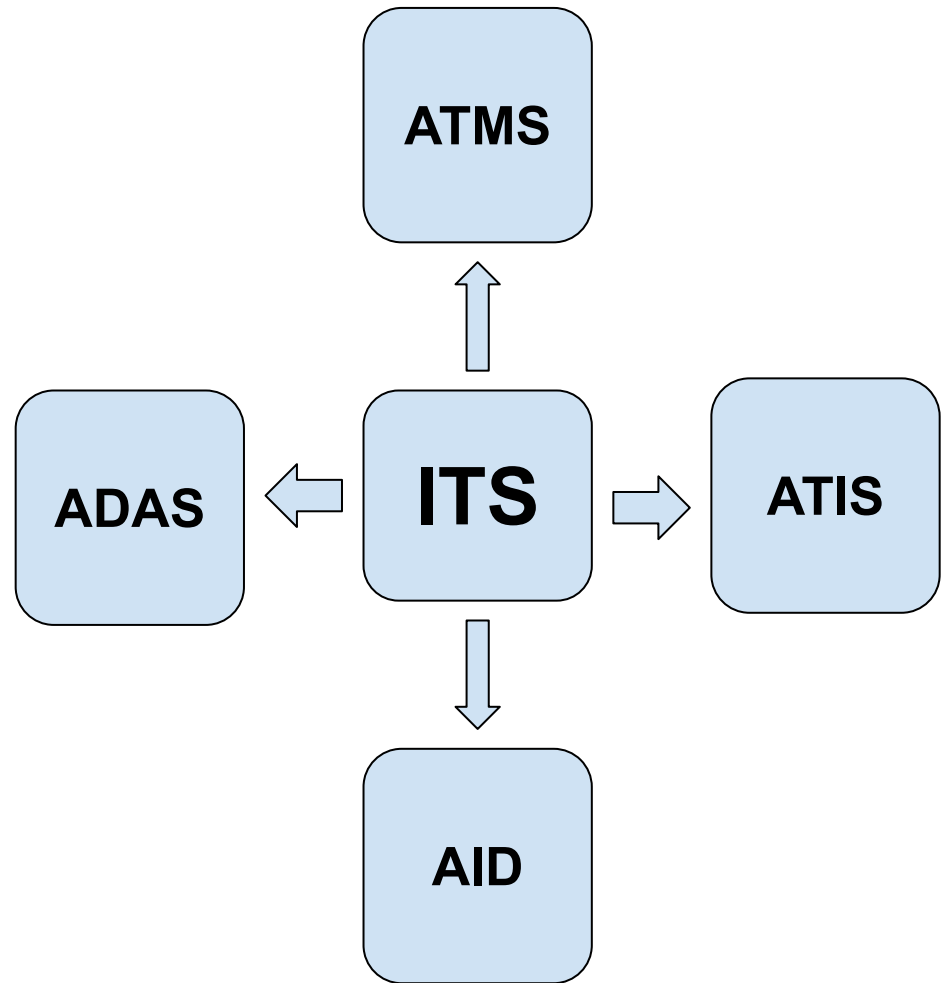
# Background: ITS and Subsystems

- Automatic Incident Detection
  - To automatically detect incidents as soon as possible to increase safety and reduce users perception of traffic disruption



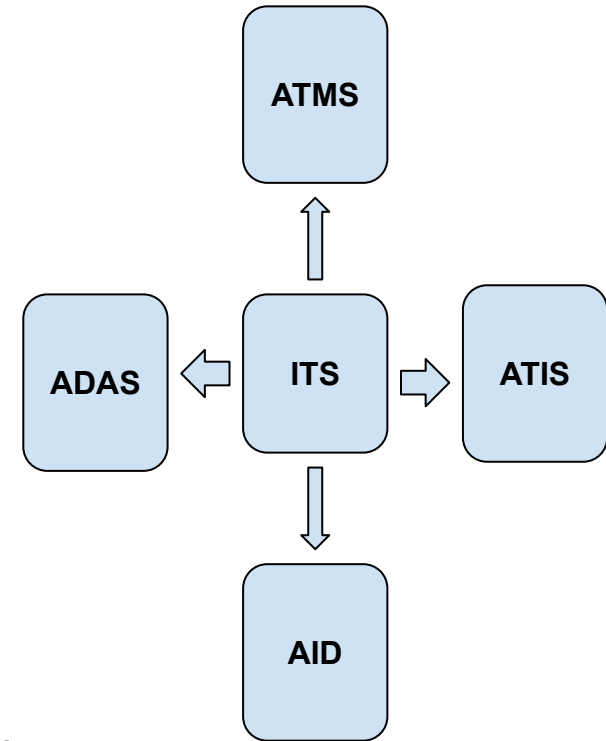
# Background: ITS and Subsystems

- Advanced Driver Assistance Systems
  - Apply technologies intended to improve driver experience and safety of the users



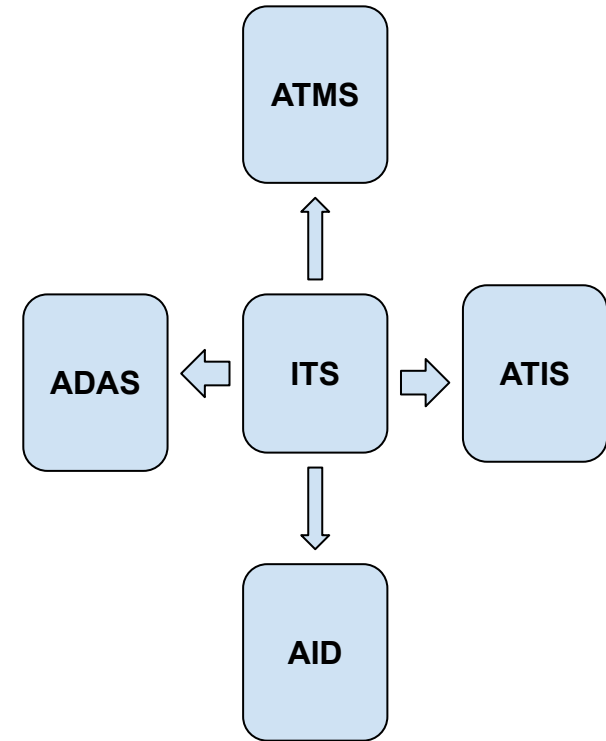
# Background: ITS and Data

- ITS are powered by data as much as possible
- Traditional traffic sensors are ineffective for ITS requirements
- Fortunately, There are available other data sources
  - Cameras, GPS, Smartphones, vehicles
  - Probe vehicles, Social Media ...
- These sensors provides timely and precise information about traffic status



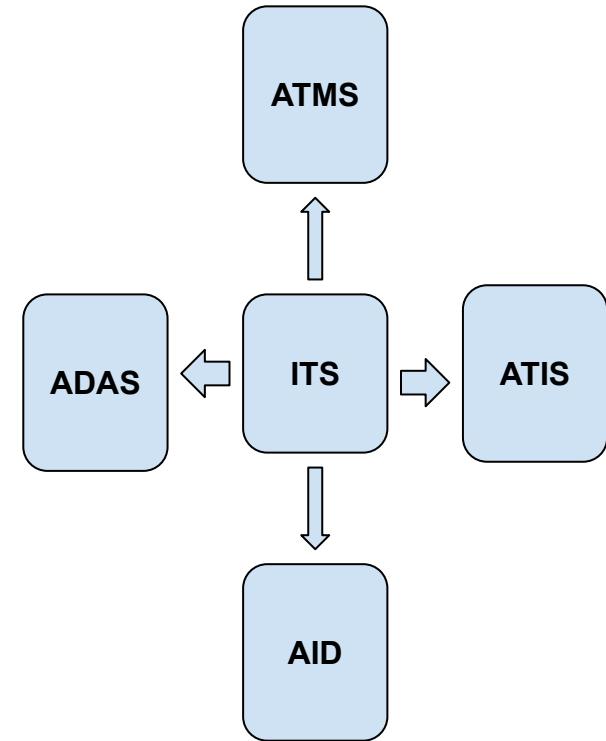
# Background: ITS and Data Fusion

- Data information is the heart of ITS
- Usually the data is heterogeneous
- Heterogeneous data fusion techniques are suitable in such situation
  - Inference, Estimation, Classification, Aggregation ...



# Background: ITS and Data Fusion

- Can we use directly data catch from sensors to ITS subsystems?
  - In general, NO
- **Data aspects** make DF process a challenging task
  - Imperfection
  - Uncertainty
  - Outliers
  - Ambiguity
  - ...



# Vehicular Data

- The sensors on a car, communicate with the Engine Control Unit (**ECU**)
- The data they output is accessible using the On-Board-Diagnostics (**OBD**) interface
- The information is collected from different sensors spread across different parts of the vehicle's body in different measuring units
- It can be considered **Heterogeneous**



# Vehicular Data



Sensors					
Engine load	Vehicle speed	Torque sensor	Fuel pressure	Oxygen sensors	Fuel Tank Level
Kilometers per litre	Intake air temperature	Ambient air temperature	Catalyst temperature	Relative throttle position	Accelerator pedal position
Fuel flow rate	CO2	Ethanol fuel %	Engine oil temperature	Fuel injection timing	O2 sensor monitor
Voltage	Distance traveled	Fuel remaining	Fuel rail pressure	Hybrid battery pack remaining life	Evap. system vapor pressure
Engine RPM	Engine coolant temperature	Fuel type	Malfunction indicator lamp	Exhaust gas recirculation error	Mass Air Flow Sensor
Altitude	GPS location	Collision sensor	Automatic brake actuator	Steering angle sensor	Rear camera
GPS speed	Gyroscope	luminosity sensor for headlights	Active park assist	Water in fuel sensor	Airbag sensor
Barometric Pressure	Acceleration	Cost per mile/km	Front object laser radar	Night pedestrian warning IR sensor	Tire pressure sensor
Microphone sensor	Pressure sensor	Drowsiness sensor	Shock sensor	Rain-Sensing Windshield Wipers	Motion sensor



# Vehicular Data



Sensors					
Engine load	Vehicle speed	Torque sensor	Fuel pressure	Oxygen sensors	Fuel Tank Level
Kilometers per litre	Intake air temperature	Ambient air temperature	Catalyst temperature	Relative throttle position	Accelerator pedal position
Fuel flow rate	CO2	Ethanol fuel %	Engine oil temperature	Fuel injection timing	O2 sensor monitor
Voltage	Distance traveled	Fuel remaining	Fuel rail pressure	Hybrid battery pack remaining life	Evap. system vapor pressure
Engine RPM	Engine coolant temperature	Fuel type	Malfunction indicator lamp	Exhaust gas recirculation error	Mass Air Flow Sensor
Altitude	GPS location	Collision sensor	Automatic brake actuator	Steering angle sensor	Rear camera
GPS speed	Gyroscope	luminosity sensor for headlights	Active park assist	Water in fuel sensor	Airbag sensor
Barometric Pressure	Acceleration	Cost per mile/km	Front object laser radar	Night pedestrian warning IR sensor	Tire pressure sensor
Microphone sensor	Pressure sensor	Drowsiness sensor	Shock sensor	Rain-Sensing Windshield Wipers	Motion sensor

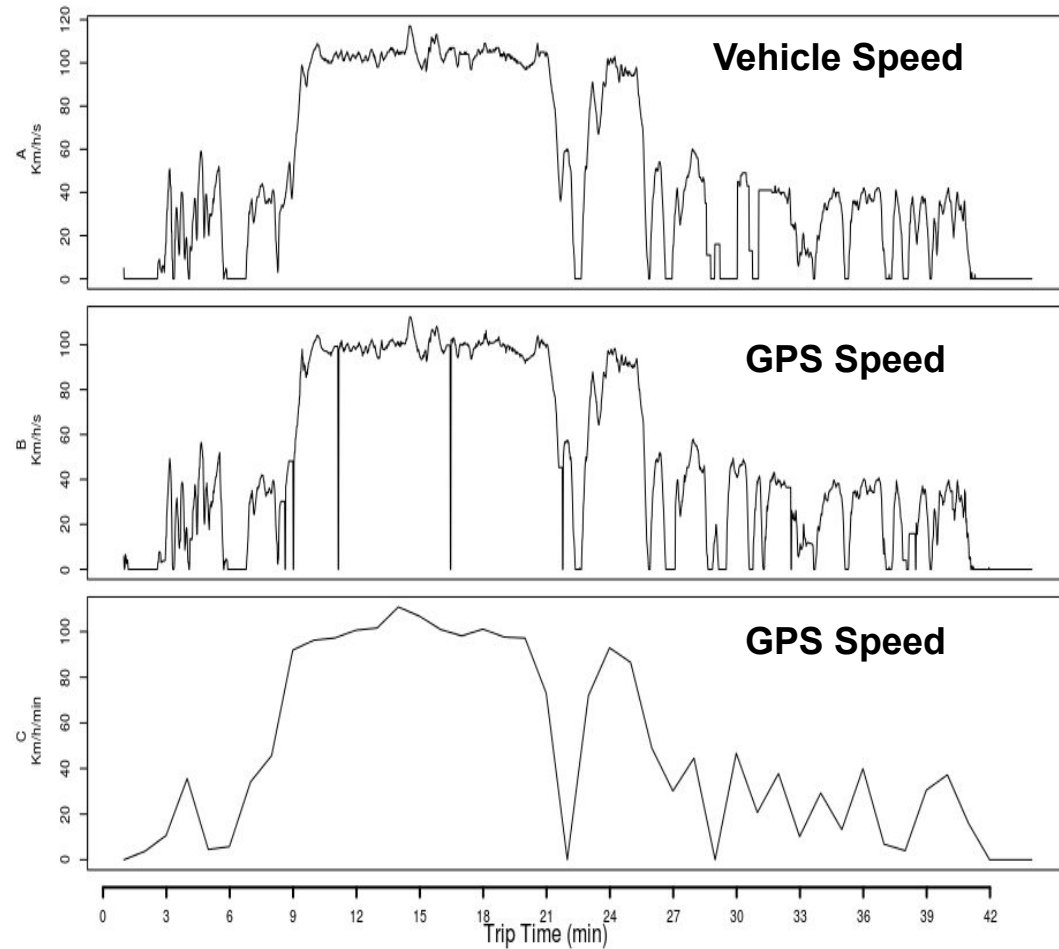
# Case Study

- **Problems of heterogeneous data fusion**
  - A practical and comprehensive data analyses in vehicular sensor data
- We used an **OBD** Bluetooth adapter to collect data from a car
- The logs of these vehicles consist of **55 trips** of **40 km** with an average time of **50 minutes each**

# Case Study

- Granularity

Is related to the ability to derive **valuable** information about entities of **interest** on a data set.



# Case Study

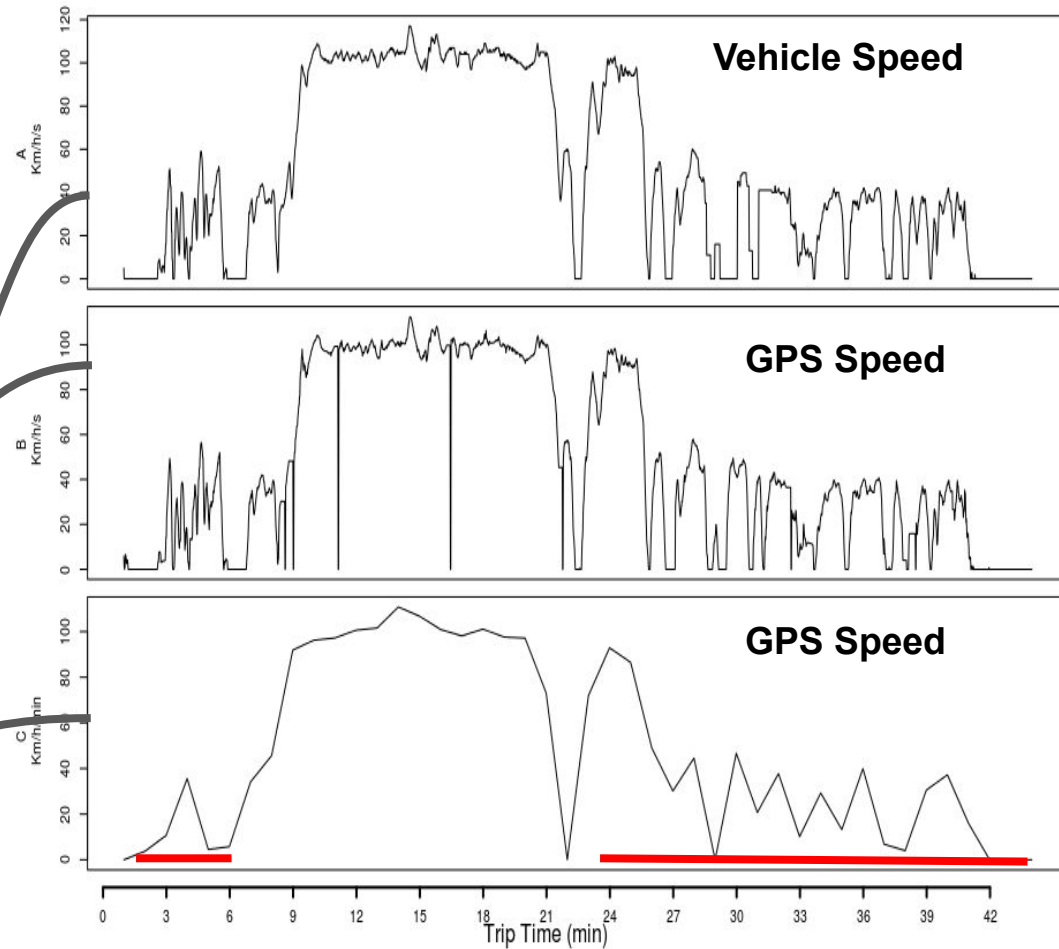
- Granularity

Is related to the ability to derive **valuable** information about entities of **interest** on a data set.

1 seg

60 seg

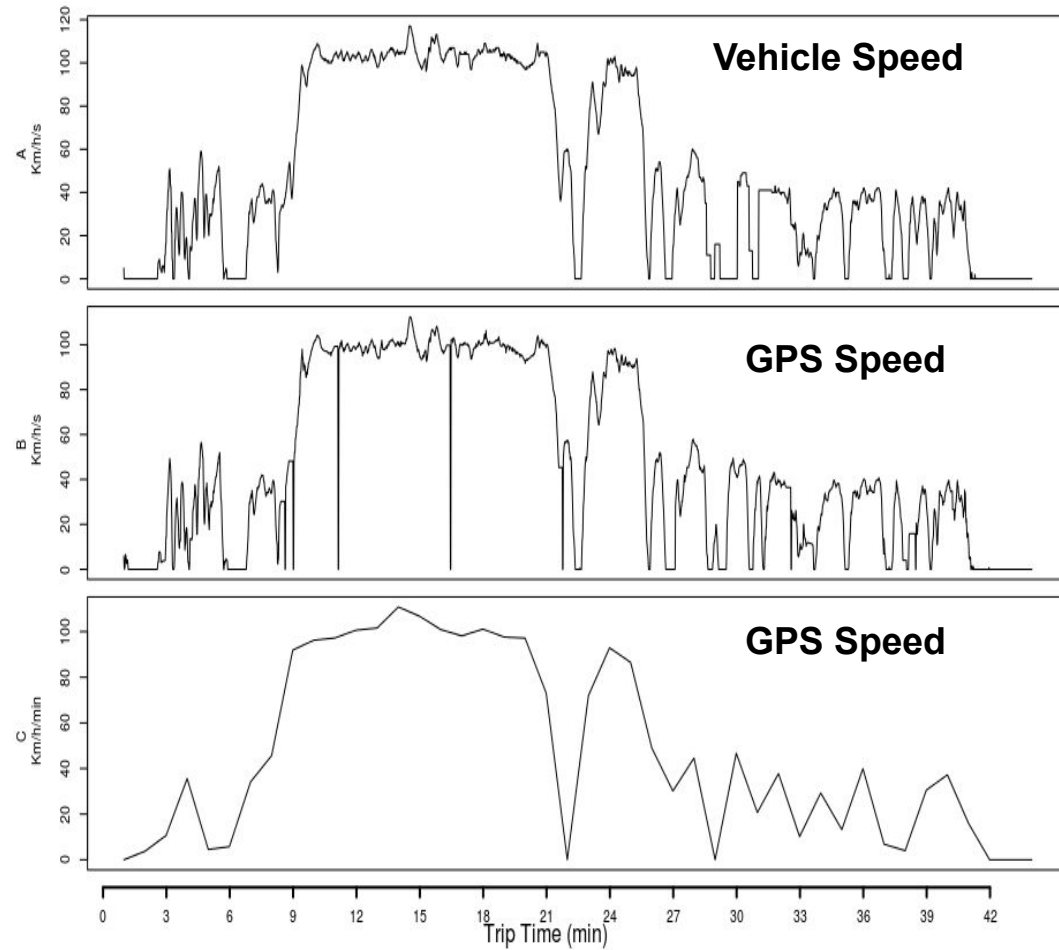
Lose Information



# Case Study

- Vagueness

Occurs in data sets where attributes are **not well defined**



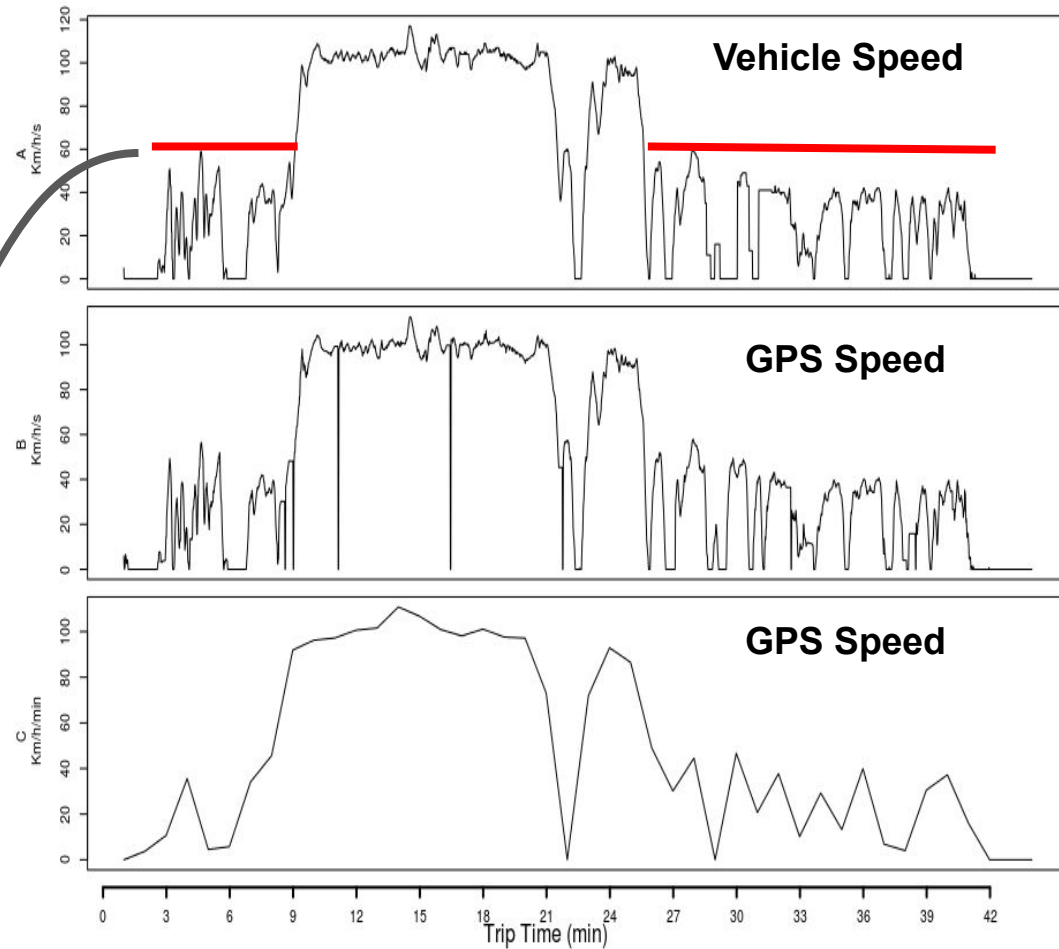
# Case Study

- Vagueness

Occurs in data sets where attributes are **not well defined**

FAST or SLOW ?

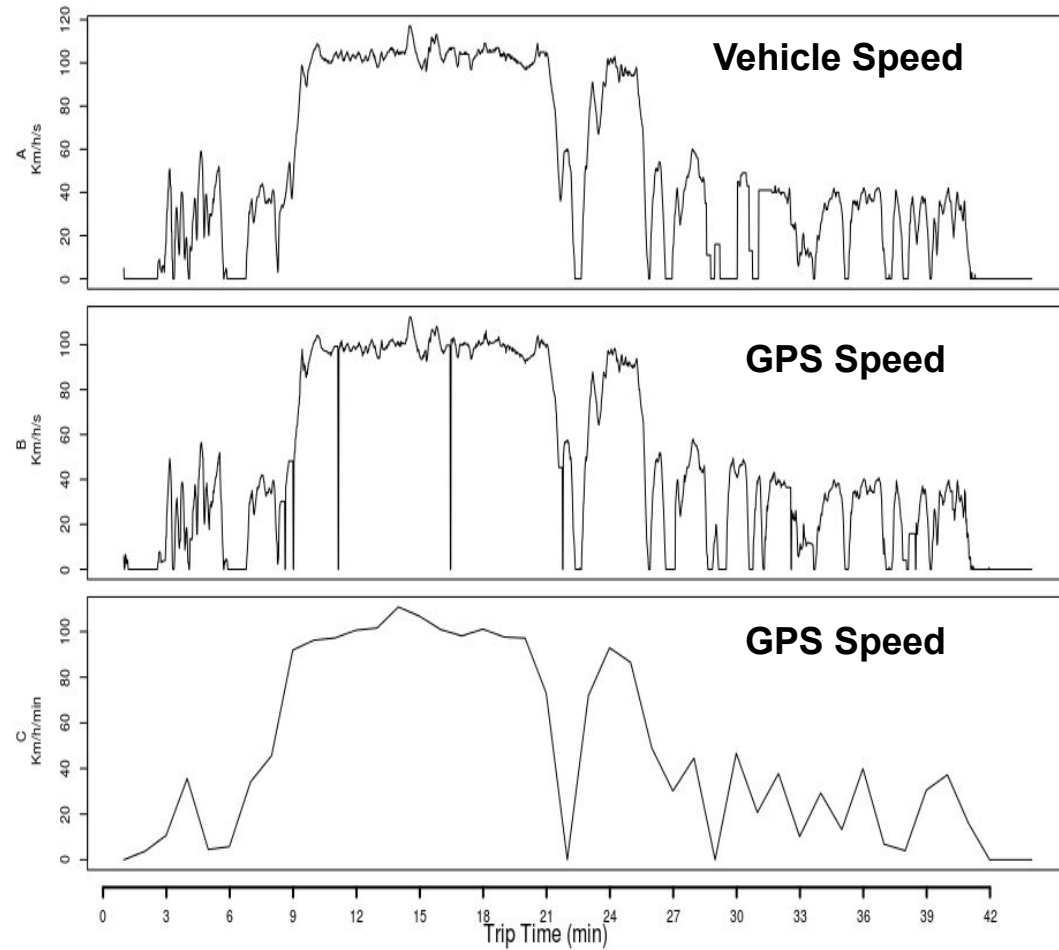
Depend on the context



# Case Study

- Outlier

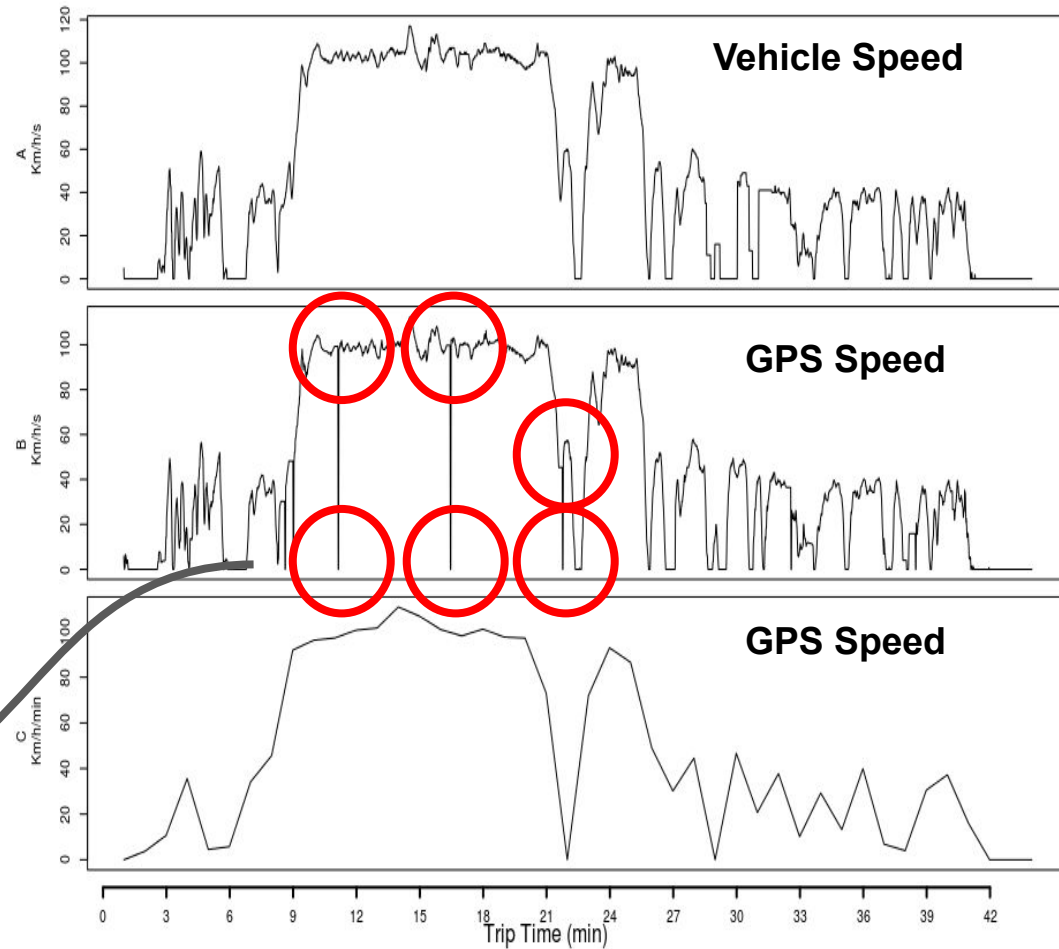
There are points whose data is **incorrect**



# Case Study

- Outlier

There are points whose data is **incorrect**



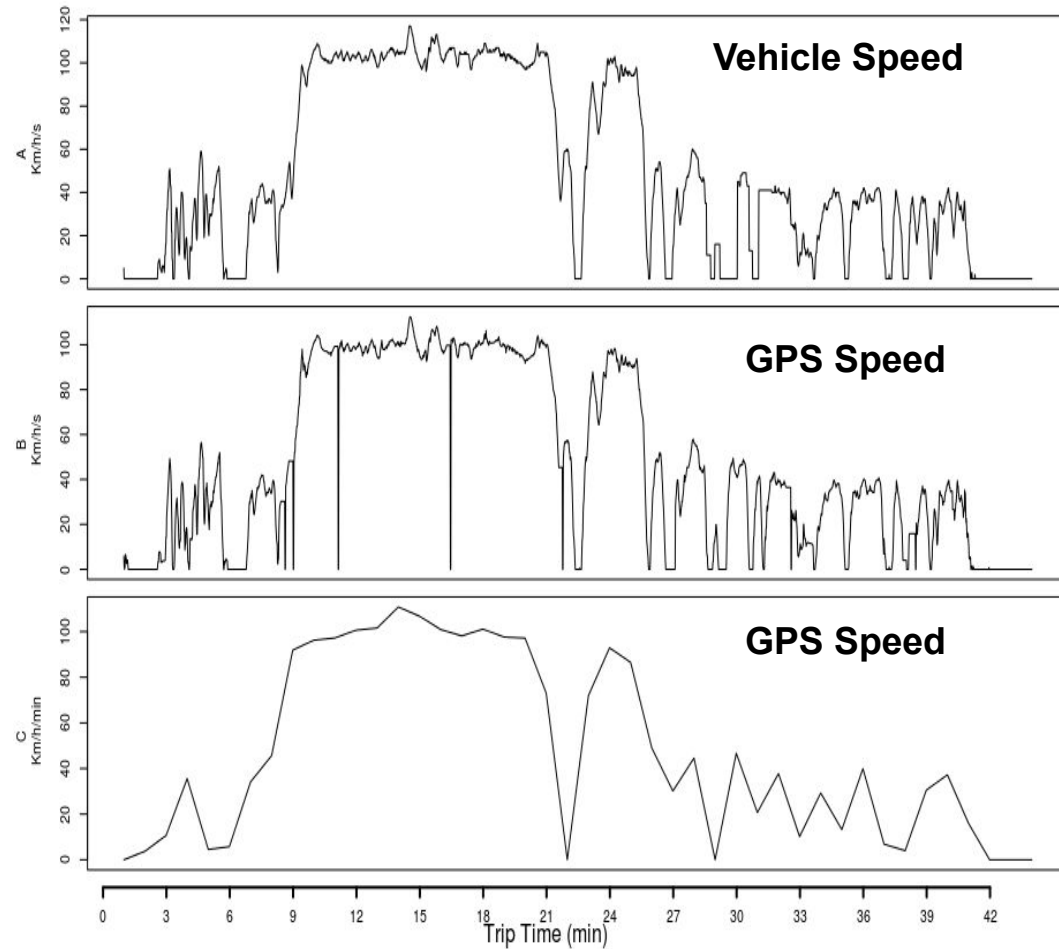
Distorter points with 0  
(zero) values



# Case Study

- Conflict

**Divergent** specialists' opinions or punctual **errors** in sensor readings happen and cause **conflicts** in data observations.



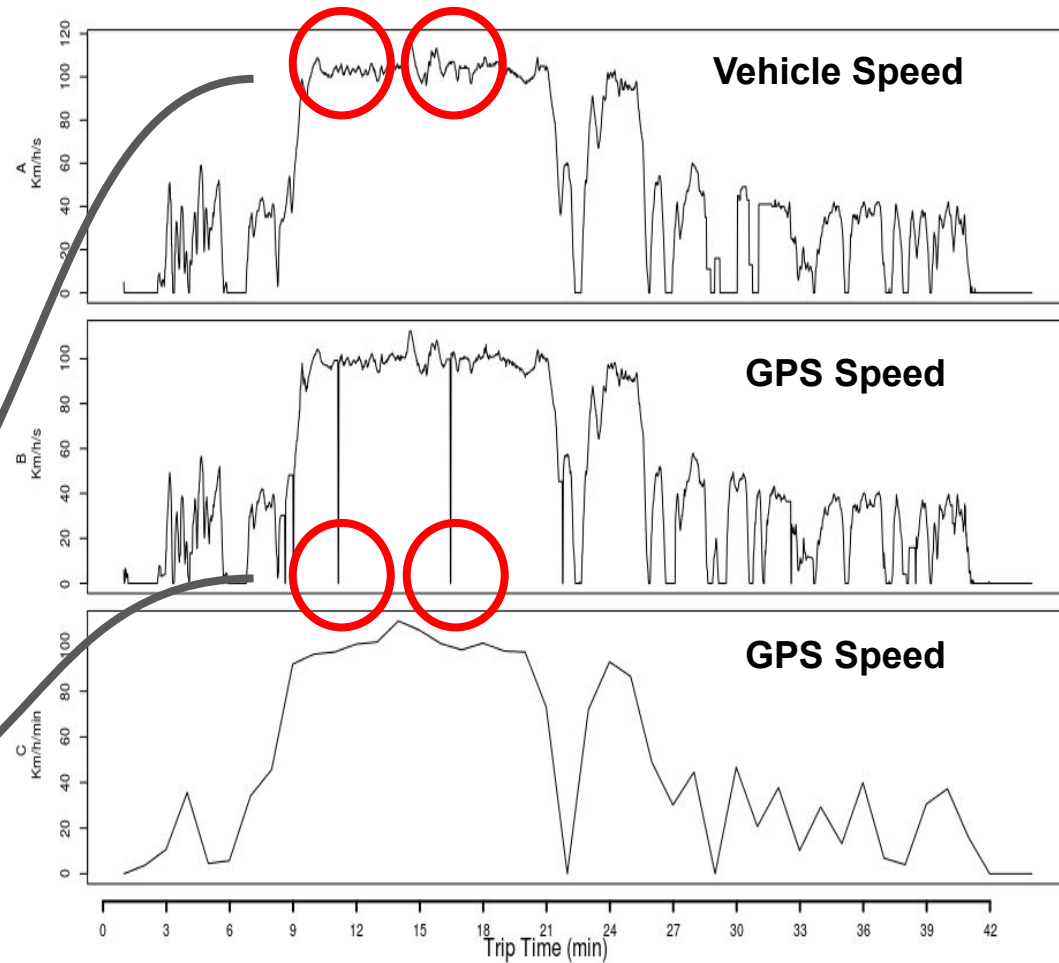
# Case Study

- Conflict

**Divergent** specialists' opinions or punctual **errors** in sensor readings happen and cause **conflicts** in data observations.

Points with 100 km/h

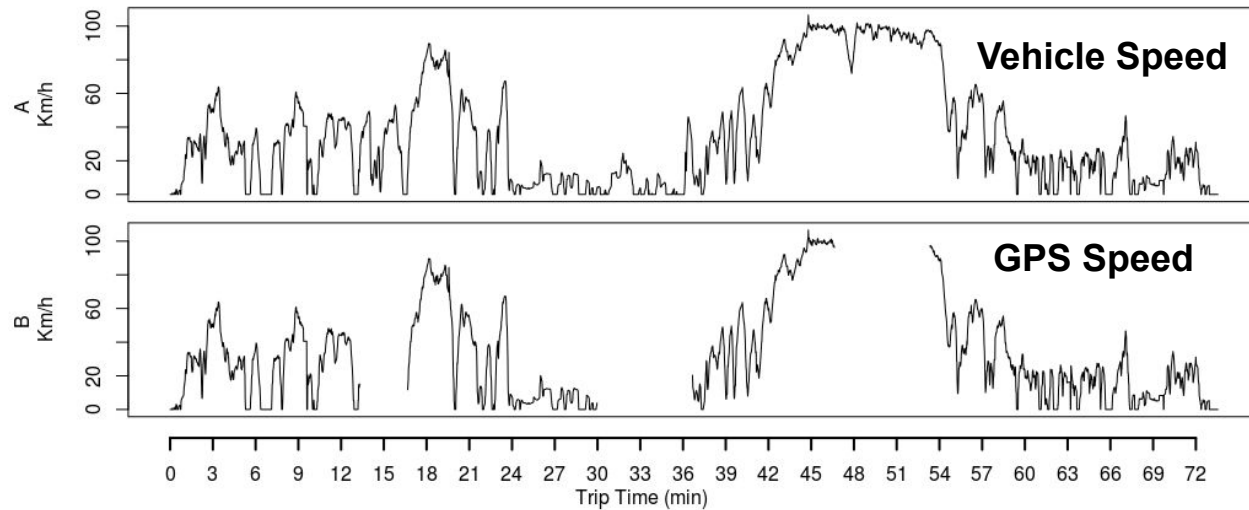
Points with 0 km/h



# Case Study

- Incompleteness

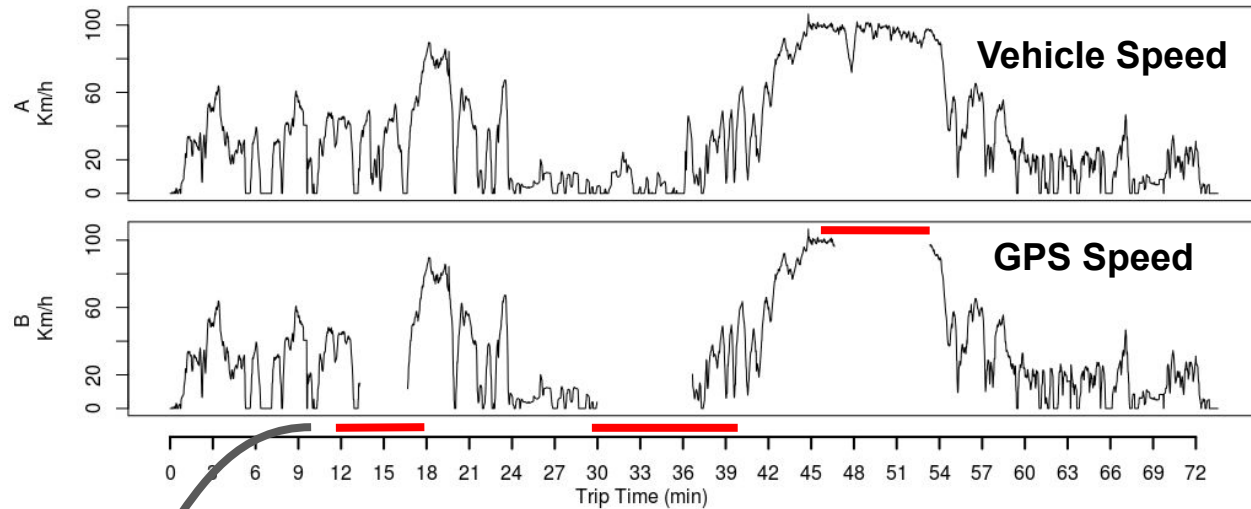
Data with missing parts



# Case Study

- Incompleteness

Data with missing parts

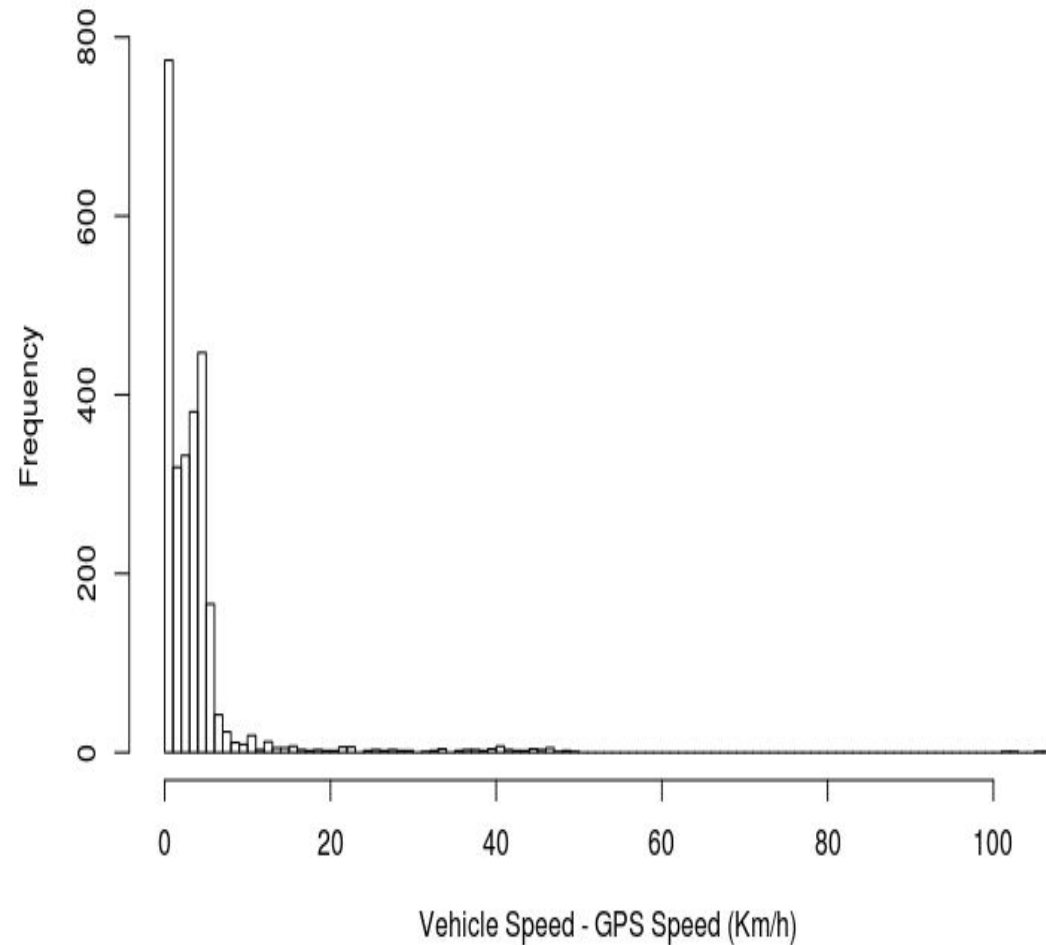


Barriers in the environment  
as tunnels

# Case Study

- Ambiguity

When occurrences in the data set are assumed to express the same information, however they differ from each other.



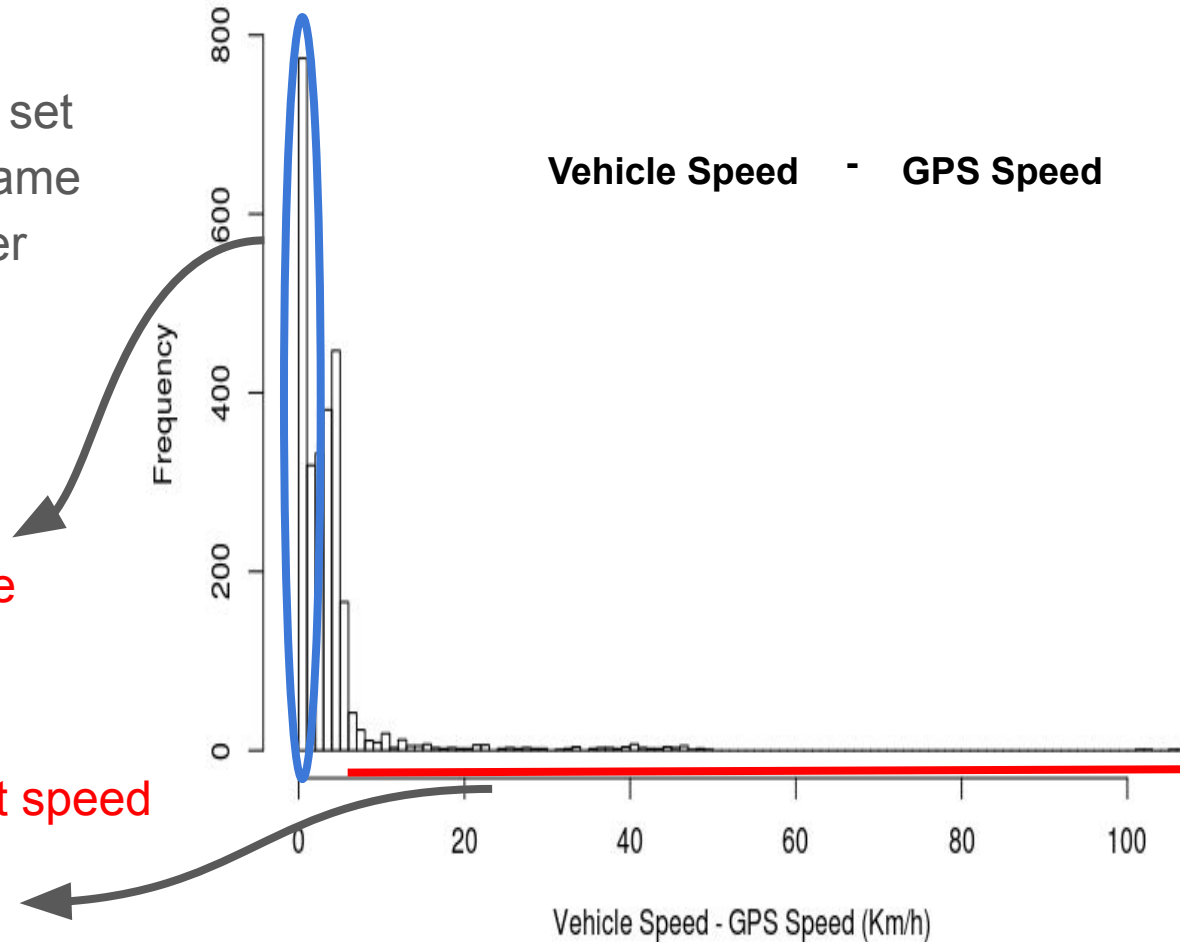
# Case Study

- Ambiguity

When occurrences in the data set are assumed to express the same information, however they differ from each other.

Both sensors collected the same speed

vehicle speed shows the current speed and GPS speed a different or conflicted value



# Case Study

- Uncertainty

In the case of **sensors**, the uncertainty is **always present**, in other words, it is inherent a property of any sensor

- Disorder

This problems is **not common in our scenarios**, because the process to data collect is **synchronous**

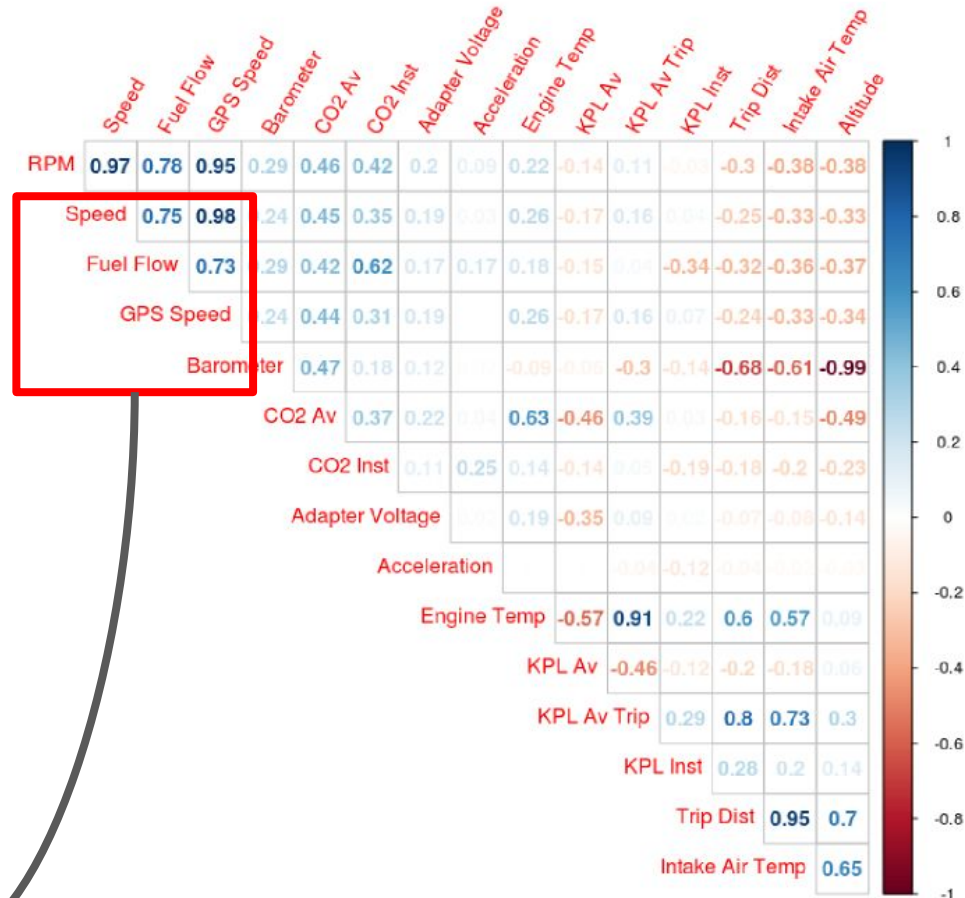




# Case Study

- Correlation

It is problematic since it can either enhance or attenuate some aspects due the data is fed multiple times in the data fusion system, multiplying its importance on the final result.

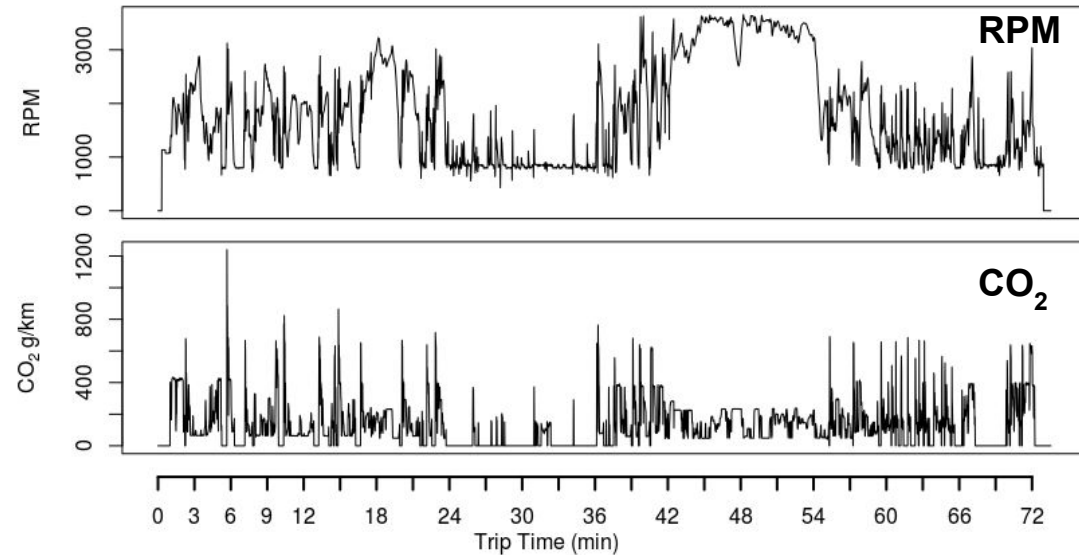


High correlations. So that, these data can be reduced to only one variable as Speed, for instance

# Case Study

- Disparateness

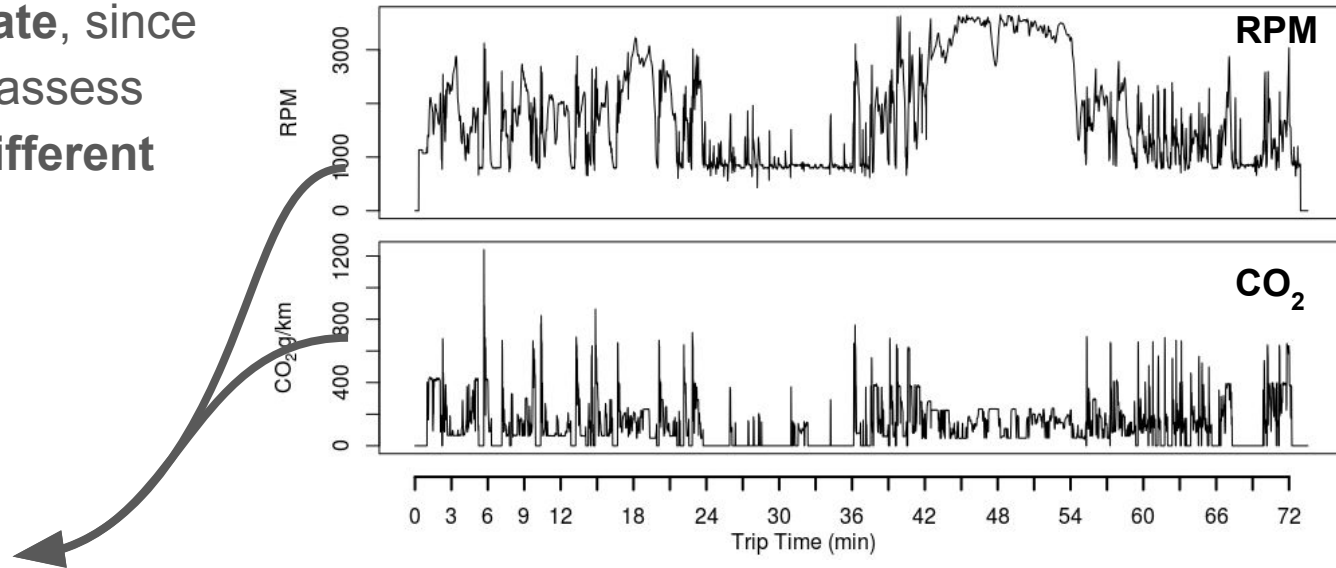
It is **inherently disparate**, since there are sensors that assess **different aspects** in **different units and scales**.



# Case Study

- Disparateness

It is **inherently disparate**, since there are sensors that assess **different aspects** in **different units and scales**.



Dissimilarity between  
two sensors

# Conclusion

- ITS can be boosted by take in account heterogeneous data collected from several sources as much as possible
- However, in general, the data comes with some issues making difficult heterogeneous data fusion process
- Therefore they must be treated before fusion process
- This work can guide beginner researchers to better understand the data, mainly in vehicular context, and some problems they possibly have to deal.

Thanks