Data Handling in PEGASUS

2nd PEGASUS Symposium

Dr.-Ing. Adrian Zlocki, Univ.-Prof. Dr.-Ing. Lutz Eckstein
Data Handing in the PEGASUS Database

1. Integration by database mechanics
2. Storage in database
3. Generation of complete scenario space
4. Output generation & test concept

Input Data

- Source FOT
- Source Test drives
- Source xy
- Source Accident data

Data with traffic events

(Logical) scenario

(Logical) scenario with parameter distributions

Concrete scenario

Usage testing ground
Usage driving simulator
Usage simulation

Data with traffic events

© PEGASUS | Final Event 13.05.2019 | fka
Signal definitions and quality check

- **Signal definitions** for description traffic constellations that are relevant for the motorway chauffeur

- **Different signal types** are defined, for example
  - Vehicle specific signals (dimensions, object type, …)
  - Signals for the motion state (position, velocity, angle, acceleration, …)
  - Signals to describe the environment (lanes, lane markings, …)
  - Signals to describe the driver and the driver-vehicle-interaction (control elements, …)

- Definition of three data quality levels
  - Minimum: With even lower quality, data would be useless within the database
  - Recommended: Data should have this quality to extract the most relevant information
  - Optimum: Highest standard to describe the full scenario with high precision

- Automated checks with feedback to uploader
  - Signal timings, Signal availability, Signal plausibility, …

- **Signals** are defined through JSON-files
- **Measurements** are saved as MAT or HDF5 files
## Input Data Sources

<table>
<thead>
<tr>
<th>Scenario Description</th>
<th>Scenario Relevance</th>
<th>Scenario Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Real Traffic Data (uninfluenced driving)</strong></td>
<td>Complete (depending on sensor setup)</td>
<td>Frequency of scenarios</td>
</tr>
<tr>
<td><strong>FOT with active ADAS/HAF function</strong></td>
<td>Complete (depending on sensor setup)</td>
<td>Frequency of scenarios with HAD/ADAS-function</td>
</tr>
<tr>
<td><strong>FOT/NDS measurement data scenarios</strong></td>
<td>Realistic Parameter Combinations</td>
<td>Limited, only with statistical population</td>
</tr>
<tr>
<td><strong>Proving ground (test track)</strong></td>
<td></td>
<td>Identification of human performance</td>
</tr>
<tr>
<td><strong>Simulation</strong></td>
<td>Identification of physical boundaries of the scenarios</td>
<td>Theoretical performance</td>
</tr>
<tr>
<td><strong>Accident data</strong></td>
<td>Limited (just two main accident participants)</td>
<td>Limited, only with statistical population</td>
</tr>
<tr>
<td><strong>Driving simulator</strong></td>
<td></td>
<td>Identification of human performance</td>
</tr>
</tbody>
</table>
Comparison between Real Traffic Data Collection Methodologies

### Preparation

<table>
<thead>
<tr>
<th>FOT/NDS</th>
<th>Infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase of data collection vehicle</td>
<td>Purchase of drone</td>
</tr>
<tr>
<td>Purchase of sensor systems (today as many sensors as possible)</td>
<td>Permit to operate drone</td>
</tr>
<tr>
<td>Data collection vehicle modification</td>
<td>Infrastructure modification</td>
</tr>
<tr>
<td>Data logger and data synchronisation</td>
<td></td>
</tr>
</tbody>
</table>

### Data Collection

<table>
<thead>
<tr>
<th>FOT/NDS</th>
<th>Infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permanent operation of vehicle and all sensors</td>
<td>Naturalistic driving/ no influence of road users and drivers</td>
</tr>
<tr>
<td>Data collection based on vehicle kilometres and speed</td>
<td>Complete 360° view and surroundings</td>
</tr>
<tr>
<td>Influence on driver and surrounding traffic by highly instrumented vehicle</td>
<td>No data privacy issues</td>
</tr>
<tr>
<td>Only detection of surrounding vehicles, occlusion of other road users</td>
<td>Measurement of all objects at flexible location</td>
</tr>
<tr>
<td>130 km per hour on highways</td>
<td>420 m coverage with 1 drone at highways</td>
</tr>
<tr>
<td></td>
<td>Typically about 3,000 km per hour on highways</td>
</tr>
</tbody>
</table>

### Data Evaluation

<table>
<thead>
<tr>
<th>FOT/NDS</th>
<th>Infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario data with perspective of the measurement vehicle</td>
<td>Highly accurate behaviour and scenario data</td>
</tr>
<tr>
<td>High quality raw sensor data and vehicle dynamics data</td>
<td>Raw sensor data of the test vehicle must be recorded separately</td>
</tr>
<tr>
<td>Measurement data quality depends on sensors and sensor processing algorithms</td>
<td>Different scenarios possible</td>
</tr>
<tr>
<td>High flexibility regarding location and conditions</td>
<td>Fast data processing due to AI algorithms</td>
</tr>
<tr>
<td>Slow data processing due to large data quantity</td>
<td>Data processing depending on number of collection locations</td>
</tr>
<tr>
<td></td>
<td>Accurate perception due to static infrastructure</td>
</tr>
<tr>
<td></td>
<td>Data processing depending on sensor set-up and number of collection locations</td>
</tr>
</tbody>
</table>

---

© PEGASUS | Final Event 13.05.2019 | fka
Process to Upload Data into the Database

Vehicle cutting in and braking

- Signals according to JSON definitions
- Minimum requirements on dataset
- Format: Mat or HDF5

Input Data
Data Handling in the PEGASUS Database

1. Integration by database mechanics

2. Storage in database

3. Generation of complete scenario space

4. Output generation & test concept

- Usage testing ground
- Usage driving simulator
- Usage simulation

Data with traffic events

- Source FOT
- Source Test drives
- Source xy
- Source Accident data

(Logical) scenario

(Logical) scenario with parameter distributions

Concrete scenario
Data Layer Model for Scenario Description

**Digital information:**
- e.g. V2X information on traffic signals, digital map data
- \( \Rightarrow \) Availability and quality of information communicated to ownship

**Environmental conditions**
- Light situation, weather (rain, snow, fog...) temperature
- \( \Rightarrow \) environmental influences on system performance

**Moving objects**
- Vehicles, pedestrians moving relatively to ownship
- \( \Rightarrow \) relevant traffic participants and their motion incl. dependencies

**Temporal modifications and events**
- Road construction, lost cargo, fallen trees, dead animal
- \( \Rightarrow \) temporary objects minimizing / influencing the driving space

**Road furniture and Rules**
- Traffic signs, railguards, lane markings, bot dots, police instructions
- \( \Rightarrow \) including rules, where to drive how

**Road layer**
- Road geometry. Road uneveness (openCRG)
- \( \Rightarrow \) physical description, no scenario logics

[1] Bock et al. 2018: Data Basis for Scenario-Based Validation of HAD on Highways

© PEGASUS | Final Event 13.05.2019 | fka
Main challenge for scenario concepts (Layer 4):
Which dynamic objects are relevant:
- Which objects need to be described?
- Which objects must be described how accurate?
**Scenario Concept**

- **A challenging vehicle induces a reaction of the subject vehicle to prevent an accident [1]**
  - Description based on accident reconstruction
  - Relational description from the subject vehicle perspective with relative paths
  - Considering the potential impact location (front, side, rear) and the initial position of a challenger vehicle

- **Challenger Vehicle**
  - 9 Scenario Types for influenced driving
  - 1 (non-) Scenario for uninfluenced driving

- **Further Vehicles:**
  - Occlude relevant information (“dynamic occlusion”)
  - Constrain possible actions of subject vehicle (“action constraints”)
  - Challenge the subject vehicle at the same time
  - Cause the challenger’s action (“challenger cause”)

- **Every other vehicle is not relevant (enough) for scenario-description**

---

[1] Bock et al. 2018: Data Basis for Scenario-Based Validation of HAD on Highways
## Scenario Types: (Potential) Frontal Impact

<table>
<thead>
<tr>
<th>Impact</th>
<th>Initial position</th>
<th>Path</th>
<th>Indication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front</td>
<td>1</td>
<td>A</td>
<td>Lead vehicle challenger</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>B</td>
<td>Slower turn into path challenger</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>C</td>
<td>Overtaking turn into path challenger</td>
</tr>
</tbody>
</table>

![Diagram showing vehicles and potential impact scenarios]
Action Constraints

- Further surrounding vehicles **constraint the possibilities to react**
- Distinguish between **Object, Gap and Blockage** for each location around the vehicle (front/rear/left/right)
Scenario Types: Overview

<table>
<thead>
<tr>
<th>Impact</th>
<th>Initial position</th>
<th>Path</th>
<th>Indication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front</td>
<td>1</td>
<td>A</td>
<td>Lead vehicle challenger</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>B</td>
<td>Slower turn into path challenger</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>C</td>
<td>Overtaking turn into path challenger</td>
</tr>
<tr>
<td>Side</td>
<td>2</td>
<td>D</td>
<td>Slower side swipe challenger</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>E</td>
<td>Side swipe challenger</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>F</td>
<td>Overtaking side swipe challenger</td>
</tr>
<tr>
<td>Rear</td>
<td>2</td>
<td>G</td>
<td>Slower Rear End Challenger</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>H</td>
<td>Rear end turning into path challenger</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>I</td>
<td>Rear end challenger</td>
</tr>
<tr>
<td>Non</td>
<td>-</td>
<td>-</td>
<td>Uninfluenced/Free driving</td>
</tr>
</tbody>
</table>
Data Processing

Determination of Scenario Affiliation
Data Handing in the PEGASUS Database

1. Integration by database mechanics
2. Storage in database
3. Generation of complete scenario space
4. Output generation & test concept

Output Data

Source FOT
Source Test drives
Source xy
Source Accident data

Data with traffic events
(Logical) scenario
Concrete scenario

(Logical) scenario with parameter distributions

Usage testing ground
Usage driving simulator
Usage simulation
Scenario output generation: Variants

- Measurement scenario generation:
  - Scenarios with complete parameter distributions
  - Concrete scenarios
  - Parameterizable scenarios (Concrete scenario where some scenario parameters are replaced by distributions)

- Filtering scenario data, e.g. taking the ODD of the function into account:
  1. Scenario Type
  2. Data source (NDS, FOT, Accident, Aerial data, ...)
  3. Scenario parameters (e.g. restricting velocity ranges)
  4. Sorting according to one scenario parameter: e.g. driving demand metric (criticality)

- Combinatorial scenario generation:
  - based on scenario concept (data-independent, but verified with data)

→ Conversion of scenario output into OpenScenario: OSC-Transpiler
  - Higher level description and parametrization of scenarios is transpiled (~converted) into OpenScenario
Testing of a Concrete Scenario in Simulation

- The selected concrete scenario can be reproduced in the simulation. A HAD-function integrated in the simulation can be tested.
- Here: “Slower turn into path challenger” (see screen 1)
Testing of a Concrete Scenario on the Test Track

- The selected concrete scenario can be reproduced on the test track. A HAD-function integrated in VUT can be tested.
- Here: “Slower turn into path challenger” (see screen 1)
Thank you!

© PEGASUS | Final Event 13.05.2019 | fka