Big Engineering Data Analytics with MATLAB

Dmitrij Martynenko
Application Engineer
dmitrij.martynenko@mathworks.de

MathWorks
AUTOMOTIVE CONFERENCE 2015
How do you define Big Data?

“Any collection of data sets so large and complex that it becomes difficult to process using … traditional data processing applications.”

(General Definition)

“Any collection of data sets so large that it becomes difficult to process using traditional MATLAB functions, which assume all of the data is in memory.”

(MATLAB)
# Your Big Data Sources

<table>
<thead>
<tr>
<th>Test or Simulation Method</th>
<th>Engine Control Module (ECM)</th>
<th>Engine</th>
<th>Exhaust Gas Aftertreatment</th>
<th>Vehicle and Driver</th>
<th>Topographic Drive Route</th>
<th>Variability of Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fleet of Test Vehicles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehicle in Loop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aftertreatment in Loop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engine in Loop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECM in Loop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pure Simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Vehicle Component**

- **Hardware**
- **Software**

---

*MathWorks AUTOMOTIVE CONFERENCE 2015*
Challenges Analyzing Fleet Data

**More Data** ⇒ Better Understanding of Field Conditions
More Interesting Events

**Challenges:**
- Big Data
- Needle in the Haystack
- Testing Ideas
- Knowledge Transfer
Data Analytics with MATLAB

- Merging/aligning data
- Working with Big Data
  - MapReduce
  - Hadoop
- Taking analytics to production
Fleet Data Event Detection

- Parse data, find sudden deceleration
- MapReduce workflow

<table>
<thead>
<tr>
<th>Device Time</th>
<th>Engine RPM</th>
<th>Fuel flow rate</th>
<th>Torque</th>
<th>Speed (ODB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-Jan-XXXX</td>
<td>1445.75</td>
<td>0.0537</td>
<td>5.26</td>
<td>3.10</td>
</tr>
<tr>
<td>15-Jan-XXXX</td>
<td>1445.75</td>
<td>0.0537</td>
<td>5.74</td>
<td>0.48</td>
</tr>
<tr>
<td>15-Jan-XXXX</td>
<td>1458.5</td>
<td>0.0107</td>
<td>11.10</td>
<td>6.21</td>
</tr>
<tr>
<td>15-Jan-XXXX</td>
<td>1458.5</td>
<td>0.0107</td>
<td>11.10</td>
<td>0.59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Acceleration</th>
<th>Deceleration</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-Jan-XXXX</td>
<td>-2.62</td>
<td>0</td>
</tr>
<tr>
<td>15-Jan-XXXX</td>
<td>4.49</td>
<td>0</td>
</tr>
<tr>
<td>15-Jan-XXXX</td>
<td>1.24</td>
<td>0</td>
</tr>
<tr>
<td>15-Jan-XXXX</td>
<td>-5.62</td>
<td>1</td>
</tr>
</tbody>
</table>

Acceleration: -5.62
Analysis Domains

Statistics
- Summary Statistics
- Regression, ANOVA, Machine Learning

Signal Processing
- Sound quality analysis
- LIDAR analysis

Image Processing
- Active Safety

Location/Mapping
- Analyzing GPS Data
- Custom Visualizations
Taking MATLAB to Production

- Code Generation
- Report Generation
- Application Packaging

- Embedded System
- Reports
- Applications
- Business System
Analyzing Big Data using MATLAB

- Operate on Big Data from MATLAB with MapReduce
- Quickly try out an idea, then iterate
- Transfer your results by taking MATLAB to production
Additional Resources

Machine Learning
mathworks.com/machine-learning

Parallel Computing

MapReduce
Questions?