

Henry Clausen, David Aspinall

Examining traffic micro-structures for model probing

WTMC 2021



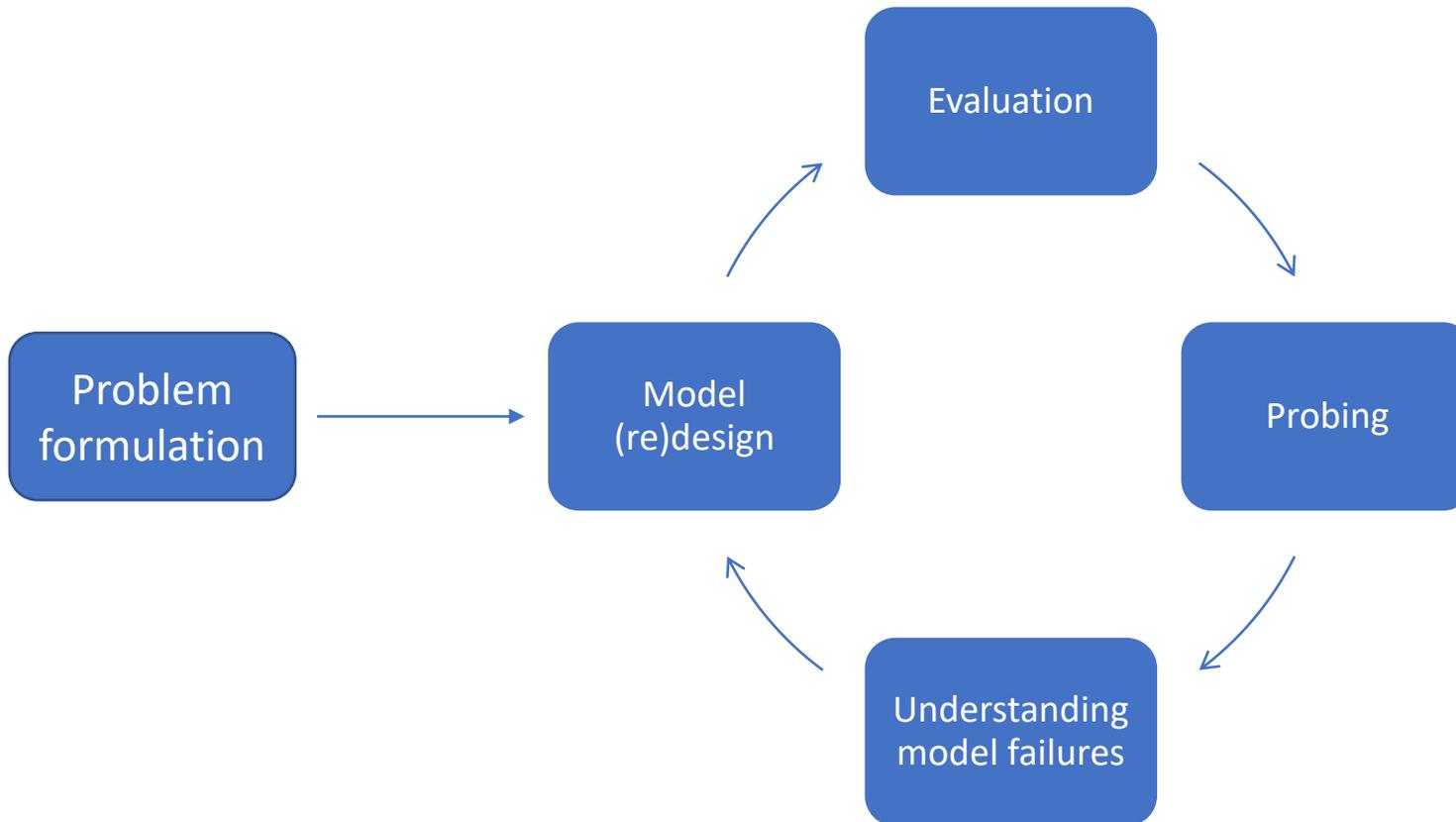
THE UNIVERSITY *of* EDINBURGH
informatics



**The
Alan Turing
Institute**



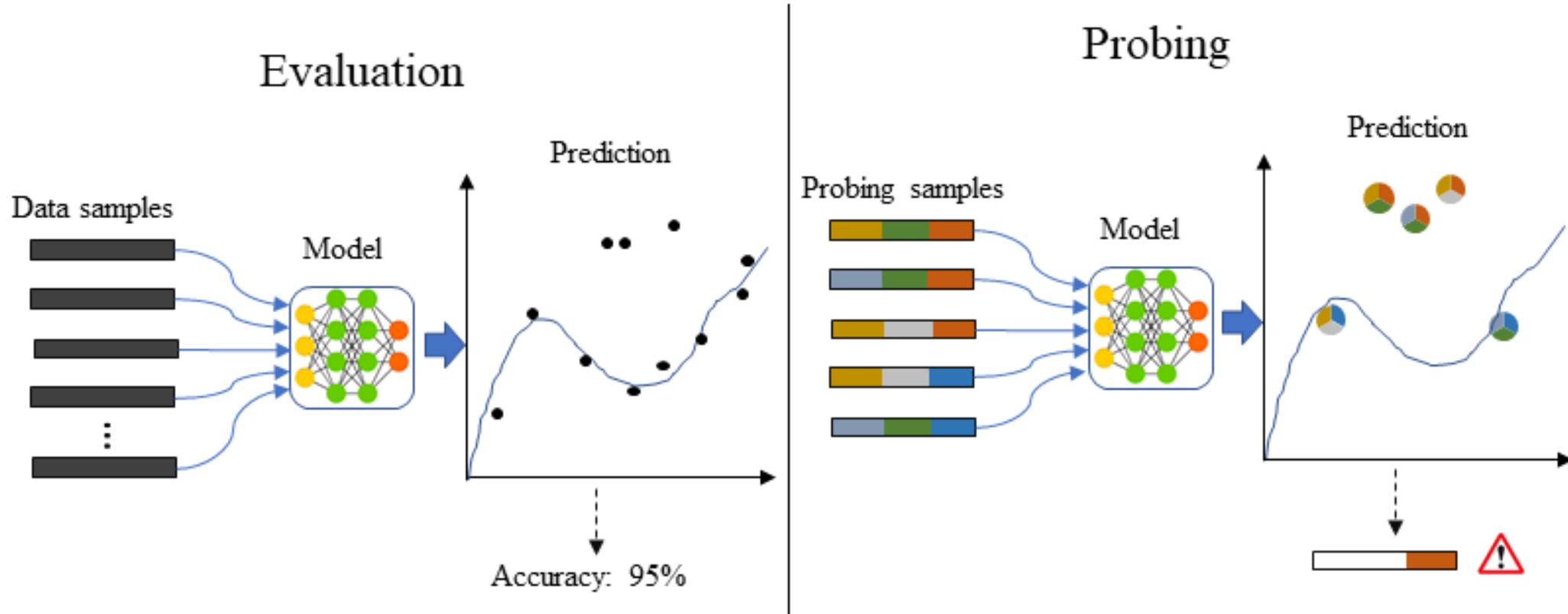
Machine learning progress



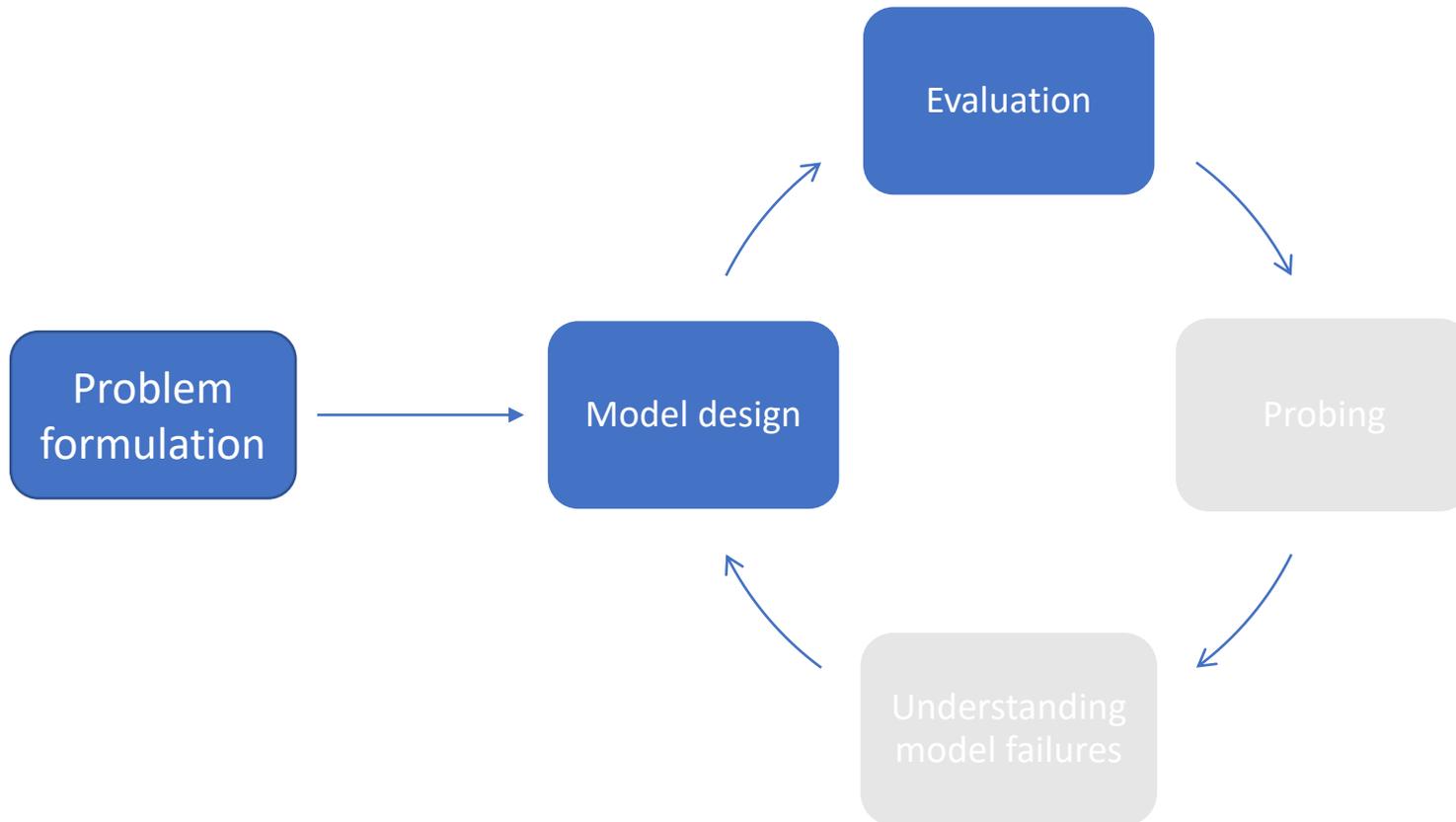
Prominent failures

- Ambiguous words in Translations
→ Attention layer
- Object sizes in video enhancement
→ Multi-scale encoders

Model evaluation vs probing



Machine learning progress in NID

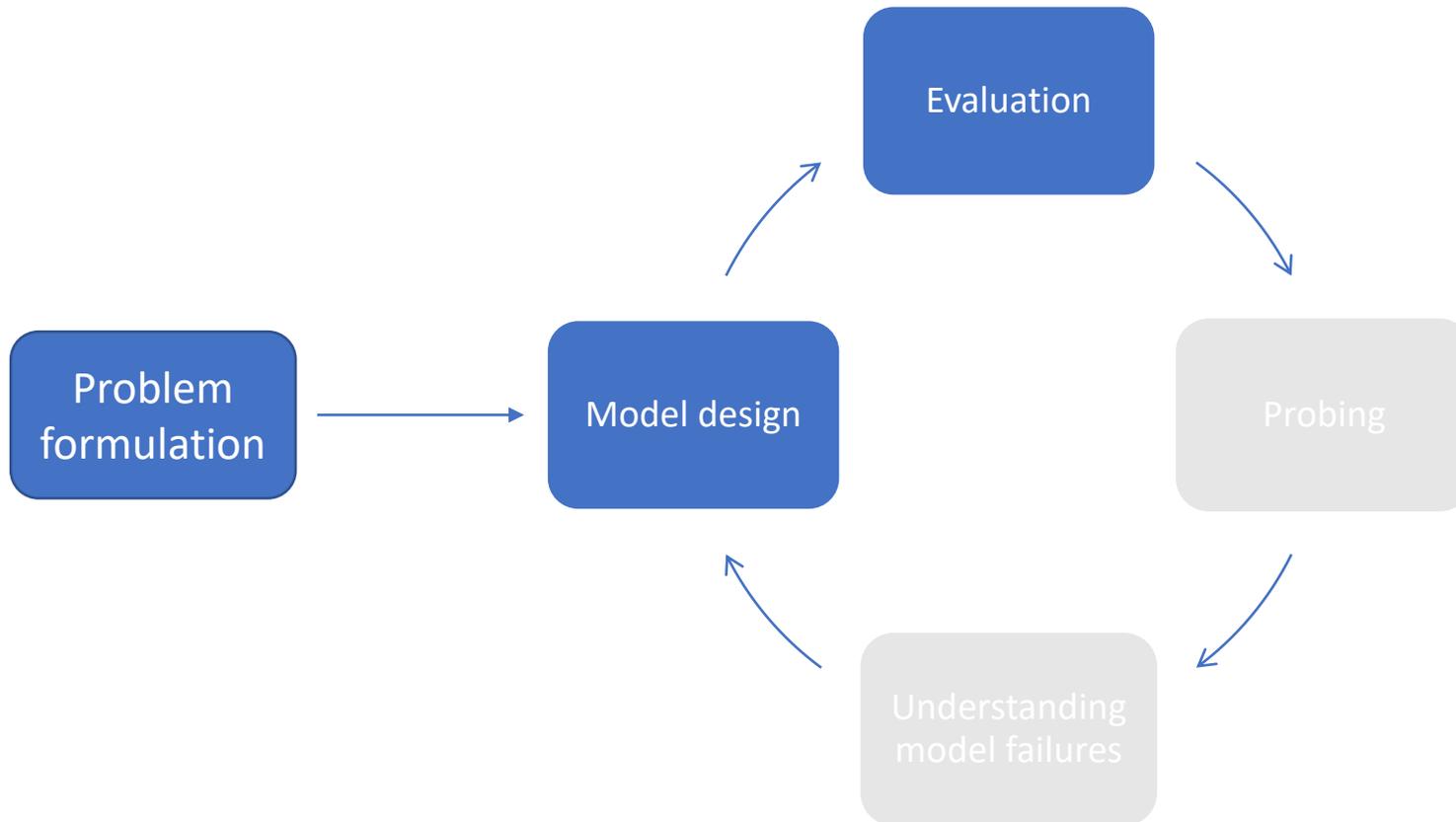


NID-datasets

- Sparse labelling
- Difficult to read
- Hard to alter specific structures



Machine learning progress in NID



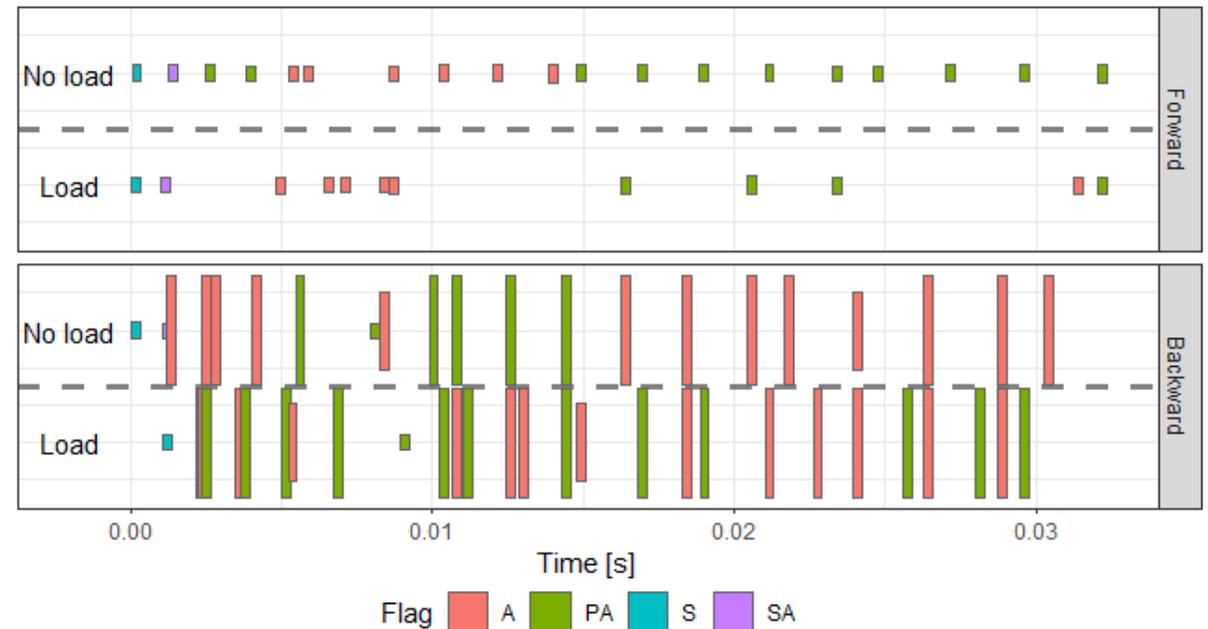
Our case-study

- Probe two NID-methods
- Identify microstructures related to failures
- Measure improvements

Traffic microstructures

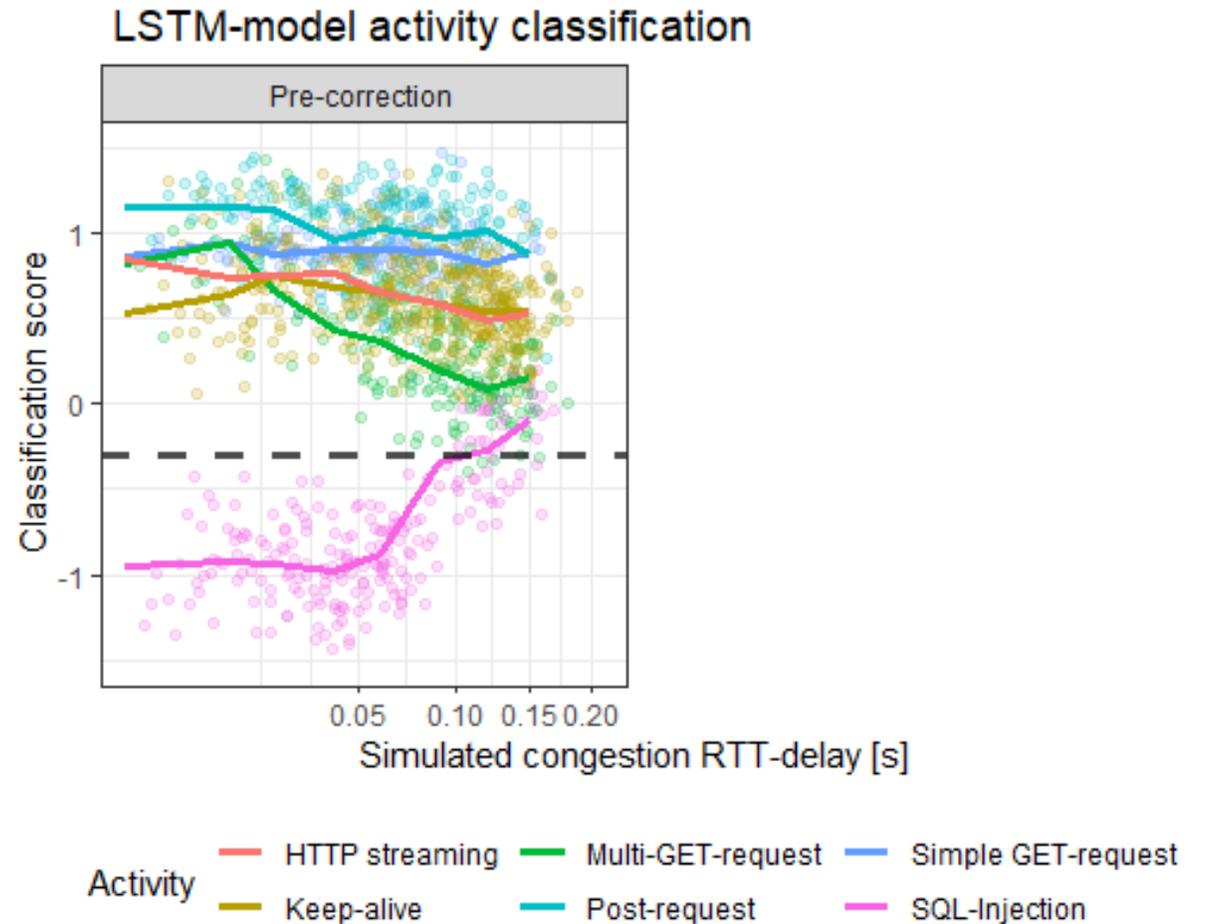
- Short-term structures at packet or connection level
- manifest in IATs, frame sizes, flags etc.
- Altered by factors such as protocols, congestion, implementation
- Control with DetGen-tool
Clausen et al., SecureComm 2021

FTP-connection comparison under load



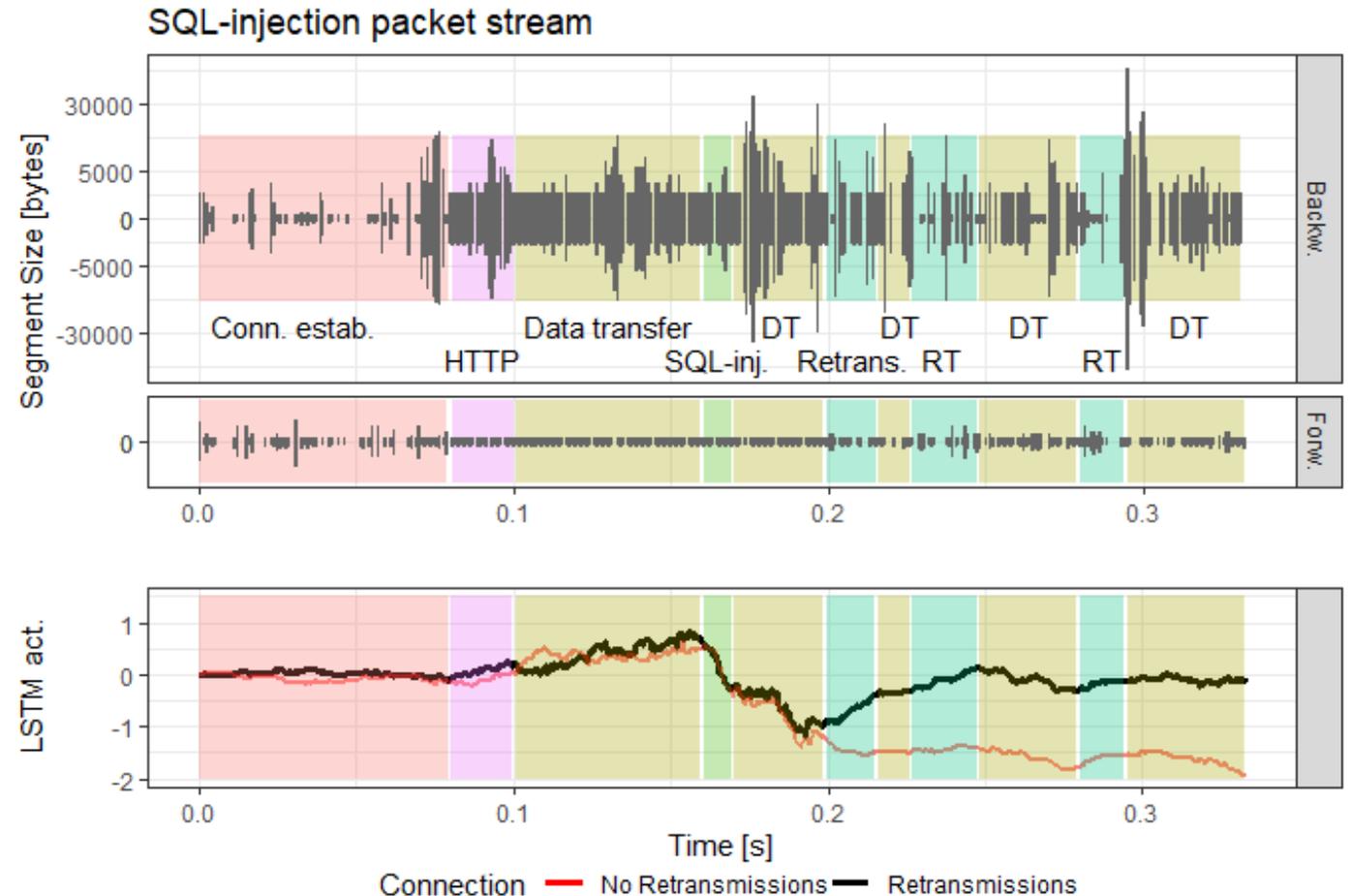
Examining a traffic classifier

- Packet-stream LSTM-classifier by Hwang et al. 2019
 - Detect SQL-injections
- Train on CICIDS-17 data (85%) + DetGen traffic (15%)
 - 96% DR, 2.7% FPR
- Probe with randomized traffic + structure labels
 - Correlation between misclassifications and latency



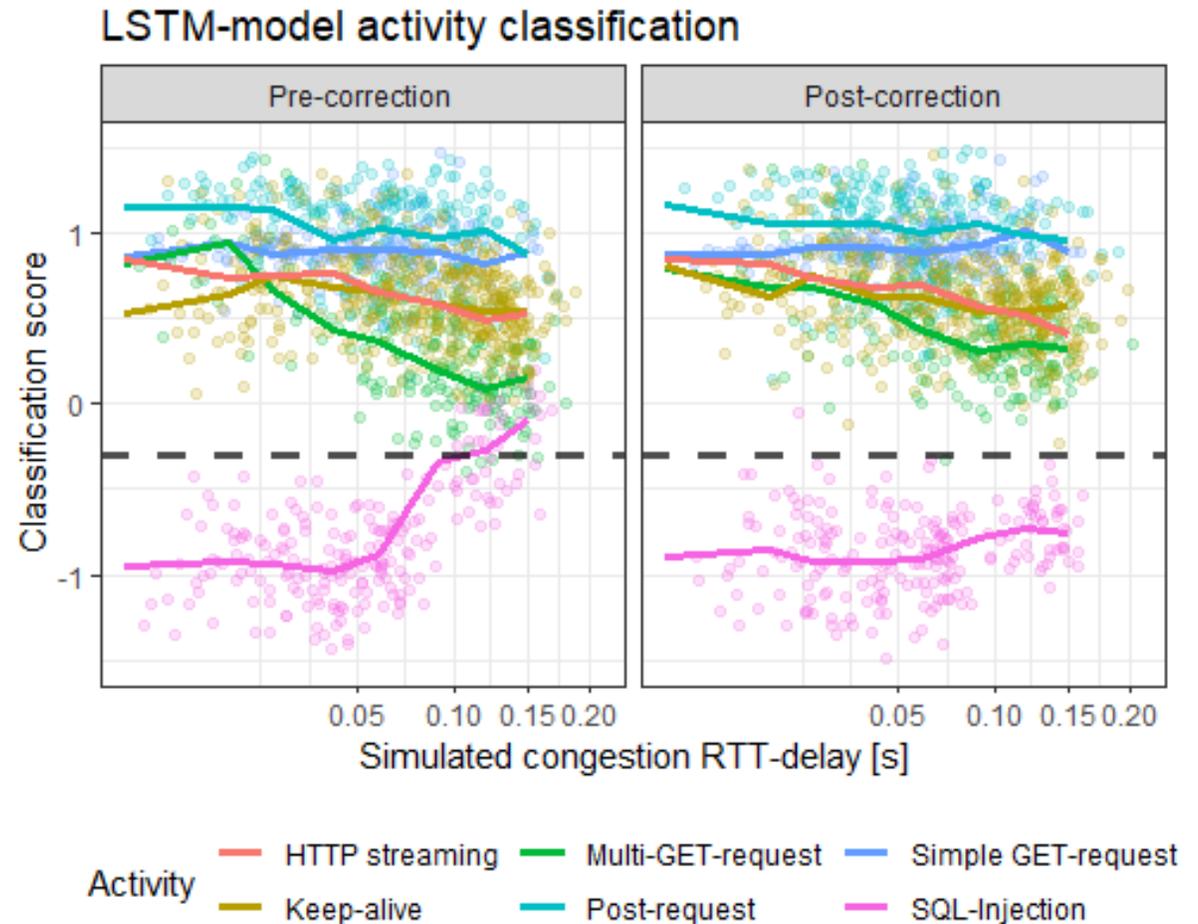
Examining a traffic classifier

- Generate two SQL-injection connections
 - Constant microstructures
 - One with high latency
- Retransmission sequences deplete activation
- Filter RT-sequences
 - 98% DR and 0.4% FPR



Examining a traffic classifier

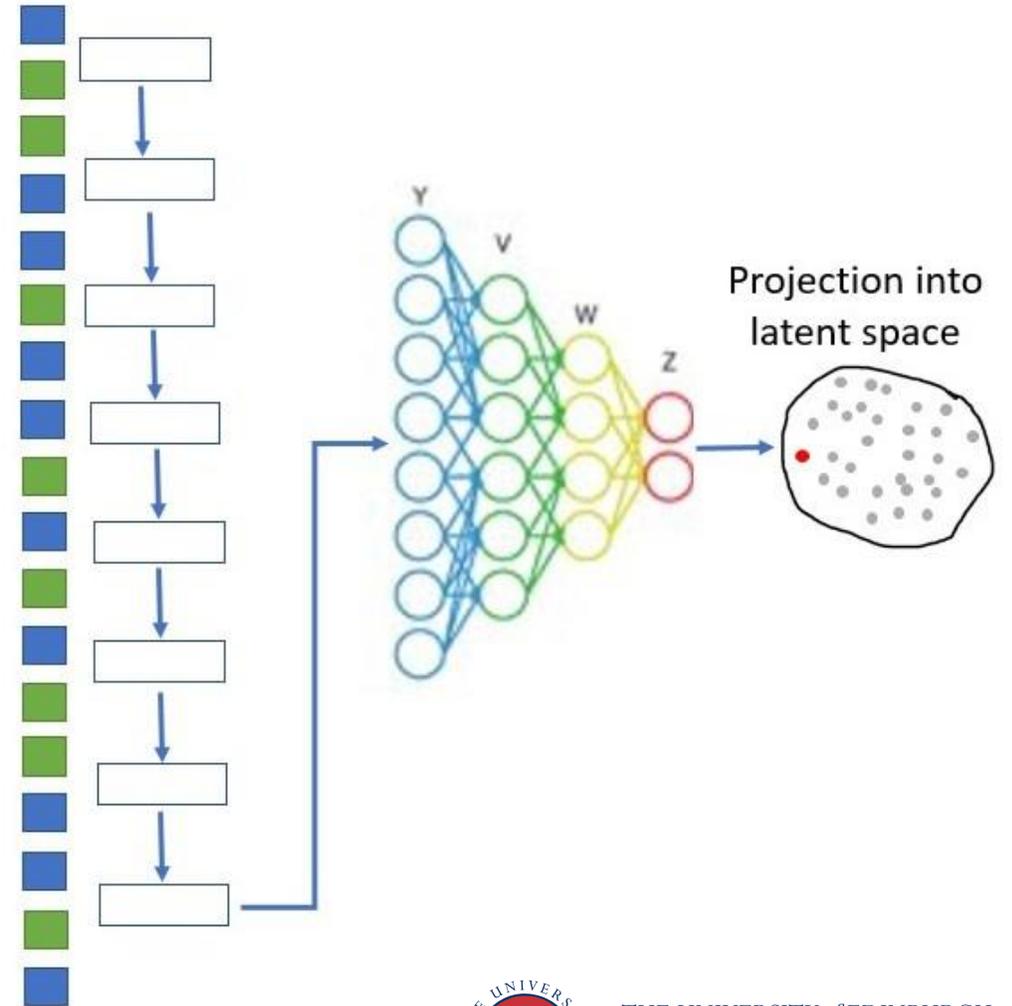
- Generate two SQL-injection connections
 - Constant microstructures
 - One with high latency
- Retransmission sequences deplete activation
- Filter RT-sequences
 - 98% DR and 0.4% FPR



Projection sensitivity

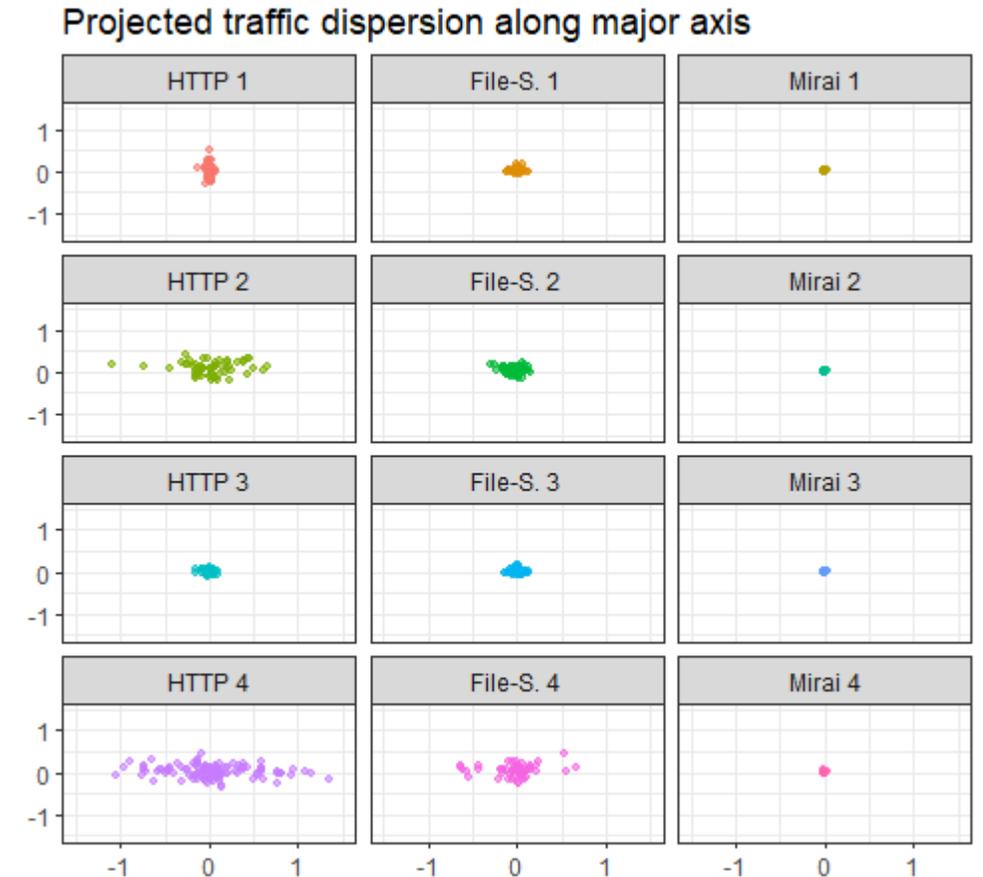
- Kitsune 2018
 - Seq-encoding for anomaly detection
 - Botnet, man-in-middle, Brute-force,...
- Traffic groups with constant settings
- Projections should be consistent
- Sensitive to
 - connection IATs
 - half-open connections

TCP-connection



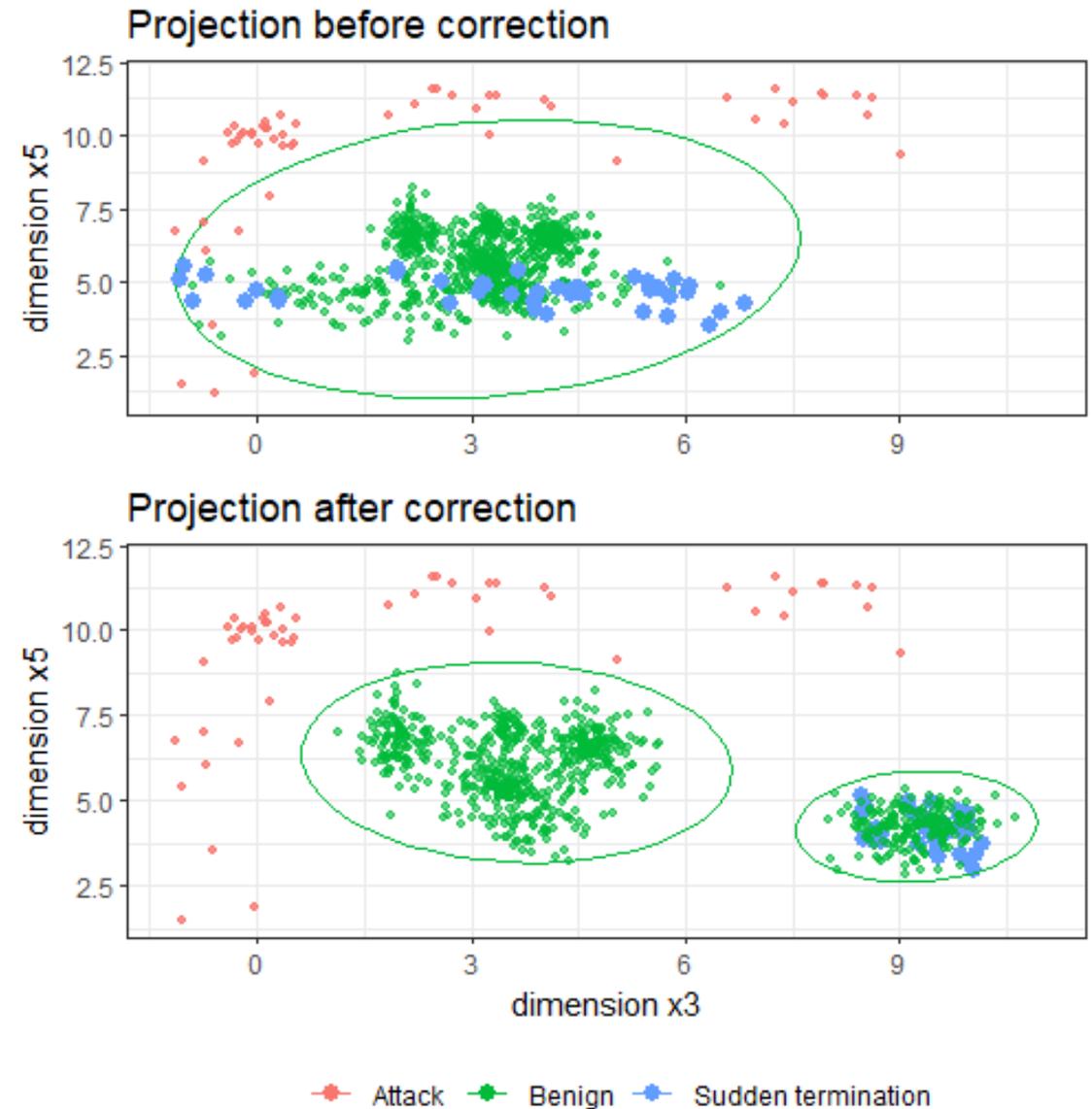
Projection sensitivity

- Kitsune 2018
 - Seq-encoding for anomaly detection
 - Botnet, man-in-middle, Brute-force,...
- Traffic groups with constant settings
- Projections should be consistent
- Sensitive to
 - connection IATs
 - half-open connections



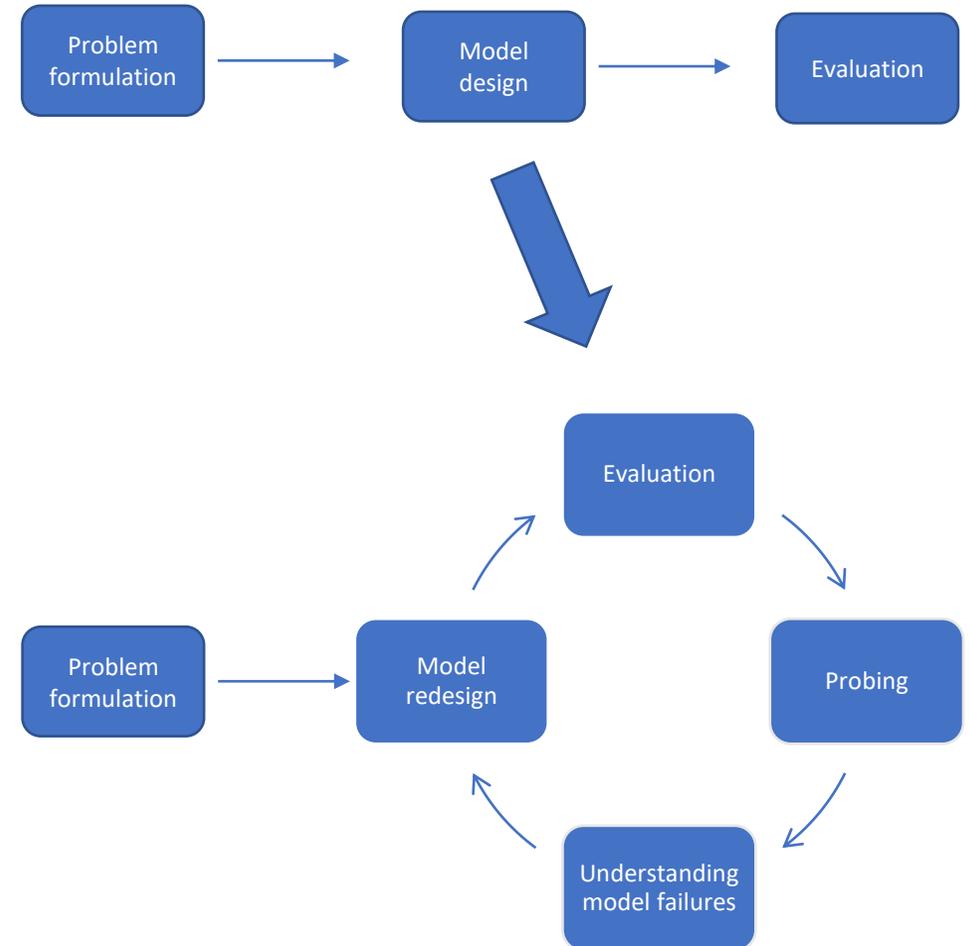
Projection sensitivity

- Kitsune 2018
 - Seq-encoding for anomaly detection
 - Botnet, man-in-middle, Brute-force,...
- Traffic groups with constant settings
- Projections should be consistent
- Sensitive to
 - connection IATs
 - half-open connections



Conclusion

- Targeted probing can identify model failures
- Labelling for misclassification correlation
- Control traffic microstructures
 - Randomise for broad probing
 - Reduce variations for close examination
- github.com/detlearsom/DetGen



Controlling traffic microstructures

DetGen Clausen et al., SecureComm 2021

- Traffic generation tool
- Controlling and labelling microstructures:
 - Performed task/application
 - Implementations
 - Congestion
 - Failures
 - ...
- github.com/detlearsom/DetGen

