Scalable Clustering of Traffic Scenarios Without a Mental Model

Bachelor's Thesis or Master's Thesis

Examiner: Prof. Dr. Alexander Pretschner
Supervisor: Florian Huber, Simon Speth
Email: `florian.maximilian.huber@tum.de`
Phone: `+49 (89) 289 - 17334`

Starting date: As soon as possible

Informatics 4 – Chair of Software and Systems Engineering
Prof. Dr. Alexander Pretschner

TUM School of Computation, Information and Technology
Technical University of Munich

Boltzmannstr. 3
85748 Garching b. München

`https://www.cs.cit.tum.de/en/sse/`

Context

Scenario-based testing is an established system-level testing technique in the safety argumentation of automated driving systems [1]. For the scenario term, three levels of abstraction are distinguished: *functional*, *logical* and *concrete* scenarios [2]. Similar and recurring traffic constellations are grouped into *functional scenarios*, so-called **scenario types**. For test case generation, these (textual) scenario types can be parameterized to so-called *logical scenarios*. Each instantiation in this huge space corresponds to a **concrete scenario instance**.
Test scenario generation requires knowledge about the recurring traffic cases, i.e., scenario types. This can be done either in a manual expert-driven or a data-driven way (see [3] for an overview). However, the manual expert-driven determination of scenario types poses the threat of overlooking certain types, e.g., due to an inadequate expert's mental model [3]. Therefore, complementary to the expert-driven scenario type derivation, clustering approaches are applied to determine scenario types from scenario instances in recorded real traffic data. However, these either do not scale (e.g., due to a quadratic complexity in the similarity measure and the determination of clusters [3]) or pre-determine the number of clusters (e.g., [4]), which is unknown *a priori*. To continually reduce the number of unknown scenario types (required for certification [1]) and converge to a *complete* list of scenario types [5] (which may evolve over time [6]), a continuous collection of scenarios is required. As the continuous collection results in an ever-growing number of scenario instances of tremendous scale, we aim for a **scalable approach for the derivation of scenario types**.

Goal

To converge over time to a complete catalog of scenario types, as well as to a representative collection of real scenario instances for each scenario type, a scalable approach is necessary. **Autoencoders** show promising results in **learning latent-space representation of time-series** when used for **anomaly, resp., novelty, detection** (see [7] for an overview), i.e., detecting patterns that deviate from the normal (known traffic) behavior. Therefore, we propose a solution based on an autoencoder that (1) does not require a pre-determination of the (*a priori* unknown) number of scenario types but instead determines the number of clusters, resp., scenario types in a data-driven way using latent-space clustering and (2, optional depending on thesis type) that scales to a large number of scenario instances by filtering out *uninteresting* scenario instances, i.e., scenario instances which either do not belong to a novel scenario type (a), are not at the border of a cluster and therefore, not helpful in limit testing (b), or belong to an underrepresented cluster (c). To showcase the approach, it is applied to a real traffic dataset (e.g., inD [8] or highD [9]) and evaluated using scientific methods.

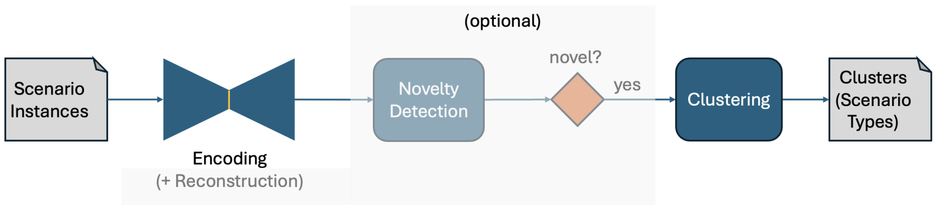

Figure 1: Simplified data flow of proposed approach

Prerequisites

Participation in courses covering relevant knowledge on deep learning is required, participation in courses covering software testing is a plus. Knowledge in deep learning frameworks (e.g., PyTorch or TensorFlow) is required.

Informatics 4 – Chair of Software and Systems Engineering
Prof. Dr. Alexander Pretschner

TUM School of Computation, Information and Technology
Technical University of Munich

Boltzmannstr. 3
85748 Garching b. München

`https://www.cs.cit.tum.de/en/sse/`

Working Plan

1. Familiarize yourself with the literature on scenario-based testing of automated and autonomous driving systems in general, and specifically with the literature on scenario clustering (optional: and novelty detection).
2. Extract scenario instances from recorded real traffic dataset, e.g., highD [9] or inD [8] dataset.
3. Train autoencoder on extracted scenario instances, including:
   (a) Select a suitable neural network architecture for time-series trajectory data (see, e.g., [7] for an overview).
   (b) Apply data augmentation to allow for a generalization in the learning process.
   (c) Select hyper-parameters systematically.
4. Select clustering technique (e.g., k-means), determine number of clusters in a data-driven way (e.g., using knee/elbow detector [3]), and apply clustering to traffic dataset.
5. Optional: Determine threshold to detect novel scenarios, e.g., using reconstruction threshold based on clustering results (from step 4) or threshold regarding the distance to the centroid in the latent space.
6. Evaluate, e.g., regarding:
   (a) Manual inspection of derived clusters, e.g., regarding their number of clusters for each scenario type.
   (b) Compare clustering results to traditional similarity-measure-based baseline [3].
   (c) Computational performance (benefit) compared to traditional clustering approach.
   (d) Optional: Accuracy in the detection of novel scenario types, e.g., measured based on recall, precision, etc.
7. Write the thesis report.

Deliverables

- Source code of the implementation.
- Technical report with comprehensive documentation of the implementation, i.e. report, wiki, examples, tests, usage instructions.
- Final thesis report written in English and in conformance with TUM guidelines.
- Presentation of the work at the chair (2-3 weeks after submission).

References

[1] *ISO 21448:2022 – Road vehicles — Safety of the intended functionality*, ISO Standard, Geneva, Switzerland, 2022.

[2] T. Menzel, G. Bagschik, and M. Maurer, "Scenarios for development, test and validation of automated vehicles," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, Changshu: IEEE, Jun. 2018, pp. 1821–1827. DOI: 10.1109/IVS.2018.8500406.

[3] F. Hauer, I. Gerostathopoulos, T. Schmidt, and A. Pretschner, "Clustering traffic scenarios using mental models as little as possible," in *2020 IEEE Intelligent Vehicles Symposium (IV)*, Las Vegas, NV, USA: IEEE, Oct. 19, 2020, pp. 1007–1012. DOI: 10.1109/IV47402.2020.9304636.

[4] N. Roßberg, M. Neumeier, S. Hasirlioglu, M. E. Bouzouraa, and M. Botsch, "Assessing the completeness of traffic scenario categories for automated highway driving functions via cluster-based analysis," in *2025 IEEE Intelligent Vehicles Symposium (IV)*, Cluj-Napoca, Romania: IEEE, Jun. 22, 2025, pp. 1787–1794. DOI: 10.1109/IV64158.2025.11097767.

[5] F. Hauer, T. Schmidt, B. Holzmuller, and A. Pretschner, "Did we test all scenarios for automated and autonomous driving systems?" In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Auckland, New Zealand: IEEE, Oct. 2019, pp. 2950–2955. DOI: 10.1109/ITSC.2019.8917326.

[6] A. Pretschner, F. Hauer, and T. Schmidt, "On scenario-based testing of cyber-physical systems," in *Applicable Formal Methods for Safe Industrial Products: Essays Dedicated to Jan Peleska on the Occasion of His 65th Birthday*, A. E. Haxthausen, W.-l. Huang, and M. Roggenbach, Eds. Cham: Springer Nature Switzerland, 2023, pp. 68–82. DOI: 10.1007/978-3-031-40132-9_5.

[7] Z. Zamanzadeh Darban, G. I. Webb, S. Pan, C. Aggarwal, and M. Salehi, "Deep learning for time series anomaly detection: A survey," *ACM Computing Surveys*, vol. 57, no. 1, pp. 1–42, Jan. 31, 2025. DOI: 10.1145/3691338.

[8] J. Bock, R. Krajewski, T. Moers, S. Runde, L. Vater, and L. Eckstein, “The ind dataset: A drone dataset of naturalistic road user trajectories at german intersections,” in *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 1929–1934. DOI: 10.1109/IV47402.2020.9304839.

[9] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, “The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems,” in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 2118–2125. DOI: 10.1109/ITSC.2018.8569552.

Informatics 4 – Chair of Software and Systems Engineering
Prof. Dr. Alexander Pretschner

TUM School of Computation, Information and Technology
Technical University of Munich

Boltzmannstr. 3
85748 Garching b. München

https://www.cs.cit.tum.de/en/sse/