EXPERT SYSTEM BASED FAULT DIAGNOSIS FOR RAILWAY POINT MACHINES

Susanne Reetz[1*], Thorsten Neumann[1], Gerrit Schrijver[2], Arnout van den Berg[2], Douwe Buursma[2]

[1]Institute of Transportation Systems, German Aerospace Center (DLR), Germany
[2]Strukton Rail, The Netherlands
*susanne.reetz@dlr.de

ABSTRACT

To meet the increasing demands for availability at reasonable cost, operators and maintainers of railway point machines are constantly looking for innovative techniques for switch condition monitoring and prediction. This includes automated fault root cause diagnosis based on measurement data (such as motor current curves) and other information. However, large, comprehensive sets of labeled data suitable for popular supervised learning approaches are not yet available. Existing data-driven approaches focus only on distinguishing a few large, broad fault categories that can be distinguished with limited measurement data. There is great potential in hybrid models that combine expert knowledge with multiple sources of information to automatically identify failure causes at a much more detailed level. This poster presents a Bayesian network diagnostic model for determining the root causes of faults in point machines, based on expert knowledge and few labeled data examples from the Netherlands. Human-interpretable current curve features and other information sources (e.g., past maintenance actions) are used as evidence. The result of the model is a ranking of the most likely failure causes with associated probabilities in the form of a fuzzy multi-label classification, which is directly aimed at providing decision support to maintenance engineers. The validity and limitations of the model are demonstrated through a scenario-based evaluation and a brief analysis using information theoretic measures. The approach can be generalized to the development of similar models for various complex technical assets under similarly challenging conditions.

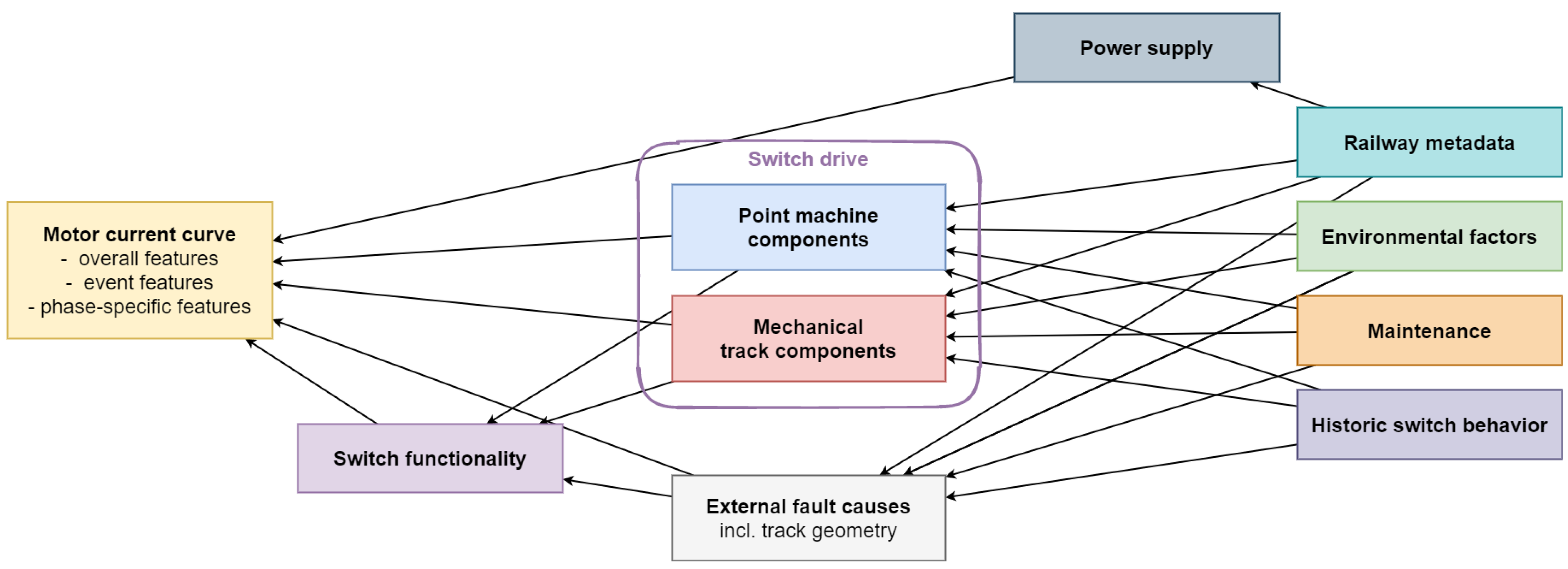

Figure 6. Overview of the node groups of the diagnostic model and their (causal) relations. The full model comprises 66 nodes, 178 states, 105 links and 661 free parameters.

CHALLENGE

Perform fault diagnosis in the sense of fuzzy multi-label classification for railway point machines, using motor current measurement data and various other information, to support maintenance operators in their decisions.

METHODOLOGY

Bayesian networks are probabilistic graphical models with powerful reasoning processes (see Figure 1), that are traceable and provide interpretable results. They can be developed both manually or data-driven.

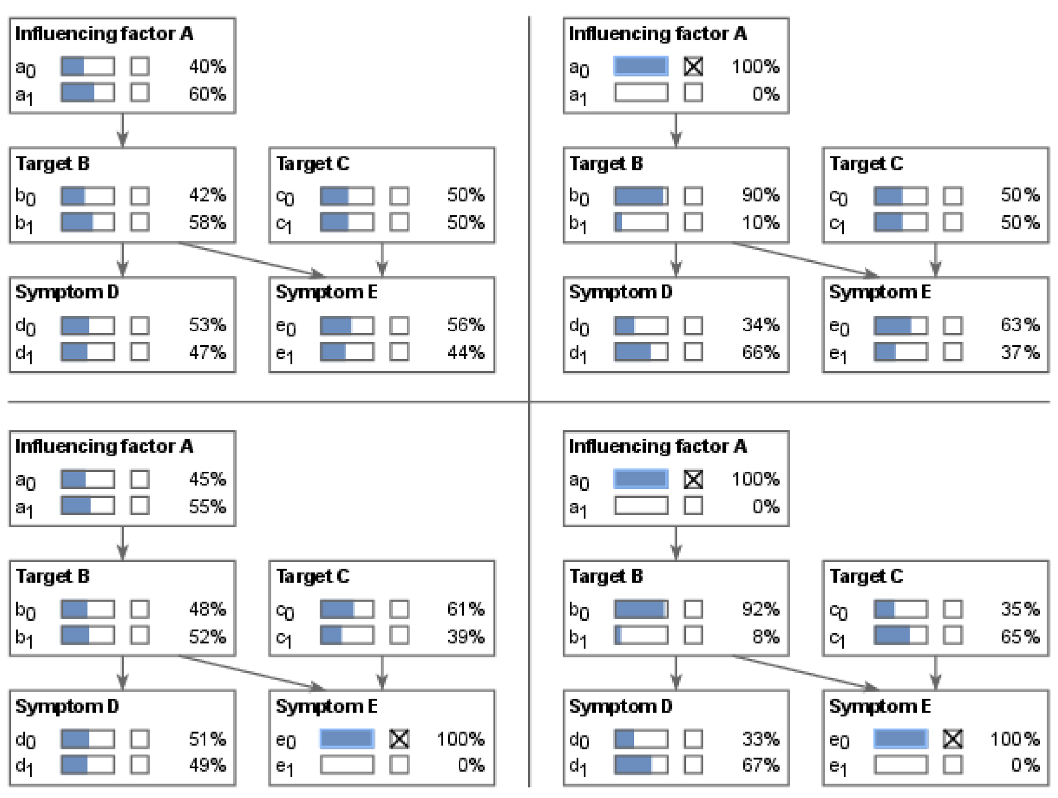

Figure 1. Reasoning processes in Bayesian networks. No evidence case (upper left), causal reasoning (upper right), diagnostic reasoning (lower left) and intercausal reasoning (lower right).

INFORMATION SOURCES

Due to a lack of larger amounts of labeled data, the development process is based on the following:

- Historical data (labeled and unlabeled), see Figure 2
- Experimental data, see Figure 3 & 4
- Expert knowledge, FMECA, manuals and literature

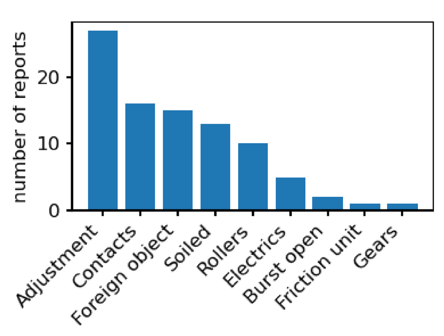

Figure 2. Historical fault statistics.

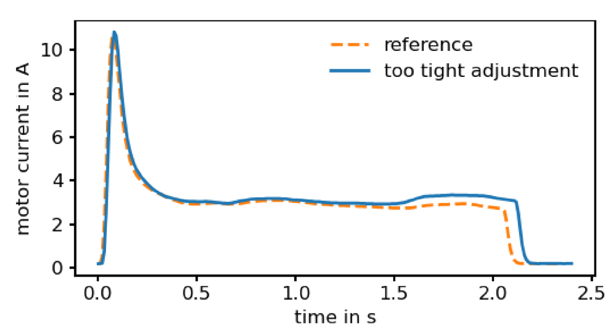

Figure 3. Too tight adjustment.

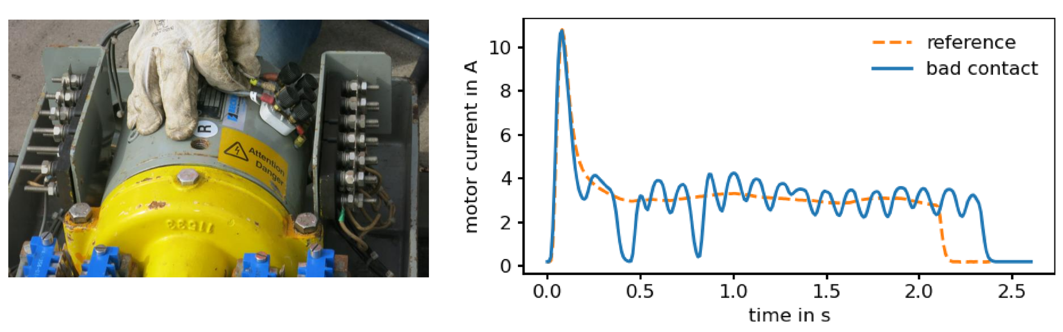

Figure 4. Experimental simulation of a bad contact.

DEVELOPMENT PROCESS

Figure 5 (violet box) shows the generalized manual development process. Main considerations are:

- Structural design:
  - Component-based vs. functional view
  - Granularity: "As fine as necessary, as broad as possible."
- Temporal scope: dynamic Bayesian networks (short-term dependencies) or specific nodes that represent aggregated historical development (long-term dependencies)
- Parametrization: reduce number of free parameters by using local probability distributions such as NOISY-MAX whenever possible
- Scenario-based evaluation: evaluate model diagnosis on various evidence sets that cover a wide range of real-life fault scenarios

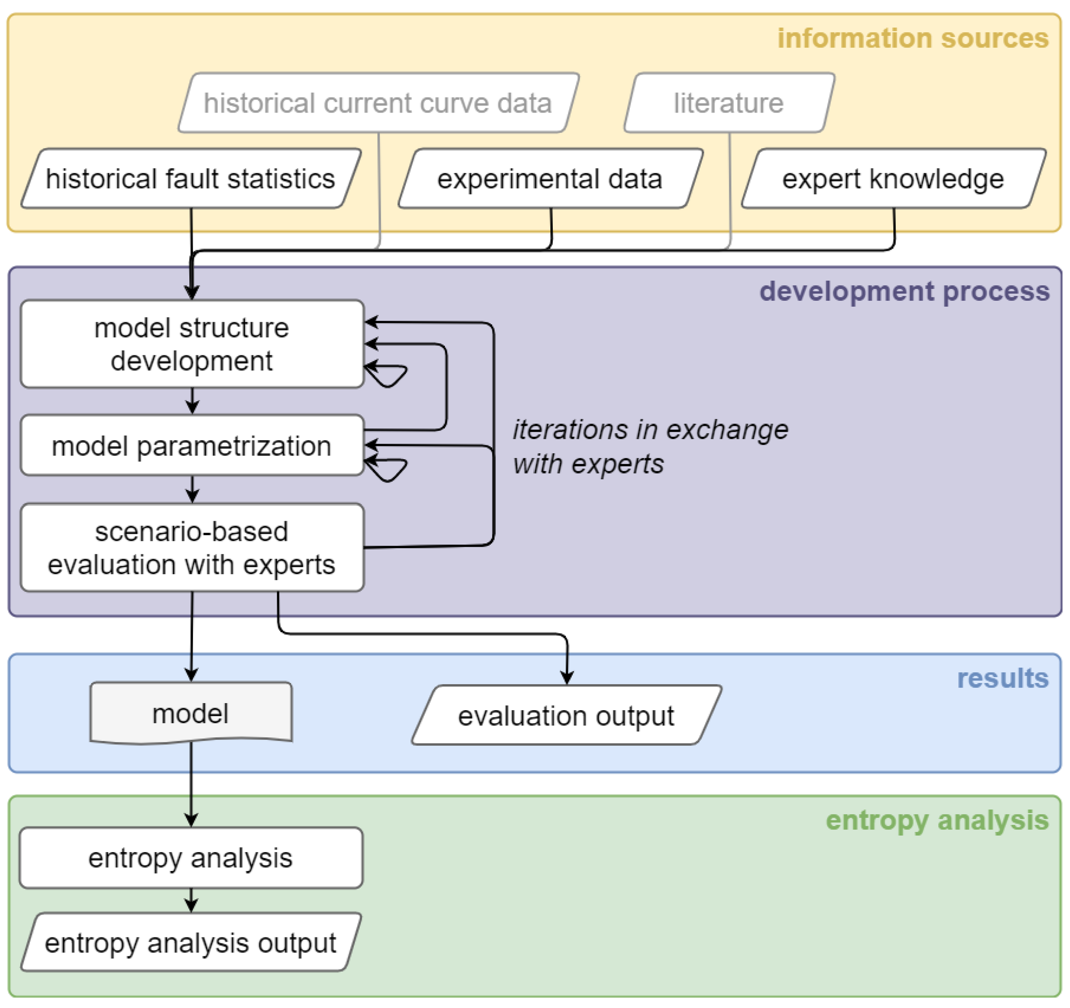

Figure 5. Development process with information sources, resulting model and subsequent analysis.

MODEL

Figure 6 shows an overview of the node groups of the Bayesian network diagnostic model and their (causal) relations. The node groups are split into the following layers (from right to left) :

- Influencing factors: railway metadata, environmental factors, maintenance, historic switch behavior and power supply (the latter could be considered a target)
- Targets: point machine components, mechanical track components, external fault causes and switch functionality
- Symptoms: motor current curve

In application, evidences for the influencing factor and symptom nodes are collected and fed into the network. The resulting probabilities for the fault states in the target nodes are presented to maintenance operators in the form of a ranking (fuzzy multi-label classification).

EXEMPLARY SCENARIOS

Figure 7 shows the motor current curves and probabilities of relevant fault states in target nodes of two scenarios. In the left column:

- Base scenario: Evidence on motor current curve features, including an overall hump in the movement phase.
- Variant (a): Base scenario plus information on recent maintenance on the mechanical parts of the track.
- Variant (b): Base scenario plus no recent tamping, high load, and both bad substructure and bad ballast quality.

In the right column :

- Base scenario: Evidence on motor current curve features, with a humped inrush and locking phase.
- Variant (a): Base scenario plus additional information on build-up precipitation inside the point machine.
- Variant (b): Variant (a) plus historical heavy locking of the switch.

The more evidences for different sources available, the more precise and reliable the model's results are.

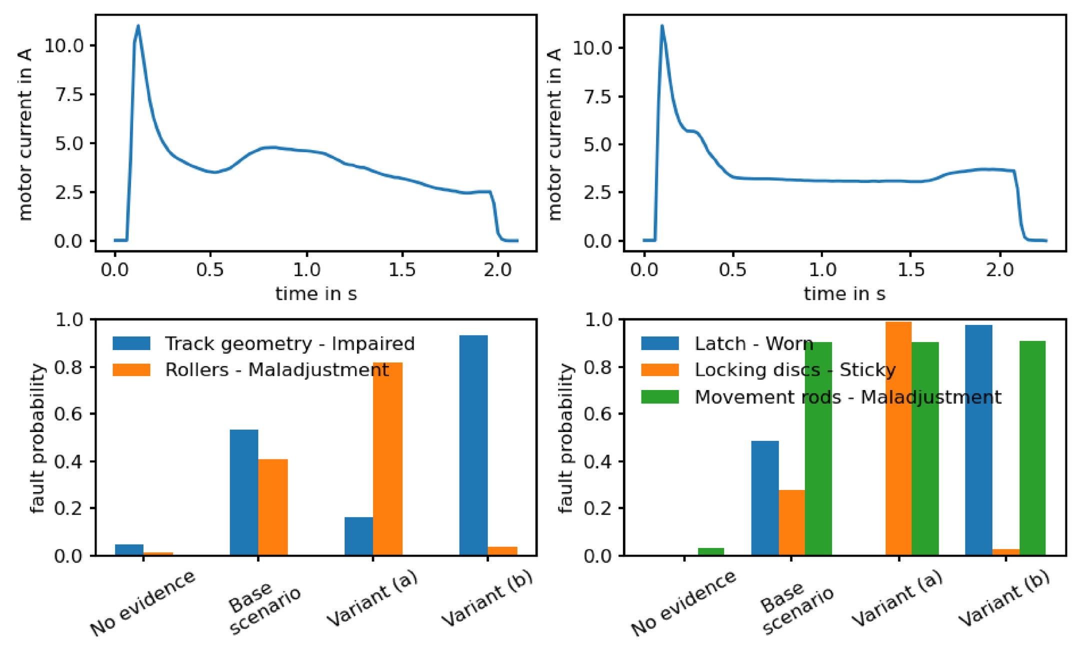

Figure 7. Exemplary scenarios.

ENTROPY ANALYSIS

Supplementary to scenario-based evaluation, the relative entropy reduction per evidence node for each target node can be calculated.

$$\frac{H(X)-H(X|Y)}{H(X)},\ H(X)=-\sum_{x} p(x)\log\big(p(x)\big)\geq 0,\ H(X|Y)=-\sum_{x,y} p(x,y)\log\big(p(x|y)\big)\leq H(X)$$

For the presented model, for each target node, the evidence nodes that have the most impact in the scenario-based evaluation also result in the highest entropy reduction.

INFO BOX

Switches and crossings are responsible for 19% of the infrastructure faults of category 1-4 in Germany in 2022, and 21% of the faults in category 1-2, which have to be addressed immediately. Most of the infrastructure faults are caused by the control, command and signaling system (55% for cat. 1-4 and 52% for cat. 1-2). Delay minutes per train kilometer attributed to the network in 2022 are the highest since records began in 2009. Ref: IZB 2022, DB AG.

CONCLUSIONS

The presented approach has the following key characteristics:

- Fuzzy multi-label classification
- Interpretability and traceability
- No extensive dataset required, but thus diagnoses are only qualitative
- Balance between a component-based and functional view
- Focus on scenario-based evaluation, close to intended application

For railway point machines, combining measured data and various influencing factors is key to achieving high diagnostic performance.

References

This work has been published under Reetz S, Neumann T, Schrijver G, van den Berg A, Buursma D. Expert system based fault diagnosis for railway point machines. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. 2024; 238 (2): 214-224.

Funding

This project has received funding from the Shift2Rail Joint Undertaking (JU) under grant agreement No 881574. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Shift2Rail JU members other than the Union. This publication reflects only the author's view. The Shift2Rail Joint Undertaking is not responsible for any use that may be made of the information it contains.

Strukton Rail

DLR