WCRR BIRMINGHAM UK | 2022

Towards the automation of anomaly detection and integrated fault identification for railway switches in a real operational environment

Daniela NAREZO GUZMAN[1], Judith HEUSEL[1], Norman WEIK[1], Susanne REETZ[1], Douwe BUURSMA[2], Arnout VAN DEN BERG[2], Gerrit SCHRIJVER[2], Thorsten NEUMANN[1], Serge VAN DEN BROEK[2], Jörn C. GROOS[1]

[1] German Aerospace Center (DLR), Institute of Transportation Systems, Berlin/Braunschweig, Germany
[2] Strukton Rail, 3600BA Maarssen (Utrecht), Netherlands
Corresponding Author: Daniela Narezo Guzman (Daniela.NarezoGuzman@dlr.de)

Abstract

We report on the latest developments regarding automatic anomaly detection and fault diagnosis for railway switches achieved in the context of the EU Shift2Rail project In2Smart2. This paper presents improvements regarding the interpretability of the output of a data-driven anomaly detection routine, thereby increasing its usability for maintenance engineers. Warning messages are generated by a new module in case certain critical engine failure types are detected or additional information on detected anomalies is available. Further developments in the diagnostic model brought it to a state in which it is ready for a first plausibility check under supervision of maintenance experts. Taken together, these developments are an important step towards the integration of anomaly detection and diagnosis into a comprehensive condition monitoring system. The advances have recently been or are projected to be implemented in an upgrade of the workflow running in near real-time at the maintenance control center of Strukton Rail, in the Netherlands. This paper gives an overview of the workflow which includes the new warning module, the extended anomaly detection pipeline, and how the diagnostic model is to be conceptually embedded, i.e. its interactions with the other modules. Moreover, methodological improvements, e.g., in data preprocessing for feature extraction and warning generation, are also described.

Keywords: asset management, railway switches, condition monitoring, condition-based maintenance.

1. Introduction

Railway switches are critical elements of the railway infrastructure since they are vital for train routing and operation. Switch failures are responsible for a large share of train delays, resulting in a negative impact on service reliability and reputation. Besides the service aspect, degraded switch moving parts can also pose a safety hazard if not maintained on time. Frequent inspection, maintenance and renewal of components, though costly, are an essential safety requirement. It is thus crucial to develop algorithms that enable condition monitoring for optimizing maintenance, increasing availability and supporting management [1, 2]. In this context, identifying the nature of a failure through fault detection and diagnosis is highly relevant for optimizing maintenance orders.

In this paper we report on further developments of the switch anomaly detection (AD) and fault diagnosis workflow in the context of EU Shift2Rail project In2Smart2. The workflow is implemented at the control center of maintenance provider Strukton Rail (SR). The workflow is connected to POSS, a condition monitoring system developed by SR. POSS provides asset condition monitoring based on the engine current measured at the point machine during blade repositioning. We refer to these measurements as current curves (CC). POSS also stores air temperature at the relay house for every CC, which is important as switch behavior (and thus CC) has a systematic temperature dependence. In the past, POSS has helped to identify degrading and failing switches. However, it heavily relies on manual work for rising and validating alarms, tying up significant manpower resources (see [3] for details on current practices at SR for switch condition monitoring and maintenance). The goal of this research is to enable condition-based maintenance (CBM) by combining domain knowledge with methods for feature engineering, AD and fault diagnosis in an automated workflow. For a first report about the integration of the AD model [4] and the diagnosis model (DM) [5] into real operation environment see [3]. Since then, research has further addressed two main challenges: 1. The automated generation of reliable alarms in real-time to enable CBM, 2. support maintenance engineers (ME) in identifying technical defects of faulty switches for optimizing maintenance. This paper describes how recent advances and additions to the workflow address these challenges – adopting a holistic perspective. Section 2.1 gives an overview of the workflow. Sections 2.2 to 2.5 explain recent additions to the workflow. Section 2.6 summarizes latest developments of the DM. Section 3 presents examples of the workflow output. Section 4 summarizes recent and future efforts.

2. Methodology

In order to provide ME with insightful information regarding the origin of faulty switch behaviour, two main additions to the workflow were implemented. Addition 1: consists of five parallel applications of the AD model, improving previous AD process. That is, in each application the model takes as input a different set of features; each set is derived from one of the four CC segments – in addition to the features derived from the CC, as a whole. Curve segmentation is based on expert knowledge; the advantage of localizing anomalies segment-wise, is that it helps ME to narrow down the compromised functionality of the switch. The AD model focuses on detecting deviations from normal switch behaviour (where "normal" denotes the opposite of abnormal and not the repositioning direction normal), while considering the systematic influence of temperature. It does not specify which input features contribute the most to the detection of an anomaly. Even if available, this information is not useful for the ME since most AD input features are statistical quantities with no physical interpretation (there are two exceptions; *length* is the total duration of the blades repositioning process, and *area under the CC* is proportional to the power consumed by the motor during this process). This motivates the development of Addition 2: the warning module (WM). It generates warning messages based on two new sets of features: the first set is designed for identifying specific faults deemed very critical by ME; the second set relates to switch operation information (not extracted from CC). The WM bridges the gap between detecting anomalies and providing details on specific faults and operation conditions to the ME by means of interpretable messages. Some WM input features shall serve as input for the DM in the long term.

2.1 Overview of Workflow

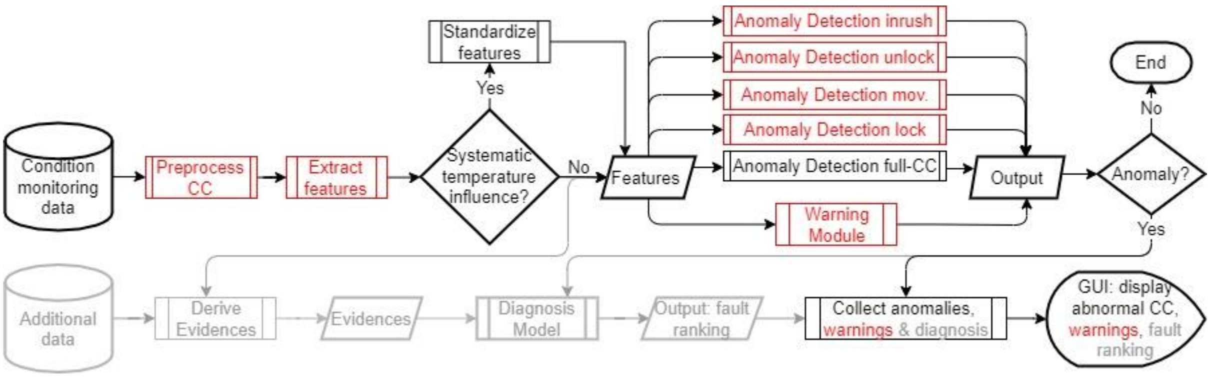

Fig. 1 Schematic representation of the workflow. Parts coloured in grey, black and red correspond to future, previously realized and most recent implementations/modifications, respectively.

In its current version, the automated workflow (see Fig. 1 – in black and in red) collects switch condition monitoring data every hour. It then deploys an algorithm for pre-processing the CC described in Section 2.2. In the next step, features are derived from the CC. Features that have a systematic temperature dependence are standardized (or normalized), see [4]. Features are stored together with information related to the operation of the switch. The features belong to six different sets, each set is input for a subsequent process in the workflow, see Section 2.3. The output of these processes, i.e. the AD pipelines (Section 2.4) and the WM (Section 2.5), are stored. Anomalies and warning messages are pushed into the GUI every hour, where ME have access to other tools and data sets. The DM (Section 2.6) is planned to be added to the workflow. As seen in Fig. 1 (in grey) additional data (e.g. power supply) and interpretable features are going to be processed into evidences i.e. valid input for the DM. It is only going to be deployed for detected anomalous CC and its output pushed to the GUI.

2.2 Pre-processing of Current Curves

CC are sampled at 50 Hz. This sampling rate is not high enough to capture the inrush peak or global maximum in detail. In fact, if the latter is located between two sample points, the feature maximum (see Table 1) is underestimated. In addition, the start and end of a CC are defined by the position of the global maximum and by the last CC sample that is higher than 0.5 A, respectively. Thus, the accuracy of the CC start and the CC considered for further processing depend on the sampling choice. This, in turn, affects the accuracy of all CC-derived features (especially length). The analysis of a test data set containing CC sampled at 1000 Hz showed that the accuracy of both position and magnitude of the global maximum for CC samples at 50 Hz can be improved by an additional pre-processing step. It consists of upsampling the CC to 1000 Hz and then quadratically interpolating the inrush segment to capture position and value of the inrush peak (the unlock segment is also quadratically interpolated, whereas movement and lock segments are linearly interpolated). Next, the CC is down sampled back to 50 Hz such that the sample of the estimated maximum is now a sample point, which defines time = 0 seconds of the CC (as in Fig. 2).

2.3 Feature Extraction

Table 1 lists all features computed in the workflow, their usage (in reference to Fig. 1) and the set they belong to. Most features are derived from the CC, however operational information is included in the table for completeness. New features were designed for the main additions to the workflow: parallel AD model deployment on CC segments and the WM. Their design involved domain knowledge as CC segmentation relies on experts' estimation on the duration of each phase of the blades repositioning. Input features to the WM are based on information provided by the ME regarding traces left on CC by specific faults. CC segments correspond to phases of the blades repositioning: inrush, unlock, movement and lock (see Fig. 2). Features for the full CC and for the segments belong to different feature sets. Given that the length of inrush, unlock and lock segments is pre-defined, length is only a relevant feature for the movement segment and the full CC. The AD model requires input features to be standardized (see Fig. 1); this step removes as best as possible their systematic temperature dependence [4]. Features in sets WM-CC and WM-operational are further processed by the WM to create warning messages, see Section 2.5. In principle, some of the features in WM-CC could be input for the AD model after proper normalization (e.g. the area under the hump could be normalized by the area under the baseline). Further analysis on the inclusion of such features into the AD pipelines is needed.

Feature set	Input for	Features
AD-inrush	AD inrush pipeline	Standardized values of: mean, standard deviation, area under CC, maximum, minimum, median, skewness & kurtosis – derived from inrush segment
AD-unlock	AD unlock pipeline	Same as in AD-inrush set - derived from unlock segment
AD-lock	AD lock pipeline	Same as in AD-inrush set - derived from lock segment
AD-movement	AD movement pipeline	Same as in AD-inrush set + standardized length - derived from movement segment
AD-full-CC	AD full-CC pipeline	Same as in AD-movement, derived from full-CC
WM-CC	Warning Module	<u>Flat spot</u>: start/end position, current level, CC length, normal CC length. <u>Hump</u>: start/end position, prominence, area under humped part of CC, area under baseline along humped part of CC. <u>CC too long/short:</u> CC length, normal CC length, mean current in movement segment from previous normal CC
WM-operational	Warning Module	End-to-start resting time (time between end of a CC and start of the next one), temperature bin occupancy in training set of AD model

Table 1: Overview of features. Features in WM-CC are differentiated based on warning message (underlined).

2.4 Anomaly Detection in Current Curve Segments

Once the model is trained with features derived from historical normal CC (as in [4]), it recursively assesses new CC and assigns an anomaly score (based on the Principal Components) to each. The user-defined anomaly score threshold is used to filter out the most anomalous CC. The workflow implements the AD model in five parallel pipelines, each taking as input the corresponding feature set (see Table 1). Anomalies detected for a CC by more than one pipeline are aggregated. The advantage of parallel pipelines is knowing which segments show anomalous behaviour. This can support the ME in narrowing down the compromised functionalities, since the segments are associated to specific parts of the switch and thus to possible primary faults.

2.5 Warning Module

Other than the AD model, the WM does not rely on training data but requires pre-defined parameters. It generates warnings based on sets WM-CC and WM-operation (see Table 1) in near real-time. These sets respectively relate to certain fault types and to operations that are relevant for ME. They interpret warning messages by combining additional information and domain knowledge. A subset of all anomalous CC detected by the AD pipelines are accompanied by a warning message. This section summarizes the domain knowledge behind each warning message (in bold letters), some examples thereof are given in Section 3.

Humps**:** typically appear in the movement or in the lock segments. They are characterized by an abnormal increase and a posterior decrease in the current. Humps are caused by a temporary additional/abnormal mechanical resistance (e.g. caused by a burr on a sliding chair) experienced by the motor as it powers the switch blades to move. ***Flat spots*****:** are observed in CC when the clutch slips, i.e. when it reaches the safety friction level that limits the transmitted torque. This is a mechanism to safeguard the motor from overloading as it prevents it from stalling, which would blow the fuse or cause it to burn. When a clutch is slipping, no mechanical parts of the switch are in motion since they experience a mechanical resistance equal or higher than the friction level of the clutch. In the best case, the stick-slip phenomenon causes intermittent motion of the mechanical parts and the blades can reach their end position. Flat spots are identified as regions located on top of a hump, where the current remains at the same level for a certain time. In principle, the length of a CC with one or more flat spots is longer than normal by the total duration of all flat spots. A degrading clutch is characterized by a systematic decrease or increase over time in the clutch slipping level. Identifying degrading clutches is highly relevant; for this purpose, the WM keeps record of the level of flat spots. Note that CC that do not reach zero at the end of the measurement do not trigger a flat spot warning, since the condition of having a hump on top is not fulfilled. ***CC is too short/long*****:** if the CC length is at least 20% shorter/longer than normal. Normal values are derived from AD model training set statistics which considers the CC temperature dependence. A too short CC is e.g. observed when the blades did not reach end position in the previous repositioning. If a CC is too long due to a slipping clutch, the mean current in the movement segment with respect to the last normal CC is provided. This quantity is indicative of the margin between recent normal behavior and the slipping level of the clutch. The longer a clutch slips, the higher are the chances of it getting damaged. Thus, identifying these CC is crucial. ***Resting time too short:*** when a switch is being maintained, it is often operated repetitively within a short time, resulting in a resting time shorter than 2 minutes. This message helps to identify anomalous CC during maintenance windows, which are of no concern to ME. ***Low temperature bin occupancy:*** this message is only pushed to the GUI for anomalous CC. The AD model is not well trained for rare temperatures (measured by temperature bin occupancy); anomalies for these temperatures should be managed carefully, as they might be false positives.

2.6 Diagnosis Model

In the process of analyzing faulty switches, ME consider the AD output, WM messages, CC historical record and additional information (e.g. prior and upcoming maintenance actions), to judge the health status of a switch and decide on the initialization of maintenance actions. This manual process is time and labor intensive, and requires highly qualified personnel. The DM aims to support the decision-making process by combining CC features, CC history, domain knowledge and additional information to estimate probabilities of all possible fault types and subsequently output the most likely sources of error. The DM is currently developed separately from the implemented workflow, but necessary interfaces are already partially put into place; especially the work on the new features lays the groundwork for future integration. Once embedded in the workflow (Fig. 1, in grey), the DM is going to be triggered by the AD, and its output visualized in the GUI. With regard to methodology, given the lack of labeled data required for purely supervised machine-learning approaches, the DM is mainly based on expert knowledge, feature engineering and exemplary data samples. The DM is a Bayesian network (BN), which is a common approach for diagnostic modelling in various disciplines [6]. The DM probabilistically models the point machine and moving mechanical parts of the switch and their relation to CC features and probabilistic influences by e.g. weather (see "Additional data" in Fig. 1). Once embedded in the workflow, the DM would be fed automatically with evidences whenever they become available, improving the estimation of the probability of related fault types. The summarized BN output is directly understandable for ME and the reasoning process behind the output can transparently be reconstructed by looking at the complete model. Before integrating the DM into the workflow, further development is necessary. For this purpose, evidences are manually entered via an interactive interface, such that the effect of each piece of information is traceable. This allows for calibration and validation in cooperation with the practitioners at SR. A prototypical stand-alone DM developed with support and checked by ME for plausibility is planned to be finished by the end of In2Smart2.

3. Results

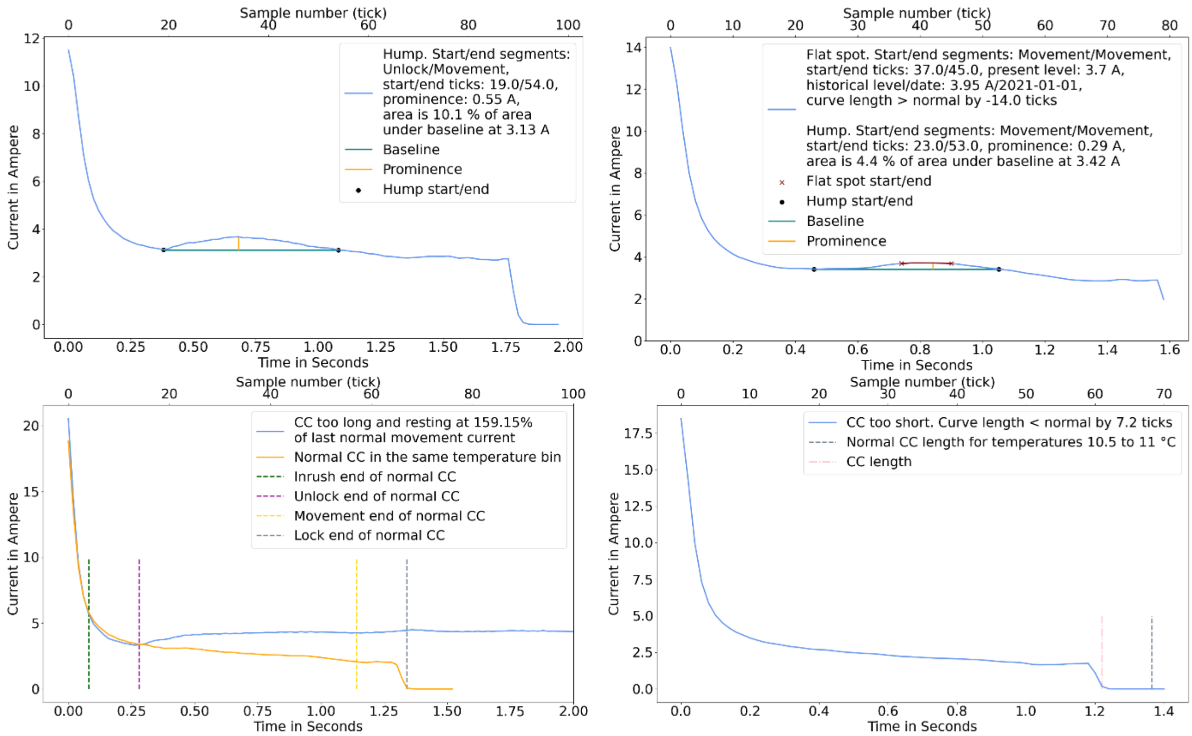

Fig. 2 Examples of CC with warning messages (in labels) related to a hump (top left), flat spot (top right), too long duration and resting at a certain level (bottom left), too short duration (bottom right).

This section presents first results of the WM. Even though the parallel AD pipelines are a recent addition to the workflow, no results are presented here due to limited space; see [4] for results of the AD model applied to full CC. Warning messages associated to faults contain features that characterize the faulty CC and that support ME in making decisions on maintenance. Fig. 2 – top left: CC with a hump; also displayed are its prominence, start/end, and baseline (approximation to the CC without hump). The message includes the additional area caused by the hump (i.e. area between CC and baseline) as a percentage of the area under the baseline. Top right: CC with a flat spot; the message contains information on its start/end, its level (present value and historical maximum) and number of ticks larger than normal. Given the condition that a flat spot must overlap with a hump, a respective warning is also provided. Bottom left: depicts two CC from the same switch. One did not reach zero Ampere before POSS stopped measuring it and is accompanied by the warning "too long and resting at x% of last normal movement current". The other one is the last normal CC prior to the abnormally long CC. The segments of the normal CC are indicated by vertical lines; the mean current value in its movement segment is used as reference in the warning message displayed. Bottom right: shows a CC that is too short with respect to normal for the particular temperature during the measurement (lengths displayed by vertical lines).

4. Conclusion and Outlook

The most recent additions to the workflow running in a real operation environment bring CBM at the SR control centre one step closer. The output of the AD pipelines and the WM support ME in maintenance planning through: 1. Interpretable warning messages related to specific fault types, 2. Warning messages related to operational information that help assess the relevance of certain anomalies, and 3. Localization of anomalies at the CC segment level, which is linked to specific switch parts. However, this additional information does not provide a full fault diagnosis. Our efforts to refine and finalize the DM continue, as the model aims to paint a comprehensive picture of possible fault causes and provide a ranking of the most likely sources of error behind anomalous behaviour in the workflow. Future work includes generalizing the pre-processing step to the lock segment in order to improve the accuracy of CC end position. Furthermore, the AD model needs to be validated for a larger number of switches than considered in [4]; for this, ME feedback and historical fault records will be considered. Also, the integrity of the AD model and the WM needs to be verified (e.g. check whether all CC with too long/short messages are detected as anomalies) in order to finetune parameters associated to the sensitivity of the models. Also, extending the WM to include further fault types is planned. Finally, to integrate the DM into the workflow, further research into translating features and other information into evidences is required. In addition, a plausibility check of the DM outcome with help of ME expertise is necessary.

Acknowledgment

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. The authors thank Henk Samson for his support in the project.

References

1. Swier, J. (2021) Ways to Improve the Performance (And Cost) of Switches. In: Galeazzi R. et al. (eds) Intelligent Quality Assessment of Railway Switches and Crossings. Springer International Publishing, pp 261–274
2. Hamadache, M. et al. (2019) On the Fault Detection and Diagnosis of Railway Switch and Crossing Systems: An Overview. Applied Sciences 9:5129. https://doi.org/10.3390/app9235129
3. Narezo Guzman, D. et al. Detecting anomalous behavior of railway switches under real operation conditions: workflow and implementation. In: TRA 2020 conference proceedings. Transport Research Arena 2020
4. Narezo Guzman, D. et al. (2020) Anomaly Detection and Forecasting Methods Applied to Point Machine Monitoring Data for Prevention of Railway Switch Failures. In: Ball, A. et al. (eds) Advances in Asset Management and Condition Monitoring, vol 166. Springer International Publishing, Cham, pp 307–318
5. Reetz, S. et al. (2021) Kombination von Messdaten und wissensbasierter Modellierung zur Fehlerdiagnose bei Weichen: Connecting measurement data and knowledge-based engineering for heavy rail switch fault diagnosis. SIGNAL + DRAHT 113:27–35
6. Cai, B. et al. (2017) Bayesian Networks in Fault Diagnosis. IEEE Trans Ind Inf 13:2227–2240