Detection of Anomaly in Train Speed for Intelligent Railway Systems

Seungmin Kang*, Sravana Sristi*, Jabir Karachiwala*, Yih-Chun Hu*[†]
*Advanced Digital Sciences Center, Singapore
[†]University of Illinois Urbana-Champaign, USA
Email: seungmin.k@adsc.com.sg, sravana.s@adsc.com.sg, jabir.k@adsc.com.sg, yihchun@illinois.edu

Abstract*—Anomaly detection has been applied to diverse critical applications or systems since anomalous behaviors could lead to fatal situations during the operation. In intelligent transportation systems, anomaly detection also plays an important role by allowing the system administrator to assess the imminent emergence of any incidents. In this paper, we address real-time anomaly detection that has not yet been thoroughly explored in railway system. We propose an online anomaly detection scheme in train speed form railway systems using machine learning approaches. We adopt the Bayesian statistical learning model to represent normal behavior of train speed changes and detect the anomaly based on the occurrence probability of each speed change observation. While the Bayesian statistical learning model can detect sudden speed changes, it may not be able to detect malicious behavior of an intelligent attacker who gradually reduces or increases the train speed to cause the collision between two subsequent trains. We thus propose a linear regression model that takes into account time duration and travel distance from the departure station to detect anomaly. We evaluate the proposed scheme through comprehensive simulations. The results show that the proposed scheme efficiently detects anomalous speed change by accurate predictions from the learning phase and it outperforms a baseline approach with an improvement in sensitivity by up to 22%.

***Index Terms*—Anomaly detection, statistical model, railway systems, intelligent transport systems.**

I. Introduction

Anomaly detection refers to identifying patterns in data that do not conform to expected behavior and it is required in diverse application domains including intrusion detection for cyber-security, fault detection in safety critical systems, etc. Since anomalous behaviors could lead to fatal situations in such systems, anomaly detection becomes an important area to address failure diagnosis or potential security issues. For example, anomalous readings from a space craft sensor could signify a fault in some components of the space craft [1] and an anomalous trace pattern in a computer network could mean that a compromised computer sends out sensitive data to an unauthorized destination [2].

In intelligent transportation systems (ITS), anomaly detection also plays very important role by allowing the system administrator to assess the imminent emergence of any incidents, i.e., detect deviations from normal situations. ITS has been successfully implemented in communication-based train control (CBTC) systems [3], which are one of key components to ensure a safe and efficient operation by using various on-board sensors. Several relevant approaches have been proposed for the vehicular networks [4], [5], however, very few solutions so far have been proposed for railway systems, leaving the problem of anomaly detection in railway systems remain challenging. In [6], the authors developed a detection system to predict potential failures by monitoring information using sensors positioned on the main train components. However, this work did not address real-time detection of anomalies on the running trains. In [7], the authors studied a data-driven predictive maintenance system that issues an alarm whenever an automatic door is predicted to suffer a failure. The proposed system can work online through the evolving models using sliding windows. Our work differs from the above works as we consider an online detection model for train speed.

Anomaly in train speed can be caused by different factors. In normal working conditions, an incident of a certain component for instance, failure of speed sensor of the train, can create abnormal acceleration or deceleration. In addition to the failure scenarios, there are cases where an attacker may influence measured speed to manipulate the system into a hazardous situation. For example, after gaining unauthorized access to the trainborne network, the attacker may stop/overspeed the train in the middle of the track to cause the collision. Regardless of root causes of abnormal behaviors, in this paper, we address the problem of anomaly detection in train speed of railway systems, allowing system administrators to avoid fatal collision. To achieve this objective, we propose an anomaly detection scheme in railway systems. We develop a Bayesian statistical model that represents train behavior in speed changes in normal working conditions. Assuming that speed changes of trains in normal working conditions follow a probability distribution, we estimate the parameter of the probability distribution by using the Bayesian statistical learning model. Given a new observation of speed change captured by sensors during operation, computing the occurrence probability of the observation with the estimated probability distribution allows us to detect the anomaly, i.e., smaller the value of the occurrence probability, more anomalous the observation is.

While the Bayesian statistical learning model can detect sudden speed changes of trains, it may not be able to detect malicious behavior of an intelligent attacker who gradually reduces or increases the speed of trains, creating the collision between two subsequent trains. We thus propose a linear regression model that takes into account time duration and travel distance from the departure station. For a given time duration after leaving the departure station, if the travel distance is not in the safe range, it can be considered as anomaly. We use OpenRails platform to simulate the operation of trains and generate data sets to evaluate the performance of our proposed

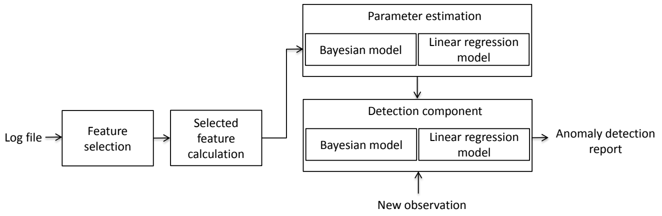

Fig. 1: Overview of anomaly detection system (ADS).

models. We evaluate the performance of the proposed scheme through comprehensive simulations.

The rest of the paper is organized as follows. We present the security threats and proposed model in Section II. We describe feature extraction for anomaly detection in Section III. We introduce our anomaly detection scheme in Section IV and its integration in railway systems in Section V. We present experimental results in Section VI. We discuss the related work in Section VII before concluding the paper in Section VIII.

II. Threat Model and Proposed Approach

A. Threat Model

Anomalous behavior needs to be examined to ensure safe railway operation of CBTC systems. Particularly, dead reckoning becomes the base of automatic train protection (ATP) system by measuring the speed and position reliably. The train determines its speed based on data from on-board sensors such as tachometer and Doppler. However, railway vehicle sensors suffer from insufficient measurement accuracy for several reasons. For example, the Doppler radar is prone to adverse weather conditions while wheel speed sensors are not sufficiently robust against wheel slip and wheel wear. In addition to such possible malfunction of on-board sensors, we also consider potential security threats as another cause of the anomalous behavior. This is because, if an attacker gains unauthorized access to the trainborne network and takes actions on control logic maliciously, i.e., manipulate of multiple measured speeds, main components of onboard CBTC subsystem become targets of attacks. The specific scenarios of potential attacks which cause the collision are: the attacker may stop the train through unauthorized active braking or overspeed the train in the middle of the track and the following/leading train is not aware of a sudden stop/overspeed of malicious leading/following train.

B. Proposed Approach

The proposed anomaly detection system (ADS) analyzes the deviations of parameters of real feature values from the estimated statistical models of those values. As shown in Fig. 1, there are mainly two stages for anomaly detection. The first step builds the Bayesian statistical model and linear regression model using historical log file in normal working conditions. This step is realized on the formerly selected and calculated features and results in the estimated parameters of the Bayesian and linear regression model parameters. Given a new observation, the second step is to assess the difference between the actual value and the calculated Bayesian/linear regression representation of the value. Based on this difference, the anomaly can be detected.

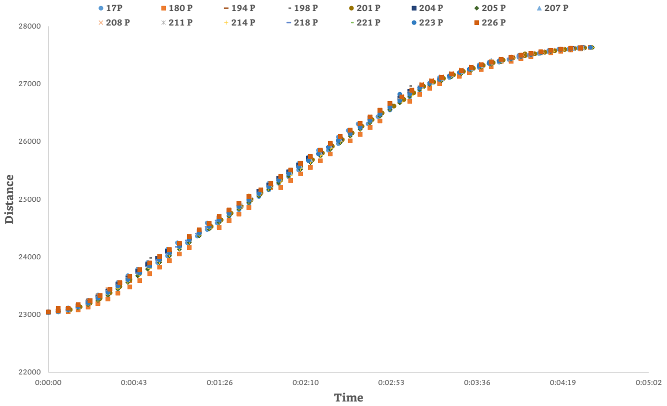

Fig. 2: Correlation between time instant and travel distance for 15 trains between Shenzen and Taoyuan stations.

For the deployment of ADS, we consider two possible ways: the train sends data, i.e., speed and travel distance, through the radio-based communication link to the wayside system, which is further connected with the central automatic train supervision (ATS) system and thus ADS can be deployed at ATS. Furthermore, when we consider a fully-automated CBTC system, a zone controller, which sends individual commands to each train under its control, uses train information to determine limit of movement authority (LMA) for trains. Therefore, ADS can also be deployed at zone controller to aid LMA determination. It is worth mentioning that when ADS is deployed at those two systems, respectively, we can address the redundancy issue by providing an additional layer of safety.

III. Feature extraction

In order to detect anomalies on train movements, we need to extract some representative features from the observed log data. We obtain log data from the OpenRails platform[1], which is an open-source platform to simulate the train motion according to a given timetable schedule. Based on the log data, we observe the fact that the speed profiles of multiple trains between stations tend to be very similar, i.e., train movement consists of three operation regimes: acceleration, run (constant speed), and deceleration. Based on this observation, we use speed differences to capture the operation state of a train. Specifically, given the extracted speeds of each train T_i $(1 \leqslant i \leqslant N)$ between two measurement time instants, the speed difference of T_i is obtained as follows:

$$|s_i(t+u) - s_i(t)|, \; i = 1, .., N \tag{1}$$

where $s_i(t)$ is the speed of T_i at time instant t and u is a predefined time interval between two measurements. We finally have a training data set, $X = \{x_{i,j}\}$, where i is the index of the train and j is the index of the measurement. This training set consists of the normal train speed differences during the interested time window and it is used for learning step.

Furthermore, since trains operate according to a predetermined schedule from the departure to the arrival station, we observe the fact that there is a correlation between time duration and travel distance. As shown in Fig. 2, we observe

[1] OpenRails: http://www.openrails.org

that given a particular time instant, the travel distances of all the trains who pass the same section of track are very similar. Hence, we also consider a pair of the time and its corresponding travel distance as another interested feature and thus have a training set, $Y = \{t_{i,j}, d_{i,j}\}$, where i is the index of the train and j is the index of measurement, representing the normal behavior of travel distance during the interested time window and will also be used for learning step.

IV. Statistical anomaly detection

We now describe the proposed anomaly detection scheme by introducing a statistical approach to infer the suspected anomaly. When there is a new observation, the probability that the estimated statistical model should generate this observation is calculated. The new observation is considered anomalous if it is improbable to occur in that model, i.e., its occurence probability is smaller than a pre-defined anomaly threshold.

A. Anomaly definition

Assuming that samples from the normal situation are generated by a known probability density $P(x|\theta)$ for a set of parameters θ, the smaller the probability of generating a new observation from the distribution, the more anomalous is it. To define how unusual a new observation z is, we represent the probability of generating a more common sample than z from the distribution as:

$$A(z|\theta) = \int_{x\in\Omega} P(x|\theta)dx \tag{2}$$

where $\Omega = \{x : P(x|\theta) > P(z|\theta)\}$.

There are several desirable properties of $A(z|\theta)$. First, $A(z|\theta)$ increases when z is more anomalous. Second, it is comparable to the anomaly of other distributions or sets of parameters. Third, it is directly connected to the rate of false alarms. If we set a threshold on $A(z|\theta)$ of $1 - \sigma$ over which an observation is determined anomalous, the probability that a normal observation is wrongly detected as anomalous is then simply σ. This corresponds to the probability of the tails beyond z and it can be expressed as:

$$\bar{A}(z|\theta) = 1 - A(z|\theta) = \int_{x\in\Omega} P(x|\theta)dx \tag{3}$$

$$\text{where } \Omega = \{x : P(x|\theta) \leq P(z|\theta)\}.$$

Therefore, we determine anomaly for z using anomaly threshold, denoted as σ:

$$\begin{cases} \text{Normality} & \text{if } \bar{A}(z|\theta) \geq \sigma \\ \text{Anomaly} & \text{if } \bar{A}(z|\theta) < \sigma \end{cases} \tag{4}$$

As we mentioned, one of challenging issues of anomaly detection is a false alarm. We will describe how we tackle the problem of false alarm to improve the accuracy of the proposed ADS in the following subsection.

B. Bayesian learning

Given the data samples in normal working conditions, we need to estimate the parameter(s) θ for the statistical model. In classical statistics, the approach would be to find the maximum likelihood estimate of the parameters of the statistical model [8], and use those parameters when calculating the probability of a new observation. However, the classical approach does not consider the uncertainty in the estimate, i.e., uncertainty regarding the true values of the parameters. This is especially true when there are few data samples and this is not an ideal property of an anomaly detector since it will give a lot of false alarms by focusing too much on the peculiarities of the data.

To address this issue, we use the Bayesian approach, which considers the parameters of the distribution as stochastic variables. The final answer is then obtained as an integral over all possible parameter values. Thus, the Bayesian approach will dampen the effect of random occurrences and instead single out the significant cases. Although early on a Bayesian approach will accept more samples as normal, the parameter estimation will become more accurate as more training data is collected, making anomaly detection more precise. In this paper, we assume that we know the parametric form of the distribution and let θ denote the unknown parameters of the distribution. Hence, all the knowledge of the distribution are the parametric form and the set of training samples X.

Under this assumption, we need to find the posterior distribution over the parameters θ conditioned on X and this can be expressed using Bayes' rule:

$$P(\theta|X) = \frac{P(X,\theta)}{P(X)} = \frac{P(X|\theta)P(\theta)}{\int_\theta P(X|\theta)P(\theta)d\theta} \tag{5}$$

Since we assume that X consists of independent sample, $x_{i,j}$, it follows:

$$P(X|\theta) = \Pi_{i=1}^{n}\Pi_{j=1}^{\frac{\Delta t}{u}} P(x_{i,j}|\theta) \tag{6}$$

$$P(\theta|X) \propto P(X|\theta)P(\theta) = \Pi_{i=1}^{n}\Pi_{j=1}^{\frac{\Delta t}{u}} P(x_{i,j}|\theta)P(\theta) \tag{7}$$

where $P(\theta)$ is the prior distribution over the parameters.

C. Bayesian anomaly model

From $A(z|\theta)$ and $P(\theta|X)$, we infer the expected anomaly of a new observation z by integrating over all possible parameter values of θ:

$$A(z|X) = \int_\theta A(z|\theta)P(\theta|X)d\theta \tag{8}$$

where $A(z|X)$ is defined as the Bayesian Anomaly in this paper. To express the posterior $P(\theta|X)$ in terms of the prior $P(\theta)$, we substitute Eq. (5) to Eq. (8) and obtain:

$$A(z|X) = \frac{\int_\theta A(z|\theta)P(X|\theta)P(\theta)d\theta}{\int_\theta P(X|\theta)P(\theta)d\theta} \tag{9}$$

The Bayesian Anomaly has many suitable properties to apply it in practice: First, the false alarm rate, which is a major problem for many anomaly detection algorithms when used in practice, can be controlled directly by adjusting the anomaly threshold, σ. Second, the Bayesian approach makes the system work when there are limited amounts of training data, as is often the case. Third, the training data used does not necessary to be absolutely clean since the method will itself test each sample and only learn those that are judged non-anomalous.

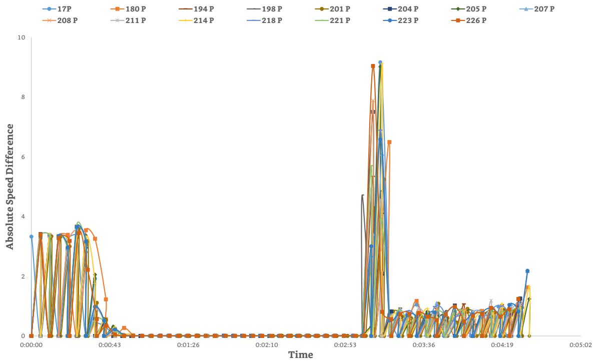

Fig. 3: Speed difference measurement for 15 trains from Shenzen station to Taoyuan station.

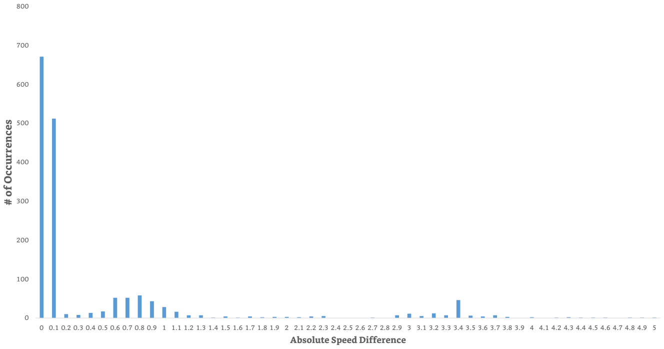

Fig. 4: Number of occurrence of speed difference for 15 trains from Shenzen station to Taoyuan station.

V. Anomaly Detection in Railway Systems

In this section, we build a statistical model from data in normal situations and describe how the proposed ADS detects anomalous behaviors in train speeds in railway systems.

A. Anomaly detection in speed change

To model a normal behavior, we observe the fact that train operation regime between two stations, i.e., acceleration, run, and braking, has a certain threshold. For example, all the speed differences are within 4 m/s as shown in Fig. 3. We also observe the number of occurrences of speed difference for multiple trains between two stations as shown in Fig. 4. Based on this, we assume that the speed difference follows an exponential distribution:

$$P(x|\lambda) = \lambda e^{-\lambda x} \tag{10}$$

where λ is the inverse of the mean value of speed differences. We assume that a set of training samples, $X = \{x_{i,j}\}, 1 \leqslant i \leqslant N, 1 \leqslant j \leqslant K$, consists of speed difference for each train between two measurement instants where N is the number of trains and K is the number of measurement instants between two stations. To calculate the mean value, we may take the simple mean from $X = \{x_{1,1}, ..., x_{1,K}, ..., x_{N,1}, ..., x_{N,K}\}$, as follows:

$$\lambda = \frac{NK}{\sum_{i=1}^{N} \sum_{j=1}^{K} x_{i,j}}. \tag{11}$$

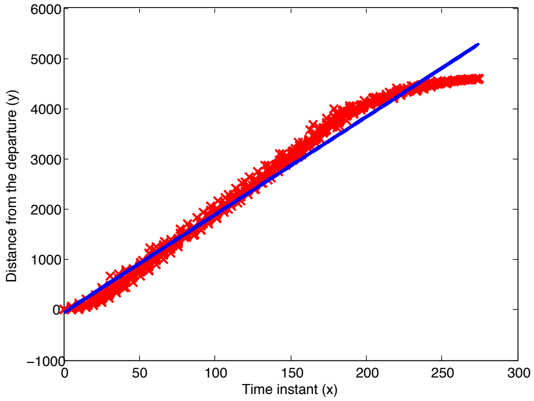

Fig. 5: Linear regression fit.

We can obtain Bayesian anomaly of our interested feature from Eq. (8) by substituting θ to λ in Eq. (7) and λ parameter is learned by the Bayesian model. It is now possible to find a train that has changed behavior recently by testing its speed differences against the average of those from a longer historical time period. For this, we observe speed differences that are generated during a certain time interval, and fit them to the probability distribution defined by the estimated parameter λ, which is learned from log files in normal working conditions. If the probability of the occurrence is too small, i.e., smaller than a certain threshold, it is considered as anomaly.

B. Anomaly detection in travel distance

Although anomalous speed differences, i.e., sudden large speed changes, are detected by our proposed ADS, it may not be able to detect more sophisticated attacking scenarios. For example, when an attacker gradually stops the train through unauthorized active braking in the middle of the track, if the speed change is still within the range of normal state, the anomaly detection based only on the speed difference could not address such scenario. Hence, we also consider the travel distance of a train with respect to the time duration since the departure instant. Without loss of generality, we assume that at the departure station, time instant is set to 0. Given a time instant, denoted as t, the travel distance, denoted as d, is then estimated by the linear regression technique as follows:

$$d = \alpha t + \beta \tag{12}$$

where α is a coefficient and β is the intercept, which are learned by executing the gradient descent on a given training dataset, which contains the samples, each being represented by (t, d) where t is the time instant and d is the distance from the departure station.

As shown in Fig. 5, the learned model based on linear regression model is fit to the training data well. Hence, anomalous events can be detected if a new observation deviates from the normal pattern: if the observed travel distance is higher/lower than the normal state, overspeed/stop event is detected. To evaluate the correctness of the learning model, we use a mean squared error (MSE) as follows:

$$err = \frac{1}{2m}[\sum_{i=1}^{m}(\hat{X}_i - X_i)^2] \tag{13}$$

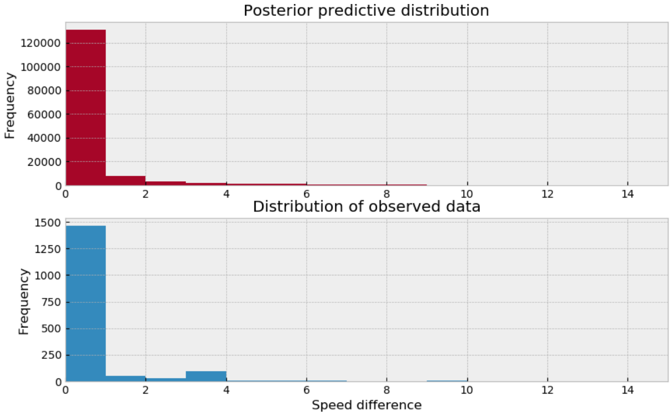

Fig. 6: Prediction by bayesian learning.

where m is the number of training samples, $\hat{X}_i$ is an estimated value, and X_i is an actual value corresponding to an input to the function which generated the predictions. Given a new observation, the anomaly is determined by comparing the deviation of the observation from the estimated value, using an anomaly threshold, $\sigma \in [0, 1]$:

$$\begin{cases} \text{Anomaly} & \text{if } \ e' \geq \sigma \\ \text{Normality} & \text{if } \ e' < \sigma \end{cases} \tag{14}$$

where the deviation is computed as follows:

$$e' = \frac{(\hat{X}_i - X_i)^2}{\hat{X}_i^{\,2}}. \tag{15}$$

VI. Experimental Results

A. Experimental Setting

We use OpenRails to generate the training data set in normal working conditions for our machine learning algorithms. The validation and test data sets are generated by randomly injecting attacking messages that change the train speed accordingly. We analyze the performance of ADS through the sensitivity of the models. The model sensitivity represents the ability of a test to correctly identify those with the anomaly and it is computed as $TP/(TP + FN)$ where TP is true positive and FN is false negative.

B. Detection of Anomalous Speed Difference

In Fig. 6, we present the predictive distribution of the speed difference predicted by Bayesian learning. The results show that posterior distribution is well predicted by Bayesian learning since the estimated distribution is almost identical with the distribution of the observed data in the training data set. This demonstrates the effectiveness of the Bayesian learning model in fitting the data samples.

In Fig. 7, we present the sensitivity of the Bayesian learning model compared to the Maximum Likelihood Estimation (MLE). The results show that the Bayesian learning model outperforms the MLE approach. With the entire testing data set, the Bayesian learning model achieves up to 78% of sensitivity compared to 48% generated by the MLE approach. Even though the Bayesian learning model requires complex computation to estimate the parameter of the probability model, this is only one-time cost and can be run offline given the training data set of the model.

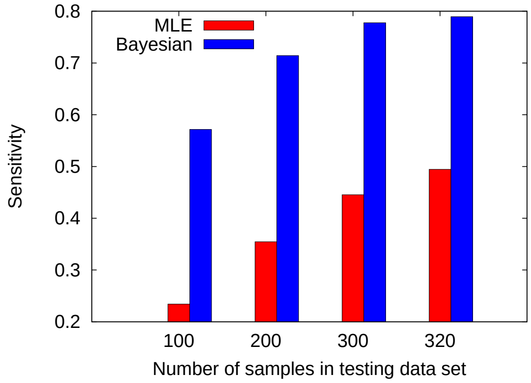

Fig. 7: Sensitivity of Bayesian learning model and LME.

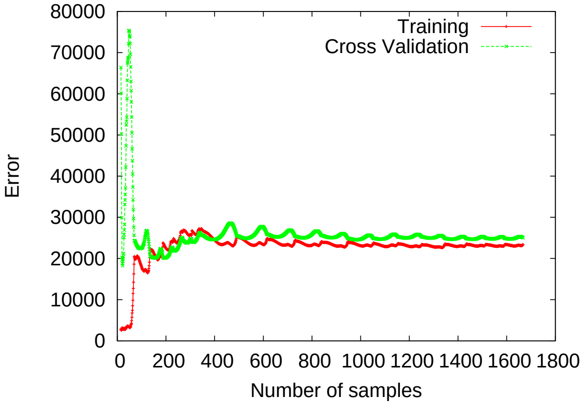

Fig. 8: Learning curve of training set and cross validation set.

C. Detection of Anomalous Travel Distance

1) Error rate of learning model: In Fig. 8, we present the learning curve of training set and cross validation set with respect to the number of training samples to show the correctness of our learning method. The results show that when we train data only based on very small sample size, the error is small since the model is likely to be biased on those samples. When the sample size is increased, the model is adjusted for all training samples, and thus the error is increased. We also observe that after the moment when the number of samples is large enough, the training error is stable. We use the cross validation set to estimate the accuracy of learning method and the number of observation in the cross validation set is much larger than the number of training samples. The results show that when the number of training sample is small, a lot of samples in the cross validation set do not appear in the training set, and thus making the error dominated. However, when the number of training sample is increased, samples in the cross validation set can be covered by the training set. Therefore, the training error decreases and attains the training error when the number of samples is large enough.

2) Accuracy of learning model: In Fig. 9, we compare the sensitivity of ADS and that of baseline with respect to the number of samples. In the baseline approach, the anomaly threshold is fixed for all the observations, whereas ADS adjusts the anomaly threshold for each observation.

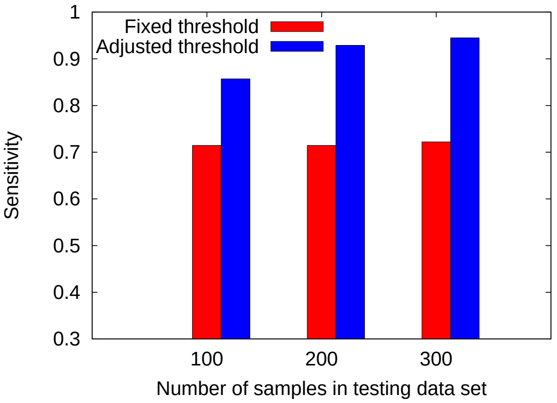

Fig. 9: Sensitivity of ADS and baseline approach.

Choosing an appropriate fixed threshold gives the fairness for the comparison, e.g., if the fixed threshold is too large, there are more false alarms. Hence, we choose the value as 0.0225, which is a bit larger than the adjusted threshold by ADS, 0.008982. The results show that the adjusted threshold of ADS has always better sensitivity than that of baseline approach. Furthermore, ADS has better sensitivity when the sample size is increased while the baseline approach has similar sensitivity regardless of the sample size. This is because ADS adjusts the threshold by comparing it with the computed error for each observation and changes the threshold as the error value if anomaly decision is not correct. Therefore, this helps to increase the sensitivity with more sample sizes. The adjusted threshold allows by up to 9% deviations as normal, whereas the fixed threshold allows by up to 15% deviations as normal. This shows the importance of adjusting threshold appropriately for accurate anomaly detection.

VII. Related Work

Typical use cases of anomaly detection include intrusion detection for cyber-security [9], fault detection in safety critical systems, and military surveillance for enemy activities. The study of [10] is based on clustering technique to identify anomalous measurements in sensor nodes. Although this technique does not require a priori knowledge of data distribution, it is difficult to determine an appropriate parameter of cluster width, which is used to compute the average inter-cluster distance. Similar to our work, a data-driven modeling approach is proposed in [11] to identify point anomalies by using the sequential information of sensor reading. The authors proposed several one-step ahead predictors, which are based on a sliding window of previous data, to predict the new output and compare it to the actual output. However, this work does not easily integrate several sensor streams to help detect anomalies. The authors in [12] apply attributed graphs by allowing for contextual data to be included within a graph structure. Specifically, they propose the algorithm to explore parts of the graph that were previously less emphasized by using additional metadata. However, the algorithm of [12] is difficult to use in real-time analytics since the estimation of their algorithm is not explored in detail.

Anomaly detection in railway systems has also gained importance to address the issue of maintenance and condition monitoring. The authors in [6] proposed an automatic detection system to identify anomalies for predicting potential failures. Similar to our work, they firstly characterize normal behavior by taking account data such as itinerary, weather conditions, etc. They then measure the compliance of new data according to extracted knowledge by classifying whether a system behavior is normal or anomalous. Although this work considers the temporal nature of sequential data which is collected by sensors, they do not meet the real-time issue, i.e., cannot detect the anomalies on the running trains. On the other hand, our work realizes the online anomaly detection system.

VIII. Conclusion

In this paper, we proposed an online anomaly detection scheme (ADS) for train speed in railway systems. We developed a Bayesian statistical model that represents speed changes of trains in normal working conditions. Assuming that this follows a probability distribution, we estimated the parameter of the probability distribution and successfully detected the anomaly by computing the occurrence probability of a new observation of speed change with the probability distribution. However, using only the proposed Bayesian model may not address more sophisticated attacking scenarios in which train speed is gradually reduced or increased, we proposed a linear regression model that takes into account travel distance to detect more advanced attacks. The validation was performed through simulations on data sets which are generated from the OpenRails platform. The results show that the proposed scheme efficiently detects anomalous speed change by accurate predictions from the learning phase of the proposed Bayesian model and it outperforms a baseline approach. The simulation results also demonstrate the accuracy of the proposed detection scheme using linear regression model by increasing the sensitivity by up to 22% compared with the baseline approach.

References

[1] R. Fujimaki, T. Yairi, and K. Machida, “An Approach to Spacecraft Anomaly Detection Problem Using Kernel Feature Space,” in *ACM SIGKDD 2005*, Chicago, USA, Aug. 2005, pp. 401–410.

[2] V. Kumar, “Parallel and distributed computing for cybersecurity,” *IEEE Distributed Systems Online*, vol. 6, no. 10, 2005.

[3] I. V. T. Society, “IEEE Standard for Communications-Based Train Control (CBTC) Performance and Functional Requirements,” *IEEE Std 1474.1-2004*, 2004.

[4] E. Kwon, S. Noh, M. Jeon, and D. Shim, “Scene Modeling-Based Anomaly Detection for Intelligent Transport System,” in *IEEE ISMS 2013*, Bangkok, Thailand, Jan. 2013, pp. 252–257.

[5] J. Raiyn and T. Toledo, “Real-Time Road Traffic Anomaly Detection,” *Journal of Transportation Technologies*, vol. 4, no. 3, pp. 256–266, 2014.

[6] J. Rabatel, S. Bringay, and P. Poncelet, “Anomaly detection in monitoring sensor data for preventive maintenance,” *Expert Systems with Applications*, vol. 38, no. 6, pp. 7003–7015, 2011.

[7] R. P. Ribeiro, P. Pereira, and J. Gama, “Sequential anomalies: a study in the Railway Industry,” *Machine Learning*, vol. 105, pp. 127–153, 2016.

[8] F. Sowell, “Maximum likelihood estimation of stationary univariate fractionally integrated time series models,” *Journal of Econometrics*, vol. 53, no. 1, pp. 165–188, 1992.

[9] S. Axelsson, “Intrusion detection systems: A survey and taxonomy,” Tech. Rep., 2000.

[10] S. Rajasegarar, C. Leckie, M. Palaniswami, and J. C. Bezdek, “Distributed Anomaly Detection in Wireless Sensor Networks,” in *IEEE ICCS 2006*, Singapore, Oct. 2006, pp. 1–5.

[11] D. J. Hill and B. S. Minsker, “Anomaly detection in streaming environmental sensor data: A data-driven modeling approach,” *Environmental Modelling and Software*, vol. 25, no. 9, pp. 1014–1022, 2010.

[12] B. A. Miller, N. Arcolano, and N. T. Bliss, “Efficient anomaly detection in dynamic, attributed graphs: Emerging phenomena and big data,” in *IEEE ISI 2013*, Seattle, USA, Jun. 2013, pp. 179–184.