Anomaly Sign Detection for Automatic Ticket Gates by Histogram Limitation Method

Ken Ueno[1], Shigeru Maya[1], and Kiyoku Endo[2]

[1] *System AI Lab., Corporate R&D Center, Toshiba Corporation, Kawasaki, Kanagawa 212-8582, Japan*
ken.ueno@toshiba.co.jp
shigeru1.maya@toshiba.co.jp

[2] *Toshiba Automation Systems Service Co., Ltd., Kawasaki, Kanagawa 210-8541, Japan*
endo.kiyoku@toshiba-tass.co.jp

ABSTRACT

It is crucial for automatic ticket gates (ATGs) on railways, also known as fare collection systems, to detect anomalies at an early stage, especially in the automatic separation module for multiple tickets. It is also required for efficient and low-cost monitoring without any additional sensors especially for old-type ATGs that need to be maintained frequently. However, the failure rate is basically very low, and monitoring data contain various kinds of normal status indicators depending on complicated mechatronics controls. In addition, it is hard to collect high quality learning data because ATGs are affected by various ticket conditions or timing when releasing tickets by users, which makes detecting anomaly signs difficult. For these reasons, conventional machine learning or deep learning methods are not suitable for anomaly detection for ATGs. In this paper, we propose a simple anomaly detection method with new anomaly sign index, called the histogram limitation method (HLM), for effective monitoring to realize preventive maintenance of ATGs based only on system log data. Despite being a quite simple and compact method, HLM provides anomaly sign scores that agree adequately with assessments by maintenance service engineers in our evaluation with real field ATGs in operation.

1. BACKGROUND

To realize smooth transportation for many people in urban cities with highly developed railway transportation networks, it is very important to maintain and monitor automatic ticket gate (ATG) operation. Nowadays, although ATGs have fewer mechanical parts as e-tickets become more popular, ATGs that use paper tickets must be maintained for at least 10 years into the future as part of the railway infrastructure. Especially for old-type ATGs, it is required for efficient, low-cost, and compact monitoring function that need to be maintained frequently without any additional sensors because it is necessary to pay additional costs for hardware as low as possible. This means it is hard to utilize rich data for diagnosis like anomaly detection for switchgear in substations (Yamaguchi, 2022) or manufacturing equipment (Maya, 2019).

There are few chances for service engineers (SEs) who maintain ATGs to obtain experience for high-level mechanical maintenance, and it is especially that younger SEs must develop their skills before the experts retire. To mitigate these problems, we propose anomaly detection for predictive maintenance based on Condition-Based Maintenance (CBM). Here, an anomaly sign refers to symptoms in the normal state before an anomaly occurs. Early detection can help prompt early action for preparing for part replacement or summoning SEs for mechanical maintenance to prevent ATGs from defects that would interrupt ATG service for several hours during repairs.

SEs adjust or repair the parts on-site when they are summoned by station attendants in addition to performing scheduled inspections. Although the original ATG failure rate is extremely low (around 1/10,000), the most frequent failure is related to the ticket separation module (Figure 1) located at the front of the ATG, for which the overall failure rate is about 30%.

Figure 1: Separation Module in an Automatic Ticket Gate

(https://www.global.toshiba/jp/products-solutions/security-automation/fare-collection/automatic-gate-system.html)

It is important to prevent ticket jamming failures beforehand because SEs need time to adjust the gap between rollers to fix jams in the separation module; in some cases, they need to replace rollers and/or belts, making the ATG unavailable. When station attendants cannot resolve a jam, they call spot maintenance, which may cause delays for parts procurement and ultimately cause long periods of ATG downtime. However, it is hard to collect high quality data for conventional machine learning or deep learning to detect slight signs of failure because extremely imbalanced data between normal and abnormal operation may lack failure signs for machine learning or deep learning due to the extremely low failure rate and complicated mechatronic behavior of ATGs even during normal operation.

For these reasons, we propose a simple but very effective histogram limitation method (HLM) that can detect slight failure signs among the various normal modes in ATGs. Until now, research has been reported on failure prediction for ATGs (JRW, 2022), rule-based diagnosis for ATGs (Shimamura, 2019), ATG state path analysis by sequential mining algorithm (Ueno, 2023), circuit board degradation evaluation for ATGs (Oki, 2022), and anomaly detection for ticket selling devices using deep learning (Xie, 2020). However, to the best of our knowledge, no existing anomaly detection method for separation module of ATGs.

In this paper, we introduce our anomaly sign detection HLM for ATGs with new anomaly sign index, in terms of practical anomaly detection system development for infrastructure in urban cities**.** We describe the basic mechanism of the separation module for ATG in Sec. 2. Next, we discuss our proposed anomaly sign detection HLM in Sec. 3. The evaluation results are reported in Sec. 4, and finally, we summarize our research and identify future work.

2. SEPARATION MODULE

ATGs can automatically separate multiple tickets by feed and reverse rollers at very high speed once these tickets reach the beginning of the separation module. First, the feed roller and reverse roller rotate in the same direction. When thickness sensors detect multiple tickets, the reverse roller rotates in the opposite directions to send back only the upper ticket. The feed roller then sends only the lower ticket farther, and after that, sends multiple tickets down the line once they are aligned in a horizontal line. This process is illustrated in Figure 2.

During maintenance, SEs adjust the gap between the reverse and feed rollers. If the gap is not appropriate, the ATG cannot separate multiple tickets correctly. Some of the ATGs have an automatic gap adjustment mechanism, but the roller may even become enlarged or worn down unevenly due to fluids like water and oil. In that case, even ATGs with automatic gap adjustment need new rollers. In other cases, when the friction between tickets is high, the separation module takes more time to separate them than normal. Furthermore, ATGs take time to separate tickets by sending tickets away from the separation module and retrying the separation them when it cannot separate tickets for some reason.

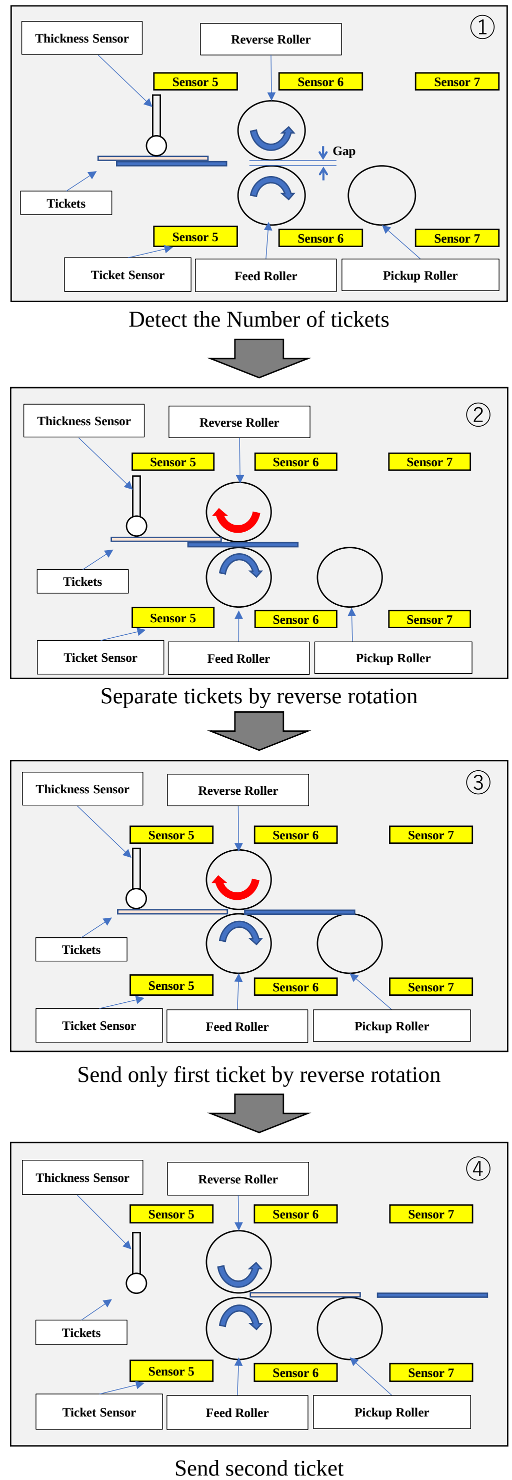

Figure 2: Ticket Processing by the Separation Module

3. HISTOGRAM LIMITATION METHOD

3.1. Basic Concept (HLM ver. 1)

Our preliminary research revealed that slight anomaly signs tend to appear in the tail of a histogram for passing time during separation (i.e., the time needed for a ticket to pass through the separator), that can be calculated by using only system log data. Based on this observation, it seems easy to detect anomaly signs by monitoring the histogram tail. However, the shape of the histogram is also be affected by ticket defects such as those caused by bending, folding, and being wet or oily, which make it difficult to decide which area of the histogram should be focused on.

For these reasons, we propose HLM ver. 1 as a first concept that enables us to calculate the rate of the anomaly score after maintenance by automatically selecting the anomaly sign area (Figure 3) that makes the passing time fastest compared with the one before maintenance in the modeling step. The score can be calculated by the ratio of calculated anomaly sign area (red area in Figure 4) to all area (blue and red areas in Figure 4) in the detection step as follows. Here we define set $\mathbf{S}_1$ and set $\mathbf{S}_2$ as defined in (1)and (2). Then calculate the anomaly sign score, ***anoscore***, as shown in (3). Note that calculated anomaly sign area exists from T_1 to T_2, U refers to upper bound for HLM, and $|\cdot|$ refers to the number of elements in the set.

$$\mathbf{S}_1 = \{i \mid 0 \leqq t_i \leqq U\} \qquad (1)$$

$$\mathbf{S}_2 = \{j \mid T_1 \leqq t_j \leqq T_2\} \qquad (2)$$

$$\boldsymbol{anoscore} = \frac{|\mathbf{S}_2|}{|\mathbf{S}_1|} \qquad (3)$$

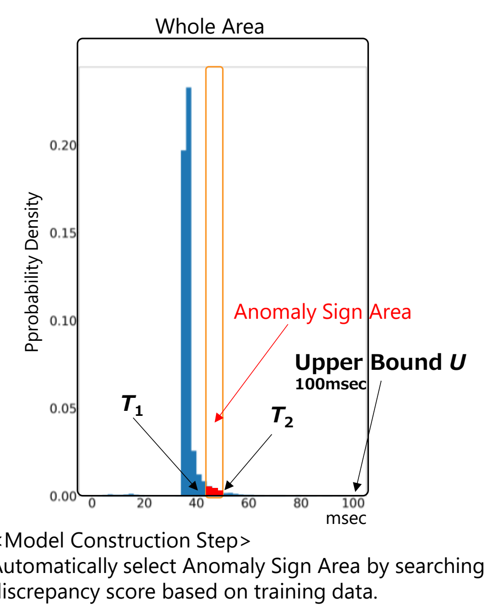

Figure 3: Anomaly Sign Calculation by the Proposed Method (HLM ver. 1) in the Model Construction Step

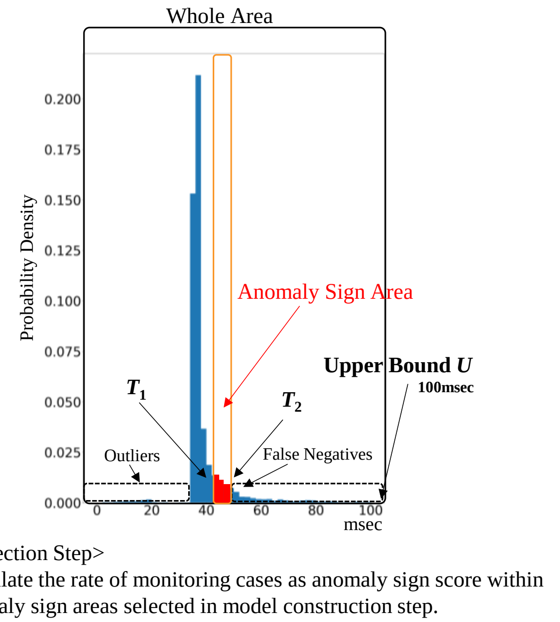

Figure 4: Anomaly Sign Calculation by the Proposed Method (HLM ver. 1) in the Detection Step

The best combination of $\mathbf{L}$ as in (4) consisting of T_1 and T_2 can be calculated as follows. Firstly, we choose samples of passing time corresponding to T_1 and T_2 for a week before maintenance and define as $\mathbf{S}_1^v$, $\mathbf{S}_2^v$. Then, we choose the samples of these corresponding to T_1 and T_2 for a week after maintenance and define as $\mathbf{S}_1^w$, $\mathbf{S}_2^w$. Here we search the best combination of $\mathbf{L}$ consisting of T_1 and T_2 that maximize the discrepancy $d(\mathbf{L})$ defined as eq (5) and calculate $\mathbf{L}_{best}$, the best combination of $\mathbf{L}$, by selecting $\mathbf{L}$ d among all candidate combination of T_1 and T_2 based on the maximum as in (6). Note that $\bar{\mathbf{S}}$ means sample mean of $\mathbf{S}$ and $\boldsymbol{\sigma}$ means standard deviation.

$$\mathbf{L} = (T_1, T_2) \qquad (4)$$

$$d(\mathbf{L}) = \frac{\overline{S_w} - \overline{S_v}}{\sigma_w - \sigma_v} \qquad (5)$$

$$\mathbf{L}_{best} = \underset{d}{\operatorname{argmax}}\, d(\mathbf{L}) \qquad (6)$$

As a result of evaluation, we found that that the ***anoscore*** can partially detect anomaly signs correctly before maintenance. However, the detection results tend to be unstable a short time (i.e., days) before maintenance, and therefore the SE assessments generally do not agree with the calculated scores. This may be because the method often selects a narrow area, especially when the anomaly sign occurs outside the calculated anomaly sign area, which results in false negatives and is affected by outliers, as shown in Figure 4.

3.2. HLM Ver. 2

To improve the stability of the anomaly score, we created HLM ver. 2, which can calculate a stable score by focusing on an area wider than that of ver. 1, including normal and abnormal sign areas on a cumulative histogram (Figure 5). HLM ver. 2 also excludes outlier values caused by complicated ATG mechatronic controls and variations of ticket insert timing and position.

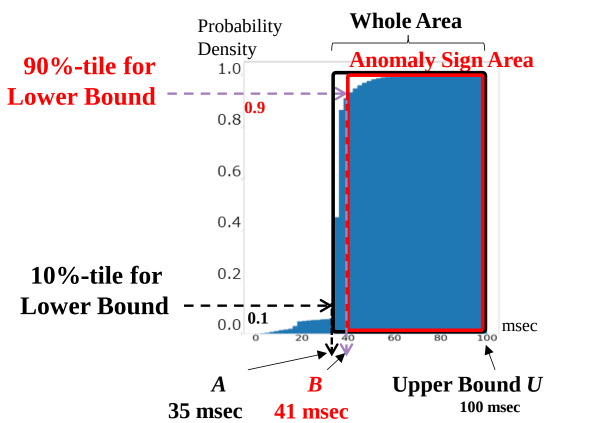

Figure 5: Anomaly Sign Calculation by the Proposed Method (HLM ver. 2) in the Model Construction Step

Specifically, first we set the lower percentile as 10% to remove outliers and set the lower percentile as 100%, and then calculated the lower values *A* and *B*. In this illustrative example, we set the upper value *U* as 90% (100 msec). In the detection step, we focus on the frequency between *A* and *B*. Figure 6 shows an illustrative example to show the difference in the focus area of cumulative histograms between a normal period (approx. 2 months after maintenance) and anomaly period (approx. 2 months before maintenance) for the same ATG. It is clear that the cumulative histogram of normal period is steeper than that of abnormal period. Based on the observation, we revised HLM as described below.

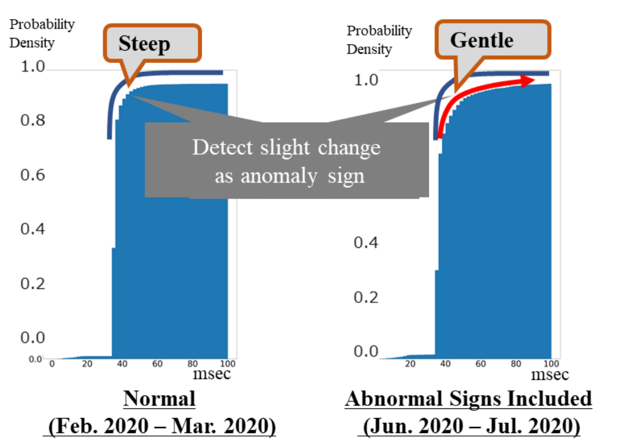

Figure 6: Anomaly Sign Calculation by the Proposed Method (HLM ver. 2) in the Detection Step

Here, we define set $\mathbf{S}_1$ and set $\mathbf{S}_2$ as defined in (7) and (8). Then calculate the ***anoscore*** as shown in (9). Note that $|\cdot|$ refers to the number of elements in each set.

$$\mathbf{S}_1 = \{i \mid A \leqq t_i \leqq U\} \tag{7}$$

$$\mathbf{S}_2 = \{j \mid B \leqq t_j \leqq U\} \tag{8}$$

$$\boldsymbol{anoscore} = \frac{|\mathbf{S}_2|}{|\mathbf{S}_1|} \tag{9}$$

In this example, we set the denominator as the frequency between lower value *A* (35 msec) and *U* (100 msec), and the numerator as the frequency between lower value *B* (41 msec) and *U* (100 msec) based on the observation that the focus area shows slightly different anomaly periods from a normal period in the cumulative histogram, and then calculate the fraction as the ***anoscore***.

The process flow of HLM ver. 2 consists of the model construction step (Figure 7) and anomaly sign detection step (Figure 8). In the model construction step, for example, the system calculates *A* and *B* based on the data for the set of passing time in the separation module containing a normal and abnormal period, before maintenance for about 2 months. Then the system memorizes the parameters *A*, *B*, and *U*, and the anomaly sign judgement threshold *TH*, which is set by the system user currently based on our experience. In the detection step, the system calculates the frequencies $|\mathbf{S}_1|$ and $|\mathbf{S}_2|$ by using *A*, *B*, and *U*. By calculating the fraction described above, the ***anoscore*** and judgement results on each day can be calculated and visualized on the web page for HLM ver. 2.

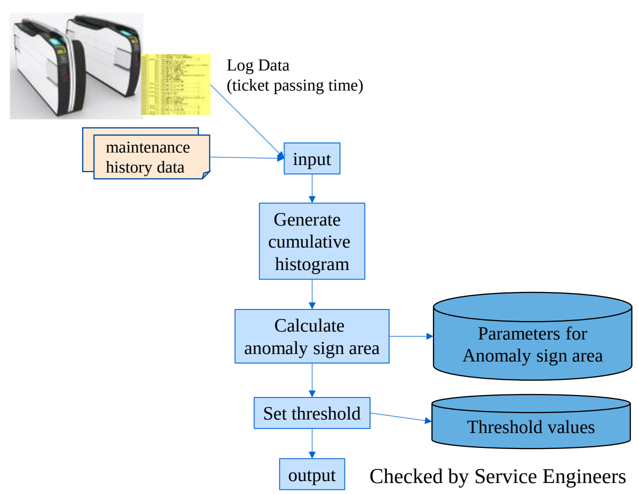

Figure 7: Process Flow of the Proposed Method (HLM ver. 2) in the Model Construction Step

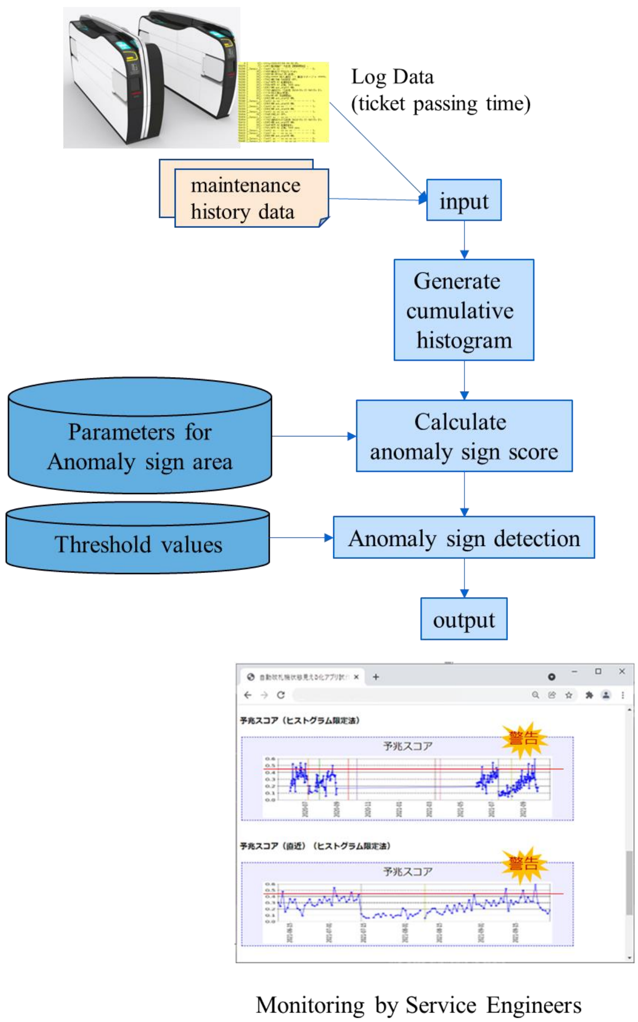

Figure 8: Process Flow of the Proposed Method (HLM ver. 2) in the Detection Step

4. Evaluation

We evaluated the HLM performance with the datasets on five real ATGs in operation at train stations. As a result of HLM ver. 1, the first proposed method detected about 3 weeks before the maintenance day as an earliest warning case (Figure 9). However, the judged anomaly signs are localized immediately before maintenance.

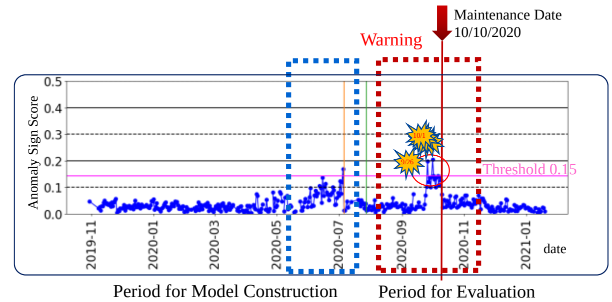

Figure 9: Results of Anomaly Sign Detection by the Proposed Method (HLM ver. 1)

As a result of HLM ver.2, Figure 10 shows that the improved method can successfully detect four anomaly signs, and the earliest warning was recognized about 5 weeks before the maintenance day for one of the real ATGs. Furthermore, we confirmed that these results matched the SEs' evaluation of the field ATG that is getting worse before maintenance.

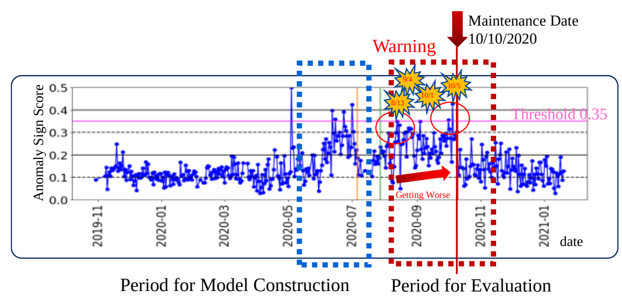

Figure 10: Results of Anomaly Sign Detection by the Proposed Method (HLM ver. 2)

5. Conclusion

In this paper, for a practical anomaly detection system development for infrastructure in urban cities, we proposed our anomaly sign detection HLM for ATGs and its performance with five real ATGs. Especially, HLM ver. 2 detected anomalies 5 weeks before the maintenance day, and furthermore, SEs agreed with the results. Despite it is quite simple and compact approach, the HLM successfully detected anomaly signs for real ATGs in operation. To grasp anomaly signs (anomaly symptoms during normal operation), SEs can perform predictive maintenance to avoid breakdowns. We are now applying the HLM for real ATG monitoring at several sites. In the future, we plan to apply the method to other infrastructure facilities in urban cities and increase the number of ATG monitoring sites for effective and preventive maintenance.

REFERENCES

Ueno., K., Ishikawa M., Kobayashi, Y., Sunaoshi T., Endo, K. (2023). *Failure Sign Detection by State Path Analysis for Fare Collection System - Evaluation by Sequential Pattern Mining with Mechatronics Knowledge -, IEICE Technical Report, Vol.122, No.420, R2022-50, pp.13-18, R2022-50, The Institute of Electronics, Information and Communication Engineers* (*IEICE*)*, Japan.* https://ken.ieice.org/ken/paper/20230310DCSz/eng/

Yamaguchi, A., Ueno, K., Uchida, K., Matsumoto, E., and Saida, T. (2022). *Development of Advanced AI Technologies for Condition Diagnosis of High Voltage Switchgear in Substations. CIGRE Science & Engineering* (*CSE*) *Journal, No.026, International Council on Large Electric Systems* (*Conseil international des grands réseaux électriques, CIGRE*). *https://cse.cigre.org/cse-n026/development-of-advanced-ai-technologies-for-condition-diagnosis-of-high-voltage-switchgear-in-substations*

JRW. (2022). *JR West Press Release on Failure Prediction for Automatic Ticket Gate.* https://www.westjr.co.jp/press/article/items/221205_00_press_infuramente.pdf

Oki. (2022). *Degradation evaluation services applied to the automatic ticket gates for JR East Mechatronics.* https://www.oki.com/jp/case/2022/img/jrem.pdf

Xie, Z., Zhu, J., Wang, F., Li, W., and Wang, T. (2020). *Long short-term memory based anomaly detection: A case study of China railway passenger ticketing system. The Institute of Engineering and Technology* (*IET*) *Intelligent Transportation Systems, Vol.15, Issue 1, pp.98-106, Wiley & Sons.* *https://ietresearch.onlinelibrary.wiley.com/doi/epdf/10.1049/itr2.12007*

Maya, S., Ueno, K., Nishikawa, T. (2019). *dLSTM: a new approach for anomaly detection using deep learning with delayed prediction. International Journal of Data Science and Analytics, vol.8, pp.137-164, Springer.* *https://rd.springer.com/content/pdf/10.1007/s41060-019-00186-0.pdf?pdf=core*

Shimamura, N., Inuiguchi, M., Seki, H., Inoue, M., Takagi, S., Kishine, D. (2019). *If-Then Rule Induction from Historical Data about Automatic Ticket Gate Machines*. In *Proceedings of the Fuzzy System Symposium 2019, vol. 35, pp.772-777, Japan Society for Fuzzy Theory and Intelligent Informatics.* https://www.jstage.jst.go.jp/article/fss/35/0/35_772/_pdf/-char/ja

Ken Ueno received the Ph.D. degree from Keio University, Japan, in 2005. He joined Corporate Research and Development Center, Toshiba corporation, Japan in 2002. He was a visiting post doctorial scholar at department of computer science and engineering, University of California, Riverside, USA, from 2005 to 2007. He is currently Fellow with System AI Laboratory, Corporate R&D Center, Toshiba Corporation, Japan. He engaged in research and development on time series data mining, machine learning, and anomaly detection especially for infrastructure, energy, and manufacturing.

Shigeru Maya received the M.Sc. degree in information science and technology from University of Tokyo, Japan, in 2015. He is currently Research Scientist with System AI Laboratory, Corporate R&D Center, Toshiba Corporation, Japan. His current research interests include machine learning, data mining, and the optimization algorithms.

Kiyoku Endo joined Toshiba Automation Systems Service Co., Ltd. in 2005. He has been working as a maintenance technician for Automatic Ticket Gates or Automatic Ticket Vending Machines for over 10 years. Then he was assigned to Technical Support Department and is now currently working as a specialist on the research and development for Condition Based Maintenance (CBM).