A Data-Driven Approach for Estimating the Effects of Station Closures in Metro Systems

MING YANG1, HUARONG QIN1, KE ZHANG2, XIN DING2, AND ZIYUE MI2

1School of Economics and Management, Ningbo University of Technology, Ningbo 315211, China
2State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Ke Zhang (19120952@bjtu.edu.cn)

This work was supported in part by the State Key Laboratory of Rail Traffic Control and Safety, China, under Grant RCS2020ZT005.

ABSTRACT To estimate the impact of station closures, a novel data-driven approach is developed. First, Symbolic Aggregate Approximation (SAX) is designed to classify stations with an anomaly in passenger flow volume. Secondly, to define the trend anomaly within each segment, the Dynamic Time Warping (DTW) approach is suggested. The method simultaneously considers the mean value and trend information for passenger flow data. To check the efficiency of the proposed model, a case analysis of the Beijing metro system is illustrated. The findings suggest that the proposed model is superior in terms of the calculation of the impact of station closures to other state-of-the-art models.

INDEX TERMS Metro system, station closures, data-driven, dynamic time warping.

I. INTRODUCTION

Metro, the most effective and energy-efficient transit option, is seen as the perfect way to relieve urban pollution and thus plays a growing role in many big cities around the world. Such a heavy dependence imposes enormous strains on metro systems and makes service disruptions hardly affordable. Even minor service disturbances, such as temporary station closures in metro systems, may result in a major loss of public safety and service efficiency, as these disturbances also have ripple effects that extend to other stations and lines [1], [2].

Take the 4-day station closure from 23th April 2019 to 26th April 2019 in Beijing’s metro network for example: train services at 10 stations were disrupted and more than 200,000 passengers were affected. It is therefore more important for operators to recognize the effect of the station closure on their passengers.

The station closures affect passenger flow according to their origins, destinations and routes, such that passenger flow shifts in different stations vary [2], [3]. The review of relevant research focuses on the following aspects: (a) changes in passenger behavior under station closure; and (b) estimates of the impact of station closure based on smart card data.

Understanding the behavior of travelers under metro closures is important for implementing appropriate remedial actions, which will be helpful to minimize impacts to the metro system. However, research on passenger flow characteristics under a metro system closure has been limited as metro station closures are not frequent and the available relevant data is limited. More recently, Pnevmatikou et al. [4] combined information on traveler experiences and perceptions to model mode choice during a long-run metro service disruption. Sarker et al. [5] proposed Affective Events Theory to investigate transit users behavioral reaction to metro service disruptions based on survey data. Saxena et al. [6] analyzed traveler’s mode choice under the two forms of public transport service disruptions using a joint revealed preference and stated preference survey method.

Fortunately, the widespread use of automated fare collection (AFC) data in metro systems offers an alternative chance to understand passenger behavior under station closures massively. However, few studies have utilized the potential of AFC data for metro station closure analysis. Younes et al. [7] investigated the effects of three public transit disruptions in Washington, D.C. on travel model choice and bikeshare demand during the SafeTrack project. This study suggested that the transit station closures had a considerable effect on bikeshare ridership. These above station closures both last for several months, while our study focuses
on temporary station closures that last for several hours or a day.

In general, it is possible to assess the impact of station closures by detecting the temporal-spatial anomaly of passenger traffic in comparison with standard scenarios. Sun et al. [1] have developed a disruption estimation model that, taking into account the spatiotemporal factors involved, affects passengers on the metro network. Zhao et al. [8] suggested statistical and clustering methods to better understand the travel habits of individual Shenzhen metro passengers and the irregular passengers were classified based on the results of the clustering. Tonneller et al. [9] established four methods using smart card logs to identify anomalies in a transport network. In order to detect passenger outflow irregularities and to alert administrators in real time, Wu et al. [2] proposed an anomaly detection approach based on ensemble algorithms. These studies are focused on statistical and clustering approaches and have the following drawbacks: the efficiency of the detection of clustering approaches is susceptible to data size and setting parameters, and the overall statistical index or aggregation results do not give rise to the interpretability of these approaches.

Current anomaly detection work on time series data can be extended to the passenger flow analysis of station closures in order to resolve these problems. A comprehensive literature review concluded that a great deal of technique was built based on time series data in cluster analysis and detection of anomalies [10]. Symbolic Aggregate Approximation (SAX), which uses the mean value of a segment as the symbol, has been widely used in time series data mining [11], [12], as well as in passenger flow analysis [13], as a major symbolic representation process. For example, Zhang et al. [13] proposed that SAX defines the characteristics of the daily passenger flow of each station based on Beijing Metro AFC data. However, in previous studies [14], two key time series characteristics, i.e. period and pattern, are not completely taken into account, as is SAX. Specifically, SAX ignores the pattern of a shift in the value of the segment, which in some cases can lead to an incorrect classification, because it is not capable of distinguishing different time series with different patterns from the same average value symbol. Dynamic Time Warping (DTW) [15], [16] is used to calculate the value change trend in the passenger flow segment, the value change trend during each passenger flow data segment is further calculated using SAX, and a new weight equivalent distance is created taking into account the mean value and trend information of passenger flow segment, the value change trend from the same average value symbol. Dynamic Time Warping of distinguishing different time series with different patterns is used to calculate the value change trend during each passenger flow segment, the value change trend during each passenger flow data segment is further calculated using SAX, and a new weight equivalent distance is created taking into account the mean value and trend information of the passenger flow segment.

SAX-DTW is a robust method for measuring the similarity between two time series with unequal lengths by taking warping alignments and the obtained distance is not sensitive to abnormal points [15], [17], and it is very better in deal with temporal drift than existing improved SAX. Moreover, our proposed method is shape-based data-driven and does not need predefined parameters, thus the estimation effect of station closures is more elastic and more intuitive than existing methods [1], [2]. Second, the real-world traffic scenarios data from Beijing subway is tested in our case study, resulted in strong stability. It shows that the proposed model can more accurately capture the abnormal changes in passenger flow under closure.

The rest of the paper is structured as follows. A theoretical approach for estimating the impact of station closures is suggested in Section 2. The characteristics of the data used from the Beijing subway AFC system are listed in section 3. Section 4 illustrates the application of SAX-DTW in the identification of the effect of passenger flow under station closure, comparing our approach with the other models in detail. The paper ends with section 5.

II. METHOD

Consider a metro network with the set of metro stations V. Let the outbound volume be the feature attribute of each station and there be N days in the sample with Z normal days. The outbound volume of each station i at day j is a time series of length T, represented as $X_{ij} = \{x_{ij1}, x_{ij2}, \ldots, x_{ijT}\}$, $i \in V, j \in N$. The historical average passenger flow of station $i$ on normal days is defined as

$$\bar{x}_i = \left\{ \frac{1}{Z} \sum_{k=1}^{Z} x_{ik1}, \frac{1}{Z} \sum_{k=1}^{Z} x_{ik2}, \ldots, \frac{1}{Z} \sum_{k=1}^{Z} x_{ikt}, \ldots, \frac{1}{Z} \sum_{k=1}^{Z} x_{ikt}, \ldots \right\}$$

(1)

where $\bar{x}_{it}$ is the outbound volume of station $i$ at interval $t$ on normal days and $\bar{x}_{it} = \frac{1}{Z} \sum_{k=1}^{Z} x_{ikt}$.

To address different time series that may have different offsets and amplitudes, time series of outbound volume are converted through Z-score standardization.

$$x^z_{it} = \frac{\bar{x}_{it} - \frac{1}{T} \sum_{t=1}^{T} \bar{x}_{it}}{\sqrt{\frac{\sum_{t=1}^{T} (\bar{x}_{it} - \frac{1}{T} \sum_{t=1}^{T} \bar{x}_{it})^2}{T}}}$$

$$x^{st}_{ijt} = \frac{x_{ijt} - \frac{1}{T} \sum_{t=1}^{T} x_{ijt}}{\sqrt{\frac{\sum_{t=1}^{T} (x_{ijt} - \frac{1}{T} \sum_{t=1}^{T} x_{ijt})^2}{T}}}$$

(3)

where $x^z_{it}$ is the standardized outbound volume at interval $t$ on normal days at station $i$, $x^{st}_{ijt}$ is the standard deviation of the
outbound volume at interval \( t \) on day \( j \) at station \( i \). Hence, the normalized passenger flow of station \( i \) on normal days is denoted as \( X_i^* = \{x_{i1}^*, x_{i2}^*, \ldots, x_{iT}^*\} \) and the normalized passenger flow of station \( i \) on day \( j \) is denoted as \( X_{ij}^* = \{x_{ij1}^*, x_{ij2}^*, \ldots, x_{ijT}^*\} \).

A. QUANTITY ANOMALY IDENTIFICATION

SAX is a technique of data adaptive reduction that transforms a time series into a symbolic sequence in two main steps: (1) transform the raw data into a representation of Piecewise Aggregate Approximation (PAA); (2) transform the representation of PAA into a sequence of symbols belonging to a predefined set of alphabets. Key steps are given as follows, and it is possible to refer to more SAX specifics [21].

First, one time series \( Y \) with length \( T \) are divided into \( w \) segments, then \( Y \) can be represented in a \( w \)-dimensional space by a vector \( \tilde{Y} = (\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_W) \). The \( w \)th element of \( \tilde{Y} \) is calculated by the following equation:

\[
\tilde{y}_w = \frac{1}{T} \sum_{k=T(w-1)}^{Tw} y_k
\]

where \( \tilde{y}_w \) represents the mean value of the outbound volume in the segment of station \( i \) and \( W \) is the number of PAA segments.

Then, the mapping from the above PAA representation \( \tilde{Y} \) to the corresponding symbolized sequence \( \hat{Y} \) is obtained as follows:

\[
\hat{y}_w = \text{alpha}_k, \text{if } \beta_{k-1} \leq \hat{y}_w \leq \beta_k
\]

where \( \hat{y}_w \) is the mapping symbol; \( \text{alpha}_k \) denotes the \( k \)th element of the alphabet, i.e., \( \text{alpha}_1 = \alpha \) and \( \text{alpha}_2 = b \); the breakpoint \( \beta_k \) depends on the alphabet set size \( \alpha \). Specifically, the breakpoints are derived using a Gaussian distribution such that the area from \( \beta_k \) to \( \beta_{k+1} \) is \( 1/\alpha \) [18]. In this way, the time series \( X_i^* \) is first converted to the symbolized sequence \( \tilde{X}_i = \{\tilde{x}_{i1}, \tilde{x}_{i2}, \ldots, \tilde{x}_{iw}\} \) of station \( i \), and then converted to a string sequence.

To further symbolize the data, a distance function of passenger flow at station \( i \) on given day \( j \) is defined to return the minimum distance \( D_t \) between historic average passenger flow \( \tilde{X}_i = \{\tilde{x}_{i1}, \tilde{x}_{i2}, \ldots, \tilde{x}_{iT}\} \) and the passenger flow on given day \( j \) \( X_{ij} = \{x_{ij1}, x_{ij2}, \ldots, x_{ijT}\} \) as

\[
D_t(\tilde{X}_i, X_{ij}) = \sqrt{\frac{1}{W} \sum_{w=1}^{W} \Omega(\tilde{x}_{iw}, \tilde{x}_{ijw})^2}
\]

where

\[
\Omega(\tilde{x}_{iw}, \tilde{x}_{ijw}) = \begin{cases} 0, & \text{if } |\tilde{x}_{iw} - \tilde{x}_{ijw}| \leq 1 \\ \beta_{\max(\tilde{x}_{iw}, \tilde{x}_{ijw})-1}, & \text{otherwise} \\ \beta_{\min(\tilde{x}_{iw}, \tilde{x}_{ijw})}, & \text{otherwise} \end{cases}
\]

B. TREND ANOMALY IDENTIFICATION

DTW is suggested to calculate the pattern correlation between the historical average and any day for each section. DTW is a robust method of calculating the similarity between two time series and overcomes Euclidean distance drawbacks, such as having equal lengths for the two time series and being unreasonably large because of its sensitivity to even slight mis-matches [15], [17], [19]. The pattern anomaly in segment \( w \) is determined by the distance \( D_d \) between the outbound volume of each segment at each station at different dates, calculated as follows:

Let the historical outbound volume of normal days and day \( j \) with station closures in the \( w \)th segment of station \( i \) be

\[
X_i^* = \{x_{w(i-1)}, x_{w(i-1)+1}^*, \ldots, x_{wT}^*\}
\]

and

\[
X_{ij}^* = \{x_{w(i-1)}, x_{w(i-1)+1}^*, \ldots, x_{wT}^*\}
\]

respectively. The warping path \( P \) of \( X_i^* \) and \( X_{ij}^* \) is

\[
P = (p_1, p_2, \ldots, p_u, \ldots, p_U), \quad 1 \leq U \leq 2l - 1
\]

where \( p_u = \min\{d(x_{im}^*, x_{ijk}^*), (x_{im}^* - x_{ijk}^*)^2\} \) represents the Euclidean distance between the \( m \)th point of \( X_i^* \) and the \( k \)th point of \( X_{ij}^* \) in the step \( u \) of the warping path \( P \) and \( l = T/W \).

Then, the DTW-distance \( D_d \) at the \( w \)th segment minimizes the distance between all possible expansions that can be derived from warping path \( P \).

\[
D_d(X_i^*, X_{ij}^*) = \min\left\{\sum_{u=1}^{U} p_u\right\}
\]

Note that computing the DTW-distance and deriving an optimal warping path is usually solved by applying a dynamic program and the pseudo-code of DTW can be referred to existing studies [17].

C. A NOVEL SAX-DTW ANOMALY IDENTIFICATION METHOD

A new distance, called the SAX-DTW distance, is generated to define the passenger flow anomaly and measured as follows, considering the quantity anomaly at different segments and the pattern anomaly within a segment simultaneously.

\[
D_{sd} = \sqrt{\frac{1}{W} \sum_{w=1}^{W} \Omega(\tilde{x}_{iw}, \tilde{x}_{ijw})^2 + \frac{W}{T} D_d(X_i^*, X_{ij}^*)}
\]

where \( \Omega(\tilde{x}_{iw}, \tilde{x}_{ijw}) \) is the SAX-based symbolization distance calculated by Eq. (7), and \( D_d(X_i^*, X_{ij}^*) \) is the DTW trend distance calculated by Eq. (9). Note that the novel distance \( D_{sd} \) is weighted with the trend distance \( D_d \) on the basis of the
existing symbolic distance $\Omega(\vec{x}_i, \vec{x}_j)$ by the weight $w/N$, which reflects the size of the segments, resulting in a more comprehensive distance than basic SAX.

III. EXPERIMENT

The Olympic Sports Center Station on Beijing Subway Line 8 will be chosen for a comparison test to verify the efficiency of the proposed model. The station at the Olympic Sports Center is situated between two transfer stations, west of Forest Park Station’s South Gate. There are more than 14000 visitors a day from Beitucheng Station to the Olympic Sports Center Station due to the major cultural and recreational activities and the many tourist attractions that also take place in the stadium.

The data of passenger flow for the possible affected stations in the Beijing metro are shown in Fig. 1, which introduces the layout of the Beijing metro network. Specifically, the possible affected stations are selected based on travel time: (1) the neighbour stations that belongs to the same Line 8 (i.e., Beitucheng, Olympic Green, South Gate of Forest Park, Lincuiqiao); (2) the stations that needs transfer once and does not the same Line 8 (i.e., Beishatan, Anlilu, and Datunlu East).

The station closure that happened from 17:00 until the last train stopped on April 24, 2017 was selected. The experiment is based on archived 5-minute intervals of historical data. Data from the first two weeks are used to model the passenger flow characteristics at normal days, and the data at the station closure day is used to identify the effects and evaluate the predicted results.

IV. IDENTIFICATION RESULTS AFFECTED BY THE STATION CLOSURE

A. SAX RESULTS

The passenger flow from 10:00 to 22:00 is considered and the possible affected period is divided into 8 intervals according to the proposed method in Section II. Let $\alpha=\{a, b, c, d, e\}$. The mean value of each segment is marked with the corresponding symbol and different colored lines. Red, green, blue, yellow, and black represent a, b, c, d, and e, respectively. The symbol sequence of the normal day and the station closure day time series can be calculated based on the proposed SAX method, as shown in Fig. 2. Note that the left subplot is the PAA symbolized sequence of the station closure day and the right subplot is the PAA symbolized sequence of the normal days.

According to Fig.2, the quantity change of passenger flow under the station closure can be obtained as follows: (1) the passenger flow of Beitucheng and South Gate of Forest Park Station change the most. For example, the PAA symbolized sequence of Beitucheng Station turned from ‘’cbbbbbceceba’’ into ‘’baccececeba’’, where nine of twelve symbols are different shown in Fig.2(d). Also, the PAA symbolized sequence of South Gate of Forest Park Station turned from ‘’ecdedebca’’ into ‘’cbcecececa’’, where seven of twelve symbols are different shown in Fig.2(g). (2) other stations have no change in the passenger flow due to their almost fixed PAA symbolized sequences.

Then, $D_e$ for each time series are calculated as shown in Table 1. Beitucheng and South Gate of Forest Park Station are clearly detected by the SAX distance colored by red in

---

**FIGURE 1.** The sketch map of study area with 8 stations in Beijing Subway.
Moreover, the values of segments are analyzed, some abnormal segments of other stations are detected shown in Table 1. The thresholds of SAX distance for segments of each station and $D_s$ are calculated according to three-sigma rule, respectively. For example, there are several different segments near the time of the station closure (i.e., the seventh segment) for Olympic Green Station (i.e., the third segment shown in Fig.2(b)), Beishatan Station (i.e., the seventh segment shown in Fig.2(c)) and Lincuiqiao Station (i.e., the eighth segment shown in Fig.2(f)). It can be concluded that the effects of the station closure are related with time distance and space distance: more abnormal segments exist in the nearer time distance or in the nearer space distance of the station closure.

**B. SAX-DTW RESULTS**

The trend changes of passenger flow under station closures are calculated by SAX-DTW to further obtain the tiny temporal-spatial range shown in Table 2. The thresholds for segments and $D_{sd}$ are calculated according to three-sigma rule, respectively. The results show that the most affected intervals are from the seventh to the ninth interval during the station closure (i.e., the eighth to the tenth columns of Table 2). During the transition segment before the station closure, the passenger flow of many affected stations decrease significantly compared with the normal days while their passenger flow suddenly increase and the rising trend are obvious after the time of station closure. What’s more, the nearby transfer stations are more likely to be affected, while the stations which have to pass two or more transfer stations are less affected. In this scenario, the most affected stations are Beitucheng and South Gate of Forest Park Stations. Moreover, Olympic Green, Beishatan and Lincuiqiao Stations are less affected.

In summary, our proposed SAX-DTW method can address the anomaly detection under station closures well not only from the view of quantity, but also from the view of trend in segments.

**C. COMPARISONS OF METHODS**

To validate the performance of our proposed SAX-DTW method, it is compared with the existing SAX method and a hybrid method mentioned by [2].

First, our proposed SAX-DTW is compared with the existing SAX method. Take the identification results of Datunlu east and Lincuiqiao Station for example.

1) Datunlu east Station anomaly detection. The change of passenger flow at this station is illustrated in Figure. 3. There was no difference in passenger flow data for the first and seventh segments of the station using the SAX method, but there were variations of passenger flow in the first and seventh segments using our proposed SAX-DTW method. To verify which of the two is rational, data on the passenger flow in the first and seventh segments of the station is combined.
for study. It clearly shows that the historical average data in the first and seventh segments in Figure.3 is relatively stable. However, the flow of passengers fluctuates (a sharp rise and a high, followed by a rapid decline) during the first segments of the station closure. The flow of passengers in the seventh segment also fluctuates considerably (the second box in the picture). In other words, the SAX-DTW algorithm proposed in this paper can reliably detect a pattern shift in passenger flow that is more reliable than the SAX process.

(2) Lincuiqiao Station anomaly detection. The shift in the number of passengers at that station is shown in Figure 4. For the eighth and ninth segments of the station, there was no difference in passenger flow data using the SAX system, but there were differences in passenger volume using our proposed SAX-DTW method. The passenger flow data in the station’s eighth and ninth segments is combined for analysis to verify which of the two is fair. It can be seen from the figure that the passenger flow is much higher than the historical average passenger flow shown in the box in Fig.4 on the closing date of the eighth and ninth segments. Therefore, there are clear variations between these two segments in the amount of passenger traffic. That is, the proposed SAX-DTW algorithm in this paper can accurately define the change in the number of valued passengers, which is more precise than the SAX method.

In short, the proposed SAX-DTW method can be inferred to have better performance in the detection of irregularities under station closures in terms of precision and stability than the current SAX method.

Second, compare the performance of our proposed SAX-DTW method with the hybrid method.

In the hybrid model, LOF algorithm, Grabus criterion, kNN algorithm are combined by weighting method, and voting method is used to integrate independent sample T test,
Wilcoxon Signed Rank Test and Mann-Whitney U test, respectively. A total of five related parameters need to be set: for KNN, Euclidean Distance for distance measurement is selected and the K value is 5, for LOF, the K value is 5 and contamination is 0.1; for Grabus, the confidence probability is 90 percent. The results are given as shown in Table 3 and the anomaly is marked by red color. Take the identification results comparison of South Gate of Forest Park Station and Olympic Green Station for example.

V. CONCLUSION

In this study, a novel approach is proposed for estimating the effects of station closures. To achieve this objective, a novel SAX-DTW method that considers the mean value and trend information of passenger flow data simultaneously is proposed to determine the affected time and station automatically. A case study of the Beijing metro system is done to verify the performance of the proposed model. The proposed model is superior to other state-of-the-art models in terms of the estimation of the effects of station closures.

In reality, the proposed method can offer an abundance of passenger information to metro managers: stations with irregular inbound or outbound volumes lead to the overcrowding of passenger flow management strategies and train service adjustment strategies [20]–[22]. In addition, the proposed method considers the difference in overall passenger flow

TABLE 3. The total results of the hybrid method.

| Segment      | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | Thresholds |
|--------------|----|----|----|----|----|----|----|----|----|----|----|----|------------|
| Anliulu      | 1.21 | 0.40 | 0.34 | 0.42 | 0.51 | 0.17 | 0.66 | 0.83 | 0.71 | 0.61 | 0.22 | 0.16 | 1.21       |
| Olympic Green| 0.76 | 0.33 | 0.46 | 0.35 | 0.30 | 1.40 | 1.08 | 0.87 | 0.71 | 0.33 | 0.86 | 0.37 | 1.64       |
| Beishatan    | 0.42 | 0.70 | 0.28 | 0.81 | 0.51 | 0.29 | 0.30 | 0.96 | 0.47 | 0.74 | 0.26 | 0.56 | 1.33       |
| Beitucheng   | 0.16 | 0.29 | 0.40 | 0.63 | 0.48 | 0.42 | 0.46 | 1.37 | 0.48 | 0.32 | 0.39 | 1.14 | 0.74       |
| Datunlu East | 0.82 | 1.06 | 0.67 | 0.28 | 0.65 | 0.64 | 0.67 | 1.51 | 0.38 | 1.06 | 0.94 | 0.44 | 1.54       |
| Lincuiqiao   | 0.30 | 0.69 | 0.19 | 0.37 | 0.65 | 0.71 | 0.54 | 0.99 | 0.37 | 1.19 | 0.35 | 0.27 | 1.23       |
| South Gate of Forest Park | 0.23 | 1.35 | 0.94 | 1.77 | 2.12 | 0.23 | 0.51 | 1.15 | 1.01 | 0.35 | 0.44 | 0.53 | 2.37       |
during station closure and normal time, but also the trend of passenger flow. In future, more cases and methods should be compared.

REFERENCES

[1] H. Sun, J. Wu, L. Wu, X. Yan, and Z. Gao, “Estimating the influence of common disruptions on urban rail transit networks,” Transp. Res. A, Policy Pract., vol. 94, pp. 62–75, Dec. 2016.

[2] Y. Wu, B. Huang, X. Li, Y. Zhang, and X. Xu, “A data-driven approach to detect passenger flow anomaly under station closure,” IEEE Access, vol. 8, pp. 149602–149615, 2020.

[3] R. Silva, S. M. Kang, and E. M. Airoldi, “Predicting traffic volumes and estimating the effects of shocks in massive transportation systems,” Proc. Nat. Acad. Sci. USA, vol. 112, no. 18, pp. 5643–5648, May 2015.

[4] A. M. Pnevmatikou, M. G. Karlaftis, and K. Kepaptsoglou, “Metro service disruptions: How do people choose to travel?” Transportation, vol. 42, no. 6, pp. 933–949, Nov. 2015.

[5] R. I. Sarker, S. Kaplan, M. Mailer, and H. J. P. Timmermans, “Applying affective event theory to explain transit users’ reactions to service disruptions,” Transp. Res. A, Policy Pract., vol. 130, pp. 593–605, Dec. 2019.

[6] N. Saxena, T. Hossein Rashidi, and J. Auld, “Studying the tastes effecting mode choice behavior of travelers under transit service disruptions,” Travel Behav. Soc., vol. 17, pp. 86–95, Oct. 2019.

[7] H. Younes, A. Nasiri, G. Baiocchi, and L. Zhang, “How transit service closures influence bikesharing demand; lessons learned from SafeTrack project in Washington, D.C. Metropolitan area,” J. Transp. Geography, vol. 76, pp. 83–92, Apr. 2019.

[8] J. Zhan, Q. Qu, F. Zhang, C. Xu, and S. Liu, “Spatio-temporal analysis of passenger travel patterns in massive smart card data,” IEEE Trans. Intell. Transp. Syst., vol. 18, no. 11, pp. 3135–3146, Nov. 2017.

[9] E. Tonnelier, N. Baskiotis, V. Guigue, and P. Gallinari, “Anomaly detection in smart card logs and distant evaluation with Twitter: A robust framework,” Neurocomputing, vol. 298, pp. 109–121, Jul. 2018.

[10] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection for discrete sequences: A survey,” IEEE Trans. Knowl. Data Eng., vol. 24, no. 5, pp. 823–839, May 2012.

[11] H. Yahyaoui and R. Al-Daihani, “A novel trend based SAX reduction technique for time series,” Expert Syst. Appl., vol. 130, pp. 113–123, Sep. 2019.

[12] H. Ruan, X. Hu, J. Xiao, and G. Zhang, “TSAX—An improved time series symbolic representation for classification,” IS & Trans., vol. 100, pp. 387–395, May 2020.

[13] L. Zhang, B. Meng, and Q. Yin, “Classification of urban rail transit stations based on SAX,” J. Geo-Inf. Sci., vol. 18, no. 2, pp. 1597–1607, 2016.

[14] Y. Liu, Z. Liu, and R. Jia, “DeepPF: A deep learning based architecture for metro passenger flow prediction,” Transp. Res. C, Emerg. Technol., vol. 101, pp. 18–34, Apr. 2019.

[15] R. J. Kate, “Using dynamic time warping distances as features for improved time series classification,” Data Mining Knowl. Discovery, vol. 30, no. 2, pp. 283–312, Mar. 2016.

[16] B. J. Jain, “Making the dynamic time warping distance warping-invariant,” Pattern Recognit., vol. 94, pp. 35–52, Oct. 2019.

[17] H. Li, “Time works well: Dynamic time warping based on time weighting for time series data mining,” Inf. Sci., vol. 547, pp. 592–608, Feb. 2021.

[18] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, “A symbolic representation of time series, with implications for streaming algorithms,” in Proc. 8th ACM SIGMOD Workshop Res. Issues Data Mining Knowl. Discovery - DMKD, 2003, pp. 2–11.

[19] H. Kim, C. R. Ahn, D. Engelhardt, and S. Lee, “Application of dynamic time warping to the recognition of mixed equipment activities in cycle time measurement,” Autom. Construct., vol. 87, pp. 225–234, Mar. 2018.

[20] H. Li, Y. Wang, X. Xu, L. Qin, and H. Zhang, “Short-term passenger flow prediction under passenger flow control using a dynamic radial basis function network,” Appl. Soft Comput., vol. 83, Oct. 2019, Art. no. 105620.

[21] C. Yu, H. Li, X. Xu, and J. Liu, “Data-driven approach for solving the route choice problem with traveling backward behavior in congested metro systems,” Transp. Res. E, Logistics Transp. Rev., vol. 142, Oct. 2020, Art. no. 102037.

[22] X.-Y. Xu, J. Liu, H.-Y. Li, and J.-Q. Hu, “Analysis of subway station capacity with the use of queuing theory,” Transp. Res. C, Emerg. Technol., vol. 38, pp. 28–43, Jan. 2014.