

# Improving Accuracy of Anomaly Detection in Spatial-Temporal Population Data through SHAP Values of Reconstruction Errors\*

Ryo Koyama<sup>1</sup>, Tomohiro Mimura<sup>1</sup>, Shin Ishiguro<sup>1</sup>, Keisuke Kiritoshi<sup>2</sup>, Takashi Suzuki<sup>1</sup> and Akira Yamada<sup>1,\*</sup>

<sup>1</sup>NTT DOCOMO, INC, Sanno Park Tower, 2-11-1 Nagatacho, Chiyoda-ku, Tokyo, Japan

<sup>2</sup>NTT Communications Corporation, Otemachi Place West Tower, 2-3-1 Otemachi, Chiyoda-ku, Tokyo, Japan

## Abstract

When accidents, disasters, or other large-scale events occur, they significantly disrupt traffic, leading to congestion and reduced mobility. To effectively address these issues, it is crucial to precisely detect the underlying causes of these disruptions through the analysis of human mobility data. A common approach in anomaly detection is to employ dimensionality reduction techniques to compute reconstruction errors. However, the reconstruction errors generated by traditional methods are influenced by the correlations among features, which may obscure the true causes of anomalies. To overcome this limitation, we introduce an approach that calculates the SHAP (SHapley Additive exPlanations) values associated with the reconstruction errors resulting from dimensionality reduction. We conducted experiments using a dataset of human mobility patterns to evaluate the effectiveness of this method. The results demonstrate that our approach provides a more accurate explanation of anomalies compared to conventional methods.

## Keywords

Anomaly Detection, Spatial-Temporal, SHAP, Dimensionality Reduction, Reconstruction Error

## 1. Introduction

Anomaly detection in spatial-temporal population data has gained significant attention in recent years due to its potential applications in urban planning, disaster management, and public safety. Accurate identification of unusual patterns in human mobility can help authorities respond more effectively to disruptive events such as accidents and disasters. However traditional anomaly detection methods often face challenges in capturing the complex spatial and temporal dependencies in high dimensional population data.

Previous studies have explored various approaches to anomaly detection in spatial temporal data. Ochiai et al.[1] proposed a method that utilizes mesh-based population data derived from mobile communication records to detect non-designated evacuation centers during disasters. Their approach relies on significant reconstruction errors in anomaly scenarios, which are trained only with data representing normal conditions. While this method demonstrates potential, it may not effectively capture the underlying causes of anomalies.

On the other hand, Takeishi[2] demonstrated the effectiveness of using Shapley values to explain the causes of anomalies in dimensionality reduction scenarios. By applying Shapley values to one-dimensional health data, such as myocardial infarction and breast cancer records, Takeishi's method provides a more interpretable understanding of anomaly detection results. However, the applicability of this approach to high-dimensional spatial-temporal data has not been explored.

Building upon the insights from Ochiai et al. and Takeishi, we propose a novel anomaly detection framework that combines the strengths of both approaches. Our method integrates SHAP (SHapley Additive exPlanations) values with dimensionality reduction to identify and explain anomalies in spatial-temporal population data. By leveraging the explanatory power of SHAP values, we aim to improve the

accuracy and interpretability of anomaly detection results, while also extending the applicability of Takeishi's approach to high-dimensional data.

The main contributions of this paper are as follows:

- We introduce a novel anomaly detection method that integrates dimensionality reduction with SHAP values to identify anomalies in spatialtemporal population data.
- We evaluate the effectiveness of our approach using a real-world dataset of human mobility patterns in a major urban area, demonstrating its superiority compared to traditional reconstruction error-based methods.
- We extend the applicability of Takeishi's Shapley value-based approach to high-dimensional spatial-temporal data, enhancing its utility for real-world scenarios.

The remainder of this paper is organized as follows. Section 2 provides an overview of related work in anomaly detection and spatial-temporal data analysis. Section 3 describes our proposed methodology in detail. Section 4 presents the experimental setup. Section 5 discusses the experimental results. Finally, Section 6 concludes the paper and discusses future research directions.

## 2. Related Work

The detection of anomalies in urban population flows has been extensively explored using diverse data sources, including road sensors[3], surveillance cameras[4], and social media data[5]. While road sensors and surveillance cameras prove effective for identifying local abnormalities, their broader application for city-wide anomaly detection is hampered by high installation and maintenance costs. On the other hand, social media data facilitates multimodal anomaly detection methods, such as the integration of bike-sharing and taxi usage history[6], and the semantic interpretation of location-based anomalies identified through social media analytics[7].

STRL'24: Third International Workshop on Spatio-Temporal Reasoning and Learning, 5 August 2024, Jeju, South Korea

\* Current affiliation: NTT Corporation

✉ ryou.koyama.aw@nttdocomo.com (R. Koyama)

📞 0009-0007-4157-1634 (R. Koyama)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

This research leverages population data extracted from communication logs between mobile devices and base stations to enhance anomaly detection capabilities. In contrast to road sensors and surveillance cameras, mobile device data captures a wide array of individual behaviors throughout the entire city, thus offering a more comprehensive solution for detecting anomalies. For instance, Yabe et al.[8] utilized statistical methods to detect the emergence of non-designated shelters during disasters, although their studies lacked a quantitative assessment of accuracy. Similarly, Ochiai et al. used mobile phone-based population data to detect non-designated evacuation sites during disasters by focusing on reconstruction errors. However, as Takeishi has pointed out, these reconstruction errors, significantly influenced by feature interactions, may not accurately pinpoint anomaly locations. To overcome this limitation, Takeishi introduced a method that employs Shapley Values to elucidate the origins of anomalies within dimensionality reduction models.

### 3. Methodologies

This section details our proposed methodology, SHAP Values of Reconstruction Error, for anomaly detection, incorporating SHAP values derived from reconstruction errors. We begin by describing Principal Component Analysis (PCA) as the foundation for dimensionality reduction. Subsequently, we outline the traditional method based on reconstruction errors. Then, we introduce our enhanced approach that integrates SHAP values to improve the accuracy and effectiveness of anomaly detection. Finally, we explain the calculation of SHAP values for multi-dimensional objective variables, extending the methodology to more complex scenarios.

#### 3.1. Principal Component Analysis

PCA is a widely-used technique for reducing the dimensionality of high-dimensional data while preserving the most significant features. By projecting the data onto a lower-dimensional space, PCA identifies the principal components that capture the maximum variance in the data.

Given a dataset  $\mathbf{X} \in \mathbb{R}^{n \times d}$  with  $n$  samples and  $d$  features, the PCA process involves the following steps:

1. **Standardize the Data:** Subtract the mean of each feature from the dataset to center the data around the origin.
2. **Compute the Covariance Matrix:** Calculate the covariance matrix  $\mathbf{C} = \frac{1}{n} \mathbf{X}^T \mathbf{X}$ .
3. **Perform Eigenvalue Decomposition:** Decompose the covariance matrix into eigenvalues and eigenvectors:  $\mathbf{C} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ , where  $\mathbf{\Lambda}$  is a diagonal matrix containing the eigenvalues, and  $\mathbf{V}$  is a matrix whose columns are the corresponding eigenvectors.
4. **Select Principal Components:** Choose the top  $p$  eigenvectors corresponding to the largest eigenvalues to form the principal components. These components maximize the variance retained in the lower-dimensional space.

In this study, we set the threshold for the variance to be retained at 90%. This means that we select the number of principal components  $p$  such that the cumulative variance explained by these components is at least 90%. By retaining the principal components that explain the majority of the

variance, PCA ensures that the most important features of the data are preserved, allowing for effective dimensionality reduction and subsequent analysis.

#### 3.2. Reconstruction Error

Using the principal components obtained from PCA, we can perform dimensionality reduction and reconstruction. Consider a test data vector  $\mathbf{y} \in \mathbb{R}^d$ . The reduced representation  $\mathbf{y}_{red} \in \mathbb{R}^p$  is obtained using the mapping function  $reduce : \mathbb{R}^d \rightarrow \mathbb{R}^p$ , and the reconstructed vector  $\hat{\mathbf{y}} \in \mathbb{R}^d$  is computed using the reconstruction function  $reconstruct : \mathbb{R}^p \rightarrow \mathbb{R}^d$  as follows:

$$\hat{\mathbf{y}} = reconstruct(reduce(\mathbf{y})) \quad (1)$$

The reconstruction error  $e_{\mathbf{y}}$ , a measure of fidelity of reconstruction, is defined as the squared Euclidean distance between the original and reconstructed vectors:

$$e_{\mathbf{y}} = \|\mathbf{y} - \hat{\mathbf{y}}\|_2^2 = \sum_{i=1}^d (y_i - \hat{y}_i)^2 \quad (2)$$

This error metric helps identify significant deviations from normal patterns, which are potential indicators of anomalies.

Additionally, the reconstruction error for each feature  $i$  is calculated as:

$$e_{y_i} = |y_i - \hat{y}_i| \quad (3)$$

This feature-specific error is employed to identify which specific features are exhibiting anomalies.

#### 3.3. SHAP Values of Reconstruction Error

Typically, the SHAP value  $\phi_i(x)$  for each feature  $i$  is calculated using all features of the instance  $x$  as explanatory variables and the prediction  $\hat{y}$  as the objective variable as follows:

$$\phi_i(x) = shap\_value(i; \hat{y}, x) \quad (4)$$

This formulation allows for measuring the impact of feature  $i$  on the prediction  $\hat{y}$ , providing a concrete metric for understanding the significance of each feature in the model.

Similarly, the SHAP value of reconstruction error  $\psi_i(x)$  is calculated using all features of the instance  $x$  as explanatory variables and the reconstruction error  $e_{\mathbf{y}}$  as the objective variable using the following function:

$$\psi_i(x) = shap\_value(i; e_{\mathbf{y}}, x) \quad (5)$$

This formulation measures the impact of feature  $i$  on the reconstruction error  $e_{\mathbf{y}}$ , providing a concrete metric for evaluating the significance of each feature in anomaly detection.

#### 3.4. SHAP Values of Reconstruction Error for Multidimensional Objective Variables

The SHAP value of Reconstruction Error  $\psi_i(x)$  for feature  $i$ , when the objective variable is represented in one dimension, is calculated using the following equation [9]:

$$\psi_i(x) = \sum_{S \subseteq \{1, \dots, d\} \setminus i} \frac{(d - |S| - 1)! |S|!}{d!} [f(x_{S \cup i}) - f(x_S)] \quad (6)$$

In this equation,  $f(x_{S \cup i})$  denotes the model's predicted value when the feature set  $S$  includes feature  $i$ , and  $f(x_S)$  represents the predicted value when the set  $S$  is used without feature  $i$ .  $|S|$  denotes the number of elements in the feature subset  $S$ , and  $d$  is the total number of features.

Subsequently, the SHAP Values of Reconstruction Error  $\psi_i^{(k)}(x)$  for feature  $i$  impacting the  $k$ -th dimension of the objective variable is calculated as the average difference between the model predictions with and without feature  $i$ , across all combinations of features, thus extending equation 6 into the multidimensional context of SHAP Values of Reconstruction Error as follows:

$$\psi_i^{(k)}(x) = \frac{1}{K} \sum_{k=1}^K \left[ \sum_{S \subseteq \{1, \dots, d\} \setminus i} \frac{(d - |S| - 1)! |S|!}{d!} [f^{(k)}(x_{S \cup i}) - f^{(k)}(x_S)] \right] \quad (7)$$

This formulation allows for measuring the impact of feature  $i$  across different combinations of features on each dimension of the reconstruction error. By doing so, it provides a concrete metric for elucidating the significance of feature  $i$  in anomaly detection, offering detailed insights into the causes of anomalies in specific dimensions.

## 4. Preliminaries

### 4.1. Definition: Spatial-Temporal Population Data

This study utilizes Mobile Spatial Statistics (MSS) [10], representing population counts recorded across a two-dimensional geographic grid. Each record, denoted as  $(g, t, pop_{g,t})$ , indicates the population count  $pop_{g,t}$  at grid  $g$  and timestamp  $t$ .

### 4.2. Problem Statement

The goal of this study is to assess the accuracy of anomaly detection in spatial-temporal population data. We compare two methodologies: a traditional approach using reconstruction errors, and a novel approach using SHAP Values of Reconstruction Errors.

#### 4.2.1. Anomaly Insertion Methodology

During the test phase, artificial anomalies are introduced by altering population figures within selected grids. For each timestamp  $t$ , a grid  $g$  is chosen randomly, and  $pop_{g,t}$  is modified to the maximum or minimum value seen during the training period, defined as:

$$pop_{g,t}^{new} = \begin{cases} \max(pop_{g,t'} : t' \in T_{train}), & \text{if max anomaly} \\ \min(pop_{g,t'} : t' \in T_{train}), & \text{if min anomaly} \end{cases}$$

#### 4.2.2. Evaluation Methodology

The efficacy of each detection method is quantified using the Hits@ $k$  metric, which determines if the true anomalous grid is among the top  $k$  ranks based on anomaly scores. The scores are calculated using the following equations:

$$e_{y_g} = |y_g - \hat{y}_g| \quad (\text{as defined in Equation 3}) \quad (8)$$

$$\psi_g^{(k)}(x) = \frac{1}{K} \sum_{k=1}^K \left[ \sum_{S \subseteq \{1, \dots, d\} \setminus g} \frac{(d - |S| - 1)! |S|!}{d!} [f^{(k)}(x_{S \cup g}) - f^{(k)}(x_S)] \right] \quad (\text{as defined in Equation 7}) \quad (9)$$

Rankings for each grid are obtained by comparing these scores against all others in the dataset.

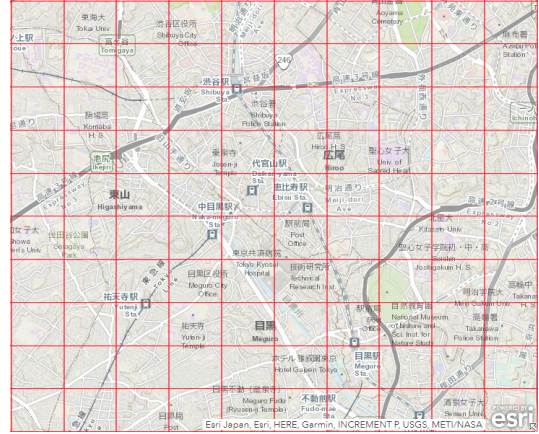
**Hits@ $k$  Calculations** For both methods, Hits@ $k$  is defined and calculated separately for each method to assess the efficacy in identifying the true anomalies within the top  $k$  ranks of predicted anomalies. The total number of test instances, denoted as  $N$ , is used to normalize the calculations, ensuring that the results are proportional to the size of the test dataset. The calculations are as follows:

$$\text{Hits@}k_{\text{reconst}} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}(\text{rank}(e_{y_{g_i}}) \leq k) \quad (10)$$

$$\text{Hits@}k_{\text{SHAP}} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}(\text{rank}(\psi_{g_i}^{(k)}(x)) \leq k) \quad (11)$$

where  $\mathbf{1}(\cdot)$  is the indicator function, and  $N$  represents the total number of test instances. These metrics facilitate a direct comparison of the methods' effectiveness in accurately detecting anomalies.

## 5. Experiments



**Figure 1:** The study area encompassing a total area of 5km x 5km around Shibuya Station, divided into a grid with each cell measuring 500m x 500m.

### 5.1. Dataset

This study utilizes Mobile Spatial Statistics data generated from communication records between NTT DOCOMO's base stations and mobile devices. This data is divided into mesh units across Japan in accordance with the Regional Mesh standards provided by the Ministry of Internal Affairs and Communications Statistics Bureau [11]. Population counts for each grid are estimated at 10-minute intervals, considering factors such as number of devices accessing each base station, market share rates, residential areas, age,

**Table 1**

Examples of Anomaly Detection Results for Randomly Sampled Grids: This table compares the detection rankings from reconstruction error and SHAP value methods across three timestamps. Ranks close to 1 indicate higher detection accuracy.

Timestamp $t$	Randomly Sampled Grid $g$	$pop_{g,t}$	$pop_{g,t}^{new}$	$\text{rank}(e_g)$	$\text{rank}(\psi_g^{(k)}(x))$
2022/10/31 19:10	5339-3588-4	1.378	-1.257	1	1
2022/10/31 10:50	5339-3574-1	1.429	2.187	9	8
2022/10/31 9:10	5339-3584-1	1.536	1.682	6	2

and gender. To ensure privacy, the data is prepared in accordance with guidelines published by NTT DOCOMO[12]. The experimental area, as shown in figure 1, comprises 100 grids of 500 square meters each, centered around Shibuya Station. The population data is treated as 100-dimensional feature vectors and standardized for each dimension.

## 5.2. Training Phase

The training data consists of population records with a 10-minute resolution from October 17 and October 24, 2022, totaling 288 instances ( $6 \times 24 \times 2$ ), were prepared. A PCA model was trained with these data, setting the dimensionality reduction to retain 90% of the variance.

## 5.3. Testing Phase

The test data consists of 144 population records ( $6 \times 24$ ) from October 31, 2022, matching the same month and day of the week as the training data. Anomalies were inserted using the method described in Section 4.2.1. For each instance, one grid was randomly selected, and its population count was replaced with either the maximum or minimum population observed for that grid. A total of 288 tests were conducted to determine if the grid with the altered population could be accurately identified.

## 5.4. Experimental Results

This section presents the results of anomaly detection experiments conducted using both the traditional reconstruction error method and the proposed SHAP value method. The performance of each method is illustrated through selected examples at various timestamps, as detailed in Table 1.

The analysis shows varying levels of detection accuracy for each method, with lower ranking values indicating higher precision in anomaly detection. Specifically, at 19:10 on October 31, 2022, both methods accurately detected the anomaly in grid 5339-3588-4, achieving the lowest possible rank of 1. This instance demonstrates the effectiveness of both approaches in scenarios where there is a substantial change in population, from 1.378 to -1.257.

Conversely, at 10:50 on the same day, the anomaly in grid 5339-3574-1 was detected with lower accuracy, resulting in ranks of 9 and 8 for the reconstruction error and SHAP methods, respectively. This indicates a reduced effectiveness of both methods in detecting anomalies associated with smaller changes in population, from 1.429 to 2.187.

Moreover, at 9:10, the SHAP method outperformed the reconstruction error method by more accurately identifying the anomaly in grid 5339-3584-1, with a rank of 2 compared to 6. This demonstrates the SHAP method’s enhanced ability to detect subtle yet significant changes in population, from 1.536 to 1.682.

A comprehensive evaluation using the Hits@ $k$  metric, which assesses performance under scenarios of maximum

**Table 2**

Comparison of Hits@ $k$  Metrics for Reconstruction Error and SHAP Methods Across All Test Instances: This table presents the performance of both anomaly detection methods under conditions of maximum and minimum population changes. Hits@ $k$  values range from 0 to 1, where values closer to 1 indicate higher accuracy in detecting anomalies.

	MAX		MIN	
	$k = 1$	$k = 3$	$k = 1$	$k = 3$
Hits@ $k_{\text{reconst}}$	0.382	0.458	0.340	0.410
Hits@ $k_{\text{SHAP}}$	<b>0.417</b>	<b>0.472</b>	<b>0.375</b>	<b>0.438</b>

and minimum population changes, is summarized in Table 2. Hits@ $k$  values range from 0 to 1, with higher values indicating more effective anomaly detection.

For maximum population changes, the SHAP method demonstrates superior performance with Hits@ $k$  scores of 0.417 at  $k = 1$  and 0.472 at  $k = 3$ , exceeding the scores of the reconstruction error method, which are 0.382 at  $k = 1$  and 0.458 at  $k = 3$ . Similarly, in scenarios of minimum population changes, the SHAP method achieves better scores of 0.375 at  $k = 1$  and 0.438 at  $k = 3$ , outperforming the reconstruction error method’s scores of 0.340 at  $k = 1$  and 0.410 at  $k = 3$ . These findings confirm the effectiveness of the SHAP method in consistently identifying anomalies under varied conditions, highlighting its superiority over the traditional method.

## 6. Conclusion

This paper evaluated the performance of established reconstruction error techniques and the SHAP value method for anomaly detection in spatio-temporal population datasets. The study highlighted the SHAP method’s enhanced capability for precise anomaly identification, which is crucial for high-accuracy applications such as urban planning and emergency management. The experimental datasets were synthetically modified to include anomalies, offering a controlled setting to examine and contrast the performance of these methods. Future research aims to extend the application of these techniques to real-world data, particularly in scenarios impacted by events like accidents, disasters, or significant public gatherings.

## References

- [1] K. Ochiai, M. Terada, M. Hanashima, H. Sano, Y. Usuda, Detection of non-designated shelters by extracting population concentrated areas after a disaster (industrial paper), in: Proceedings of the 30th International Conference on Advances in Geographic Information Systems, 2022, pp. 1–9.
- [2] N. Takeishi, Shapley values of reconstruction errors



- of pca for explaining anomaly detection, in: Proceedings of the International Conference on Data Mining Workshops, Beijing, China, 2019, pp. 793–798.
- [3] C. Xu, W. Wang, P. Liu, A genetic programming model for real-time crash prediction on freeways, *IEEE Transactions on Intelligent Transportation Systems* 14 (2013) 574–586.
  - [4] A. Adam, E. Rivlin, I. Shimshoni, D. Reinitz, Robust real-time unusual event detection using multiple fixed-location monitors, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (2008) 555–560.
  - [5] X. Lin, A. L. Kenneth, P. R. Spence, Exploring extreme events on social media: A comparison of user reposting/retweeting behaviors on twitter and weibo, *Computers in Human Behavior* 65 (2016) 576–581.
  - [6] Y. Zheng, H. Zhang, Y. Yu, Detecting collective anomalies from multiple spatio-temporal datasets across different domains, in: Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems, 2015, pp. 1–10.
  - [7] B. Pan, Y. Zheng, D. Wilkie, C. Shahabi, Crowd sensing of traffic anomalies based on human mobility and social media, in: Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2013, pp. 344–353.
  - [8] T. Yabe, K. Tsubouchi, A. Sudo, Y. Sekimoto, A framework for evacuation hotspot detection after large scale disasters using location data from smartphones: case study of kumamoto earthquake, in: Proceedings of the 24th SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2016, pp. 1–10.
  - [9] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 4768–4777.
  - [10] M. Terada, T. Nagata, M. Kobayashi, Population estimation technology for mobile spatial statistics, *NTT DOCOMO Technical Journal* 14 (2013).
  - [11] Statistics Bureau of Japan, Area mesh statistics, [https://www.stat.go.jp/data/mesh/m\\_tuite.html](https://www.stat.go.jp/data/mesh/m_tuite.html), 2024.
  - [12] NTT DOCOMO, Mobile spatial statistics guideline, [https://www.docomo.ne.jp/english/binary/pdf/service/world/inroaming/inroaming\\_service/Mobile\\_Kukan\\_Toukei\\_Guidelines.pdf](https://www.docomo.ne.jp/english/binary/pdf/service/world/inroaming/inroaming_service/Mobile_Kukan_Toukei_Guidelines.pdf), 2024.