Urban Network Traffic Analysis, Data Imputation, and Flow Prediction based on Probabilistic PCA Model of Traffic Volume Data

Muhammad Farhan Fathurrahman
Department of Intelligent System
PT. Pusat Riset Energi
Bandung, Indonesia
mf.fathur@rce.co.id

Herman Y. Sutarto
Department of Intelligent System
PT. Pusat Riset Energi and Institut Teknologi Harapan Bangsa
Bandung, Indonesia
hy.sutarto@rce.co.id

Ivana Semanjski
Department of Industrial Systems Engineering and Product Design
Ghent University
Ghent, Belgium
Flanders Make
Lommel, Belgium
Ivana.Semanjski@ugent.be

Abstract—**The growth of vehicle mobility in the past decades and increased traffic complexity leads to a need for traffic management systems, especially in large-scale urban traffic networks. The erroneous data problems are common problems that affect traffic management systems. The traffic management systems also relied on traffic prediction particularly in traffic signal control and route guidance. This paper investigated probabilistic principal component analysis (PPCA) methods to impute missing traffic count data and predict future data. We also investigated the resulting principal components' significance in urban traffic analysis. These methods are applied to traffic count data from vehicle detectors in the urban network of Surabaya city, Indonesia. The results show that the PPCA-based data imputation method is able to impute missing data with imputation error under 20% WMAPE. The resulting principal components analysis demonstrates that 1st principal component scores can be seen as a fundamental temporal pattern of the Surabaya urban network while the characteristic of the link can be derived from the 1st principal component coefficient. We also demonstrate that 1st principal component coefficient of the link might detect outliers or anomalies such as detector malfunction and unique temporal pattern. PPCA can also be used to predict future data based on observed data, but experiments show that even though the majority of the links can be predicted accurately, some links are having large errors that might be caused by different temporal patterns between future data and observed data.**

Keywords—PPCA, traffic count data, missing data, data quality, anomalies detection, prediction, traffic analysis.

I. Introduction

The growth of vehicle mobility in the past decades and the increased complexity of traffic over heterogeneous large-scale networks especially in urban traffic has brought tremendous pressure on urban transport. In order to improve this situation, a traffic management system in urban transport has become a necessity, especially in large-scale networks. However, most traffic management systems require complete and reliable traffic data. For example, intelligent traffic control systems such as SCATS [1] and max pressure control [2,13] require sufficient traffic flow data (i.e., headway, volumes, and speeds) to work effectively. However, a lot of traffic management systems are suffering from erroneous data problems that lead to difficulties to control and predict traffic flow which makes traffic analysis and management becoming less accurate.

The erroneous data problems are a common problem and can be caused by various reasons such as detector faults and connection problems that lead to loss of data. These reasons lead to two different main types of errors, missing data error and incorrect data error.

Missing data error is a common challenge that face many data driven systems. The reasons for this can be manyfold, sensing device malfunction, powering issues, communication error, occlusion, etc. For instance, for the traffic management system that is in the development phase, such as in Surabaya, Indonesia the ratio of missing data from vehicle detectors measuring traffic volume is around 18.3%. Some vehicle detectors do not generate traffic data at all for period of time. These problems are huge obstacles for research in traffic management systems that requires complete and valid data. Several approaches for traffic data imputation have been proposed for missing data error and compiled in survey papers such as Chang & Ge [3] and Li, Li, & Li. [4]. However, traffic data imputation methods failed to solve incorrect data errors.

Incorrect data errors are less common compared to missing data errors, but they are also more challenging to detect. However, their detection is a crucial part for intelligent traffic control systems since the failure to detect might lead to incorrect control strategies. In this context, the incorrect data error can be considered as an anomaly, or an outlier, in traffic data where the observed data is inconsistent with the traffic condition.

Traffic prediction is an important part of traffic management, particularly for route guidance and traffic control. For example, the optimization of traffic signal control can be improved by inputting the prediction of traffic conditions in near future to the traffic signal control system. Similarly, route guidance information can be more accurate if the input involves prediction data in addition to data from current traffic conditions. A problem of traffic prediction in traffic management system, especially in large-scale network, is that it requires methods that are efficient and scalable. Most prediction methods approaches that we found in the literature are univariate prediction methods [14,15]. Such approaches are usually demanding in terms of computation and not scalable for large-scale networks.

In the past years, Principal Component Analysis (PCA) [6] and its maximum likelihood reformulation, Probabilistic Principal Component Analysis (PPCA) [5], are considered as a potential method for both data imputation and traffic prediction in large-scale networks. A study of PPCA-based missing data imputation method is shown by Li & Li [7] and it shows promising results in imputing traffic flow volume data in Beijing, China, which covers 50 loop detectors and 17 intersections with root-mean-square imputation error reduced by at least 25% compared to conventional methods. Coogan, Flores, & Varaiya [9] proposes traffic predictive control based on low-rank structures which are identified using PCA that reduces delay around 7.8 veh·hr per day on a test site in Beaufort, South Carolina. Jenelius and Koutsopoulos [10] propose a PPCA-based network travel time prediction on the road network of Shenzhen, China, and shows a good performance with the assumption that PPCA has multivariate normal distributed variables. The resulting principal components from PCA can also be utilized for traffic analysis such as variability in traffic flow over a network, as done by Tsekeris and Stathopoulos [8].

In this paper, we applied both the PPCA-based data imputation method and PPCA-based network-wide prediction methods to the urban network of Surabaya, Indonesia, using two months of traffic count data from vehicle detectors in the network. The performance of both imputation methods against missing data and network-wide prediction is evaluated. The contribution of this paper is the analysis of the resulting principal component from PPCA model, especially in detecting anomalies from the resulting principal component coefficients.

The rest of this paper is organized as follows: Section II describes the PPCA-based imputation and prediction method. Section III describes the application to traffic count data in Surabaya, Indonesia, with results presented in Section IV. Section V concludes this paper and discusses potential future research from these early results.

II. Methodology

A. PPCA-based Data Imputation

PCA [6] is a standard tool in exploratory data analysis commonly used for dimensionality reduction, data compression, feature extraction, and factor analysis. It can be seen as the orthogonal projection of data to a lower dimensional linear space, called principal subspace, such that the variance in the projected space is maximized. These subspaces are ordered by the variance that is captured in the data so that the first one has the maximum variance that is possible to represent a subspace.

PPCA is a formulation of PCA as a maximum-likelihood procedure based on a probability density model of the data that gives several advantages such as the ability to deal with missing data and better scalability. Multiple PCA models can also be combined as mixtures of PPCA. This PPCA formulation was proposed by Tipping and Bishop [5].

The idea behind imputation using PPCA is that we treat missing data as random variables that have never been observed. If we can estimate the likelihood function from observed data, then we can estimate missing data from the obtained likelihood function.

PPCA, like PCA, defines a relation between observed data with its principal component. Suppose that the observed data is generated from PPCA and the relation between observed data and the principal component described as standard factor analysis mapping [11]

$$\mathbf{y} = \mathbf{W}\mathbf{x} + \boldsymbol{\mu} + \boldsymbol{\epsilon} \tag{1}$$

where $\mathbf{y}$ is a $p \times n$ matrix of observed data with p represents number of time intervals and n represents number of vehicle detectors, $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is the $k \times n$ matrix of principal component coefficients with $k < p$ and assumed to be i.i.d. $\mathbf{W}$ is a $p \times k$ principal component scores matrix that represents linear mapping between observed data $\mathbf{y}$ and principal component coefficients $\mathbf{x}$. $\boldsymbol{\mu}$ is a $p \times n$ matrix of mean value of each variable thus allows the model to have nonzero mean $\boldsymbol{\epsilon}$ is a $p \times n$ matrix represents isotropic noise assumed to be i.i.d. normal with zero mean and σ^2 variance. This noise represents errors in the observed data that are caused by measurement noises.

The number of principal components k is a design parameter in PPCA. Generally, the larger the number of k, the more variance of the observed data is preserved. Larger number of k tends to have better accuracy, however it also leads to model overfit. In order to balance between accuracy and generality, the number of k is calibrated by cross-validation.

The resulting model have following distribution:

$$\boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{W}\boldsymbol{W}^{\mathrm{T}} + \sigma^2 * \mathbf{I}(k)\right) \tag{2}$$

There is no closed-form analytical solution for $\boldsymbol{W}$ and σ^2, so their estimates are determined by iterative maximization of the corresponding log likelihood using an expectation-maximization (EM) algorithm. An efficient EM algorithm of PPCA that estimates $\boldsymbol{\mu}, \boldsymbol{W}$, and σ^2 in the presence of missing data has been derived in [12]. The interested reader can refer to that paper in detail.

B. PPCA-based Network-Wide Prediction

Let $\boldsymbol{t_i}$ be the $1 \times n$ vector of traffic count data for all vehicle detectors in time interval i in current day. All time intervals in current day are stacked in a $p \times n$ vector t. Matrix t is assumed to be generated from PPCA model (1), same as the observed data.

Let j be the current time interval. At time interval j, only the observed data at time interval $\{j - P + 1, \dots, j\}$ are available while the rest of data at time interval $\{j + 1, \dots, j + F\}$ need to be predicted. The matrix t then can be split into observed data $\boldsymbol{t}_P$ and future data $\boldsymbol{t}_F$.

$$\boldsymbol{t}_P = \begin{pmatrix} \boldsymbol{t}_{j-P+1} \\ \vdots \\ \boldsymbol{t}_j \end{pmatrix} \boldsymbol{t}_F = \begin{pmatrix} \boldsymbol{t}_{j-P+1} \\ \vdots \\ \boldsymbol{t}_j \end{pmatrix} \tag{3}$$

The principal component scores matrix $\boldsymbol{W}$ and the mean value matrix $\boldsymbol{\mu}$ then can be split similarly into $\boldsymbol{W}_P, \boldsymbol{W}_F, \boldsymbol{\mu}_P$, and $\boldsymbol{\mu}_F$.

Jenelius and Koutsopoulos [10] derived network-wide prediction based on the conditional distribution of the future data $\boldsymbol{t}_F$ given observed data $\boldsymbol{t}_P$, $\boldsymbol{t}_F | \boldsymbol{t}_P \sim \mathcal{N}(\hat{\boldsymbol{t}}_{F|P}, \boldsymbol{\Sigma}_{F|P})$, from the properties of multivariate normal distributed variables and the matrix inversion lemma where

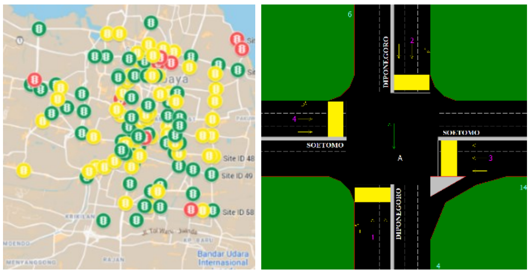
$$\hat{\boldsymbol{t}}_{F|P} = \boldsymbol{\mu}_F + \boldsymbol{W}_F(\boldsymbol{W}_P^T \boldsymbol{W}_P + \sigma^2 \boldsymbol{I})^{-1} \boldsymbol{W}_P^T (\boldsymbol{t}_P - \boldsymbol{\mu}_P) \tag{4}$$

$$\boldsymbol{\Sigma}_{F|P} = \sigma^2 \boldsymbol{W}_F (\boldsymbol{W}_P^T \boldsymbol{W}_P + \sigma^2 \boldsymbol{I})^{-1} \boldsymbol{W}_F^T + \sigma^2 \boldsymbol{I} \tag{5}$$

The $\hat{\boldsymbol{t}}_{F|P}$ matrix represents the point predictor for future data, while the covariance matrix $\boldsymbol{\Sigma}_{F|P}$ represents variability around the point predictions. This normal distribution assumes that the variables have almost the same mean and variance and this is indicated by a similar flow profile.

III. Case Study

Both data imputation and network-wide prediction methodology are applied to traffic count data of the urban network of Surabaya, Indonesia. The traffic count data used in this paper are the number of vehicles passing vehicle detectors aggregated every 15 minutes. There are 285 vehicle detectors counting vehicles that leave a link in an intersection and located at 115 intersections around the urban network of Surabaya, Indonesia, covering an area around 200 km^2 as shown in Fig. 1(a). The placement of the vehicle detectors are denoted by yellow squares in Fig. 1(b). These real-time traffic count data are collected from 1 January 2020 to 29 February 2020.



(a) (b)

Fig. 1. Vehicle detectors in Surabaya, Indonesia. Every traffic light symbol in (a) represents one intersection. The placement of vehicle detectors in an intersection is shown in (b) where each yellow square represents one vehicle detector.

To ensure that the traffic count data has a similar pattern on a day-to-day basis, we only use traffic count data collected during Monday from eight different weeks. Each day has 96 data points, so the total number of data points used is 768 data points for each of the 285 vehicle detectors which leads to 218,880 traffic count data. The resulting traffic count data used for estimating the PPCA model is 768×285 matrix of observed data $\boldsymbol{y}$.

In this case study, we compare how different numbers of k effect to data imputation performance. We also compare three different scenarios of missing data as follows:

- Original traffic count data with 1.4% missing data;
- Traffic count data with 8 hours data omitted (4% missing data); and
- Traffic count data with 64 hours data omitted (33% missing data).

We evaluate our PPCA model by cross-validation to determine the number of principal components k. The resulting k are then applied to the PPCA-based network-wide prediction method. The resulting principal components are also investigated for traffic analysis especially in detecting outliers or anomalies.

IV. Results

A. Imputation Performance Evaluation Criteria

Generally, the better the imputation method performance leads to a closer value between imputed data and missing data. The most common performance criteria for such case is the mean error, which is the average value of error between imputed data and missing data in every time interval. The following mean error are used to determine imputation performance:

1) *Mean absolute error (MAE)*

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p} |\hat{y}_{ij} - y_{ij}|}{p \times n} \tag{6}$$

2) *Root mean square error (RMSE)*

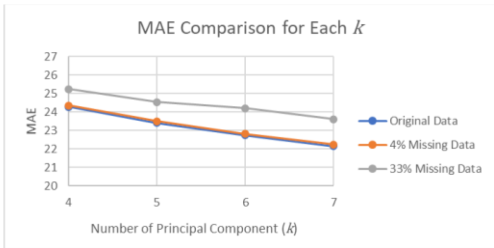
$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p} \left(\hat{y}_{ij} - y_{ij}\right)^2}{p \times n}} \tag{7}$$

3) *Weighted mean absolute percentage error (WMAPE)*

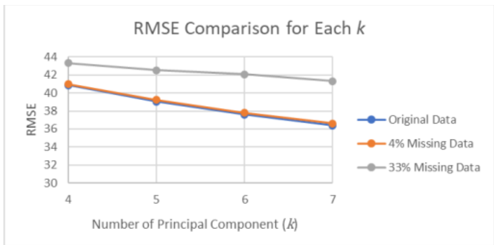
$$\mathrm{WMAPE} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p} |\hat{y}_{ij} - y_{ij}|}{\sum_{i=1}^{n}\sum_{j=1}^{p} y_{ij}} \times 100\% \tag{8}$$

B. PPCA-Based Data Imputation Performance

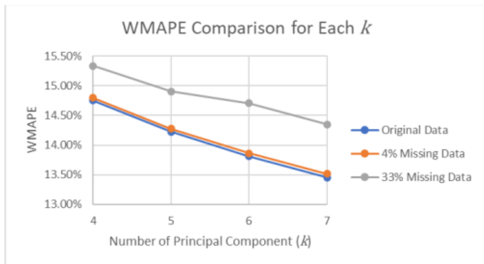
Fig. 2 summarizes the imputation errors in terms of mean absolute error (MAE), root mean square error (RMSE), and weighted mean absolute percentage error (WMAPE) across eight days of data.



(a)



(b)



(c)

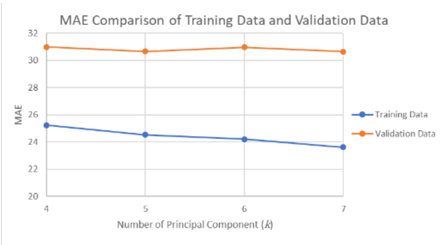
Fig. 2. PPCA-based data imputation performance in terms of (a) MAE, (b) RMSE, and (c) WMAPE

Fig. 2 shows that a higher missing ratio leads to larger error but the difference between original data that have 1.4 % missing data and data that have 33% missing data is only 1% WMAPE. The largest WMAPE, out of all cases, is for the case with $k = 4$ and 33% missing data and equals 15.34%, which is under the acceptable error margin of 20%. Although the error is still acceptable, the error of the original data (without omitted data) is still high. Hence, the error might be affected by the similarity among traffic flow patterns of all the links in the network. In this case study, the location of the vehicle detectors covers areas that have different land use characteristics, such as business districts and residential areas.

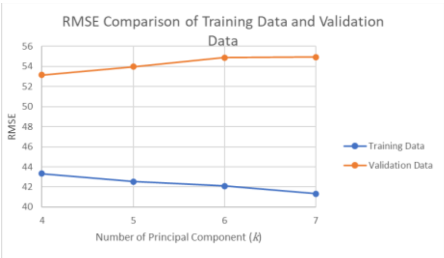
The PPCA-model error might be reduced by clustering areas with similar characteristics and constructing the PPCA model for each area separately.

It should be noted that the observed data should cover the time-of-day of the missing data for PPCA to be able to reconstruct missing data from observed data.

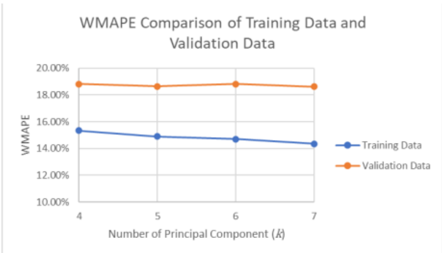
A higher number of principal components tends to give better accuracy because more variance of the observed data is preserved. But it is also known that it tends to overfit the PPCA model. Fig. 3 compares the results of data imputation for training data and validation data sets.



(a)



(b)



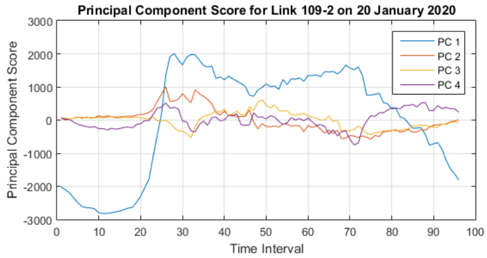
(c)

Fig. 3. PPCA-based data imputation performance comparison for training data and validation data in terms of (a) MAE, (b) RMSE, and (c) WMAPE

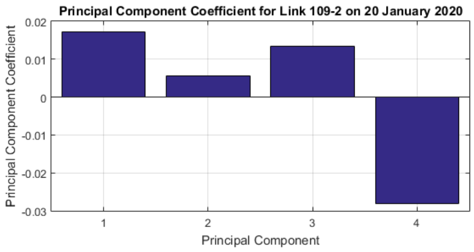
Fig. 3 shows that the error of the PPCA model tested on training data is lower for the higher number of principal components (k) while the error of the PPCA model tested on validation data is higher for the higher number of principal components (k), especially RMSE. This finding means that the higher number of principal components (k) tends to overfit the PPCA model because it doesn't perform better for the higher number of principal components k for validation data.

C. *PPCA for Traffic Analysis*

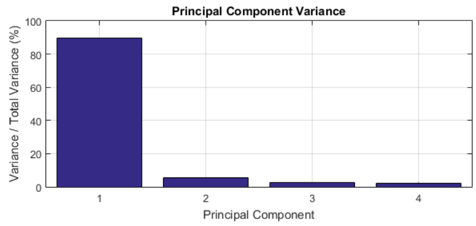
One advantage of the PPCA compared to other data imputation methods is the ability to get principal component score and coefficient of the observed data that help analysis of traffic conditions. For clarity, we only show principal components of the PPCA model for link number 109-2 on 20 January 2020 for the case with $k = 4$ and 8 hours omitted data (4% missing data) in Fig. 4.



(a)



(b)



(c)

Fig. 4. Plot of (a) principal component scores ($\boldsymbol{W}$), (b) principal component coefficients ($\boldsymbol{x}$), and (c) principal component variances of link number 109-2 on 20 January 2020 for the case with k=4 and 8 hours omitted data (4% missing data)

Tsekeris and Stathopoulos [8] presented that PCA can show temporal patterns of traffic data. Fig. 4 shows that the 1st principal component score (blue line) is similar to the common traffic flow pattern that has two peaks, during morning rush hour and evening rush hour, relatively large traffic count during noon while having minimum value during midnight and dawn. The 1st principal component variance is very dominant compared to the rest of the principal components with almost 90% variance out of the total variance of other principal components. This finding shows that the 1st principal component's score can be seen as the fundamental temporal patterns of all links observed in the Surabaya network while the rest of the principal components are minor temporal patterns that show the differences between links.

Fig. 5 shows the comparison of coefficients of the 1st principal components for all links. Most links in the Surabaya network are having a positive value coefficient for 1st principal component which means that most links in the networks possess a similar fundamental temporal pattern. We will analyze two categories of the value of the 1st principal component: the large value of the 1st principal component coefficient and the negative value of the 1st principal component coefficient.

Fig. 6 highlights links that have a large value of 1st principal component such as links number 73, 116, 151, and 158. All the mentioned links are having the same characteristic which is a relatively large mean ($\mu > 500$) compared to the mean of the network ($\mu = 164.5$). This finding is useful for clustering links based on the 1st principal component coefficient, assuming that the 1st principal component variance is dominant.

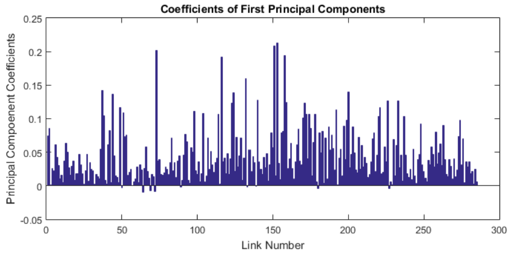


Fig. 5. Comparison of coefficients of the 1st principal components for all links

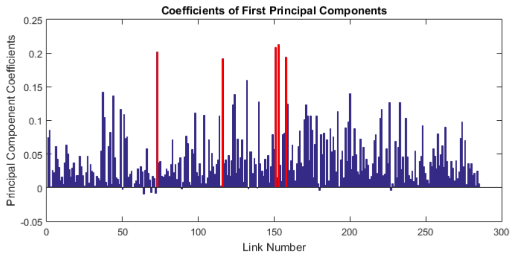


Fig. 6. Comparison of coefficients of the 1st principal components for all links and links that have large coefficient value of the 1st principal component highlighted

Fig. 7 highlights links that have a negative value of 1st principal component such as links number 50, 64, 69, 89, 133, 139, 180, and 227. These links can be considered as anomalies because they do not follow the fundamental temporal pattern and from our findings, it can be caused by links that have unique temporal patterns or events.

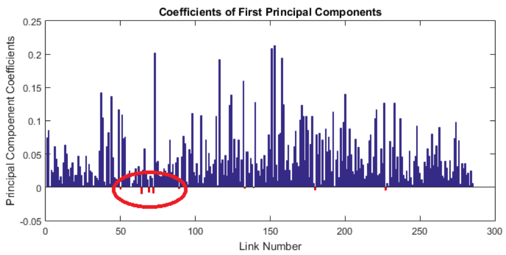


Fig. 7. Comparison of coefficients of the 1st principal components for all links and links that have a negative coefficient value of the 1st principal component highlighted

An example of a link that has a unique temporal pattern is link number 64. From Fig. 8, we can see that this link has a low traffic count during noon as opposed to the relatively large traffic count found in the fundamental temporal pattern. It also has a relatively large traffic count during nighttime. This pattern is consistent over all observed eight weeks of data, which shows that it is not caused by events and can be considered as anomalies in the form of a unique temporal pattern.

Link number 50 is an example of a link that is affected by events. For our case, we consider two events, vehicle detector malfunction, and events such as holiday, road closure, etc. The plot of traffic count data in link number 50 is shown in Fig. 9. In link number 50 case, the event is caused by a detector malfunction where traffic count data from this link number is suddenly dropped to zero during morning rush hour until evening. The vehicle detectors only show a non-zero values during midnight and dawn.

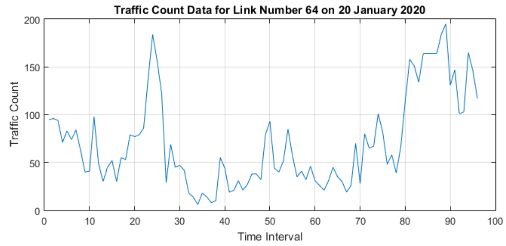


Fig. 8. Traffic count data from vehicle detectors at link number 64 on 20 January 2020. This link is considered as anomalies that caused by a unique temporal pattern.

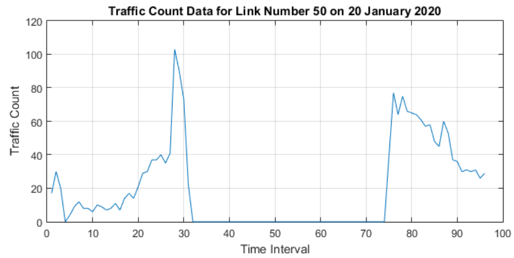
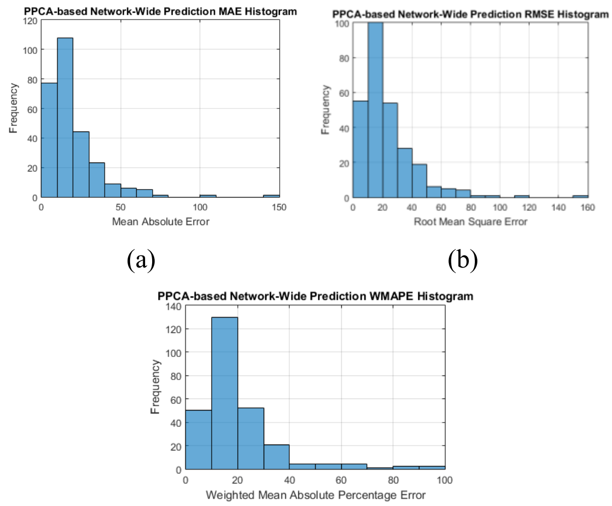


Fig. 9. Traffic count data from vehicle detectors at link number 50 on 20 January 2020. This link is considered as anomaly caused by detector malfunction.

These findings show that the 1st principal component score has the potential to detect anomalies caused by events such as detector malfunction or other events.

D. PPCA-based Network-Wide Prediction Performance

For PPCA-based prediction, we try to predict 3 hours of traffic count data (12 data points from 21.00-24.00) on every link, on 27 January 2020, using data from both 20 January 2020 (21.00-24.00) and 27 January 2020 (00.00-21.00). We choose $k = 4$ for this method based on cross-validation results. Fig. 10 summarizes the prediction errors in terms of mean absolute error (MAE), root mean square error (RMSE), and weighted mean absolute percentage error (WMAPE).

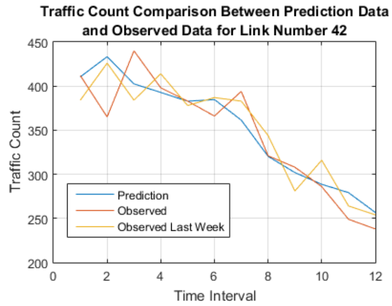


(a) (b)

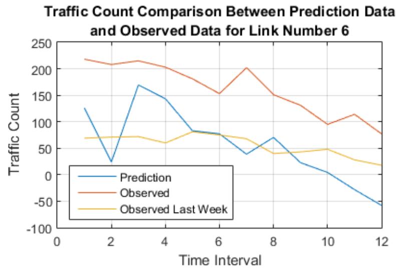
(c)

Fig. 10. PPCA-based network-wide prediction performance histogram plot in terms of (a) MAE, (b) RMSE, and (c) WMAPE. Frequency shows how many links have similar errors.

Fig. 10 shows that the majority of the links have WMAPE lower than 30% which means the majority of the links can be predicted relatively accurately, but some minorities of the links have huge errors. To show why some links are predicted accurately and some links are predicted poorly, we will plot both cases in Fig. 11.



(a)



(b)

Fig. 11. PPCA-based prediction results compared to observed data for (a) link number 42 and (b) link number 6. Link number 42 is an example of accurate prediction (WMAPE = 5.446%) and link number 6 is an example of poor prediction (WMAPE = 74.77%)

From Fig. 11 we can see that link number 42 has a similar temporal pattern and mean between data on 20 January 2020 and 27 January 2020, while link number 6 has a different pattern and large offset between those two dates. This finding means that to be able to predict accurately using PPCA-based prediction methods, the temporal pattern and mean between historical data and predicted data should be similar to some extent.

V. Conclusions

In this paper, a PPCA-based data imputation method, a PPCA-based network-wide prediction method, and urban traffic analysis based on principal components have been applied for traffic count data collected from 285 vehicle detectors in Surabaya, Indonesia. The following conclusions can be drawn from this paper:

- PPCA-based data imputation method is able to impute missing traffic count data with error under 20% using traffic count data consists of 33% missing data.
- The resulting PPCA model, especially the 1st principal components, can be used as traffic analysis tools, such as analyzing characteristics of links in a network and finding anomalies in terms of unique temporal pattern or detector faults.
- PPCA-based network-wide prediction show some promising results with the majority of the links prediction errors under 30% WMAPE, but some links show large prediction error which might be caused by the limitation of the PPCA model to predict data with different temporal pattern.

We propose some related future research regarding these early results are as follows:

- The error of both PPCA-based data imputation and network-wide prediction might be reduced by clustering the network into sub-network to ensure that the principal components obtained are able to reconstruct the original observed data and minimize outlier in each sub network data.
- Alternatively, PPCA-based models can be improved to Mixtures of PPCA-based models so different clusters are covered by multiple different Gaussian distributions, similarly to how multiple PCA models are used for different clusters.

Acknowledgment

The authors would like to thank Kusumo Purnomoputro, PT. Newtel, for helpful discussion and providing traffic count data of Surabaya, Indonesia.

References

[1] P. Lowrie, “Scats, sydney co-ordinated adaptive traffic system: A traffic responsive method of controlling urban traffic,” 1990.

[2] P. Varaiya, “Max pressure control of a network of signalized intersections,” Transportation Research Part C: Emerging Technologies, vol. 36, pp. 177–195, 2013.

[3] G. Chang and T. Ge, “Comparison of missing data imputation methods for traffic flow,” in Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE), 2011, pp. 639–642.

[4] Y. Li, Z. Li, and L. Li, “Missing traffic data: comparison of imputation methods,” IET Intelligent Transport Systems, vol. 8, no. 1, pp. 51–57, 2014.

[5] M. E. Tipping and C. M. Bishop, “Probabilistic principal component analysis,” Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 61, no. 3, pp. 611–622, 1999.

[6] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” Chemometrics and intelligent laboratory systems, vol. 2, no. 1–3, pp. 37–52, 1987.

[7] L. Qu, L. Li, Y. Zhang, and J. Hu, “PPCA-based missing data imputation for traffic flow volume: A systematical approach,” IEEE Transactions on intelligent transportation systems, vol. 10, no. 3, pp. 512–522, 2009.

[8] T. Tsekeris and A. Stathopoulos, “Measuring variability in urban traffic flow by use of principal component analysis,” Journal of Transportation and Statistics, vol. 9, no. 1, p. 49, 2006.

[9] S. Coogan, C. Flores, and P. Varaiya, “Traffic predictive control from low-rank structure,” Transportation Research Part B: Methodological, vol. 97, pp. 1–22, 2017.

[10] E. Jenelius and H. N. Koutsopoulos, “Urban network travel time prediction based on a probabilistic principal component analysis model of probe data,” IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 2, pp. 436–445, 2017.

[11] D. J. Bartholomew, Latent Variable Models and Factors Analysis. USA: Oxford University Press, Inc., 1987.

[12] J. M. Porta, J. J. Verbeek, and B. J. Kröse, “Active appearance-based robot localization using stereo vision,” Autonomous Robots, vol. 18, no. 1, pp. 59–80, 2005.

[13] S. Ramadhan, H. Sutarto, G. Kuswana, and E. Joelianto, “Application of area traffic control using the max-pressure algorithm,” Transportation planning and technology, vol. 43, no. 8, pp. 783–802, 2020.

[14] H. Y. Sutarto, R. K. Boel, and E. Joelianto, “Parameter estimation for stochastic hybrid model applied to urban traffic flow estimation,” IET Control Theory & Applications, vol. 9, no. 11, pp. 1683–1691, 2015.

[15] F. Zheng and H. Van Zuylen, “Urban link travel time estimation based on sparse probe vehicle data,” Transportation Research Part C: Emerging Technologies, vol. 31, pp. 145–157, 2013.