무선 센서 네트워크에서 장애 검출을 위한 결합 주성분분석과 적응형 임계값

Thien-Binh Dang*, Vi Van Vo*, Duc-Tai Le*, Moonseong Kim**, Hyunseung Choo*
*성균관대학교 전자전기컴퓨터공학과
**서울신학대학교 교양학부
e-mail: {dtbinh, ldati, choo}@skku.edu
moonseong@stu.ac.kr

Joint PCA and Adaptive Threshold for Fault Detection in Wireless Sensor Networks

Thien-Binh Dang*, Vi Van Vo*, Duc-Tai Le*, Moonseong Kim**, Hyunseung Choo*
* Dept. of Electrical and Computer Engineering, Sungkyunkwan University
**Dept. of Liberal Arts, Seoul Theological University

Abstract

Principal Component Analysis (PCA) is an effective data analysis technique which is commonly used for fault detection on collected data of Wireless Sensor Networks (WSN). However, applying PCA on the whole data make the detection performance low. In this paper, we propose Joint PCA and Adaptive Threshold for Fault Detection (JPATAD). Experimental results on a real dataset show a remarkably higher performance of JPATAD comparing to conventional PCA model in detection of noise which is a popular fault in collected data of sensors.

1. Introduction

In Wireless Sensor Networks (WSN), the limitation on size and cost of a sensor make it a weak device, such as low computational speed, small memory, limited energy and restricted communication bandwidth [1]. Thus, the WSN are highly vulnerable to random faults and cyber-attacks. The faulty data collected from sensors can leads data analysts to improper decisions. To ensure accuracy and reliability of the sensory data, an efficient fault detection algorithm needs to be developed. Principal Component Analysis (PCA) is an important method to analyze multivariate data obtained from WSN. It was used in many existing works to detect faulty data such as [1], [2]. However, the conventional PCA is not sensitive enough to recognize small faults whose data has small differences to the normal data. In this paper, we propose Joint PCA and Adaptive Threshold for Anomaly Detection (JPATAD) which focuses on improving the sensitiveness of conventional PCA by splitting data into small segments.

2. Joint PCA and Adaptive Threshold for Anomaly Detection

The key idea is that the data is split into smaller segments and the detection threshold adapt to the variation of data from segment to segment. The proposed scheme has two phases: training phase and testing phase. In training phase, considering the data matrix $\boldsymbol{X} = \{\boldsymbol{x^1}, \boldsymbol{x^2}, \dots, \boldsymbol{x^M}\}$ where column i contains N data samples $\boldsymbol{x^i} = \{x_0^i, x_1^i, \dots, x_{N-1}^i\}$ of sensor i collected from normal operation, M is the number of sensors. Then the data $\boldsymbol{X}$ is split into $\frac{N}{S}$ segments where S is the length of a segment (or the number of samples of a segment). Then, PCA is applied on these segments independently to compute Square Prediction Error (SPE) threshold. In testing phase, the incoming data s is first identified the segment it resides in by its collected time and the interval of $\frac{N}{S}$ segments which are established in training phase. PCA is then applied on s to calculate its SPE value. The data s is normal if its SPE value is less or equal to the SPE threshold of identified segment, otherwise s is detected as faulty data.

The advantage of PCA is that it can capture the

correlation of by projecting sensors' data into a lower dimension space which still preserves maximum variance of the original data in minimum number of dimensions. In order to apply PCA, the data matrix is normalized to zero-mean and scaled to unit variance. Let $\boldsymbol{Y_s}$ is the normalized data of a, $\boldsymbol{Y_s}$ can be expressed as:

$$\boldsymbol{Y_s} = (\boldsymbol{Y} - \bar{\boldsymbol{Y}})D^{\frac{-1}{2}}$$

where $\bar{\boldsymbol{Y}} = 1(\frac{1}{N}1^T\boldsymbol{X})$ and $D = \frac{1}{N-1}[(\boldsymbol{X} - \bar{\boldsymbol{X}})^T(\boldsymbol{X} - \bar{\boldsymbol{X}})] \circ I_M$ where $\circ$ denoting the Hadamard multiplication and I_M is the identity matrix. Then, the co-variance matrix $\boldsymbol{R}$ of matrix $\boldsymbol{Y_s}$ is constructed by $\boldsymbol{R} = \boldsymbol{Y_s^T}\boldsymbol{Y_s}$ where $\boldsymbol{Y_s^T}$ is the transpose matrix of matrix $\boldsymbol{Y_s}$. In next step, Singular Value Decomposition (SVD) is performed on $\boldsymbol{R}$ as $\boldsymbol{R} = \boldsymbol{V\Lambda V^T}$ where $\boldsymbol{\Lambda}$ is the diagonal matrix containing M eigenvalues of matrix $\boldsymbol{R}$ in descending order ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M \geq 0$) and matrix $\boldsymbol{V}$ is the collection of M eigenvectors of $\boldsymbol{R}$. SPEs of data samples are calculated for fault detection based on the loading matrix $\hat{\boldsymbol{P}}$ which is formed by l smallest eigenvectors. SPE statistics can be calculated by the following equation: $SPE = \|(\boldsymbol{I} - \hat{\boldsymbol{P}}\hat{\boldsymbol{P}}^{\boldsymbol{T}})\boldsymbol{Y_s}\|$. The data is considered normal if its $SPE \leq \delta^2$, where δ^2 is expressed as follows:

$$\delta^2 = \theta_1\left[\frac{C_\alpha\sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0 - 1)}{\theta_1^2}\right]^{\frac{1}{h_0}}$$

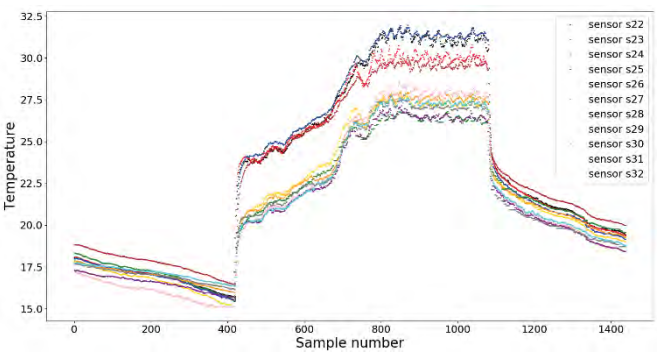
$$h_0 = \frac{2\theta_1\theta_3}{\theta_2^2},$$

$$\theta_i = \sum_{j=l+1}^{L} \lambda_j^i,$$

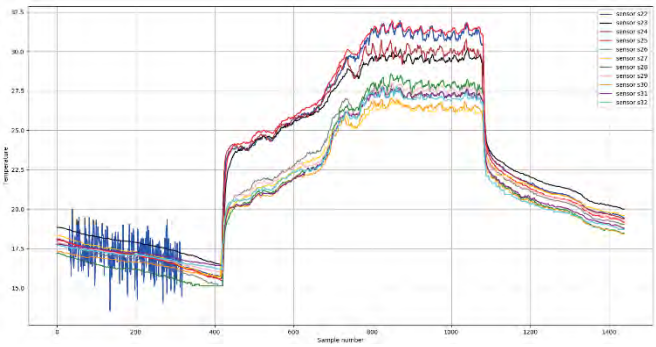
where λ_j is the eigenvalue associated with j_{th} the eigenvector, C_α is the standard normal deviation corresponding to the confident level of standard normal distribution.

3. Performance Evaluation

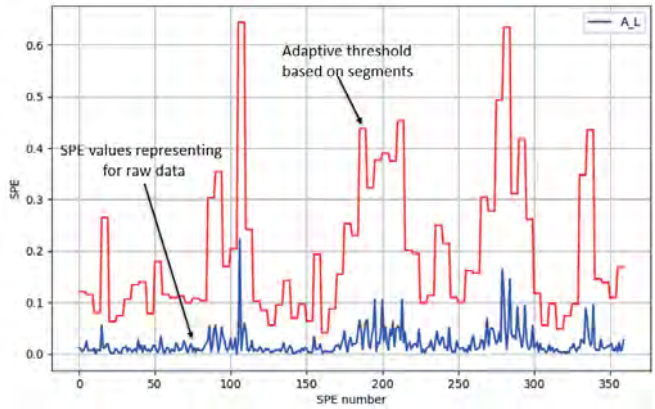
In this research, a real WSN from Intel Berkeley Research lab (IBRL) is used for evaluating the



(Figure 1) Training data



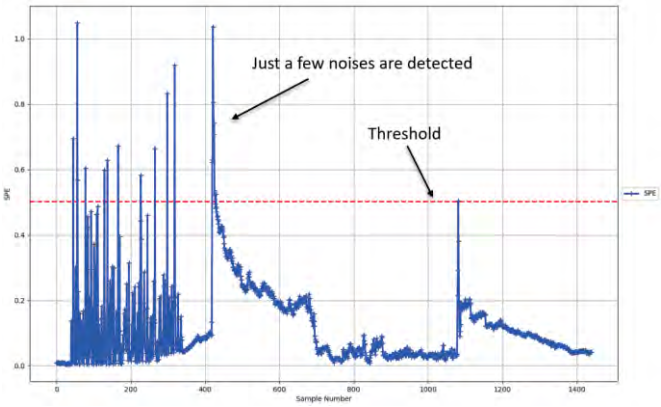
(Figure 2) Training data



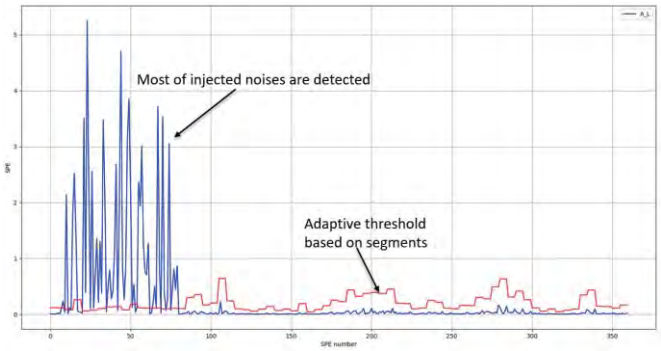
(Figure 3) Noise in sensor 22

efficiency of JPATAD. Without loss of generality, we only choose temperature measurements from eleven sensors whose IDs are 22, 23, ⋯, 32 for experiment. In our experiment, temperature readings from these eleven sensors are re-sampled every minute so a total of 1440 samples are taken in one day. We use 1400 samples of March 1st for training (Figure 1), the segment length S is set to 20 samples and we consider this training data is normal. The adaptive threshold of the training data is shown in Figure 2.

For testing, we inject noise into this normal data (i.e. an example of injected noise is shown in Figure 3) and comparing the performance of conventional PCA and JPATAD.



(Figure 4) Detection result of conventional PCA with noise



(Figure 5) Detection result of JPATAD with noise

The experiment results in the Figure 4 and 5 shown that the conventional PCA cannot detect well the injected noise but the noise is shown clearly in JPATAD depicted by the number of SPE points higher than the SPE threshold (the red dot line).

4. Conclusion

In this work, we proposed Joint PCA and Adaptive Threshold in WSN. The experiment results show that the JPATAD outperforms the conventional PCA in detection noise.

ACKNOWLEDGEMENT

본 논문은 과학기술정보통신부 및 정보통신 기획평가원의 Grand ICT연구센터지원사업 (IITP-2020-2015-0-00742), 정보통신기술 진흥센터 (No.2015-0-00567, 유무선 통합 네트워크에서 접속 방식에 독립적인 차세대 네트워킹 기술 개발), 과학기술 정보통신부 및 정보통신기획평가원의 글로벌핵심인재 양성지원사업 (2019-0-01579)의 연구결과로 수행되었음.

References

[1] D. W Carman, P. S Kruus, and B. Matt, “Constraints and approaches for distributed sensor network security (final),” 10 2000.

[2] M. Livani and M. Abadi, “Distributed pca-based anomaly detection in wireless sensor networks,” in Internet Technology and Secured Transactions (ICITST), 2010 International Conference.

[3] X. L. Zhang, F. Zhang, J. Yuan, J. lan Weng, and W. h. Zhang, “Sensor fault diagnosis and location for small and medium-scale wireless sensor networks,” in 2010 Sixth International Conference on Natural Computation.