Integrated Group-based Valuable Sensor Selection Approach for Remaining Machinery Life Estimation in the Future Industry 4.0 Era

Kun-Chih (Jimmy) Chen and Zi-Jie Gao
Department of Computer Science and Engineering, National Sun Yat-sen University, Taiwan

ABSTRACT

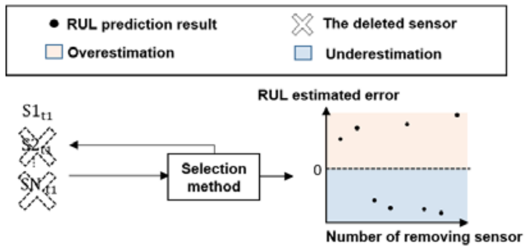
Industry 4.0 is the evolution trend for current manufacturing technology. By analyzing the real-time sensing data, the health status of each machinery is usually monitored to reduce the risk of suddenly machine failure. Although massive sensors allocation can leverage the Remaining Useful Life (RUL) estimation for each machinery, the cost for the sensor network construction will become expensive. Hence, it is necessary to have an approach to remove the redundant sensors under a certain constraint of RUL estimation. On the other hand, due to the attractive performance on the object classification, many researches apply Artificial Neural Network (ANN) to decide which allocated sensor should be removed during the training process. However, the current researches aim to remove the redundant sensors based on the sensing data at a specific time, which lacks the intrinsic feature of time-series sensing data. Therefore, the current researches suffer from the problem of sensor under-killing due to the worst-case consideration. In this paper, we consider the information of time-series sensing data to propose an integrated group-based valuable sensor selection algorithm. Because the proposed approach considers the historical data during the redundant sensor removing process, we can reduce the number of involved allocated sensors precisely and significantly. In order to verify the proposed method, we use the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) dataset and adopt Prognostics and Health Management (PHM) score to evaluate the RUL estimation performance. Compared with the conventional approach, the proposed approach can reduce 86% average PHM score and employ fewer sensors to fit the strict constraint of PHM score with less computing overhead.

I. INTRODUCTION

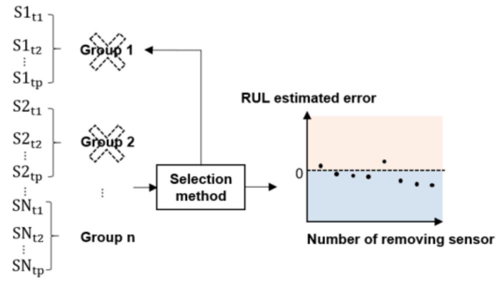
Industry 4.0 is the evolution trend for current manufacturing technology. It incorporates the automation technology, the Internet of Things (IoT) technology, the cloud computing technology, and artificial intelligence (AI) technology, to improve the efficiency and quality of the production in factories [1]. With this new trend, many methods, such as Prognostics and Health Management (PHM) [2], are proposed for the purpose of maintenance management to identify the unreliable machinery in factory. Among the different PHM methods, Conditional-Based Maintenance (CBM) is a widely used approach because it can dynamically arrange a maintenance schedule by monitoring the health status of the machinery [3]. By using the information of CBM, the maintainers of factory can not only prevent the problem that the healthy machinery is replaced too early; but also prevent the risk that the unhealthy machinery is late to replace. To get the precise CBM information, it is necessary to estimate the Remaining Useful Life (RUL) first.

Obviously, massive number of involved sensors can provide much detailed sensing information of the machinery and improve the efficiency of the RUL estimation. However, in practical way, the number of involved sensors is usually limited based on the constraint of the sensor allocation cost and installation space. Hence, it is a critical issue to select minimum number of the required sensors and maintain the precision of the RUL estimation. In [4], Liu *et* al. proposed entropy-based sensor selection algorithm to remove the redundant sensors by analyzing the entropy of time-series input. In [5], Peng *et al*. proposed maximum relevance minimum redundancy (*mRMR*) algorithm to select the relevant features and remove the redundant feature in Lymphoma tissue testing. However, these approaches do not consider the correlation between the selected feature and RUL estimated model. Hence, the selected sensors may not suitable for the involved estimated model, which leads to inefficient RUL estimation.

This work was supported by the Ministry of Science and Technology, TAIWAN, under Grant MOST 108-2218-E-110-010 and MOST 108-2218-E-110 -006



(a)



(b)

Fig. 1 (a) The conventional approach does not consider time-series sensing data; (b) the proposed approach can consider time-series data and select the proper sensors for the efficient RUL estimation.

In recent years, Artificial Neural Network (ANN) is proven as an efficient way to find the correlation between the input and the output without prior knowledge. Therefore, the ANN can be adopted to find the correlation between the input sensing data and the expected RUL result. To find the redundant sensors (*i.e.*, the sensing results are not significant to the expected RUL result), Sun *et al*. applied the property of Least Absolute Shrinkage and Selection Operator (LASSO) to propose a LASSO-based ANN model [6]. By integrating the LASSO penalty to minimizes the mean square error in the ANN training phase, some weights in the input layer of the involved ANN model will become near-zero. The near-zero weight means that the correlation between the corresponding input sensing data and the final output result is not significant. Consequently, through removing the smaller weights of the input layer, the corresponding input sensing data can be neglected (*i.e.*, the corresponding redundant sensors are neglected). However, the RUL information is relevant to the time-series sensing data, and the current LASSO-based ANN model does not consider the time-series sensing data. On the other hand, in most of the industrial applications, the RUL overestimation (*i.e.*, the remaining life of the target machinery is optimistic) may postpone the maintenance of the machinery, which results in a large production yield loss [7]. On the contrary, the RUL underestimation (*i.e.*, the remaining life of the target machinery is pessimistic) only leads to maintenance resource waste. Therefore, the conventional approach may leave some redundant sensors due to the worst-case consideration, as shown in Fig. 1(a).

978-1-7281-6083-2/20/$31.00 ©2020 IEEE

In order to solve the aforementioned problems, we first propose an integrated group-based valuable sensor selection approach. In the proposed approach, we consider several groups of successive time-series sensing data from the corresponding sensors as the inputs of the involved ANN model, as shown in Fig. 1(b). By following the proposed approach, the weights in the input layer of the involved ANN model will become sparse. If one or several weights are near-zero, the sensing data from the corresponding sensor will be neglected. In other words, the sensors, which the providing sensing data is neglected, are the redundant sensors. Furthermore, to prevent the proposed method from the RUL overestimation problem, we proposed a novel loss function in this work. In this way, we can select proper sensors to provide valuable sensing data. The contributions of this paper are summarized below

1) An integrated group-based valuable sensor selection approach is proposed to neglect the redundant sensors in the machinery.
2) A novel loss function is proposed to avoid the problem of RUL overestimation.

In order to verify the proposed approach, we employ the Commercial Modular Aero-Propulsion System Simulation dataset (C-MAPSS) [8] to estimate RUL. In this work, we adopt a widely used metric, Prognostics and health management (PHM) score, to evaluate the performance of RUL estimation. Compared with the conventional approach, the proposed approach can reduce 86% average PHM score and employ fewer sensors to fit the strict constraint of PHM score with less computing overhead.

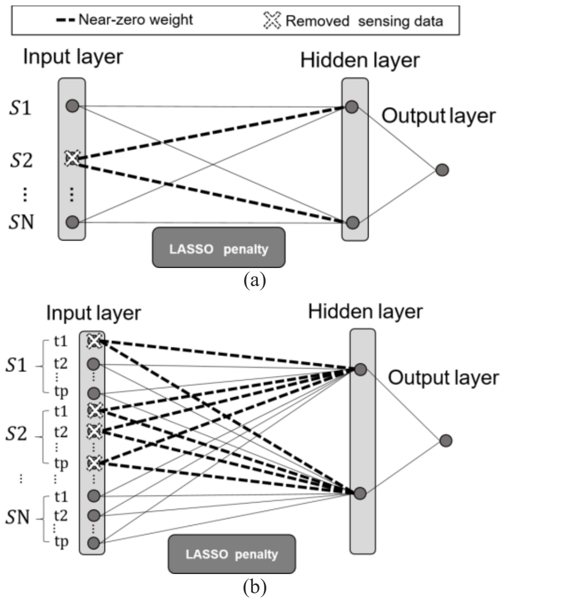


Fig. 2 (a) The conventional LASSO-based ANN model, and (b) the conventional LASSO-based ANN model cannot remove the redundant sensor efficiently.

II. RELATED WORKS

A. *Entropy-based sensor selection algorithm [4]*

In [4], the authors proposed an Entropy-based sensor selection algorithm to reduce the number of sensors for Aero-Propulsion RUL estimation. Because the RUL estimation result depends on the received sensing data, the authors proposed two kinds of sensor selection algorithms to pick the proper sensors: 1) source entropy sensor selection and 2) permutation entropy sensor selection. The first algorithm is used to determine which sensor can provide more information for the RUL estimation. The other one is used to find the proper sensors, which the trend of the sensing data is similar to the RUL trend of the target machinery. By considering the two algorithms simultaneously, the redundant sensors can be determined and neglect the corresponding sensing data. However, this selected algorithm does not consider the correlation between inputs data and the RUL estimation model, which still leaves some unnecessary sensors and receive many redundant sensing data.

B. *Maximum relevance minimum redundancy algorithm [5]*

In [5], the authors proposed a maximum relevance minimum redundancy algorithm to reduce the number of features in Lymphoma tissue testing. In this work, all of the features can be classified into relevant features and redundant features. Because the redundant features cannot provide useful information to estimate the RUL, the authors defined a statistical method to calculate the relevance and redundancy between each feature. Through sorting the result after the calculation, the unnecessary sensors can be removed. Afterward, the left sensors will be used to train the involved RUL estimation model, which is used to estimate the RUL. However, this approach does not consider the correlation between the involved RUL estimation model and the sensor selection strategy. Therefore, the selected sensors may not proper to the involved RUL estimation model, which results in imprecise RUL estimation.

C. *Lasso-based regression algorithm with ANN [6]*

In [6], the authors employed the LASSO property and proposed a kind of regression algorithm to assist with the involved ANN model to diagnose diabetes. By integrating the LASSO penalty to the involved mean square error (MSE) loss function during training the regression model, some trained weights in the ANN input layer will become near-zero. The near-zero weights mean that the correlation between the corresponding input and the final output result is not significant. Hence, those inputs with near-zero weights could be removed. However, the RUL information is relevant to the historical sensing data, and LASSO-based regression algorithm with ANN does not consider the time-series sensing data. Hence, this approach still suffers from the imprecise RUL estimation result.

III. INTEGRATED GROUP-BASED VALUABLE SENSOR SELECTION FOR RUL ESTIMATION

A. *Background of LASSO-based ANN model*

LASSO [9] is an approach to fit the linear regression model by minimizing the mean square error under a constraint on the sum of the absolute coefficient value for the regression, which can be defined as

$$\begin{aligned} &\phi = \arg\min_{\beta_i} (\hat{y} - \sum_{i=1}^{P} X_i \beta_i)^2 \\ &\text{s.t.} \sum_{i=1}^{P} |\beta_i| \le t \end{aligned} \quad , \tag{1}$$

where $\hat{y}$ is the ground truth; $\beta_1, \beta_2, \ldots$, and β_p are the P different coefficients; $X_1, X_2, \ldots$, and X_P are the P distinguished input variables; $\sum_{i=1}^{p} |\beta_i|$ is the constraint of mean square error, which cannot greater than a parameter t. If t is small sufficiently, some of β_i will become near-zero. Because the linear regression model is a kind of multiply-accumulation operation, it can be seen as an ANN operation. Besides, those inputs with near-zero weights mean that the values will not affect the results of ANN operation significantly. Hence, the corresponding input (*i.e.*, the sensing data provided by sensors) can be neglected by adjusting t.

As mentioned before, the LASSO-based linear regression model can be seen as an ANN operation. Because ANN is proven as an efficient way to find the correspondence between input and output without prior knowledge, the authors in [6] proposed a variable selection method for the LASSO-based ANN operation. The authors adopt an ANN with one hidden layer as the involved model, which could be formulated as

$$y_{\text{ANN}} = f(\sum_{j=1}^{q} \beta_j^0 f((\sum_{i=1}^{p} \beta_{ij} X_j) + B_i^h) + B^0), \quad (2)$$

where p is the number of input neurons; q is the number of hidden neurons; X_j is the input variable; β_{ij} is the weight between the X_j and the i-th hidden neurons; B_i^h is the bias of i-th hidden neurons; B^0 is the bias of the output neuron; β_j^0 presents the j-th output weight between the hidden layer and the output layer; f is the involved activation function. By changing the model, the formula (1) can be derived to

$$\theta = \arg\min_{\beta_i} \left(\hat{y} - y_{\text{ANN}}\right)^2 \quad . \quad (3)$$
$$\text{s.t.} \sum_{i=1}^{P} |\beta_i| \le t$$

By using the LASSO property, the constraint of mean square error can make some regression coefficients in the regression model (*i.e.*, the weights of the ANN model in (2)) become sparse after several iterations during the training phase. Therefore, the inputs with the near-zero weights can be seen as the redundant input data, and the corresponding sensors can be neglected, as shown in Fig. 2(a).

Although LASSO property can help to find the redundant inputs in the involved ANN model, it does not consider the intrinsic feature of the time-series sensing data, which is a common feature in the sensing data type of the factory. Hence, the conventional approach may leave some improper sensors, which provides partial redundant sensing data. Fig. 2(b) illustrates an example. Through conventional model, Sensor *S2* can be neglected because the weights of all the sensing data from Sensor *S2* are near-zero. However, we cannot remove Sensor *S1* even though some sensing data from Sensor *S1* is seen as the redundant sensing data. Therefore, the conventional approach cannot remove the redundant sensors efficiently, which results in the imprecise RUL estimation. In addition, the involved MSE loss function in the LASSO property cannot distinguish the error from the RUL overestimation or RUL underestimation. Consequently, through the conventional LASSO-based ANN approach, we may have a result of small RUL overestimation error, which still leads to dramatic production yield loss.

B. *Proposed integrated group-based valuable sensor selection algorithm and the corresponding loss function*

To solve the aforementioned problems, we propose an integrated group-based sensor selection approach in this work. Different from the conventional LASSO-based ANN model, we consider several groups of successive time-series sensing data from the corresponding sensors as the inputs of the involved ANN model for RUL estimation. The number of successive time-series sensing data in each input group should be identical, and the maintainers can define it. To consider all successive time-series sensing data simultaneously, in the proposed ANN model, we add an additional hidden layer accordingly, which is called as Historical Feature Estimation (*HFE*) layer.

Fig. 3 illustrates an example of the proposed ANN model. We assume that p successive time-series sensing data of each sensor are considered. Furthermore, we will have N neurons in the first hidden layer if the sensing data from N sensors are considered initially. By connecting each input sensing data to the corresponding neurons in *HFE* layer, the historical feature of the successive time-series sensing data from a specific sensor will be extracted. Afterward, the extracted features of the sensing data from each sensor will become the input of the second hidden layer, which is called as Group Feature Integration (*GFI*) layer. Obviously, we can neglect all the sensing data from a specific sensor by removing the corresponding neuron in the *HFE* layer. To achieve this goal, we adopt LASSO property and the (3) can be modified to

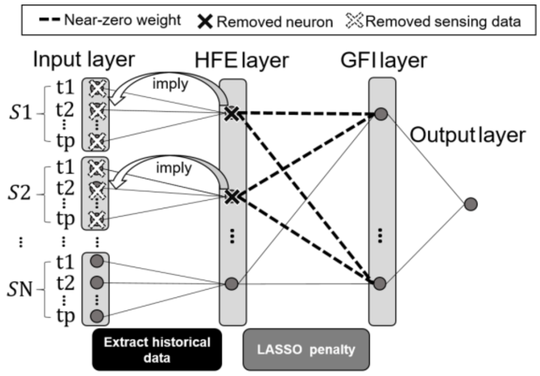


Fig. 3 Integrated group-based sensor selection model.

$$\delta = \arg\min_{\beta_i} (\hat{y}\text{-}y'_{\text{ANN}})^2 \quad , \quad (4)$$
$$s.\text{t.} \sum_{i=1}^{k} |\beta^g{}_i| \le t$$

where B_i^g is the k-th weight of GFI layer and y'_{ANN} is the result of the proposed ANN model. Afterward, the neurons, which provide the inputs to the *GFI* layer with near-zero weights, as well as the corresponding sensors will be neglected, as shown in Fig. 3.

On the other hand, to distinguish the loss between RUL overestimation and RUL underestimation, the authors in [7] proposed a metric, Prognostics and health management (PHM) score, to evaluate the quality of the RUL estimation, which is formulated as

$$\text{PHM score} = \begin{cases} \sum_{i=1}^{N} (e^{-\frac{h_i}{13}} - 1), & \text{for } h_i < 0 \\ \sum_{i=1}^{N} (e^{\frac{h_i}{10}} - 1), & \text{for } h_i \ge 0 \end{cases} , \quad (5)$$

where the N is the number of RUL estimation results. Besides, h_i is the RUL estimation error of the i-th RUL estimation, and it can be formulated as

$$h_i = (\text{Estimated RUL} - \text{True RUL}). \quad (6)$$

Obviously, with small h_i, the PHM score in (5) will become smaller. Hence, the smaller PHM score presents a better RUL estimation. In addition, the PHM score will provide more penalty (*i.e.*, the score growing rate) to the situation of RUL overestimation (*i.e.*, $h_i \ge 0$) because RUL overestimation leads to more negative impact to the factory than the RUL underestimation (*i.e.*, $h_i < 0$.) To consider the quality of the current RUL estimation, we employ the PHM score instead of the MSE in the LASSO property, and the proposed new loss function can be formulated to

$$\phi = \begin{cases} \arg\min_{\beta_i} \sum_{i=1}^{N} (e^{-\frac{h_i}{13}} - 1), \ h_i < 0 \\ \arg\min_{\beta_i} \sum_{i=1}^{N} (e^{-\frac{h_i}{10}} - 1), \ h_i \ge 0 \end{cases} , \quad (7)$$
$$s.t. \sum_{i=1}^{P} |\beta_i| \le t$$

where N is the number of RUL estimation results.

IV. EXPERIMENTAL RESULTS

To verify the proposed approach, we employ the C-MAPSS dataset [8], which adopts 21 sensors to record the process of the turbofan engines from health state to failure. Similar to [8], the

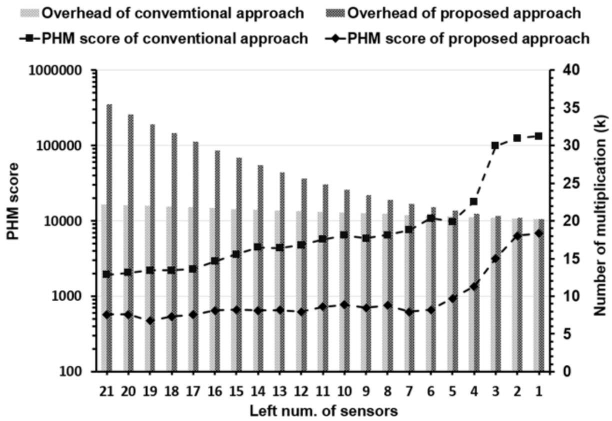


Fig. 4 The comparison of PHM score and the computing overhead under different number of removing sensors.

piecewise linear degradation labeling method is adopted. In the following experiments, we will compare the performance between the proposed approach and the conventional LASSO-based ANN [6]. To fit the involved C-MAPPS dataset, the adopted LASSO-based ANN contains four hidden layers, and 50 neurons are included in each hidden layer. According to the proposed ANN model, we add one additional *HFE* layer in front of the original four hidden layers of the LASSO-based ANN, and 21 neurons are used in the *HFE* layer because 21 sensors are considered in this dataset initially. To simplify the problem, the length of the successive time-series sensing data in each input group of the proposed approach is set to 30 as a design example.

A. The PHM score analysis for the RUL estimation

In order to compare the performance of the RUL estimation by using different sensor selection methods, Fig. 4 shows the PHM score under different number constraint of involved sensors. Obviously, the sensing information of machinery status become fewer with respect to the fewer sensors (*i.e.*, more sensors are removed.) Therefore, the PHM score will become worse gradually. Compared with the conventional approach [6], the proposed method can help to reduce 86% average PHM score. The reason is that the conventional approach estimates the RUL only adopts specific sensing data, which does not consider the time-series sensing data to monitor the damage trend of the machinery status. In the contrary, the proposed approach can select the proper sensors, which provide valuable sensing data for the RUL estimation, by considering the time-series sensing data.

B. The computational overhead analysis of RUL estimation

Because of one additional hidden layer, it seems that the computing overhead of the proposed approach may be increased, as shown in the bar chart of Fig. 4. However, under a certain constraint of PHM score, the proposed approach can reduce the computing overhead significantly by using fewer and valuable sensing data. To analyze the benefit bringing from the proposed method, we compare the number of involved multiplications under five different PHM scores, which are 500, 1000, 1500, 2000, and 2500. As shown in Table 1, the proposed approach can employ fewer sensors (*i.e.*, more sensors are removed) than the related work to fit the constraint of PHM score, which helps to reduce the number of involved multiplications. While the constraint of the PHM score becomes strict (*i.e.*, the PHM score is smaller than 1500), the proposed approach can still fit the target of PHM score with fewer sensors. On the other hand, the conventional approach cannot fit the target of PHM score even though employing all 21 sensors because it does not consider the time-series sensing data to estimate the RUL.

Table 1 The number of removing sensors and correspond multiplication between five PHM score

	Related work [6]		Proposed work	
PHM score	Left num. of sensors	Number of multiplication	Left num. of sensors	Number of multiplication
2,500	17	22,100	3	20,670
2,000	20	21,800	3	20,670
1,500	21	22,200	3	20,670
1,000	21	22,200	4	21,350
500	21	22,200	19	32,830

V. CONCLUSION

RUL is a critical factor to estimate the CBM in the modern factory, and it highly depends on the results of the sensing data provided by the installed sensors in the machinery. However, the conventional approaches do not consider the time-series sensing data, which results in imprecise RUL estimation. To solve this problem, we first propose an integrated group-based valuable sensor selection approach. By considering the successive time-series sensing data, the proposed approach can select the proper sensing data to estimate the RUL. Furthermore, we proposed a novel loss function to solve the problem of RUL overestimation. The experimental results show that the proposed approach can reduce 86% average PHM score, which is a widely used metric to evaluate the performance of RUL estimation. Besides, compared with the conventional approach, the proposed method can employ fewer sensors to fit the strict constraint of PHM score with less computing overhead.

REFERENCE

[1] J.-R. Jiang *et al.*, "Enhancing Convolutional Neural Network Deep Learning for Remaining Useful Life Estimation in Smart Factory Applications," *IEEE International Conference on Information, Communication and Engineering (ICICE)*, pp. 17-20, Nov. 2017.

[2] J. Lee *et al.*, "Introduction to data-driven methodologies for prognostics and health management," *Probabilistic Prognostics and Health Management of Energy Systems*, pp. 9-32, Apr. 2017.

[3] Udren and Eric A, "Protection system maintenance program choices-TBM, CBM, and PBM," *IEEE Annual Conference for Protective Relay Engineers*, pp. 10-22, Mar. 2014.

[4] L. Liu *et al.*, "Entropy-based sensor selection for condition monitoring and prognostics of aircraft engine," *Microelectronics Reliability*, vol. 55, pp. 2092-2096, Aug. 2015.

[5] H. Peng *et al.*, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 122-1238, Aug. 2005.

[6] K. Sun *et al.*, "Design and application of a variable selection method for multilayer perceptron neural network with LASSO," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 6, pp. 1386-1396, Jun. 2017.

[7] G. Babu *et al.*, "Meta-cognitive regression neural network for function approximation: application to remaining useful life estimation," *IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 4803-4810, Jul. 2016.

[8] X. Li *et al.*, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Reliability Engineering and System Safety*, vol. 172, pp. 1-11, Apr. 2018.

[9] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royak Statistical Society: Series B (Methodol.)*, vol. 58, no. 1, pp. 267-288, Jan. 1996.