Feature Selection Method for Life Prediction in Multiple Degradation Unit: Generalized Rank Mutual Information

Taewan Hwang[1], Keunsu Kim[2], Su J. Kim[3], Byungjoo Jeon[4], and Byeng D. Youn[5]

[1,2,3,4]*Seoul National University, Seoul, 08826, Republic of Korea*
michel4546@gmail.com
keunshu@gmail.com
hippo130@naver.com
jbj0418@gmail.com

[5]*Seoul National University & OnePredict Inc., Seoul, 08826, Republic of Korea*
bdyoun@snu.ac.kr

ABSTRACT

Many industries are making efforts to minimize the losses caused by shutdown of manufacturing facilities and to set an optimal maintenance schedule. In this context, prognostics, which predict remaining useful life (RUL) based on information extracted from sensory signals, have attracted attention. There are three methods to perform life prediction: physics-based, data-based, and hybrid. However, data-driven methods are the only way to apply them to a complex industrial facility. By assuming multiple degradation unit data, we can extract various features from the data and select the best feature to create a health index(HI). In this study, we propose a new method for the feature selection step that greatly determines the performance of RUL prediction. Proposed algorithm can automatically select features that are monotonic and have a consistent level of value in normal and failure zone. We validate our method using real degradation data acquired from bearing life testbed.

1. INTRODUCTION

Prognostics and Health Management (PHM) is a field of engineering that monitors the status of a mechanical system using sensors and captures the signs of failure to establish a condition-based maintenance strategy. It is important to predict the remaining useful life (RUL) of a system to determine the optimal maintenance time. However, the degradation characteristics of mechanical systems are non-linear, so it is difficult to predict accurately the RUL. In recent years, many studies use machine learning techniques to carry out RUL prediction since machine learning techniques are capable of learning relationship between multi-dimensional health feature and system's health state (Si, Wang, Hu, & Zhou, 2011). There is a lot of caution when using machine learning techniques, but the most important thing is not to significantly increase the dimensionality of the data. As the dimension of data increases, the size of the space to analyze increases exponentially. This is a much bigger problem in the PHM field since there are not many data to learn the degradation pattern in a real system. Therefore, it is important to reduce the dimension of the data by using either feature selection or transformation technique. However, it is more practical to use a feature selection method because dimension can be reduced without losing physical meaning (Yang, Liao, Meng & Lee, 2011).

Various indicators have been developed to select favorable feature(s) for regression analysis. Recently, Hu, Che, Zhang, Zhang, Guo and Yu (2012) developed Rank Mutual Information (RMI) to select feature which has strong monotonicity with predictive variables. Niu, Qian and Choi (2016) apply RMI to select feature(s) that monotonically increase or decrease over time, and use it for RUL prediction. However, the selected feature(s) using RMI did not guarantee the consistency of the health and failure criteria. In this study, we propose a generalized rank mutual information(GRMI) score function that is more sensitive to the feature's healthy and faulty level, by adding additional term into previous score function. Finally, the proposed methodology will be applied to the degradation data obtained from the bearing life test-bed and its performance will be analyzed in a qualitative manner. This paper is organized as follows: in section2, we review the RMI based feature selection and then the proposed GRMI metric in feature selection method. In Section 3, the details of the experimental setup and the feature selection results will be presented. At last, conclusions are summarized in section 4.

2. FEATURE SELECTION SCORE FUNCTION

The feature selection method can be classified into two types. The filter base method is to select the optimal feature by evaluating the performance of each feature with a score function. The wrapper base method is to apply the feature to the learning algorithm to evaluate the performance and select the optimal feature parameter. In this study we focused on score function in filter based method. The following section briefly review conventional RMI based score function with mathematical expression and propose score function called GRMI (Generalized Rank Mutual Information).

2.1. Conventional score function: rank mutual information

Hu et al. (2012) developed a metric called rank mutual information (RMI) to select a monotonic feature in the ordinal classification problem where the target classes are discretely ordered. RMI shows how highly ordered the features are with another feature. Hu et al. (2012) defined Upward Rank Mutual Information(URMI) which refers to degree of monotonic increase as follow:

$$URMI(B,C) = -\frac{1}{n}\sum_{i=1}^{n}\log\frac{\left|[x_i]_B^{\geq}\right| \times \left|[x_i]_C^{\geq}\right|}{n \times \left|[x_i]_B^{\geq} \bigcap [x_i]_C^{\geq}\right|} \quad (1)$$

where B and C refers to specific feature, $[x_i]_a^{\geq}$ refers to a set of elements that monotically increase respect to feature a, and $|\cdot|$ is a operator to evalue number of element in a set.

Hu et al. (2012) also defined Downward Rank Mutual Information(DRMI) which refers to of monotonic decrease as follow:

$$DRMI(B,C) = -\frac{1}{n}\sum_{i=1}^{n}\log\frac{\left|[x_i]_B^{\leq}\right| \times \left|[x_i]_C^{\leq}\right|}{n \times \left|[x_i]_B^{\leq}\right| \bigcap \left|[x_i]_C^{\leq}\right|} \quad (2)$$

where B and C refers to specific feature, $[x_i]_a^{\leq}$ refers to a set of elements that monotically decrease respect to feature a, $|\cdot|$ is a operator to evalue number of element in a set, and details are mentioned in the references (Hu et al., 2012).

If we calculate the RMI between a specific feature and time, we can evaluate the degree of monotonicity of that feature with respect to time. Niu et al. (2016) shows that RUL prediction using features with a high RMI value is more accurate, since that feature has more information about degradation.

Score function of feature i on degradation unit j is defined as follow:

$$score_{RMI}(i,j) = \max(URMI(x_{i,j}, t_{i,j}), DRMI(x_{i,j}, t_{i,j})) \quad (3)$$

where $x_{i,j}$, $t_{i,j}$ refers to value of feature i and corresponding time on degradation path j.

Score function of feature i in all degradation unit is defined as follow:

$$score_{RMI}(i) = \frac{\sum_{j=1}^{N} score_{RMI}(i,j)}{N} \quad (4)$$

where N is total number of degradation unit.

For illustration, we generate artificial degradation data with two feature as Figure 1. We can say that feature A has larger monotonicity over time than feature B in both unit 1 a nd 2. Likewise, RMI of feature A is slightly larger than RMI of feature B as shown in Table 1.

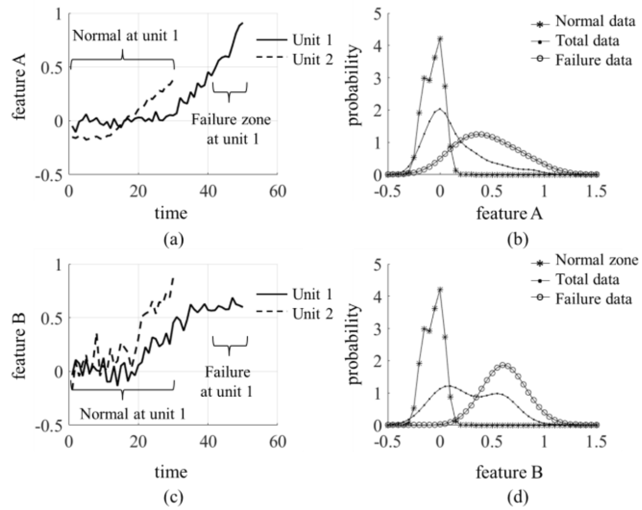

Figure 1. Artificial degradation data with two features (a) time trend of feature A (b) normal to failure histogram of feature A (c) time trend of feature B (d) normal to failure histogram of feature B

Table 1. Score function of feature A and B in artificially generated degradation data

feature	$score_{RMI}$	$score_{GRMI}$
A	0.88	1.16
B	0.73	1.23

2.2. Proposed score function: generalized rank mutual information

Considering the problem of predicting the RUL of the current unit based on the historical degradation unit, the failure (or normal) data is important, since they give us an information about failure (or normal) threshold value. If the failure (or normal) zone value of various unit has little deviation with compare to the total deviation on degradation data, we can define more accurate threshold value and RUL prediction result. However, traditional RMI metric cannot guarantee above property. So, we proposed generalized RMI score of feature i as follows:

$$dist = \frac{\sigma_{i,normal}{}^{2} + \sigma_{i,failure}{}^{2}}{\sigma_{i,total}{}^{2}} \tag{5}$$

$$score_{GRMI}(i) = score_{RMI}(i) + \exp(-dist) \tag{6}$$

where $\sigma_{i,normal}$ refers to standard deviation of feature *i* at normal zone on all degradation unit, $\sigma_{i,failure}$ refers to standard deviation of feature *i* at failure zone on all degradation unit and $\sigma_{i,total}$ refers to standard deviation of feature *i* on all degradation unit. Variable *dist* means the normalized distance between normal and failure zone data distribution on all degradation unit. In Figure 1, the failure data distribution of feature A is difficult to distinguish from the total data distribution. However, the distribution of failure data of feature B is relatively easy to discern from that of total data. In the same manner, consistency score of feature B is larger than that of feature A, and we can say that the feature B is much more consistence in failure threshold.

3. RESULT AND ANALYSIS

3.1. Dataset description

We performed 3 acceleration life tests on the rolling element bearing tester and acquire acceleration signal with DAQ system. Each set of degradation data is divided into two stage, normal and degradation stage as shown in figure 2. The failure was defined as the moment when the moving average of vibration RMS exceeded failure threshold. We obtained 21 filtered signals through various signal processing combinations (e.g. frequency filter, Hilbert transform), and extracted total 357 features by evaluating time and frequency features for each signal.

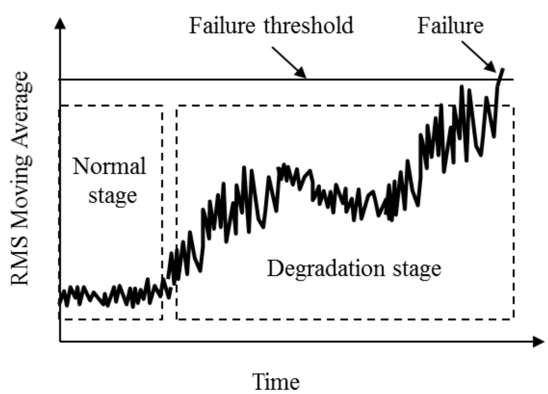

Figure 2. Illustration of normal stage, degradation stage and failure definition of experiment data.

3.2. Feature selection result

For each feature extracted from the bearing life test data, RMI and GRMI score function were calculated. To illustrate the results, we select four features, e.g. 189, 169, 119, 13. Time trend of that features over three run-to-failure test are plotted is Figure 3. In Figure 3, the magnitude of the data is scaled between 0 and 1 for convenience. Also score functions are presented in Table 2.

As we expected, the features with high RMI score, e.g. feature 223, 169, tend to show strong monotonicity over time. Note that feature 223 and 169 both have a high RMI value, whereas feature 223 has a constant value at normal and failure level across test 1 through 3, whereas feature 169 has significantly different values at failure level across test 1 through 3. These results can be seen from GRMI score of feature 223 and 169, because feature 223 has larger consistency level than feature 169. Therefore, we have to select a feature considering GRMI score like Table 2. In this case, we can select an optimal feature such as feature 223.

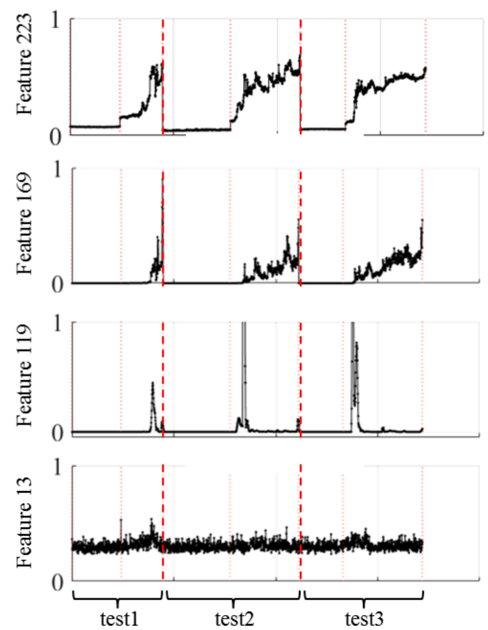

Figure 3. Three bearing run-to-failure data (test1-3) time trend of 4 features with rescaled magnitude

Table 2. Score metrics of specific features in real bearing run-to-failure data

Feature number	$score_{RMI}$	$score_{GRMI}$
223	0.752	1.454
169	0.782	1.051
119	0.393	1.268
13	0.174	0.312

4. CONCLUSION

In this paper, we studied about feature selection metric to select best feature for prognostic purpose. Based on RMI, which represented monotonicity over time, we added failure (or normal) threshold consistency based term. We evaluated both metric in case of artificial degradation data and real bearing degradation data. And we confirmed that the better features can be selected if GRMI score are considered. With GRMI score function we can choose a feature with high monotonicity over time and high similar failure (or normal) threshold. Future study will quantitatively analyze that the features selected considering both metrics further increase the accuracy of RUL prediction.

ACKNOWLEDGEMENT

This work was supported by the Technology Innovation Program (10050980, System Level Reliability Assessment and Improvement for New Growth Power Industry Equipment) funded by the Ministry of Trade, Industry & Energy (MI, Korea).

REFERENCES

Si, X. S., Wang, W., Hu, C. H., & Zhou, D. H. (2011). Remaining useful life estimation–A review on the statistical data driven approaches. *European journal of operational research*, vol. 213(1), pp. 1-14. doi:10.1016/j.ejor.2010.11.018

Yang, Y., Liao, Y., Meng, G., & Lee, J. (2011). A hybrid feature selection scheme for unsupervised learning and its application in bearing fault diagnosis. *Expert Systems with Applications*, vol. 38(9), pp. 11311-11320. doi: 10.1016/j.eswa.2011.02.181

Hu, Q., Che, X., Zhang, L., Zhang, D., Guo, M., & Yu, D. (2012). Rank entropy-based decision trees for monotonic classification. *IEEE Transactions on Knowledge and Data Engineering*, vol. 24(11), pp. 2052-2064. doi: 10.1109/TKDE.2011.149

Niu, G., Qian, F., & Choi, B. K. (2016). Bearing life prognosis based on monotonic feature selection and similarity modeling. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 230(18), pp.3183-3193. doi: 10.1177/0954406215608892

BIOGRAPHIES

Taewan Hwang received the B.S degree from Seoul National University (SNU), Seoul, Korea, in 2013. He is a Ph.D. student at the Department of Mechanical and Aerospace Engineering in SNU. His current research area is prognostics and health management for bearing. He received a PHM Society Data Challenge Competition Winner award (2014).

Keunsu Kim received his B.S. degree from Seoul National University, Seoul, Republic of Korea, in 2013. He is a Ph.D. student at the Department of Mechanical and Aerospace Engineering in Seoul National University. His research area includes prognostics and health management for machine tools. He received an award including the IEEE PHM Data Challenge Competition Winner (2014) and the PHM Society Data Challenge Competition Winner (2015).

Su J. Kim This Su J. Kim received the B.S degree from Inha University, Incheon, South Korea, in 2015. He is a Ph.D. student at the Department of Mechanical and Aerospace Engineering in Seoul National University (SNU). His current research area is failure diagnosis, prognostics and health management (PHM).

Byungjoo Jeon received the B.S. degree from Seoul National University (SNU), Seoul, Korea, in 2015. He is a M.S. student at the Department of Mechanical and Aerospace Engineering in SNU. His current research area is prognostics and health management for bearing.

Byeng D. Youn received the B.S. degree from Inha University, Incheon, South Korea, in 1996, the M.S. degree from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 1998, and the Ph.D. degree from the University of Iowa, Iowa City, IA, USA, in 2001. He is currently an Associate Professor of mechanical and aerospace engineering with Seoul National University, Seoul, South Korea. Dr. Youn has garnered substantive peer recognition resulting in notable awards, including the ASME IDETC Best Paper Awards in 2001 and 2008, respectively, the ISSMO/Springer Prize for a Young Scientist in 2005, the IEEE PHM Competition Winner in 2014, the PHM Society Data Challenge Competition Winner in 2014 and 2015, respectively, etc.