

センサモデルと投票に基づく異常検知

菊地啓† 近藤和樹† 堀田政二† 渋谷久恵†† 前田俊二††

† 東京農工大学大学院 情報工学専攻 ††(株) 日立製作所 生産技術研究所

1 はじめに

社会インフラ事業における設備の安定稼働のため、産業用機器の劣化に伴う異常予兆を速やかに検知することが求められている。予兆の検知によって劣化した機器を故障が起こる前に交換または修理することが可能となり、設備の安定性の向上につながる。本稿ではこの様な課題に対し、複数のセンサデータを用いて故障の前に機器の異常を検知するための手法を提案する。

設備内の機器の稼働状況を把握するために提案手法では機器に装着してある複数のセンサに着目する。これらのセンサは冷却水の温度や油圧といった機器の状態を出力する。これらのセンサを用いて異常の予測を行うにあたり重要なことは予兆が起きた際にその原因を検知結果から読み取ることである。香川ら [1] は機器に装着したセンサ出力値を用い、機器の状況を監視するシステムを提案しているが、センサ出力は一つのベクトルに統合されてしまうため、どのセンサが異常の原因となっているかが不明瞭である。また、この手法は故障を検出する手法であり、故障の予兆を検出するものではない。

本稿では故障の予兆とその原因を得るため手法としてセンサ毎のモデルと投票に基づく異常検知について検討する。本手法ではまず、各センサの正常時のモデルをセンサ出力の平均値を用いて作成する。次に入力データとなるセンサ出力値が正常稼働時のモデルからどれだけ外れているかを投票を用いて計算する。この投票により、ある日の各センサの出力が正常な状態と比較してどれだけ異なるかを定量的に評価することが可能となる。

2 提案手法

以下では、複数のセンサの一日毎の平均的な出力値と実際のセンサ出力を比較することで、機器の異常を予測する手法について述べる。

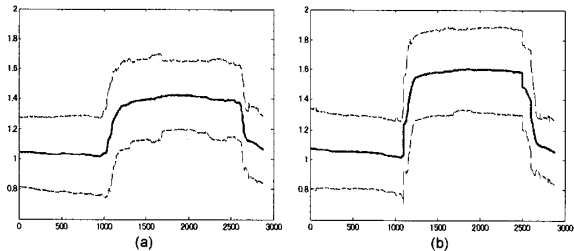


図 1: (a) 月曜日のセンサモデル. (b) 日曜日のセンサモデル. 実線が平均値, 破線は平均土標準偏差を表す.

2.1 曜日毎センサモデルの作成

各センサの出力値は一日毎に 2880 個の区間で記録されている。本稿で用いるデータはある設備内の機器から得られたセンサ出力値となっており、センサの数が 26 個、記録期間が 519 日となっている。調査の結果、センサ出力値は曜日毎に平均的な形状が異なることが分かったため、日曜日から月曜日までのそれぞれの曜日に分けて 2880 個の各区間での平均と標準偏差を計算する。作成した平均と標準偏差を合わせて本稿ではセンサモデルと呼ぶ。あるセンサから算出した月曜日と日曜日のセンサモデルを図 1 に示す。

2.2 センサモデルを用いた投票による各日の異常度の算出

作成したセンサモデルを用いて各日の異常度を算出するために、入力データと入力データと同じ曜日のセンサモデルとの比較を行う。具体的には $I_k^{day}(j, t)$ を第 j 日目 (day 曜日) におけるセンサ k の区間 t の出力値とし、 $M_k^{day}(t)$ を入力データと同じ曜日の区間 t におけるセンサ k の出力値の平均値、 $S_k^{day}(t)$ をそれに対応する標準偏差とする。ここで t は整数であり、 $1 \leq t \leq 2880$ の値をとる。この時、センサ k の第 j 日の異常度 $D(k, j)$ を次式で評価する。

$$D(k, j) = \sum_{t=1}^{2880} \delta(|M_k^{day}(t) - I_k^{day}(j, t)| > n S_k^{day}(t)) \quad (1)$$

ここで、 $\delta(\cdot)$ は (\cdot) が真であれば 1 を、そうでなければ 0 を返す関数である。また、 n は平均と平均土標準偏差との間の距離を制御するパラメータであり、実験では $n = 2$ を用いた。 $D(k, j)$ は第 j 日の入力データと入力データと同じ曜日のセンサモデルとを比較し、

Anomaly Detection Using Sensor Models and Voting

† Kei KIKUCHI

† Kazuki KONDO

† Seiji HOTTA (s-hotta@cc.tuat.ac.jp)

†† Hisae SHIBUYA

†† Shunji MAEDA

Department of Computer and Information Sciences, Graduate School of Tokyo University of Agriculture and Technology (†)

2-24-16 Naka-cho, Koganei-shi, Tokyo 184-8588, Japan

Production Engineering Research Laboratory, Hitachi, Ltd.

(††)

入力データの出力値がセンサモデルの平均土 $n \times$ 標準偏差外に存在した回数, すなわち異常とみなした投票回数を表している.

2.3 投票数の累積値に基づく異常予測

次に, 各日の異常度を利用して異常予測を行う方法について述べる. 式 (1) を用いると第 j 日にセンサ k が平均的な出力値と比較してどの程度異なった値を出力しているかを知ることが出来る. この $D(k, j)$ が著しく大きい値をとった場合は機器になんらかの異常が起きていると予想できる. したがって, $D(k, j)$ がある閾値を超えたら, 第 j 日は異常度が高いとみなし, 危険を知らせる予測アラートと呼ばれる信号を出力する. 本稿では, 正常稼働日における $D(k, j)$ の平均値を閾値とし, 5つのセンサが閾値を超えたら予測アラートを出力するよう設定した.

3 実験

提案手法がどの程度正確に異常を予測できるかを検証するため, 実際のセンサデータを用いて異常を予測する実験を行った. 実験に用いたデータセットにはセンサデータの他に, 実際に機器に異常が発生した場合に出力されるアラート (図 2 上段) が含まれている.

提案手法の予測精度を測るため, 実験では *detect* と *purity* という 2 種類の尺度を定義した. これらの算出方法を示す.

$$detect = \frac{\text{実際の異常を予測できた件数}}{\text{実際に異常が発生した件数}} \times 100 \quad (2)$$

$$purity = \left(1 - \frac{\text{誤って出力した予測アラート数}}{\text{出力した予測アラートの総数}}\right) \times 100 \quad (3)$$

detect は実際の異常をどの程度予測できたかを示し, *purity* は予測アラートの精度を示している. すなわち, 両方の値が高いほど, 予測精度が高いことを示す. 本稿において, 実際のアラートが出力された日から 5 日前までに予測アラートが出力された場合, 実際のアラートは正しく予測できたとみなし, それ以外の予測アラートは誤りとした. 予測精度の推定には, 5 fold-cross validation を利用した. すなわち, データの区間を均等に 5 つに分割した後, その内の 1 区間をテスト用のデータとし, 残りの 4 区間をモデル作成用のデータとして利用した. なお, 常に一定の値を出力し続けているようなセンサは異常予測に適切でないと判断し, モデル作成の際に利用しなかった.

実験の結果, *detect* の平均値が 79.6%, *purity* の平均値が 58.7% となった. *purity* が *detect* に比べて低い値となっており, 異常が発生する前には予測アラートを出力しているが, 正常に稼働している区間でも予

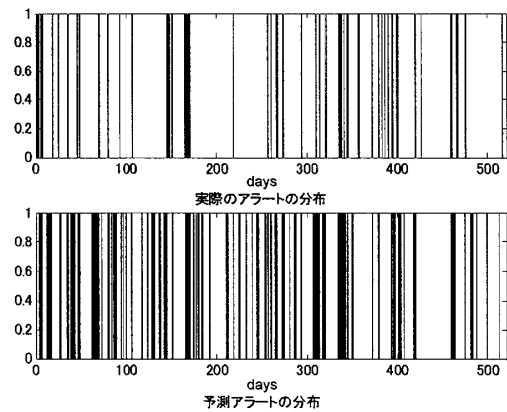


図 2: 実際のアラートと予測アラートの分布.

測アラートを出力していることが分かる. ちなみに, 同じ方式でリアルタイムな予測に改めた場合, *detect* が 80.1%, *purity* が 38.9% となった.

予測アラートの出力頻度を確認するため, 図 2 に実際のアラートと予測アラートの分布を示す. 横軸がデータの記録期間を表し, 上段が実際のアラートの分布, 下段が予測アラートの分布を表している. 図から, 実際のアラートに対して, 予測アラートが多く出力されていることが分かる.

4 まとめ

本稿では, 産業用機器の劣化に伴う異常予兆を速やかに検知するため, 複数のセンサモデルを用いて, 機器の異常を予測する手法を提案した. 提案手法を用いて実験を行った結果, 実際の異常が発生する前に, 危険を知らせる予測アラートを出力できることを示した. しかし, 正常稼働時にも予測アラートを出力してしまうという問題点があることが分かった. 今後は予測精度の向上のため, センサの選択 [2] やセンサ間の相関 [3] を利用する予定である.

参考文献

- [1] 香川敏昌, 平井雅人, 小山茂則, 豊田佳克 “生産支援のための故障診断システムの開発,” 徳島県立工業技術センター研究報告, vol. 14, pp. 9-14, 2005.
- [2] 久保田和人, 森田千絵, 渡辺経夫 “データマイニングを用いたプラントの知的メンテナンス技術,” 東芝レビュー, vol. 61, No. 21, pp. 8-12, 2006.
- [3] 井出剛 “近傍保存原理による異常検知,” 10th Workshop on Information-Based Induction Sciences, pp. 212-217, 2007.