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

Intelligent Plant Maintenance by Data Mining

久保田　和人　■KUBOTA Kazuto
森田　千絵　■MORITA Chie
渡辺　経夫　■WATANABE Tsuneo

プラント事業は，設備を売るだけの時代から，顧客設備の運用・保守サービスをもメーカーのメニューに加える時代へと変化しつつある。この新しいサービスを実現するためには，実際のプラント運転データを効率良く分析する必要があり，情報処理技術を活用した知的メンテナンス技術が今後ますます重要になってくると考えられる。

東芝は，発電プラントの運転管理における異常発見モデルの生成に，データマイニングの適用を試みた。その結果，その有望性が確認できたので，この知的メンテナンス技術を更に発展させ，サービス事業に展開していく。

The plant business is currently changing from the provision of equipment to the provision of services; namely, subcontracted operational and maintenance work.　In order to carry out these services safely and efficiently, information processing technology to analyze plant data will be a critical requirement.

Toshiba has applied data mining to power plant operation management and achieved promising results.　We will further develop intelligent maintenance technology to promote plant service businesses.

1　まえがき

製造業のプラント事業は，設備の提供だけではなく，運用・保守サービスも一括請け負うサービス，更には設備が生み出す価値の向上も提供する時代へと変化しつつある[(1), (2)]。例えば，発電事業では発電設備を納めるだけではなく，運用及び保守を一括請け負うサービスや，プラントのライフサイクル価値を増大させるサービスをも提供していくという考え方に変わりつつある。また，航空機エンジン事業では，エンジンを納めるという考え方から，推進力を生み出すサービスを提供するという考え方に変わりつつある。このような変化を実現するためのキーとなる技術が，知的メンテナンス技術である。

ここでは，知的メンテナンス技術について整理するとともに，データマイニングを用いたプラントの異常発見についての試みを，技術的な側面から述べる。

2　知的メンテナンス技術

サービスを安定的かつ継続的に提供するためには，メンテナンス技術が重要な役割を果たす。タイムリーなメンテナンスの実施あるいはメンテナンス戦略の策定には，プラントの運転状況を把握し，将来の故障や性能劣化を予測できる高度な情報処理が必要となる。このとき活用可能な情報は，プラントの運転履歴データや設計データ，運転員のノウハウなどが挙げられる。“知的メンテナンス”は，これらの情報を総合的に分析し，この結果に基づいてメンテナンスを実施する。これにより，プラント運転時の安全性の向上やメンテナンスコストの削減が図れるほか，経営戦略との融合によりプラントのライフサイクル価値を増大させることが可能になる(**図1**)。

知的メンテナンス技術は，その構成要素ごとに様々な技術的課題を持つ。運転履歴データの利用に関しては，センシング技術や取得されたデータの分析技術が挙げられる。設計データの利用に関しては，物理モデルの生成やこれを利用した監視のための特徴量の生成，シミュレーション技術などが挙げられる。また，運転員の知識の利用に関しては，暗黙知の形式知化や知識のデータベース化が挙げられる。この中でも，非常に多くのセンサから長時間蓄積された膨大な

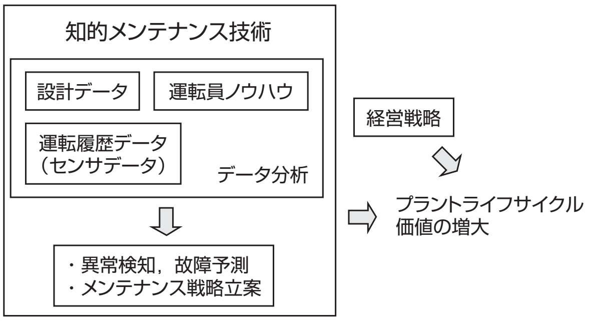

図1．知的メンテナンス技術と波及効果 ― 様々な情報を総合的に分析することで異常検知や故障予測に役だてられるほか，その結果をプラントライフサイクル価値の増大へと波及させることが可能となる。

Intelligent maintenance and its effects

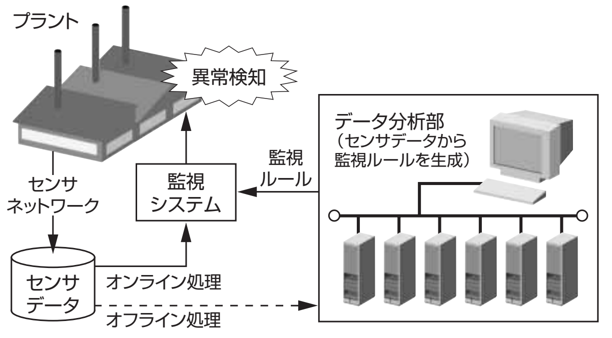

図2．知的メンテナンスを実施するための環境 ― センサデータはデータ分析部で分析され，生成された監視ルールはプラントの監視システムへと組み込まれる。

Environment required for intelligent maintenance

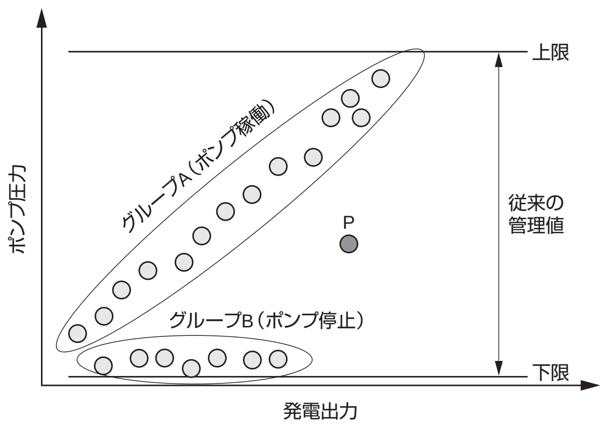

図3．発電所における圧力値と発電出力の関係 ― ポンプが稼働するか否かでセンサ間の関係が異なり，二つのグループが存在する。P点は異常であるが，正常とみなされてしまう。

Pump pressure of two operation groups vs. power output in power plant

データの履歴を分析し，プラントの異常やその予兆を発見する技術は，メンテナンスを実施するうえでもっとも基礎的かつ重要な技術と位置づけることができる。

運転履歴データの一つであるセンサデータを用いた知的メンテナンスを実施するためのシステム構成を**図2**に示す。取得されたセンサデータはデータ分析部へと送られ，オフライン処理で監視ルールが生成される。監視ルールはプラントの監視システムへと組み込まれ，異常の検知が行われる。知的メンテナンス技術の一例として，データマイニングを利用して，発電プラントのセンサデータの監視ルールである異常発見モデルを生成する試みについて，以下に述べる。

3　従来のセンサ値の異常発見手法の課題

センサ値の異常を監視するためのルールは，通常，センサの上限と下限を用いて設定されることが多い。しかしこの方法では，次に示す二つの問題が生じる。

3.1　適正範囲の設定の困難さ・煩雑さに対する課題

センサ値の適正範囲は，エンジニアがプラントの系統図やシステムの物理的特性を考慮し，実際にセンサから取得されたデータと突き合わせながら設定される。しかし，大規模なプラントではセンサの数は数万点にも及び，その一つ一つについて専門家が適正範囲を設定することは事実上不可能になりつつある。

したがって，適正範囲設定のための何らかの技術的サポートが必要となる。

3.2　未検出異常の存在に対する課題

センサがとりうる値は，そのセンサが設置された設備の稼働状況に大きく依存し，上限と下限を規定するだけでは異常を見逃してしまう可能性がある。その具体例を以下に示す。

図3は発電出力とポンプ圧力の関係をプロットしたもので，ポンプ圧力の異常を検出することを目的としている。図中の上下限が適正範囲を示している。昨今の大規模な発電所では複数のポンプが設置されている場合があり，発電出力が小さいとポンプは稼働している場合としていない場合の二つの状態をとりうる。グループＡが稼働状態の集合，グループＢが停止状態の集合に相当する。ここで，Ｐという時刻の点を考える。上下限で判定する方法では，圧力の値が従来の管理値の間に入っているため正常とみなされる。しかし，この点はグループＡからもＢからも離れており，このような圧力値と発電出力の値の組合せを取ることはまれで，異常である可能性が高いと考えられる。

したがって，このような異常を検出できるような監視ルール，すなわち異常発見モデルを生成する仕組みが必要になる。

4　センサ値の組合せを考慮した異常発見方式

3章のような課題を解決する手法として，複数のセンサ値の組合せからプラントの異常性を判定する方法，及びそのための監視ルールの生成手法について述べる。組合せによる異常判定では，あるセンサ（以下，ターゲットセンサと呼ぶ）の値をほかの一つ以上のセンサ（以下，モデルセンサと呼ぶ）から推定するモデルを生成し，推定モデルと実際の値のかい離度から異常性を判定する。ゴールとなる異常発見モデルを**図4**に示す。図3の二つのグループのデータについて，それぞれ推定モデルＡ，Ｂと，この例ではデータのばらつきから計算された3σ値である管理幅が設定される。この管理幅から外れたデータは異常性が高いと判断される。

3.1節で述べたように，すべてのセンサについて手作業でこのようなモデルを作成することは困難である。したがってここでは，データマイニングを用いることでセンサ履歴デー

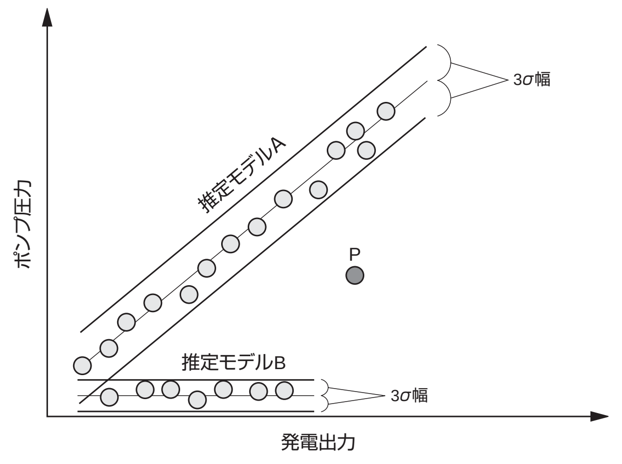

図4．ゴールとなる異常発見モデル及び管理幅 ― グループごとに推定モデルと管理幅を生成することで，高精度に異常を発見することが可能になる。

Fault detection model with three-sigma control individually established for each operation group

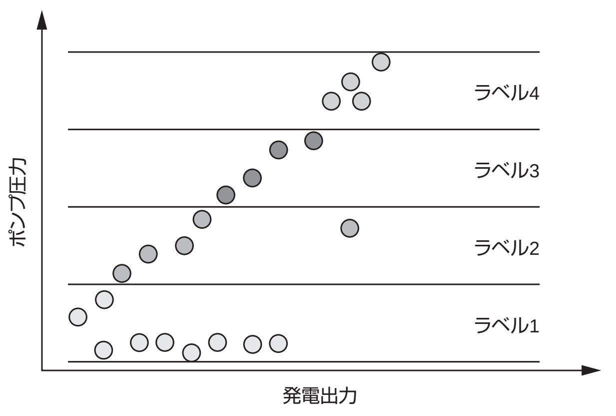

図5．ターゲットセンサの区間分割とラベル付け ― ポンプ圧力の変動範囲を4等分し，それぞれのデータに1から4までのラベルが割り当てられる。

Segmentation and labeling of target sensors

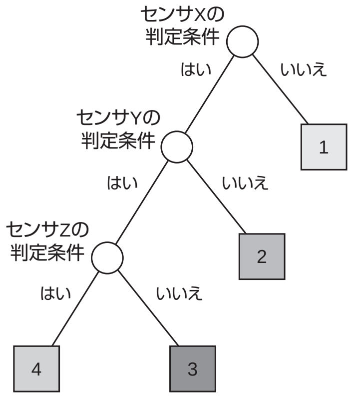

図6．モデルセンサ選択のための決定木生成 ― ポンプ圧力の値をほかのセンサから推定する決定木である。ノードに表れるセンサはモデルセンサの候補となりうる。

Decision tree generation to select model sensors

タからの異常発見モデルを自動的に生成する。そのためには二つの技術が必要となる。まず，多数のセンサの中からモデルセンサとして適切なセンサ（この例では発電出力）を見つける必要がある。次に，モデルセンサを用いて推定モデルを生成する必要がある。このとき，図4のように二つのモデルを生成したほうが精度の高いモデルが生成される場合は，分布を考慮しながら点のグルーピングを行い，二つのモデルを生成しなければならない。

高次元のデータ集合からクラスタ及びクラスタの存在する部分空間を抽出する手法は，一般に部分空間クラスタリング（Subspace Clustering）[(3)]と呼ばれる。ここで述べる手法はこの範ちゅうに含まれるが，既存の手法の多くは点どうしの距離を基準に点をグルーピングしているのに対し，点が同じモデル上に乗るかという基準でグルーピングしている点に特徴がある。

5　決定木を利用したモデルセンサ選択

ターゲットセンサをモデルセンサから推定するためには，両者がプラント動作時に連動する必要がある。相関係数を見れば，ある程度両者が連動することはわかるが，逆に相関係数が低いからといって連動していないとは言えない。図4の例では，グループA，B内でのポンプ圧力と発電出力の相関は高いが，全体の相関係数はそれほど高くない。

ここでは，モデルセンサ選択にデータマイニングの一手法である決定木[(4)]を利用する。決定木は，表形式のデータが存在したとき，ある属性の値をほかの属性から推定するツリー状のルールを生成する手法である。モデルセンサ選択の手順は，まず，**図5**に示すようにポンプ圧力の変動範囲を区間に分割して各区間のデータにラベルを付加し，続いて，このラベルをほかのセンサから推定する決定木を作成する（**図6**）。ここで，決定木に出現するセンサはポンプ圧力と連動する可能性が高いと考え，モデルセンサの候補とする。図6の場合，センサX，Y，Zがモデルセンサの候補となる。なお，この段階ではモデルセンサを一つに絞らず，複数のセンサを候補としてそれぞれ推定モデルを生成し，その中の最良のものを最終的な推定モデルとする。候補を選択して個々に推定モデルを生成するという手順を踏んでいるのは，すべてのセンサに対して推定モデルを生成するのは計算のコストが大きいためである。センサの数が限られている場合は候補選択のステップを省略し，すべてのセンサを候補として次節で述べる異常発見モデルの生成を試みればよい。

6　異常発見モデルの生成手法

異常発見モデルの生成には，重回帰クラスタリング（RC：Regression Clustering）[5]と呼ばれる手法を利用する。通常のクラスタリングでは点どうしの距離の近いものを一つのクラスタへとまとめるが，RCでは空間上の点を回帰モデルへの当てはまりを考えながらクラスタリングする。RCには様々な手法があるが，ここでは，データを再帰的に分割して行う。

RCが実施されるようすを**図7**に示す。縦横のグリッドがクラスタを分割するための分割線の候補である。ここでは，軸と直交する直線を候補とする。分割線の候補の一つsで入力データDを分割して，クラスタD1s，D2sを作成する。続いて，クラスタD1s，D2sについて回帰モデル（この例では線形モデル）を求め，評価値を計算する。これらの操作をすべての分割線の候補について行い，評価が最大となるsを求めてDを分割する。以上の処理を再帰的に行うことでクラスタを生成する。同時に，クラスタ内の点を回帰する線分，すなわち，異常発見モデルを生成する。

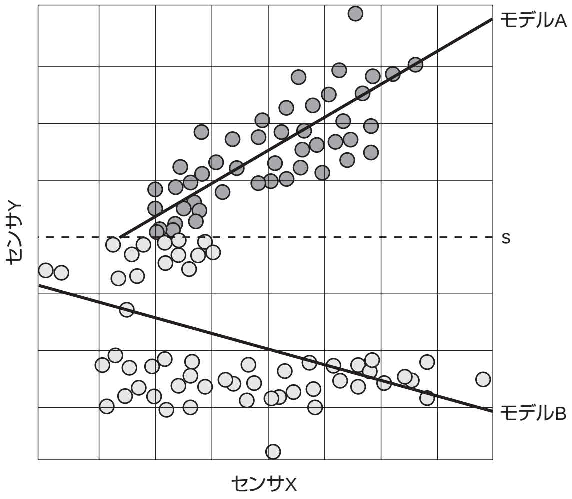

図7．重回帰クラスタリングの実施例 ― 空間を再帰的に分割することでクラスタリングを行っていく。回帰モデルへの当てはまりの良さが分割の際の指標となる。

Regression clustering (RC)

7　発電所データへの適用例

この手法を発電所のセンサデータに対して適用した結果を示す。対象をプラントの特定のブロックに属する約2,500センサとした。これらのセンサは1分刻みで10か月分のデータが取得されており，そのデータ量は膨大である。この中から圧力センサAと振動センサAをターゲットセンサとし，残りのセンサから推定モデルを生成した。

まず，ターゲットのセンサの変動範囲を10区間に分割してラベル付けし，そのラベルを推定する決定木を生成した。圧力センサAの場合は128種類，振動センサAの場合は181種類のセンサが分類に利用された。続いて，これらのモデルセンサの候補の一つと発電出力をモデルセンサとしてRC処理を繰り返し，その中の最良のものを異常発見モデルとした。分割面の候補は，軸と直交し，各センサの変動範囲を100等分するような平面を選んだ。**図8**と**図9**は，圧力セン

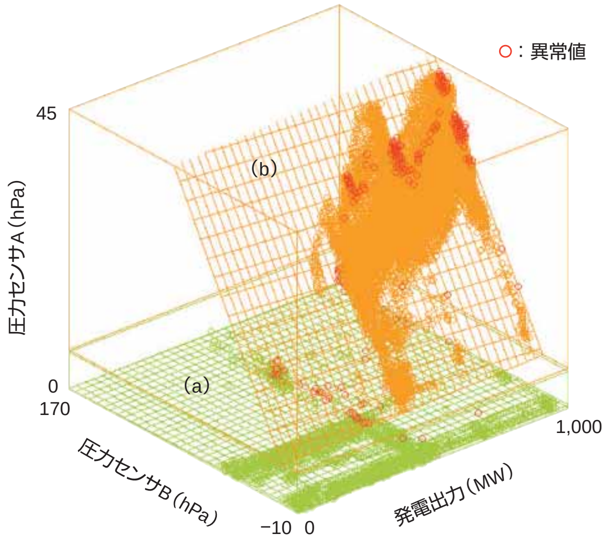

図8．圧力センサAの異常発見モデル ― 二つのモデル（a），（b）が生成されている。（a）は圧力センサAが0の状態で，（b）は圧力センサA，Bが連動する状態である。

Fault detection model for pressure sensor A

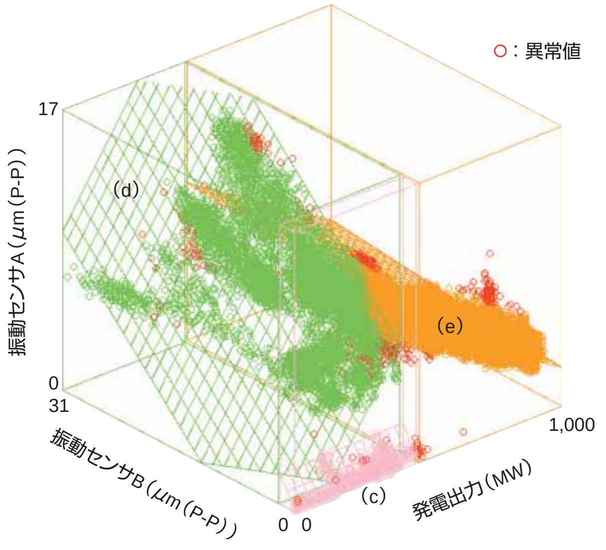

図9．振動センサAの異常発見モデル ― 三つのモデル（c），（d），（e）が生成されている。（c）は振動センサA，Bがほぼ0の状態で，（d）と（e）は発電出力は異なるが，振動センサA，Bが連動する状態である。

Fault detection model for vibration sensor A

サＡと振動センサＡに関して評価値がもっとも高かった異常発見モデルである。

圧力センサＡは，モデルセンサとして発電出力のほかに圧力センサＢが選択された。二つのクラスタが生成され，圧力センサＡが0の状態（図８の（a））と，圧力センサＢと連動する部分（図８の（b））に分けられた。振動センサＡに関しては，振動センサＢが選択された。クラスタは三つ生成され，図９の（c）の部分は両者が0で動いていない部分，（d）と（e）は両者が連動している部分であるが，発電出力の大小によって，その関係が異なっていることがわかる。

赤い部分が異常値である。ここで自動的に抽出されたモデルセンサ及び生成された異常発見モデルは，いずれもエンジニアの直感に合うものであり，データマイニングを用いた異常発見モデル生成の有望性を示すものとなった。実用的には，ここで生成されたモデルとプラント設計時の情報，運転員のノウハウをすり合せ，最終的な異常発見モデルを構築していくことになる。

8　あとがき

ここでは，知的メンテナンス技術について整理するとともに，データマイニングを用いたプラントの異常発見の試みについて，技術的な側面から述べた。知的メンテナンス技術は，製造業が設備だけを売る時代からサービスをも提供する時代へと変化するためのキーとなる技術であり，その基礎を担うのがデータマイニングに代表されるデータ分析技術である。

今後は，ここで述べた異常発見技術に加え，プラント事業における諸問題に対してデータ分析を実施し，製造業のサービス事業への展開をサポートしていく。

文　献

(1) 福田　浩．火力発電プラントのライフサイクルにおけるITソリューション．東芝レビュー．**58**，1，2003，p.32－36．
(2) 下田理文，ほか．上下水道事業における維持管理サービスへの取組みとソリューション展開．東芝レビュー．**61**，5，2006，p.21－24．
(3) J. R. キンラン．AIによるデータ解析．トッパン，1995，293p．
(4) L. Parsons, et al. Subspace Clustering for High Dimensional Data: A Review. SIGKDD Explorations. **6**，1，2004，p.90－105.
(5) B. Zhang. Regression Clustering. Proc of ICDM. 2003, p.451－458.

久保田　和人　KUBOTA Kazuto, D.Eng.
研究開発センター システム技術ラボラトリー研究主務，工博。データマイニング及び高性能計算システムに関する研究・開発に従事。情報処理学会，電子情報通信学会会員。
System Engineering Lab.

森田　千絵　MORITA Chie
研究開発センター システム技術ラボラトリー研究主務。データマイニングシステムに関する研究・開発に従事。人工知能学会，情報処理学会会員。
System Engineering Lab.

渡辺　経夫　WATANABE Tsuneo
産業システム社 府中事業所 情報制御システム部主幹。発電所管理計算機システムの開発に従事。情報処理学会会員。
Fuchu Complex