

# 時系列波形異常検知技術の変電設備診断への適用と改良

山口 晃広<sup>†</sup> 植野 研<sup>†</sup> 内田 和徳<sup>††</sup> 松本 詠治<sup>††</sup> 才田 敏之<sup>††</sup>

<sup>†</sup> 株式会社東芝 研究開発センター システム AI ラボラトリー 〒212-8582 川崎市幸区小向東芝町 1

<sup>††</sup> 東芝エネルギーシステムズ株式会社 〒212-0013 川崎市幸区堀川町 72-34

E-mail: <sup>†</sup>akihiro5.yamaguchi@toshiba.co.jp

**あらまし** 電力会社では高信頼で効率的なインフラ設備の保守運用に向けて IoT 技術の活用が進んでおり、開閉装置などの変電設備では設置されたセンサから時系列データを収集し機械学習を用いて状態監視を行う設備診断技術に注目が集まっている。我々は、正常な時系列データのみを入力として学習し正常な波形パターンからのズレとして異常の根拠を提示できる時系列波形異常検知技術 OCLTS (One-Class Learning Time-Series Shapelets) を開発してきた。本研究では、設備診断向けの時系列波形異常検知技術の課題を明らかにし、変電設備の開閉動作をセンシングしたストローク波形データに OCLTS を適用し僅かな異常兆候でも説明性を持って捉えられることを確認する。更に、設備が正常であっても動作の度に変動し不安定な箇所のある時系列データに対して、正しく異常兆候を捉えられるように OCLTS を改良し、このような特性を持つモーター電流波形データに適用しその有効性を確認する。

**キーワード** 時系列波形異常検知, 説明性, スマートファクトリー, 設備診断, shapelets

## 1 はじめに

近年 IoT の普及に伴い、電力会社では送電・配電に関わるインフラ設備の効率的な維持・運用のために様々なセンサを取り付けて機器を監視するシステムが普及しはじめている [1, 2, 3]. 特に開閉装置を始めとする変電所設備については、その重要な役割、設備量の多さ、設備の老朽化などの観点から、異常の有無を感度よく具体的な障害が生じるより早期に検出し、以降の保守運用や設備維持・更新計画に有効活用できる設備診断技術の確立が求められており、この分野への機械学習技術の適用が期待されている。

その中でも重要なアプリケーションが変電所設備の異常兆候を早期かつ高精度に検出する設備診断技術である。従来の開閉装置では、その開閉動作を時系列データとしてセンシングしたストローク波形から、開閉時間をはじめ特定箇所の動作時間や速度など、専門家が事前に設計した特徴量を用いて、閾値判定やトレンド分析をとおして設備診断を行っていた。一方、近年の機械学習技術の発展とその応用により、より多くの設備情報が含まれると考えられる測定波形自体を対象として分析することが可能となっており、一層高度な設備診断に応用できる可能性がある。とりわけ、今後年々台数の増加が見込まれている高経年設備に対しては、設備診断を行う有効性が高いものの、従来技術で設備診断を行う場合の判断基準の取得が不十分なケースもあり、取得した時系列データから機械学習により特徴量自体を学習し、以降の設備診断を実施できれば大きな有効性を示すものと言える。

我々が開発した One-Class Learning Time-series Shapelets (OCLTS) 技術 [4, 5] は、正常な時系列データのみを入力として学習し、異常検知時には異常と判定した根拠も提示できる。

本研究では、(1) 機械学習の説明性、(2) 異常データの希少性、(3) 時系列データの変動不安定性からなる、設備診断技術の 3 つの重要な課題を明らかにする。次に、そのうち課題 (1)(2) を解決する従来技術 OCLTS を変電所の設備診断に適用し、課題 (3) への対応が不要なストローク波形で有効性を確認する。更に、3 つの課題全てを解決するように OCLTS を改良する。そして、課題 (3) への対応も必要なモーター電流波形を用いて、改良版 OCLTS の設備診断技術における有効性を評価する。

### 1.1 本研究の主な貢献

本論文の主な貢献を以下に示す。なお、本内容は基本的に文献 [6, 7] に準じるが、文献 [6, 7] よりも機械学習技術の部分の詳細に述べる。

- 時系列データを用いた機械学習の設備診断技術において、工場出荷前試験で特に重要となる 3 つの課題を明らかにする。
- そのうち 2 つの課題を解決する従来技術 OCLTS を変電所の設備診断に適用しその有効性を確認する。
- 設備が正常でも変動したり不安定な箇所を持つ時系列データに対応できるように OCLTS を次の 2 点で改良する。
  - (a) 正常な時系列データの不安定な時間的急変部分にロバストな特徴量を導入する。
  - (b) 正常事例からなる学習データの中で時系列データの変動が大きい事例の影響を低減するように学習する。
- 3 つの課題全てが必要となる変電所の設備診断に改良版 OCLTS を適用し有効性を確認する。

## 2 研究背景

本章では、機械学習による時系列データを用いた設備診断の 3 つの課題を明確にする。その後、1 つ目の課題を解決する

shapelets 学習法と 2 つの課題を解決する OCLTS を説明する。

## 2.1 設備診断向け時系列波形異常検知技術の課題

特にインフラ設備を製造工場で出荷前に検査する場合など、時系列データを用いた設備診断に機械学習を適用する際には、単純に精度を向上させるだけでなく、次の 3 つの重要な課題を解決する必要がある。

1 つ目の課題は機械学習の説明性である。変電所開閉装置などのインフラ設備の異常検知や診断では、専門家の知識に基づき従来の特徴量を設計していることから分かるように、専門家は時系列波形と機械的な現象との関係を専門知識として保有している。そのため、専門家の知識と機械学習の判定結果とを照らし合わせて納得・安心して利用したいという要求がある。また、専門家は時系列データを見て異常が発生した原因を究明するため、機械学習が異常と判定した根拠となる波形パターンの位置や崩れ方を知りたいという要求もある。そのため、判定根拠を提示できないブラックボックスな機械学習技術では専門家の理解が得られず導入の障壁となる。

2 つ目の課題は異常データの希少性である。変電所開閉装置など信頼性の高いインフラ設備は殆どの場合正常に稼働しており異常が測定されるのは稀である。また、多数の異常モードを網羅的に把握することは難しく、異常モードによっては発生メカニズムが不明で再現が難しい場合もある。このような課題から、データの収集時に発生しなかった異常も検知する必要がある。異常事例の蓄積を待っていると機械学習技術の導入が遅れてしまう。

3 つ目の課題は時系列データの変動不安定性である。センサの種類によっては動作時の環境条件や計測ノイズなどの影響により、設備としては正常であってもセンサで計測した時系列データは変動を伴う場合や動作の都度変化する不安定な箇所を含む場合がある。例えばバネ蓄勢モーター電流波形のインラッシュ電流など、モーターの動作開始時点の条件により設備が正常であっても動作の都度変化する時間的急変部分が含まれる場合がある。波形の僅かな概形の違いで正常と異常を判別する場合には、このような時系列データの変動不安定性は機械学習の正しい判断を妨げる。

## 2.2 関連研究：Shapelets 学習法

事例数  $I \in \mathbb{N}$  で長さ  $Q \in \mathbb{N}$  の時系列データセット  $\mathbf{T} \in \mathbb{R}^{I \times Q}$  から、 $i = 1, 2, \dots, I$  における  $i$  番目の時系列データ  $\mathbf{t}_i \in \mathbb{R}^Q$  のクラスを機械学習により分類する時系列分類技術において、クラス分類に用いる短い波形パターン (shapelets) を機械学習により特定する技術がデータマイニングや機械学習の分野で研究されている [8, 9]。このような shapelets 学習法は、高い分類性能を達成するだけでなく、その分類器に用いられる特徴量が shapelets と時系列データとの非類似度 (距離) であるため専門家に理解しやすい。専門家は、クラス分類に用いる shapelets を確認し専門知識と照合することで、ねじの緩みなど設備故障に関係する機械的な現象と、直感的に理解できる。

しかしながら、通常の shapelets 学習法 [8, 9, 10] では学習

時に異常事例を収集する必要があり、2 つ目の課題 (異常データの希少性) を解決できない。コスト考慮型学習を導入し正常事例に対して異常事例が少ない場合に対応する shapelets 学習法 [11] は提案されたが、異常事例は学習時にやはり必要となりこの課題を解決できない。

## 2.3 関連研究：時系列波形異常検知技術 OCLTS

我々は説明性と異常データの希少性の両方の課題を解決するため、正常な時系列データセット  $\mathbf{T} \in \mathbb{R}^{I \times Q}$  及び、ハイパーパラメータである shapelets の個数  $K \in \mathbb{N}$  や長さ  $L \in \mathbb{N}$  を与えると、shapelets  $\mathbf{S} \in \mathbb{R}^{K \times L}$  を学習する技術 OCLTS を開発した [4, 5]。時系列分類技術である従来の shapelets 学習法とは異なり、異常検知技術 OCLTS では正常データしか学習時に参照しないため異常波形が時系列データにどのように表れるか学習時には分からない。そのため、OCLTS における shapelets は学習データを構成する典型的な複数の短い波形パターンと定義される。そして、第  $i$  番目の時系列データ  $\mathbf{t}_i \in \mathbb{R}^Q$  の特徴ベクトル  $\mathbf{x}_i \in \mathbb{R}^K$  は、時系列データ  $\mathbf{t}_i$  に対して shapelets  $\mathbf{S}$  を隙間無くかつズレが最小となるように割り当てて計算される。この割り当てによる shapelet  $\mathbf{s}_k$  の  $\mathbf{t}_i$  上の位置集合を  $\mathcal{J}_{i,k}$  とする<sup>1</sup>。ここで、 $\mathcal{J}_{i,k}$  の要素  $j$  は  $[1, Q - L + 1]$  の範囲の整数値をとる。OCLTS の  $k$  番目の特徴量  $x_{i,k} \in \mathbb{R}$  は二乗ユークリッド距離  $d_{i,k,j}$  を介して次で測られる。

$$x_{i,k}(\mathbf{s}_k, \mathbf{t}_i) := \max_{j \in \mathcal{J}_{i,k}} d_{i,k,j}(\mathbf{s}_k, \mathbf{t}_i), \quad (1)$$

$$d_{i,k,j}(\mathbf{s}_k, \mathbf{t}_i) := \frac{1}{L} \sum_{l=1}^L (t_{i,j+l-1} - s_{k,l})^2. \quad (2)$$

この shapelets  $\mathbf{S}$  に基づく特徴ベクトル  $\mathbf{x}_i \in \mathbb{R}^K$  を再生核ヒルベルト空間 (RKHS)  $\mathcal{F}$  へマッピングする関数を  $\phi(\cdot)$  とし、この空間  $\mathcal{F}$  上での内積を  $\langle \cdot, \cdot \rangle$  とし、 $[\cdot]_+$  をヒンジ関数とすると、OCLTS は shapelets  $\mathbf{S}$  と 1 クラス分類器  $\mathbf{w}$  を同時に学習する数理最適化問題として次で定式化される。

$$\begin{aligned} \underset{\substack{\mathbf{S} \in \mathbb{R}^{K \times L} \\ \mathbf{w} \in \mathcal{F}}}{\text{minimize}} \quad & \frac{\langle \mathbf{w}, \mathbf{w} \rangle}{\alpha} + \sum_{i=1}^I [1 - \langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle]_+ \\ & + \sum_{i=1}^I \sum_{k=1}^K \sum_{j \in \mathcal{J}_{i,k}} f(d_{i,k,j}). \end{aligned} \quad (3)$$

ここで、shapelets  $\mathbf{S}$  を固定した場合 (つまり、特徴ベクトル  $\mathbf{x}_i$  を固定した場合)、目的関数の第 1 項と第 2 項は One-class support vector machine (OCSVM) の定式化 [12, 13] と等しい。 $f(\cdot)$  は単調増加関数であり、目的関数の第 3 項の最小化は shapelets とそれらが割り当てられた部分時系列との誤差の最小化に対応しており、これにより shapelets を正常データの典型的な波形パターンに近づける。なお  $\alpha > 0$  は正則化パラメータである。この非凸最適化問題は、内積  $\langle \cdot, \cdot \rangle$  を Radial basis function (RBF) kernel に置き換えてから確率的劣勾配降下法を用いて効率的に計算される [4, 5]。

1: 任意の  $i = 1, 2, \dots, I$  及び  $k = 1, 2, \dots, K$  で  $\mathcal{J}_{i,k}$  は空集合ではない。

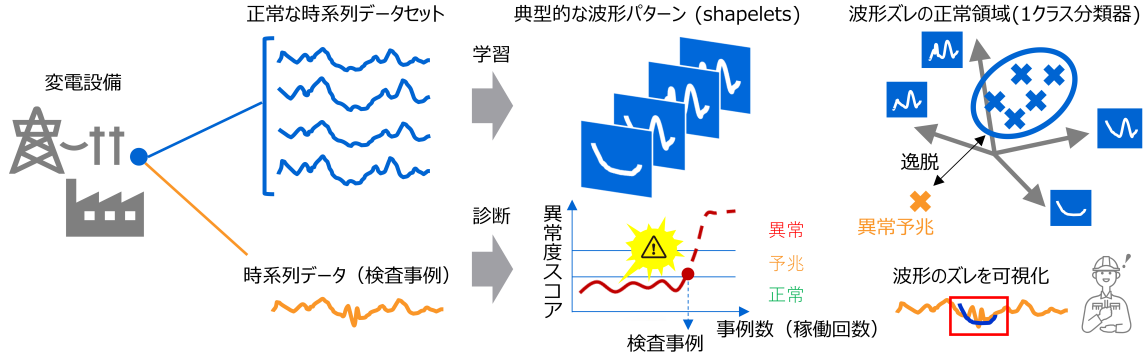


図 1 OCLTS を用いた設備診断技術 (上: 学習時, 下: 運用時)。

しかしながら, 2.1 節で述べた 3 つの課題を解決する従来技術は我々の知る限り無い. OCLTS [4, 5] は 1 つ目の課題 (機械学習の説明性) と 2 つ目の課題 (異常データの希少性) を解決するが 3 つ目の課題を解決しない. 一方, 3 つ目の課題 (時系列データの変動不安定性) に対応する shapelets 学習法 [10] では, 異常事例が学習時に必要となり 2 つ目の課題を解決しない [10]. 更に, OCLTS を含む shapelets 学習法をインフラの設備診断に適用する研究自体も, 我々の研究 [3, 14] 以外に殆どない. 文献 [15, 16] では, 製造装置やインフラ設備の診断に shapelets 学習法を適用するが, 1 つ目の課題のみを対象とし, 残りの課題を解決しない.

そこで, 本研究では OCLTS [4, 5] の設備診断への適用を 3 章で紹介し, 4 章では 2.1 節で前述した 3 つの課題全てを解決するように OCLTS を改良し設備診断へ適用する.

### 3 基本版 OCLTS の設備診断への適用

本章では, 改良前の OCLTS [4, 5] による設備診断を説明した後, 変電設備で収集される変動不安定箇所が少ないストローク波形の場合 (つまり, 3 つ目の課題を解決する必要がない場合) で, OCLTS の異常検知性能と説明性を評価する.

#### 3.1 OCLTS を用いた設備診断技術

OCLTS による設備診断技術の概要を図 1 に示す. 設備が正常であると想定できる時系列データセットを事前に収集し, それを入力データとして OCLTS を式 (3) で学習する. 特徴量が shapelet と時系列データとのズレであることから, 学習された 1 クラス分類器  $\mathbf{w}$  は, 正常な時系列データが shapelets  $\mathbf{S}$  からどの程度ずれて良いかの許容領域を表している. 実際に設備を診断する運用時には, 検査用の時系列データ  $\mathbf{t} \in \mathbb{R}^Q$  を学習済みの OCLTS に入力し, 1 クラス分類器として学習した正常領域からどの程度外れているかを次の異常度スコア  $a(\cdot)$  で測る.

$$a(\mathbf{t}; \mathbf{S}, \mathbf{w}) := \frac{\tilde{a}(\mathbf{t}; \mathbf{S}, \mathbf{w}) - \mu}{\sigma}, \quad (4)$$

$$\tilde{a}(\mathbf{t}; \mathbf{S}, \mathbf{w}) := 1 - \langle \mathbf{w}, \phi(\mathbf{x}) \rangle.$$

ここで,  $\mu$  と  $\sigma$  は学習データにおける  $\tilde{a}(\cdot)$  の平均と標準偏差であり,  $\mathbf{x}$  は時系列データ  $\mathbf{t}$  の shapelets  $\mathbf{S}$  に基づく特徴ベクトルである. 異常度スコアが上昇した際には診断システム [17] はアラートを出し, 専門家は時系列データの shapelets からの波

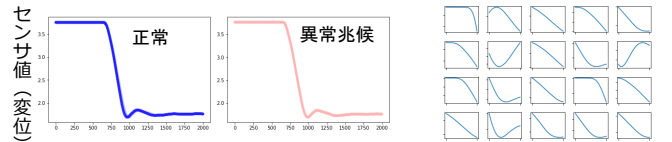


図 2 左: 正常/異常兆候時のストローク波形, 右: 学習した shapelets.

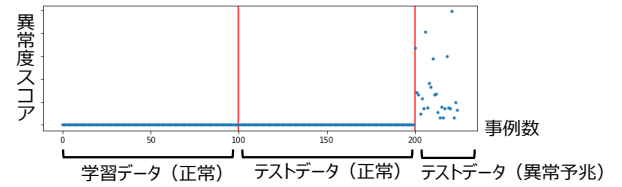


図 3 ストローク波形における正常/異常兆候時の異常度スコア

形の崩れを確認することで異常検知の判定根拠を確認できる.

#### 3.2 ストローク波形を用いた有効性評価

ガス絶縁開閉装置の設備診断に OCLTS を適用する. 対象とする時系列データは変動不安定箇所が少ないストローク波形であり, 1 回分の装置の動作が 1 事例に対応する. まず, 100 回分の正常時系列データを用いて OCLTS を学習する. その後, 引き続き 100 回分の正常時系列データと設備異常の兆候が出現する 30 回分の時系列データとを OCLTS で診断する. 図 2 (左) では, 正常時の時系列データと異常兆候が出現する時系列データとをそれぞれ重ねて描いている. 時系列波形全体から目視で違いを見つけることは難しく, 従来の専門家の知識に基づく特徴量でも明確な判別は難しい. Shapelets の個数  $K$  を 20 個とし, shapelet の長さ  $L$  を時系列データ長  $Q$  の 10% として, shapelets  $\mathbf{S}$  を学習する.

図 2 (右) は OCLTS で学習された 20 個の shapelets であり, ストローク波形を構成する典型的な波形パターンが学習されていることが分かる. 図 3 では, 横軸が動作回数 (つまり事例数) を表し, 縦軸が各事例における式 (4) の異常度スコアを表している. 異常兆候が出現すると異常度スコアが大きく上昇しており, OCLTS が設備の僅かな異常兆候を正確に検知できることを確認できる.

図 4 では, 異常兆候が出現する時系列データの中で最も波形が崩れたもの (異常度スコアが最大の時系列データ) を黒線で描いている. その時系列データの中で shapelets から最も波形

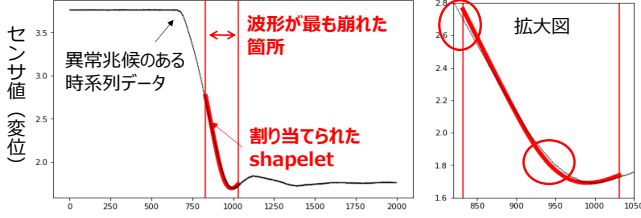


図 4 異常兆候のあるストローク波形における説明性

が乖離した箇所を図 4 (左) の赤線で囲っている。太い曲線は、OCLTS が学習した shapelets の中でその箇所に割り当てられた shapelet を表している。図 4 (右) はその赤線で囲われた部分の拡大図である。拡大することで異常兆候が始めるとストローク波形の傾きが僅かに緩やかになることが分かる。これは部品の不具合により開閉動作が緩慢になるという専門家の知見と一致する。このように、専門家は判定根拠に納得して機械学習を利用することができる。

## 4 OCLTS の改良と設備診断への適用

3 章で述べたように、変電所のガス遮断器におけるストローク波形のように正常時の時系列データに不安定箇所や揺らぎが少ない場合には、基本版 OCLTS [4, 5] は高精度に説明性をもって設備を診断できる。しかし、変電所の断路器におけるバネ蓄勢モーター電流波形のような時系列データでは、設備が正常に稼動していても、動作の度に時間的急変部分が変化したり時間軸上で波形が揺らいだりする。本提案の診断技術を様々な設備診断に拡充していくには、このような変動不安定箇所がある時系列データに対しても正しく診断できる必要がある。そのため、本章では時系列データの変動不安定性の課題に対応するための 2 つの改良を提案し、それらをバネ蓄勢モーター電流波形に適用し有効性を評価する。

### 4.1 正常波形の不安定箇所にロバストにする改良

この改良では、例えば動作の都度変化しスパイク状の波形として現れる時間的急変部分など、正常な時系列データの不安定箇所に対して、OCLTS の特徴量をロバストにする。基本版 OCLTS の特徴量では、式 (2) のように shapelet とそれが割り当てられる部分時系列との二乗ユークリッド距離を用いて、shapelet と時系列データとのズレを測っているため、学習時の時系列データに不安定箇所が含まれると、従来の OCLTS では shapelets がその不安定箇所にひきずられてしまい正常波形からの逸脱が大きくなってしまふ。その結果、異常波形からの逸脱が相対的に減少することで正常データと異常データとの異常度スコアの差が不明瞭となり異常検知の性能が下がる問題がある。この課題を解決するため、改良版 OCLTS では shapelet と時系列データとのズレをロバストな距離で測る。

具体的には、特にスパイク状の不安定箇所にロバストとなるように、shapelet とそれが割り当てられる部分時系列との間で大きく外れた点をスキップさせる。Shapelet の長さ  $L$  の中でスキップする点数  $r \in [0, L-1]$  を与えて、集合  $\{(t_{i,j+l-1} - s_{k,l})^2\}_{l=1}^L$

の中で値が大きい上位  $r$  個の位置  $l$  の集合を  $\mathcal{R}$  とすると、式 (2) の代わりに次の式を用いる。

$$d_{i,k,j}(\mathbf{s}_k, \mathbf{t}_i) := \frac{1}{L} \sum_{l=1}^L \mathbb{I}(l \notin \mathcal{R})(t_{i,j+l-1} - s_{k,l})^2. \quad (5)$$

ここで、インジケータ関数  $\mathbb{I}(\cdot)$  は引数が真のとき 1 を返し偽のとき 0 を返す。スキップする点数  $r$  を 0 と設定すると、基本版 OCLTS の式 (2) と等しくなる。このロバストな距離は、異常波形がスパイク波形として現れる場合には異常を見逃してしまうため不適ではあるが、波形パターンの概形の違いにより異常を検知したい場合には有効であり、図 7 や図 8 で後述するようにスキップした点を可視化することで直感的にも理解しやすい。式 (1) における shapelet に基づく特徴量を式 (5) を介して計算することで、特徴量が不安定箇所の影響を受けにくくなりその結果、学習される shapelets や 1 クラス分類器がその不安定箇所にロバストになる。

### 4.2 変動が大きい学習データの悪影響を低減する改良

この改良では、学習データの信頼度を自動算出しながら、環境条件や計測ノイズなどの影響で信頼性の低い学習データの悪影響を受けにくくするように OCLTS を学習する。基本版 OCLTS では、確率的劣勾配降下法という学習アルゴリズムに基づき、学習データから一様にランダムに事例を選び shapelets や 1 クラス分類器を更新するため、低信頼な事例が偶然選ばれてしまうと性能が劣化する。このような課題への対策として Self-Paced Learning (SPL) という学習アルゴリズムがある [18]。人間や動物の学習過程と同様に、SPL では学習データの各事例の信頼度を自動的に算出しながら、まず簡単に信頼できる事例を優先的に選んで学習し徐々に変動が大きく信頼性の低い事例を含めて学習する。我々の先行研究 [10] ではクラス分類技術の shapelets 学習法 [9] に SPL を導入したが、本研究では異常検知技術の OCLTS [4, 5] に SPL を導入する。

正常な時系列データセット  $\mathbf{T} \in \mathbb{R}^{I \times Q}$  に対して、信頼度  $\mathbf{v} \in [0, 1]^I$  を付与し、式 (3) の数理最適化問題に組み込み、動的かつ自動的に信頼度  $\mathbf{v}$  を学習する。式 (3) とは異なり、1 クラス分類損失や shapelets とのずれが大きい時系列データに対して信頼度を減らすように数理最適化問題の目的関数を定式化する。つまり、次のように shapelets  $\mathbf{S}$  と 1 クラス分類器  $\mathbf{w}$  と信頼度  $\mathbf{v}$  とを同時に最適化する。

$$\begin{aligned} \underset{\mathbf{S} \in \mathbb{R}^{K \times L}, \mathbf{w} \in \mathcal{F}, \mathbf{v} \in [0, 1]^I}{\text{minimize}} \quad & \frac{\|\mathbf{w}\|_2^2}{\alpha} + \sum_{i=1}^I (v_i g_i(\mathbf{S}, \mathbf{w}) + \lambda h(v_i)), \\ g_i(\mathbf{S}, \mathbf{w}) := & [1 - \langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle]_+ + \sum_{k=1}^K \sum_{j \in \mathcal{J}_{i,k}} f(d_{i,k,j}). \end{aligned} \quad (6)$$

ここで、SPL パラメータ  $\lambda > 0$  は信頼度の正則化パラメータであり、SPL 関数  $f(\cdot)$  は狭義単調減少関数<sup>2</sup>である。時系列データ  $\mathbf{t}_i$  の信頼度  $v_i$  は 0 のとき最も低く 1 のとき最も高い。

2: つまり、関数  $f(\cdot)$  では  $0 \leq v < \bar{v} \leq 1$  であれば  $f(v) > f(\bar{v})$  が成り立つ。



式 (6) において, SPL パラメータ  $\lambda$  が増加すると信頼度  $v_i$  の最適値も上昇するので, より変動の大きい時系列データが学習過程に含まれるようになるため, 学習の進行状況に応じて SPL パラメータ  $\lambda$  を徐々に増加させる. これにより, 特に学習過程の初期段階で変動の大きい学習データの悪影響を避けることができる. SPL 関数  $f(\cdot)$  の定義や SPL パラメータ  $\lambda$  の更新は, 先行研究 [10] と同様である. 式 (6) おける shapelets  $\mathbf{S}$  に基づく特徴ベクトル  $\mathbf{x}_i$  は, 4.1 節の改良も含めると式 (5) を介して式 (1) で定義される.

式 (6) の解法は, 先行研究 [10] と同様に交互最適化アルゴリズムに基づく. 最適化する変数を 2 つのブロックに分割し, 片方のブロックに含まれる変数を固定してもう片方のブロックに含まれる変数を最適化し, このプロセスを交互に繰り返すことで全変数を学習する. 各ブロックの最適化で目的関数の値が減少していくことで, この交互最適化は局所的最適解に収束する. 片方のブロックには学習データの信頼度  $\mathbf{v}$  が含まれ, 残りのブロックには shapelets  $\mathbf{S}$  と 1 クラス分類器  $\mathbf{w}$  が含まれる. 先行研究 [10] と同様の導出方法で信頼度  $v_i$  の最適値は閉形式で求まり<sup>3</sup>, shapelets  $\mathbf{S}$  と 1 クラス分類器  $\mathbf{w}$  の最適化は OCLTS [4, 5] と同様に確率的劣勾配降下法により求まる.

### 4.3 改良版 OCLTS のアルゴリズム

4.1 節と 4.2 節の両方を実施した改良版 OCLTS のアルゴリズムを Algorithm 1 に示す. 外側の  $M$  回のループは交互最適化の繰り返しに対応し, 内側の  $\tilde{M}$  回のループは勾配降下法の繰り返しに対応する. 1 行目では, 学習データの信頼度を  $\mathbf{v} = (1, 1, \dots, 1)$  と初期化し, 変数  $\mathbf{S}$  と  $\mathbf{w}$  については基本版 OCLTS [4, 5] と同じであり,  $\mathbf{X}$  は式 (5) を介して式 (1) で計算される. 7 行目では, 基本版 OCLTS [4, 5] と同じように, OCSVM に基づく 1 クラス分類器の更新には Pegasos [13] のカーネル版アルゴリズム (文献 [13] の Fig. 3) を用いる.

### 4.4 モーター電流波形を用いた設備診断

GSRK-300FP という機種のパネ蓄勢モーター電流波形を用いた変電所設備診断に改良版 OCLTS を適用することでその有効性を評価する. ここでは, モーターの周囲温度が 20 度の場合を正常としてその温度が 40 度の場合を異常と模擬して実験する. 図 5 は時系列データを正常時と異常模擬時でそれぞれ重ねて描いている. 正常時であってもモーター起動時の条件で都度変化するインラッシュと呼ばれるスパイク状の波形が含まれる. 波形の概形によって正常と異常を判別する場合, このスパイク波形は正常と異常の判別に寄与せず不安定箇所となる. また, 図 2 のストローク波形と比べて正常時の横軸方向の揺らぎも大きい. 事前に収集した全事例の数は正常と異常模擬とでそれぞれ 50 個である. 正常な時系列データセットのうち 20 個を学習データに使用して, 残りの 80 個で診断性能を検証する.

基本版と改良版の OCLTS で診断性能を比較する. 全ての技術で, shapelet の長さ  $L$  を時系列データ長  $Q$  の 2% の 80 点と設定し, shapelets の個数を 60 個と設定する. 改良版 OCLTS に

#### Algorithm 1 改良版 OCLTS のアルゴリズム

**Input:** 正常な時系列データセット:  $\mathbf{T} \in \mathbb{R}^{I \times Q}$ ; Shapelet の長さ:  $L \in \mathbb{N}$ ; Shapelets の個数:  $K \in \mathbb{N}$ ; スキップする点数:  $r \in [0, L - 1]$ ; 交互最適化の繰り返し回数:  $M \in \mathbb{N}$ ; 勾配降下法の繰り返し回数:  $\tilde{M} \in \mathbb{N}$

**Output:** Shapelets:  $\mathbf{S} \in \mathbb{R}^{K \times L}$ ; 1 クラス分類器:  $\mathbf{w} \in \mathcal{F}$

- 1: 各変数 ( $\mathbf{S}, \mathbf{w}, \mathbf{v}, \mathbf{X}$ ) を初期化する.
- 2: **for**  $m = 1, 2, \dots, M$  **do**
- 3:   SPL パラメータ  $\lambda$  を設定する.
- 4:   **for**  $\tilde{m} = 1, 2, \dots, \tilde{M}$  **do**
- 5:     **for**  $i = 1, 2, \dots, I$  **do**
- 6:       **if**  $v_i > 0$  **then**
- 7:          Shapelets  $\mathbf{S}$  と 1 クラス分類器  $\mathbf{w}$  を更新する.
- 8:          特徴ベクトル  $\mathbf{x}_i$  を式 (5) を介して式 (1) に設定する.
- 9:       **end if**
- 10:    **end for**
- 11:   **end for**
- 12:   各  $i = 1, 2, \dots, I$  に対して  $v_i$  を最適値に設定する.
- 13: **end for**
- 14: **return**  $\mathbf{S}, \mathbf{w}$

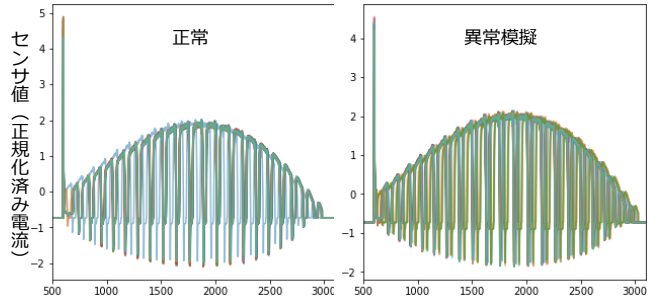


図 5 正常／異常模擬時のパネ蓄勢モーター電流波形.

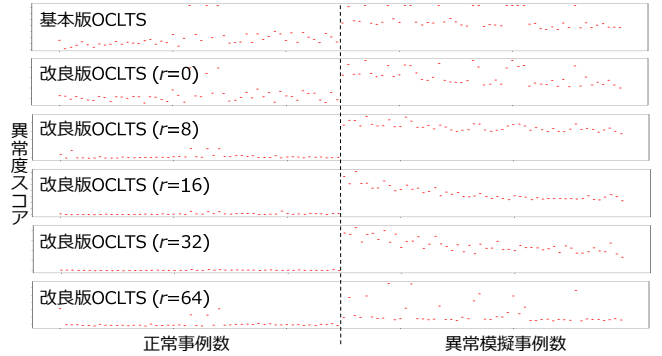


図 6 モーター電流波形における正常／異常模擬時の異常度スコア.

における外側と内側ループの繰り返し回数を  $M = 30$  と  $\tilde{M} = 20$  と設定し, 基本版の繰り返し回数を 600 回と設定する. 改良版の OCLTS では, スキップする点数  $r$  を 0, 8, 16, 32, 64 点まで変えて実験した. 特に  $r = 0$  の場合, 4.1 節の改良を実施せず, 4.2 節の改良のみを実施する場合に相当する. 異常検知の定量的な性能を正常時から異常模擬時への異常度スコアの跳ね上がりで評価する. 異常度スコアの跳ね上がりは, テストデータにおける異常度スコアの平均値を異常模擬時と正常時とでそれぞれ算出し, それらの平均値の差として計算される. この跳ね上

3: この閉形式の詳細な導出は文献 [10] の Proposition 3 を参照されたい.

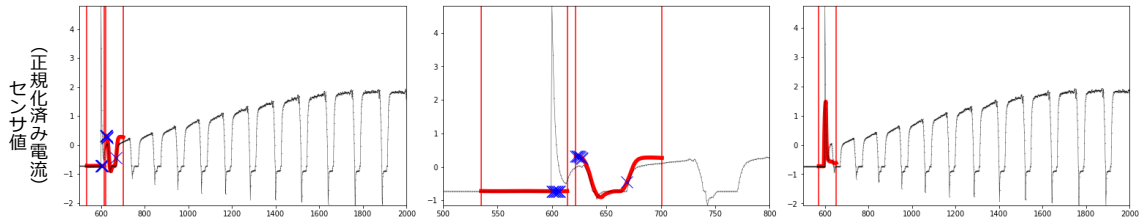


図 7 不安定箇所における shapelets (左：改良版 OCLTS, 中央：左の拡大図, 右：基本版 OCLTS)。

がりが高いほど、異常検知の性能が良いことを意味する。

図 6 は基本版と改良版の OCLTS を設備診断に適用した場合のテストデータの異常度スコアを表している。異常度スコアの跳ね上がりは、基本版 OCLTS で 3.3 であり、改良版 OCLTS ではスキップする点数  $r = 0, 8, 16, 32, 64$  でそれぞれ 8.2, 16.9, 85.8, 126.7, 2.7 である。これより、改良版 OCLTS におけるスキップする点数  $r$  が極端に大きく設定されない限りは、基本版 OCLTS に比べて全ての改良版 OCLTS の方が異常度スコアの跳ね上がりが高く異常検知性能が良いことが分かる。また、4.1 節と 4.2 節の両方の改良を実施した場合に性能が最も向上することも分かる。特に  $r = 32$  のとき、異常度スコアの跳ね上がりが基本版 OCLTS の場合と比べて 38 倍向上した。また、基本版 OCLTS では正常であるにも関わらず異常度スコアが上がってしまう事例が 3 つあったが、改良版 OCLTS では 1 つも無くなっているのが分かる。以降の評価では、異常度スコアの跳ね上がりが最良であった  $r = 32$  の場合で改良版 OCLTS を評価する。

図 7 では、4.1 節で述べたロバストな距離の導入による有効性を shapelets を可視化して評価する。左は改良版 OCLTS の結果であり、中央はその拡大図であり、右は基本版 OCLTS の結果である。黒線は、正常であるにも関わらず基本版 OCLTS で異常度スコアが上昇してしまった時系列データである。赤線は、その時系列データから最も波形が逸脱した shapelets である。青色の×印は、改良版 OCLTS で不安定箇所とみなしスキップした点である。改良版 OCLTS では、期待通りスパイク波形を不安定箇所として自動的にスキップすることで学習した shapelets は時系列波形の概形を表しており不安定箇所の影響を受けていない。一方、基本版 OCLTS ではスパイク波形に無理にフィッティングするように shapelet を学習してしまい、そこで波形がズレてしまっている。これにより、基本版 OCLTS では正常にも関わらず異常度スコアが上昇してしまったことが分かる。

図 8 では、改良版 OCLTS で異常度スコアが最大となった時系列データの波形の崩れを確認する。正常データで学習した shapelets と最も波形が逸脱した箇所を赤色で囲み、図の左側ではその部分を拡大している。これより、異常模擬の時系列データでは該当箇所の概形の高さが正常時よりも高くなることが分かる。この箇所は専門家が温度変化による波形の変形に対して注視している箇所と一致するため、専門家はこの判定根拠に納得して改良版 OCLTS を利用できる。本節の実験評価によ

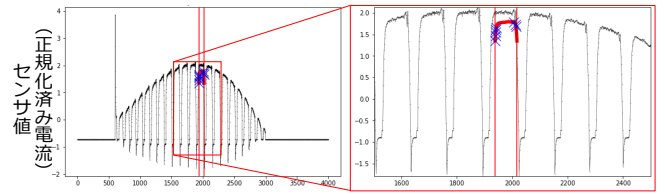


図 8 異常模擬モーター電流波形における説明性。

り、改良版 OCLTS の有効性を確認した。

## 5 ま と め

本研究では、インフラ設備から収集される時系列データを用いて機械学習により異常検知や設備診断を実施する際に、機械学習の説明性、異常データの希少性、時系列データの変動不安定性という 3 つの重要な課題があることを明確にし、それらを解決する時系列波形異常検知技術を提案した。また、変電所の開閉装置から得られる時系列データを用いて本技術による設備診断の有効性を確認した。本技術は特に、設備診断の必要性が高い既設・高経年機器への適用に親和性が高く、将来の効率的で高度な開閉装置の保守運用の実現に非常に有効と考えられる。今後の研究の一つの方向性としては、設備が正常であっても周囲の温度や圧力が変化する場合に波形が変化するため、shapelets の変形を予測する shapelets 学習法 [19] と統合したより高度な設備診断の確立があげられる。

## 文 献

- [1] Y. Matsushita, H. Chida, S. Noguchi, and K. Sasamori. "Practical use of monitoring/diagnostic systems to realize more efficient maintenance, stable electric power supply and optimized refurbishment". In: *CIGRE Paris*. CIGRE, 2014.
- [2] K. Uchida, Y. Abe, T. Koike, Y. Shimogouchi, K. Baba, and M. Nakahata. "GIS On-line Condition Monitoring Solution Using Real-Time Data Collected from GIS-Mounted Sensors". In: *CIGRE - AORC Technical Meeting*. CIGRE, 2020.
- [3] E. Matsumoto, K. Uchida, M. Saito, A. Yamaguchi, and T. Maekawa. "Recent Digitization of GIS and Sophistication of Equipment Condition Monitoring and Diagnosis

- applying AI Technologies”. In: *CIGRE Paris*. CIGRE, 2022.
- [4] A. Yamaguchi and T. Nishikawa. “One-Class Learning Time-Series Shapelets”. In: *Big Data*. IEEE, 2018.
- [5] A. Yamaguchi and T. Nishikawa. “One-Class Learning Time-Series Shapelets”. In: *International Journal of Data Mining Science* (2019).
- [6] A. Yamaguchi, K. Ueno, K. Uchida, E. Matsumoto, and T. Saida. “Development of advanced AI technologies for condition diagnosis of high voltage switchgear in substations”. In: *CIGRE Kyoto Symposium*. CIGRE, 2022.
- [7] A. Yamaguchi, K. Ueno, K. Uchida, E. Matsumoto, and T. Saida. “Development of advanced AI technologies for condition diagnosis of high voltage switchgear in substations”. In: *CIGRE Science and Engineering* (2022).
- [8] L. Ye and E. Keogh. “Time Series Shapelets: A New Primitive for Data Mining”. In: *KDD*. ACM, 2009.
- [9] J. Grabocka, N. Schilling, M. Wistuba, and L. Schmidt-Thieme. “Learning Time-series Shapelets”. In: *KDD*. ACM, 2014.
- [10] A. Yamaguchi, S. Maya, and K. Ueno. “RLTS: Robust Learning Time-series Shapelets”. In: *ECML PKDD*. Springer, 2020.
- [11] S. Roychoudhury, M. Ghalwash, and Z. Obradovic. “Cost Sensitive Time-Series Classification”. In: *ECML PKDD*. Springer, 2017.
- [12] B. Schölkopf, J. C. Platt, J. C. Shawe-Taylor, A. J. Smola, and R. C. Williamson. “Estimating the Support of a High-Dimensional Distribution”. In: *Neural Comput.* (2001).
- [13] S. Shalev-Shwartz, Y. Singer, N. Srebro, and A. Cotter. “Pegasos: Primal Estimated Sub-gradient Solver for SVM”. In: *Math. Program.* (2011).
- [14] 山口晃広. “AI 判断の根拠を説明する XAI を使いこなす : 4. Shapelets 学習によるインフラ・製造分野向け時系列波形の異常診断技術 -異常の検知や診断に有効な波形パターンを発見する AI-”. In: *情報処理* (2022).
- [15] O. P. Patri, A. B. Sharma, H. Chen, G. Jiang, A. V. Panangadan, and V. K. Prasanna. “Extracting discriminative shapelets from heterogeneous sensor data”. In: *Big Data*. IEEE, 2014.
- [16] O. P. Patri, A. V. Panangadan, C. Chelms, R. G. McKee, and V. K. Prasanna. “Predicting failures from oil-field sensor data using time series shapelets”. In: *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2014.
- [17] 内田和徳, 森裕哉, 松本詠治, 山口晃広, and 植野研. “開閉装置用 AI 診断システムの開発”. In: *令和 4 年電気学会全国大会*. 2022.
- [18] M. P. Kumar, B. Packer, and D. Koller. “Self-Paced Learning for Latent Variable Models”. In: *International Conference on Neural Information Processing Systems*. Curran Associates Inc., 2010.
- [19] A. Yamaguchi, K. Ueno, and H. Kashima. “Learning Evolvable Time-series Shapelets”. In: *ICDE*. IEEE, 2022.