

# 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

Time-Series Waveform Anomaly Diagnostic Methods Utilizing Learning Shapelets for Infrastructure and Manufacturing Fields

## 誤検出・見逃しのリスク、及び学習データの信頼度を考慮して 判定根拠を提示できる、時系列波形の異常診断技術を開発

インフラ・製造分野では、IoT(Internet of Things)の普及に伴って時系列波形データが大量に蓄積されており、AIを用いた波形診断のニーズが高まっています。これらの分野にAIを適用するには、専門家に判定根拠を提示できることだけでなく、誤検出のリスクや学習データの信頼度を考慮できることが求められます。

そこで東芝は、判定根拠の提示を必須要件とした上で、代表的な波形パターン(Shapelets)学習をベースに、誤検出・見逃しのリスクに応じた学習技術LTSpAUC(Learning Time-series Shapelets for Optimizing Partial AUC(Area Under the ROC(Receiver Operating Characteristic) Curve)), 及びデータの信頼度に応じた学習技術RLTS(Robust Learning Time-series Shapelets)を開発しました。

### 背景

近年、IoTの普及に伴い、インフラ設備や生産現場の製造装置などの状態監視をするために、センサーを取り付けて時系列波形データを収集できる環境が整ってきました。これに伴って、メンテナンスが必要な設備や装置の状況を自動で検知・診断する、AIを用いた時系列波形の診断技術が注目されています。

インフラ・製造分野にAIを適用するには、判定性能を向上させるだけでなく、①判定根拠の提示、②誤検出のリスク管理、③データの信頼度に応じた学習、も求められます。①は、専門家に異常の判定根拠となる波形部分やその崩れ方を提示できないと、原因究明が難しくなります。②は、対象製品・サービスのコストや危機管理の性質に応じて、誤検出と見逃しのどちらを重要視するかが決まります。③は、人手により入力した異常と正常の間違いやノイズが混入した波形データなど、信頼度の低い学習データによる影響を抑える必要があります。しかし従来技術では、これらの要求に対応できませんでした。

東芝は、①に応える時系列波形の診断技術であるShapelets学習をベースに、これを更に拡張した独自の技術を開発しました<sup>(1), (2)</sup>。ここでは、まず①に応える既存のShapelets学習について述べ、次に①と②に応えるLTSpAUC、及び①と③に応えるRLTSについて説明します。

### 判定根拠を提示できる時系列波形の診断技術

インフラ・製造分野では、専門家が時系列波形の専門知識を保有しているため、専門家の知識とAIの判定結果を照合してAIを納得・安心して利用したいという要求や、異常の原因究明の際にAIが異常と判定した根拠となる部分波形の位置や崩れ方を知りたいという要求があります。このような高い説明性は、変電所開閉装置の異常予兆検知や、エレベーターの保守、工場の製造装置の異常検知などで、強く求められます。

Shapelets学習は、診断に有効な特徴が短い波形パターンに現れることに着目し、そのShapeletsを機械学習で抽出する技術です。図1のように、正常時や異常時に収集された時系列波形データを入力すると、正常と異常を見分けるために有効なShapeletsと、正常と異常を判別する判別モデルを同時に学習します。そして、正常か異常か不明な検査対象の時系列波形データを判別モデルに入力すると、判定根拠となるShapeletsを提示するとともに、正常/異常を判定します。専門家は、判別に有効なShapeletsを確認し専門知識と照合することで、AIの判定根拠をねじの緩みなど故障に関係する機械的な現象と、直感的に理解できます。

### 誤検出・見逃しのリスクに応じた学習技術LTSpAUC

診断対象の製品やサービスなどによって、AIの誤検出・見逃しのコストは大きく異なります。例えば、半導体製造装

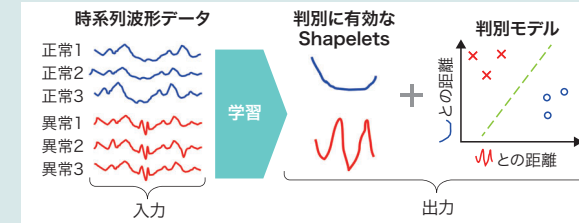


図1. 判定根拠を提示できる波形診断手法

正常や異常といった診断結果を判別するだけでなく、診断に有効なShapeletsも提示できます。

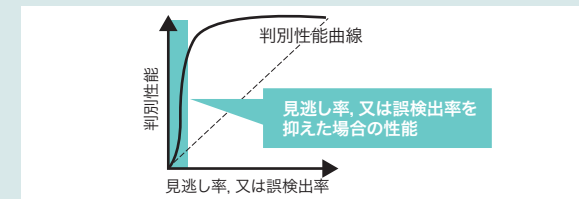


図2. 異常の見逃し率、又は誤検出率を一定に抑えた場合の判別性能

LTSpAUCでは、見逃し率、又は誤検出率の割合を一定に抑えて、判別性能を最大化するように学習できます。

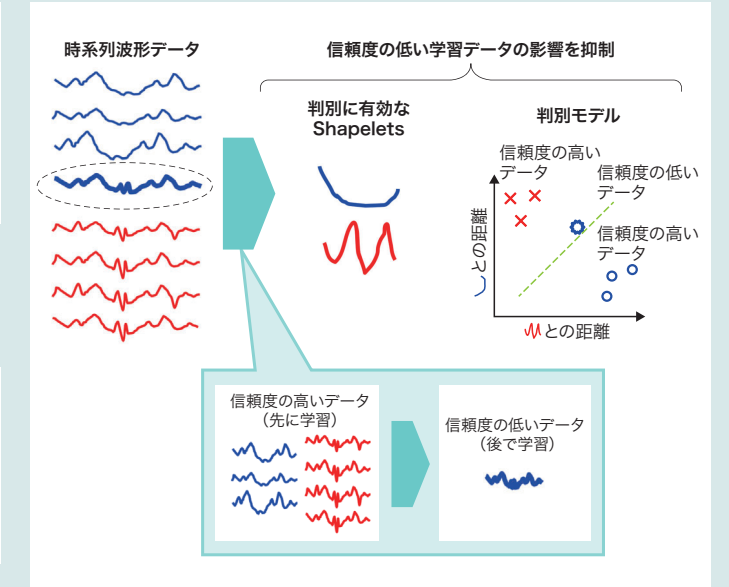


図3. 学習データの信頼度を考慮できる波形診断手法

RLTSでは、学習データの信頼度を学習しながらその信頼度に基づいて学習を進めることで、信頼度の低い学習データによる影響を抑えます。

置で装置が停止するような異常の見逃しはリスクが大きいため、見逃しを低く抑えた上で誤検出を低減することが必要です。一方、インフラ保守で保守員が現場に赴いて部品を分解・調査すべき異常かどうかを事前に判断するような予知保全は、誤検出のコストが大きいため、誤検出を低く抑えた上で見逃しを低減することが必要です。このように、サービスや製品のコスト構造に応じた適切な誤検出・見逃しの低減が重要となります。

そこで図2のように、見逃し率、誤検出率のいずれか一方の範囲をユーザーが指定すると、その範囲内でもう一方を最適化して判別性能を最大化するようにShapeletsと判別モデルを同時に学習できる、LTSpAUCを開発しました。性能検証実験で、製造・インフラ保守に関する公開データなどを用いて、その有効性を確認しました。また、見逃し率を低く抑えると、まれにしか発生しない異常も検知できました。

### データの信頼度に応じた学習技術RLTS

インフラ・製造分野では、過去にほかの用途で設置したセンサーからデータを取得して異常検知を行うことがあり、センサー性能が不十分な場合があります。また、部品交換の履歴を基にAIに正常と異常を学習させる場合、予防的に部品交換を行うと異常であっても波形データにその兆候が現れないことがあります。そのため、信頼度の低い学習データが含まれていても、信頼度を自動的に学習し診断性能を悪化させない技術が求められます。

そこで図3のように、学習データの信頼度を考慮したロバストな学習アルゴリズム(SPL: Self-Paced Learning)をShapelets学習に導入した、RLTSを開発しました。SPLは、学習データの信頼度を学習しながら、その信頼度に基づいて学習を進める手法です。人間や動物の学習過程のように、まず信頼度の高いデータで学習し、徐々に信頼度が低く診断が難しいデータで学習することで、信頼度の低い学習データによる影響を抑える効果があります。RLTSをダイカストマシンや変電所開閉装置の異常検知に適用し、その有効性を確認しました。

### 今後の展望

今後、変電所開閉装置などに開発した技術の適用を進めるとともに、更なる新技術を創出していきます。

### 文献

- (1) Yamaguchi, A. et al. LTSpAUC: Learning Time-series Shapelets for Partial AUC Maximization. Big Data. 2020, 8, 5, p.391-411.
- (2) Yamaguchi, A. et al. "RLTS: Robust Learning Time-series Shapelets". European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2020). Virtual Conference, 2020-09, ECML PKDD. 2020, p.595-611.

### 山口 晃広

研究開発センター 知能化システム研究所 システムAIラボラトリー 博士(情報科学)