

令和 6 年 6 月 19 日現在

機関番号：32644
研究種目：基盤研究(C) (一般)
研究期間：2020～2023
課題番号：20K11832
研究課題名(和文) IoTセンサーデータ解析のための時系列パターンの時間発展とセグメンテーション

研究課題名(英文) Temporal Evolution and Segmentation of Time Series Patterns for IoT Sensor Data Analysis

研究代表者
今村 誠 (Imamura, Makoto)

東海大学・情報通信学部・教授

研究者番号：30780291

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：IoTとビッグデータの普及に伴い、工場/公共設備の予知保全から人の行動分析/健康管理に至るまで、センサーデータの解析ニーズが高まっている。しかし、従来の周波数解析や時系列解析は定常性を仮定するのでセンサーデータを扱えない場合があった。そこで、ハイパーパラメータとして従来の「時系列の時間幅」ではなく「値の変動幅」とする点をオリジナリティとする新たな時系列分解手法Spikeletを提案した。そして、センサーデータの構造を発見する「文法的モチーフ」と、劣化や故障などの時間変化を検出する「モチーフの時間発展」を開発した。この成果は、トップレベル国際会議のACM-KDDやIEEE-ICDMに採録された。

研究成果の学術的意義や社会的意義
四次産業革命やSociety 5.0を実現するために、IoT(Internet of Things)普及に伴って飛躍的に増加するセンサーデータを分析する基盤技術を確立する

研究成果の概要(英文)：With the proliferation of IoT and big data, there is an increasing need for sensor data analysis, ranging from predictive maintenance in factories and public facilities to human behavior analysis and health management. However, traditional frequency analysis and time series analysis assume stationarity, making them often unsuitable for sensor data. To address this, we propose a novel time series decomposition method called Spikelet, which uses "the range of fluctuation of time series values" as a hyperparameter instead of "the time width of the time series." Additionally, we developed "grammatical motifs" to discover the structure of sensor data and "temporal evolution of motifs" to detect changes such as degradation or failure. This work has been accepted at top international conferences, including ACM-KDD and IEEE-ICDM.

研究分野：情報学 マルチメディア・データベース ビッグデータ分析・活用

キーワード：時系列データマイニング Time series motifs Symbolic Representation DTW Motifs Predicative Maintenance

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

IoT で全ての人とモノがつながり、様々な知識や情報が共有され、今までにない新たな価値を生み出す **Society5.0** が第 5 期科学技術基本計画として提唱された。**Society5.0** 実現に必要な中核技術の一つが、IoT の急速な普及により生じる機器や人のセンサーデータの解析である。応用範囲は、「工場（半導体・化学・自動車など）や公共設備（交通・電力・水道・ビルなど）の予知保全」から「車やスマートフォンに内蔵されたセンサを用いた人の行動分析・健康管理」に至るまで多岐にわたる。従来の物理的な振動・波動や経済や制御向けの時系列解析は、主に定常時系列を対象としていたが、工場の機器などのセンサーデータは、図 1 に示すように、一つの工程中に、(a)位置決め、(b)吹付、(c)洗浄などの複数のレジュームからなる非定常データなので、定常を仮定するフーリエ変換や時系列解析ではうまく解析できない場合が多いことが知られていた。

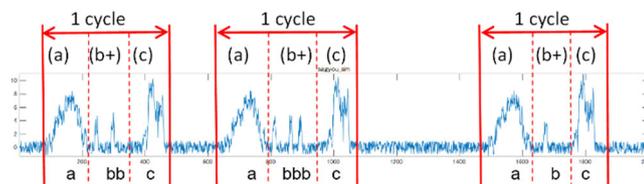


図1: 工場機器のセンサーデータの部分構造の例

2. 研究の目的

(1) 時系列の表現方法の確立

我々は、非定常時系列であるセンサーデータを解析するための要素技術として、定型的に反復出現する未知のパターンを発見するモチーフ発見に注目した。しかし従来のモチーフ発見技術は、ハイパーパラメータとして固定長の窓サイズを指定する必要があるため、一つの工程中に異なる窓サイズをもつパターンが混在する時系列を扱うことが難しいという課題があった。そこで、本研究では、モチーフ発見の重要な未解決問題である「モチーフの部分列の長さを推定する問題（部分列長さの推定）」を解決することにより、「長さの異なるモチーフを一度に発見する技術を確立すること」を第一の目的とした。

(2) 文法的モチーフとモチーフの時間発展

ついで、機器の予知保全や人の動作解析へ応用するために、この部分列長推定技術をベースとして、センサーデータの構造を発見するための「文法的モチーフ発見」と、劣化や故障などのセンサーデータの時間変化を検出するための「モチーフの時間発展」技術を開発することを第二の目的とした。

ここで、文法的なモチーフとは、モチーフ部分列を単語と見立て、モチーフの並びにより構成される句や文に相当する上位のパターンを指す。文法的モチーフの発見技術は、時系列を構成するレジュームを同定したり、レジュームを特徴づけるモチーフパターンを発見するのに役立つ。

また、モチーフの時間発展とは、モチーフが少しずつ変化し最終的には異なったパターンになるモチーフ部分列の系列を指す。モチーフの時間発展を分析することにより、機器の劣化のプロセスや、故障にいたる状態変化を検出するのに役立つ。

3. 研究の方法

(1) 時系列の表現方法の確立

従来の時系列解析では、ハイパーパラメータとして窓サイズを与えて、時系列の特徴量を抽出していた。この方式では、部分列長を推定するには、いろいろな部分列長で特徴量を抽出した上で、一番よい部分列を選ぶというグリッドサーチを用いていた。しかし、この方式では、時系列に対して一様に一つの部分列長を選択することになるので、時系列の場所毎に異なる部分列長を扱うことができない。また、異なった部分列長間の関係がわからないので、「部分列の部分構造を扱うことができない」という課題がある。そこで、「時系列の時間幅」をハイパーパラメータにするのではなく、「時系列の値の変動幅」をハイパーパラメータとする時系列の解析方法を検討することにした。言い換えると、「時系列がどの程度の幅でどの程度の時間幅で変動しているかを解析する」というアイデアである。このアイデアに基づき、新たな時系列の表現方法として、「時系列の値の変動幅をハイパーパラメータとして、与えた値以上の変動幅をもつスパイク形状をもれなく抽出する Spikelet 分解方式」を提案した。

(2) 文法的モチーフとモチーフの時間発展

文法的モチーフに対しては、Spikelet 分解方式によって得られたスパイク系列に対して、スパイクに文字を対応させることにより、時系列を文字列で表現する。そして、文字列のパターンの構造を抽出することにより、文法的モチーフを発見していく。例えば、図 1 に示すように、まず、時系列を `abbcabbbabc` という文字列パターンに変換し、次いで、`(ab+c)` が反復している構造を抽出する。このとき、まず、文字 `a, b, c` に対応する部分列がそれぞれモチーフとして同定され、ちで、その上位構造である `(ab+c)` もモチーフとして同定される。

(2) モチーフの時間発展

モチーフの時間発展に対しては、まず、モチーフ部分列の候補を列挙し、ついで、抽出した部分列の特徴量が規則性をもって変動しているような部分列のシーケンスを発見するアプローチをとる。そして、この方式の検討にあたっては、現実の機器のセンサーデータや動物(昆虫, 人間を含む)の動作データを対象として、機器の劣化や動物の反復動作を分析することにより、モチーフ部分列における有効な特徴量を選択していく。

4. 研究成果

(1) 時系列の表現方法の確立

モチーフの部分列長さを推定するための基盤技術として、時系列の上下変動幅を適応的に解析することにより、オーバーラップを許容したスパイク(凸形状)の系列に時系列を分解する表現方式(Spikelet 分解)と呼ぶ新たな時系列の表現を提案した[1]。Spikelet 分解を用いることで、時間方向の伸び縮みを許容する DTW(Dynamic Time Warping)距離に近い部分列が存在する部分列長を推定し、長さの異なる部分列の対をモチーフとして抽出できる。また、時系列の文字列表現を用いることで、モチーフ発見を劇的に高速化することもできる。結果として、従来技術は同じ長さの部分列のモチーフしか抽出できないのに対して、提案技術により長さの異なるモチーフを高速に発見できるようになった。本開発成果は、データマイニングの最難関会議である IEEE の ICDM(International Conference on Data Mining)に採録されるなど、国際的に高く評価されている。

[1] Imamura, M. and Nakamura, T., "An Adaptive Symbolic Approximation for Finding Higher-Level Structure in Time Series," IEEE ICDM, pp. 1120-1125 (2021).

(2) 文法的モチーフとモチーフの時間発展

Spikelet 分解を用いることにより時系列を文字列に変換し、DTW 距離としての部分列の類似性を文字列としての類似性として近似する文法的なモチーフ技術[2]を提案した。さらに、モチーフの時間発展の精度を向上される技術[3]も開発した。これらの成果も、ACM の KDD(Knowledge Discovery and Data Mining)に採録されるなど、国際的に高く評価されている。

[2] Imamura, M. and Nakamura, T., "Parameter-free Spikelet: Discovering Different Length and Warped Time Series Motifs using an Adaptive Time Series Representation," ACM SIGKDD, pp. 857-866 (2023).

[3] Imamura, M., Nakamura, T. and Keogh, E., "Matrix Profile XXI: A Geometric Approach to Time Series Chains Improves Robustness," ACM SIGKDD, pp. 1114-1122 (2020).

5. 主な発表論文等

〔雑誌論文〕 計15件（うち査読付論文 13件 / うち国際共著 5件 / うちオープンアクセス 10件）

1. 著者名 Imamura Makoto, Nakamura Takaaki	4. 巻 1
2. 論文標題 Parameter-free Spikelet: Discovering Different Length and Warped Time Series Motifs using an Adaptive Time Series Representation	5. 発行年 2023年
3. 雑誌名 ACM KDD(Knowledge Discovery and Datamining)	6. 最初と最後の頁 1067-1072
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3580305.3599310	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Nakamura Takaaki, Mercer Ryan, Imamura Makoto, Keogh Eamonn	4. 巻 37
2. 論文標題 MERLIN+: parameter-free discovery of time series anomalies	5. 発行年 2023年
3. 雑誌名 Data Mining and Knowledge Discovery	6. 最初と最後の頁 670 ~ 709
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s10618-022-00876-7	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 Makoto Imamura, Nakamura Takaaki	4. 巻 印刷中
2. 論文標題 Parameter-free Spikelet: Discovering Different Length and Warped Time Series Motifs using an Adaptive Time Series Representation	5. 発行年 2023年
3. 雑誌名 ACM KDD 2023	6. 最初と最後の頁 857-866
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3580305.3599310	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Imamura Makoto, Nakamura Takaaki	4. 巻 1
2. 論文標題 Spikelet: An Adaptive Symbolic Approximation for Finding Higher-Level Structure in Time Series	5. 発行年 2021年
3. 雑誌名 IEEE ICDM (International Conference on Data Mining)	6. 最初と最後の頁 1120-1125
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICDM51629.2021.00132	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Hara Shigeomi, Douzono Hiroshi, Imamura Makoto, Yoshioka Tatsuya	4. 巻 12
2. 論文標題 Estimation of Photovoltaic Cell Parameters Using Measurement Data of Photovoltaic Module String Currents and Voltages	5. 発行年 2022年
3. 雑誌名 IEEE Journal of Photovoltaics	6. 最初と最後の頁 540 ~ 545
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/JPHOTOV.2021.3135262	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Terada Masahiro, Minakawa Takumi, Anzai Hiroto, Imamura Makoto	4. 巻 141
2. 論文標題 Unsupervised Discovery of Repetitive Activity Patterns in Acceleration Time Series Data	5. 発行年 2021年
3. 雑誌名 IEEJ Transactions on Electronics, Information and Systems	6. 最初と最後の頁 1039 ~ 1047
掲載論文のDOI (デジタルオブジェクト識別子) 10.1541/ieejieiss.141.1039	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Imamura Makoto, Nakamura Takaaki, Keogh Eamonn	4. 巻 1
2. 論文標題 Matrix Profile XXI: A Geometric Approach to Time Series Chains Improves Robustness	5. 発行年 2020年
3. 雑誌名 26th ACM SIGKDD international conference on Knowledge discovery and data	6. 最初と最後の頁 1114-1122
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3394486.3403164	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Nakamura Takaaki, Imamura Makoto, Mercer Ryan, Keogh Eamonn	4. 巻 1
2. 論文標題 MERLIN: Parameter-Free Discovery of Arbitrary Length Anomalies in Massive Time Series Archives	5. 発行年 2020年
3. 雑誌名 20th IEEE International Conference on Data Mining	6. 最初と最後の頁 1190-1195
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICDM50108.2020.00147	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Shein Thi Thi, Puntheeranurak Sutheera, Imamura Makoto	4. 巻 8
2. 論文標題 Discovery of Loose Group Companion From Trajectory Data Streams	5. 発行年 2020年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 85856 ~ 85868
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2020.2992596	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Shein Thi Thi, Puntheeranurak Sutheera, Imamura Makoto	4. 巻 62
2. 論文標題 Discovery of evolving companion from trajectory data streams	5. 発行年 2020年
3. 雑誌名 Knowledge and Information Systems	6. 最初と最後の頁 3509 ~ 3533
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s10115-020-01471-2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Tatemichi Syouta, Yamauchi Azusa, Shimada Takeru, Imamura Makoto	4. 巻 140
2. 論文標題 Concept Hierarchy Network based Information Retrieval for Restaurant Recommendation Dialogue	5. 発行年 2020年
3. 雑誌名 IEEJ Transactions on Electronics, Information and Systems	6. 最初と最後の頁 1230 ~ 1238
掲載論文のDOI (デジタルオブジェクト識別子) 10.1541/ieejeiss.140.1230	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Terada Masahiro, Yago Tadao, Matsubara Naoya, Miyake Hiroki, Imamura Yudai, Inoue Mao, Imamura Makoto	4. 巻 140
2. 論文標題 An Analysis Method of Baggage Lifting Motion for Lower Backache Prevention	5. 発行年 2020年
3. 雑誌名 IEEJ Transactions on Electronics, Information and Systems	6. 最初と最後の頁 1369 ~ 1379
掲載論文のDOI (デジタルオブジェクト識別子) 10.1541/ieejeiss.140.1369	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Masaki Endo, Munenori Takahashi , Masaharu Hirota, Makoto Imamura, Hiroshi Ishikawa	4. 巻 Vol. 12, No: 3
2. 論文標題 Analytical Method using Geotagged Tweets Developed for Tourist Spot Extraction and Real-time Analysis	5. 発行年 2020年
3. 雑誌名 International Journal of Informatics Society (IJIS)	6. 最初と最後の頁 157-165
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 遠藤 雅樹, 佐藤 崇志, 今村 誠, 谷口 雄治, 佐々木 建太, 菅沼 啓, 高橋 秀誠	4. 巻 Vol. 37, No. 1
2. 論文標題 ドイツにおける職業訓練におけるデジタル化の取り組み	5. 発行年 2020年
3. 雑誌名 技能科学研究	6. 最初と最後の頁 1- 6
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 今村 誠	4. 巻 Vol 65, No. 4
2. 論文標題 予知保全のための機械学習	5. 発行年 2021年
3. 雑誌名 システム/制御/情報	6. 最初と最後の頁 119-125
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計14件 (うち招待講演 2件 / うち国際学会 0件)

1. 発表者名 今村 誠
2. 発表標題 予知保全のための機械学習
3. 学会等名 日本材料学会, 腐食防食部門委員会 第346回例会 (招待講演)
4. 発表年 2023年

1. 発表者名 田上樹, 小西一輝, 今村誠
2. 発表標題 深度センサを用いた関節角度に基づく運動動作フォームの解析
3. 学会等名 電気学会電子・情報・システム部門大会
4. 発表年 2022年

1. 発表者名 濱岡弘, 小西一輝, 今村誠
2. 発表標題 適応的記号近似を用いた時系列からの反復作業パターンの教師なし発見方式
3. 学会等名 電気学会電子・情報・システム部門大会
4. 発表年 2022年

1. 発表者名 小西一輝 王雲傑 五月女龍二 河内広樹 田上樹 今村誠
2. 発表標題 スポーツ動画からの深層学習を用いた膝関節角度の抽出方式
3. 学会等名 日本体育測定評価学会第22回大会
4. 発表年 2023年

1. 発表者名 坂口慧 塚本博字 佐久間拓海 濱岡弘生 今村誠
2. 発表標題 健康的な歩行判定のための時系列からの着地時点の抽出方式
3. 学会等名 日本体育測定評価学会第22回大会
4. 発表年 2023年

1. 発表者名 今村 誠
2. 発表標題 時系列モチーフの構造と時間発展について
3. 学会等名 電子情報通信学会 情報通信マネジメント研究会 信学技報, vol. 121, no. 354, ICM2021-41, pp. 46-50, (招待講演)
4. 発表年 2021年～2022年

1. 発表者名 南 直輝, 皆川 拓海, 小西 一輝, 今村 誠
2. 発表標題 動作時系列の分類のための関節位置に基づく特徴量について
3. 学会等名 情報処理学会 第33回コンシューマ・デバイス&システム研究会(CDS), CDS33-10
4. 発表年 2021年～2022年

1. 発表者名 安齋博人・皆川拓海・今村誠
2. 発表標題 動作解析における画像センサーと深度センサーのロバスト性の比較実験
3. 学会等名 電気学会C部門大会, TC8-12
4. 発表年 2021年～2022年

1. 発表者名 皆川拓海, 濱岡弘生, 今村誠
2. 発表標題 時系列モチーフ解析を用いた加速度データからの動作検出方式
3. 学会等名 電気学会C部門大会, TC8-11
4. 発表年 2021年～2022年

1. 発表者名 島田建, 濱岡弘生, 今村誠
2. 発表標題 畳み込みニューラルネットを用いた時系列からのパターン抽出法
3. 学会等名 電気学会C部門大会, PS3-1
4. 発表年 2021年～2022年

1. 発表者名 今村誠
2. 発表標題 現場から学ぶ! データサイエンスからみた予知保全手法の体系化
3. 学会等名 MathWorksライブイベント
4. 発表年 2021年～2022年

1. 発表者名 寺田 昌弘, 皆川 拓海, 安齋 博人, 今村 誠
2. 発表標題 加速度時系列データからの反復作業パターンの教師なし発見方式
3. 学会等名 情報処理学会 研究報告情報システムと社会環境
4. 発表年 2020年

1. 発表者名 川島寛隆, 今村誠
2. 発表標題 多次元時系列に対する深層表現学習における比較実験
3. 学会等名 電気学会 情報システム研究会
4. 発表年 2020年

1. 発表者名 寺田 昌弘, 井上 真生, 島田 建, 南 直輝, 今村 誠
2. 発表標題 モチーフ発見と文法的推論に基づく繰り返し作業時系列のセグメンテーション方式
3. 学会等名 人工知能学会全国大会論文集
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関