情報処理学会論文誌　データベース　Vol.19　No.2　70–81 (Apr. 2026)

時間変化する因果関係の抽出に基づいた高速将来予測

千原 直己[1,2,a)]　松原 靖子[1,b)]　藤原 廉[1,2,c)]　櫻井 保志[1,d)]

受付日 2025年9月10日, 採録日 2025年10月28日

概要：本論文では，大規模時系列データストリーム中の時間変化する因果関係の抽出および将来予測を同時に行うための最新手法 ModePlait を提案する．提案手法は以下の優れた特性をすべて満たす．(a) 時々刻々と変化する環境の移り変わりに従って変化する因果関係を明らかにする．(b) 時間変化する因果関係の抽出および将来予測を同時かつ正確に行う．(c) 計算時間は時系列データストリーム全体の長さに依存せず，高速に処理を行う．人工データおよび実データを用いた評価実験により，提案手法が最新の既存手法に比べて因果探索，将来予測の両方の観点において高精度であること，そして，計算効率の良い高速な処理が可能であることを明らかにした．

キーワード：時系列データ分析，ストリームデータ，動的システム，因果探索

Modeling Time-evolving Causality and Forecasting in Data Streams

Naoki Chihara[1,2,a)]　Yasuko Matsubara[1,b)]　Ren Fujiwara[1,2,c)]　Yasushi Sakurai[1,d)]

Received: September 10, 2025, Accepted: October 28, 2025

Abstract: We study the novel problem of discovering the time-varying cause-and-effect relationships across transitions of dynamical patterns in multivariate co-evolving data streams. To solve such a problem, we present a streaming method, ModePlait, which is designed for modeling such causal relationships (i.e., time-evolving causality) and forecasting their future values. ModePlait has the following desirable properties: (a) *Effective*: it discovers the time-evolving causality in multivariate co-evolving data streams by detecting the transitions of distinct dynamical patterns adaptively. (b) *Accurate*: it enables both the discovery of time-evolving causality and the forecasting of future values in a streaming fashion. (c) *Scalable*: our algorithm does not depend on data stream length and thus is applicable to very large sequences. Extensive experiments on both synthetic and real-world datasets demonstrate that our proposed model outperforms state-of-the-art methods in terms of discovering the time-evolving causality as well as forecasting.

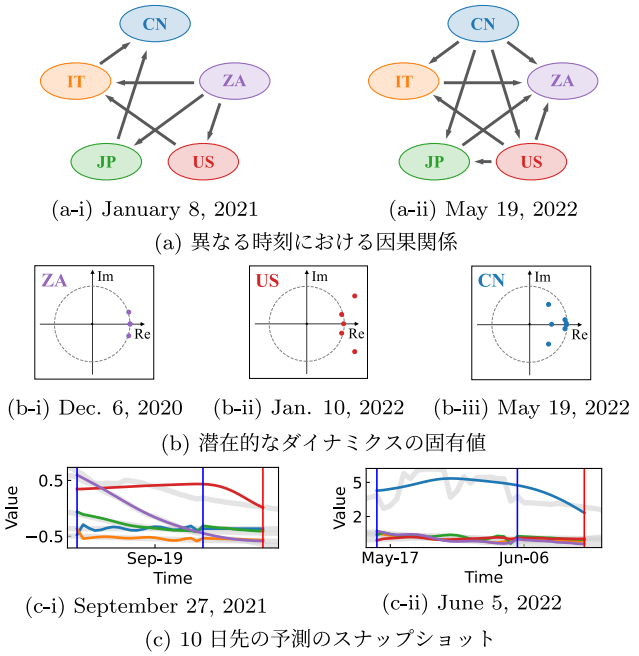
Keywords: time series analysis, data stream, dynamical system, causal discovery

1. はじめに

時系列データは，Internet of Things（IoT）[1], [2]，ソーシャルネットワーク [3], [4]，医療情報分析 [5], [6]，ユーザー行動分析 [7] など多岐にわたるイベントおよびアプリケーションから大量に生成されている．特に，実世界においてはこれらのデータは高速かつ継続的に生成され続けるため，ストリーミング方式で処理することの需要が増している．

また，時系列データの観測値間にはさまざまな関係性（e.g., 相関，独立性）が存在し，それらはクラスタリング [8], [9]，将来予測 [10]，欠損値補完 [11] など [12]，幅広い時系列分析にとって重要な特徴である．その中でも因果関係 [13], [14] は特に価値のある洞察を提供することで知られており，多くの研究がその調査に注力している [15], [16]．また，因果関係を帰納バイアスとして取り入れることで，より汎化性の高い表現を学習し，下流タスクの改善を図る

[1] 大阪大学産業科学研究所産業科学 AI センター
SANKEN, The University of Osaka, Ibaraki, Osaka 567–0047, Japan
[2] 大阪大学大学院情報科学研究科
IST, The University of Osaka, Suita, Osaka 565–0871, Japan
[a)] naoki88@sanken.osaka-u.ac.jp
[b)] yasuko@sanken.osaka-u.ac.jp
[c)] r-fujiwr88@sanken.osaka-u.ac.jp
[d)] yasushi@sanken.osaka-u.ac.jp



(a-i) January 8, 2021　(a-ii) May 19, 2022

(a) 異なる時刻における因果関係

(b-i) Dec. 6, 2020　(b-ii) Jan. 10, 2022　(b-iii) May 19, 2022

(b) 潜在的なダイナミクスの固有値

(c-i) September 27, 2021　(c-ii) June 5, 2022

(c) 10 日先の予測のスナップショット

図 1　疫病データストリームに対する ModePlait の結果

Fig. 1　Modeling power of ModePlait over an epidemiological data stream (i.e., #1 *covid19*).

研究もさかんである [17], [18]．しかし，大半の研究では多次元時系列データ中の因果関係は定常であると仮定している．新しい原因を迅速に検出し，ストリーミング形式で正確に予測するには，時間変化する関係性の発見が重要である．たとえば，感染症の流行では，ある国で新規のウイルス株が出現すると，その国内での感染者数が急増するだけでなく，国境を越える移動などといった特定の活動が他国の感染者数の増加を引き起こす可能性が存在する．また，このような原因となる国は時間の経過とともに変化する．

時間変化する因果関係をとらえるために，これらの変化の要因ともいえる，時々刻々と変化する環境の変化をとらえる必要がある．これらは，特徴的な時系列パターンの遷移という形で現れる．具体的には，web の検索データの場合，時系列パターンの遷移は新商品の発売のようなさまざまな理由で生じる．そして，これらのイベントは既存の商品の売り上げなどに影響を与える，すなわち，因果関係の時間変化にまでつながるといえる．本論文ではこのような時系列パターンのことを “レジーム” と呼ぶ．

本研究では，上記で述べた時間変化する因果関係を構造方程式モデル [19] に基づいて表現し，時系列データストリームをリアルタイムに予測する手法として ModePlait を提案する．本論文では次の問題を扱う．

問題.　d 次元の観測データで構成される半無限長の時系列データストリーム $\boldsymbol{X} = \{\boldsymbol{x}(1), \ldots, \boldsymbol{x}(t_c), \ldots\}$ が与えられ，現時刻を t_c としたとき，(a) 潜在的な時系列パターン（レジーム）の遷移の検出，(b) その遷移に従って時間変化する因果関係の抽出，(c) l_s ステップ先の値 $\boldsymbol{v}(t_c + l_s)$ の予測をすべて達成する．

1.1　具体例

図 1 は疫病データストリーム (#1) *covid19* に対する ModePlait を適用した際の出力の様子を示している．このデータセットは，主要 5 か国（日本，米国，中国，イタリア，南アフリカ共和国）における COVID-19 感染者数の推移を記録したもので構成される．

図 1 (a) は ModePlait が抽出した時間変化する因果関係を示している．図中の各矢印の基部は原因を，先端は結果を意味する．ModePlait は，疫病データストリームから国間の時間変化する因果関係を正確に抽出する．たとえば，図 1 (a-i) は，南アフリカ共和国（ZA）が他国に影響を及ぼしていたことを示している．これは，南アフリカ保健省が 2020 年 12 月 18 日に南アフリカ共和国で COVID-19 の新たな変異株，つまり，501.V2 の発見を発表したという事実 [20] に対応している．ゆえに，新たな変異株による影響を ModePlait が適応的に検出したといえる．さらに，図 1 (a-ii) は，図 1 (a-i) とは対照的に，中国（CN）が他の国々に影響を与えていたことを示している．これは，2022 年 4 月初旬から 2022 年 6 月 1 日まで続いた上海で最も長く厳しいロックダウンの期間と一致している [21]．これは，ModePlait が上海での厳格で長期的なロックダウンにつながる程の COVID-19 の感染拡大の影響を検出したことを意味する．

図 1 (b) は，外生変数における潜在的なダイナミクスの固有値を示す．これらの図は複素平面であり，灰色の点線は単位円で，色付きの点は潜在的ダイナミクスの固有値であり，それぞれが特定のモードの減衰率と時間周波数を示している．具体的には，固有値の絶対値が 1 より大きい場合，対応するモードは増幅を，1 より小さい場合は減衰を示す．図 1 (b-i) は，南アフリカ共和国の外生変数の弱い増幅モードを示しており，南アフリカ共和国での感染者数が前述の新たな変異株によって増加したことを意味する．図 1 (b-ii) は，米国（US）における外生変数の強い増幅モードを示しており，新規感染者数が 1 日で初めて 100 万人を超えた状況を反映している [22]．図 1 (b-iii) は，中国に対応する外生変数の減衰モードを示している．この期間は上海のロックダウンの終わり頃で，感染の拡大が緩和し始めている事実があり，提案手法はこれを正確にとらえているといえる．

図 1 (c) は，現在のウィンドウが与えられた場合の $l_s = 10$ ステップ先の将来予測のスナップショットを示している．青い縦軸は，現在のウィンドウの開始時点 t_m と現在の時点 t_c を，赤い縦軸は l_s ステップ先の時点 $t_c + l_s$ を示している．また，推定された値は色付きの太線で示し，元の値は灰色の線で示す．ModePlait は現在の特徴的な時系列パターンをとらえ，任意の時点で連続的に将来の値を生成

する．

本論文の貢献．　ModePlait [*1]の貢献点を以下に示す．

- 時系列データ中のレジームの推移を逐次的にとらえ続けることにより，時間変化する因果関係を発見する．
- 因果関係を理論的に発見し，高精度に将来を予測する．
- 計算時間はデータストリーム全体の長さに依存しない．

論文の構成．　本論文の構成は以下のとおりである．2 章では，関連研究を概観し，本研究の位置付けを明確にする．3 章では，時間変化する因果関係を表現するための提案モデル ModePlait の定式化を行う．4 章では，提案モデルを学習し，ストリーミング環境下で逐次的に更新・予測を行うための具体的なアルゴリズム，およびその理論的な解析について述べる．5 章では，人工データおよび実データを用いた評価実験により，提案手法の有効性を実証する．最後に，6 章で本論文を総括し，今後の課題について述べる．

2. 関連研究

時系列モデリングと将来予測．　時系列モデリングおよび将来予測は，多くの分野で大きな関心を集めている非常に重要な分野である．自己回帰和分移動平均（ARIMA）[24] やカルマンフィルター（KF）[25] は，従来のモデリングおよび予測方法の代表的な例であり，それらの派生型に関する研究も数多く行われている [26], [27], [28], [29]．TICC [8] は，マルコフ確率場に基づいて異なる観測間の相互依存性を特徴付けるが，因果関係をとらえることはできない．また，ストリーミング形式に対応可能な手法は，時間やメモリの制約下で膨大なデータを処理するために不可欠な存在となっており，データマイニングやデータベース分野において非常に重要であることが実証されている [3], [30], [31], [32]．OrbitMap [33] は，ストリーム予測に焦点を当てた最新の手法で，主要な動的時系列パターン間の遷移を見つけることができる．ただし，観測間の因果関係を発見することはできない．また，深層学習モデルを用いた時系列予測の研究が近年非常に活発である [10], [34], [35], [36], [37]．深層学習モデルは高い表現能力を有しているものの，時系列データの学習にともなう計算コストが非常に高いため，最新の観測データによるモデルの継続的な更新が困難となり，ストリーミング形式の将来予測への適用は限定的である．

因果推論および探索．　因果推論や因果探索 [15], [16], [38], [39], [40] そして，課題解決のための因果の概念の応用 [17], [41], [42] などに関する幅広い研究が長年行われている．NOTEARS [43] は，有向非巡回グラフ（DAG）の推定問題を，非巡回制約項を組み込んだ滑らかな連続最適化問題として定式化した微分可能なフレームワークである．グレンジャー因果 [44] は，時間的な因果関係を解析するために広く利用されている．しかしながら，グレンジャー因果は変数間の予測的因果関係を示すものであり，典型的な因果関係とは異なる [45]．具体的には，従来の因果関係はある観測が原因となって別の観測を引き起こすかどうかを表すのに対し，グレンジャー因果性はある観測が別の観測の予測に役立つかどうかを示す [46]．本論文では，時系列データストリーム中の時間変化する因果関係の抽出に焦点を当てる．

[*1] この論文は文献 [23] を発展させ，追加実験を行ったものをまとめたものである．

3. 提案モデル

本章では，提案モデルを紹介する．本題に入る前に，ModePlait の基本的な概念について説明する．提案モデルは $\boldsymbol{X}_{\text{sem}} = \boldsymbol{B}_{\text{sem}}\boldsymbol{X}_{\text{sem}} + \boldsymbol{E}_{\text{sem}}$ のように記述される構造方程式モデル（SEM）[19] に基づいて設計される．ここで，$\boldsymbol{X}_{\text{sem}}$ は観測変数，$\boldsymbol{B}_{\text{sem}}$ は因果隣接行列，$\boldsymbol{E}_{\text{sem}}$ は非ガウス分布を持つ相互に独立した外生変数である．本論文では，データ生成過程は線形であり，因果関係を有向非巡回グラフで表現し，未観測共通原因が存在しないことを仮定する．構造方程式モデルは典型的な因果を表現できるが，実社会に存在する特徴的な時系列パターンの遷移に応じて時間変化する因果は表現できない．そこで，提案モデルは外生変数が動的システムとして振舞うと仮定する．しかし，外生変数は互いに独立しているため，それらを単一の動的システムとして考えるのは不適切である．これらをふまえて，本研究では，さまざまな特徴的な時系列パターン（レジーム）を含む時系列データストリームが与えられたとき，時間変化する因果関係を逐次的に発見することを目指す．

3.1 ModePlait モデル

本節では提案モデルの詳細を述べる．初めに必要な概念の定義について説明する．

定義 1（固有信号：$\boldsymbol{E}$）．非ガウス分布に従う d 個の相互に独立した信号 $\boldsymbol{E} = \{\boldsymbol{e}_{(i)}\}_{i=1}^{d}$ を固有信号と呼ぶ．ただし，$\boldsymbol{e}_{(i)} = \{e_{(i)}(1), \ldots, e_{(i)}(t)\}$ は i 番目の単変量時系列である．これは，時間の経過に従って変化するという特徴がある．

図 2 は提案モデルの全体図である．提案手法は，以下のような特徴をとらえることで目的を達成する．

(P1) 固有信号の潜在的な時間ダイナミクス

(P2) 単一レジーム内の特徴的な時系列パターン

(P1) は，固有信号を基底ベクトル（モード）の重ね合わせで表現する．そして，上記の要素を基に **(P2)** をとらえる．

3.1.1 固有信号中の潜在的な時間ダイナミクス (P1)

初めに，i 番目の固有信号 $\boldsymbol{e}_{(i)} = \{e_{(i)}(1), \ldots, e_{(i)}(t)\}$ から潜在的な時間ダイナミクスをとらえる方法について説明する．問題点としては，システム内の潜在的なダイナミクスが一般に多次元であるため，システムを十分に表現するためには，単次元なデータではしばしば不十分であることがあげられる．この問題点を補うために，状態空間の拡張手法を活用する．特に，本研究

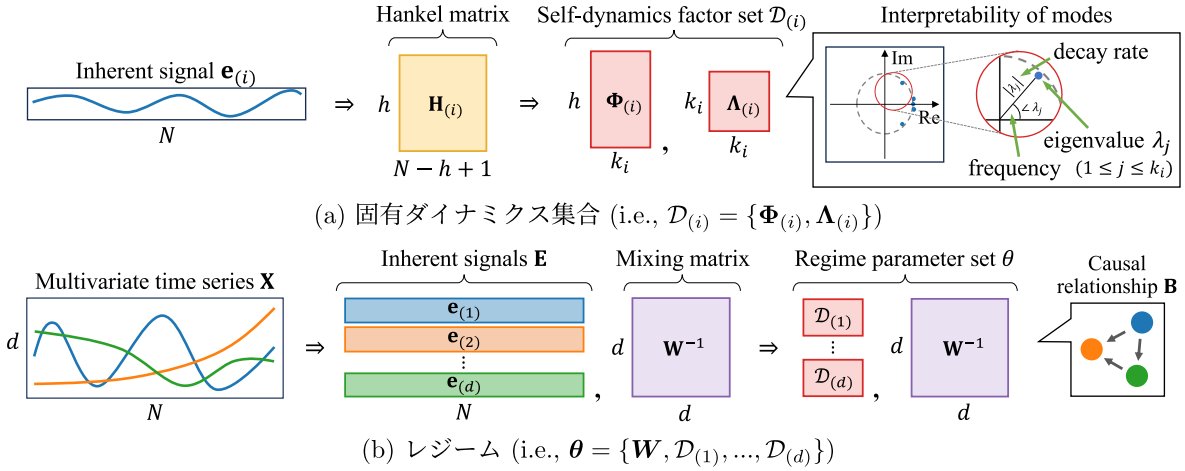


図 2　ModePlait のモデル概要図
Fig. 2　Illustration of ModePlait.

では非線形なダイナミクスの抽出に有効な時間遅れ埋め込みを採用する．具体的には，これは一般的な観測量 $g(e_{(i)}(t)) := (e_{(i)}(t), e_{(i)}(t-1), \ldots, e_{(i)}(t-h+1)) \in \mathbb{R}^h$ に基づいており，非線形システムのアトラクタを幾何学的に再構成するための確立された手法である．ただし，h は埋め込み次元である．上記の $g(\cdot)$ を用いてハンケル行列を形成する．

$$\boldsymbol{H}_{(i)} = \begin{bmatrix} | & | & & | \\ g(e_{(i)}(h)) & g(e_{(i)}(h+1)) & \cdots & g(e_{(i)}(t)) \\ | & | & & | \end{bmatrix} \tag{1}$$

式 (1) のとおり，各状態ベクトルは過去情報を付与して拡張されている．さらに，Takens の埋め込み定理 [47] によれば，特定の条件下において，時間遅れ埋め込みによって生成されるベクトルは，元の状態と微分同相なダイナミクスを持つことが保証される．直感的に説明をすると，この再構成は元の力学系の特性を理論的に保つ，つまり，ハンケル行列 $\boldsymbol{H}_{(i)}$ の解析を通じて，元のデータからは直接抽出できない重要な特徴を明らかにすることを可能にする．多くの場合，微分同相写像を犠牲にすることなく埋め込み次元を選択できる．

ここで，i 番目の固有信号 $\boldsymbol{e}_{(i)}$ の動的システムのために，k_i 次元の複素数値の潜在状態 $\boldsymbol{s}_{(i)}(t) \in \mathbb{C}^{k_i}$ を導入する．ただし，k_i はモードの数である．したがって，i 番目の固有信号 $\boldsymbol{e}_{(i)}$ は以下の式で記述される．

モデル 1.　$\boldsymbol{s}_{(i)}(t)$ を時刻 t における k_i 次元の潜在状態とする．i 番目の単変量固有信号 $\boldsymbol{e}_{(i)}$ は次の式で表現される．

$$\begin{aligned} \boldsymbol{s}_{(i)}(t+1) &= \mathbf{\Lambda}_{(i)} \boldsymbol{s}_{(i)}(t) \\ e_{(i)}(t) &= g^{-1}(\mathbf{\Phi}_{(i)} \boldsymbol{s}_{(i)}(t)) \end{aligned} \tag{2}$$

ここで，$g^{-1}(\cdot)$ は観測量 $g(\cdot)$ の逆写像，$\mathbf{\Phi}_{(i)}$ の各列が各モードで，$\mathbf{\Lambda}_{(i)}$ が k_i 個の固有値である．

$\boldsymbol{s}_{(i)}(t)$ は k_i 個のモードの重ね合わせで表現される．そして，固有値 $\mathbf{\Lambda}_{(i)} \in \mathbb{C}^{k_i \times k_i}$ が時間ダイナミクスを示し，モード $\mathbf{\Phi}_{(i)} \in \mathbb{C}^{h \times k_i}$ および $g^{-1}(\cdot)$ は時刻 t における i 番目の固有信号 $e_{(i)}(t)$ を生成するための射影を示す．まとめると，以下を得る．

定義 2（固有ダイナミクス集合：$\mathcal{D}_{(i)}$）．モード $\mathbf{\Phi}_{(i)}$ と固有値 $\mathbf{\Lambda}_{(i)}$ による集合 $\mathcal{D}_{(i)} = \{\mathbf{\Phi}_{(i)}, \mathbf{\Lambda}_{(i)}\}$ を固有ダイナミクス集合と呼ぶ．これは，i 番目の単変量固有信号 $e_{(i)}$ の潜在的な時間ダイナミクスを表現する．

3.1.2　単一レジーム内の特徴的な時系列パターン (P2)

続いて，時系列データストリーム中の時間変化する因果関係を考慮した特徴的な時系列パターン（レジーム）を表現する方法について述べる．時刻 t における推定値 $\boldsymbol{v}(t) \in \mathbb{R}^d$ を生成するためのモデルを，d 個の固有ダイナミクス集合 $\mathcal{D}_{(1)}, \ldots, \mathcal{D}_{(d)}$ で構築する．また，d 個の潜在状態を $\boldsymbol{S}(t) = \{\boldsymbol{s}_{(i)}(t)\}_{i=1}^{d}$ と表記する．したがって，多変量時系列データはモデル 1 を拡張した以下の式で記述される．

モデル 2.　$\boldsymbol{s}_{(i)}(t)$ を時刻 t の i 番目の固有信号 $e_{(i)}(t)$ のための k_i 次元の潜在状態，$\boldsymbol{e}(t)$ を時刻 t の d 次元の固有信号 $(\boldsymbol{e}(t) = \{e_{(i)}(t)\}_{i=1}^{d})$，$\boldsymbol{v}(t)$ を時刻 t の d 次元の推定値とする．レジームは次の式で表現される．

$$\begin{aligned} &\boldsymbol{s}_{(i)}(t+1) = \mathbf{\Lambda}_{(i)} \boldsymbol{s}_{(i)}(t) \quad (1 \le i \le d) \\ &e_{(i)}(t) = g^{-1}(\mathbf{\Phi}_{(i)} \boldsymbol{s}_{(i)}(t)) \quad (1 \le i \le d) \\ &\boldsymbol{v}(t) = \boldsymbol{W}^{-1} \boldsymbol{e}(t) \end{aligned} \tag{3}$$

モデル 2 のために，新たなパラメータである分離行列 $\boldsymbol{W}$ を導入する．これは，d 個の固有信号間の関係性を表現し，因果関係の特定のために重要な役割を果たす．$\boldsymbol{W}$ から $\boldsymbol{B}$ を特定するためのアルゴリズムについては 4.2.3 項にて説明する．まとめると，以下を得る．

定義 3（レジーム：$\boldsymbol{\theta}$）．$\boldsymbol{\theta} = \{\boldsymbol{W}, \mathcal{D}_{(1)}, \ldots, \mathcal{D}_{(d)}\}$ をレジームを表現するパラメータ集合とする．ここで，$\boldsymbol{W}$ は因果隣接行列 $\boldsymbol{B}$ を生成するための基盤となる要素である．

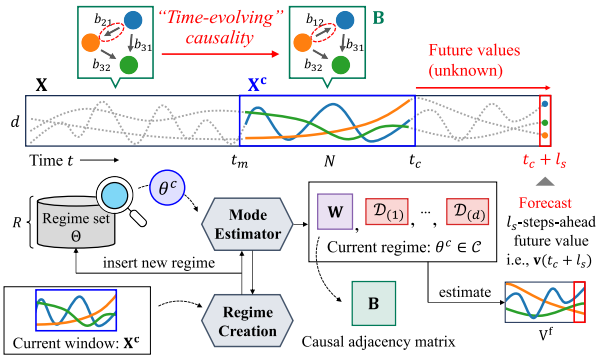


図 3　ModePlait のアルゴリズム概要図：毎時刻新たなデータ $\boldsymbol{x}(t_c)$ を受け取るたびに，最適なレジーム $\boldsymbol{\theta}^c$ を取得し，l_s ステップ先の値 $\boldsymbol{v}(t_c + l_s)$ を予測する

Fig. 3　Overview of ModePlait algorithm: Given a new value $\boldsymbol{x}(t_c)$ sequentially, it searches for the best regime $\boldsymbol{\theta}^c$ and forecasts the l_s-steps-ahead future value, i.e., $\boldsymbol{v}(t_c + l_s)$.

最終的な目的は，複数のレジーム $\boldsymbol{\theta}$ が時々刻々と遷移する様子をとらえることである．なお，このレジームの遷移が因果関係の時間変化を誘発する．時刻 t までの適切なレジーム数を R とすると，時系列データストリーム $\boldsymbol{X}$ は R 個のレジーム $\{\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^R\}$ によって要約される．これらをふまえて，レジームセットおよび時間変化する因果関係を以下のように定義する．

定義 4（レジームセット：$\boldsymbol{\Theta}$）．複数のレジームで構成されるモデルパラメータ集合 $\boldsymbol{\Theta} = \{\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^R\}$ として定義し，レジームセットと呼ぶ．これは，時系列データストリーム全体の特徴的な複数の時系列パターンを表現する．

定義 5（時間変化する因果関係：$\mathcal{B}$）．因果隣接行列の集合を $\mathcal{B} = \{\boldsymbol{B}^1, \ldots, \boldsymbol{B}^R\}$ として定義し，時間変化する因果関係と呼ぶ．ここで，$\boldsymbol{B}^i$ は i 番目のレジーム $\boldsymbol{\theta}^i$ に対応し，レジームの変化に応じて異なる値を取る．

4. アルゴリズム

本章では，時間変化する因果関係 $\mathcal{B}$ およびレジームセット $\boldsymbol{\Theta}$ を推定するための効率的なアルゴリズムを提案する．**図 3** は，提案アルゴリズムの概要図である．初めに，単一のレジームのみを持つと仮定された多変量時系列データからレジームを算出する効果的な方法を提示する．その後，複数の異なる特徴的な時系列パターンを含む時系列データストリームに対して，$\mathcal{B}$ を同定しながら，$\boldsymbol{\Theta}$ を逐次的に更新するアルゴリズムを説明する．

4.1 RegimeCreation

初めに，レジーム $\boldsymbol{\theta} = \{\boldsymbol{W}, \mathcal{D}_{(1)}, \ldots, \mathcal{D}_{(d)}\}$ を推定するためのアルゴリズム RegimeCreation を提案する．このアルゴリズムは，以下の 2 つの主要な手順から構成される．(i) $\boldsymbol{X}$ を分離行列 $\boldsymbol{W}$ と固有信号 $\boldsymbol{E}$ に分解する．(ii) 式 (2) に従って，d 個の固有ダイナミクスセット $\{\mathcal{D}_{(1)}, \ldots, \mathcal{D}_{(d)}\}$ を計算する．最適な因果関係をとらえるために，独立成分分析（ICA）を使用して $\boldsymbol{X}$ を分解する．次に，i 番目の固有ダイナミクスセット $\mathcal{D}_{(i)}$ の計算に関しては，ハンケル行列 $\boldsymbol{H}_{(i)}$ に基づく以下のデータ行列を使用する．

$$\boldsymbol{L}_{(i)} = \begin{bmatrix} g(e_{(i)}(h+1)) & \cdots & g(e_{(i)}(t)) \end{bmatrix} \in \mathbb{R}^{h \times (t-h)}$$
$$\boldsymbol{R}_{(i)} = \begin{bmatrix} g(e_{(i)}(h)) & \cdots & g(e_{(i)}(t-1)) \end{bmatrix} \in \mathbb{R}^{h \times (t-h)}$$

これを基に，ここでは以下の重み付き損失関数を使用する．

$$\min_{\boldsymbol{A}} \sum_{t'=h}^{t-1} \mu^{2(t-1-t')} ||g(e_{(i)}(t'+1)) - \boldsymbol{A}_{(i)} g(e_{(i)}(t'))||_2^2$$
$$= \min_{\boldsymbol{A}} ||(\boldsymbol{L}_{(i)} - \boldsymbol{A}_{(i)} \boldsymbol{R}_{(i)}) \boldsymbol{M}||_F^2 \qquad (4)$$

ここで，$\boldsymbol{A}_{(i)}$ は遷移行列であり，これは固有値分解が与えるモード $\boldsymbol{\Phi}_{(i)}$ と対応する固有値分解 $\boldsymbol{\Lambda}_{(i)}$ により構成される．また，$\boldsymbol{M} = \mathrm{diag}(\mu^{t-h-1}, \ldots, \mu^0) \in \mathbb{R}^{(t-h)\times(t-h)}$ は忘却行列である．これは逐次最小二乗法の原理に基づいている．加えて，Koopman 作用素論 [48] によると，古典的な線形時不変システムのモード分解とは異なり，遷移行列は線形であるが非線形動的システムにも適用可能である．まとめると，以下を得る．

I. 独立成分分析を使用して $\boldsymbol{X} = \boldsymbol{W}^{-1}\boldsymbol{E}$ を計算する．
II. 式 (1) に従い，ハンケル行列 $\boldsymbol{H}_{(i)}$ を構成する．
III. データ行列のペア $(\boldsymbol{L}_{(i)}, \boldsymbol{R}_{(i)})$ を計算する．
IV. 特異値分解（SVD）を使用して $\boldsymbol{R}_{(i)}\boldsymbol{M} = \boldsymbol{U}_{(i)}\boldsymbol{\Sigma}_{(i)}\boldsymbol{V}_{(i)}^\top$ を計算する．特異値の数は [49] に従って決定する．
V. 遷移行列 $\boldsymbol{A}_{(i)}$ を左特異値ベクトル $\boldsymbol{U}^{(i)}$ が張る k_i 次元部分空間に射影する．

$$\tilde{\boldsymbol{A}}_{(i)} = \boldsymbol{U}_{(i)}^\top \boldsymbol{A}_{(i)} \boldsymbol{U}_{(i)} = \boldsymbol{U}_{(i)}^\top \boldsymbol{L}_{(i)} \boldsymbol{M} \boldsymbol{V}_{(i)} \boldsymbol{\Sigma}_{(i)}^{-1} \in \mathbb{R}^{k_i \times k_i}$$

VI. 固有値分解 $\tilde{\boldsymbol{A}}_{(i)}\boldsymbol{Z}_{(i)} = \boldsymbol{Z}_{(i)}\boldsymbol{\Lambda}_{(i)}$ を計算する．ここで，$\boldsymbol{U}^{(i)}$ は直交行列であるため，固有値行列 $\boldsymbol{\Lambda}_{(i)}$ は $\boldsymbol{A}_{(i)}$ の支配的な k_i 個の固有値と一致する．
VII. モード $\boldsymbol{\Phi}_{(i)} = \boldsymbol{U}_{(i)}\boldsymbol{Z}_{(i)}$ を計算する．

4.1.1 安定した将来予測

RegimeCreation によって推定されるモデルは特徴的な時系列パターンを正確に表現する点においては優れている．その一方で，学習データに含まれる観測ノイズやデータの偏りといった要因に過剰に適合してしまう結果，予測値が時間とともに発散したり，現実的ではない不自然な振動を示したりするなど，不安定な挙動を呈することがある．この問題を緩和し，より安定した予測を実現するために，以下のような正則化付き損失関数を導入する．

$$\min_{\boldsymbol{A}} ||(\boldsymbol{L}_{(i)} - \boldsymbol{A}_{(i)}\boldsymbol{R}_{(i)})\boldsymbol{M}||_F^2 + \alpha ||\boldsymbol{A}_{(i)}||_F^2 \qquad (5)$$

ここで，α は正則化パラメータである．本正則化はスペクトル半径（固有値の絶対値の最大値）の上界を間接的に抑制させる働きがある．スペクトル半径とはシステムの安定

性を特徴付けるもので，値が大きいほどシステムが不安定であることを示す．したがって，本正則化は，学習したモデルの予測の発散を抑え，安定した将来予測に寄与する．

4.2　ストリーミングアルゴリズム

続いて，提案モデルを用いてストリーミング方式で分離行列 $\boldsymbol{W} \in \boldsymbol{\theta}$ から因果隣接行列 $\boldsymbol{B}$ を特定し，予測する方法を提案する．はじめに，いくつかの重要な概念を定義する．

定義 6（更新用パラメータ：$\boldsymbol{\omega}$）．レジーム $\boldsymbol{\theta}$ を更新するためのパラメータ集合を $\boldsymbol{\omega} = \{\{\boldsymbol{P}_{(i)}\}_{i=1}^{d}, \{\boldsymbol{\epsilon}_{(i)}\}_{i=1}^{d}\}$ と定義し，更新用パラメータと呼ぶ．ただし，$\boldsymbol{P}_{(i)} = (\boldsymbol{R}_{(i)}\boldsymbol{M}\boldsymbol{R}_{(i)}^{\top})^{-1}$，$\boldsymbol{\epsilon}_{(i)}$ はエネルギーである．

定義 7（モデルパラメータ集合：$\mathcal{F}$）．$\mathcal{F} = \{\boldsymbol{\Theta}, \boldsymbol{\Omega}\}$ を ModePlait のパラメータ集合とする．ただし，$\boldsymbol{\Theta}$ および $\boldsymbol{\Omega}$ は R 個のレジーム $\boldsymbol{\Theta} = \{\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^R\}$，更新用パラメータ集合 $\boldsymbol{\Omega} = \{\boldsymbol{\omega}^1, \ldots, \boldsymbol{\omega}^R\}$ によって構成される．

これらの定義に基づくと，問題定義は以下のとおりである．

問題 1．時刻 t_c における新規データを $\boldsymbol{x}(t_c)$ として，時系列データストリーム $\boldsymbol{X}$ が与えられたとき，

- 最適なモデルパラメータ集合 $\mathcal{F} = \{\boldsymbol{\Theta}, \boldsymbol{\Omega}\}$ を発見する，
- 時間変化する因果関係 $\mathcal{B}$ を抽出する，
- l_s ステップ先の値 $\boldsymbol{v}(t_c + l_s)$ を予測する．

ここで，カレントウィンドウ $\boldsymbol{X}^c = \boldsymbol{X}[t_m : t_c]$ のレジームを $\boldsymbol{\theta}^c$ と，$\boldsymbol{\theta}^c$ に対応する更新パラメータを $\boldsymbol{\omega}^c$ と呼ぶ．さらに，l_s ステップ先の値 $\boldsymbol{v}(t_c + l_s)$ を予測するには，現在の時刻 t_c での潜在ベクトル $\boldsymbol{S}(t_c)$ が必要であるため，これを $\boldsymbol{S}_{en}^c$ として保持する．要約すると，提案アルゴリズムは，これらをモデル候補 $\mathcal{C} = \{\boldsymbol{\theta}^c, \boldsymbol{\omega}^c, \boldsymbol{S}_{en}^c\}$ として保持する．

4.2.1　全体像

ModePlait は以下のアルゴリズムによって構成される．

- ModeEstimator：最適なモデルパラメータ集合 $\mathcal{F}$ およびモデル候補 $\mathcal{C}$ を推定する．
- ModeGenerator：現在の $\mathcal{C}$ より，l_s ステップ先の値 $\boldsymbol{v}(t_c + l_s)$ を予測し，因果隣接行列 $\boldsymbol{B}$ を抽出する．
- RegimeUpdater：現在の更新用パラメータ $\boldsymbol{\omega}^c$ と新規データ $\boldsymbol{x}(t_c)$ より，現在のレジーム $\boldsymbol{\theta}^c$ を更新する．

4.2.2　ModeEstimator

現在の時刻 t_c の観測値 $\boldsymbol{x}(t_c)$ が与えられたとき，最初にモデルパラメータ集合 $\mathcal{F}$ およびカレントウィンドウ $\boldsymbol{X}^c$ を最も表現するモデル候補 $\mathcal{C}$ を逐次的に更新する．ここで，$f(\boldsymbol{X}^c; \boldsymbol{S}_0^c, \boldsymbol{\theta}^c)$ はカレントウィンドウ $\boldsymbol{X}^c$ と推定ウィンドウ $\boldsymbol{V}^c$ の誤差を最小化することによって，最適なモデルパラメータ集合を算出する（i.e., $f(\boldsymbol{X}^c; \boldsymbol{S}_0^c, \boldsymbol{\theta}^c) = \sum_{t=t_m+h-1}^{t_c} ||\boldsymbol{x}(t) - \boldsymbol{v}(t)||$）．式 (3) に基づくと，$\boldsymbol{S}_0^c$ を計算する最も簡便な方法は $\{\boldsymbol{\Phi}_{(i)}^{\dagger} g(e_{(i)}(t_m + h - 1))\}_{i=1}^{d}$ を用いることである．しかし，過度なノイズが含まれた初期値は適切な予測が達成できない．これを対処するために，LM（Levenberg-Marquardt）アルゴリズム [50] を使用して $\boldsymbol{S}_0^c$ を最適化し，観測におけるノイズの影響を除去する．まとめると，ModeEstimator は次の手順に従う．

I. カレントウィンドウ $\boldsymbol{X}^c$ と現在のレジーム $\boldsymbol{\theta}^c$ 間の誤差を最小化するように初期値 $\boldsymbol{S}_0^c$ を最適化する
II. $f(\boldsymbol{X}^c; \boldsymbol{S}_0^c, \boldsymbol{\theta}^c) > \tau$ のとき，最適な $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ を得る．
III. $f(\boldsymbol{X}^c; \boldsymbol{S}_0^c, \boldsymbol{\theta}^c) > \tau$ のとき，RegimeCreation で新たなレジームを生成し，レジームセット $\boldsymbol{\Theta}$ に追加する．

4.2.3　ModeGenerator

続いて，逐次的に因果隣接行列 $\boldsymbol{B}$ を抽出し，l_s ステップ先の値 $\boldsymbol{v}(t_c + l_s)$ を予測するアルゴリズム ModeGenerator を提案する．予測については，式 (3) に従って $\boldsymbol{v}(t_c + l_s)$ を推定する．一方で，因果隣接行列 $\boldsymbol{B}$ については前述のとおり分離行列 $\boldsymbol{W} \in \boldsymbol{\theta}^c$ から抽出する．独立成分分析で得られる混合行列（分離行列の逆行列）には独立成分の順序および尺度という 2 つの主要な不定性が存在する．しかし，最適な因果隣接行列を抽出するためには，これらの問題を解決しなければならない．上記の不定性を解消し，因果隣接行列 $\boldsymbol{B}$ を特定するアルゴリズムは以下のとおりである．

I. $\boldsymbol{W}$ の行を並べ替えることで，主対角線上にゼロを含まない行列 $\tilde{\boldsymbol{W}}$ を得る．
II. $\tilde{\boldsymbol{W}}$ の各行を対応する対角要素で割ることで，主対角線上にすべて 1 を持つ新たな行列 $\tilde{\boldsymbol{W}}'$ を得る．
III. $\boldsymbol{B}$ の推定値を $\hat{\boldsymbol{B}} = \boldsymbol{I} - \tilde{\boldsymbol{W}}'$ によって算出する．
IV. 最後に，因果順序を得るために，$\hat{\boldsymbol{B}}$ の置換行列 $\boldsymbol{K}$ を用いて $\tilde{\boldsymbol{B}} = \boldsymbol{K}\hat{\boldsymbol{B}}\boldsymbol{K}^{\top}$ を計算する．これは，$\tilde{\boldsymbol{B}}$ の上三角行列の要素の総和を最小化する．

4.2.4　RegimeUpdater

最後に，既存レジームの表現力向上のために，新規データを使用したレジームの更新方法について説明する．RegimeUpdater は主に，(i) 分離行列 $\boldsymbol{W}$ の更新，および (ii) 固有ダイナミクス集合 $\mathcal{D}$ の更新の 2 つの手順から構成される．手順 (i) では，適応フィルタに基づいたアルゴリズムを使用する [51], [52]．これは，計算とメモリの両方の観点で非常に効率的である．更新手順は以下のとおりである．

I. 現在時刻 t_c において，更新前の $\boldsymbol{W}$ の i 番目の行ベクトル $\boldsymbol{w}_i$ に $\boldsymbol{x}(t_c)$ を射影することで，i 番目の固有信号 $g(e_{(i)}(t_c))$ を算出する．
II. $g(e_{(i)}(t_c))$ を用いて，復元誤差およびエネルギー $\boldsymbol{\epsilon}_{(i)}$ を計算する．
III. 誤差およびエネルギー $\boldsymbol{\epsilon}_{(i)}$ を用いて $\boldsymbol{w}_i$ を更新する．

一方で，手順 (ii) では以下の再帰式を用いる．

$$\boldsymbol{A}_{(i)}^{new} = \boldsymbol{A}_{(i)}^{prev} + (g(e_{(i)}(t_c)) - \boldsymbol{A}_{(i)}^{prev} g(e_{(i)}(t_c - 1)))\boldsymbol{\gamma}_{(i)}$$

$$\boldsymbol{\gamma}_{(i)} = \frac{g(e_{(i)}(t_c - 1))^{\top} \boldsymbol{P}_{(i)}^{prev}}{\mu + g(e_{(i)}(t_c - 1))^{\top} \boldsymbol{P}_{(i)}^{prev} g(e_{(i)}(t_c - 1))}$$

$$\boldsymbol{P}_{(i)}^{new} = \frac{1}{\mu}(\boldsymbol{P}_{(i)}^{prev} - \boldsymbol{P}_{(i)}^{prev} g(e_{(i)}(t_c - 1))\boldsymbol{\gamma}_{(i)})$$

ここで，$\boldsymbol{\Phi}_{(i)}$ および $\boldsymbol{\Lambda}_{(i)}$ はそれぞれ $\boldsymbol{A}_{(i)}$ の固有ベクト

ル，固有値を表す．この再帰式は，式 (4) に示した重み付き損失関数を最小化し，新規データ $\boldsymbol{x}(t_c)$ に着目することで，時系列パターンの変化に適応する．まとめると，RegimeUpdater は以下のとおりである．

I. 手順 (i) で述べたアルゴリズムに従い，新規データ $\boldsymbol{x}(t_c)$ を用いて分離行列 $\boldsymbol{W}$ を更新する．
II. 更新後の分離行列 $\boldsymbol{W}$ を用いて，カレントウィンドウ $\boldsymbol{X}^c$ から現在の固有信号を $\boldsymbol{E}^c$ を計算する．
III. 手順 (ii) で述べた更新式に従い，固有ダイナミクス集合 $\mathcal{D}_{(i)}$ を更新する．

4.3 理論的な分析

最後に，ModePlait の時間計算量および因果関係の識別可能性について議論する．

定理 1. RegimeCreation の計算量は $O(N(d^2+h^2)+k^3)$ である．ただし，$k=\max_i(k_i)$ である．

証明 1. RegimeCreation における主要な手順は I，IV，VI である．独立成分分析を用いて $\boldsymbol{X}$ を $\boldsymbol{W}^{-1}$ と $\boldsymbol{E}$ に分解するために必要な計算量は $O(d^2N)$ である．各観測に対して，$\boldsymbol{R}_{(i)}\boldsymbol{M}$ の特異値分解は $O(h^2N)$ の計算時間を要し，$\tilde{\boldsymbol{A}}_{(i)}$ の固有値分解には $O(k_i^3)$ の計算時間が必要となる．IV および VI の処理を簡便な方法で実行する場合，計算を d 回逐次的に行うことになるため，全体で $O(dh^2N+\sum_i k_i^3)$ の時間計算量を要する．しかし，これらの操作は互いに干渉しないため，並列処理による計算が可能である．したがって，RegimeCreation の計算量は $O(N(d^2+h^2)+k^3)$ となる．ただし，$k=\max_i(k_i)$ である．□

定理 2. ModePlait における因果探索は，ModeGenerator での因果隣接行列 $\boldsymbol{B}$ の抽出と同値である．

証明 2. まず，因果構造を定式化する必要がある．本論文では，構造方程式モデル（structural equation model, SEM）[19] を利用し，$\boldsymbol{X}_{\mathrm{sem}}=\boldsymbol{B}_{\mathrm{sem}}\boldsymbol{X}_{\mathrm{sem}}+\boldsymbol{E}_{\mathrm{sem}}$ と表現する．このモデルは因果性の一般的な定式化として知られており，このモデルにおける $\boldsymbol{B}_{\mathrm{sem}}$ が特定できれば，因果性を発見したといえる．いい換えれば，提案するアルゴリズムがこのモデルに準拠する因果隣接行列 $\boldsymbol{B}$ を特定できることを証明する必要がある．構造方程式モデルを $\boldsymbol{X}_{\mathrm{sem}}$ について解くと $\boldsymbol{X}_{\mathrm{sem}}=\boldsymbol{W}_{\mathrm{sem}}^{-1}\boldsymbol{E}_{\mathrm{sem}}$ を得る．ここで，$\boldsymbol{W}_{\mathrm{sem}}=\boldsymbol{I}-\boldsymbol{B}_{\mathrm{sem}}$ である．観測データが非ガウス独立成分の線形かつ可逆な混合である場合，独立成分分析（ICA）[53] により上記の式における $\boldsymbol{W}_{\mathrm{sem}}$ が，独立成分の順序とスケーリングを除いて識別可能であることが示されている．したがって，提案アルゴリズムである ModeGenerator が混合行列 $\boldsymbol{W}^{-1}$（すなわち，$\boldsymbol{W}\in\boldsymbol{\theta}^c$ の逆行列）の 2 つの不定性を正確に解決することを示せば証明は完了する．なぜなら，$\boldsymbol{W}$ は RegimeCreation 内で独立成分分析により計算されるからである．

まず，提案アルゴリズムが順序の不定性を解決できることを示す．因果隣接行列 $\boldsymbol{B}$ は，非巡回性の仮定により下三角行列に並べ替え可能である [13]．したがって，正しく並べ替えられ，スケーリングされた $\boldsymbol{W}$ は対角成分がすべて 1 である下三角行列となる．さらに，上記の条件を満たすように $\boldsymbol{W}$ を並べ替える方法は一意であるとされている [38]．以上より，ModeGenerator は混合行列の行を並べ替え，主対角線にゼロを含まない行列を得るというステップ I のプロセスを通じて，混合行列の順序を特定できる．次に，尺度の不定性に関しては，並べ替えられスケーリングされた $\boldsymbol{W}$ の対角成分がすべて 1 であることを考慮すれば，注目すべきは対角成分のみであることは明らかである．したがって，提案アルゴリズム ModeGenerator が混合行列 $\boldsymbol{W}^{-1}$ の順序および尺度に関する不定性を解決できることを示した．□

この定理は，提案アルゴリズムが因果関係を正確に抽出できることを理論的に保証する．

定理 3. 定理 1 に基づくと，各プロセスにおける ModePlait の計算時間量は少なくとも $O(N\sum_i k_i+dh^2)$ であり，たかだか $O(RN\sum_i k_i+N(d^2+h^2)+k^3)$ である．

証明 3. 各時点において，ModePlait はまず ModeEstimator を実行し，現在のウィンドウ $\boldsymbol{X}^c$ に対して，最適なモデルパラメータ集合 $\mathcal{F}$ とモデル候補 $\mathcal{C}$ を推定する．現在のレジーム $\boldsymbol{\theta}^c$ が適合する場合，計算時間は $O(N\sum_i k_i)$ である．一方，適合しない場合は，$\boldsymbol{\Theta}\in\mathcal{F}$ 内でより良いレジームを探索するために $O(RN\sum_i k_i)$ の計算時間を要する．さらに，ModePlait が未知のパターンに遭遇した場合，RegimeCreation を実行し，$O(N(d^2+h^2)+k^3)$ の計算時間を要する．その後，ModePlait は ModeGenerator を実行して因果隣接行列を特定し，l_s ステップ先の未来値を予測する．この計算時間量はそれぞれ $O(d^2)$ および $O(l_s)$ であり，l_s は小さい定数値であるため，無視できる．最後に，ModePlait が新しいレジームを作成しない場合，RegimeUpdater を実行するが，この処理の計算時間は $O(dh^2)$ である．したがって，1 プロセスあたりの全体的な計算量は，少なくとも $O(N\sum_i k_i+dh^2)$ であり，最大で $O(RN\sum_i k_i+N(d^2+h^2)+k^3)$ である．□

これは提案手法がデータストリーム全体の長さ t_c に対して一定の計算時間しか必要としないことを示す．したがって，ModePlait は実行速度の観点から半無限長のデータストリームに対して実用的であるといえる．

5. 評価実験

本論文では，ModePlait の有効性を検証するため，人工データおよび実データを用いた実験を行った．本章では以下の項目について検証する．

Q1 提案手法がとらえる時間変化する因果関係の有効性
Q2 因果探索および予測に対する提案手法の精度の検証
Q3 データストリームの予測に対する計算時間の検証

すべての実験において Intel Xeon Platinum 8268 2.9 GHz 24 core CPU，512 GB DDR4 RAM，NVIDIA RTX A6000 GPU を搭載した Linux マシンを使用した．また，カレントウィンドウの長さは $N = 50$ とした．

データセット． 使用した人工データおよび実データは以下のとおりである．また前処理として，各データセットは平均値と分散で正規化（z-normalization）して使用した．

- (#0) *synthetic*：構造方程式モデル [19] に基づいて生成した．詳細は次節を参照されたい．
- (#1) *covid19*：Google COVID-19 Open Data [54] から収集され，日本，アメリカ，中国，イタリア，南アフリカ共和国の感染者数によって構成される．日ごとに記録された 900 日以上のデータが含まれている．
- (#2) *web-search*：Google Trends [55] から 10 年間にわたり毎週収集される，ビールに関連する検索クエリの検索件数で構成される．
- (#3) *chicken-dance*，(#4) *exercise*：CMU motion capture database [56] のデータであり，左脚，右脚，左腕，右腕に対応する 4 次元ベクトルで構成される．

人工データの生成手順． ランダムに生成された人工の多変量時系列データストリームを用意し，各時系列データストリームには特定の因果関係を持つ複数のクラスタを含めた．各クラスタにおいて，因果隣接行列 $\boldsymbol{B}$ は Erdös-Rényi（ER）モデル [57] を使用して生成し，エッジ密度を 0.5，観測変数の数 d を 5 とした．データの生成過程は，構造方程式モデル [19] に基づいており，因果隣接行列 $\boldsymbol{B}$ の各値は一様分布 $\mathcal{U}(-2,-0.5)\cup(0.5,2)$ からサンプリングした．さらに，時間変化性のある外生変数を表現するために，固有信号の分散 $\sigma_{i,t}^2$（すなわち，$e_i(t) \sim \mathrm{Laplace}(0, \sigma_{i,t}^2)$）を時間とともに変化させた．具体的には，$\log(\sigma_{i,t}^2)$ を $h_{i,t}$ と定義し，自己回帰モデルに従って時間変化させた．この自己回帰モデルの係数およびノイズの分散はそれぞれ $\mathcal{U}(0.8, 0.998)$ および $\mathcal{U}(0.01, 0.1)$ からサンプリングした．時系列データストリーム全体は，クラスタセグメントの時間的なシーケンスを構築することで生成し，各セグメントは 500 の観測値を含む（例：「1, 2, 1」は 2 種類の因果関係を含む 3 つのセグメントで構成され，合計のサンプルサイズは 1,500 である）．実験では，さまざまな現実世界のシナリオを反映するために，「1, 2, 1」，「1, 2, 3」，「1, 2, 2, 1」，「1, 2, 3, 4」，および「1, 2, 3, 2, 1」の 5 種類の時間的シーケンスを使用した．

比較手法． 最新の手法を含む 7 つの因果探索手法および 5 つの時系列予測手法，計 12 手法を採用した．

因果探索手法

- CASPER [40]：因果探索における最先端の手法であり，スコア関数にグラフ構造を組み込み，推定された因果構造と真の因果構造との因果距離を反映する．
- DARING [15]：明示的な残差独立制約を課す敵対的学習戦略を使用する．正則化ペナルティ $\{\alpha, \beta, \gamma\}$ について，$\{0.001, 0.01, 0.1, 1.0, 10\}$ の範囲で試行した．
- NoCurl [58]：初期段階で閉路を含む解を生成し，その後にグラフの Hodge 分解を適用する．
- NOTEARS-MLP [59]：NOTEARS [43]（後述）の非線形拡張版であり，生成的構造方程式モデルを MLP で近似することで因果関係を得る．
- NOTEARS [43]：非巡回性を正則化項として導入した微分可能な最適化手法で因果隣接行列を推定する．
- LiNGAM [38]：データの非ガウス性を利用して因果関係の方向性を特定する手法である．
- GES [60]：スコアベースのベイジアンアルゴリズムとして広く知られる手法で，貪欲法を用いて因果関係を探索する．スコア関数として BIC を採用した．

時系列予測手法

- TimesNet [37]：時間的畳み込みネットワーク（TCN）に基づく最先端の手法である．過去のシーケンス長は，カレントウィンドウの長さに合わせて 16 に設定した．
- PatchTST [10]：トランスフォーマーベースの最先端の時系列予測手法である．過去のシーケンス長は，カレントウィンドウの長さに合わせて 16 に設定した．
- DeepAR [34]：RNN に基づく時系列予測モデルであり，将来の値が従う予測分布の生成により予測を行う．
- OrbitMap [33]：ストリーム予測における重要な時間変化パターンを特定する手法である．予測誤差を最小化するように遷移強度 ρ を最適化した．
- ARIMA [24]：線形方程式に基づいた古典的な時系列予測モデルである．最適なパラメータセットを決定するために AIC を使用した．

5.1 Q1：提案手法の有効性

疫病データストリーム (#1) に対する提案手法の結果は 1 章の図 1 に示したとおりである．ModePlait は複雑な社会の情勢の変化が反映された時間変化する因果関係の探索を実現している．加えて，この特徴を用いて高精度な将来予測を達成した．

5.2 Q2：提案手法の精度

因果探索精度． 提案手法の因果探索の精度を検証するために，人工データ (#0) を用いた実験を行った．評価指標については構造ハミング距離（SHD: structural Hamming distance）および構造介入距離（SID: structural intervention distance）[61] を採用した．SHD は因果隣接行列の構造的な距離を定量化する指標であり，欠落した辺，余分な辺，逆転した辺の数を計測する．一方，SID は因果探索の精度を評価するために特に適しており，推定された因果隣接行列を使用した場合に，介入分布 $p(x_j \mid \mathrm{do}(X_i = \bar{x}))$ が誤って計算されるような組 (i, j) の数を測定する．どちらの評価指標も，値が小さいほど推定された因果関係が良い

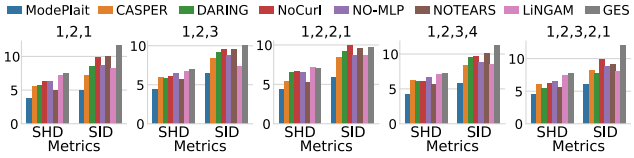


図 4　人工データに対する ModePlait の因果探索の精度

Fig. 4　Causal discovery accuracy of ModePlait.

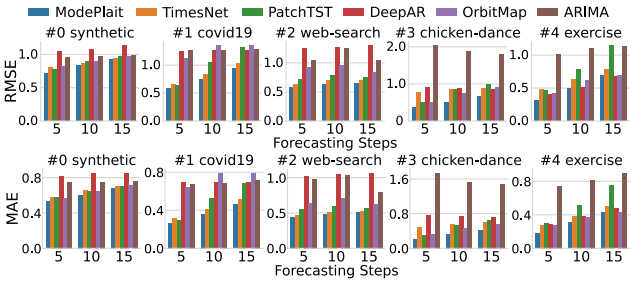


図 5　ModePlait の多次元時系列予測の精度

Fig. 5　Multivariate forecasting accuracy of ModePlait.

ことを意味する．**図 4** は，複数の人工データセットに対する各手法の因果探索の精度を示している．提案手法はすべてのデータセットにおいて一貫して比較手法を上回る性能を示す．この結果は定理 2 の解析結果と一致している．比較手法については，いずれの手法も時系列データストリーム内の時間変化する因果関係を処理できないため精度が低下したといえる．

予測精度．　続いて，ModePlait の l_s ステップ先の予測精度を検証する．評価指標は，推定値の二乗平均誤差（RMSE: root mean square error）および平均絶対誤差（MAE: mean absolute error）の 2 種類を利用した．どちらの評価指標も，値が小さいほど予測精度が良いことを意味する．**図 5** は，複数の予測ステップ幅（i.e., $l_s \in \{5, 10, 15\}$）に対する ModePlait および比較手法の時系列データストリームにおける予測精度を示している．ModePlait は最新の比較手法を上回る予測精度を示している．深層学習モデルは高い表現能力を有しているが，逐次的なパラメータ更新を達成できないため，予測精度が低下する．OrbitMap は複数の特徴的な時系列パターンを扱うことが可能であるが，時間変化する因果関係をとらえることはできないことが性能低下の原因であると考えられる．

アブレーション研究．　因果関係が時系列予測の精度に与える影響を定量的に評価するために，分離行列 $\boldsymbol{W}$ を単位行列に固定した *w/o causality* との比較に基づいてアブレーション研究を行った．**図 6** は，人工データおよび実データの両方を用いて行ったアブレーション研究の結果である．*w/o causality* はすべての実験設定において予測精度が低下していることが確認された．したがって，時系列データストリーム内の時間変化する因果関係を抽出することが予測精度を向上させるといえる．

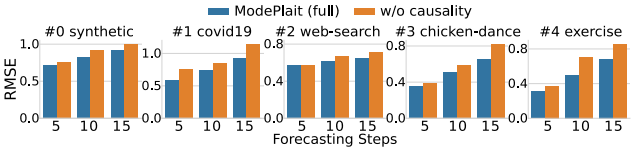


図 6　アブレーション研究の結果

Fig. 6　Ablation study results.

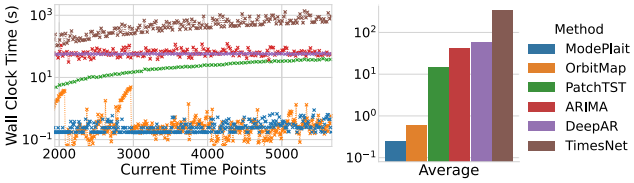


図 7　各時刻 t_c における計算コスト（左）と平均値（右）：これらの図の y 軸は対数スケールで表示されている

Fig. 7　Scalability of ModePlait.

5.3　Q3：提案手法の計算時間

最後に，提案手法の計算コストについて検証する．**図 7** は，ModePlait と比較手法の計算効率を比較したものである．具体的には，図 7 の左図は (#4) *exercise* に対する各時点 t_c における計算コストを，右図は時系列データストリーム全体の計算時間の平均値を示している．これらの図の y 軸は対数スケールで表示している．逐次更新のおかげで，提案手法は比較手法より高速に動作することが可能であり，これは定理 3 の内容と一致している．

6.　むすび

本論文では，大規模時系列データストリーム中の時間変化する因果関係の抽出および将来予測を効率的かつ適応的に実現するための時系列解析技術として ModePlait を提案した．ModePlait は冒頭で確認した以下の優れている特性をすべて達成している．

- 時系列データストリーム中の時間変化する因果関係という重要な特徴を逐次的に抽出する．
- 評価実験において，提案手法が時間変化する因果関係の抽出および将来予測を正確に行うことを実証した．
- 計算コストはデータストリームの長さに依存することなく，高速に処理が可能である．

今後の課題．　ModePlait は 3 章で述べたように，データ生成過程は線形であり，因果関係を有向非巡回グラフで表現し，未観測共通原因が存在しないことを仮定しているため，今後は非線形関係や潜在的な交絡因子を含むより一般的な設定への拡張が重要である．

謝辞　本研究の一部は，JSPS 科研費，JP25KJ1729, JP24KJ1618, JST CREST JPMJCR23M3, JST START JPMJST2553, JST CREST JPMJCR20C6, JST K Program JPMJKP25Y6, JST COI-NEXT JPMJPF2009, JST COI-NEXT JPMJPF2115, 大阪大学博士課程教育リーディングプログラムの助成を受けたものです．

参考文献

[1] De Francisci Morales, G., Bifet, A., Khan, L., Gama, J. and Fan, W.: Iot big data stream mining, *Proc. 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.2119–2120 (2016).

[2] Mahdavinejad, M.S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P. and Sheth, A.P.: Machine learning for Internet of Things data analysis: A survey, *Digital Communications and Networks*, Vol.4, No.3, pp.161–175 (2018).

[3] Kawabata, K., Matsubara, Y., Honda, T. and Sakurai, Y.: Non-Linear Mining of Social Activities in Tensor Streams, *Proc. 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.2093–2102 (2020).

[4] Nakamura, K., Matsubara, Y., Kawabata, K., Umeda, Y., Wada, Y. and Sakurai, Y.: Fast and Multi-aspect Mining of Complex Time-stamped Event Streams, *Proc. ACM Web Conference*, pp.1638–1649 (2023).

[5] Kimura, T., Matsubara, Y., Kawabata, K. and Sakurai, Y.: Fast Mining and Forecasting of Co-evolving Epidemiological Data Streams, *Proc. 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.3157–3167 (2022).

[6] Matsubara, Y., Sakurai, Y., Van Panhuis, W.G. and Faloutsos, C.: FUNNEL: Automatic mining of spatially coevolving epidemics, *Proc. 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.105–114 (2014).

[7] Matsubara, Y., Sakurai, Y., Faloutsos, C., Iwata, T. and Yoshikawa, M.: Fast mining and forecasting of complex time-stamped events, *Proc. 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.271–279 (2012).

[8] Hallac, D., Vare, S., Boyd, S. and Leskovec, J.: Toeplitz inverse covariance-based clustering of multivariate time series data, *Proc. 23rd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.215–223 (2017).

[9] Obata, K., Kawabata, K., Matsubara, Y. and Sakurai, Y.: Dynamic Multi-Network Mining of Tensor Time Series, *Proc. ACM on Web Conference*, pp.4117–4127 (2024).

[10] Nie, Y., Nguyen, N.H., Sinthong, P. and Kalagnanam, J.: A Time Series is Worth 64 Words: Long-term Forecasting with Transformers, *11th International Conference on Learning Representations* (2023).

[11] Wang, D., Yan, Y., Qiu, R., Zhu, Y., Guan, K., Margenot, A. and Tong, H.: Networked time series imputation via position-aware graph enhanced variational autoencoders, *Proc. 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.2256–2268 (2023).

[12] Tozzo, V., Ciech, F., Garbarino, D. and Verri, A.: Statistical models coupling allows for complex local multivariate time series analysis, *Proc. 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.1593–1603 (2021).

[13] Bollen, K.A.: *Structural equations with latent variables*, John Wiley & Sons (1989).

[14] Spirtes, P., Glymour, C.N. and Scheines, R.: *Causation, prediction, and search*, MIT press (2000).

[15] He, Y., Cui, P., Shen, Z., Xu, R., Liu, F. and Jiang, Y.: Daring: Differentiable causal discovery with residual independence, *Proc. 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.596–605 (2021).

[16] Fujiwara, D., Koyama, K., Kiritoshi, K., Okawachi, T., Izumitani, T. and Shimizu, S.: Causal discovery for non-stationary non-linear time series data using just-in-time modeling, *Conference on Causal Learning and Reasoning*, pp.880–894, PMLR (2023).

[17] Dai, E. and Chen, J.: Graph-Augmented Normalizing Flows for Anomaly Detection of Multiple Time Series, *10th International Conference on Learning Representations* (2022).

[18] Cheng, Y., Yang, R., Xiao, T., Li, Z., Suo, J., He, K. and Dai, Q.: CUTS: Neural Causal Discovery from Irregular Time-Series Data, *11th International Conference on Learning Representations* (2023).

[19] Pearl, J.: *Causality: Models, Reasoning, and Inference*, Cambridge university press (2009).

[20] Fink, S.: South Africa announces a new coronavirus variant, *New York Times* (2020).

[21] Buckley, C.: Relief, Reunions and Some Anxiety as Shanghai (Mostly) Reopens, *New York Times* (2022).

[22] Ellyatt, H.: U.S. reports over 1 million new daily Covid cases as omicron surges, *CNBC* (2022).

[23] Chihara, N., Matsubara, Y., Fujiwara, R. and Sakurai, Y.: Modeling Time-evolving Causality over Data Streams, *Proc. 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (2025).

[24] Box, G.E. and Jenkins, G.M.: *Time series analysis: Forecasting and control* (*Revised Edition*), John Wiley & Sons (1976).

[25] Durbin, J. and Koopman, S.J.: *Time series analysis by state space methods*, Vol.38, OUP Oxford (2012).

[26] Papadimitriou, S., Brockwell, A. and Faloutsos, C.: Adaptive, hands-off stream mining, *Proc. 29th International Conference on Very Large Data Bases*, pp.560–571, Elsevier (2003).

[27] Li, L., Prakash, B.A. and Faloutsos, C.: Parsimonious linear fingerprinting for time series, *Proc. VLDB Endowment*, Vol.3, No.1-2, pp.385–396 (2010).

[28] De Livera, A.M., Hyndman, R.J. and Snyder, R.D.: Forecasting time series with complex seasonal patterns using exponential smoothing, *Journal of the American Statistical Association*, Vol.106, No.496, pp.1513–1527 (2011).

[29] Shi, Q., Yin, J., Cai, J., Cichocki, A., Yokota, T., Chen, L., Yuan, M. and Zeng, J.: Block Hankel tensor ARIMA for multiple short time series forecasting, *Proc. AAAI Conference on Artificial Intelligence*, Vol.34, No.4, pp.5758–5766 (2020).

[30] Aggarwal, C.C. (Ed.): *Data Streams - Models and Algorithms*, Advances in Database Systems, Vol.31, Springer (2007).

[31] Hahsler, M. and Bolaños, M.: Clustering data streams based on shared density between micro-clusters, *IEEE Trans. Knowledge and Data Engineering*, Vol.28, No.6, pp.1449–1461 (2016).

[32] Matsubara, Y. and Sakurai, Y.: Regime shifts in streams: Real-time forecasting of co-evolving time sequences, *Proc. 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.1045–1054 (2016).

[33] Matsubara, Y. and Sakurai, Y.: Dynamic modeling and forecasting of time-evolving data streams, *Proc. 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.458–468 (2019).

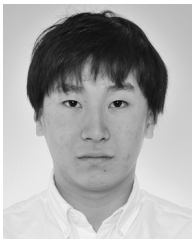
[34] Salinas, D., Flunkert, V., Gasthaus, J. and Januschowski, T.: DeepAR: Probabilistic forecasting with autoregressive recurrent networks, *International Journal of Forecasting*, Vol.36, No.3, pp.1181–1191 (2020).
[35] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H. and Zhang, W.: Informer: Beyond efficient transformer for long sequence time-series forecasting, *Proc. AAAI Conference on Artificial Intelligence*, Vol.35, No.12, pp.11106–11115 (2021).
[36] Zeng, A., Chen, M., Zhang, L. and Xu, Q.: Are Transformers Effective for Time Series Forecasting?, *Proc. AAAI Conference on Artificial Intelligence* (2023).
[37] Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J. and Long, M.: TimesNet: Temporal 2D-Variation Modeling for General Time Series Analysis, *11th International Conference on Learning Representations* (2023).
[38] Shimizu, S., Hoyer, P.O., Hyvärinen, A., Kerminen, A. and Jordan, M.: A linear non-Gaussian acyclic model for causal discovery, *Journal of Machine Learning Research*, Vol.7, No.10 (2006).
[39] Jiang, S., Huang, Z., Luo, X. and Sun, Y.: CF-GODE: Continuous-Time Causal Inference for Multi-Agent Dynamical Systems, *Proc. 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.997–1009 (2023).
[40] Liu, F., Ma, W., Zhang, A., Wang, X., Duan, Y. and Chua, T.-S.: Discovering Dynamic Causal Space for DAG Structure Learning, *Proc. 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.1429–1440 (2023).
[41] Richens, J.G., Lee, C.M. and Johri, S.: Improving the accuracy of medical diagnosis with causal machine learning, *Nature Communications*, Vol.11, No.1, p.3923 (2020).
[42] Wu, C., Wang, X., Lian, D., Xie, X. and Chen, E.: A Causality Inspired Framework for Model Interpretation, *Proc. 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp.2731–2741, ACM (2023).
[43] Zheng, X., Aragam, B., Ravikumar, P.K. and Xing, E.P.: Dags with no tears: Continuous optimization for structure learning, *Advances in Neural Information Processing Systems*, Vol.31, pp.9492–9503 (2018).
[44] Granger, C.W.: Investigating causal relations by econometric models and cross-spectral methods, *Econometrica: Journal of the Econometric Society*, pp.424–438 (1969).
[45] Peters, J., Janzing, D. and Schölkopf, B.: *Elements of causal inference: Foundations and learning algorithms*, The MIT Press (2017).
[46] Granger, C.W.J. and Newbold, P.: *Forecasting economic time series*, Academic press (2014).
[47] Takens, F.: Detecting strange attractors in turbulence, *Dynamical Systems and Turbulence, Warwick 1980: Proc. Symposium Held at the University of Warwick 1979/80*, pp.366–381, Springer (2006).
[48] Koopman, B.O.: Hamiltonian systems and transformation in Hilbert space, *Proc. National Academy of Sciences*, Vol.17, No.5, pp.315–318 (1931).
[49] Gavish, M. and Donoho, D.L.: The optimal hard threshold for singular values is $4/\sqrt{3}$, *IEEE Trans. Information Theory*, Vol.60, No.8, pp.5040–5053 (2014).
[50] Moré, J.J.: The Levenberg-Marquardt algorithm: Implementation and theory, *Numerical Analysis: Proc. Biennial Conference Held at Dundee, 1977*, pp.105–116, Springer (2006).
[51] Yang, B.: Projection approximation subspace tracking, *IEEE Trans. Signal Processing*, Vol.43, No.1, pp.95–107 (1995).
[52] Haykin, S.: Adaptive Filter Theory, *Prentice Hall Google Schola*, Vol.2, pp.67–94 (2002).
[53] Comon, P.: Independent component analysis, a new concept?, *Signal Processing*, Vol.36, No.3, pp.287–314 (1994).
[54] Google COVID-19 Open Data Repository, available from ⟨https://health.google.com/covid-19/open-data/⟩.
[55] Google Trends, available from ⟨https://trends.google.co.jp/⟩.
[56] CMU Graphics Lab Motion Capture Database, available from ⟨http://mocap.cs.cmu.edu/⟩.
[57] Erdos, P., Rényi, A., et al.: On the evolution of random graphs, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, Vol.5, No.1, pp.17–60 (1960).
[58] Yu, Y., Gao, T., Yin, N. and Ji, Q.: DAGs with no curl: An efficient DAG structure learning approach, *Proc. 38th International Conference on Machine Learning*, pp.12156–12166, PMLR (2021).
[59] Zheng, X., Dan, C., Aragam, B., Ravikumar, P. and Xing, E.P.: Learning sparse nonparametric DAGs, *International Conference on Artificial Intelligence and Statistics* (2020).
[60] Chickering, D.M.: Optimal structure identification with greedy search, *Journal of Machine Learning Research*, Vol.3, No.Nov, pp.507–554 (2002).
[61] Peters, J. and Bühlmann, P.: Structural intervention distance for evaluating causal graphs, *Neural Computation*, Vol.27, No.3, pp.771–799 (2015).

千原 直己

2023 年大阪大学工学部電子情報工学科卒業．2025 年同大学院博士前期課程修了．同年同大学院博士後期課程に進学．2025 年より日本学術振興会特別研究員（DC1）．時系列データマイニング，因果推論に関する研究に従事．

松原 靖子　（正会員）

2007 年お茶の水女子大学理学部情報科学科卒業．2009 年同大学大学院博士前期課程修了．2012 年京都大学大学院情報学研究科社会情報学専攻博士後期課程修了．博士（情報学）．2012 年 NTT コミュニケーション科学基礎研究所 RA．2013 年日本学術振興会特別研究員（PD）．2014 年熊本大学大学院自然科学研究科助教．この間，カーネギーメロン大学客員研究員．2016 年国立研究開発法人科学技術振興機構さきがけ研究者．2019 年大阪大学産業科学研究所准教授．2025 年より大阪大学産業科学研究所教授．2016 年度日本データベース学会上林奨励賞，情報処理学会山下記念研究賞，2018 年度 IPSJ/ACM Award for Early Career Contributions to Global Research，2020 年度情報処理学会マイクロソフト情報学研究賞，電気通信普及財団第 36 回テレコムシステム技術賞，令和 4 年度科学技術分野の文部科学大臣表彰「若手科学者賞」，令和 6 年度科学技術分野の文部科学大臣表彰「科学技術賞（研究部門）」，第 57 回市村学術賞貢献賞等受賞．2018〜2019 年度日本データベース学会理事．大規模時系列データマイニングに関する研究に従事．ACM，IEEE，電子情報通信学会，日本データベース学会各会員．



藤原 廉

2021 年大阪大学基礎工学部情報科学科卒業．2023 年同大学院博士前期課程修了．同年同大学院博士後期課程に進学．2024 年より日本学術振興会特別研究員（DC2）．非線形ダイナミクスに基づく時系列データマイニングの研究に従事．

櫻井 保志　（正会員）

1991 年同志社大学工学部電気工学科卒業．1991 年日本電信電話（株）入社．1999 年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了．博士（工学）．2004〜2005 年カーネギーメロン大学客員研究員．2013〜2019 年熊本大学大学院自然科学研究科教授．2019 年より大阪大学産業科学研究所教授．2022〜2025 年日本学術振興会学術システム研究センター主任研究員．本会平成 18 年度長尾真記念特別賞，平成 16 年度および平成 19 年度論文賞，電子情報通信学会平成 19 年度論文賞，日本データベース学会上林奨励賞，ACM KDD best paper awards（2008，2010 年），電気通信普及財団テレコムシステム技術賞（第 27 回，第 36 回），令和 6 年度科学技術分野の文部科学大臣表彰「科学技術賞（研究部門）」等受賞．データマイニング，データストリーム処理，センサーデータ処理，Web 情報解析技術の研究に従事．ACM，IEEE，電子情報通信学会，日本データベース学会各会員．

（担当編集委員　横山 友也）