

研究報告書

「時系列データの自動解析技術の実現」

研究期間：2018年4月～2020年3月

研究者番号：50207

研究者：小林 亮太

1. 研究のねらい

これまで時系列データの解析には時系列モデルが使われてきた。しかし、このアプローチでは、対象やデータが変わるたびに時系列モデルを構築する必要がある。時系列モデルの構築には多くの時間とコストが必要となるため、時系列データの活用は十分に進んでいない。

本研究では、時系列データの中でも「イベント時系列」（点過程データ）、すなわち、あるイベントが起きた時刻についてのデータに着目する。イベント時系列は、Twitter・YouTubeなどのオンライン上の行動履歴（Web情報学）、金融市場における注文履歴（ファイナンス）、神経スパイク（脳科学）、地震（地震学）など分野横断的にあらわれるデータ形式である。

ACT-I 研究期間（2016年12月～2018年3月）では、1つのイベント時系列を対象としSNSやインターネットから得られるイベント時系列の予測技術を開発した。一方、本加速フェーズでは、複数のイベントから得られたイベント時系列（多次元データ）へ対象を拡大させる。加速フェーズのねらいは、

① イベント間の因果関係を抽出する技術を開発する、

② 脳や社会から得られる実データに適用して開発技術の有効性を検証する、

の2点である。一連の研究で開発される技術は、これまであまり着目されてこなかったイベント時系列のデータ活用を促進させ、科学・産業界に変革をもたらすことが期待される。

2. 研究成果

(1) 概要

本研究課題では、複数のイベント（複数の神経細胞のスパイク信号、複数のオンライン動画の閲覧履歴など）に関するイベント時系列データに着目し、以下の2つの研究テーマを進めた。

研究テーマ A 「因果関係を推定する技術の開発」

研究テーマ B 「高精度神経細胞モデルを推定する技術の開発」

以下、個々のテーマの概要について説明する。研究テーマ A では、複数のイベント（複数の神経細胞のスパイク信号、複数のオンライン動画の閲覧履歴など）に関するイベント時系列からイベント間の因果関係を推定する技術を開発した（図1）。

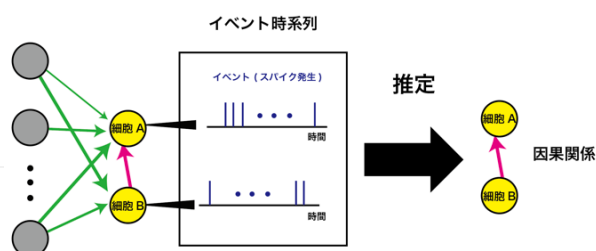


図1: イベント間の因果関係をデータ(イベント時系列)から推定す

さらに、脳の大規模シミュレーションデータやラット海馬の実験データに適用し、開発技術が従来手法に比べてはるかに高い推定精度を有することを確認した。データ解析のプログラム (Python) は研究者が自由に使えるよう公開しており、Web アプリ (JavaScript) も提供している (主な研究成果リスト 5. (3): <http://www.ton.scphys.kyoto-u.ac.jp/~shino/GLMCC/>)。研究テーマ B では、イベント時系列 (神経スパイクデータ) から神経細胞の高精度シミュレーションモデルを得るための方法論を開発した。

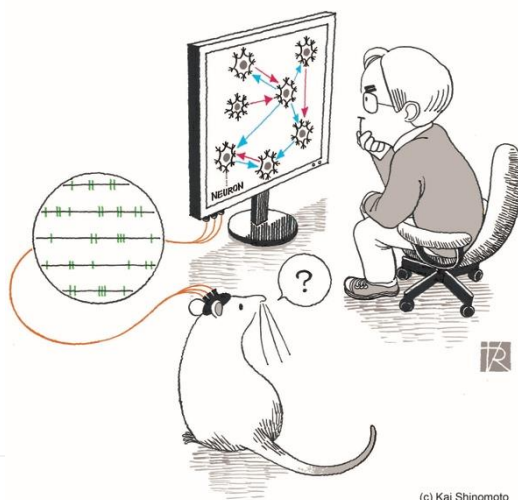
(2) 詳細

上で紹介した 2 つの研究テーマの詳細を説明する。

研究テーマ A 「因果関係を推定する技術の開発」

脳の回路構造、すなわちニューロン (神経細胞) 間の電氣的なつながり (シナプス結合) を知ることは、脳における情報処理・記憶・学習メカニズムの理解にとって極めて重要である。しかし、生きた動物のシナプス結合を直接計測することが非常に困難であるため、大規模な分析はされてこなかった。一方、近年の計測技術の進展によって数多くのニューロンが発生するスパイク信号を長時間にわたって記録できるようになり、それらの信号を分析すれば多数のニューロン間の電氣的なつながりを推定できることが期待されている。ニューロン A からニューロン B に正 (興奮性) の神経結合があれば、ニューロン A が活動した後は B が活動する頻度が上がり、負 (抑制性) の神経結合があれば活動頻度が下がるので、活動頻度の変化を逆にたどることで神経結合を推定できるはずである。このような考え方は 50 年以上前から提案されていたが、神経信号にはニューロン間の結合による直接の影響だけでなく、観測されていない多数のニューロンを介した間接的な影響や外部信号の影響も加わっていて現象が複雑であり、信頼性のある推定をすることができなかった。

本研究では、多数のニューロンから計測されたスパイク時系列からニューロン間の神経結合を高精度に推定するデータ解析技術を開発した (図 2: 主な研究成果リスト 5. (1) Kobayashi et al., Nature Communications 2019)。



(c) Kai Shinomoto

図 2: 研究テーマ A で開発した技術
本研究では、神経スパイク信号から脳の回路図 (神経回路構造) を推定する技術 GLMCC を開発した。神経スパイク信号は、神経細胞がスパイクを生成した時刻のデータ (緑) である。神経回路構造は、ニューロンがどのようにつながっているか (矢印) を示す。

具体的にはスパイク信号から相関イベント時系列（CC: Cross Correlation）を計算し、その後、一般化線形モデル（GLM: Generalized Linear Model）を適用して外部からの入力の影響を消すことによりニューロン間の神経結合を抽出した（図 3A）。GLMCC（Generalized Linear Model for Cross Correlation）と名付けられたこの解析法は、シナプス結合の強さをシナプス膜電位の単位で推定できる。また、推定する膜電位に応じて必要な計測時間の見積もりを与える公式も導出した。

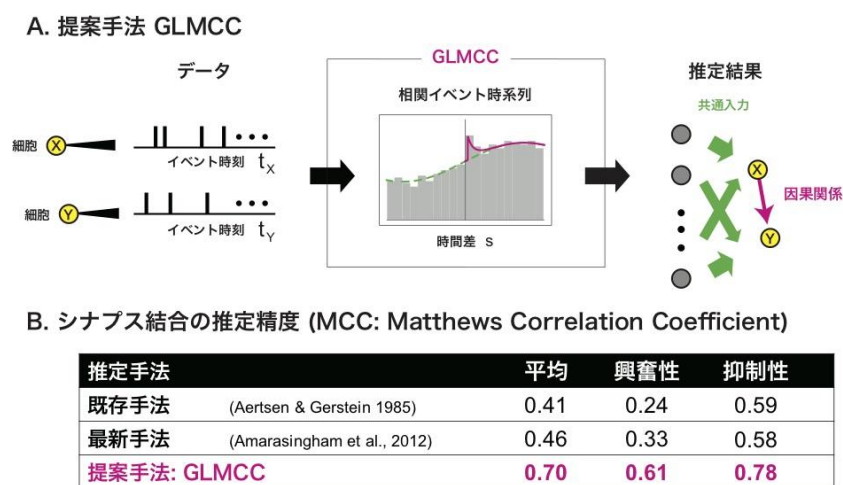


図 3: 因果関係を推定する技術 GLMCC の説明。

A. GLMCC の模式図。GLMCC では、まず 2 つのニューロンのスパイク時系列の相関イベント時系列を計算する。次に、一般化線形モデル（GLM）を相関イベント時系列に適用し、外部からの入力（緑）を取り除きながら、ニューロン間の因果関係（シナプス結合: マゼンタ）を抽出する。

B. 推定精度の比較。MCC は -1 から 1 までの値をとり、大きいほど性能が高いことを示

次に、シナプス結合の推定精度を評価するため、1,000 個のモデルニューロンからなる神経回路のシミュレーションを行った。シミュレーションデータを用いて検証した結果、提案手法は従来手法に比べてはるかに高い推定精度（精度: 97%, MCC: 0.7）を有することを確認した（図 3B）。

最後にラットから計測された実験データに適用し、海馬のニューロン間の神経結合を推定した。神経細胞には、興奮性の結合だけを出力する興奮性細胞と抑制性の結合だけを出力する抑制性細胞がある。スパイク波形や神経活動の様子などから専門家が判定した興奮性・抑制性細胞と推定された興奮性・抑制性細胞は一致していることを確認した。また、ニューロン間の結合確率はこれまでの知見と整合している（結合確率は約 3%）ことを確認した。データ解析のプログラム（Python）は研究者が自由に使えるよう公開しており、Web アプリ

(JavaScript) も提供している (図 4)。本研究は、栗田 修平 (理研 AIP)、北野 勝則 (立命館大)、水関 健司 (大阪市立大)、Anno Kurth, Markus Diesmann (ユーリッヒ研究所、ドイツ)、Barry J. Richmond (NIH、米国)、篠本 滋 (京都大) との国際共同研究であり、Web アプリは内藤 雅博、春名 純一 (京都大: ACT-I 予算で雇用した学生)、篠本 滋と共同で開発した。

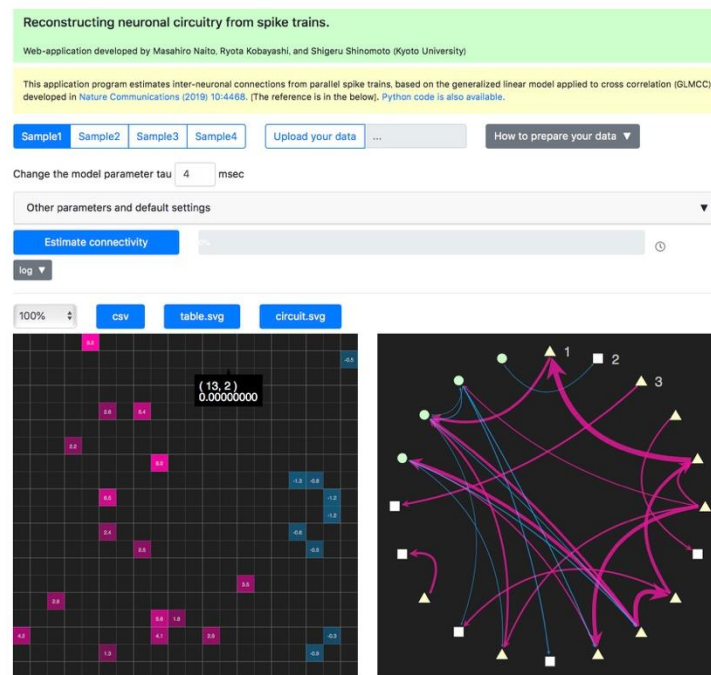


図 1: 内藤 雅博、篠本 滋と共同で開発したシナプス結合を推定する Web アプリ

(<http://www.ton.scphys.kyoto-u.ac.jp/~shino/GLMCC/>)

研究テーマ B「高精度神経細胞モデルを推定する技術の開発」

このテーマでは、イベント時系列 (神経スパイクデータ) から神経細胞の高精度シミュレーション モデルを得るための方法論を開発した。まず、進化的戦略の探索アルゴリズムの 1 つである CMA-ES を用いることにより、神経細胞モデルのパラメータ探索効率を向上させた (主な研究成果リスト 5. (1) Han et al., ICASSP 2019)。本研究は、Han Xu、篠崎 隆宏 (東京工業大) との共同研究である。また、昆虫の嗅覚神経細胞に焦点を当て、神経スパイクデータを高精度に再現するシミュレーションモデルを得るための方法を開発した (主な研究成果リスト 5. (1) Levakova et al., Journal of the Royal Society Interface, 2019)。本研究は、Marie Levakova (コペンハーゲン大、デンマーク)、Lubomir Kostal (チェコ科学アカデミー、チェコ)、Christelle Monsempès, Philippe Lucas (INRA、フランス) との国際共同研究である。

3. 今後の展開

本研究で開発した GLMCC により、神経イベント時系列から脳の神経回路構造を高精度に推定できるようになった。また得られた公式によって、研究者は、実験を行う前に推定に必要な計測時間を見積もることができる。今後は、最新の計測技術によって、視覚野、運動野など様々な脳領域から多くの神経ビッグデータが得られるため、この解析法によって脳内の異なる機能領域における情報処理や情報の流れの違いが明らかになり、脳科学の研究が大きく進展するだろう。

また、本研究で開発した技術は、脳科学だけでなく、Web データ（Twitter, Youtube のログデータ）や人々のコミュニケーション（メール、SMS、携帯電話の通話など）から得られたビッグデータの分析にも応用できる（Aoki et al., Phys Rev E, 2016）。今後は、これらのデータに応用するための方法論の開発を進めていき、炎上やデマ拡散などの社会問題が起こるメカニズムの理解を目指して研究を進めたい。

4. 自己評価

① 研究目的の達成状況

本研究の目的は、イベント間の因果関係を抽出する技術を開発し、実データに適用して開発技術の有効性を検証することであった。本研究では、イベント時系列から因果関係を推定するデータ解析技術である GLMCC を開発した。そして、この成果を論文発表（Kobayashi et al., Nature Communications 2019）した。このことから、研究目的は達成できたと考えている。また、時系列分析の専門知識がなくても気軽に使えるようにするために Web アプリ（<http://www.ton.scphys.kyoto-u.ac.jp/~shino/GLMCC/>）や Python コード（<https://github.com/NII-Kobayashi>）を開発して公開した。これらの反響も大きいことから、研究当初のねらいであった「イベント時系列のデータ活用の促進」についても達成できた。

② 研究の進め方

研究体制としては、研究者（小林）が中心となって進めた。後述するように、本課題は非常に挑戦性が高かったため、篠本滋をはじめ、国内、米国、ドイツの共同研究者と密に議論を行いながら進めた。また、Web アプリや Github の開発を進めるため、3 名の学生を雇用した。研究費は、共同研究者との打ち合わせ旅費、学生の雇用、研究成果発表（論文のオープンアクセス費、国際会議での発表）などに執行した。

③ 研究成果の科学技術及び学術・産業・社会・文化への波及効果

主な研究成果を Nature Communications 誌で発表することができ（Kobayashi et al., 2019）、学術（脳神経科学）に大きなインパクトを与えることができた。また、開発したイベント時系列の解析技術は Web・社会データ（Kobayashi & Lambiotte 2016）への展開も可能であるため、データサイエンスの基盤技術としても重要である。今後は、Web マーケティングや金融データ解析などへの波及効果も期待される。

④ 本研究の独創性・挑戦性

本課題の独創性は、これまであまり着目されてこなかったイベント時系列に着目し、イベント時系列から因果関係を絞り込む技術を開発した点にある。この例として、ニューロン間のシナプス結合の推定問題に取り組んだ。この問題は、脳科学で重要であり、50 年以上も前から研究がされてきた (Perkel et al., 1967) もの、これまで信頼性のある推定ができなかった極めて挑戦性が高い課題であった。本研究 では一般化線形モデル (GLM) と相関イベント時系列解析 (Cross-Correlation) を融合させて GLMCC を新たに提案し、この問題を解決できることを示した。提案技術 GLMCC そのものも独創性が高いと考えている。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. Kobayashi R, Kurita S, Kurth A, Kitano K, Mizuseki K, Diesmann M, Richmond BJ, & Shinomoto S. “Reconstructing Neuronal Circuitry from Parallel Spike Trains”, Nature Communications, 2019, vol. 10: 4468.
2. Levakova M, Kostal L, Monsempès C, Lucas P, Kobayashi R. “Adaptive integrate-and-fire model reproduces the dynamics of olfactory receptor neuron responses in moth” Journal of the Royal Society Interface, 2019, vol. 16: 20190246.
3. Han X, Shinozaki T, and Kobayashi R. “Effective and Stable Neuron Model Optimization Based on Aggregated CMA-ES”, 2019, ICASSP 2019:1264–1268.

(2) 特許出願

研究期間累積件数: 0 件

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

主要な学会発表

1. [招待講演] 小林 亮太, 「イベント時系列モデリング: SNS データ分析や脳科学への応用」, ネットワーク科学セミナー2018, 2018.
2. Kobayashi R, Kurita S, Kitano K, Mizuseki K, Richmond BJ, & Shinomoto S, “Estimation of synaptic connections from parallel spike trains”, Neural Coding 2018, 2018.
3. Kobayashi R, Kurita S, Kitano K, Mizuseki K, Richmond BJ, & Shinomoto S, “Estimating Synaptic Connectivity from Parallel Spike Trains”, The 28th Annual Conference of the Japanese Neural Network Society, 2018 (口頭発表に選ばれた: 10/71 = 14.1%).
4. [招待講演] 小林 亮太, 「イベント時系列解析とその応用」, 第 18 回情報科学技術フォーラム, 2019.
5. [招待講演] 小林 亮太, “Estimating synaptic connections from parallel spike trains”, 神経ダイナミクスミニワークショップ, 2019.

著作物

6. 小林 亮太, 岡本 洋, 山川 宏, 特集「物理学と AI」にあたって, 人工知能 33(4), 2018 年
7. 小林 亮太, 木村 睦, 三宅 陽一郎, 市瀬 龍太郎, 特集「マテリアルズインフォマティクス」にあたって, 人工知能 34(3), 2019 年
8. Kobayashi R., Kurita S, Kurth A, Kitano K, Mizuseki K, Diesmann M, Richmond BJ, & Shinomoto S. “Synthetic spike data generated by a network of 1,000 Hodgkin–Huxley type neurons”, figshare 2019.
9. Web アプリ: Masahiro Naito, Ryota Kobayashi, and Shigeru Shinomoto, “Reconstructing neuronal circuitry from spike trains.”, <http://www.ton.scphys.kyoto-u.ac.jp/~shino/GLMCC/>
10. 小林 亮太, “Universal AI”, AI 事典 第 3 版 (編集: 中島秀之, 浅田稔, 橋田浩一, 松原仁, 山川宏, 栗原聡, 松尾豊). (2019 年 12 月).

プレスリリース

11. 神経信号からニューロンのつながりを推定-神経活動データから脳の回路図を描く-(2019 年 10 月).