

2020年度 発表文献集

Publication List 2020



東京理科大学 工学部情報工学科 池口研究室

Ikeguchi Laboratory,
Department of Information and Computer Technology,
Faculty of Engineering,
Tokyo University of Science

2021年03月作成

目次

1	原著論文	1
	(1) Yutaka Shimada, Kohei Yamamoto and Tohru Ikeguchi, “Detecting prediction limit of marked point processes using constrained random shuffle surrogate data,” Chaos: An Interdisciplinary Journal of Nonlinear Science, Vol. 31, No. 1, 013122, 2021. https://doi.org/10.1063/5.0005267	2
	(2) Ryo Yamazaki, Yutaka Shimada and Tohru Ikeguchi, “Chaos MIMO system with efficient use of information of bits,” Nonlinear Theory and Its Applications, IEICE, Vol. 11, No. 4, pp. 546–560, 2020. https://doi.org/10.1587/nolta.11.546	14
	(3) Kazuya Sawada, Yutaka Shimada, Tohru Ikeguchi, “Network structure detection using convergent cross mapping on multivariate time series,” Nonlinear Theory and Its Applications, IEICE, Vol. 11, No. 4, pp. 422–432, 2020. https://doi.org/10.1587/nolta.11.422	29
2	解説論文	40
	(1) 池口徹, デジタルコンピュータ上での非線形力学系の実装について:−ベルヌーイシフト写像 の場合−, 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, Vol. 14, No. 3, pp. 229–238, 2021. https://doi.org/10.1587/essfr.14.3_229	41
3	査読あり国際会議	51
	(1) Kazuya Sawada, Yutaka Shimada, and Tohru Ikeguchi, “Reconstruction of Dynamical System from Marked Point Process and its Appli- cation to Causality Detection,” Proceedings of 2021 RISP International Workshop on Nonlinear Circuits, Com- munications and Signal Processing (NCSP’21), 1AM1–1–4, pp. 13–16, 2021. . . .	52

(2)	Shiki Kanamaru, Nina Sviridova, Yutaka Shimada, and Tohru Ikeguchi “Discrimination of deterministic chaos and noisy periodicity by the recurrence plot with variable threshold,” Proceedings of 2021 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP’21), 1AM1–1–5, pp. 17–20, 2021. . . .	56
(3)	Kihei Magishi, Tomoko Matsumoto, Yutaka Shimada, and Tohru Ikeguchi, “Effects of Document Characteristics on Language Network Structure,” RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP’21), 1AM2–1–2, pp. 45–48, 2021.	60
(4)	Luyan Xu, Kazuya Sawada, Yutaka Shimada and Tohru Ikeguchi, “Influence of Contact Characteristics to Information Diffusion on Temporal Network,” RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP’21), 1PM1–1–1, pp. 73–76, 2021.	64
(5)	Masayoshi Tosa, Nanfu Miya, Honami Tsushima, and Tohru Ikeguchi “A Method for Searching Edge Candidates of TSP,” Proceedings of 2021 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP’21), 2AM1–1–4, pp. 206–209, 2021. . .	68
(6)	Nina Sviridva, and Tohru Ikeguchi “Application of Recurrence Quantification Analysis to Hypertension Photoplethysmograms,” Proceedings of 2020 International Symposium on Nonlinear Theory and its Applications(NOLTA2020), Communications and Signal Processing (NCSP’21), 2AM1–1–4, pp. 206–209, 2021.	72
4	査読なし国内学会	75
4.1	学生発表会	
(1)	金丸志生, Sviridova Nina, 島田裕, 池口徹, ロジスティック写像のカオス応答とノイズな周期解の判別, 電子情報通信学会 東京支部学生会 第 26 回研究発表会, 19, 2021.	76
(2)	澤田和弥, 島田裕, 池口徹, 因果性検出手法を用いたニューロン間の結合性推定, 電子情報通信学会 東京支部学生会 第 26 回研究発表会, 21, 2021.	77
4.2	研究会	
(1)	金丸志生, Sviridova Nina, 島田裕, 池口徹, リカレンスプロット密度の変化による非線形時系列の特性抽出, 電子情報通信学会 技術研究報告, vol. 120, No. 330, NLP2020-41, pp. 7–12, 2021.	78
(2)	三浦英, 池口徹, 発火間隔時系列を用いたニューロンへの共通入力への再構成, 電子情報通信学会 技術研究報告, Vol. 120, No. 330, pp. 1–6, 2021.	84
(3)	真鍋歩未, 澤田和弥, 宮南風, 池口徹, 感染症数理モデルを用いた感染者隔離効果の定量的評価, 電子情報通信学会 技術研究報告, Vol. 120, No. 214, pp.15–20, 2020.	90

	(4)	對馬帆南, 松浦隆文, 池口徹, 大規模自転車再配置問題に対する動的な重み係数を用いた解法の性能調査, 電子情報通信学会 技術研究報告, Vol. 120, No. 55, NLP2020-14, pp. 19–22, 2020.	96
	(5)	ジョロエン, 澤田和弥, 島田裕, 池口徹, ネットワーク構造が情報拡散に与える影響の解析, 電子情報通信学会 技術研究報告, Vol. 120, No. 55, NLP2020–15, pp. 23–26, 2020.	100
	(6)	眞岸祈平, 松本朋子, 島田裕, 池口徹, 新約聖書を対象とした言語ネットワークの解析, 電子情報通信学会 技術研究報告, Vol. 120, No. 55, NLP2020–16, pp. 27–32, 2020.	104
	(7)	澤田和弥, 島田裕, 池口徹, 決定論的非線形力学系からのマーク付点過程の生成と力学系の特徴保存に関する検討, 電子情報通信学会 技術研究報告, Vol. 120, No. 165, NLP2020–22, pp. 1–6, 2020.	110
4.3		総合大会, ソサイエティ大会, NOLTA ソサイエティ大会等	
	(1)	ジョロエン, 澤田和弥, 島田裕, 池口徹, コンタクトの特徴がテンポラルネットワーク上での情報拡散に与える影響の解析, 2021 年電子情報通信学会総合大会講演論文集, N–1–6, 2021.	116
	(2)	三浦英, 池口徹, 複数の再構成アトラクタを用いたニューロンへの共通入力への推定, 電子情報通信学会 総合大会, N–1–7, 2021.	117
	(3)	澤田和弥, 島田裕, 池口徹, マーク付点過程からの状態空間再構成について, 2021 年電子情報通信学会総合大会講演論文集, N–1–20, 2021.	118
	(4)	眞岸祈平, 松本朋子, 島田裕, 池口徹, 単語の使用頻度と言語ネットワークの構造的特徴の関係, 2021 年電子情報通信学会総合大会講演論文集, N–1–21, 2021.	119
	(5)	金丸志生, Sviridova Nina, 島田裕, 池口徹, カオス時系列とノイズな準周期解の判別, 2021 年電子情報通信学会総合大会講演論文集, N–1–25, 2021.	120
	(6)	對馬帆南, 松浦隆文, 池口徹, 自転車再配置問題の実データに対するタブーサーチ法を用いた解法の性能調査, 電子情報通信学会 2021 年総合大会講演論文集, D–1–3, 2021.	121
	(7)	土佐真義, 宮南風, 對馬帆南, 池口徹, TSP の近似解法に与える枝候補作成手法について, 2021 年電子情報通信学会総合大会講演論文集, D–1–11, 2021.	122
	(8)	眞岸祈平, 松本朋子, 島田裕, 池口徹, 文書内容の違いが言語ネットワークの構造に与える影響, 電子情報通信学会 2020 年ソサイエティ大会講演論文集, N–1–14, 2020.	123
	(9)	ジョロエン, 澤田和弥, 島田裕, 池口徹, テンポラルネットワーク上での情報拡散の解析, 電子情報通信学会 2020 年ソサイエティ大会講演論文集, N–1–11, 2020.	124

	(10)	宮南風, 島田裕, 藤原寛太郎, 池口徹, 興奮性/抑制性ニューロン比率がシナプス伝達情報量に与える影響, 電子情報通信学会 2020 年ソサイエティ大会講演論文集, N-1-10, 2020.	125
	(11)	澤田和弥, 島田裕, 池口徹, マーク付き点過程データに対する因果性解析, 電子情報通信学会 2020 年ソサイエティ大会講演論文集, N-1-6, 2020.	126
	(12)	對馬帆南, 松浦隆文, 池口徹, 自転車再配置問題に対する Or-opt と CROSS-exchange の近傍解を削減した解法の 性能調査, 電子情報通信学会, 2020 年ソサイエティ大会講演論文集, N-1-4, 2020.	127
5	その他発表会		128
5.1	2020 年度 4 研究室合同セミナー		
	(1)	對馬帆南, 自転車再配置問題に対する計算時間を短縮させるための探索戦略	129
	(2)	金丸志生, リカレンスプロット密度を変化させることによる決定論的カオスと非線形確率時系列 の特性抽出	130
	(3)	澤田和弥, 非線形力学系理論に基づいたマーク付点過程間の因果性検出	131
	(4)	毛福佳, マーク付点過程に対する Visibility Graph の有効性の検討	132
	(5)	宮南風, ニューロン比率と物理的距離がシナプスの結合性へ与える影響	133
	(6)	ジョロエン, テンポラルネットワークの構造が情報拡散に与える影響の解析	134
	(7)	土佐真義, TSP の枝候補作成に用いる巡回路誤差率の比較	135
	(8)	眞岸祈平, 単語の使用頻度が言語ネットワークの構造的特徴に与える影響	136
	(9)	岡安高輝, マスクの着用が生体信号に与える影響の調査	137
	(10)	田所知也, 決定論的非線形力学系から生成するマーク付き点過程と力学系の特徴保存についての 調査	138
	(11)	眞鍋歩未, 感染症数理モデルを用いた「感染者隔離」の定量的評価	139
	(12)	相原里江, 複雑ネットワーク理論を用いた文章構造の時代的变化の解析	140
	(13)	大矢早紀, ブートストラップリサンプリング法による時系列データの予測	141
	(14)	澤田幸輝, スペクトルグラフ距離を用いた WS モデルの構造解析	142

	(15)	三浦英, ニューロンへの入力時系列の再構成	143
5.2		GMT Morning Workshop on Nonlinear Dynamics and Statistics	
	(1)	Kazuya Sawada, Yutaka Shimada, Tohru Ikeguchi, “Detecting Causality for Marked Point Processes,” GMT Morning Workshop on Nonlinear Dynamics and Statistic, 2020.	(*)
5.3		東京理科大学 脳学際研究部門 第4回公開シンポジウム	
	(1)	宮南風, 島田裕, 藤原寛太郎, 池口徹, 数理モデルを用いたニューロン比率不均衡による脳機能障害の原因解明, 東京理科大学研究推進機構総合研究院 脳学際研究部門 (BIRD) 第4回公開シンポジ ウム, 2020.	144
	(2)	三浦英, 池口徹, 発火間隔時系列を用いたニューロンへの共通入力の再構成, 東京理科大学研究推進機構総合研究院 脳学際研究部門 (BIRD) 第4回公開シンポジ ウム, 2020.	145
6		非線形ワークショップ	146
6.1		2021年12月13日, 非線形ワークショップ	
	(1)	對馬帆南, 自転車再配置問題の実データに対する解法の性能調査	(*)
	(2)	金丸志生, リカレンスプロット密度の変化による非線形時系列の特性抽出	(*)
	(3)	澤田和弥, マーク付点過程に対する因果性解析手法の検討	(*)
	(4)	毛福佳, マーク付点過程をネットワークに変換する手法の検討と楽曲構造解析への応用	(*)
	(5)	宮南風, 数理モデルを用いたニューロン比率不均衡による脳機能障害の原因解明	(*)
	(6)	ジョロエン, コンタクトの特徴が情報拡散に与える影響の解析	(*)
	(7)	土佐真義, TSPの枝候補に含まれる枝の長さの分布	(*)
	(8)	眞岸祈平, 文書の特徴が言語ネットワークの構造的特徴に与える影響	(*)
	(9)	澤田幸輝, 有向グラフとスペクトルグラフ距離	(*)
	(10)	三浦英, 発火間隔時系列を用いたニューロンへの共通入力の再構成	(*)
6.2		2021年2月27日, 非線形ワークショップ	
	(1)	對馬帆南, 自転車再配置問題の中規模サイズ実データに対するタブーサーチ法を用いた解法の性 能調査	147

	(2)	澤田和弥, 点過程データ間の因果性検出手法の検討	148	
	(3)	ジョロエン, コンタクトの順番が情報拡散に与える影響の解析	149	
	(4)	土佐真義, TSP の近似解法に与える枝候補作成手法の検討	150	
	(5)	眞岸祈平, 新約聖書を対象とした言語ネットワークの解析	151	
	(6)	澤田幸輝, スペクトルグラフ距離における累積寄与率と構造的特徴	152	
7		学位論文	153	
	7.1	修士論文		
		(1)	金丸志生, リカレンスプロットの閾値変化による非線形時系列の特性抽出法, 東京理科大学 大学院工学研究科 経営工学専攻, 修士論文, 2021.	154
		(2)	澤田和弥, 非線形力学系理論に基づいた点過程データに対する因果性検出, 東京理科大学大学院工学研究科 経営工学専攻, 修士論文, 2021.	156
		(3)	毛福佳, “Effectiveness of Visibility Graph for Marked Point Processes and its Application to Analyzing Structural Features of Musical Composition,” 東京理科大学 大学院工学研究科 経営工学専攻, 修士論文, 2021.	158
		(4)	宮南風, “A Study on Emerging of Brain Disorders Induced by Neuron Ratio Imbalance Using a Mathematical Model of Spiking Neural Networks,” 東京理科大学 大学院工学研究科 経営工学専攻, 修士論文, 2021.	160
	7.2	卒業論文		
		(1)	岡安高輝, マスクの着用の有無を対照条件として測定した光電脈波に対する非線形時系列解析, 東京理科大学 工学部 情報工学科, 学士論文, 2021.	162
		(2)	真鍋歩未, 隔離を導入した感染症数理モデルを用いた感染拡大抑制策の定量的評価, 東京理科大学 工学部 情報工学科, 学士論文, 2021.	166
		(3)	相原里江, 複雑ネットワーク理論を用いた文章構造の時代的变化の解析, 東京理科大学 工学部 情報工学科, 学士論文, 2021.	170
		(4)	大矢早紀, ブートストラップリサンプリング法によるリアプノフ指数の推定, 東京理科大学 工学部 情報工学科, 学士論文, 2021.	174
		(5)	澤田幸輝, 固有値の累積寄与率がスペクトルグラフ距離に与える影響, 東京理科大学 工学部 情報工学科, 学士論文, 2021.	178

	(6)	三浦英, 発火間隔時系列から再構成したアトラクタの合成によるニューロンへの共通入力の推定, 東京理科大学 工学部 情報工学科, 学士論文, 2021.	182
8	受賞その他		
	(1)	宮南風, 東京理科大学研究推進機構総合研究院 脳学際研究部門 (BIRD) 第4回公開シンポジウム 優秀賞	186
	(2)	Kazuya Sawada, NCSP'21Student Paper Award	187
	(3)	Shiki Kanamaru, NCSP'21Student Paper Award	188

(*) 印: 口頭発表, スライドは本稿に未収録

原著論文

Detecting prediction limit of marked point processes using constrained random shuffle surrogate data

Cite as: Chaos **31**, 013122 (2021); <https://doi.org/10.1063/5.0005267>

Submitted: 17 April 2020 . Accepted: 18 December 2020 . Published Online: 11 January 2021

 Yutaka Shimada, Kohei Yamamoto, and  Tohru Ikeguchi



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[A generalized permutation entropy for noisy dynamics and random processes](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 013115 (2021); <https://doi.org/10.1063/5.0023419>

[Brain rhythm bursts are enhanced by multiplicative noise](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 013117 (2021); <https://doi.org/10.1063/5.0022350>

[Effect of rate of change of parameter on early warning signals for critical transitions](#)

Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 013116 (2021); <https://doi.org/10.1063/5.0025533>



Detecting prediction limit of marked point processes using constrained random shuffle surrogate data

Cite as: Chaos 31, 013122 (2021); doi: 10.1063/5.0005267
 Submitted: 17 April 2020 · Accepted: 18 December 2020 ·
 Published Online: 11 January 2021



Yutaka Shimada,^{1,a)} Kohei Yamamoto,² and Tohru Ikeguchi^{2,3}

AFFILIATIONS

¹Department of Information and Computer Sciences, Graduate School of Sciences and Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama-shi, Saitama 338-8570, Japan

²Department of Management Science, Graduate School of Engineering, Tokyo University of Science, 6-3-1 Nijuku, Katsushika, Tokyo 125-8585, Japan

³Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, 6-3-1 Nijuku, Katsushika, Tokyo 125-8585, Japan

^{a)} Author to whom correspondence should be addressed: yshimada@mail.saitama-u.ac.jp

ABSTRACT

Marked point processes refer to time series of discrete events with additional information about the events. Seismic activities, neural activities, and price movements in financial markets are typical examples of marked point process data. In this paper, we propose a method for investigating the prediction limits of marked point process data, where random shuffle surrogate data with time window constraints are proposed and utilized to estimate the prediction limits. We applied the proposed method to the marked point process data obtained from several dynamical systems and investigated the relationship between the largest Lyapunov exponent and the prediction limit estimated by the proposed method. The results revealed a positive correlation between the reciprocal of the estimated prediction limit and the largest Lyapunov exponent of the underlying dynamical systems in marked point processes.

© 2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0005267>

Even though the prediction limit is a fundamental and important property of time series data, only a few methods are available for estimating the prediction limits of marked point processes in terms of nonlinear dynamics. In this paper, we propose a method for estimating the prediction limit of marked point processes that are generated from nonlinear dynamical systems, by utilizing random shuffle surrogate data with time constraints. The idea underlying the proposed method is very simple, but it can evaluate an important feature of time series, namely, the largest Lyapunov exponent of the marked point process, from the estimated prediction limit.

I. INTRODUCTION

Various complex phenomena are ubiquitous in the real world. When time series data observed in these real complex phenomena

exhibit complex and irregular behavior, the possibility that these irregular fluctuations might be produced by nonlinear deterministic, possibly chaotic, dynamical systems should be considered. If the time series data are considered to be generated from nonlinear dynamical systems, a variety of methods for analyzing complex and irregular time series data in terms of the nonlinear dynamical systems theory are available when the time series data are observed at regular intervals of time, such as methods for detecting determinism, estimating the largest Lyapunov exponent (LLE), and evaluating the predictability of the time series data.

On the other hand, there are time series data that can be observed only at irregular intervals, which are called marked point process data; for example, the occurrence time of seismic activities, spike timings of neurons, and price movements in financial markets constitute marked point process data. In contrast to regularly sampled time series data, only a few methods have been proposed to analyze marked point process data in terms of deterministic

nonlinear dynamics.^{1,2} In particular, to the best of our knowledge, no methods for estimating the prediction limits of marked point process data in terms of deterministic nonlinear dynamics are available. Namely, it is a fundamental and important research challenge to develop such a method.

In this study, we attempt to estimate the prediction limit of the marked point process data that are generated from deterministic nonlinear dynamical systems. Our method is based on a simple idea: if the observed marked point process data are generated from nonlinear deterministic dynamical systems, the data exhibit short-term correlations, then the time when the correlations disappear can be related to a prediction limit of the observed data. To investigate the time when the short-term correlation disappears, we use a surrogate data method that determines if marked point process data exhibit short-term serial dependence, thereby evaluating the prediction limit of the marked point process data. Although various surrogate data methods have been proposed^{3–6} for analyzing nonlinearity in regularly sampled time series data, these methods cannot be directly applied to marked point process data in our approach. We thus also propose a simple surrogate data method to investigate the prediction limit of the marked point process data, which is an extended random shuffle surrogate method for marked point process data.⁷ In our method, a marked point process is divided into multiple segments, and then surrogate data are generated by shuffling the order of M neighboring segments, where M represents the strength of the constraints on the shuffle of the segments. Using this method, we can generate surrogate data such that local temporal structures (short-term correlations) in the original marked point process data are destroyed. The method developed in this study is called *constrained random shuffle surrogates* (CRSS). Further, we estimate the minimum value of M such that the distribution of discriminating statistics of the CRSS data does not differ significantly from that of the surrogate data generated by shuffling the order of all segments. This minimum value of M represents the approximate time at which the short-term correlation disappears, thereby providing information about the prediction limits of the marked point process data.

We applied the proposed method to marked point process data generated from several nonlinear deterministic dynamical systems and investigated the relationship between the reciprocal of the prediction limit estimated by the proposed method and the LLE of the dynamical systems that generate the marked point process data. The results obtained by conducting numerical simulations indicate that the reciprocal of the estimated prediction limits and the LLE are positively correlated. This result strongly suggests that our method can detect the prediction limit of marked point process data.

II. METHODS

In our method, we first generate CRSS data from the given marked point process data and perform a statistical test to estimate its prediction limits using a measure of recurrence plot quantification analysis⁸ as a discriminating statistic. To obtain a recurrence plot of the marked point process data, we divide the data into smaller segments (time windows) and calculate the distances between the time windows. The marked spike train metric (MSTM) is used to calculate the distances between the marked point process data.⁷ This

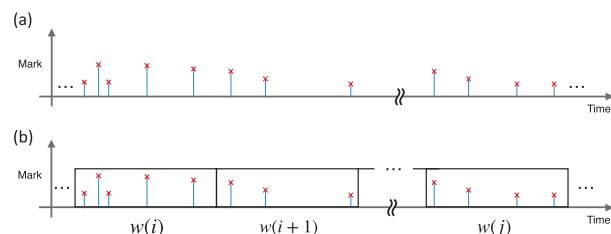


FIG. 1. Example of dividing marked point process data into small segments (time windows). (a) Original data and (b) divided data.

procedure for analyzing marked point process data is grounded in the recurrence-plot-based method introduced by Hirata and Aihara,² where the i th time window, $w(i)$, is defined to be the i th state value, as depicted in Fig. 1. The rest of this section is organized as follows: the MSTM is introduced in Sec. II A; the recurrence plot and a measure of determinism called DET, which is used to quantify the deterministic features of the obtained recurrence plots, are introduced in Sec. II B; and the proposed method is described in Sec. II C.

A. Marked spike train metric (MSTM)

In the MSTM, the distance between two marked point processes is defined to be the total cost of transforming one marked point process into another process through two operations.⁷ The first operation is deletion or insertion of a single event, and its cost is unity. The second operation is a shift in a mark direction or temporal direction of a single event, the cost of which is proportional to the shift length. The distance between two marked point processes is then defined by

$$D(i, j) = \min_C \left\{ I + J - 2P + \sum_{(m, n) \in C} g_{ij}(m, n) \right\}, \quad (1)$$

$$g_{ij}(m, n) = \Lambda_t |t_i(m) - t_j(n)| + \Lambda_p |p_i(m) - p_j(n)|, \quad (2)$$

where i and j are the indices of the marked point process data, or the time window to be compared; $g_{ij}(m, n)$ shows the cost of shifting the m th event in the i th time window to the n th event in the j th time window; m is an event in the i th window; n is an event in the j th window; C is a set of pairs of events (m, n) ; I and J correspond to the total number of events in the i th and j th time windows; and P is the number of elements in the set C . In Eq. (2), $t_i(m)$ is the occurrence time of the m th event in the i th time window, $p_i(m)$ is the mark value of the m th event in the i th time window, the coefficients Λ_t and Λ_p determine the weights of the corresponding terms, and the terms $\Lambda_t |t_i(m) - t_j(n)|$ and $\Lambda_p |p_i(m) - p_j(n)|$ on the right-hand side of Eq. (2) describe the cost of shifting one event to another in the temporal and mark directions, respectively. An example of MSTM calculation between two marked point processes or two time windows, $w(i)$ and $w(j)$, is depicted in Fig. 2.

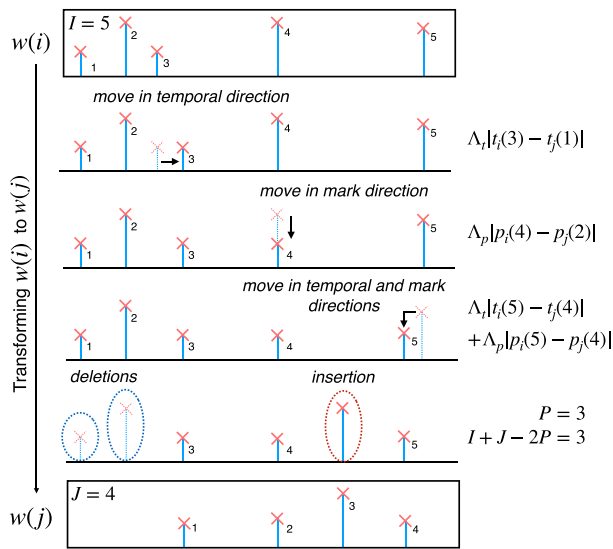


FIG. 2. Example of calculating the MSTM between two marked point process data or the time windows $w(i)$ and $w(j)$.

B. Recurrence plot and DET

Recurrence plots⁸ are used as a tool to visualize recurrences in nonlinear dynamical systems, which represent nonstationarity and nonperiodicity in time series data. A recurrence plot is a two-dimensional image. The value of the (i, j) th pixel of the recurrence plot is defined by

$$R_{i,j} = \begin{cases} 1 & (D(i,j) < \theta), \\ 0 & (\text{otherwise}), \end{cases} \quad (3)$$

where i and j ($i, j = 1, \dots, N$) denote the indices of the i th and j th time windows (Fig. 1), N is the total number of time windows in the marked point process data, θ is a threshold, and $D(i, j)$ is the MSTM between the i th time window $w(i)$ and the j th time window $w(j)$.

Recurrence quantification analysis is an analysis method that quantifies the patterns appearing in recurrence plots. In this study, we use a measure of determinism called DET,⁸

$$\text{DET} = \frac{\sum_{l=l_{\min}}^{N-1} IU(l)}{\sum_{l=1}^{N-1} IU(l)}, \quad (4)$$

where

$$U(l) = \sum_{i=1}^{N-l+1} \sum_{j=i+1}^{N-l+1} (1 - R_{i-1, j-1})(1 - R_{i+l, j+l}) \prod_{k=0}^{l-1} R_{i+k, j+k}. \quad (5)$$

In Eq. (4), $U(l)$ is the number of diagonal lines in the plot whose length is l (pixels) ($1 \leq l \leq N-1$) and l_{\min} is the minimum length of the diagonal line. In this study, l_{\min} is set to two. In Eq. (5), we define that $R_{0,j} = 0$ ($j = 1, \dots, N+1$) and $R_{i, N+1} = 0$ ($i = 1, \dots, N$). Diagonal lines usually disappear in recurrence plots generated from

stochastic processes with no serial dependence or processes with weak determinism. On the other hand, if the recurrence plot is generated from deterministic systems or processes with strong determinism, longer diagonal lines appear. The length of the diagonal lines is closely related to the second order Rényi entropy H_2 (see, for example, Sec. 3.6 in Ref. 8). Therefore, the DET value can contain information about the determinism or predictability.

Because the percentage of plotted points in the recurrence plot directly affects the value of DET, when we compare two or more recurrence plots, recurrence plots need to be generated that have the same ratio of the number of plotted points to the size ($N \times N$). In this study, the threshold θ in Eq. (3) is determined such that the number of plotted points in each recurrence plot is $0.1 \times N \times N$.

C. Estimating prediction limits by constrained random shuffle surrogate data

We attempt to estimate the prediction limit of marked point process data generated from nonlinear deterministic dynamical systems. If the observed marked point process data are generated by nonlinear dynamical systems, the data exhibit short-term correlations, and the time when the correlations disappear can be related to the prediction limit of the observed marked point process data.

To examine whether there is a short-term correlation or not, we first propose a method for generating random shuffle surrogate data with time window constraints, namely, CRSS data, thereby ascertaining the time when the short-term correlations disappear. We generate CRSS data using the following algorithm:

1. A marked point process is divided into multiple time windows whose lengths are L .
2. M neighboring time windows are grouped, and the order of time windows in each group is randomly shuffled.

In this method, we randomly shuffle the temporal order of M neighboring time windows, thereby destroying the short-term correlation of marked point process data. The above algorithm preserves the distribution of distances between time windows. This method is an extended method of the random shuffle surrogate for the marked point process data,⁷ where CRSS corresponds to the random shuffle surrogate when $M = N$. To check whether the marked point process data have short-term correlations, we employ the null hypothesis that the observed marked point process has no short-term serial dependence. If the marked point process has no short-term correlations, the value of DET can be almost the same as the original marked point process even when the order of time windows is randomly shuffled.

The examples are depicted in Fig. 3. When $M = 3$, the data are CRSS data [Fig. 3(a)], where the order of three neighboring time windows is randomly shuffled. This random shuffling partially changes the distribution of inter-event intervals (IEIs), whereas the distribution of marks and the distances between time windows do not change. When $M = 1$, the surrogate data are equivalent to the original data [Fig. 3(b)]. When M is the total number of time windows N , the data are fully random shuffled surrogate (FRSS) data [Fig. 3(c)], where the order of all time windows is randomly shuffled. In the numerical experiments, we increased M gradually from 1 to N (the number of time windows) and estimated the minimum value of

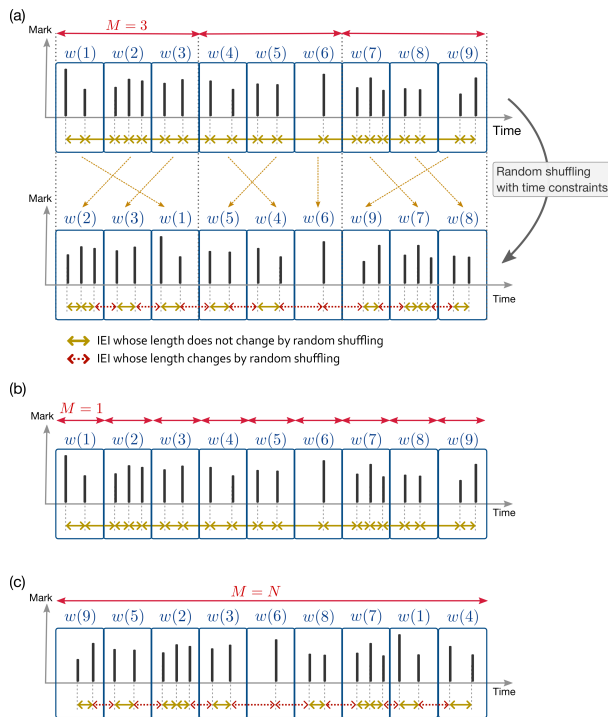


FIG. 3. Examples of constrained random shuffling. In case that (a) the number of randomly shuffled time windows $M = 3$ (CRSS data), (b) the number of randomly shuffled time windows $M = 1$, which corresponds to the original data, and (c) the number of randomly shuffled time windows $M = 9$ ($= N$) (FRSS data). The process of random shuffling is depicted in (a), while only the resultant marked point process data are depicted in (b) and (c).

M , represented by M_{\min} , such that the distribution of DETs obtained from CRSS data does not differ significantly from that obtained from FRSS data.

If M_{\min} is large, it indicates that similar patterns appear repeatedly in the marked point process data. In this case, the data can be predicted easily. On the other hand, if M_{\min} is small, the state of the marked point process data changes rapidly with time, and the data are not likely to contain similar patterns. In such cases, the data might not be predictable. In this sense, the obtained M_{\min} is fundamentally related to the time at which the deterministic features disappear and produces information about the relevant prediction limits.

In this study, M_{\min} is determined using the Welch t-test between the CRSS data and the FRSS data. The algorithm for calculating M_{\min} is as follows:

1. Calculate the values of DET of N_f FRSS data.
2. Set M to unity and γ to zero. γ represents the number of times that the null hypothesis is not rejected.
3. Calculate the values of DET of N_c CRSS data with the constraint parameter M .

4. Calculate a significance level α' using the value of M .
5. Perform the Welch t-test to check whether the mean values of DET calculated from CRSS data are significantly larger than those calculated from FRSS data. When the p -value is larger than α' , $\gamma \leftarrow \gamma + 1$. If $\gamma \geq \Theta_\gamma$, $M_{\min} = M$ and the algorithm terminates, where Θ_γ is the parameter of our method. When $\gamma < \Theta_\gamma$, go to step 6.
6. Increase the value of M by $M \leftarrow M + 1$ and go to step 3.

Figure 4 depicts the flow chart of the above-mentioned algorithm. In our method, statistical tests are performed repeatedly, which might create the problem of multiple comparisons. To counteract this problem, we determine the significance level α' in step 5 using the Šidák correction⁸ as follows:

$$\alpha' = 1 - (1 - \alpha)^{\frac{1}{M}}, \quad (6)$$

where α is the desired overall significance level and $\alpha = 0.01$ in this study.

We here note that if we employ the small shuffle surrogate data⁴ instead of the proposed CRSS, the surrogate data whose short-term correlation is destroyed can be obtained. Let $x(t)$ ($t = 1, 2, \dots, \tau$) be the t th data point in a regularly sampled time series data and $\mathcal{J}(t)$ be the temporal index of $x(t)$. The small shuffle surrogate data⁴ are generated using the following three steps:

- (i) $\tilde{\mathcal{J}}(t)$ is defined by $\tilde{\mathcal{J}}(t) = \mathcal{J}(t) + \xi_t$, where ξ_t is a Gaussian random number with appropriate mean and variance.
- (ii) By ranking $\tilde{\mathcal{J}}(t)$ ($t = 1, 2, \dots, \tau$) in ascending order of their values, the rank of $\tilde{\mathcal{J}}(t)$, which is described by $\text{rank}(\tilde{\mathcal{J}}(t))$ ($1 \leq \text{rank}(\tilde{\mathcal{J}}(t)) \leq \tau$), is adopted as the new index of the t th data point in the original time series.
- (iii) The small shuffle surrogate data are obtained by $x(\text{rank}(\tilde{\mathcal{J}}(t)))$ ($t = 1, 2, \dots, \tau$).

If this method is applied directly to the marked point process data by considering $x(t)$ to be the t th event, the surrogate data whose short-term correlation is destroyed can be generated. However, this method does not preserve the distances between time windows, which indicates that whenever the surrogate data are generated, the MSTM has to be repeatedly recalculated. These repeated calculations of the MSTM involve an extremely high computational cost. On the other hand, in our method, even if the order of time windows is randomly shuffled, the distribution of the distances between time windows holds before and after performing a random shuffle of M neighboring time windows. Thus, we need to calculate the distances between time windows only once.

III. EXPERIMENTAL SETTINGS

We used two types of dynamical systems to generate the marked point process data. The first dynamical system is the Lorenz equations¹⁰ described by

$$\begin{cases} \dot{x} = -\sigma x + \sigma y, \\ \dot{y} = -xz + rx - y, \\ \dot{z} = xy - \beta z, \end{cases} \quad (7)$$

where σ , β , and r are parameters. In this study, $\sigma = 10$, $\beta = \frac{8}{3}$, and $0 \leq r \leq 220$, and Eq. (7) is numerically integrated using

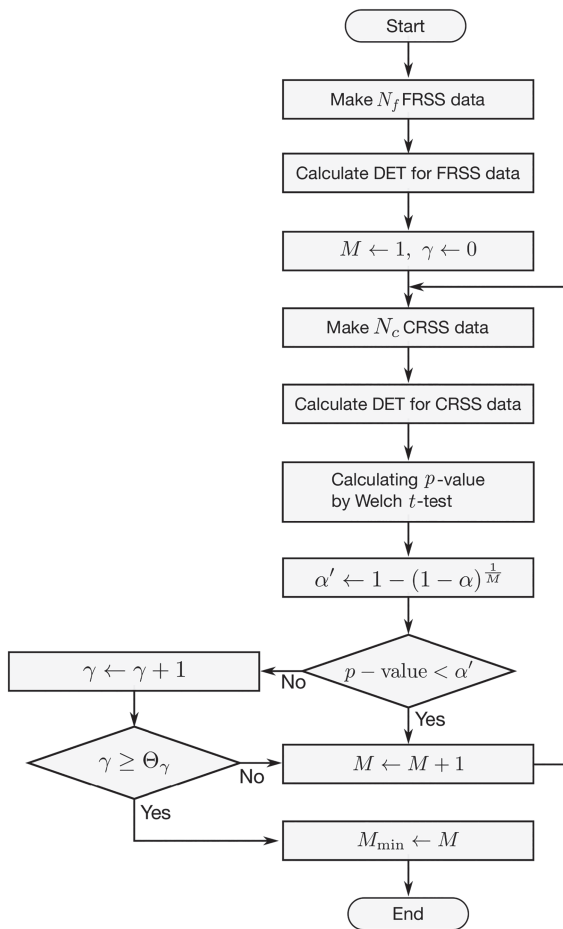


FIG. 4. Flow chart of the proposed method.

the fourth-order Runge–Kutta method with a step size of $h = 0.001$. The marked point process data are generated by extracting the local maxima from the time series of $x(t)$ of the Lorenz equations. The mark values of the marked point process are the local maxima, and the event timings are the times at which the local maxima occur.¹¹ Figures 5(a)–5(c) show examples of marked point process data generated from the Lorenz equations, where the red vertical lines correspond to events.

The second dynamical system is the neuron model proposed by Izhikevich,¹² which is described by

$$\begin{cases} \dot{v} = 0.04v^2 + 5v + 140 - u + I, \\ \dot{u} = a(bv - u), \end{cases} \quad (8)$$

$$\text{if } v \geq 30 \text{ [mV]} \quad \text{then} \quad \begin{cases} v \leftarrow c, \\ u \leftarrow u + d, \end{cases}$$

where v corresponds to the membrane potential, u corresponds to the recovery variable of a neuron, and I corresponds to the input to a neuron. In Eq. (8), a , b , c , and d are parameters. The variables v and u are reset to c and $u + d$, respectively, when the value of v reaches 30 mV. In this study, $a = 0.02$, $b = 0.2$, $c = -55$, and $I = 10$, unless otherwise stated. The Izhikevich simple neuron model can produce rich firing patterns exhibited by real biological neurons. The marked point process data are generated by extracting the value of $u(t)$ when the neuron fires. In this study, Eq. (8) is numerically integrated using the fourth-order Runge–Kutta method with a step size of $h = 0.025$, where the time when a neuron fires is calculated with the bisection method so that $|v(t) - 30| < 10^{-6}$. Figures 5(d)–5(g) show examples of marked point process data generated from the Izhikevich neuron model, where the red vertical lines correspond to events.

To compare the marked point process data generated from the deterministic dynamical systems with those generated from a stochastic process, we used the Poisson process with marks as the underlying model for point process data. The inter-event intervals of the Poisson process are independent and identically distributed with the exponential distribution whose probability density function is described by $Q(x) = \ell e^{-\ell x}$. The mark of each event is independent and identically distributed with the Gaussian distribution whose average and variance are zero and unity, respectively. We call this point process *marked Poisson process* in this paper. In the experiments, $\ell = 10$ and $\Lambda_t = \Lambda_p = 1$.

In this study, we first investigate the relationships between the values of DET and the parameter values of M that determine the strength of the random shuffle of time windows. Next, we investigate the relationship between the reciprocal of M_{\min} and the LLE. Positive correlation between the reciprocal of M_{\min} and the sum of positive Lyapunov exponents implies that M_{\min} obtained by our method can plausibly evaluate the prediction limit of the given marked point process data, because the Kolmogorov–Sinai entropy, which can be a measure of predictability, is estimated by the sum of positive Lyapunov exponents. In this paper, both the Lorenz system [Eq. (7)] and the Izhikevich neuron model [Eq. (8)] have one positive Lyapunov exponent, and consequently we investigate the relationship between LLE and the reciprocal of M_{\min} . The Lyapunov exponents of the Lorenz equations were estimated by the QR decomposition-based method.¹³ Let $(\lambda_1, \lambda_2, \lambda_3)$ be the Lyapunov spectrum of the Lorenz equations, where $\lambda_1 > \lambda_2 > \lambda_3$. Figure 6(a) depicts the relationships between the values of the parameter r and the corresponding Lyapunov exponents, where the value of r is increased from 0 to 220 with a step size of 0.5. On the other hand, in the Izhikevich neuron model, whenever the neuron fires, the variables v and u are reset. If the QR decomposition-based method is directly used, this after-spike resetting mechanism makes the estimated Lyapunov spectra inaccurate. We thus use the method for estimating the Lyapunov spectrum based on saltation matrices.¹⁴ Let (μ_1, μ_2) be the Lyapunov spectrum of the neuron model described in Eq. (8), where $\mu_1 > \mu_2$. Figure 6(b) depicts the relationships between the values of the parameter d and the corresponding Lyapunov exponents, where the value of d is increased from 0.86 to 0.95 with a step size of 0.0005.

In the numerical experiments, we also directly estimated the LLE based on the method proposed by Sato *et al.*¹⁵ and compared it

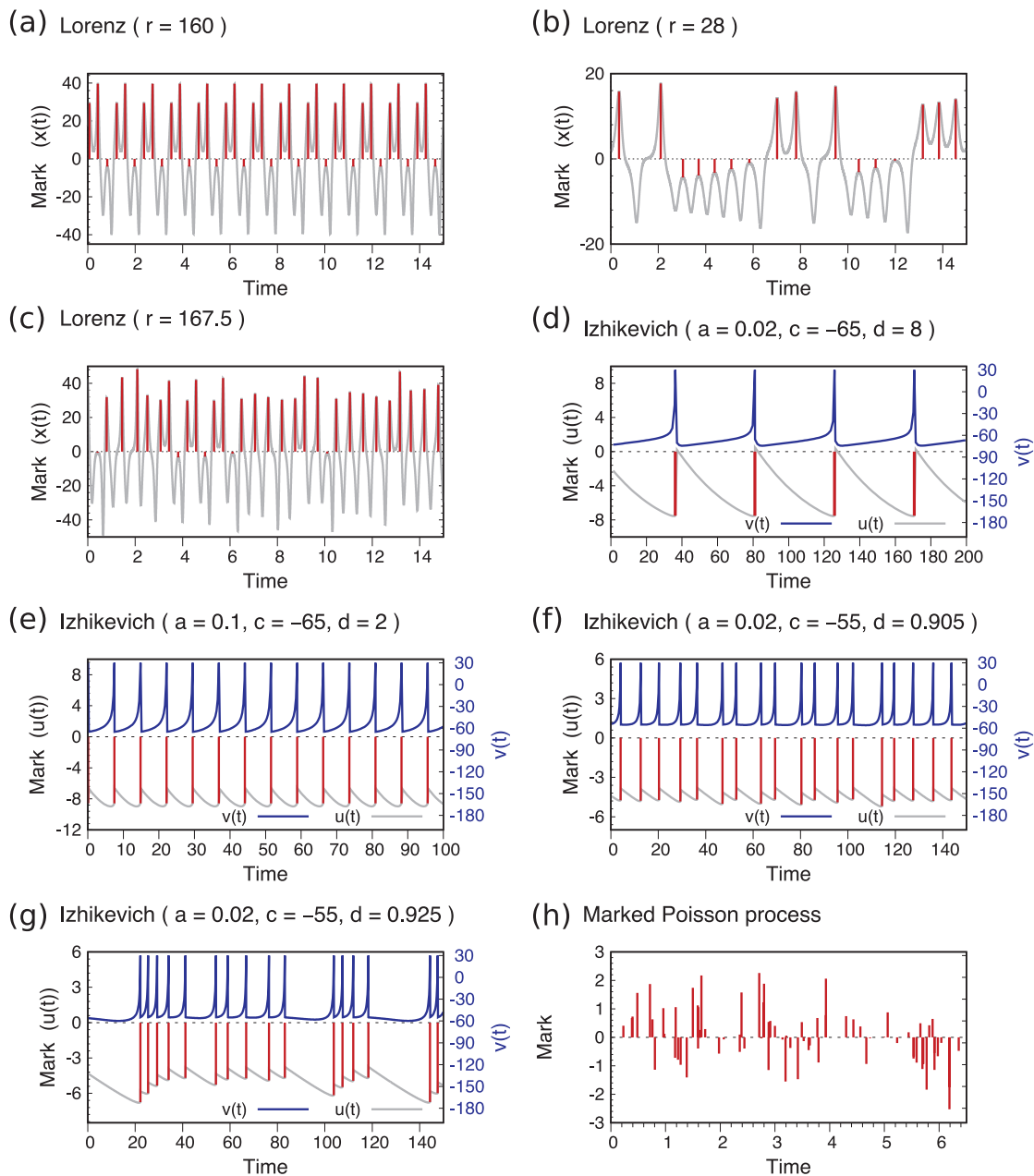


FIG. 5. Examples of marked point process data obtained from (a)–(c) the Lorenz equations, (d)–(g) the Izhikevich neuron model in Eq. (8), and (h) the marked Poisson process. (a) A periodic orbit ($r = 160$), (b) a chaotic orbit ($r = 28$) whose largest Lyapunov exponent $\lambda_1 \approx 0.90$, and (c) a chaotic orbit ($r = 167.5$) whose largest Lyapunov exponent $\lambda_1 \approx 1.95$. (d) A regular-spiking neuron ($a = 0.02, b = 0.2, c = -65, d = 8, l = 10$), (e) a fast-spiking neuron ($a = 0.1, b = 0.2, c = -65, d = 2, l = 10$), (f) a chaotic-spiking neuron ($a = 0.02, b = 0.2, c = -55, d = 0.905, l = 10$) whose largest Lyapunov exponent $\mu_1 \approx 0.04$, and (g) a chaotic-spiking neuron ($a = 0.02, b = 0.2, c = -55, d = 0.925, l = 10$) whose largest Lyapunov exponent $\mu_1 \approx 0.07$. (h) The marked Poisson process ($\ell = 10$). The red vertical lines show generated marked point process data. The gray lines in (a)–(g) show the original time series obtained from each of the dynamical systems. In (d)–(g), the blue lines show the time series of $v(t)$, and the gray lines show the time series of $u(t)$.

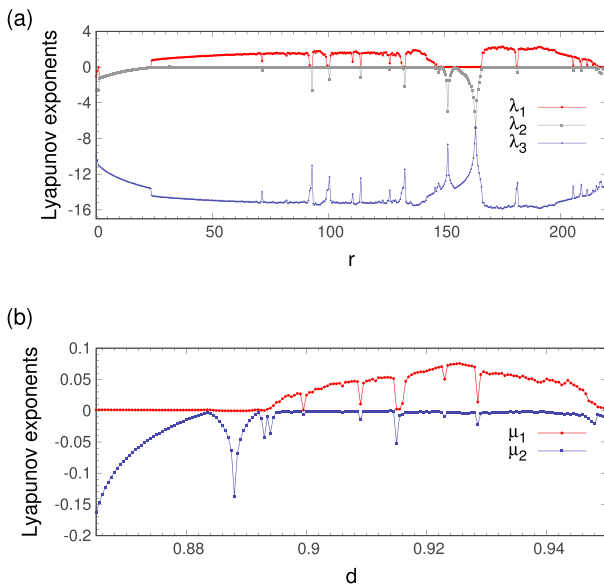


FIG. 6. Lyapunov spectra of (a) the Lorenz equations and (b) the Izhikevich neuron model. In (a), the red line represents λ_1 , the gray line represents λ_2 , and the blue line represents λ_3 . In (b), the red line represents μ_1 and the blue line represents μ_2 .

with M_{\min} obtained using our method. The LLE can be estimated by

$$\text{LLE}(s) \propto \frac{1}{s} \frac{1}{N-s} \sum_{i=1}^{N-s} \ln \frac{D(i+s, \kappa(i+s))}{D(i, \kappa(i))}, \quad (9)$$

where $\kappa(i)$ is the index of the nearest time window of the i th time window and $D(i+s, \kappa(i+s))$ represents the distance between the i th and $\kappa(i)$ th time windows after s steps, where $s = 1$ in this study.

In the experiment for investigating the relationship between

TABLE I. Parameter values used in the experiments.

Parameter	Lorenz equations [Eq. (7)]	Neuron model [Eq. (8)]
The number of time windows N	200	200
Length of time windows L	5	200
Weight of MSTM Λ_t in Eq. (2)	1	1
Weight of MSTM Λ_p in Eq. (2)	1	1
Step size of the Runge–Kutta method h	0.001	0.025
Threshold Θ_y	3	3
Significance level α	0.01	0.01
The number of FRSS data N_f	100	100
The number of CRSS data N_c	100	100
Step size s in Eq. (9)	1	1

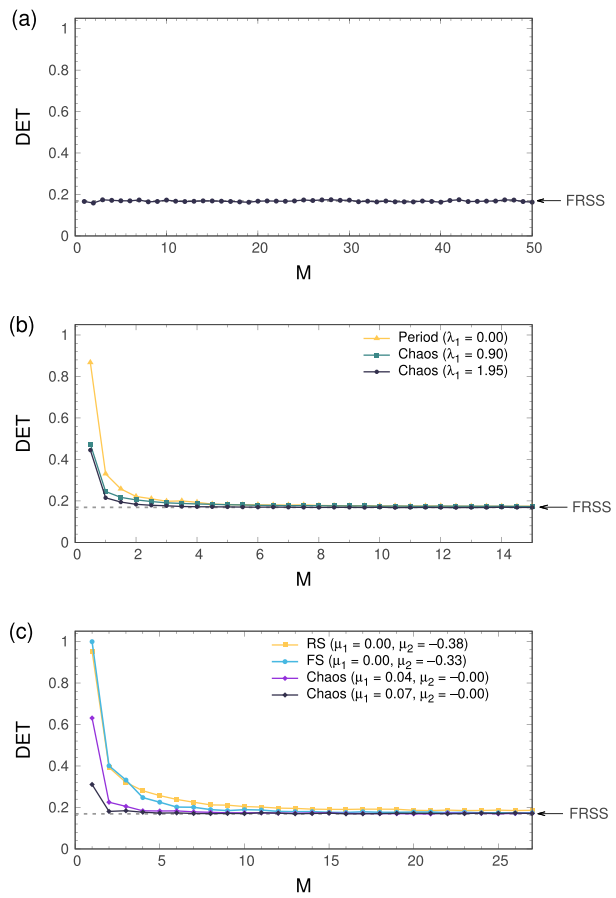


FIG. 7. DETs obtained from CRSS data. (a) DETs obtained by the marked Poisson process, where $L = 2$. (b) DETs for the marked point process data generated from the periodic Lorenz system ($r = 160$; represented by the yellow line) and the chaotic Lorenz systems ($r = 28$, $\lambda_1 \approx 0.90$ and $r = 167.5$, $\lambda_1 \approx 1.95$; represented by green and dark blue lines, respectively), where $L = 1$. (c) DETs for the marked point process data generated from the Izhikevich neuron model [Eq. (8)] whose patterns of spiking are regular-spiking ($a = 0.02$, $b = 0.2$, $c = -65$, $d = 8$; represented by the yellow line), fast-spiking ($a = 0.1$, $b = 0.2$, $c = -65$, $d = 2$; represented by the light blue line), and chaotic-spiking ($a = 0.02$, $b = 0.2$, $c = -55$, $d = 0.905$, $\mu_1 \approx 0.04$ and $a = 0.02$, $b = 0.2$, $c = -55$, $d = 0.925$, $\mu_1 \approx 0.07$; represented by the purple and dark blue lines, respectively), where $L = 200$. The horizontal dashed lines represent the DETs of FRSS data.

M_{\min} and the LLE, we generated 100 marked point process data with different initial values for each parameter set of the Lorenz equations and the Izhikevich neuron model. We then calculated the mean values of M_{\min} and the LLE. The value of r is increased from 25 to 220 with a step size of 1 in the Lorenz equations. In the Izhikevich neuron model, the value of d is increased from 0.89 to 0.95 with a step size of 0.001. In both dynamical systems, we use each of the values of r and d only when λ_1 and μ_1 are larger than 10^{-2} , because

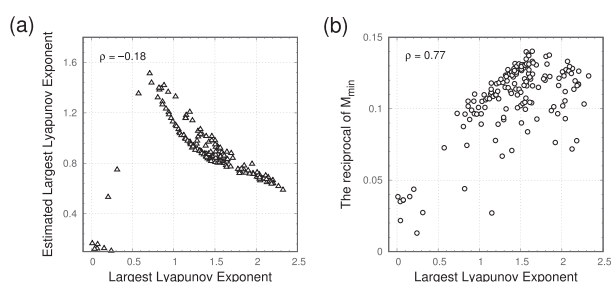


FIG. 8. Results for the Lorenz equations. (a) The relationship between the LLE estimated by Eq. (9) and the LLE. (b) The relationship between the reciprocal of M_{\min} and the LLE. The value of the correlation coefficient ρ is shown in each figure.

the recurrence plots become extremely sensitive to the length of time windows when the dynamical systems exhibit periodic behavior (see Appendix A for details). The parameter values used in these experiments are those listed in Table I, unless otherwise stated.

IV. RESULTS

First, we investigated the relationship between DET and the value of the parameter M . Figure 7(a) is the result of the marked Poisson process, where the inter-event intervals are described as identical and independently distributed random variables. The marked Poisson process has no short-term correlation. From Fig. 7(a), even when the value of M is small, there is no difference between the DET values of the FRSS data and the CRSS data. This result indicates that our method correctly detects the property that the marked Poisson process does not have a short-term correlation. Figure 7(b) shows the results for the Lorenz equations, where three values of r are used: when $r = 160$, the Lorenz system exhibits periodic behavior, and when $r = 28$ and 167.5 , the Lorenz system exhibits chaotic behavior, where the LLE of the Lorenz system with $r = 28$ is smaller than that with $r = 167.5$. From Fig. 7(b), when the

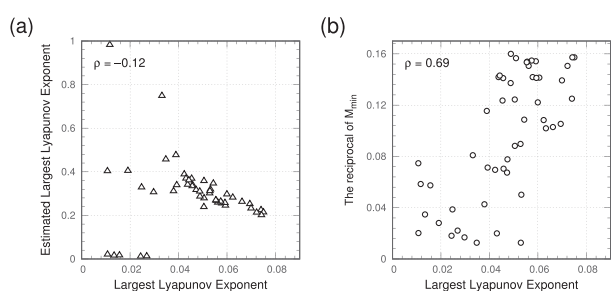


FIG. 9. Results for the Izhikevich neuron model. (a) The relationship between the LLE estimated by Eq. (9) and the LLE. (b) The relationship between the reciprocal of M_{\min} and the LLE. The value of the correlation coefficient ρ is shown in each figure.

marked point process data are generated from the periodic Lorenz system, the DET values gradually decrease with increasing value of M , compared with the results for the marked point processes generated from chaotic Lorenz systems. In addition, when the LLE is large, the decrease in the value of DET accelerates as the value of M increases.

The same applies to the results of the Izhikevich neuron model shown in Fig. 7(c), which shows the results for three types of behavior of neurons: regular-spiking (RS), fast-spiking (FS), and chaotic-spiking (Chaos) behavior. From Fig. 7(c), if the patterns of spiking of neurons are RS or FS, the DET values gradually decrease as the value of M increases. The speed of the decrease of the FS neuron is faster than that of the RS neuron, which appears to be proportional to the values of μ_2 . On the other hand, if the pattern of spiking exhibits chaotic behavior, the DET values quickly decrease, and the speed of the decrease is proportional to the LLE (μ_1).

Next, we compared M_{\min} and the LLE in the cases where $\lambda_1 > 10^{-2}$. Figures 8(a) and 9(a) show the relationships between the LLE estimated from the marked point process data by Eq. (9) and the LLE of the original dynamical systems. Figures 8(b) and 9(b) show the relationship between the reciprocal of M_{\min} and the LLE of the original dynamical systems. From Fig. 8(a), the LLE estimated from the marked point process data by Eq. (9) correlates positively with the LLE of the original dynamical systems when λ_1 is less than approximately 0.7. However, this positive correlation disappears when λ_1 becomes large, and the LLE estimated by Eq. (9) appears to correlate inversely with the LLE of the original dynamical systems.

This tendency can also be found in Fig. 9(a), which shows the results for the Izhikevich neuron model; the LLE estimated by Eq. (9) is slightly inversely proportional to the LLE of the original dynamical systems. These results indicate that Eq. (9) cannot evaluate the LLE precisely when the LLE of the original dynamical systems becomes large. On the other hand, from Figs. 8(b) and 9(b), the reciprocal of M_{\min} correlates positively with the LLE even when the LLE of the original dynamical systems takes large values, where the correlation coefficient ρ between the reciprocal of M_{\min} and the LLE is approximately 0.7 in both cases. These results imply that the prediction limits of the marked point process data generated from the Lorenz equations and the Izhikevich neuron model can be estimated by M_{\min} obtained via the proposed method.

V. DISCUSSION

In Sec. IV, we show that the prediction limits of the marked point process data generated from several dynamical systems can be estimated by M_{\min} and the reciprocal of M_{\min} is closely related to the LLE by numerical simulations. As discussed in Ref. 8, the distribution of the diagonal lines is directly related to the second order Rényi entropy, but not to the LLE. Consequently, the DET values might not be directly related to the LLE, because the DET value is proportional to the average diagonal line length. However, our results imply the possibility that the prediction limit estimated by using the DET values, namely, the diagonal line length, can be indirectly related to the LLE. Although the results obtained by the numerical simulations [Figs. 8(b) and 9(b)] clarify that the reciprocal of M_{\min} and the LLE of

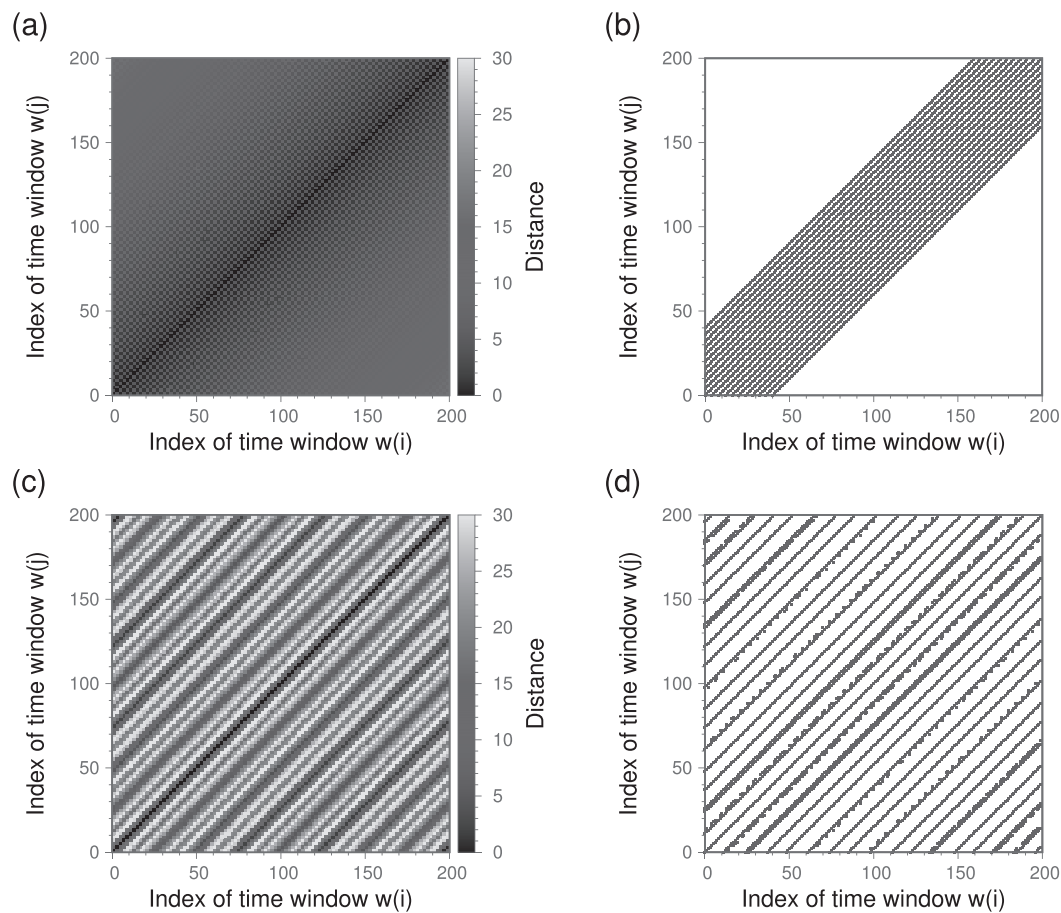


FIG. 10. Distance matrices and recurrence plots when the parameter $r = 146.9$, $\sigma = 10$, $b = \frac{8}{3}$ in the Lorenz equations. (a) The distance matrix when $L = 15.67$. (b) The recurrence plot corresponding to (a). (c) The distance matrix values when $L = 15.77$. (d) The recurrence plot corresponding to (c).

the original dynamical systems have a positive correlation, the theoretical evidence thus far is insufficient. Therefore, it is important to obtain an analytical relationship among M_{\min} , the LLE, and the sum of positive Lyapunov exponents in the future.

Moreover, it is necessary to discuss the range of applicability of the proposed method. Our method primarily aims at estimating the prediction limits of the marked point process data generated from nonlinear dynamical systems, which implicitly assumes that IEIs and marks of the marked point processes data include the information of the dynamical systems that generate them. Thus, in a realistic situation, we need to assess whether the marked point processes are stochastic or not before using the proposed method. Although the proposed method uses the CRSS data to determine whether the observed marked point process has no short-term serial dependence or not, it might not be sufficient to distinguish the marked point processes generated from nonlinear dynamical systems and those from stochastic processes. Because if an observed marked point process

consists of colored noise and contains significant short-term correlations, the null hypothesis employed in the CRSS method can be false.

In the case where the information of original dynamical systems is encoded into sequences of the IEIs, several methods based on surrogate data can be useful to assess whether the observed marked point process data are stochastic or not.^{1,16–18} However, even if the sequences of IEIs are not stochastic, there is a possibility that the marks of events can be described by stochastic variables. In such a case, in the proposed method, it might be needed to adjust the weight value of marks Λ_p in Eq. (2) so that the weight value of marks (Λ_p) is smaller than that of IEIs (Λ_t). In addition to these cases, we should note the possibilities such as (i) a sequence of marks is not stochastic, but a sequence of IEIs is stochastic; (ii) both sequences of IEIs and marks are stochastic; and (iii) both sequences of IEIs and marks are not stochastic, but they are independently generated from distinct dynamical systems. In this paper, we mainly focused on the

case where both sequences of marks and IEIs include information about the identical nonlinear dynamical system that generates them. Consequently, the proposed method lacks the considerations of cases (i)–(iii). In this sense, to avoid misinterpretations of observed data, it is not appropriate to apply our method to the data with properties (i)–(iii), and thereby it is also important to examine these possibilities before applying the method proposed in this paper to real data.

VI. CONCLUSION

We proposed a method for estimating the prediction limits of the marked point process data generated from nonlinear dynamical systems, using constrained random shuffle surrogate data. Using two dynamical systems, namely, the Lorenz equations and the Izhikevich neuron model, we numerically investigated the validity of the proposed method. From the results, we concluded that a positive correlation exists between the LLE and the reciprocal of M_{\min} , which is the time when the short-term correlation disappears. These results of numerical simulations show that the prediction limit of the marked point process can be detected by the proposed method, although the theoretical evidence is not yet sufficient.

In this study, we mainly investigated several deterministic nonlinear dynamical systems and fully stochastic point processes. However, it is also important to evaluate how the proposed method works for various types of stochastic dynamical systems, including the case where the marked point process whose IEIs (marks) have some kind of dynamics but the marks (IEIs) have no dynamics. Then, it is an important avenue for future research studies to investigate marked point processes generated from such systems.

ACKNOWLEDGMENTS

This research is partially supported by the JSPS Grants-in-Aid for Scientific Research (Nos. JP15KT0112, JP17K00348, JP17K00348, JP18K18125, and JP20H00596).

APPENDIX A: UNDESIRABLE EFFECTS OF TIME WINDOW LENGTHS

We also investigated the effect of the time window lengths and observed that they occasionally produce undesirable effects on the distances calculated by the MSTM. Figure 10 presents examples of distance matrices whose (i, j) th element (pixel) is the distance between the time windows $w(i)$ and $w(j)$ in a marked point process generated from a time series of x in Lorenz equations with strong periodicity. Both distance matrices [Figs. 10(a) and 10(c)] are calculated from the same marked point process by using the common parameter values of the MSTM. The only difference between them is the time window length L : $L = 15.67$ in Fig. 10(a) and $L = 15.77$ in Fig. 10(c). Despite the small difference in the time window length, the distance matrices, shown in Fig. 10, are observed to be completely different. The values of DET obtained from these distance matrices are also quite different. This difference due to the window length becomes prominent when the largest Lyapunov exponent is small. From these results, we consider the case in which $\lambda_1 > 10^{-2}$ in our experiments.

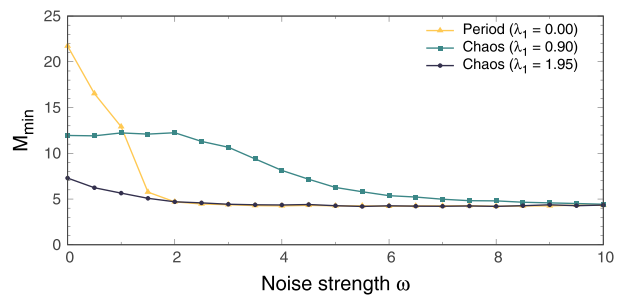


FIG. 11. Relationship between M_{\min} and ω for the Lorenz equations. The values of M_{\min} are calculated from the marked point process data generated from the Lorenz system with dynamical noise [Eq. (B1)], whose behavior is periodic ($r = 160$; represented by the yellow line) and chaotic ($r = 28$, $\lambda_1 \approx 0.90$ and $r = 167.5$, $\lambda_1 \approx 1.95$; represented by green and dark blue lines, respectively), where $L = 2$.

APPENDIX B: EFFECT OF DYNAMICAL NOISE

We further investigated the effect of dynamical noise on M_{\min} , where the dynamical noise was added to the Lorenz equations [Eq. (7)] and the Izhikevich neuron model [Eq. (8)], as follows. The Lorenz equations with dynamical noise are described by

$$\begin{cases} \dot{x} = -\sigma x + \sigma y + \omega \xi_x, \\ \dot{y} = -xz + rx - y + \omega \xi_y, \\ \dot{z} = xy - \beta z + \omega \xi_z, \end{cases} \quad (\text{B1})$$

where $\omega \xi_x$, $\omega \xi_y$, and $\omega \xi_z$ are the noise terms. The Euler–Maruyama method with the step size 0.0001 was used to obtain the numerical solutions of Eq. (B1). The Izhikevich neuron model with dynamical

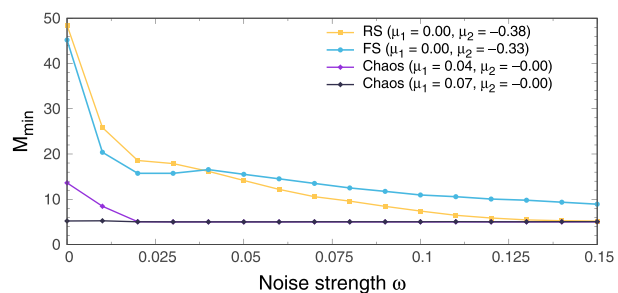


FIG. 12. Relationship between M_{\min} and ω for the Izhikevich neuron model. The values of M_{\min} are calculated from the marked point process data generated from the Izhikevich neuron model with dynamical noise [Eq. (B2)], whose patterns of spiking are regular-spiking ($a = 0.02$, $b = 0.2$, $c = -65$, $d = 8$; represented by the yellow line), fast-spiking ($a = 0.1$, $b = 0.2$, $c = -65$, $d = 2$; represented by the light blue line), and chaotic-spiking ($a = 0.02$, $b = 0.2$, $c = -55$, $d = 0.905$, $\mu_1 \approx 0.04$ and $a = 0.02$, $b = 0.2$, $c = -55$, $d = 0.925$, $\mu_1 \approx 0.07$; represented by the purple and dark blue lines, respectively), where $L = 100$.

noise is described by

$$\begin{cases} \dot{v} = 0.04v^2 + 5v + 140 - u + I + \omega\xi_v, \\ \dot{u} = a(bv - u), \end{cases}$$

$$\text{if } v \geq 30 \text{ [mv]} \text{ then } \begin{cases} v \leftarrow c, \\ u \leftarrow u + d, \end{cases} \quad (\text{B2})$$

where $\omega\xi_v$ is the noise term. The Euler–Maruyama method with the step size 0.0025 is used to obtain the numerical solutions of Eq. (B2). In Eqs. (B1) and (B2), ω is the strength of noise. In this paper, ξ_x , ξ_y , ξ_z , and ξ_v represent white Gaussian noise whose average and variance are zero and unity, respectively.

For the Izhikevich neuron model, the marked point process data are generated by extracting the value of $u(t)$ when the neuron fires. In contrast, for the Lorenz equations, the marked point process data are generated by the following procedure: (i) a time series of x obtained by the Euler–Maruyama method with the step size 0.0001 is resampled every 10 steps and then the time series $x(t)$ is generated. (ii) The local maxima are detected by checking whether $x(t)$ satisfies both $x(t) > x(t - \tilde{T})$ and $x(t) > x(t + \tilde{T})$ for all t and \tilde{T} ($1 \leq \tilde{T} \leq T$), where $T = 0.25$ in this paper. (iii) The marked point process data are generated by extracting these local maxima. Steps (i) and (ii) were incorporated to reduce the number of events included in the marked point process data. If the local maxima are simply extracted from the time series obtained by the Euler–Maruyama method, the number of events in the marked point process data generated from Eq. (B1) considerably increases when $\omega > 0$, which simultaneously increases the computational cost of the MSTM. Steps (i) and (ii) can counteract this effect.

Figures 11 and 12 show the relationship between M_{\min} and the strength of noise ω , where the parameter values used in the experiments are the same as those listed in Table I except for the values of L ; $L = 2$ in the Lorenz equations and $L = 100$ in the Izhikevich neuron model. In addition, the Euler–Maruyama method was used instead of the Runge–Kutta method.

From Figs. 11 and 12, it can be seen that the value of M_{\min} gradually decreases with increasing noise strength. In addition, Fig. 11 indicates that M_{\min} obtained from the periodic system can be potentially smaller than that obtained from the chaotic system with dynamical noise. These results not only show that the dynamical noise decreases M_{\min} , namely, the estimated prediction limit, but also imply that the estimated prediction limit might not reflect the true predictability of the underlying dynamical systems in the marked point processes.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

- Y. Hirata, Y. Katori, H. Shimokawa, H. Suzuki, T. A. Blenkinsop, E. J. Lang, and K. Aihara, “Testing a neural coding hypothesis using surrogate data,” *J. Neurosci. Methods* **172**(2), 312–322 (2008).
- Y. Hirata and K. Aihara, “Timing matters in foreign exchange markets,” *Physica A* **391**(3), 760–766 (2012).
- J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer, “Testing for nonlinearity in time series: The method of surrogate data,” *Physica D* **58**(1), 77–94 (1992).
- T. Nakamura and M. Small, “Small-shuffle surrogate data: Testing for dynamics in fluctuating data with trends,” *Phys. Rev. E* **72**(5), 056216 (2005).
- Y. Hirata and M. Shiro, “Detecting nonlinear stochastic systems using two independent hypothesis tests,” *Phys. Rev. E* **100**(2), 022203 (2019).
- G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska, “Surrogate data for hypothesis testing of physical systems,” *Phys. Rep.* **748**, 1–60 (2018).
- S. Suzuki, Y. Hirata, and K. Aihara, “Definition of distance for marked point process data and its application to recurrence plot-based analysis of exchange tick data of foreign currencies,” *Int. J. Bifurcat. Chaos* **20**(11), 3699–3708 (2010).
- N. Marwan, M. Carmen Romano, M. Thiel, and J. Kurths, “Recurrence plots for the analysis of complex systems,” *Phys. Rep.* **438**(5), 237–329 (2007).
- Z. Šidák, “Rectangular confidence regions for the means of multivariate normal distributions,” *J. Am. Stat. Assoc.* **62**(318), 626–633 (1967).
- E. N. Lorenz, “Deterministic nonperiodic flow,” *J. Atmos. Sci.* **20**(2), 130–141 (1963).
- N. Yabuta and T. Ikeguchi, “Prediction of high-dimensional multivariate information as an amplitude-event dynamical system,” in *Proceedings of International Symposium on Nonlinear Theory and Its Applications (NOLTA2007)* (IEICE, 2007), pp. 188–191.
- E. M. Izhikevich, “Simple model of spiking neurons,” *IEEE Trans. Neural Networks* **14**(6), 1569–1572 (2003).
- H. F. von Bremen, F. E. Udawadia, and W. Proskurowski, “An efficient QR based method for the computation of Lyapunov exponents,” *Physica D* **101**(1–2), 1–16 (1997).
- F. Bizzarri, A. Brambilla, and G. S. Gajani, “Lyapunov exponents computation for hybrid neurons,” *J. Comput. Neurosci.* **35**, 201–212 (2013).
- S. Sato, M. Sano, and Y. Sawada, “Practical methods of measuring the generalized dimension and the largest Lyapunov exponent in high dimensional chaotic systems,” *Prog. Theor. Phys.* **77**, 1–5 (1987).
- Y. Minghao *et al.*, “Identifying distinct stochastic dynamics from chaos: A study on multimodal neural firing patterns,” *Int. J. Bifurcat. Chaos* **19**, 453–485 (2009).
- S. Grün, “Data-driven significance estimation for precise spike correlation,” *J. Neurophysiol.* **101**, 1126–1140 (2009).
- L. Ricci, M. Castelluzzo, L. Minati, and A. Perinelli, “Generation of surrogate event sequences via joint distribution of successive inter-event intervals,” *Chaos* **29**, 121102 (2019).

Paper

Chaos MIMO system with efficient use of information of bits

Ryo Yamazaki^{1a)}, Yutaka Shimada², and Tohru Ikeguchi^{1,3}

¹ *Department of Management Science, Graduate School of Engineering, Tokyo University of Science, Niijuku 6-3-1, Katsushika-ku, Tokyo 125-8585, Japan*

² *Department of Information and Computer Sciences, Graduate School of Science and Engineering, Saitama University, Shimo-Okubo 255, Sakura-ku, Saitama-shi, Saitama 338-8570, Japan*

³ *Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, Niijuku 6-3-1, Katsushika-ku, Tokyo 125-8585, Japan*

^{a)} *yamazaki@hisenkei.net*

Received January 31, 2020; Revised May 25, 2020; Published October 1, 2020

Abstract: The multiple-input multiple-output (MIMO) system is one of the wireless communication methods that use multiple transmit and receive antennas. To ensure security on the physical layer and also to enhance channel coding efficiency, a chaos MIMO (C-MIMO) system was previously proposed. In this system, a chaotic dynamical system is used for modulation. In this paper, we revealed that the original C-MIMO system does not effectively use the information of bits that are used for modulation, which results in a difficulty in distinguishing encrypted symbols. To solve this issue, we propose a new modulation method for the C-MIMO system. We evaluated the performance of the proposed C-MIMO system and showed that the proposed C-MIMO system significantly improves block error rates.

Key Words: chaos communication, MIMO, Bernoulli shift map

1. Introduction

Owing to the widespread use of the Internet, the development of high capacity and secure communication systems is required. To deal with this problem, it is acknowledged that the multiple-input multiple-output (MIMO) system [1] is effective owing to its large channel capacity. The MIMO system has already played an essential role in the fourth generation of cellular mobile communications (4G) and the fifth generation of cellular mobile communications (5G). Although the current MIMO system provides security for the layers higher than the physical layer, the security of the physical layer is not taken into consideration. However, as the number of users increases, the security of the current MIMO system is insufficient and encryption in the physical layer should be considered.

To ensure the security of the physical layer, the chaos-based communication method [2] is effective.

One of the well-known chaos-based communication methods is chaotic code division multiple access (C-CDMA) [3], where a chaos-based modulation is incorporated into the code division multiple access (CDMA); the C-CDMA ensures the security of the physical layer. However, the CDMA was not included in the 4G and 5G standards. Conversely, Okamoto *et al.* proposed a chaos MIMO (C-MIMO) system that ensures the security of the physical layer and also obtains channel coding gain by introducing a chaotic dynamical system in the signal modulation part [4–6]. The C-MIMO system is one of the common key cryptosystems, which uses the feature of chaotic dynamics such as a sensitive dependence on initial conditions. Although the original C-MIMO system shows high performance, it is still important to reveal why the chaotic dynamics leads to good performance from the perspective of nonlinear dynamical systems.

In this study, we numerically analyzed the dynamics of the chaotic system used in the modulation part of the original C-MIMO system [4–6] and clarified that there is a possibility that the original C-MIMO system [4–6] does not take full advantage of the information of bits used for signal modulation. Then, we further proposed a chaos-based modulation method for the modulation part in [6] and showed that the proposed C-MIMO system improves bit error rate (BER) and block error rate (BLER) performances, because the proposed C-MIMO system can effectively use the information of bits.

2. Chaos MIMO system

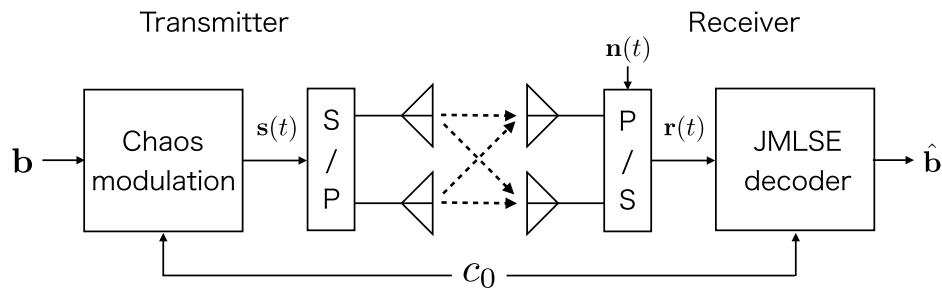


Fig. 1. Original chaos multi-input multi-output (C-MIMO) system [4–6].

In this section, we review a configuration of the original C-MIMO system [4–6]. Figure 1 shows the configuration of the original C-MIMO system.

Let N_t be the number of transmit antennas and N_r be the number of receive antennas. Figure 1 shows an example when $N_t = N_r = 2$. Let \mathbf{b} be a transmit block which is given by

$$\begin{aligned} \mathbf{b} &= [b_0, b_1, \dots, b_{L-1}], \\ b_m &\in \{0, 1\}, (m = 0, 1, \dots, L-1), \end{aligned} \quad (1)$$

where L is the number of symbols.

Let $\mathbf{s}(t)$ be the t th MIMO transmit vector which is given by

$$\mathbf{s}(t) = [s_1(t), \dots, s_{N_t}(t)]^T, \quad (2)$$

where $s_i(t)$ ($i = 1, \dots, N_t$) corresponds to the t th symbol transmitted from the i th transmit antenna and T represents the transpose. In Eq. (2), t represents the temporal order with which $\mathbf{s}(t)$ is transmitted from the transmit antennas. The transmitted symbol $s_i(t)$ is generated from the transmit block \mathbf{b} by the chaos-based modulation, which is described later in this section. Let \mathbf{S}_B be the MIMO transmit block consisting of B MIMO transmit vectors defined by

$$\mathbf{S}_B = [\mathbf{s}(0), \dots, \mathbf{s}(B-1)]. \quad (3)$$

Note that $L = N_t B$. Then, the t th MIMO receive vector is given by

$$\mathbf{r}(t) = \mathbf{H}(t)\mathbf{s}(t) + \mathbf{n}(t), \quad (4)$$

where $\mathbf{n}(t)$ is the t th noise vector whose elements obey the zero-mean Gaussian distribution. The matrix $\mathbf{H}(t)$ is the t th channel matrix given by

$$\mathbf{H}(t) = \begin{bmatrix} h_{11}(t) & \cdots & h_{1N_t}(t) \\ \vdots & \ddots & \vdots \\ h_{N_r,1}(t) & \cdots & h_{N_r,N_t}(t) \end{bmatrix}, \quad (5)$$

where $h_{ij}(t)$ is the t th channel component between the i th transmit antenna and the j th receive antenna. Values of $h_{ij}(t)$ follow antenna and symbol independent and identically distributed (i.i.d.) flat Rayleigh fading. In the receiver, the transmit block \mathbf{b} is estimated from $\mathbf{r}(t)$.

In the original C-MIMO system, a common key c_0 ($0 < \text{Re}[c_0] < 1$ and $0 < \text{Im}[c_0] < 1$) is randomly generated and shared between the transmitter and the receiver in advance. The value of c_0 is used to generate the initial values of the chaotic dynamical systems. Figure 2 shows a configuration of the chaos modulation part [4–6].

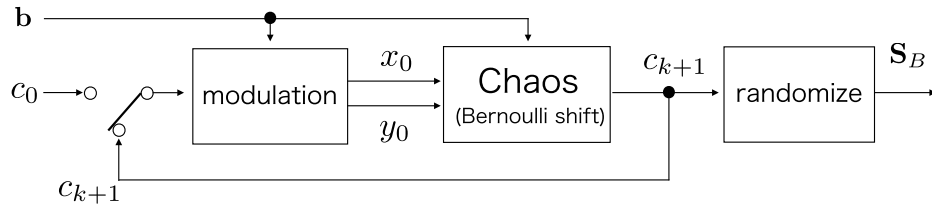


Fig. 2. Configuration of the chaos modulation part in the original C-MIMO system [4–6].

The original C-MIMO system generates the MIMO transmit block \mathbf{S}_B by using a transmit block \mathbf{b} and the k th element symbol represented by c_k ($k = 1, \dots, L$). The element symbol c_k is generated from c_0 by a chaos-based modulation described in the following procedure.

- (I) Set k to 0.
- (II) Two initial values of chaotic dynamical systems, x_0 and y_0 , are generated from the element symbol c_k and the transmit block \mathbf{b} as follows:

$$x_0 = \begin{cases} \text{Re}[c_k] & (b_k = 0), \\ 1 - \text{Re}[c_k] & (b_k = 1 \text{ and } \text{Re}[c_k] > 1/2), \\ \text{Re}[c_k] + 1/2 & (b_k = 1 \text{ and } \text{Re}[c_k] \leq 1/2), \end{cases} \quad (6)$$

and

$$y_0 = \begin{cases} \text{Im}[c_k] & (b_{(k+1) \bmod L} = 0), \\ 1 - \text{Im}[c_k] & (b_{(k+1) \bmod L} = 1 \text{ and } \text{Im}[c_k] > 1/2), \\ \text{Im}[c_k] + 1/2 & (b_{(k+1) \bmod L} = 1 \text{ and } \text{Im}[c_k] \leq 1/2). \end{cases} \quad (7)$$

In Eqs. (6) and (7), when $k = 0$, the common key c_0 is used.

- (III) Two chaotic series are generated by the Bernoulli shift map:

$$\begin{cases} x_{l+1} = 2x_l \bmod 1, \\ y_{l+1} = 2y_l \bmod 1, \end{cases} \quad (8)$$

where x_0 and y_0 are used as the initial values. After iterating Eq. (8) n times, the $(k+1)$ th element symbol c_{k+1} is calculated by

$$\begin{cases} \text{Re}[c_{k+1}] = x_{n+b_{(k+L/2) \bmod L}}, \\ \text{Im}[c_{k+1}] = y_{n+b_{\{(k+1)+L/2\} \bmod L}}, \end{cases} \quad (9)$$

where n is set to 100, as described in [6].

(IV) By using c_{k+1} , the transmitted random Gaussian symbol s_{k+1} is generated using the Box-Muller method [7] as follows:

$$s_{k+1} = \sqrt{-\log(u_{x,k+1})} \{ \cos(2\pi u_{y,k+1}) + j \sin(2\pi u_{y,k+1}) \}, \quad (10)$$

where j is the imaginary unit, $u_{x,k+1}$ and $u_{y,k+1}$ are defined by

$$\begin{aligned} u_{x,k+1} &= \frac{1}{\pi} \arccos[\cos\{37\pi(\operatorname{Re}[c_{k+1}] + \operatorname{Im}[c_{k+1}])\}], \\ u_{y,k+1} &= \frac{1}{\pi} \arcsin[\sin\{43\pi(\operatorname{Re}[c_{k+1}] - \operatorname{Im}[c_{k+1}])\}] + \frac{1}{2}. \end{aligned} \quad (11)$$

(V) After repeating the steps (II)–(IV), the MIMO transmit block \mathbf{S}_B is defined as

$$\mathbf{S}_B = \begin{bmatrix} s_1 & s_{N_t+1} & \cdots & s_{(B-1)N_t+1} \\ \vdots & & \ddots & \vdots \\ s_{N_t} & s_{2N_t} & \cdots & s_{BN_t} \end{bmatrix}, \quad (12)$$

where the symbols in the i th column are transmitted from the i th antenna. Therefore, the MIMO transmit vector $\mathbf{s}(t)$ in Eq. (2) is given by

$$\mathbf{s}(t) = [s_{tN_t+1}, \dots, s_{(t+1)N_t}]^T. \quad (13)$$

In the receiver, the joint maximum likelihood sequence estimation (JMLSE) [5] is performed to decode \mathbf{S}_B . The decoded block $\hat{\mathbf{b}}$ is given by

$$\hat{\mathbf{b}} = \arg \min_{\mathbf{b} \in \Omega} \sum_{t=0}^{B-1} \|\mathbf{r}(t) - \mathbf{H}(t)\hat{\mathbf{s}}(\mathbf{b}, c_0)\|^2, \quad (14)$$

where Ω is a set of all possible bit sequences of length L and $\hat{\mathbf{s}}(\mathbf{b}, c_0)$ is a MIMO transmit vector generated from the block \mathbf{b} and common key c_0 in the receiver side by the chaos-based modulation described in the above-mentioned steps (I)–(V). The set of possible bit sequences Ω and the vectors $\hat{\mathbf{s}}(\mathbf{b}, c_0)$ ($\mathbf{b} \in \Omega$) are generated in the receiver in advance.

In numerical simulations, when the double-precision floating-point number is adopted, the chaotic series generated by Eq. (8) converges to zero with a finite number of iterations [9]. To avoid this zero convergence, the modulo operator $(\bmod 1)$ in Eq. (8) is replaced by $\bmod(1 - 10^{-11})$, similar to [10].

In the original C-MIMO system, even if an eavesdropper estimates a key close to the common key c_0 , the communication is still secure except when the eavesdropper estimates the common key whose squared error is smaller than 10^{-26} [8]. The value 10^{-26} is significantly small; therefore, the security is considered to be sufficient.

3. Issues in the original C-MIMO system

In this section, we discuss the case where $N_t = 2$, namely $L = 2B$, for the sake of simplicity, but the same discussion is valid for the case where $N_t > 2$.

The following two issues exist in the original C-MIMO system:

- I. Equations (6) and (7) have no effect when $0 < \operatorname{Re}[c_k] < 0.5$ and $0 < \operatorname{Im}[c_k] < 0.5$.
- II. Although \mathbf{b} is composed of L bits, s_{k+1} is generated consisting of only four bits of the total L bits, namely b_k , b_{k+1} , b_{k+B} , and b_{k+B+1} . This use of a small number of bits in \mathbf{b} reduces possible values of s_{k+1} , which might lead to poor performance.

To resolve issue I, in Fig. 3, we show the flow for generating the transmitted symbol s_{k+1} from the element symbol c_k . We first present an analysis of Eqs. (6) and (8).

First, the initial values of the chaotic dynamical systems x_0 are generated from the element symbol c_k according to Eq. (6), depending on the value of b_k . After generating x_0 , x_0 is mapped to x_1 by the

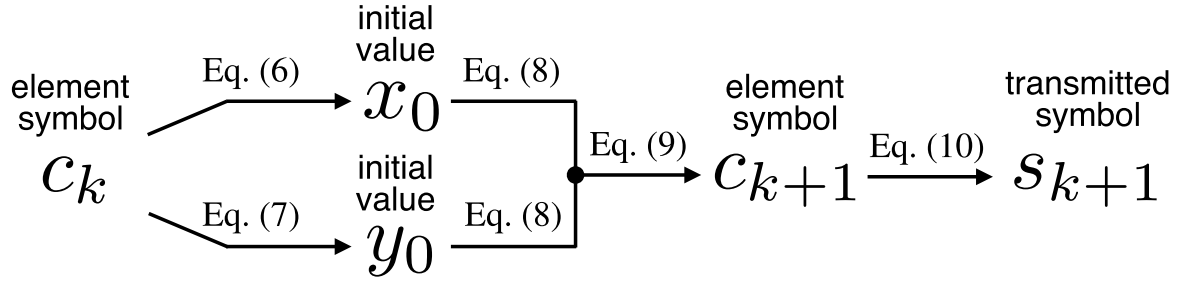


Fig. 3. Schematic diagram demonstrating the generation of the transmitted symbol s_{k+1} from the element symbol c_k in the original C-MIMO system.

Bernoulli shift map (Eq. (8)). If the bit b_k affects s_{k+1} , the value of x_1 must vary when the value of bit b_k varies.

However, regardless of whether $b_k = 0$ or $b_k = 1$, the values of x_1 are the same when $0 < \text{Re}[c_k] < 0.5$. The following reason causes the values of x_1 to remain the same: if $b_k = 0$, the value of x_1 is $2 \text{Re}[c_k] \bmod 1$. Conversely, if $b_k = 1$, the value of x_1 is $\{2 \text{Re}[c_k] + 1\} \bmod 1 = 2 \text{Re}[c_k] \bmod 1$, which is identical to the value when $b_k = 0$. This is because the modulo operation is defined by $\alpha \bmod \beta \equiv \alpha - \beta \lfloor \frac{\alpha}{\beta} \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. This indicates that Eq. (6) has no effect on s_{k+1} when $0 < \text{Re}[c_k] < 0.5$.

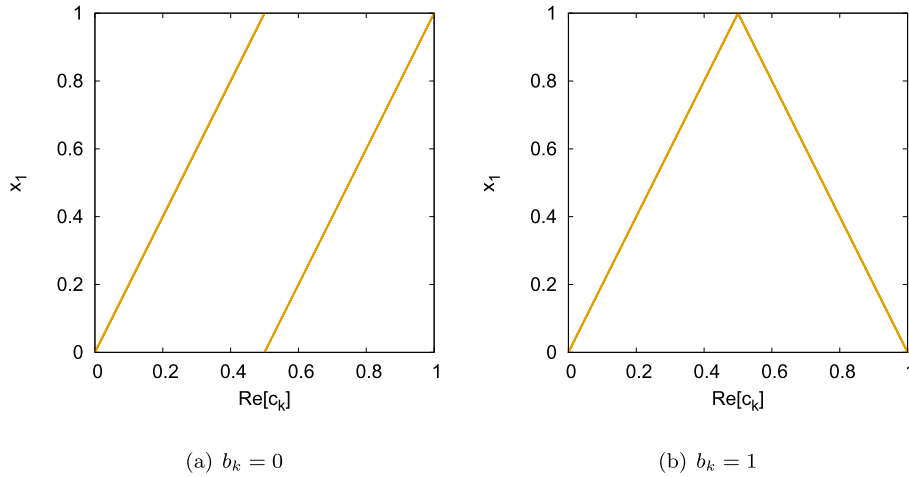


Fig. 4. Relation between $\text{Re}[c_k]$ and x_1 for (a) $b_k = 0$ and (b) $b_k = 1$.

To confirm the above discussion that Eq. (6) has no effect on s_{k+1} when $0 < \text{Re}[c_k] < 0.5$, we plotted the value of x_1 as a function of $\text{Re}[c_k]$ in Fig. 4. From Fig. 4, it can be observed that although the value of b_k is different, these two relations are equivalent when $0 < \text{Re}[c_k] < 0.5$. Thus, regardless of whether $b_k = 0$ or $b_k = 1$, the values of x_1 are the same when $0 < \text{Re}[c_k] < 0.5$.

It should be noted that the shapes of these two return maps are not important (Fig. 4(a) indicates a Bernoulli shift map and Fig. 4(b) demonstrates a tent map). The same discussion also applies to Eq. (7) when $0 < \text{Im}[c_k] < 0.5$.

For issue II, we first show the constellation diagrams of the original C-MIMO system when the MIMO transmit block length $B = 4$ and the number of transmit and receive antennas $N_t = N_r = 2$. In Fig. 5, we generated and plotted $2^L = 2^8$ symbols (s_{k+1}), which are all possible bit patterns of the transmit block \mathbf{b} . Even though the maximum number of symbols is 2^8 , the number of symbols in Fig. 5(a) is only 16, because certain symbols take exactly the same value. Thus, it is difficult to distinguish one symbol from another. The same discussion applies to Fig. 5(b).

When generating the symbol s_k , the original C-MIMO system [4–6] uses the element symbol c_k and the following four bits in \mathbf{b} : b_k in Eq. (6), $b_{(k+1) \bmod L}$ in Eq. (7), $b_{(k+B) \bmod L}$ and $b_{((k+1)+B) \bmod L}$ in Eq. (9). When $k = 0$, although the number of possible bit patterns of \mathbf{b} is $2^L (= 2^8$ in Fig. 5(a)),

the number of patterns of symbols s_1 is $2^4(= 16)$. When $k = 1$, the bits in \mathbf{b} used to generate the symbols s_2 are shifted from the bits that were used to generate the symbols s_1 by one bit. Hence, in addition to the four bits used in the case that $k = 0$, $b_{(k+2) \bmod L}$ and $b_{((k+2)+B) \bmod L}$ are used to generate the symbol s_2 . Therefore, six bits are involved in total. In Fig. 5(b), it can be observed that the number of patterns of the symbols s_2 is $2^6(= 64)$; however, the maximum number of bit patterns of \mathbf{b} is $2^8(= 256)$.

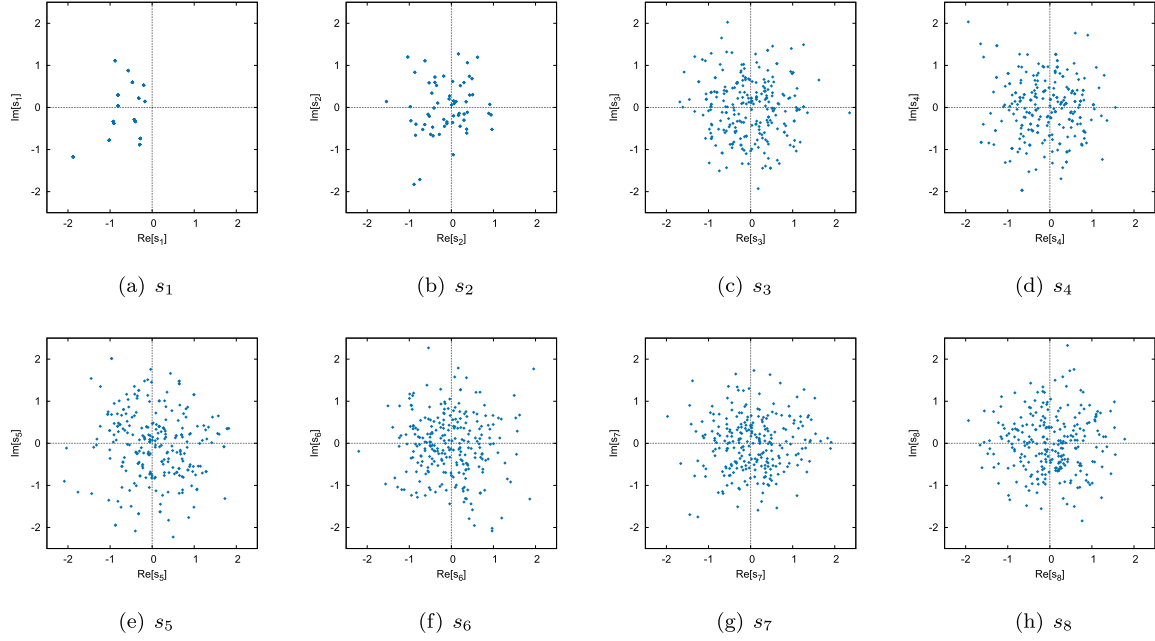


Fig. 5. Examples of constellation diagrams for the k th transmitted symbol s_k obtained using the original C-MIMO system, when $B = 4$ and $n = 100$.

4. Proposed system

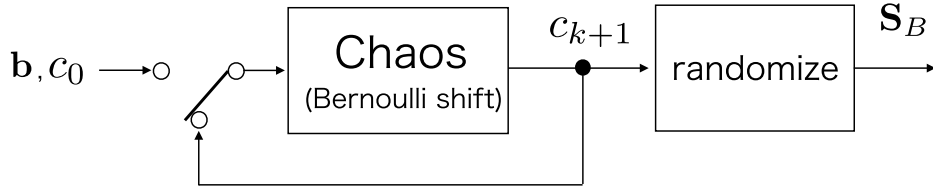


Fig. 6. Configuration of the modulation part in the proposed C-MIMO system.

We propose an improved C-MIMO system that addresses the above-mentioned issues I and II discussed in Sec. 3. Figure 6 shows a configuration of the proposed chaos-based modulation for the C-MIMO transmission system.

Let us first address issue I where the third equations in Eqs. (6) and (7) have no effect on the value of c_k as shown in Sec. 3. On the basis of this fact, we simply use the following Eqs. (15) and (16) instead of Eqs. (6) and (7).

$$x_0 = \text{Re}[c_k], \quad (15)$$

$$y_0 = \text{Im}[c_k]. \quad (16)$$

Although issue I exists in Eqs. (6) and (7) in the original C-MIMO system from a theoretical perspective, this issue seems to be resolved indirectly by replacing the modulo operator $(\bmod 1)$ in Eq. (8) with $\bmod(1 - \delta)$ in the original C-MIMO system, where $\delta (< 1)$ is a small real number and $\delta = 10^{-11}$ in [10]. This replacement of the modulo operator is introduced to prevent the values of

x_n and y_n from converging to zero in numerical simulations, but not to address issue I. Moreover, this value of δ is empirically determined; further, no reports systematically clarify how the value of δ affects the dynamics of the Bernoulli shift map and the randomness of generated transmitted symbols. In this sense, the proposed method (Eqs. (15) and (16)) can remove the uncertainty caused by the replacement of the modulo operator and simplify the original C-MIMO system.

By this change, the initial values, x_0 and y_0 , are generated only from c_k , and the initial values do not depend on \mathbf{b} anymore. Instead, \mathbf{b} is used in another part, which resolves issue II.

Second, let us address issue II. We changed Eq. (9) to

$$c_{k+1} = x_{n+i_x} + jy_{n+i_y}, \quad (17)$$

where $n = 100$ and we adaptively change the number of iterations i_x and i_y , using the transmit block \mathbf{b} . When the $(k+1)$ th element symbol c_{k+1} is generated, the values of i_x and i_y are determined by

$$i_x = \begin{cases} 0 & (k \geq 1), \\ \sum_{m=0}^{B-1} 2^m b_m & (k = 0), \end{cases} \quad (18)$$

and

$$i_y = \begin{cases} 0 & (k \geq 1), \\ \sum_{m=B}^{2B-1} 2^{m-B} b_m & (k = 0). \end{cases} \quad (19)$$

Equations (18) and (19) indicate that i_x and i_y take values calculated by the second equations in Eqs. (18) and (19) only when the first transmitted symbol s_1 is generated ($k = 0$). We emphasize that all bits in \mathbf{b} are used in the proposed C-MIMO system. In addition, because the other transmitted symbols $s_k (k \geq 1)$ are generated from the first element symbol c_1 , it is not necessary to calculate i_x and i_y for these transmitted symbols. Based on this idea, i_x and i_y are set to 0 when $k \geq 1$. These changes increase the number of patterns of symbols to $2^{N_t B}$, which is the maximum number of bit patterns of \mathbf{b} .

5. Numerical experiments

To confirm the performance of the proposed C-MIMO system, we conducted three numerical experiments. In the first experiment, we checked the number of transmitted symbols generated from all bit patterns of \mathbf{b} . Figure 7 shows the constellation diagrams obtained by the proposed C-MIMO system when the MIMO transmit block length $B = 4$.

In Fig. 7, we conducted the same experiments corresponding to Fig. 5, but used the proposed C-MIMO system instead of the original one. From Fig. 7(a) and 7(b), we can observe that the number of symbols increases substantially in comparison with Fig. 5(a) and 5(b).

Figure 8 shows distributions of inter-symbol distances obtained from the original and proposed C-MIMO systems. In the case of the original C-MIMO system, because the number of generated patterns of the symbol s_1 is small, the inter-symbol distances are localized around zero (Fig. 8(a)). Conversely, in the case of the proposed C-MIMO system, because the number of generated patterns of the symbol s_1 is 2^8 , the inter-symbol distances are widely distributed (Fig. 8(a)). From Fig. 8(b), we can observe that the tendency is almost the same. Specifically, the variation of the inter-symbol distance becomes small in the case of the original C-MIMO system. Therefore, the proposed C-MIMO system works well in increasing the number of patterns of the transmitted symbols.

In the second experiment, to show quantitatively that the proposed C-MIMO system can generate a wide variety of patterns of the transmitted symbols, we conducted the following experiment: we first calculated the transmitted symbol $s_i (i = 1, \dots, 8)$ from \mathbf{b} using the proposed C-MIMO system. The transmitted symbols were calculated for all possible bit sequences of \mathbf{b} , where the total number of patterns of \mathbf{b} is $2^{N_t B} (= 2^8)$. Consequently, the number of patterns of each transmit symbol (s_i) is at most 256 patterns. Next, 256 symbols were generated randomly, where the real and imaginary parts

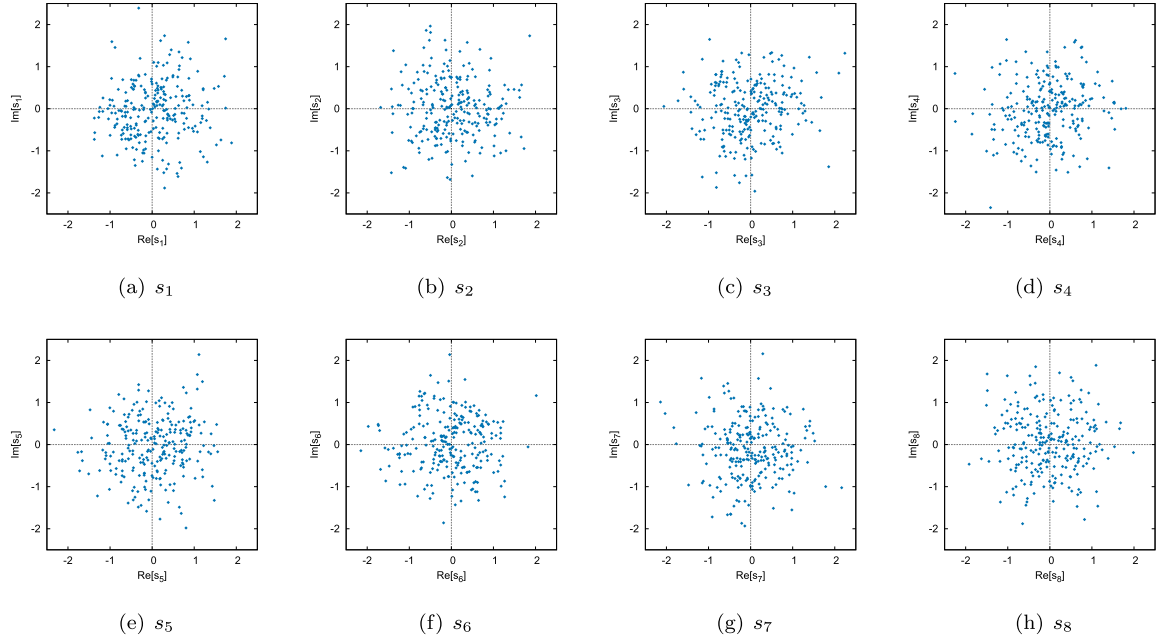


Fig. 7. Examples of constellation diagrams of the k th transmitted symbol s_k obtained using the proposed C-MIMO system when $B = 4$ and $n = 100$.

of each symbol obey a Gaussian distribution having the same average and variance as the symbols generated by the proposed C-MIMO system. We then calculated the distance between the distribution of inter-symbol distances obtained by the proposed C-MIMO system and that of random symbols as follows:

$$A_i = \int_0^{\infty} |P_i(z) - Q(z)| dz, \quad (20)$$

where $P_i(z)$ is a cumulative distribution of the distances between symbols corresponding to s_i generated by the proposed C-MIMO system and $Q(z)$ is a cumulative distribution of distances between random symbols. Because the probability that the randomly generated symbols are identical is extremely low, the 256 randomly generated symbols are predominantly different from each other. In this sense, the randomly generated symbols can be ideally placed transmitted symbols and thus the distance A_i can be a good indicator of the variation in the patterns of symbols obtained by the proposed C-MIMO system. In this experiment, by calculating Eq. (20) for the original chaos-based modulation method described in Sec. 2, we compared the proposed C-MIMO system with the original one. Figure 9 shows the averaged values of A_i calculated for 100 values of the common key c_0 . From Fig. 9, it can be observed that the proposed C-MIMO system can generate symbols that are more similar to the random symbols than the symbols generated by the original C-MIMO system.

In the third experiment, we evaluated the BER and BLER of the original and proposed C-MIMO systems. In this experiment, we also investigated the BER and BLER of two MIMO systems: the first system is a MIMO system with binary phase-shift keying (BPSK) modulation and the second system is that with quadrature phase-shift keying (QPSK) modulation. We compared the proposed C-MIMO system with these two modulation schemes. In this paper, we call the MIMO system with BPSK modulation *BPSK-MIMO* and that with QPSK modulation *QPSK-MIMO*. In the BPSK- and QPSK-MIMO systems, MIMO receive vectors were decoded by the maximum likelihood detection (MLD). Table I lists the simulation conditions. We assumed that the channel information follows the antenna and symbol i.i.d. one-path Rayleigh fading. In addition, the channel information is perfectly estimated in the receiver side.

Figures 10–12 show the BER and BLER against E_b/N_0 per receive antenna, where E_b/N_0 is the signal to noise ratio (SNR) per bit. In this experiment, we randomly generated 10^6 transmit blocks \mathbf{b} and calculated their decoded blocks $\hat{\mathbf{b}}$. The BER is the ratio of the number of incorrect bits to the total number of received bits. Conversely, the BLER is the rate at which $\hat{\mathbf{b}}$ is not decoded to \mathbf{b} .

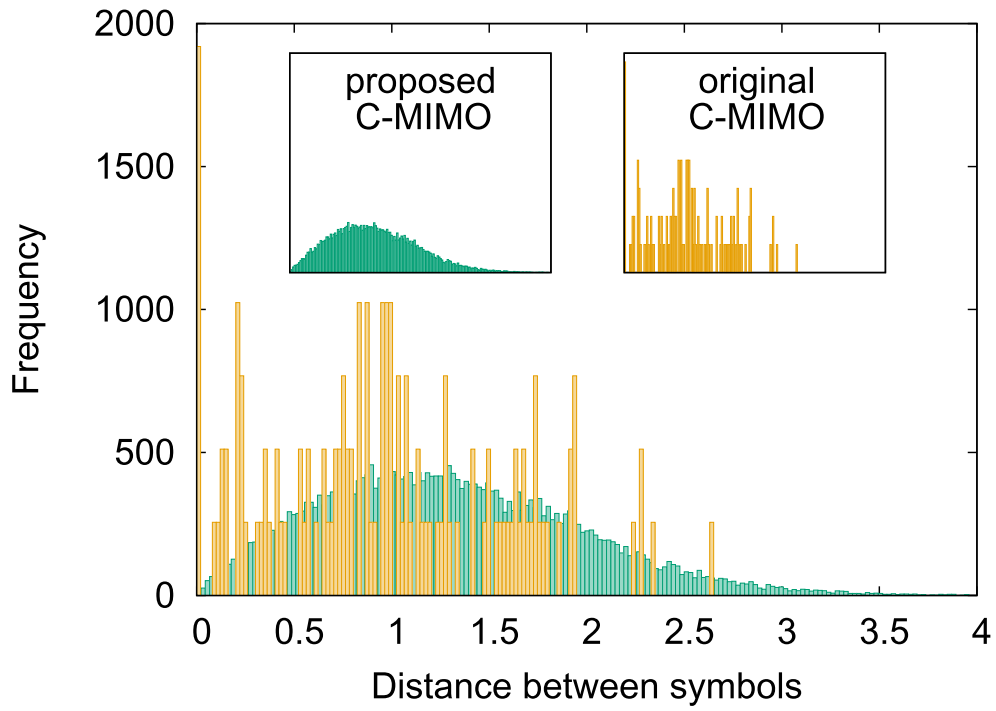
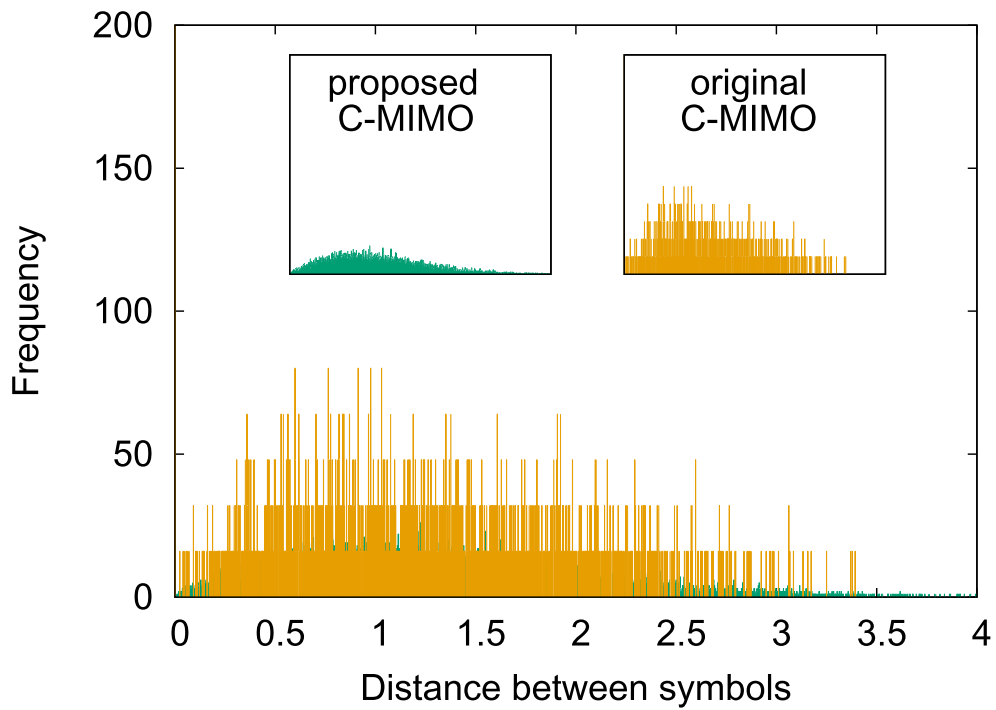
(a) s_1 (b) s_2

Fig. 8. Examples of distributions of inter-symbol distances when $B = 4$ and $n = 100$. The distributions illustrated in the insets in each figure are separately plotted distributions for the proposed and original C-MIMO systems.

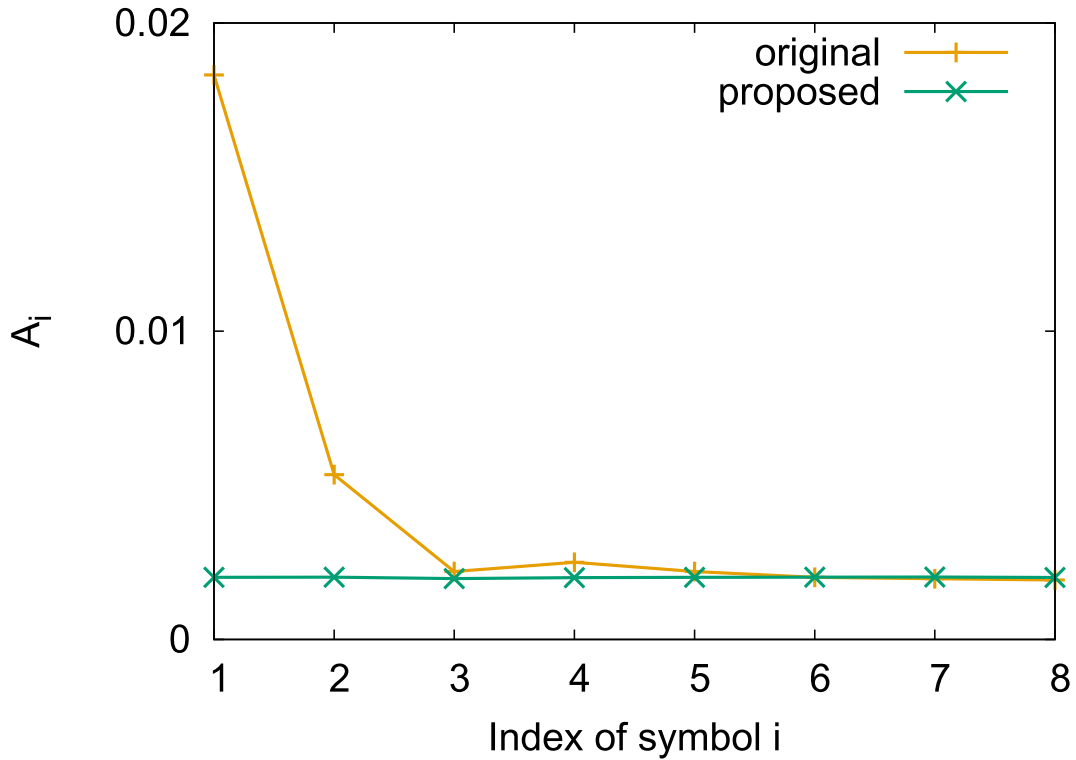


Fig. 9. Results for A_i (Eq. (20)) obtained using the original and proposed C-MIMO systems.

Table I. Conditions of the numerical experiments.

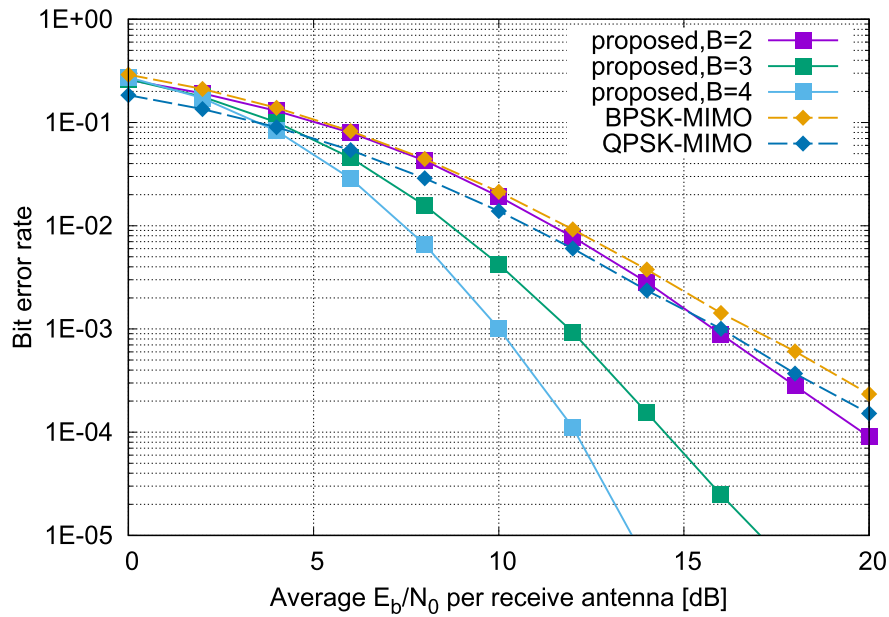
Modulation	BPSK, QPSK	Chaos (original and proposed)
Number of transmit and receive antennas	$N_t = N_r = 2$ (Figs. 10, 11(a), and 12(a)) $N_t = N_r = 2, 3,$ and 4 (Figs. 11(b) and 12(b))	
MIMO Block length	N/A	$B = 2, 3$ and 4
Chaos map	N/A	Bernoulli shift map
Number of iterations for the Bernoulli shift map	N/A	$n = 100$
Synchronization of common keys	Perfect	
Channel	Antenna and symbol i.i.d. one-path Rayleigh fading	
Receive channel state information	Perfect	
Algorithm for decoding	MLD	JMLSE

We changed the common key c_0 every 10^4 blocks (the total number of common keys c_0 is 100). The real and imaginary parts of the common key c_0 obey a uniform distribution in the range $(0, 1)$. The elements of $\mathbf{n}(t)$ are additive white Gaussian noises. Then, SNR is calculated as follows:

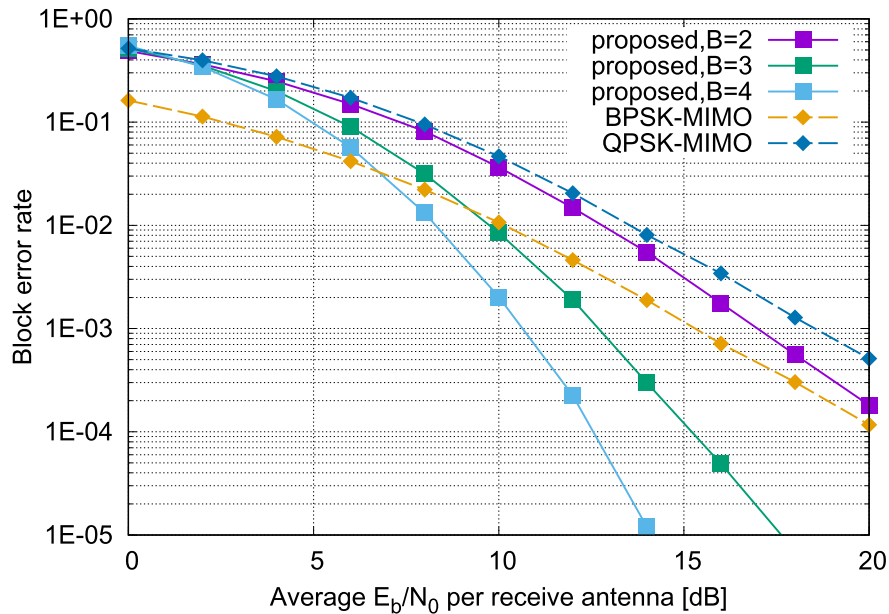
$$\text{SNR} = 10 \log_{10} \frac{\sigma_r^2}{\sigma_n^2}, \quad (21)$$

where σ_r^2 is the average power per receive antenna and σ_n^2 is the average power of noises. In the BPSK-MIMO, the original C-MIMO, and the proposed C-MIMO systems, E_b/N_0 is identical to SNR. In the case of the QPSK-MIMO system, $E_b/N_0 = \text{SNR}/2$.

We first compared the proposed C-MIMO system with the BPSK- and QPSK-MIMO systems, as illustrated in Fig. 10. From Fig. 10, it can be observed that the BPSK- and QPSK-MIMO systems exhibit better performance than the proposed C-MIMO system only when the value of SNR (E_b/N_0) is low ($E_b/N_0 \ll 10$). In contrast, especially in the region where the values of E_b/N_0 are high, the proposed C-MIMO system ($B > 2$) substantially decreases BER/BLER and shows better BER/BLER



(a) BER

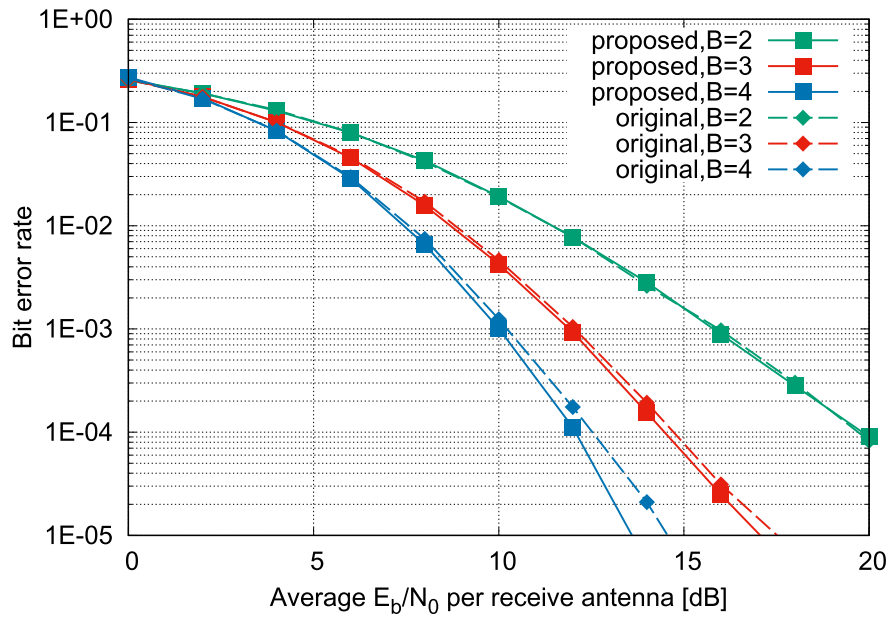
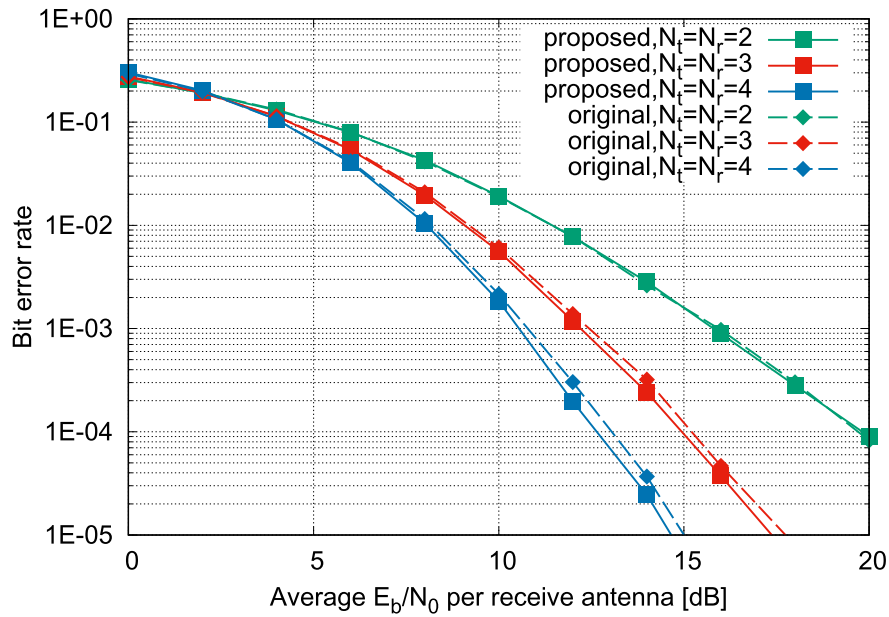


(b) BLER

Fig. 10. Bit error rate (BER) and block error rate (BLER) of the BPSK-, the QPSK-, and the proposed C-MIMO systems, where $N_t = N_r = 2$.

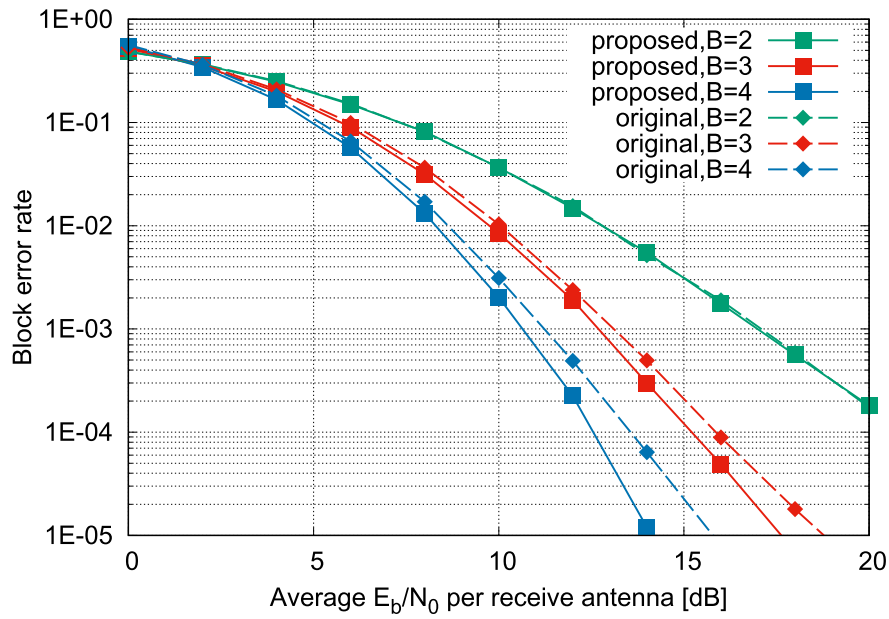
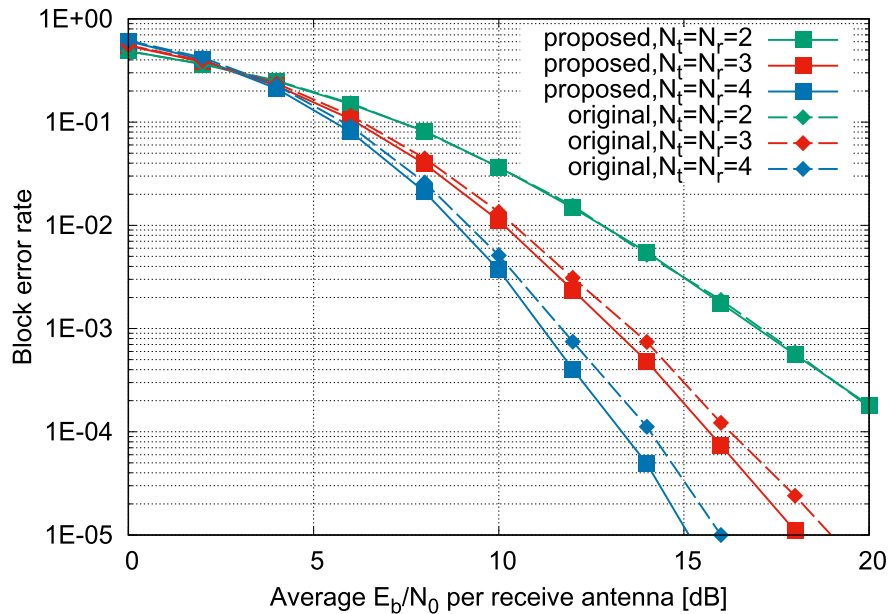
performance than the BPSK- and QPSK-MIMO systems.

Next, we compared the proposed and the original C-MIMO systems, as illustrated in Figs. 11(a) and 12(a). From Figs. 11(a) and 12(a), it can be observed that when $B = 3$ and 4, the BER and BLER performances of the proposed system are improved when compared to the original C-MIMO system. From Figs. 11(b) and 12(b), when $N_t = N_r = 3$ and 4, the BER and BLER performances of the proposed C-MIMO system are also improved when compared to the original C-MIMO system. The reason is that it is easy to distinguish transmitted symbols from each other because the number of patterns of the transmitted symbols has increased. When $B = 2$ and $N_t = 2$, the BER and BLER performances of the proposed C-MIMO system are equivalent to those of the original C-MIMO

(a) $N_t = N_r = 2$ (b) $B = 2$ **Fig. 11.** BER of the original and proposed C-MIMO systems.

system. This is because both the proposed C-MIMO system and the original one generate the same number of block patterns, i.e., 2^4 block patterns, in the case where $B = 2$.

Although these results of numerical simulations reveal that the proposed C-MIMO system exhibits higher performance than the BPSK-, the QPSK-, and the original C-MIMO systems, we should specify that the computational cost of the proposed C-MIMO system is higher than that of the BPSK- and the QPSK-MIMO systems because of the high computational cost of the chaos modulation and the decoding of the modulated symbols. In addition, the computational cost of the proposed C-MIMO system is also higher than the original C-MIMO system, because the number of iterations (i_x and i_y) is large. Consequently, reducing the computational cost of the proposed C-MIMO system is one of

(a) $N_t = N_r = 2$ (b) $B = 2$ **Fig. 12.** BLER of the original and proposed C-MIMO systems.

our important future works.

In this paper, we primarily considered the single-user MIMO; however, the multi-user MIMO (MU-MIMO) [11] and massive MIMO [12–14] are currently receiving considerable attention. Besides, several extended C-MIMO systems have been proposed, where the original C-MIMO system is extended to the MU-MIMO [15] and massive MIMO systems [16]. These extended C-MIMO systems also incorporate the chaos-based modulation which is almost identical to the chaos-based modulation used in the original C-MIMO system for a single user described in Sec. 2. This fact implies that our approach can be applied to these extended C-MIMO systems. Thus, the application of our method to these extended C-MIMO systems and the evaluation of their performance are important future works.

6. Summary

In this paper, we investigated the modulation part of the C-MIMO system [4–6] from the perspective of nonlinear dynamical systems. Then, we described the issues pertaining to the initial values of the chaotic dynamical systems used in the original C-MIMO system, i.e., the constellation diagrams of the original C-MIMO system show that the same transmitted symbols are generated from different symbols because only four bits of the total L bits in the transmit block were used to generate the transmitted symbols in the original C-MIMO system.

Based on these results, we proposed a new modulation method. In the proposed C-MIMO system, all bits in the transmit block are used to generate the transmitted symbols and the number of iterations of chaotic maps was adaptively controlled by the information of the transmit block. We showed that the BER performance of the proposed C-MIMO system was superior to that of the original C-MIMO system.

Although we only evaluated the BER (BLER) performance of the proposed C-MIMO system in this paper, throughput is also an important performance measure for the proposed C-MIMO system, which should be investigated in future work.

Acknowledgments

The authors would like to thank Prof. M. Hasegawa (Tokyo University of Science) for his valuable comments and suggestions. This research is partially supported by the JSPS Grant-in-Aid for Scientific Research (Nos. 15KT0112, 17K00348 and 18K18125).

References

- [1] G.J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [2] T.L. Carroll and L.M. Pecora, “Synchronizing chaotic circuits,” *IEEE Transactions on Circuits and Systems*, vol. 38, no. 4, pp. 453–456, 1991.
- [3] G. Mazzini, G. Setti, and R. Rovatti, “Chaotic complex spreading sequences for asynchronous DS-CDMA. I. System modeling and results,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 44, no. 10, pp. 937–947, 1997.
- [4] E. Okamoto, “A chaos MIMO transmission scheme for channel coding and physical-layer security,” *IEICE Transactions on Communications*, vol. E95-B, no. 4, pp. 1384–1392, 2012.
- [5] E. Okamoto and Y. Inaba, “Multilevel modulated chaos MIMO transmission scheme with physical layer security,” *NOLTA*, vol. 5, no. 2, pp. 140–156, 2014.
- [6] E. Okamoto and N. Horiike, “Performance improvement of chaos MIMO scheme using advanced stochastic characteristics,” *IEICE Communications Express*, vol. 5, no. 10, pp. 371–377, 2016.
- [7] G.E.P. Box and M.E. Muller, “A note on the generation of random normal deviates,” *The Annals of Mathematical Statistics*, vol. 29, pp. 610–611, 1958.
- [8] E. Okamoto, “Performance improvement of wireless communications system using nonlinear signal processing,” *IEICE Fundamental Review*, vol. 9, no. 4, pp. 350–363, 2016, In Japanese.
- [9] G. Alvarez and S. Li, “Breaking an encryption scheme based on chaotic baker map,” *Physics Letters A*, vol. 352, nos. 1–2, pp. 78–82, 2006.
- [10] Y. Masuda, E. Okamoto, K. Ito, and T. Yamamoto, “An uplink non-orthogonal multiple access scheme having physical layer security based on chaos modulation,” *Proceedings of 2019 International Conference on Information Networking (ICOIN)*, pp. 136–140, 2019.
- [11] Q.H. Spencer, C.B. Peel, A.L. Swindlehurst, and M. Haardt, “An introduction to the multi-user MIMO downlink,” *IEEE Communications Magazine*, vol. 42, no. 10, pp. 60–67, 2004.
- [12] F. Rusek et al., “Scaling up MIMO: opportunities and challenges with very large arrays,” *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.
- [13] A.L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, “Millimeter-wave massive MIMO: the next wireless revolution?,” *IEEE Communications Magazine*, vol. 52, no. 9, pp. 56–62, 2014.

- [14] E.G. Larsson, O. Edfors, F. Tufvesson, and T.L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.
- [15] Y. Inaba and E. Okamoto, “Multi-user chaos MIMO-OFDM scheme for physical layer multi-access security,” *NOLTA*, vol. 5, no. 2, pp. 172–183, 2014.
- [16] N. Horiike, E. Okamoto, and T. Yamamoto, “Uplink large-scale chaos MIMO transmission scheme using gaussian belief propagation,” *Journal of Signal Processing*, vol. 22, no. 4, pp. 169–173, 2018.

Invited Paper

Network structure detection using convergent cross mapping on multivariate time series

Kazuya Sawada^{1a)}, Yutaka Shimada², and Tohru Ikeguchi^{1,3}

¹ *Department of Management Science, Graduate School of Engineering, Tokyo University of Science, 6-3-1 Niijuku, Katsushika, Tokyo 125-8585, Japan*

² *Department of Information and Computer Sciences, Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakuraku, Saitama-shi, Saitama 338-8570, Japan*

³ *Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, 6-3-1 Niijuku, Katsushika, Tokyo 125-8585, Japan*

^{a)} *sawada@hisenkei.net*

Received January 31, 2020; Revised May 26, 2020; Published October 1, 2020

Abstract: The recent developments in measurement techniques have allowed us to observe multidimensional time series data in various fields. Thus, detecting causal relations between elements from multidimensional time series data is useful for prediction, model generation, and system control. In addition, causal relations between elements can be detected to estimate network connections. In other words, the network structure can be estimated from multidimensional time series data by using causality detection methods. Among the several methods for detecting causality, Granger causality is a well-known method that is widely used for causal estimation between time series. On the other hand, a method called convergent cross mapping (CCM) has been proposed, which can distinguish causality from pseudo-correlation. It is important to investigate whether CCM will be effective with increased number of elements, even though the evaluation of its performance with a few elements has been reported in the literature. Moreover, it is important to evaluate the performance with changing dynamics of the elements. In this study, to estimate the connectivity between elements in complex networks, we apply CCM to mathematical models of complex networks, or the Watts-Strogatz model. In particular, we investigate how complex network structures affect causal estimation, by applying CCM to multidimensional time series data produced from complex networks. According to the results, we find the connectivity estimation accuracies in the regular ring-lattice network to be slightly higher than those in random networks. Furthermore, we reveal that it is easier to perform connectivity estimation for a network with a community structure than a random structure.

Key Words: nonlinear time series analysis, complex network, causality, connectivity

1. Introduction

Causal relations exist between various events, and thus, it is important to identify them. In recent years, with the advances in observation technology and information science, considerable time series data can be measured easily. Thus, various methods have been proposed for estimating and detecting causality for time series data. For example, Granger causality is a commonly used method in causal estimation between time series signals [1]. However, several problems remain. Although Granger causality works well for a certain type of time series, it might not be effective for cases in which time series signals are generated from nonlinear deterministic dynamical systems.

On the other hand, a method called convergent cross mapping (CCM) [2] has been proposed as an effective method for estimating the causal relation between nonlinear deterministic time series signals. In [2], it was shown that CCM is effective for detecting the causality for a simple network with relatively few vertices. However, in the real world, the causality between events might be produced from complex networks and real-world networks are known to have a complex network structure [3, 4], for example, neural networks, power grids, the Internet, friendships, the spread of disease, and the spread of information. Estimating these complex connections is generally a difficult problem, but an important topic. That is, we need to evaluate the CCM performance for complex networks with relatively more vertices. Then, our purpose is two-fold. The first one is to investigate the ability of CCM to perform causality detection and coupling estimation in complex networks with relatively more vertices. The second one is to investigate whether the performance of coupling estimation by CCM depends on the network structure. Therefore, this study investigates how the causality detection accuracy resulting from the difference in the network structure appears when component elements in a network are connected by a complex network model.

In addition, coupling estimation using CCM has some problems. In CCM, after we mutually predict time series in-sample through cross mapping, we need to manually decide the threshold that determines whether a coupling exists. Therefore, we propose the use of CCM in combination with the Otsu method [5]. The proposed method can automatically decide the threshold for coupling estimation.

2. Method

Let us refer to a time series that influences as “the time series of cause,” and one that is influenced as “the time series of result.” The information in the time series of cause can be accumulated in the time series of result, and as the length of the time series increases, the amount of information of the time series of cause accumulated in the time series of result increases. In fact, such a concept is used in this paper as a common basis for causality estimation methods of CCM [2]. In this section, we review CCM, or a causality estimation method, based on the nonlinear dynamical systems theory.

The basic idea of CCM is that the higher the mutual prediction accuracy of a time series, the higher the possibility that causality exists. First, we consider a dynamical system comprising two variables, x_t and y_t , where t is time. Let x_t be a cause of y_t . Originally, CCM estimates the causality between multidimensional time series data, but for simplicity, we use two variables here.

When applying CCM to two time series data x_t and y_t , we first reconstruct the attractors using a time-delay coordinate system [6] from the original time series x_t and y_t . Embedding can be achieved by transforming an observed time series into a time-delay coordinate with a sufficiently large dimension [7]. Even if we cannot observe all variables, we can have the information on the original dynamical system only from an observed time series by reconstructing the attractors.

Let us define a reconstructed attractor from the time series of x_t by a time-delay vector $\mathbf{x}(t)$, as shown in Eq. (1).

$$\mathbf{x}(t) = (x_t, x_{t+\tau}, x_{t+2\tau}, \dots, x_{t+(m-1)\tau}), \quad (1)$$

where τ is the time delay and m is the dimension of the reconstructed state space. Similarly, $\mathbf{y}(t)$ is defined by

$$\mathbf{y}(t) = (y_t, y_{t+\tau}, y_{t+2\tau}, \dots, y_{t+(m-1)\tau}). \quad (2)$$

Next, to predict the original time series x_t and y_t , we use the above-mentioned time-delay vectors $\mathbf{x}(t)$ and $\mathbf{y}(t)$. This method is called cross mapping because y_t is predicted using $\mathbf{x}(t)$ and x_t is predicted using $\mathbf{y}(t)$. Now, we define \hat{x}_t and \hat{y}_t as the predicted time series. In the cross mapping, \hat{y}_t is calculated by

$$\hat{y}_t = \sum_{i=1}^K w_{t^i} y_{t^i}. \quad (3)$$

Here, K is the number of neighbors for predicting the time series. We take the K neighbors of $\mathbf{x}(t)$, which are $\mathbf{x}(t^1)$, $\mathbf{x}(t^2)$, \dots , $\mathbf{x}(t^K)$, where $\mathbf{x}(t^i) = (x_{t^i}, x_{t^i+\tau}, x_{t^i+2\tau}, \dots, x_{t^i+(m-1)\tau})$. Next, we take the corresponding $\mathbf{y}(t^1)$, $\mathbf{y}(t^2)$, \dots , $\mathbf{y}(t^K)$, where $\mathbf{y}(t^i) = (y_{t^i}, y_{t^i+\tau}, y_{t^i+2\tau}, \dots, y_{t^i+(m-1)\tau})$ and predict \hat{y}_t by Eq. (3). In what follows, we basically use $K = m + 1$ in numerical experiments. However, when investigating the effects of the number of neighbors K on predictability in numerical experiments, we use $K = m + 1$ or $2m$. In Eq. (4), w_{t^i} is defined by

$$w_{t^i} = \frac{u_{t^i}}{\sum_{j=1}^K u_{t^j}}, \quad (4)$$

where w_{t^i} is the weight calculated from the time-delay vector of the i th nearest neighbor $\mathbf{x}(t^i)$ in the reconstructed state space $\mathbf{x}(t)$. In Eq. (4), u_{t^i} is defined by

$$u_{t^i} = \exp \left\{ \frac{-d[\mathbf{x}(t), \mathbf{x}(t^i)]}{d[\mathbf{x}(t), \mathbf{x}(t^1)]} \right\}, \quad (5)$$

where $d[\mathbf{x}(t), \mathbf{x}(s)]$ is the Euclidean distance between $\mathbf{x}(t)$ and $\mathbf{x}(s)$ and $\mathbf{x}(t^i)$ is the i th nearest neighbor of the time-delay vector $\mathbf{x}(t)$ in the reconstructed state space. When $d[\mathbf{x}(t), \mathbf{x}(t^i)] = 0$, u_{t^i} is defined by

$$u_{t^i} = \begin{cases} 1 & (i = 1), \\ 0 & (i \neq 1). \end{cases} \quad (6)$$

Then, using Eqs. (3)–(6), the predicted state \hat{y}_t is obtained:

$$\hat{y}_t = \sum_{i=1}^K \frac{\exp\{-d[\mathbf{x}(t), \mathbf{x}(t^i)]/d[\mathbf{x}(t), \mathbf{x}(t^1)]\}}{\sum_{j=1}^K \exp\{-d[\mathbf{x}(t), \mathbf{x}(t^j)]/d[\mathbf{x}(t), \mathbf{x}(t^1)]\}} y_{t^i}. \quad (7)$$

The accuracy of cross mapping is evaluated by the correlation coefficient ρ_{xy} between the actual time series y_t and \hat{y}_t obtained by Eq. (7). Namely, ρ_{xy} is defined by

$$\rho_{xy} = \frac{\sum_{t=1}^T (\hat{y}_t - \bar{\hat{y}}_t)(y_t - \bar{y}_t)}{\sqrt{\sum_{t=1}^T (\hat{y}_t - \bar{\hat{y}}_t)^2} \sqrt{\sum_{t=1}^T (y_t - \bar{y}_t)^2}}, \quad (8)$$

where T is the length of time series and \bar{y}_t is the average of y_t . Here, let us assume that y_t is affected by x_t and x_t is not affected by y_t (x_t is cause and y_t is result). In this case, the correlation coefficient ρ_{xy} between y_t and \hat{y}_t predicted from $\mathbf{x}(t)$ is expected to take a low value. On the other hand, the correlation coefficient ρ_{yx} between x_t and \hat{x}_t predicted from $\mathbf{y}(t)$ is expected to be a high value. This is because x_t does not contain the information of y_t but y_t contains the information of x_t . From this fact, we estimate the existence and direction of causality from y_t to x_t . The prediction accuracy of the time series increases and converges when the time series length is increased while performing the cross mapping. Therefore, it is called convergent cross mapping [2].

3. Network models and time series models

3.1 Network models

In this study, we investigate how complex network structures affect the accuracy of connectivity estimation by CCM. Then, we use two complex network structures.

The first one is generated from the Watts–Strogatz model [8], which is a network model that can easily reproduce the small-world structure [8]. We use a network with 20 and 100 vertices generated from the WS model in numerical experiments. The reason why we use the WS model is that it is one of the most famous network models that can change the network randomness by changing the rewiring probability. We first create a regular ring-lattice network of arbitrary average degree when creating a network using the WS model. Here, we set the average degree to four in case of 20 vertices, and six in case of 100 vertices. Next, we rewire the edges of a regular ring-lattice network by changing the rewiring probability p . When $p = 1$, it is a random network.

The second one is a network with community structure. As shown in Fig. 1, we create two networks: a network with a community structure (Fig. 1(a)) and that with a random structure (Fig. 1(b)). Here, we describe how to create a network with a community structure. When we create such a network, we first create two complete graphs of 11 vertices. Next, we select two edges from each complete graph and create a network with a community structure by rewiring these edges. In addition, we use the following procedure to create a fixed-degree random network. We first create a regular ring-lattice network with 22 vertices and an average degree of 10, using the WS model [8]. Next, we rewire the edges randomly so that the degree of each vertex of the created network does not change [9].

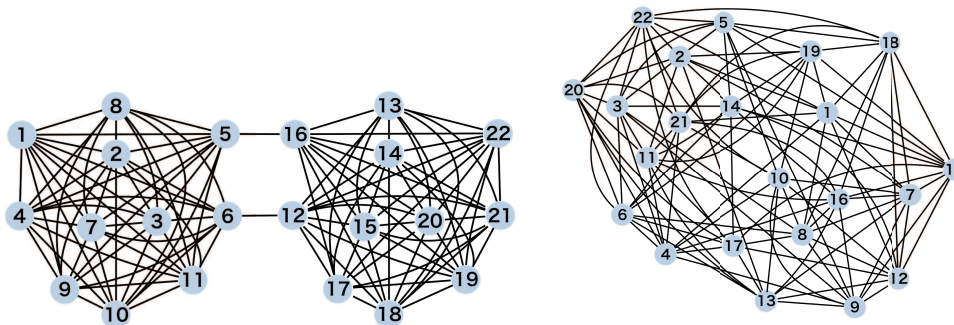


Fig. 1. Networks with (a) community structure and (b) random structure.

Creating these two complex network structures, we introduce two nonlinear dynamical systems for the vertex dynamics in the networks. The first nonlinear dynamical system is a coupled logistic map described in Sec 3.2 and the second one is a chaotic neural network described in Sec 3.3.

The reason why we use these dynamical systems is two-fold. The first reason is that [2] used a coupled logistic map consisting of two logistic maps to evaluate the CCM. That is, [2] did not try more than two logistic maps to evaluate the CCM. However, as it is well-known that the relation among more than three elements is complex, it is important to evaluate the CCM on more than three vertices with the same dynamics. The second reason is that functional shapes of elements used in these dynamical systems are different: the logistic map is a unimodal map, while a chaotic neuron is a bimodal map. Namely, it is important to evaluate the performance of CCM when the dynamics changes.

3.2 Coupled logistic map

The coupled logistic map [10, 11] is defined by Eq. (9), where $z_i(t)$ is the state value of the vertex i at time t , ε is the coupling strength between vertices, k_i is the degree of the vertex i , A_{ij} is the (i, j) th component of the adjacency matrix obtained from the WS model, and N is the number of vertices in the coupled logistic map. The function f describes a logistic map: $f(\varphi) = a\varphi(1 - \varphi)$, where a is a parameter. In this paper, we use $a = 4.0$ and $\varepsilon = 0.1$.

$$z_i(t+1) = (1-\varepsilon)f(z_i(t)) + \frac{\varepsilon}{k_i} \sum_{j=1}^N A_{ij}f(z_j(t)). \quad (9)$$

The reason why we set $\varepsilon = 0.1$ is that the state values could be a periodic solution depending on the value of ε . If the state value is a periodic solution, the prediction obtained with the cross mapping is almost unity, even if there is no causality between the vertices. Here, a bifurcation diagram visualizing the relation between ε and the state values when we use the WS model on $p = 0$ is shown in Fig. 2.

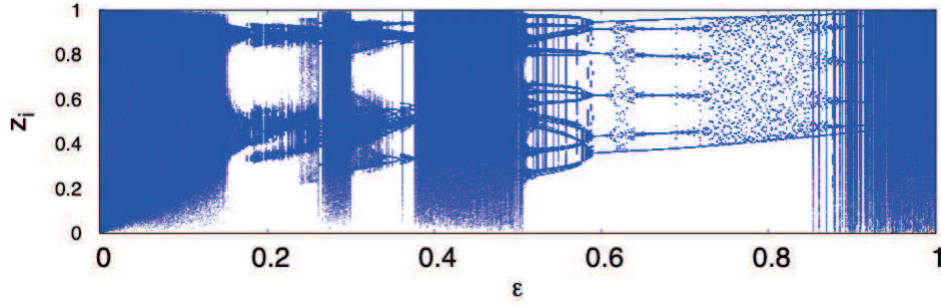


Fig. 2. Bifurcation diagram of the coupled logistic map ($a = 4.0$) on the WS model with the rewiring probability $p = 0$.

Figure 2 shows the values of $z_i (1 \leq i \leq 20)$ as the y -axis; we exclude 100,000 transient states and use 100 as the time series length. Figure 2 shows that the state value partially becomes a periodic solution when ε is more than 0.5. Therefore, we use the parameter $\varepsilon = 0.1$ with which the state value does not become a periodic solution.

3.3 Chaotic neural network

The chaotic neural network [12] is defined by Eq. (10), where $v_i(t)$ is a state value of the chaotic neuron i at time t , κ_i is the refractory time decay constant, α_i is a coefficient for refractory terms, a_i is a bias, N is the number of neurons in the network, and w_{ij} is the coupling strength between the neurons i and j ($i, j = 1, 2, \dots, N$). The function g is a sigmoid function $g(\psi) = \frac{1}{1 + \exp(-\psi/\varepsilon')}$, where ε' is a parameter. In this paper, we use $\kappa_i = 0.8$, $\alpha_i = 1.0$, $a_i = 0.2$, $\varepsilon' = 0.04$, and $w_{ij} = 0.2$ ($i \neq j$). When $i = j$, $w_{ij} = 0$.

$$v_i(t+1) = \kappa_i v_i(t) - \alpha_i g(v_i(t)) + a_i + \sum_{j=1}^N w_{ij} g(v_j(t)) \quad (10)$$

The reason why we set $w_{ij} = 0.2$ is that the state value could be a periodic solution depending on the value of w_{ij} . Here, a bifurcation diagram visualizing the relation between w_{ij} and the state values when we use the WS model of $p = 0$ is shown in Fig. 3.

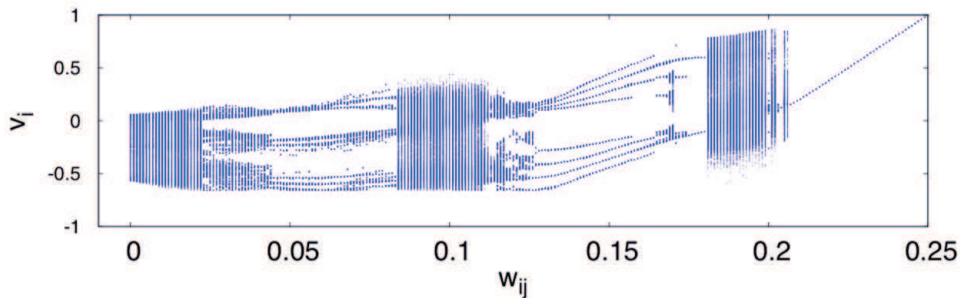


Fig. 3. Bifurcation diagram of the chaotic neural network.

In Fig. 3, we show the values of $v_i (1 \leq i \leq 20)$ as the y -axis, exclude 100,000 transient states and use 100 as the time series length. This figure indicates that the state value becomes a periodic solution when the w_{ij} is greater than 0.2. Therefore, we use a parameter whose state value does not become a periodic solution, $w_{ij} = 0.2$.

4. Numerical experiments

4.1 Estimation of connectivity for the coupled logistic map and the chaotic neural network

In this section, we explain the results of the numerical experiments. First, we use the coupled logistic map of Eq. (9) as the dynamics of each vertex of the WS model. Then, we change the edge rewiring probability p in the WS model and generate ten networks for each edge rewiring probability. We vary p from 0.0 to 1.0, in steps of 0.05. Next, we apply CCM to the time series generated from the network and estimate the connectivity between the vertices. To evaluate the accuracy of connectivity estimation, we calculate the average values of accuracies for 10 networks. Then, we can obtain 400 prediction accuracies for a single network because it has 20 vertices, and we predict the time series mutually in CCM. With sufficient data points in the time series, we obtain relatively high prediction accuracies in the case where vertices are connected and low accuracy in the case where the vertices are not connected. In other words, when a sufficient length of the time series is used, the distribution of the prediction accuracies is expected to be bimodal.

In connectivity estimation, the vertices are found to be connected if the prediction accuracy is greater than the threshold, and not connected if the prediction accuracy is less than the threshold. Then, we use the Otsu method for deciding the threshold of the prediction accuracies. The Otsu method is an effective method for deciding the threshold in cases where the distribution is bimodal. Normally, we can obtain an appropriate threshold of connectivity estimation using the Otsu method. In the Appendix, we describe the algorithm of the Otsu method.

In the following numerical experiments, the length of the time series used in the CCM is set to 4,000, the dimension m of the reconstructed state space is two, and τ is unity. Figure 4 shows the relation between the edge rewiring probability and the average of the connectivity estimation accuracy in a network generated from the WS model. In Fig. 4, the purple line with circles indicates the average values of the estimation accuracy, while the gray-shaded area indicates 95% confidence interval when we use the t -distribution.

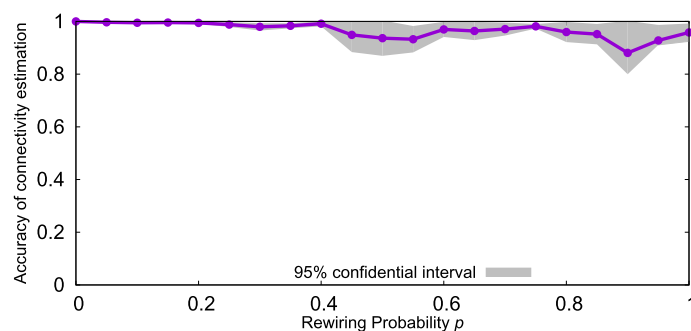


Fig. 4. Relation between the edge rewiring probability p and estimation accuracy of connectivity when dynamics of vertices are the coupled logistic map (Eq. (9)) in case of 20 vertices. The gray-shaded area indicates 95% confidence interval when we use the t -distribution.

Figure 4 shows that the estimation accuracies of connectivity with varying edge rewiring probability take higher values than 0.8. In addition, the connectivity estimation accuracies in a regular ring-lattice network are relatively higher than those in random networks. However, the estimation accuracies of connectivity take almost the same value even when the edge rewiring probability in the WS model is changed. The confidence intervals for the random network are wider than those for regular ring-lattice networks, and the variance in the accuracy of connectivity estimation is large.

Next, we perform a similar numerical experiment for the 100-vertex case. The average degree is set to six and the other parameters are the same as those in the case of 20 vertices. Figure 5 shows the relation between the edge rewiring probability and the average of the connectivity estimation accuracy. Figure 5 shows that the accuracy of connectivity estimation is better than about 0.8 for all edge rewiring probabilities. In addition, the overall connectivity estimation accuracy is slightly lower for random networks than for regular networks, and the confidence intervals are wider. This may be because as p increases, high-degree vertices emerged. Furthermore, the characteristics of the results are not much different from those of the case with 20 vertices.

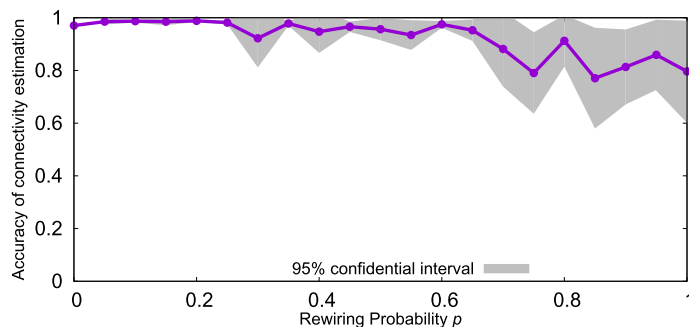


Fig. 5. Relation between the edge rewiring probability p and estimation accuracy of connectivity when the dynamics of the vertices are the coupled logistic map (Eq. (9)) in case of 100 vertices. The gray-shaded area indicates 95% confidence interval when we use the t -distribution.

Next, we apply the CCM to the chaotic neural network (Eq. (10)). Figure 6 shows the results of the numerical experiments conducted for the chaotic neural network. As shown in Fig. 6, the prediction accuracies take values between 0.6 to 0.9, which indicates that the accuracies are totally lower than the results of the logistic map. However, as shown in Fig. 4, the connectivity estimation accuracies in the regular ring-lattice network are relatively higher than those in the random networks and the confidence intervals for a random network are wider than for regular ring-lattice networks. These tendencies are almost the same as the case of the coupled logistic map shown in Fig. 4.

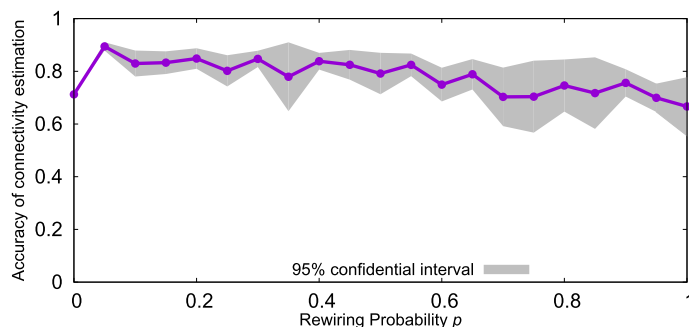


Fig. 6. Relation between the edge rewiring probability p and estimation accuracy of connectivity when the dynamics of vertices are the chaotic neural network (Eq. (10)). The gray-shaded area indicates 95% confidence interval when we use the t -distribution.

4.2 Effects of goodness of predictors and edge degrees

In causal estimation by CCM, we detect causality based on mutual prediction accuracy. In other words, the higher the prediction accuracy of the time series, the higher is the possibility that causality exists which means that it is important to create a good predictor. In particular, when we predict the time series using CCM in this paper, we calculate the weights using the distance information in the reconstructed state space (Eq. (4)). If we use this method for calculating the weights, the prediction point should be surrounded by neighboring points. Thus, if the prediction accuracy increases with

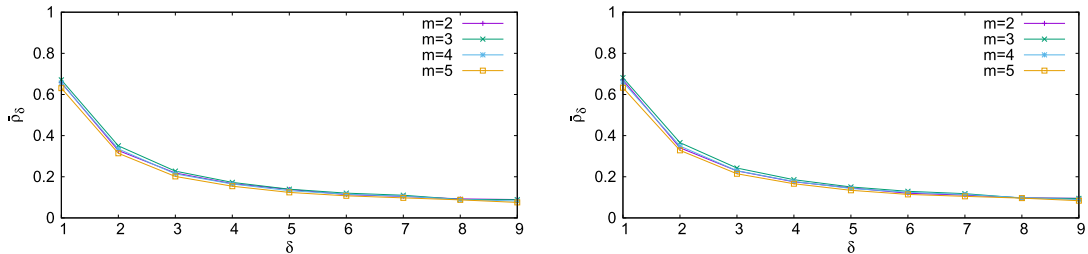
the number of neighbors used for the prediction, the prediction point might not be surrounded by neighboring points. On the other hand, the prediction point is surrounded by neighboring points with a high probability when the prediction accuracy does not increase even with an increase in the number of neighboring points used for prediction. Therefore, we investigate the mutual prediction accuracies by changing the number of neighboring points used for time series prediction, and compared their accuracies.

In addition, it is highly likely that the mutual prediction accuracy decreases in the case where the vertex has many edges, because vertices with a high degree are affected by many other vertices and the created predictor is likely to be bad. Therefore, we calculate the average of the time series prediction accuracy of the vertices of individual degrees when we change the number of neighboring points. We call it the average prediction accuracy, which is defined by Eq. (11).

$$\bar{\rho}_\delta = \frac{1}{N} \frac{\sum_{j=1}^N \sum_{i \in \Omega} A_{ij} \rho_{ij}}{\sum_{j=1}^N \sum_{i \in \Omega} A_{ij}}. \quad (11)$$

In Eq. (11), $\bar{\rho}_\delta$ is the average prediction accuracy of the degree δ , Ω is the set of indices of vertices defined by $\Omega \equiv \{i \mid k_i = \delta, i = 1, 2, \dots, N\}$, k_i is the degree of vertex i , A_{ij} is the (i, j) th element of the adjacency matrix, and ρ_{ij} is the prediction accuracy of vertex j predicted by the time series observed from vertex i .

To surround the predicted points in an m -dimensional state space, we need at least $m + 1$ points. Namely, we need at least $K = m + 1$ neighboring points when the dimension of the reconstructed state space is m [2]. Figure 7 shows the relation between $\bar{\rho}_\delta$ and degree δ when the number of neighboring points used for prediction K is $m + 1$ and $2m$. In Fig. 7, we use ten random networks ($p = 1$) with 20 vertices and an average degree of four. In addition, we change the dimension of the reconstructed state space for prediction in the range of $2 \leq m \leq 5$.



(a) The number of neighbors $K = m + 1$ ($2 \leq m \leq 5$).

(b) The number of neighbors $K = 2m$ ($2 \leq m \leq 5$).

Fig. 7. Relation between degree δ and average prediction accuracy $\bar{\rho}_\delta$ on ten random networks $p = 1$.

As shown in Fig. 7, the time series prediction accuracy decreases as the degree increases. In addition, from Figs. 7(a) and (b), there is no significant change in the prediction accuracy even when the number of neighboring points used for the prediction and the reconstruction dimension are changed. Therefore, it is sufficient that the number of neighboring points is $m + 1$ in the numerical experiments using the coupled logistic map. Then, we set the number of neighboring points K to $m + 1$ in the numerical experiments conducted for the coupled logistic map. However, in both Figs. 7(a) and (b), the prediction accuracy is slightly higher when the dimension of reconstruction is three. To make this result easier to understand, we replace the horizontal axis δ and legend m in Fig. 7. This is shown in Fig. 8. In other words, Fig. 8 shows the relation between the average prediction accuracy $\bar{\rho}_\delta$ and the reconstruction dimension m for each degree δ .

Figure 8 shows that although the average prediction accuracies $\bar{\rho}_\delta$ take almost the same value even when the dimension of the reconstructed state space is changed, the average prediction accuracies $\bar{\rho}_\delta$ are slightly improved in the case of $m = 3$. Although we investigate this reason in various ways, we cannot clarify the reason. Thus, we will investigate the reason for a future work. Figures 7 and 8 shows

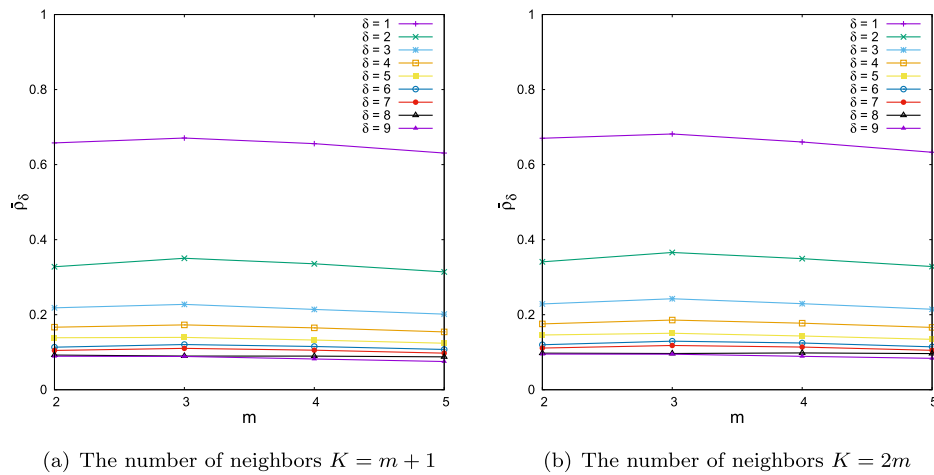


Fig. 8. Relation between dimension of reconstructed state space m and average prediction accuracy $\bar{\rho}_\delta$ for each degree δ .

that the mutual prediction accuracy for vertices with a higher degree is low. Vertices with higher degrees receive information from many other vertices, which makes it difficult to predict. In other words, the accuracy of causality detection using CCM depends on how the time series are affected.

4.3 Connectivity of community structure

In this section, using CCM, we compare the difference in the connectivity estimation accuracy between a network with a community structure and a random network with the same number of vertices and edges. We use the network shown in Fig. 1. Figure 9 shows the results of comparing the prediction accuracy. The connectivity estimation accuracy of a network with a community structure is higher than 0.8, whereas that of a random network is less than 0.6. Since the degrees of all vertices in the two networks are the same, the degree has no effect on the connectivity estimation accuracy. In other words, it is relatively easier to estimate a network having a community structure than a random network when we perform connectivity estimation by CCM.

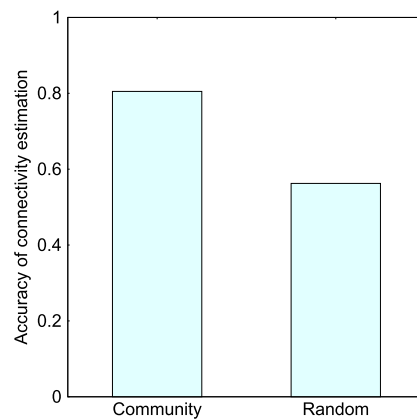


Fig. 9. Comparison of estimation accuracy of connectivity in a network with a community structure (Fig. 1(a)) and a random network structure (Fig. 1(b)).

5. Conclusion

We performed the connectivity estimation of networks generated from complex network models using CCM [2], which is a method for estimating causality. Here, causality means connectivity, and thus, estimating causality can be considered as estimating connectivity. In particular, we investigated whether there is a difference in the accuracy of connectivity estimation by CCM when the edge

rewiring probability of the WS model is changed. Thus, we used two nonlinear dynamical systems, a logistic map and a chaotic neuron model, as the network vertices.

As a result, we found that the connectivity estimation accuracies in the ring-lattice regular network were slightly higher than those in random networks. In addition, the variance in the accuracy of connectivity estimation in random networks was larger than that in regular ring-lattice networks. We also found that it is difficult to estimate the connectivity for high-degree vertices by CCM. This result supports the fact that the variance in the accuracy of connectivity estimation in a random network is large.

Furthermore, we compared the accuracies of connectivity estimation in a network with communities and its randomized networks with the same conditions (the numbers of vertices, edges, and degrees). As a result, we found that it is easier to perform connectivity estimation for a network with a community structure than that for random networks. In other words, these results imply that networks with a community structure have less disturbed information of time series than random networks.

In addition, to estimate the coupling using CCM, we need to manually decide the threshold that determines whether a coupling exists. However, the appropriate value of the threshold varies depending on the conditions. Therefore, in this paper, we proposed a method to estimate the network structure by automatically determining the threshold of the prediction accuracy and detecting the causality by combining the Otsu method and CCM. The results of the numerical experiments show that the coupling estimation accuracy obtained by the proposed method works well, which implies that it is a useful method.

In this paper, the WS model, which can easily change the network randomness, was used as a complex network model. However, it is also important to use other network models, such as the BA model [13]. Then, it is an important future work to investigate the performance of CCM using several network models including the BA model.

In addition, we would like to extend the method so that it is possible to estimate the network structure even if there are unobserved time series data, because it is not often the case that all the time series data generated from a network are observed in the real world. Therefore, it is also an important issue to analyze the case where we cannot observe all variables from target dynamical systems.

Acknowledgments

This research is partially supported by the JSPS Grant-in-Aids for Scientific Research (Nos. 15KT0112, 17K00348, 18K18125).

Appendix

In this appendix, we describe the Otsu method. The Otsu method of binarization is a discriminant analysis method often used in the field of image processing [5]. In this paper, the estimated adjacency matrix is generated by applying Otsu binarization to the prediction accuracy matrix obtained by CCM. In case that the time series is obtained from connected vertices, the prediction accuracy of time series signals obtained by CCM converges to a high value if the length of the time series signals used for prediction is sufficient. On the other hand, for unconnected vertices, the prediction accuracy of time series signals obtained by CCM converges to 0 when the length of the time series signals used for the prediction is sufficient. Therefore, if this length is sufficiently long, the prediction accuracy of the time series signals obtained by CCM is expected to be a bimodal distribution. Therefore, the Otsu binarization method is effective in estimating the causality from the prediction accuracy of time series signals obtained by CCM. The algorithm of the Otsu binarization method is as follows.

- (1) The average value μ_0 , maximum value I_{\max} , and minimum value I_{\min} of all obtained prediction accuracies are calculated.
- (2) The threshold θ is defined in the range of $I_{\min} \leq \theta \leq I_{\max}$, and the data are classified into two classes, class 1 and class 2, by the threshold θ .

- (3) The number of data in class 1, n_1 , the average value of data in class 1 μ_1 , the variance of data in class 1 σ_1^2 , the number of data in class 2, n_2 , the average value of data in class 2 μ_2 , and the variance of data in class 2 σ_2^2 are calculated.
- (4) The within-class variance σ_w^2 , between-class variance σ_b^2 , and separation rate S are calculated.

$$\sigma_w^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2}, \quad (\text{A-1})$$

$$\sigma_b^2 = \frac{n_1(\mu_1 - \mu_0)^2 + n_2(\mu_2 - \mu_0)^2}{n_1 + n_2}, \quad (\text{A-2})$$

$$S = \frac{\sigma_b^2}{\sigma_w^2}. \quad (\text{A-3})$$

- (5) Steps (2) to (4) are repeated by changing the threshold value θ in the range of $I_{\min} \leq \theta \leq I_{\max}$. Then, the threshold value, when the separation rate S takes the maximum value S_{\max} , is used for discrimination of two classes: coupled and uncoupled.

References

- [1] C.W.J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, vol. 37, no. 3, pp. 424–438, 1969.
- [2] G. Sugihara, R. May, H. Ye, C. Hsieh, E. Deyle, M. Fogarty, and S. Munch, "Detecting causality in complex ecosystems," *Science*, vol. 338, no. 6106, pp. 496–500, 2012.
- [3] L. da F. Costa, F.A. Rodrigues, G. Travieso, and P.R. Villas Boas, "Characterization of complex networks: A survey of measurements," *Advances in physics*, vol. 56, no. 1, pp. 167–242, 2007.
- [4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, nos. 4–5, pp. 175–308, 2006.
- [5] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [6] N. Packard, J. Crutchfield, D. Farmer, and R. Shaw, "Geometry from a time series," *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.
- [7] F. Takens, "Detecting strange attractors in turbulence," *Lecture Notes in Mathematics*, pp. 366–381, 1981.
- [8] D.J. Watts and S.H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [9] S. Maslov and K. Sneppen, "Specificity and stability in topology of protein networks," *Science*, vol. 296, no. 5569, pp. 910–913, 2002.
- [10] A.L. Lloyd, "The coupled logistic map: a simple model for the effects of spatial heterogeneity on population dynamics," *Journal of Theoretical Biology*, vol. 173, no. 3, pp. 217–230, 1995.
- [11] S. Jalan and A. Singh, "Cluster synchronization in multiplex networks," *EPL (Europhysics Letters)*, vol. 113, no. 3, p. 30002, 2016.
- [12] K. Aihara, T. Takabe, and M. Toyoda, "Chaotic neural networks," *Physics Letters A*, vol. 144, nos. 6–7, pp. 333–340, 1990.
- [13] A.L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

解説論文

デジタルコンピュータ上での非線形力学系の実装について

—ベルヌーイシフト写像の場合—

How to Implement Nonlinear Dynamical Systems on Digital Computers:
The Bernoulli Shift Map Case

池口 徹 Tohru IKEGUCHI

アブストラクト カオス応答を示す非線形力学系の一つにベルヌーイシフト写像がある。ベルヌーイシフト写像は、傾きが2の線形写像を組み合わせた区分解非線形力学系であるが、その記述の単純さにより、様々な場面で用いられている。ベルヌーイシフト写像では、状態値を2倍し、その結果に応じて $\text{mod } 1$ 演算を施す。そのため、2進体系を用いる現在のデジタルコンピュータでベルヌーイシフト写像の実装するには注意が必要となる。本稿では、まず、デジタルコンピュータを用いた数値計算により、ベルヌーイシフト写像の実装の際に生じる問題について述べる。続いて、この問題を解決する方法として、ベルヌーイシフト写像により状態値を更新する際に、写像の傾きの大きさに微小な変位を与える方法と状態値に変位を与える方法を導入し、これらの性能について、解軌道、リアプノフ指数、不変測度に対する実験的な調査を行った結果に基づいて議論する。更に、本稿では、微小変位を与えることで実装したベルヌーイシフト写像を擬似乱数生成器とみなすことでカオス乱数を生成し、得られたカオス乱数の性能という側面からの検討も行う。具体的には、ベルヌーイシフト写像の応答をしきい値を用いて変換した0,1の2値乱数系列に対して、NIST検定による解析を行った結果を示している。

キーワード カオス、ベルヌーイシフト写像、コンピュータ上での実装、カオス乱数

Abstract The Bernoulli shift map is one of the nonlinear dynamical systems that show chaotic responses. The Bernoulli shift map is a piecewise linear dynamic system, but because of its simplicity of description, it is used in various situations. On the other hand, the essence of the dynamics of the Bernoulli shift map is that it doubles the state value. In other words, we must be careful when we implement the Bernoulli shift map on the current digital computers that use binary systems. In this article, we first describe an issue in the implementation of the response of the Bernoulli shift map with a digital computer. To solve this issue, we introduced two methods and investigated their performance: the first one is to give a small displacement to the slope of the map and the second one is to give a small displacement to the state value of the map. We evaluated these two methods by estimating Lyapunov exponents and invariant measures. In addition, we generated chaotic time series generated from the Bernoulli shift map by the methods and investigated the performance of the obtained chaotic time series as pseudorandom numbers, presenting the results of applying the NIST test to the chaotic times series.

Key words Chaos, Bernoulli shift map, Implementation on digital computers, Chaotic random numbers

1. はじめに

2019年3月に開催された電子情報通信学会総合大会で事件は発生した。NOLTA ソサイエティ、N-2 複雑コミュニケーションサイエンスの一般セッションである。座長は、法政大学の鳥飼弘幸先生で、発表者は、著者の研究室で2020年3月に修士号を取得した山崎凌氏であった。

このときの発表タイトルは、「カオス MIMO システムにおけるビットの利用法の改良とその性能評価」である。内容は、名古

屋工業大学の岡本英二先生が提案されているカオス MIMO システム^{(1)~(4)}の改良についてであった。カオス MIMO システムでは、変調部に代表的なカオス力学系の一つであるベルヌーイシフト写像が用いられており、通信性能と秘匿性を兼ね備えた優れたシステムである。

発表が終わるといつもどおり質疑応答に移り、会場から数件の質問が寄せられた。質疑応答の時間もそろそろ終わりに近づいた頃に、座長の鳥飼先生から以下の質問を頂いた^(注1)。

鳥飼先生: 僕も似たような質問なんですけど、ベルヌーイシフトは、どうやって実装されたんですか?

池口 徹 正員: フェロー 東京理科大学工学部情報工学科

E-mail tohru@rs.tus.ac.jp

Tohru IKEGUCHI, Fellow, Member (Dept. of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, 6-3-1 Niijuku, Katsushika-ku, Tokyo, 125-8585 Japan).

電子情報通信学会 基礎・境界ソサイエティ

Fundamentals Review Vol.14 No.3 pp.229-238 2021 年 1 月

©電子情報通信学会 2021

(注1): 以下のやり取りだけを読むと、誤解を生みかねないのであえて申し上げるが、山崎凌氏はとても優秀で、修士課程修了の最後の最後まで努力を続け、改良カオス MIMO システムの提案とその性能評価に関する内容を纏め、原著論文と投稿した。この論文は、NOLTA Journal に採録されている⁽⁵⁾。

山崎凌 : ベルヌーイシフトは、プログラムで実装しています...

会場 : 笑...

鳥飼先生: (すかさず、フォローで) 傾きは2のまま?

山崎凌 : 傾きは... あ、傾きの、えっと、mod 1をとる部分に微小な誤差を乗せることで、収束することを防いでいます。

鳥飼先生: でも、えっと、傾きは2のまままで...?

山崎凌 : はい。

鳥飼先生: 収めるときに、ちょっと動かしている感じがすかね? でも、そうだとすると... 系列を100回くらい回していましたよね。

山崎凌 : 用いた方法でも頻度分布を求めて、一様分布になることは確認しています。

鳥飼先生: そうすると、 2^{-100} くらいの精度が初期値にはないといけないと思うのですが、その辺りはどうなのでしょう。

鳥飼先生: ご自身も含めて、皆さん、どうされているんですかね。

2. ベルヌーイシフト写像

2.1 ベルヌーイシフト写像の定義

ベルヌーイシフト写像のダイナミクスは、式(1)で与えられる。式(1)において、 t は離散時間、 $x(t)$ は、時刻 t における状態値である。

$$\begin{aligned} x(t+1) &= 2x(t) \pmod{1} \\ &= \begin{cases} 2x(t) & (0 < x(t) < 0.5) \\ 2x(t) - 1 & (0.5 \leq x(t) < 1) \end{cases} \end{aligned} \quad (1)$$

ベルヌーイシフト写像のリターンマップを図1に示す。図1を見ると分かるように、ベルヌーイシフト写像は区分別形な、すなわち、非線形な力学系である。

2.2 ベルヌーイシフト写像のダイナミクス

式(1)を用いて、ベルヌーイシフト写像を時間発展させることを考える。まず、時刻 t におけるベルヌーイシフト写像の状態値を2進表記し、これを $x_B(t)$ とすると、初期状態は、

$$x_B(0) = 0.b_1b_2b_3 \cdots b_i b_{i+1} \cdots b_{i-1} b_i \cdots \quad (2)$$

と表すことができる。式(2)において、 $b_i = 0, 1$ である。この初期値 $x_B(0)$ を式(1)に従って写像する。まず、状態値を2倍するが、これは、2進数では状態値を左に1ビットシフトすることに相当するので、

$$x_B(1) = 2x_B(0) = b_1.b_2b_3 \cdots b_{i+1}b_{i+2} \cdots$$

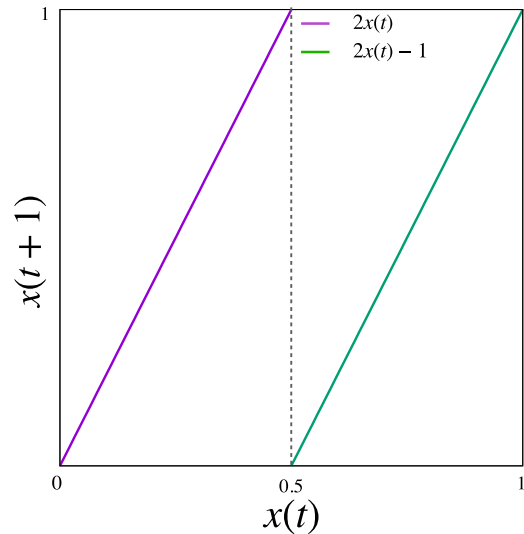


図1 ベルヌーイシフト写像(式(1))のリターンマップ

となる。 $b_1 = 0$ の場合は、 $0.b_2b_3 \cdots b_{i+1}b_{i+2} \cdots$ となる。一方、 $b_1 = 1$ の場合は、 $1.b_2b_3 \cdots b_{i+1}b_{i+2} \cdots$ となるが、これに対して mod 1を施すので、結局、ともに、 $0.b_2b_3 \cdots b_{i+1}b_{i+2} \cdots$ となる。 $t > 0$ も同様の操作が行われる。

ベルヌーイシフト写像は、 $[0, 1]$ の区間力学系であるが、初期値として $[0, 1]$ の有理数を採用した場合、上記のようなベルヌーイシフト写像のダイナミクスにより、最終的には、0か周期的な解が得られることになる。一方、初期値として $[0, 1]$ の無理数が採用された場合、 $x_B(0)$ が無限周期のビットパターンとなるため、これを初期値とした場合のベルヌーイシフト写像の解の系列も非周期的な値を取る。

$[0, 1]$ における有理数集合の濃度と無理数集合の濃度がそれぞれ可算無限、非可算無限であることを考えると、初期値として採用されるのは、ほとんどの場合、無理数ということになる。この結果、ベルヌーイシフト写像の解の系列は非周期的となる。

2.3 ベルヌーイシフト写像における初期値鋭敏依存性とリアプノフ指数

カオス力学系の特徴の一つに初期値鋭敏依存性がある。初期値鋭敏依存性とは、カオス力学系に与えた二つの初期値の微小な差が、時間発展に従って、指数関数的に拡大する性質である。

ベルヌーイシフト写像の場合は、以下のように考えることができる⁽⁶⁾。初期値 $t = 0$ においてベルヌーイシフト写像に与える二つの初期値を $x^{(1)}(0)$ 、 $x^{(2)}(0)$ とし、これらの二つの初期値の差が非常に小さく、例えば、 10^{-30} オーダであったとする。すなわち、

$$|x^{(1)}(0) - x^{(2)}(0)| \sim 10^{-30}$$

である。 $10^{-30} \approx 2^{-100}$ なので、状態値を2進表記で考えると、 $x_B^{(1)}(0)$ と $x_B^{(2)}(0)$ は、101ビットめで初めて異なるとしてよいだろう。つまり、

$$x_B^{(1)}(0) = 0.b_1b_2b_3b_4 \cdots b_{100} \mathbf{0} b_{102}b_{103} \cdots$$

$$x_B^{(2)}(0) = 0.b_1b_2b_3b_4 \cdots b_{100} \mathbf{1} b'_{102}b'_{103} \cdots$$

である。

既に説明したように、ベルヌーイシフト写像のダイナミクスは、左にビットシフトし、最上位ビットを切り捨てることと等価である。したがって、 $x_B^{(1)}(0)$ 、 $x_B^{(2)}(0)$ は、100 の写像後には、

$$x_B^{(1)}(100) = 0.\mathbf{0} b_{102}b_{103}b_{104} \cdots$$

$$x_B^{(2)}(100) = 0.\mathbf{1} b'_{102}b'_{103}b'_{104} \cdots$$

となり、これらの値の差は、

$$|x^{(1)}(100) - x^{(2)}(100)| \sim \frac{1}{2}$$

となる。すなわち、初期値においては、 10^{-30} オーダで我々には見えないであろう非常に微小な差が、時間発展とともに 10^{-1} のオーダにまで拡大し、その差を観測できるようになるということである。

さて、このようなカオス力学系に特有の特徴である初期値鋭敏依存性を定量化する指標がリアプノフ指数である。以下のようない離散時間力学系 F を考える。

$$x(t+1) = F(x(t)) \quad (3)$$

ベルヌーイシフト写像の場合、

$$F(x(t)) = 2x(t) \pmod{1}$$

となる。

いま、式 (3) における、時刻 t の状態値 $x(t)$ に対して微小な変位 $\delta(t)$ を与えると、この $\delta(t)$ は、時刻 $t+1$ においては、 $\delta(t+1)$ となるので、

$$x(t+1) + \delta(t+1) = F(x(t) + \delta(t)) \quad (4)$$

となる。 $\delta(t)$ は微小であるから、式 (4) の右辺をテラー展開し、二次以上の項を無視すると、

$$x(t+1) + \delta(t+1) = F(x(t)) + F'(x(t))\delta(t)$$

を得る。式 (3) を用いると、

$$\delta(t+1) = F'(x(t))\delta(t) \quad (5)$$

となる。これは、時刻 t から $t+1$ への時間発展により、微小変位 $\delta(t)$ は $F'(x(t))$ 倍されるということである。これが $t=0$ から $t=T$ まで繰り返されるが、その大きさを考えればよい。すなわち、

$$\left| \frac{\delta(T)}{\delta(0)} \right| = \prod_{t=0}^{T-1} |F'(x(t))|$$

である。

この伸びの大きさ $|\delta(T)/\delta(0)|$ が、指数関数的であるとする、

$$\prod_{t=0}^{T-1} |F'(x(t))| = e^{\lambda T} \quad (6)$$

と書くことができる。式 (6) の伸び率 λ が、リアプノフ指数と呼ばれる指標である。リアプノフ指数 λ を用いることで、カオス力学系の特徴の一つである軌道不安定性を定量的に評価することができる。

式 (6) の両辺の対数を取ることに、リアプノフ指数 λ は、以下の式 (7) で求めることができる。

$$\lambda = \frac{1}{T} \log \prod_{t=0}^{T-1} |F'(x(t))| = \frac{1}{T} \sum_{t=0}^{T-1} \log |F'(x(t))| \quad (7)$$

式 (1) を見ると分かるようにベルヌーイシフト写像の場合、 $|F'(x(t))| = 2$ である。したがって、式 (7) を用いると、ベルヌーイシフト写像のリアプノフ指数は、 $\lambda = \log 2$ と求めることができる。

2.4 ベルヌーイシフト写像が収束するメカニズム

本稿で対象としているベルヌーイシフト写像のようなカオス力学系の応答をデジタルコンピュータで実装する場合、注意すべき点がある。デジタルコンピュータでは、有限精度の状態値を用いざるを得ないため、状態値が最終的には 0 に収束してしまうからである。

図 2 は、浮動小数点数の精度を、単精度 (以下、float 型)、倍精度 (以下、double 型)、拡張倍精度 (以下、long double 型) で求めたベルヌーイシフト写像の時系列である。初期値は 1×10^{-1} としている。横軸は時刻 t 、縦軸が時刻 t における状態値 $x(t)$ である。図 2 を見ると分かるように、いずれの場合においても、数回の写像後に状態値が 0 に収束していることが分かる。

このようにデジタルコンピュータを用いて、式 (1) のベルヌーイシフト写像を時間発展させると、図 1 のように収束する。これは以下の理由によるものである。

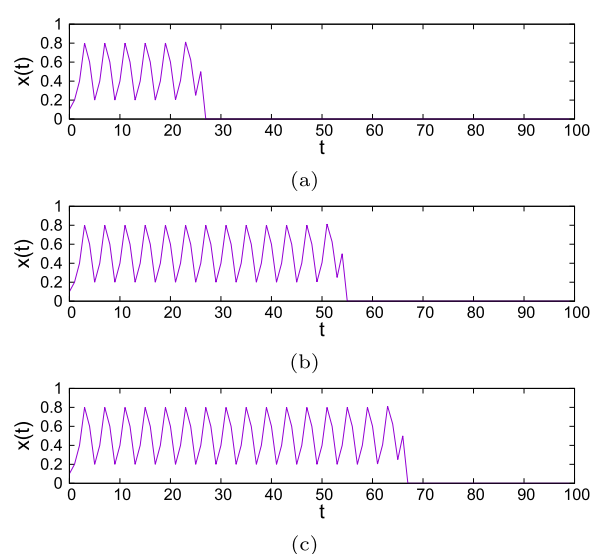


図 2 浮動小数点数の精度を (a) float 型、(b) double 型、(c) long double 型として、ベルヌーイシフト写像を定義どおりに実装した場合に得られる $x(t)$ の時系列 (初期値は、 1×10^{-1} としている)

まず、時刻 t におけるベルヌーイシフト写像の状態値を 2 進表記し、これを $x_B(t)$ とすると、初期状態は、

$$x_B(0) = 0.b_1b_2b_3 \cdots b_i b_{i+1} \cdots b_{n-1}b_n \quad (8)$$

と表すことができる。式 (8) における n は、状態値を表現するために用いられている浮動小数点数の仮数部のビット長に相当する。

この初期値 $x_B(0)$ を、式 (1) に従って写像するので、

$$x_B(1) = 2x_B(0) \pmod{1} = 0.b_2b_3 \cdots b_{i+1}b_{i+2} \cdots b_n0$$

となる。 $t > 0$ も同様の操作が行われるので、 n 回の写像を行うと、

$$x_B(1) = 2x_B(0) \pmod{1} = 0.b_2b_3 \cdots b_{i+1}b_{i+2} \cdots b_n0$$

$$x_B(2) = 2x_B(1) \pmod{1} = 0.b_3b_4 \cdots b_{i+2}b_{i+3} \cdots b_n00$$

⋮

$$x_B(n) = 2x_B(n-1) \pmod{1} = 0.00 \cdots 000$$

となる。すなわち、ベルヌーイシフト写像の繰り返しによって状態値は上記のように、最大 n 回の写像で 0 に収束する。

2.5 変位を与えたベルヌーイシフト写像

ベルヌーイシフト写像は理論的にはカオス応答を示すが、図 2 のように、数値計算により、ベルヌーイシフト写像を実装すると、有限回の写像の繰り返しで状態値は 0 に収束する。これを防ぐために、

- 傾きの大きさ 2
- 状態値に対する mod 1 演算における 1

に微小な変位を加えるという方法を考えることができる。ベルヌーイシフト写像では、現在の状態値を 2 倍することがビットシフトに対応するため、状態値を 2 倍する際に微小変位を加え、実際の演算時には、傾きが 2 からずれた値にしてこれを防ぐというのが、前者の方法の主張である。一方、後者の考えは、ベルヌーイシフト写像の状態値を 2 倍した後に施す mod 1 演算において、微小な変位を加えることで 0 への収束を抑えるというものである。なお、以下では、前者を傾きに変位を与えた、後者を状態値に変位を与えたと呼ぶことにする。前述のカオス MIMO システムの研究^{(1)~(4)}では、ベルヌーイシフト写像を実装する際には、状態値に変位を与える方法が採用されている。

本稿では、ベルヌーイシフト写像の傾き、状態値に変位を与える場合についての調査結果を述べる。まず、ベルヌーイシフト写像の傾き 2 に対して、微小変位 a を与え、 $2-a (\equiv a')$ とした場合は以下のようなになる。

$$x(t+1) = a'x(t) \pmod{1} \quad (9)$$

$$= \begin{cases} a'x(t) & \left(0 < x(t) < \frac{1}{a'}\right) \\ a'x(t) - 1 & \left(\frac{1}{a'} \leq x(t) \leq 1\right) \end{cases}$$

この場合のリターンマップは、図 3(a) である。次に、状態値を mod 1 演算する際に微小変位 ε を与え、 $1-\varepsilon (\equiv \varepsilon')$ で mod 演

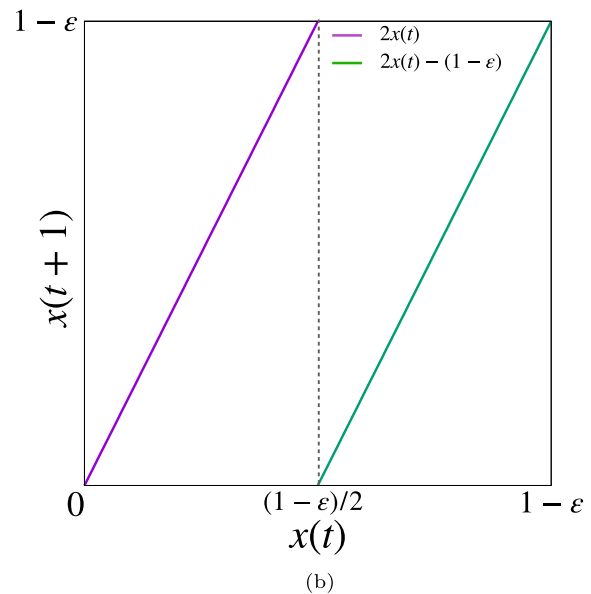
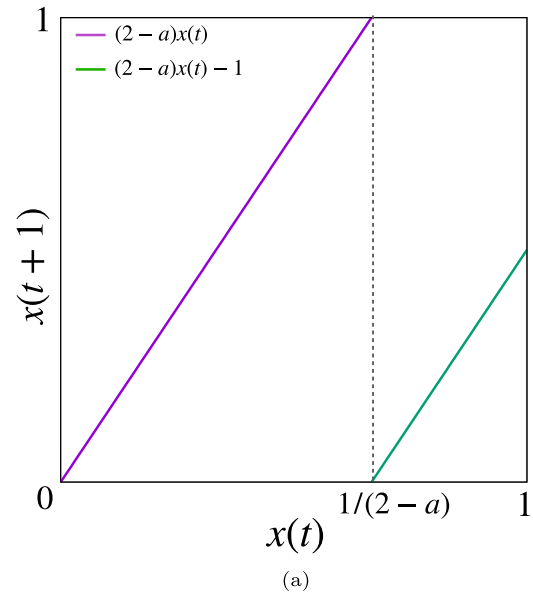


図 3 (a) 傾き、(b) 状態値に変位を与えたベルヌーイシフト写像のリターンマップ

算する場合は以下のようなになる。

$$x(t+1) = 2x(t) \pmod{\varepsilon'} \quad (10)$$

$$= \begin{cases} 2x(t) & \left(0 < x(t) < \frac{\varepsilon'}{2}\right) \\ 2x(t) - \varepsilon' & \left(\frac{\varepsilon'}{2} \leq x(t) < \varepsilon'\right) \end{cases}$$

この場合のリターンマップは、図 3(b) である。

これに加えて、傾きと状態値の mod 演算の両者に変位を与えることも可能であるが、本稿では、傾きに変位を与える場合と状態値に変位を与える場合の比較について考察するために、式 (9) と式 (10) を用いた場合の結果を述べている。具体的には、これら 2 種類の変位の与え方について、その大きさを変えた場合に得られる解軌道、不変測度、リアプノフ指数を用いた評価

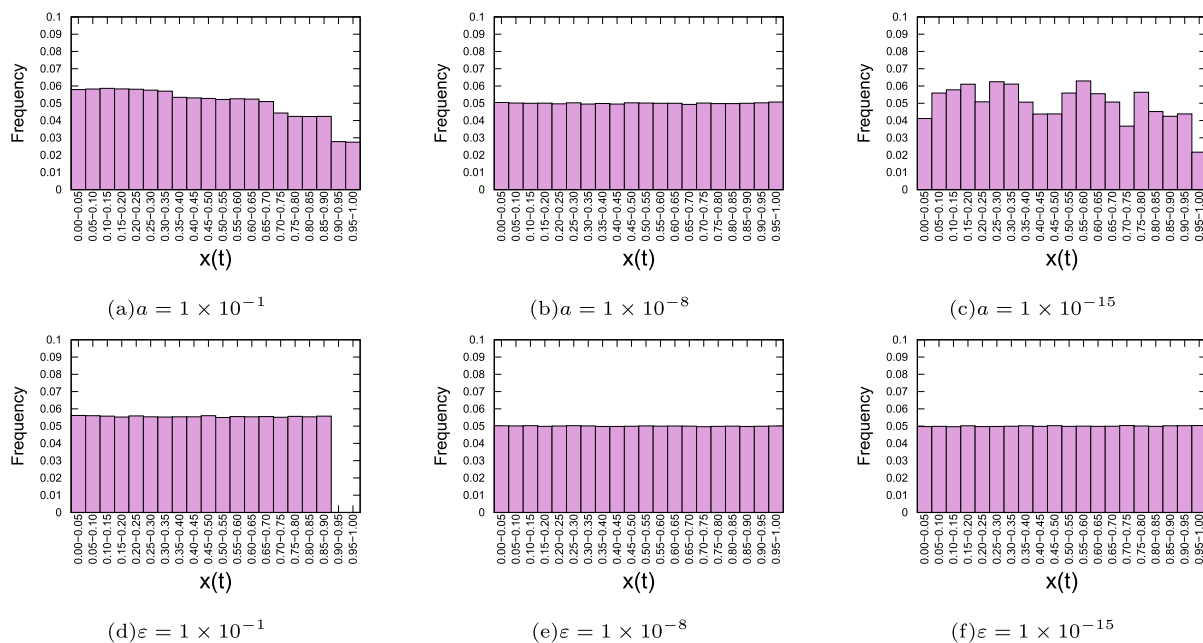


図4 浮動小数点数の精度を double 型, 初期値 1×10^{-1} とした場合に, 傾きと状態値に変位 a, ϵ を与えたベルヌーイシフト写像の不変測度の例

を行う. 更に本稿では, ベルヌーイシフト写像を擬似乱数生成器と考え, しきい値を用いて解軌道から得られる状態値を 0, 1 の 2 値に変換した擬似乱数の統計的性質という観点から評価を行っている.

3. 変位を与えたベルヌーイシフト写像の性能

まず, 傾きと状態値に変位を与えた場合について, どのような応答となるのかを不変測度を観測することにより調査する. 解軌道 $(x(t), t = 0, 1, 2, \dots)$ が与えられたときに, その密度関数を考えることができるが, これを不変測度という. ベルヌーイシフト写像の不変測度は, $[0, 1]$ の一様分布となることが知られている.

図 4 に, 浮動小数点数の精度を double 型とした場合の不変測度の結果例を示す. 図 4 (a), (b), (c) は, ベルヌーイシフト写像の傾きを 2 から $2-a$ とした場合, 図 4 (d), (e), (f) は, ベルヌーイシフト写像の状態値に対する mod 演算を, $1-\epsilon$ で行った場合である. 図 4 では, 初期値を $x(0) = 1.0 \times 10^{-10}$ とし, 過渡状態として 100,000 回を省いた後, 1,000,000 回を算出している.

まず, 図 4 (a), (b), (c) を見ると, 傾きに変位を与え, $2-a$ とすると, 変位の大きさが大きすぎると ($a = 1 \times 10^{-1}$), 不偏測度に偏りが生じることが分かる. これは図 3 (a) を見ると分かるように, 写像の区分点が $2-a$ となるためである. また, $a = 1 \times 10^{-16}$ とすると, 変位を与えない場合と同様に, 状態値 $x(t)$ が, 十分な時間経過とともに 0 に収束する結果となった. これは, double 型の場合の有効桁数は, 小数点以下第 15 桁であるため, a に 1×10^{-16} の変位を与えても, 実際は変位を与えていない場合の結果と同じになるためである.

次に, 図 4 (d), (e), (f) を見ると, 状態値に変位を与えた場

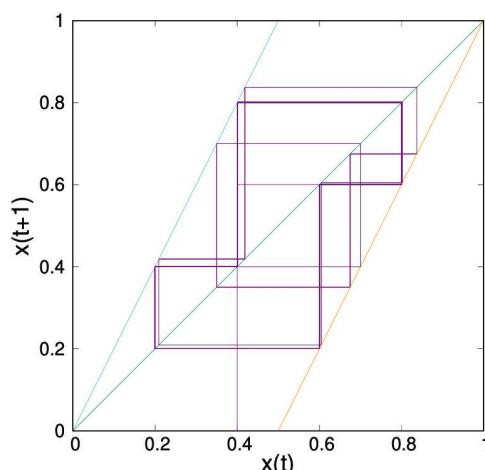


図5 状態値に変位 $\epsilon = 1 \times 10^{-16}$ を与えた場合のリターンマップ

合は, 一様な不偏測度が得られている. 図 4 (d) において, 区間 $[0.90, 1.00]$ 状態値の頻度が 0 となっているが, これは状態値に変位 ϵ を与えることにより, $x(t)$ の定義域が, $[0, 1]$ から $[0, 1-\epsilon]$ になったためである.

なお, $\epsilon = 1 \times 10^{-16}$ の場合は, 頻度分布が著しく偏る結果を得た. これは, 変位を加えているにもかかわらず, 周期応答となってしまったためである. 実際, $\epsilon = 1 \times 10^{-16}$ の場合について, 数値実験により得られた $x(t)$ を用いて, 図式解法として示した結果が図 5 である. この結果からも周期応答となることが確認できる. 図 4 は, 浮動小数点数の精度を double 型とした結果であるが, float 型, long double 型とした場合でも, 各精度の有向桁に応じた微小な ϵ を与えると同様の結果となる.

また, 初期値 1×10^{-1} , 浮動小数点数の精度を double 型,

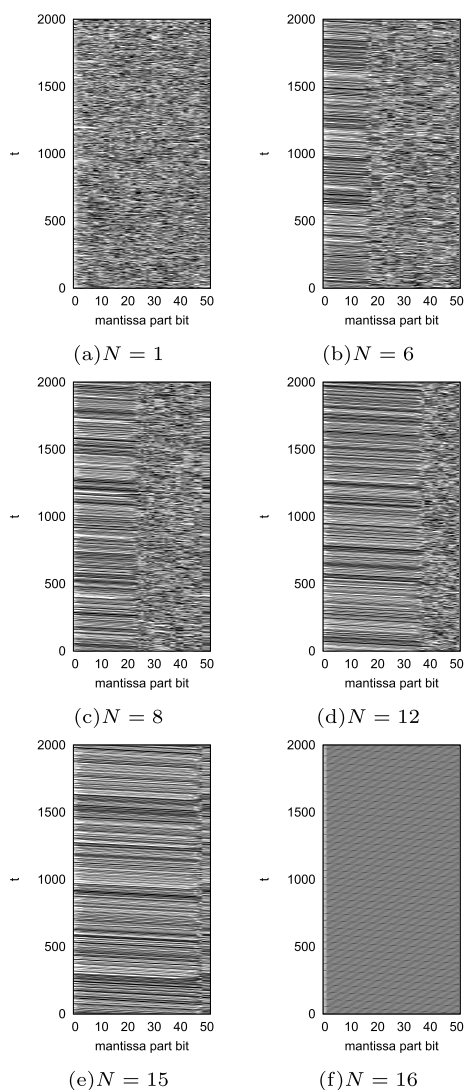


図6 浮動小数点数の精度 double 型, 初期値 1×10^{-1} とし, 状態値に大きさが 1×10^{-N} の変位を与えた場合の状態値の時間変化のビット表現

状態値に変位を与えた場合について, 状態値のビットパターンの時間変化を調査した. 図6に結果を示す. 図6では, 縦軸に写像回数, 横軸に仮数部のビット番号とし, 当該ビットが0ならば白, 1ならば黒としてプロットしている.

図6より, 状態値に変位を与えることで非周期的な応答が得られる場合は, 図の左側は斜線が出ており, 右側はモザイク状になっていることが分かる. これは, 図の斜線部では左1ビットシフトにより状態値の遷移が斜め線として現れるためである. また, モザイク状の部分では変位を与えたここにより, ビットパターンにランダムな変化が生じることを示している. 例えば, 変位 $\varepsilon = 1 \times 10^{-6}$ を与えた場合(図6(b)), $1 \times 10^{-6} \approx 0.1 \dots \times 2^{-19}$ であるため, すなわち, 状態値に変位 $\varepsilon = 1 \times 10^{-6}$ を与える場合, 仮数部の上位19ビットには変位が乗らず, 単なる左1ビットシフトとなる. 一方, 20ビット目以降には変位が加えられるため, ビットシフトした値が不規則なビットパターンとなる. また, 変位 $\varepsilon = 1 \times 10^{-16}$ を与え

表1 状態値のビット数が6ビット, 状態値の変化が周期的になる一例

t	状態値 $x_B(t)$
0	0.100101
1	0.001110
2	0.011100
3	0.111000
4	0.110100
5	0.101100
6	0.011100
7	0.111000
8	0.110100
\vdots	\vdots

た場合(図6(f))では, 規則的な斜め線が出現しており, 周期的な状態変化となっていることが分かる.

これらの結果は, 傾きと状態値に変位を加えた場合に, 非周期的な応答が得られるかどうかを調査したものである. これらの結果を見ると, 傾きに変位を与えるよりも, 状態値に変位を与える場合に安定して非周期的な応答が生み出されることが分かる. しかし, 状態値に変位を加える場合にも注意すべき点がある. それは, 状態値に与える変位の大きさとして $\varepsilon = 1 \times 2^{-i}$ (2進数表現で第*i*ビット目が1でそれ以外は0)を与えると応答が周期的となるということである.

簡単のために, 仮数部が6ビット, 変位の大きさが $\varepsilon = 6.25 \times 10^{-2} = 0.000100_{(2)}$ の場合を考える. 式(10)より, 写像を1回適用する際には,

- (i) 状態値 $x(t)$ を2倍する
- (ii) $\text{mod}(1 - \varepsilon)$ 演算する

という二つの操作を行う. (i)の操作は, 2進数表現した場合に, 状態値 $x_B(t)$ を1ビット左シフトする. (ii)の操作は, ビットシフト前の $x_B(t)$ の小数第一位のビットが1のとき実行され, ビットシフト後に一の位へ移動したビットを反転(0)し, ε を加算する. 一方, ビットシフト前の $x_B(t)$ の小数第一位のビットが0のときは(ii)の操作は行われぬ.

今, ある状態値 $x_B(0) = 0.100101$ に対して, (i), (ii)の操作を行うことを考える. $x(0) = 0.100101_{(2)}$ は小数第一位のビットが1なので, (i)の操作後に(ii)の操作を行う. 今, $\varepsilon = 0.000100_{(2)}$ なので, (i)により左に1ビットシフトされ, $1.001010_{(2)}$ となった後, (ii)により, $0.001010_{(2)} + \varepsilon = 0.001010_{(2)} + 0.000100_{(2)} \rightarrow 0.001110_{(2)}$ となる. この一連の操作により状態値が $x_B(0) = 0.100101$ から $x_B(1) = 0.001110$ に遷移した.

次に, $x_B(1) = 0.001110$ では, 小数第一位のビットは0なので, (i)の操作のみを行い, $x_B(2) = 0.011100$ に遷移する. 以降, 同様の操作で状態値が遷移していく. これをまとめたものが, 表1である.

表1を見ると, 状態値 $x_B(2) = x_B(6)$ となること, すなわち4周期となることが分かる. 今, 仮数部の上位4ビットに注目すると, 見かけ上は小数第一位のビットが, 写像によって小数第四位の位置に移動することになる. これは, $x_B(t)$ の小数第

表 2 浮動小数点数の精度を double 型、ベルヌーイ写像の初期値を 1×10^{-1} とした場合に、傾き、状態値に与えた変位の大きさとリアプノフ指数の関係（結果は小数第三位まで示している）

変位の大きさ	変位の与え方	
	傾き	状態値
1×10^{-1}	0.642	0.693
1×10^{-2}	0.688	0.693
1×10^{-3}	0.693	0.693
1×10^{-4}	0.693	0.693
1×10^{-5}	0.693	0.693
1×10^{-6}	0.693	0.693
1×10^{-7}	0.693	0.693
1×10^{-8}	0.693	0.693
1×10^{-9}	0.693	0.693
1×10^{-10}	0.693	0.693
1×10^{-11}	0.693	0.693
1×10^{-12}	0.693	0.693
1×10^{-13}	0.693	0.693
1×10^{-14}	0.693	0.693
1×10^{-15}	0.693	0.693

一位のビットが 0 の場合は、(i) により小数第四位のビットに 0 が立ち、1 の場合は (i), (ii) により 1 が立つためである。

最後に、表 2 に、傾きと状態値に与えた変位の大きさと推定したリアプノフ指数の関係を示す。なお、表 2 では応答が周期的とならない場合についての結果をまとめている。

既に述べたように一次元写像のリアプノフ指数は、写像関数 F (式 (7)) の微分値により求めることができる。すなわち、ベルヌーイシフト写像の傾きに変位 a を与え、 $2 - a$ とすると、 F' の大きさも変化するが、 $\log |2 - 1 \times 10^{-1}| \approx 0.642$ 、 $\log |2 - 1 \times 10^{-2}| \approx 0.688$ 、 $\log |2 - 1 \times 10^{-3}| \approx 0.693$ となるためである。一方、状態値に変位を与える場合は、傾きは 2 のままであるため、リアプノフ指数の大きさは、 $\log 2 = 0.693$ となる。この結果は、浮動小数点数の精度が double 型の結果であるが、float 型、long double 型でも同様の結果となっている。

4. NIST 検定

乱数列のランダム性を評価するために、国際標準的な統計検定方式がアメリカ国立標準技術研究所 (National Institute of Standards and Technology, NIST) から提供されている。本稿では、NIST が公開している NIST Special Publication 800-22⁽⁷⁾ (以下、NIST 検定) を使用して、ベルヌーイシフト写像から生成した乱数列の性能評価を行う。NIST 検定は、全 15 項目の統計検定から構成されている。具体的な各検定項目とその内容を表 3 に示す⁽⁷⁾。

NIST 検定では検定対象の乱数列長に応じて、パラメータを設定することが推奨されている。これは、乱数列長に応じて適切な検定結果を得るために必要な処理である。本稿で用いたパラメータ設定を表 4 に示す。表 4 の有意水準 α とは、検定対象となる乱数列がランダムでない場合の確率である。また、ブロッ

表 3 NIST 検定の項目内容⁽⁷⁾

No.	検定項目	検定内容
1	頻度検定	2 値乱数 (0, 1) のうち、1 の出現確率が 1/2 に近いに近いかどうかを検定
2	ブロック単位の頻度検定	あるブロック長に対して、1 の出現確率が 1/2 に近いに近いかどうかを検定
3	累積和検定	2 値乱数を (+1, -1) に変換し、その累積和で定義されるランダムウォークの 0 からの最大偏移を検定
4	連の検定	連続した 1 の数 (連) を検定
5	ブロック単位の最長連検定	あるブロック長に対して、最長の連を検定
6	2 値行列ランク検定	数列全体の互いに疎な行列のランクを検定
7	離散フーリエ変換検定	周期性の調査のため、数列の離散フーリエ変換によるピークの高さを検定
8	重なりのないテンプレート適合検定	あらかじめ用意されたビット列 (テンプレート) の出現回数を検定 (重なりなしで検定)
9	重なりのあるテンプレート適合検定	あらかじめ用意されたビット列 (テンプレート) の出現回数を検定 (重なりありで検定)
10	マウラーのユニバーサル統計検定	数列内で一致するパターンを検出し、圧縮することができるかを検定
11	近似エントロピー検定	あるビット長の出現パターンの頻度を検定
12	ランダム偏差検定	累積和とランダムウォークで K 回訪れるサイクル数を検定
13	種々のランダム偏差検定	累積和とランダムウォークで特定の状態を訪れる回数合計を検定
14	系列検定	あるビット長のすべての重複ビットパターンの出現頻度を検定
15	線形複雑度検定	乱数を生成するための線形フィードバックシフトレジスタの長さを検定

表 4 NIST 検定で用いたパラメータ項目

項目	パラメータ値
有意水準 α	0.01
ブロック単位の頻度検定 - ブロック長	16,384
重なりのないテンプレート適合検定 - ブロック長	9
重なりのあるテンプレート適合検定 - ブロック長	9
近似エントロピー検定 - ブロックの長さ	10
系列検定 - ブロックの長さ	16
線形複雑度検定 - ブロックの長さ	500

ク長とは、各検定を行う際の単位ブロック長である。

また、NIST 検定では、1 Gbit 以上の乱数長を要求されるが、1 Gbit 全体を検定するわけではなく、1 Mbit \times 1,000 のように、長さ 1,000,000 (= 1 Mbit) の系列を 1,000 個に分けて、全検定項目を実施する。このとき、各検定結果に対してそれぞれ 1,000 個の p 値が求まる。一つの検定項目に対して、1,000 個の p 値を用いた適合度検定より P 値が算出される。この P 値により、1,000 個の p 値のばらつきについて、一様性を評価する。同時に、1,000 個の p 値のうち、 p 値が有意水準 α について、 p 値 $> \alpha$ となる比率も評価する。本稿では、各検定項目について、1,000 個の p 値の一様性と比率ともに合格した場合のみ、その検定項目に合格したとしている。

本稿で行った、具体的な検定の手順は以下のとおりである。まず、式 (9), (10) により、傾きと状態値に変位を与えたベルヌーイ写像より算出した $x(t)$ に対して、以下により 2 値乱数を生成する。

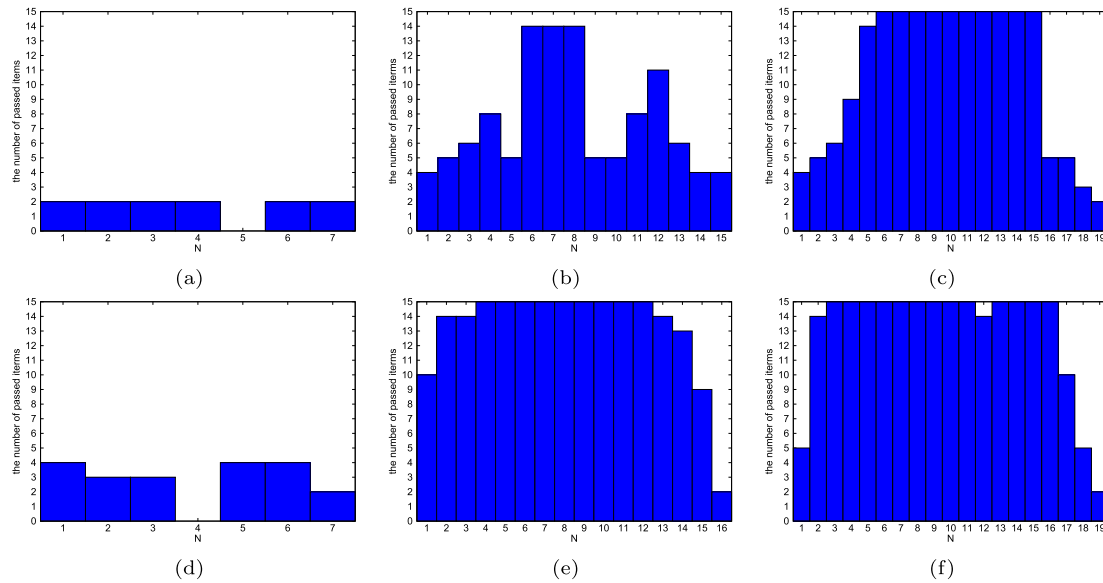


図7 初期値 1.0×10^{-3} とした場合のベルヌーイ写像の解軌道 $x(t)$ から生成した 0-1 擬似乱数に対する NIST 検定合格数結果 (左列 (a), (d) は float 型, 中列 (b), (e) は double 型, 右列 (c), (f) は long double 型を用いて, 上段 (a), (b), (c) は傾きに変位を与えた場合, 下段 (d), (e), (f) は状態値に変位を与えた場合である)

$$B(t) = \begin{cases} 0 & x(t) \leq \theta \\ 1 & x(t) > \theta \end{cases} \quad (11)$$

ここで, $\theta = (1 - \varepsilon)/2$ である.

ベルヌーイシフト写像の解軌道を生成する際には,

- 初期値 $x(0) = 1 \times 10^{-k}$ ($k = 1, 2, \dots, 16$)
- 浮動小数点数精度: float 型, double 型, long double 型
- 過渡状態: $0 \leq t \leq 100,000$
- 乱数長: 2^{30}
- 変位: $a = 1 \times 10^{-N}, \varepsilon = 1 \times 10^{-N}$ ($N = 1, 2, \dots, 16$)

とした.

変位の値の大きさと NIST 検定合格数の関係を, 図 7 に示す. 図 7 の各図において, 横軸は変位の大きさ (1×10^{-N}), 縦軸は NIST 検定合格数である. 図 7 を見ると, まず, 浮動小数点数として float 型を採用した場合 (図 7 (a), (d)), どのように変位を与えても, 検定の合格数は少ないことが分かる. 次に, 浮動小数点数として double 型と long double 型を採用した場合は (図 7 (b), (c), (e), (f)), N が増加するにつれて検定合格数も増加し, $N = 7 \sim 9$ 付近で最大となり, 更に N が増加すると検定合格数は減少する傾向を示している. 特に, 状態値に変位を与えた場合 (図 7 (f)), その大きさが $1 \times 10^{-4} \sim 1 \times 10^{-12}$ であれば, ほぼ全ての検定項目に合格していることが分かる. また, long double 型の場合は, $N = 2 \sim 13$ の広い範囲で検定合格数は 14 または 15 となっており, 乱数の性能が安定している.

この結果は, ベルヌーイシフト写像の初期値を 1.0×10^{-3} とした場合のものであるが, 初期値をほかの値とした場合も, これらの結果の傾向に変化は見られなかった. これらの結果より, ベルヌーイシフト写像を擬似乱数生成器として用いる場合, 変位を状態値のみに与える方法が最も安定した性能を有する乱数を生成できると考えられる.

5. おわりに

本稿では, 非線形力学系の応答を解析する際にデジタルコンピュータを用いた数値計算により生じる問題とその解決策に関する議論を行った. まず, 非線形力学系の例としてベルヌーイシフト写像を導入し, ベルヌーイシフト写像のダイナミクスに起因して, 有限精度の状態値を用いざるを得ないデジタルコンピュータでの実装では, 写像の状態値が最終的には 0 に収束することを示した.

次に, この問題を解決する方法として, ベルヌーイシフト写像の傾きと状態値に微小な変位を与える手法を導入し, そのダイナミクスを定式化した. 更に, 変位を与えたベルヌーイシフト写像より得た解軌道, 不変測度, リアプノフ指数を調査した. その結果, 傾き, あるいは, 状態値に変位を加えることによって, 作り出される応答は非周期的とできることを確認した. しかし, 変位の大きさによっては, 周期的となる場合があることも示した. また, その理由を明らかにするため, 状態値を 2 進数で表現し, 各時刻での状態値のビット列が, ベルヌーイシフト写像によりどのように遷移するかも調査した. その結果, 状態値に変位を与える場合に周期的応答となる仕組みが, ビットシフトと mod 演算により状態値の上位 i ビットでビットパターンのループにより生じること, したがって, 変位 ε が 1×2^{-i} となる場合に応答が周期的になることを示した. これより, デジタルコンピュータを用いた数値計算においてベルヌーイシフト写像を実装するためには, 状態値へ与える変位の大きさを $\varepsilon \neq 1 \times 2^{-i}$ とすることが必須であることが分かる.

また, 本稿では, 変位を与えることで実装したベルヌーイシフト写像の性能を, 擬似乱数の性能という観点からも解析した. 具体的には, 変位を与えることで実装したベルヌーイシフト写

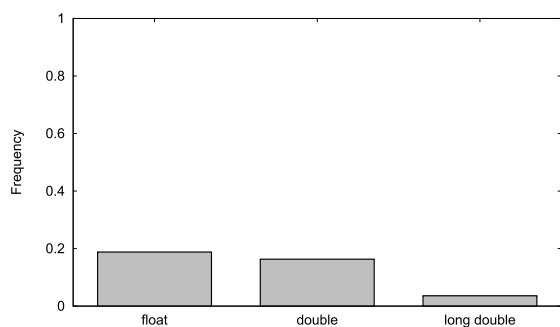


図8 ロジスティック写像に対して、float 型、double 型、long double 型を用いた場合にカオス応答が継続する割合

像の解軌道 $x(t)$ を 0, 1 の二値乱数に変換した系列を擬似乱数とみなし、乱数としての性能を NIST 検定を用いて調査した。その際、ベルヌーイシフト写像について、浮動小数点数の精度を float 型、double 型、long double 型として実装した場合について、得られる擬似乱数の性能に変化が見られるかを調査した。

その結果、浮動小数点数の精度が float 型の場合は、double 型、long double 型の場合と比較して、検定合格数が極端に減少すること、また、double 型、long double 型の場合は、状態値に変位を与えると検定に合格しやすいことが明らかとなった。具体的には、double 型の場合、状態値に与える変位の大きさを $1 \times 10^{-5} \sim 1 \times 10^{-12}$ としたとき、検定合格数がピークに達することを明らかにした。

一方、浮動小数点数の精度が long double 型の場合は、変位を傾き、状態値に与えた場合ともに検定合格数が増加することが分かった。ただし、傾きに変位を与えた場合の検定合格数がピークとなる範囲は $1 \times 10^{-7} \sim 1 \times 10^{-15}$ であったのに対し、状態値に与えた場合は、 $1 \times 10^{-3} \sim 1 \times 10^{-17}$ と広い範囲で検定合格数が高い値となった。

以上の擬似乱数の立場の解析結果から、ベルヌーイシフト写像を擬似乱数生成器として用いる場合、浮動小数点数の精度を double 型または long double 型とし、状態値へ与える変位の大きさを $1 \times 10^{-7} \sim 1 \times 10^{-13}$ とすると、NIST 検定に合格しやすい安定した擬似乱数を得ることができる。

本稿では、ベルヌーイシフト写像をデジタルコンピュータ上で実装する際、理論どおりに実装すると状態値が 0 に収束することをいかに防ぐか、という観点からの議論を行った。今回はベルヌーイシフト写像を対象としたが、ベルヌーイシフト写像と位相共役なロジスティック写像⁽⁸⁾

$$x(t+1) = ax(t)(1-x(t)) \quad (12)$$

でも同様な収束を観測できる。

例えば、図8は、式(12)のロジスティック写像において、 $a = 4$ とした際に、浮動小数点数の精度を float 型、double 型、long double 型とした場合に、状態値がどの程度収束しないのかを調査した結果である。初期値 $x(0)$ を 100 万個用意し、写像を 10 億回繰り返した場合に、応答が収束した割合を示している。100 万個の初期値は $[0, 1]$ に等間隔で配置して実験を行った。図8をみると、浮動小数点数の精度を float 型にした場合で約 18.8%、

double 型にした場合で約 16.3%、また、long double 型にした場合でも約 3.6%が収束してしまうことが分かる。このような問題に対しては、整数化による精度保証^{(9)~(14)}などの対応がある。今後も様々な側面から検討すべき重要な課題の一つであると考えられる。

本稿の内容に関して、様々ご議論をいただいた山崎凌氏、金丸志生氏、鈴木大樹氏に感謝します。

文 献

- (1) E. Okamoto, "A chaos MIMO transmission scheme for channel coding and physical-layer security," *IEICE Trans. Commun.*, vol. E95-B, no. 4, pp. 1384–1392, 2012.
- (2) E. Okamoto and Y. Inaba, "Multilevel modulated chaos MIMO transmission scheme with physical layer security," *NOLTA*, vol. 5, no. 2, pp. 140–156, 2014.
- (3) E. Okamoto and N. Horiike, "Performance improvement of chaos MIMO scheme using advanced stochastic characteristics," *IEICE Commun. Express*, vol. 5, no. 10, pp. 371–377, 2016.
- (4) 岡本英二, "無線通信システムに置ける非線形信号処理を用いた性能向上," *IEICE Fundamental Review*, vol. 9, no. 4, pp. 350–363, 2016.
- (5) R. Yamazaki, Y. Shimada, and T. Ikeguchi, "Chaos MIMO system with efficient use of information of bits," *NOLTA*, vol. 11, no. 4, pp. 546–560, 2020.
- (6) 合原一幸, *カオス学入門*, 放送大学教材, 2001.
- (7) A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray, and S. Vo, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," *National Institute of Standards and Technology, Special Publication 800-22*, 2010.
- (8) R. May, "Simple mathematical model with very complicated dynamics," *Nature*, vol. 261, no. 5560, pp. 459–467, 1976.
- (9) 董際国, 森田啓義, "整数ロジスティック写像と攪拌演算による乱数生成," *信学論 (A)*, vol. J94-A, no. 12, pp. 923–931, 2011.
- (10) 董際国, 森田啓義, "整数ロジスティック写像の諸性質: 発散, 収束, 周期性," *信学論 (A)*, vol. J96-A, no. 2, pp. 90–99, 2013.
- (11) T. Miyazaki, S. Araki, and S. Uehara, "Some properties of logistic maps over integers," *IEICE Trans. Fundamentals*, vol. E93-A, no. 11, pp. 2258–2265, 2010.
- (12) T. Miyazaki, S. Araki, Y. Nogami, and S. Uehara, "Rounding logistic maps over integers and the properties of the generated sequences," *IEICE Trans. Fundamentals*, vol. E94-A, no. 9, pp. 1817–1825, 2011.
- (13) S. Araki, T. Miyazaki, and S. Uehara, "A study on precision of pseudorandom number generators using the logistic maps," *Proc. ISITA2012*, pp. 740–743, 2012.
- (14) 金丸志生, 島田裕, 藤原寛太郎, 池口徹, "整数ロジスティック写像を用いたカオス乱数と NIST 検定による性能評価," *信学技報*, vol. 118, no. 413, NLP2018–100, pp. 23–28, 2019.

(NLP 研究会提案, 2020 年 8 月 6 日受付,
2020 年 10 月 5 日再受付)



池口 徹 (正員：フェロー)

1990 東京理科大学大学院理工学研究科電気工学専攻了。同年，同大学基礎工学部助手。1997 ルイ・バスツール大学 (現ストラスブール第一大学，フランスストラスブール市) 客員教授。2000 埼玉大学大学院理工学研究科助教授，2005 同教授を経て，2014 東京理科大学工学部第一部教授。博士 (工学)。非線形時系列解析，計算論的脳神経科学，複雑ネットワーク理論，カオスダイナミクスを用いた離散最適化問題のメタ・ヒューリスティック解法の開発，非線形振動子の同期現象などの研究に従事。NOLTA 2012 共同実行委員長，2013 非線形問題研究専門委員会委員長。2014 NOLTA ソサイエティ庶務幹事，NOLTA 2016 実行委員長，2017 NOLTA ソサイエティ会長。

査読あり国際会議

Reconstruction of Dynamical System from Marked Point Process and its Application to Causality Detection

Kazuya Sawada¹, Yutaka Shimada² and Tohru Ikeguchi¹

¹ Tokyo University of Science
 6-3-1 Nijjuku, Katushika-ku, Tokyo 125-8585, Japan
 Phone/FAX:+81-03-5876-1717
 E-mail: sawada@hisekei.net

² Saitama University
 255 Shimo-Okubo, Sakura-ku, Saitama-shi,
 Saitama 338-8570, Japan

Abstract

We have already proposed a method for detecting causality for point process data, especially for marked point process data. Our proposed method needs to reconstruct a state space from an observed marked point process, but it has not been revealed that we can reconstruct the state space from the marked point process. Therefore, in this paper, we numerically verified whether the state space could be reconstructed from the marked point process using a delayed coordinate system. In particular, we evaluated a similarity between the inter-point distance distributions in the state space of the original dynamical system that generated the marked point process and the state space reconstructed from the marked point process. As a result, we can reconstruct the state space from the marked point process using the delay coordinate system. In addition, we showed that the proposed method could be effective for detecting causality from marked point processes.

1. Introduction

In recent years, it is possible to observe various time series data with high precision because observational technologies have been much improved. Then, we can observe point process data that is the time series data observed at irregular intervals. Typical examples of point process data include spike trains derived from neural activities and marked point processes derived from seismic activities and economic activities. If we can identify causal relations only from observed data, it is useful to understand various phenomena. For example, through a method for detecting causality, it is expected to estimate the network structure only from multi-dimensional observed time series [1]; then, we use obtained knowledge to control the system. Also, various methods to estimate causality from observed time series, such as the Granger causality [2] and transfer entropy [3], have been proposed. However, it has been pointed out that these methods cannot distinguish actual causality from spurious causality and that they require a large number of data [4]. On the other hand, a method called convergent cross mapping [4] has been proposed to solve these problems.

For these reasons, we have already proposed a causality detection method for point processes by modifying convergent cross mapping [4] based on nonlinear dynamical systems

theory [5, 6, 7]. Since the proposed method detects causality based on information in the reconstructed state space, we need to reconstruct the state space from point process. However, the state space reconstruction from event intervals has already been discussed [8, 9]. Then, it has also been discussed that the data obtained by extracting the local maxima of an observed time series might preserve a part of the characteristics in the original dynamical system [10]. The marked point process is considered to be a multi-dimensional time series, because the marked point process has inter-event intervals and marked values. Moreover, it has already been revealed that the state space can be reconstructed from the multi-dimensional time series [11, 12]. However, the state space reconstruction from marked point processes has not been fully explored. Therefore, in this paper, we examine the validity of a method for reconstructing the state space from marked point processes and evaluate the proposed method for detecting causality based on its reconstruction method.

2. Method to generate a marked point process

In this section, we describe how to generate a marked point process. We generated a marked point process by extracting the local maxima of the time series (Fig. 1) [13]. In Fig. 1,

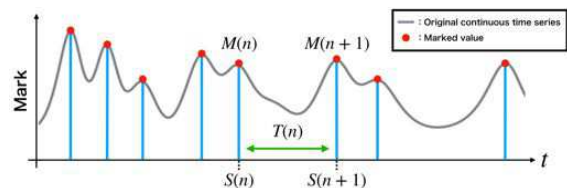


Figure 1: Schematic diagram of the method for generating a marked point process by extracting local maxima of an original continuous time series.

the n th event timing $S(n)$ is defined as when the time series takes the local maximum, and the marked value $M(n)$ is its state value. We define inter-event interval (IEI) as $T(n) = S(n+1) - S(n)$. To generate the time series, we use the Lorenz system [10] and a coupled Lorenz system [14]. We show the Lorenz system in Eq. (1).

$$\begin{cases} \dot{x} = \sigma(-x + y), \\ \dot{y} = -xz + rx - y, \\ \dot{z} = xy - bz, \end{cases} \quad (1)$$

where we use parameter values $r = 28$, $\sigma = 10$ and $b = 8/3$

in this paper. Then, the coupled Lorenz system is shown in Eq. (2).

$$(2) \quad \begin{cases} \dot{x}_1 = \sigma_1(y_1 - x_1) + C_{12}(x_2 - x_1), \\ \dot{y}_1 = r_1 x_1 - y_1 - x_1 z_1, \\ \dot{z}_1 = -b_1 z_1 + x_1 y_1, \\ \dot{x}_2 = \sigma_2(y_2 - x_2) + C_{21}(x_1 - x_2), \\ \dot{y}_2 = r_2 x_2 - y_2 - x_2 z_2, \\ \dot{z}_2 = -b_2 z_2 + x_2 y_2, \end{cases}$$

where the parameter values are set to $\sigma_1 = \sigma_2 = 10$, $b_1 = b_2 = 8/3$ and $r_1 = r_2 = 28$. Here, the causal relationship between each variable in the coupled Lorenz system is shown in Fig. 2. Figure 2 shows that Lorenz 1 and Lorenz 2 are cou-

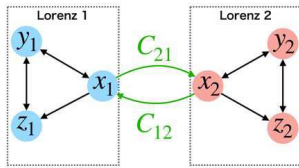


Figure 2: The causal relationship between variables in the coupled Lorenz system.

pled with the coupling strength C_{12} and C_{21} . When Lorenz 1 and Lorenz 2 are bidirectionally coupled, we set parameters as $C_{12} = C_{21} = 1$. When Lorenz 1 and Lorenz 2 are unidirectionally coupled, we set parameters as $C_{12} = 0$ and $C_{21} = 1$. When Lorenz 1 and Lorenz 2 are not coupled, we set parameters as $C_{12} = C_{21} = 0$.

3. Evaluation of reconstructed state space from marked point process

This section describes and evaluates a method to reconstruct state space from a marked point process.

3.1 State space reconstruction from marked point process

First, we describe how to reconstruct a state space from marked point processes. Let $T(n)$ be the n th inter-event interval (IEI) of a marked point process and $M(n)$ be the mark value. The state space can be reconstructed using delay coordinate system [15], as shown in Eq. (3).

$$\mathbf{V}(n) = (M(n), T(n), \dots, M(n + (m + 1)\tau), T(n + (m + 1)\tau)), \quad (3)$$

where $2m$ is the reconstruction dimension and τ is the time delay. Then, we evaluate the validity of the reconstruction of the state space from marked point processes of the Lorenz system. If the state space of the dynamical system can be reconstructed from the marked point process, the reconstructed state space should preserve the topological characteristics such as the Lyapunov spectrum and the fractal dimension of the state space of the original continuous dynamical system that generated the marked point process. In addition, if the topological characteristics of the dynamical system are preserved, the inter-point distance distribution should be similar. Therefore, to evaluate whether the original dynamical system can be reconstructed, we calculate the Jensen-Shannon (JS) divergence [16] between the inter-point distance distributions of the original state space and the reconstructed state space

of the Lorenz system. Here, the overview of the numerical experiment is shown in Fig. 3.

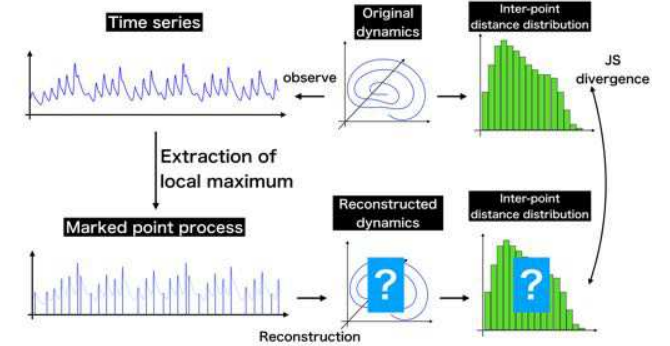


Figure 3: Evaluation procedure for the state space reconstruction from marked point processes.

3.2 Jensen-Shannon (JS) divergence

The JS divergence is defined as the sum of the Kullback-Leibler (KL) divergences. Then, first, we explain the KL divergence. The KL divergence is a measure to quantify dissimilarity of two probability distributions. The KL divergence of a probability distribution P from a probability distribution Q is defined by Eq. (4).

$$D_{KL}(P||Q) = \sum_{i=1}^N P(k_i) \log \left(\frac{P(k_i)}{Q(k_i)} \right). \quad (4)$$

In Eq. (4), $P(k_i)$ and $Q(k_i)$ are the probabilities of the i th class of the probability distribution P and Q in the case of ascending order, and N is the total number of classes. From Eq. (4), the KL divergence is an asymmetric measure, namely $D_{KL}(P||Q) \neq D_{KL}(Q||P)$. Therefore, in this paper, we use the JS divergence because it is symmetric. The JS divergence for probability distributions P and Q is defined by Eq. (5).

$$D_{JS}(P, Q) = \frac{1}{2} (D_{KL}(P||M) + D_{KL}(Q||M)). \quad (5)$$

Here, the probability distribution M is the average of the probability distributions P and Q , and is defined by $M(k_i) = \frac{1}{2} (P(k_i) + Q(k_i))$, where k_i is the i th class of distributions in ascending order.

3.3 Result

We investigated the relationship between the JS divergence and the reconstruction dimension. Figure 4 shows the results of the JS divergence between the inter-point distance distribution of the original dynamical system and the inter-point distance distribution reconstructed from the marked point process by varying the reconstruction dimension. We used 10,000 data points to calculate the inter-point distance distribution. Also, we varied the reconstruction dimension in the range of [2, 20] at intervals of 2.

In Fig. 4, the horizontal and the vertical axes represent the reconstruction dimension and the JS divergence. Furthermore, $D_{JS}(P_o||P_i)$ represents the JS divergence between the original state space of the Lorenz system and the state space

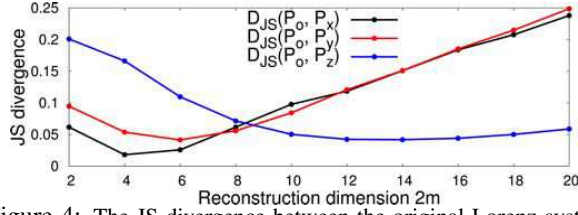


Figure 4: The JS divergence between the original Lorenz system and the system reconstructed from marked point process when m changes.

reconstructed from the marked point process of the variable i . From Fig. 4, although appropriate values of the reconstruction dimension depend on which variable is used, we found that the JS-divergence has a small value of less than 0.05. In other words, the inter-point distance distribution reconstructed from the marked point process is similar to the inter-point distance distribution of the original state space. This result is not a sufficient condition for reconstructing the state space, but it may provide evidence that the state space can be reconstructed from marked point process. In future work, we should investigate whether the reconstructed state space has characteristics satisfied the sufficient condition for embedding.

4. Detecting causality for marked point processes

In this section, we evaluate the effectiveness of the proposed method for detecting causality between marked point processes. The proposed method uses the method for reconstructing the state space from a marked point process (Eq. (3)).

4.1 Proposed method for detecting causality for multidimensional marked point processes

We describe the proposed method for detecting causality for marked point processes. Our method generates time delay vectors from two-dimensional time series of mark and IEL, as shown in Eqs. (6) and (7). Then, we predict the mark and IEL time series from the information in the reconstructed state space independently.

$$\mathbf{V}_x(n) = (M_x(n), T_x(n), \dots, M_x(n + (m + 1)\tau), T_x(n + (m + 1)\tau)), \quad (6)$$

$$\mathbf{V}_y(n) = (M_y(n), T_y(n), \dots, M_y(n + (m + 1)\tau), T_y(n + (m + 1)\tau)), \quad (7)$$

where $M_x(n)$ and $T_x(n)$ are the marked value and IEL of marked point process obtained from the variable x (MPP $_x$), and $M_y(n)$ and $T_y(n)$ are the marked value and IEL of marked point process obtained from the variable y (MPP $_y$). Next, to predict the IEL time series, $T_x(n)$ and $T_y(n)$, and the marked values, $M_x(n)$ and $M_y(n)$, we use the above-mentioned time delay vectors: $\mathbf{V}_x(n)$ and $\mathbf{V}_y(n)$. In CCM, we calculate weights from the distance information of neighboring points in the reconstructed state space. Then, we use the weights to cross-predict of in-sample and detect causality based on the prediction accuracy. Here, we only describe the case of predicting the IEL time series $T_x(n)$ and the mark time series $M_x(n)$ from the information in the reconstructed state

space of y . Now, we define $\hat{T}_x(n)$ as the predicted time series of $T_x(n)$ and $\hat{M}_x(n)$ as the predicted time series of $M_x(n)$. Here, $\hat{T}_x(n)$ is calculated by

$$\hat{T}_x(n) = \sum_{i=1}^K w_y(n_i^*) T_x(n_i^*), \quad (8)$$

and $\hat{M}_x(n)$ is calculated by

$$\hat{M}_x(n) = \sum_{i=1}^K w_y(n_i^*) M_x(n_i^*), \quad (9)$$

where the meaning of $*$ is explained in Eq. (12). In Eqs. (8) and (9), K is the number of neighbors for predicting time series. In the numerical experiments, we used $K = 2m + 1$. Then, $w_y(n_i^*)$ is defined by

$$w_y(n_i^*) = \frac{u_y(n_i^*)}{\sum_{j=1}^K u_y(n_j^*)}, \quad (10)$$

where $w_y(n_i^*)$ is the weight calculated from the i th nearest neighbor of the time delay vector $\mathbf{V}_y(n^*)$ in the reconstructed state space. In Eq. (10), $u_y(n_i^*)$ is defined by

$$u_y(n_i^*) = \exp \{-d[\mathbf{V}_y(n^*), \mathbf{V}_y(n_i^*)]\}, \quad (11)$$

where $d[\mathbf{V}_y(n^*), \mathbf{V}_y(n_i^*)]$ is the Euclidean distance between $\mathbf{V}_y(n^*)$ and $\mathbf{V}_y(n_i^*)$. Then, $\mathbf{V}_y(n_i^*)$ is the i th nearest neighbor of the time delay vector, $\mathbf{V}_y(n^*)$, in the reconstructed state space. Here, the prediction flow is shown in Fig. 5. As shown in Fig. 5, we first consider the case where $M_x(n)$ and $T_x(n)$ are the target of in-sample prediction. We show the prediction steps (i)~(v) in the following.

(i) We define the n th event timing of the MPP $_x$ and MPP $_y$ as $S_x(n)$ and $S_y(n)$. We refer to the marked value $M_y(n^*)$ and IEL $T_y(n^*)$ which corresponds to the nearest event timing $S_y(n^*)$ for $S_x(n)$. Namely, n^* is calculated by the Eq. (12).

$$n^* \in G_n = \arg \min_{n'} |S_y(n') - S_x(n)|. \quad (12)$$

In other words, n^* is the index of event on MPP $_y$ whose event timing is the nearest event against $S_x(n)$ on MPP $_x$.

- (ii) We identify the i th nearest neighbor of $\mathbf{V}_y(n^*)$, $\mathbf{V}_y(n_i^*)$, on the reconstructed state space. In Fig. 5, we show an example of the case where $m = 1$ and $\tau = 1$, that is, the reconstruction dimension is 2.
- (iii) Using Eqs. (10) and (11), we calculate $u_y(n_i^*)$ and $w_y(n_i^*)$. In addition, we refer to $S_y(n_i^*)$ which corresponds to the time of i th nearest neighbor.
- (iv) We refer to the IEL $T_x(n_i^*)$ and the marked value $M_x(n_i^*)$ that corresponds to the event with the nearest timing of $S_y(n_i^*)$, that is, $n_i^* = (n_i^*)^*$. Now, by setting $n_i^* = n_i$, we make the meaning of $*$ the same as the case of step (i).

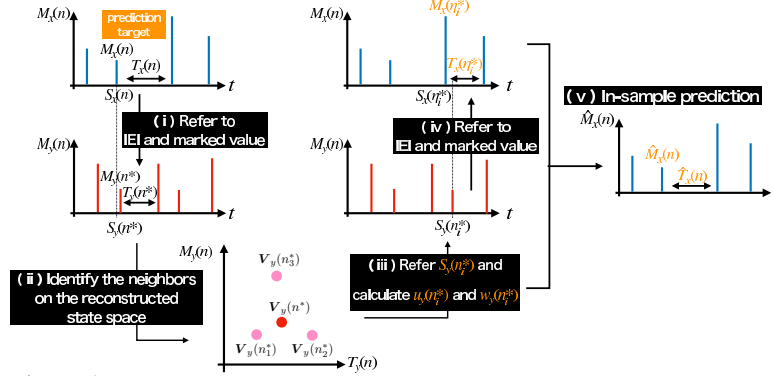


Figure 5: In-sample prediction flow for marked point process in the proposed method.

- (v) We calculate the weighted average of $T_x(\eta_i^*)$ and $M_x(\eta_i^*)$ referenced by (iv) using the weight $w_y(n_i^*)$ calculated by (iii). Namely, we predict $M_x(n)$ and $T_x(n)$ such as Eqs. (8) and (9).

We perform the above procedure for all events to generate prediction time series of the IEIs and the marked values. Here, the prediction accuracy is evaluated by the correlation coefficient in the same manner as Ref. [4]. When marked point processes have causality and the data length is sufficiently long, the prediction accuracy of the time series increases and converges. We can detect the causality by using this characteristic. The difference between the proposed method and the original CCM is two-folds. The first one is that the temporal matching of the index for in-sample prediction is realized based on the nearest time when the event occurred. The second one is that the state space is reconstructed from the marked values and IEIs.

4.2 Result

We applied the proposed method to the marked point processes obtained from the coupled-Lorenz system described in Eq. (2). The results of detecting causality are shown in Fig. 6. The horizontal and the vertical axes represent the data length used for the prediction and the prediction accuracy. In addition, $\rho_{\text{target}}[\text{Lorenz } i \mid \text{Lorenz } j]$ in the legend is the average of the prediction accuracy for predicting Lorenz i ($i = 1, 2$) using Lorenz j ($j = 1, 2$), where the “target” is IEI or mark. In other words, we average nine prediction accuracies of predicting marked point processes generated from each variable in Lorenz i using each variable in Lorenz j .

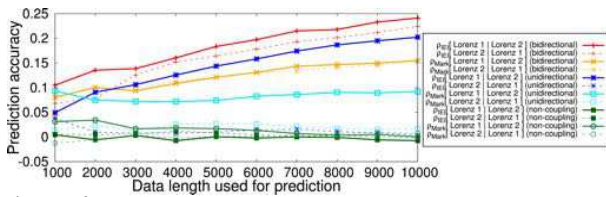


Figure 6: Results of applying the proposed method for detecting causality to marked point processes obtained by Eq. (2).

Figure 6 shows that in the case of bidirectional coupling, all prediction accuracies increase with increasing the data length. In contrast, in the case of unidirectional coupling, the prediction accuracies only increase when we predict Lorenz 1 using Lorenz 2. Furthermore, in the case of non-coupling, all prediction accuracies converge to almost zero. Therefore, we found that the proposed method is effective in detecting causality for marked point processes.

5. Conclusion

We have investigated whether a dynamical system can be reconstructed from a marked point process using a delay-coordinate system. In particular, we calculated the JS divergence between the inter-point distance distribution of the original state space and the reconstructed state space. As a result, it is suggested that the state space can be reconstructed from the marked point process. Then, we applied the proposed method for detecting causality to marked point processes generated from the coupled Lorenz system. As a result, we found that the proposed method effectively detect causality for marked point processes.

Acknowledgment

This research is partially supported by the JSPS Grant-in-Aids for Scientific Research (Nos. JP15KT0112, JP17K00348, JP18K18125, and JP20H00596).

References

- [1] K. Sawada et al., *NOLTA, IEICE*, Vol. 11, No. 4, pp. 422–432, 2020.
- [2] C. W. J. Granger, *Econometrica*, Vol. 37, No. 3, pp. 424–438, 1969.
- [3] T. Schreiber, *Phys. Rev. Lett.*, Vol. 85, No. 2, pp. 461–464, 2000.
- [4] G. Sugihara et al., *Science*, Vol. 338, No. 6106, pp. 496–500, 2012.
- [5] K. Sawada et al., *Proc. of ICMNS2019*, p. 80, 2019.
- [6] K. Sawada et al., *Proc. of NOLTA2019*, pp. 591594, 2019.
- [7] K. Sawada et al., *Proc. of IEICE Soc. Conf.*, N–1–6, 2020 (in Japanese).
- [8] T. Sauer, *Phys. Rev. Lett.*, Vol. 72, No. 24, pp. 3811–3814, 1994.
- [9] T. Sauer, *Chaos*, Vol. 5, No. 1, pp. 127–132, 1995.
- [10] E. N. Lorenz, *J. Atmos. Sci.*, Vol. 20, No. 2, pp. 130–141, 1963.
- [11] T. Sauer et al., *J. Stat. Phys.*, Vol. 65, Nos. 3–4, pp. 579–616, 1991.
- [12] E. Deyle et al., *PLOS ONE*, Vol. 6, No. 3, e18295, 2011.
- [13] N. Yabuta et al., *Proc. of NOLTA2007*, pp. 16–19, 2007.
- [14] I. Belykh et al., *Chaos*, Vol. 16, No. 1, 015102, 2006.
- [15] N. H. Packard et al., *Phys. Rev. Lett.*, Vol. 45, No. 9, pp. 712–716, 1980.
- [16] J. Lin, *IEEE Trans. Inf. Theory*, Vol. 37, No. 1, pp. 145–151, 1991.
- [17] S. Kullback et al., *Ann. Math. Stat.*, Vol. 22, No. 1, pp. 79–86, 1951.

Discrimination of deterministic chaos and noisy periodicity by the recurrence plot with variable threshold

Shiki Kanamaru^{1,(a)}, Nina Sviridova¹, Yutaka Shimada² and Tohru Ikeguchi¹

¹ Tokyo University of Science
6-3-1 Nijjuku, Katsushika, Tokyo 125-8585, Japan
Phone/FAX:+81-03-5876-1717

^(a)Email: shikimaru@hisenkei.net

² Saitama University
255 Shimo-okubo, Sakura, Saitama city,
Saitama 338-8570, Japan

Abstract

A recurrence plot (RP) is an effective method to capture characteristics of time series qualitatively. When applying RP, we need to set the threshold to plot the correlation between two points. However, it is hard to set an optimal threshold according to the time series. Therefore, we have already proposed a method for capturing the characteristic of time series using a deterministic index calculated from multiple RPs created by changing the threshold. This article examined the effectiveness of the proposed method for discriminating deterministic chaos and noisy periodicity. The results show that the proposed method is effective in the discrimination of deterministic chaos and noisy periodicity.

1. Introduction

In this article, we use a method called recurrence plot (RP) [1, 2] to visualize dynamical and structural properties of attractors. RP visualizes the correlation between points on an attractor into a two-dimensional binary image while preserving the time evolution information. The two-dimensional image is a square image with both axes corresponding to the time indices of the original time series. The pixel (i, j) is set to unity if the state values at the time i and j are in the neighborhood of each other, and zero if they are not. Thus, a two-dimensional binary image consists of 0s and 1s, which is called RP. RP is useful in detecting non-stationarity [1]. It is also considered that RPs contain topological information about attractors [3]. Moreover, there are some typical patterns [4] on RPs that can be used to understand the properties of dynamical systems behind time series. For these reasons, RP is currently one of the most promising methods for analysing nonlinear time series.

In addition, many extensions of RP have been proposed [4]. For example, it is essential to use the neighborhood relation of the points on the attractor and information of the trajectory vector in chaotic time series analysis. Therefore, the iso-directional recurrence plot (IDRP) [5] has been proposed as

an extended RP method. IDRP considers the information of trajectory vector. While RP visualizes the neighborhood relation between two points on the attractor, IDRP visualizes the same direction relation of trajectory vectors. Furthermore, based on RP and IDRP, the iso-directional neighbors plot (IDNP) [5] has been proposed to visualize the neighborhood and same direction relationships. It is expected that the number of plots that are near and in the same direction will be small on IDNP created from stochastic time series. Therefore, IDNP is considered to be effective in detecting determinism in time series [5].

However, a threshold needs to be carefully set when applying RP, IDRP, and IDNP. RP needs a threshold to determine whether or not to plot the pixel corresponding to the time index for the distance between two points on the attractor. IDRP needs a threshold for the difference between the transition vectors obtained from any two points on the attractor. Although, there is no direct threshold in the case of IDNP because the number of plots of IDNP depends on the plots on RP and IDRP, the thresholds of RP and IDRP are essential parameters in creating IDNP. If these thresholds are set incorrectly, it may not capture characteristics of the time series. For this reason, it has been studied how to set the threshold on RP [2, 6-9]. On the other hand, the best threshold of RP strongly depends on the target structure [4, 10]. In addition, instead of a two-dimensional binary image, an analysis method using a recurrence matrix with a distance between two points given to each pixel has been proposed [11]. This method eliminates the need to set a threshold and creates a unique recurrence matrix for a given attractor. However, quantifying the recurrence matrix is more complicated than quantifying the RP, so it is common to use the binarized RP [3].

Therefore, we have proposed a method to extract a time series property by creating multiple RPs, IDRPs, and IDNPs with varying thresholds instead of adopting a single threshold [12-15]. This method focused on the transition of patterns appearing on RP, IDRP, and IDNP with the threshold change. To quantify the patterns appearing on these images, we used

the occurrence probability R [5]. Therefore, we could obtain the trend of R for the variable threshold. Our proposed method [12–15] captures this trend as a property of the time series.

However, the proposed method has not been sufficiently investigated for a system, such as a nonlinear stochastic dynamical system. Therefore, in this article, we investigate that the proposed method can discriminate between the deterministic chaotic time series and the noisy periodic time series. Using the proposed method, we can obtain different properties for deterministic chaotic time series and noisy periodic time series, thereby discriminating them.

2. Methods

In this article, we assume that the time series to be analyzed is one-dimensional. We first reconstruct the state space by transforming the observed time series x_t ($t = 1, \dots, n$) into time-delay coordinates [16, 17]. The reconstructed state space vector is described by

$$\mathbf{X}_t = (x_t, x_{t+\tau}, \dots, x_{t+(m-1)\tau}), \quad (1)$$

where $t = 1, \dots, N$ with $N = n - (m-1)\tau$, m is the dimension of the reconstructed state space and τ is the time delay.

Using the reconstructed state space obtained from Eq. (1), we create an $N \times N$ two-dimensional image, or RP [1, 2] by the following procedure. The first step is to calculate the distance $D_{i,j}$ between any two points, \mathbf{X}_i and \mathbf{X}_j , in the reconstructed state space as

$$D_{i,j} = |\mathbf{X}_i - \mathbf{X}_j|, \quad (2)$$

where $|\cdot|$ denotes Euclidean distance. The (i, j) th pixel value $RP_{i,j}$ of the RP can be calculated using Eq. (2) as

$$RP_{i,j} = \begin{cases} 1 & (\text{if } D_{i,j} \leq \theta_r), \\ 0 & (\text{otherwise}). \end{cases} \quad (3)$$

In Eq. (3), θ_r is the threshold that determines the closeness between two points. Then, let M_{RP} be the set of pixels whose values are unity in RP.

Next, we will explain IDRP [5], which visualizes the difference of trajectory vectors between two points. In comparison, RP is a method to visualize the neighborhood relationship between two points. In other words, by changing Eq. (2) to Eq. (4), we can create an $(N-T) \times (N-T)$ two-dimensional image, namely IDRP.

$$DV_{i,j} = |(\mathbf{X}_{i+T} - \mathbf{X}_i) - (\mathbf{X}_{j+T} - \mathbf{X}_j)|, \quad (4)$$

where T is the transition time length of the trajectory vector. The (i, j) th pixel value of the IDRP, $IDRP_{i,j}$, can be calculated using Eq. (4) as

$$IDRP_{i,j} = \begin{cases} 1 & (\text{if } DV_{i,j} \leq \theta_d), \\ 0 & (\text{otherwise}). \end{cases} \quad (5)$$

In Eq. (5), θ_d is the threshold that determines whether two trajectory vectors are in the same direction or not. Then, let M_{IDRP} be the set of pixels whose values are unity in IDRP.

Since IDRP plots the time evolution of two arbitrary points \mathbf{X}_i and \mathbf{X}_j , it does not take into account whether they are close to each other or not. Then, the set of pixels whose values are unity on IDNP is defined as the common set of the set M_{RP} and the set M_{IDRP} :

$$M_{IDNP} = M_{RP} \cap M_{IDRP}. \quad (6)$$

Therefore, the (i, j) th pixel value of IDNP, $IDNP_{i,j}$, is determined to be

$$IDNP_{i,j} = RP_{i,j} \times IDR_{i,j}. \quad (7)$$

In other words, a relation is plotted on the IDNP if the distance between any two points ($\mathbf{X}_i, \mathbf{X}_j$) is within θ_r and the norm of the difference between the trajectory vectors is within θ_d . Moreover, the size of IDNP is $(N-T) \times (N-T)$, as in IDRP.

3. Occurrence Probability R

In this article, we use the occurrence probability R [5] to quantify the IDNP. The IDNP is the common set of the RP and IDRP as shown in Eq. (6). Then R denotes the probability of the occurrence of pairs of neighboring points whose trajectory vectors are in the same direction, which is defined by

$$R = \frac{|M_{IDNP}|}{|M_{RP}|}, \quad (8)$$

where $|M_{RP}|$ and $|M_{IDNP}|$ denote the number of elements in the sets M_{RP} and M_{IDNP} , respectively. If the target system is deterministic, it is considered that neighboring sets tend to have a similar direction and maintain a neighborhood relationship even after the transition. The occurrence probability R reflects this property in the ratio of the number of points in the denominator and numerator of Eq. (8). Therefore, R tends to be high when the time series is deterministic.

4. Variable Threshold

For RP and IDRP, it is necessary to set the thresholds θ_r and θ_d . To select the threshold of RP, the following methods have been proposed: 10% of the average or maximum distance between two points of the attractor [2], a few percent of the maximum distance between two points of the attractor [6], 1% of the total number of pixels to be plotted [7], namely the proportion called a recurrence rate (RR) [4] be 1%, or a system-specific setting method [8, 9]. However, it is difficult to determine the optimal threshold for any given time series uniquely because a threshold selection of RP strongly depends on the target system [4, 10].

Therefore, in this article, instead of selecting an optimal threshold, we created various RPs for a single time series by varying the threshold. The method of changing the threshold is explained using RP as an example. In RP, the distances between all two points in the attractor are calculated, and if $D_{i,j}$ is less than or equal to the threshold θ_r , (i, j) th pixel is colored black. In this article, we first find the shortest distance $D_{\min} = \min_{i < j} D_{i,j}$ and the longest distance $D_{\max} = \max_{i < j} D_{i,j}$. Then, the width of the range $[D_{\min}, D_{\max}]$ is $D_r = D_{\max} - D_{\min}$, and θ_r is given by

$$\theta_r = D_r \theta + D_{\min}. \quad (9)$$

In Eq. (9), θ is a parameter that manipulates the threshold θ_r with a range of $0 \leq \theta \leq 1$. For example, when $\theta = 0$, the points are plotted on the RP only when satisfying $D_{i,j} \leq D_{\min}$. On the other hand, when $\theta = 1$, all pixels are set to unity. In IDR, $DV_r = DV_{\max} - DV_{\min}$ is calculated for $DV_{i,j}$, and θ manipulates θ_d . By varying this parameter θ , we create multiple RPs and IDRs with different threshold values. When creating an IDNP, we used a common value of θ for both RPs and IDRs. In this article, $\theta = 2^{-k}$ ($k = 1, 2, \dots, 10$).

5. Model

In this article, the stochastic Rössler model (SRM) is used to analyze deterministic chaotic time series and noisy periodic time series. We define an SRM as

$$\begin{cases} \dot{x} = -y - z + \sigma \dot{\omega}_1, \\ \dot{y} = x + ay + \sigma \dot{\omega}_2, \\ \dot{z} = b + z(x - c) + \sigma \dot{\omega}_3, \end{cases} \quad (10)$$

where a , b , and c are parameters, $\omega_i(t)$ ($i = 1, 2, 3$) are independent standard Wiener processes with Gaussian increments, and σ is a parameter for the noise intensity. When $\sigma = 0$, the SRM is equivalent to the Rössler equations [18]. In this article, $a = b = 0.2$, $c = 5.7$ for the chaotic response, and $c = 3.5$ for the two-period response. In addition, 10 types of noise intensity σ ($\sigma = 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09$ and 0.1) were used to generate noisy periodic time series. These numerical calculations were performed using the fourth-order Runge-Kutta method with a step width of $h = 10^{-4}$. The time series to be analyzed has a length of 4,000 points that are sub-sampled every 500 steps from the time series of the response of 2×10^6 steps after omitting 10^6 steps as a transient state.

6. Numerical Experiment

The conditions for the numerical experiments are described below. First, we generated 100 time series with fixed parameters and different initial values. Therefore, the occurrence

probability R is averaged over these 100 time series. The transition time length T used in IDR is set to the same value of the time delay τ ($T = \tau$). The chaotic time series and the noisy periodic time series are the time series of the first variable x generated from the SRM with the parameter c and the noise intensity σ , respectively. The dimensionality m of the reconstructed state space was set to $m = 7$. The time delay τ was set to the time when the autocorrelation function of the time series to be analyzed first became $1 - 1/e$ [19].

Figure 1 shows the trend of the occurrence probability R for the chaotic time series ($c = 5.7$ and $\sigma = 0$ in SRM) and that for the noisy periodic time series ($c = 3.5$ and $\sigma > 0$ in SRM). In Fig. 1, the horizontal axis is the parameter θ in log scale. The vertical axis is the occurrence probability R . The triangles connected by the solid black line show the results of the chaotic time series ($c = 5.7$ and $\sigma = 0$). The circles connected by dashed lines show the noisy periodic time series ($c = 3.5$ and $\sigma > 0$). The difference in color indicates the difference in the noise intensity σ .

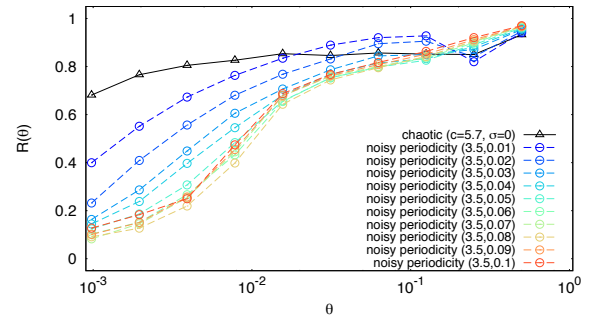


Figure 1: We generated 100 time series for various values of c and σ by changing the initial values. Solid lines are deterministic time series, and dashed lines are noisy periodic time series with the noise intensity σ . The triangles denote chaotic parameters, and the circles denote periodic parameters. Plots represent the mean of 100 runs.

Our purpose is to discriminate between deterministic chaos and noisy periodicity. From Fig. 1, we can see that the chaotic time series has high R even when θ is small, and the variation of R is small even when θ increases. On the other hand, for noisy periodic time series, R is low when θ is small, and R becomes large when θ increases. From these differences in the trend of R , it is suggested that the proposed method can discriminate between deterministic chaos and noisy periodicity.

Another finding from this result is that the proposed method of setting thresholds of RP so far has a risk of false discriminating between deterministic chaos and noisy periodic time series. First, the threshold of a few percent to 10% of the attractor size [2, 6] corresponds to the range of

$0.01 < \theta \leq 0.1$ in Fig. 1. The range of threshold values that detects high determinism ($R > 0.6$) despite the strong noise intensity σ . Moreover, θ is close to 0.1, the deterministic chaos and noisy periodic time series have a similar R , and it is difficult to discriminate between them. Next, the relationship between θ and the recurrence rate (RR) on the RPs is shown in Fig. 2. The differences from Fig. 1 are the vertical axis is the log scale, and the indicator is RR . Then, RR is calculated

as follows: $RR = \frac{2}{N^2 - N} \sum_{i < j}^N RP_{i,j}$. Figure 2 shows that θ ,

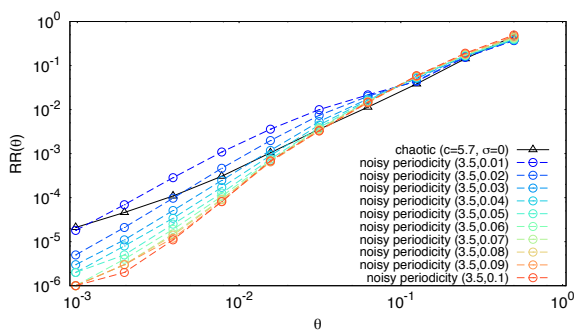


Figure 2: The relationship between θ and RR on the RPs.

where the RR is around 1%, is approximately in the range of 2^{-5} ($= 0.03125$) $\leq \theta \leq 2^{-4}$ ($= 0.0625$). In Fig. 1, looking at the occurrence probability R in the range of θ where the RR is around 1%, it can be seen that the noisy periodic time series has a high R at any noise intensity σ , and it is difficult to distinguish from the deterministic chaotic time series. These emphasize the necessity of varying the threshold rather than using a single threshold in the analysis using the conventional threshold selecting method.

7. Conclusion

We investigated the discrimination between chaotic and noisy periodic time series by feature extraction from the RP, the IDRP, and the IDNP when the threshold is varied. From the result, we found that as we changed the threshold, the occurrence probability R does not change significantly for chaotic time series, whereas it increases for noisy periodic time series. These results suggest that we can discriminate deterministic chaos and noisy periodicity using the proposed method. In addition, if we used thresholds that have been decided empirically, 1 ~ 10% of the attractor's size or the recurrence rate [2, 6, 7], it is clarified that even noisy periodicities with high noise intensity are misrecognized to have high determinism (i.e., false positives). Namely, it is not only effective but also essential to change the thresholds.

Acknowledgment

This research is partially supported by the JSPS Grant-in-Aid for Scientific Research (Nos. JP17K00348, JP18K18125, JP19K14589, and JP20H00596).

References

- [1] J.-P. Eckmann *et al.*, Europhys. Lett., **4**, pp. 937–977, 1987.
- [2] J. P. Zbilut *et al.*, Phys. Lett. A, **171**, Nos. 3–4, pp. 199–203, 1992.
- [3] M. Thiel *et al.*, Phys. Lett. A, **330**, No. 5, pp. 343–349, 2004.
- [4] N. Marwan *et al.*, Phys. Rep., **438**, Nos. 5–6, pp. 237–329, 2007.
- [5] S. Horai *et al.*, IEEJ Trans. Electron. Inf. Syst., **122**, No. 1, pp.141–147, 2002, in Japanese.
- [6] G. B. Mindlin *et al.*, Physica D, **58**, Nos. 1–4, pp. 229–242, 1992.
- [7] J. P. Zbilut *et al.*, Phys. Lett. A, **297**, Nos. 3–4, pp. 173–181, 2002.
- [8] M. Thiel *et al.*, Physica D, **171**, No. 3, pp. 138–152, 2002.
- [9] L. Matassini *et al.*, Phys. Rev. E, **65**, No. 2, p. 021102, 2002.
- [10] N. Marwan, Int. J. Bifurc. Chaos Appl. Sci. Eng., **21**, No. 4, pp. 1003–1017, 2011.
- [11] G. McGuire *et al.*, Phys. Lett. A, **237**, Nos. 1–2, pp. 43–47, 1997.
- [12] A. Ogawa *et al.*, Proc. IEICE Gen. Conf. '19, N-1-33, 2019.
- [13] S. Kanamaru *et al.*, IEICE Technical Report, **119**, No. 19, NLP2019-5, pp. 23–28, 2019.
- [14] S. Kanamaru *et al.*, Proc. NLS Conf. IEICE '19, N-1-13, 2019.
- [15] S. Kanamaru *et al.*, Proc. of NOLTA2019, pp. 536–539, 2019.
- [16] F. Takens, In Dynamical systems and turbulence, Lecture Notes in Mathematics, **898**, pp. 366–381, 1981.
- [17] T. Sauer *et al.*, J. Stat. Phys., **65**, Nos. 3–4, pp. 579–616, 1991.
- [18] Otto E. Rössler, Phys. Lett. A, **57**, No. 5, pp. 397–398, 1976.
- [19] M. T. Rosenstein *et al.*, Physica D, **65**, Nos. 1–2, pp. 117–134, 1993.

Effects of Document Characteristics on Language Network Structure

Kihei Magishi^{†,1}, Tomoko Matsumoto^{††}, Yutaka Shimada^{†††} and Tohru Ikeguchi[†]

[†] Tokyo University of Science
 6-3-1 Nijuku, Katsushika,
 Tokyo 125-8585, Japan
 Phone/FAX: +81-03-5876-1717
¹E-mail: magikihe@hisenkei.net

^{††} Tokyo University of Science
 1-3 Kagurazaka, Shinjuku,
 Tokyo 162-0081, Japan

^{†††} Saitama University
 255 Shimo-okubo, Sakura, Saitama city,
 Saitama 338-8570, Japan

Abstract

Complex network theory has revealed several characteristics of languages. However, despite recent attempts of comparison in different languages, it has not been fully examined which characteristics of languages make the differences in network structure between languages. In this paper, we investigated whether the frequency of words and part of speech affect language network, by using complex network theory. For the analysis, we used text data of the Gospel of Matthew translated into nine languages. Our main results are two folds: first, there is correlation between several statistical measures of language network structure and the frequency of words. Second, features of part of speech affect several statistical measures of language network structure. These results suggest that the frequency of words and part of speech are key factors that shape structural differences of language networks.

1. Introduction

Many different languages are used in the world, and these languages have been analyzed in various approaches. In recent years, a lot of research on quantitative analysis of languages from the viewpoint of complex network theory has been reported. As a result, it has been suggested that different languages have common features and specific features. For example, the local structures of a language network have common systems—connection of words, positions of words, and basic word orders—in different languages [1]. Also, even if languages belong to the same language family, these languages could have different characteristics [2]. Although these previous studies are interesting, these previous studies do not use text data of the same content of documents across different languages, nor use various languages of the text data.

We have already demonstrated that language networks transformed from documents have different characteristics of network structure, by using the New Testament [3–5]. Using the New Testament as text documents for analyses, we can analyze the same documents by different languages. However, we have not investigated which characteristics of documents affect the structural characteristics of language networks when the documents transformed to networks.

Therefore, in this paper, we first investigated the relationship between statistical measures of language network structure and the frequency of words, because the frequency of words could be one of the essential characteristics of text documents. As a result, we found a universal feature that the frequency of words affects the degree centrality [6], the betweenness centrality [6], and the closeness centrality [7]. However, it does not affect the clustering coefficient [8] and the square clustering coefficient [9].

Next, we investigated which characteristics of documents affect the clustering coefficient and the square clustering coefficient. Then, we found that part of speech (PoS) affects the clustering coefficient and the square clustering coefficient. These results suggest that the frequency of words and PoS are key factors that shape the structural differences between language networks.

2. Text documents

We used the Gospel of Matthew in the New Testament [10–15] as text documents in this paper. The data are advantageous for the following reasons: first, it has been translated into many different languages. Second, it unifies the content of documents between different languages.

Our data set consists of nine different languages and these languages are classified into four language families: Greek, Latin, English, German, French, Russian (Indo-European language family), Japanese (Japanese language family), Finnish (Uralic language family), and Chinese (Sino-Tibetan language family). Table 1 summarizes the basic information of the dataset used in this paper.

Table 1: Summary information of the data set (Versions of the Gospels of Matthew in the New Testament).

Language	Family	Version	Year	Words	Vertices	Edges
Greek1	Indo-European	NA28	2012	16,293	7,442	13,511
Greek2		UBS5	2014	15,546	7,296	12,946
Latin		Bublia Sacra Vulgata	1955	16,435	2,849	9,499
English1		King James Version	1611	23,665	2,394	8,889
English2		English Standard Version	2001	22,663	2,471	8,901
German		Lutherbible	1984	21,789	2,352	8,438
French1		Louis Segond	1910	23,263	2,454	8,494
French2		Segond 21	2007	23,432	2,502	9,123
Russian		Russian Synodal Version	1876	16,954	4,378	9,020
Finnish1		Uralic	Vanha Kirkkoraamattu	1642	17,805	4,253
Finnish2	Uusi Kirkkoraamattu		1992	16,749	4,690	9,809
Japanese1	Japanese	Colloquial translation	1954	31,055	2,204	7,811
Japanese2		New common translation	1987	29,311	2,177	7,653
Chinese	Sino-Tibetan	Chinese Standard Bible Simplified	2009	22,846	2,628	9,198

3. Method for transforming documents into co-occurrence language network

First, we applied a morphological analysis to divide a sentence into prototypes of words and obtain their PoSs. The morphological analysis tool that we used is Mecab [16] for Japanese, and Tree-tagger [17] for other languages. Then, we treated the pair of the prototype and the PoS of a word as a vertex. Next, we created an unweighted and undirected network by connecting the vertices with edges if they were adjacent in sentences. Here, the duplication of words or connections between words was not allowed, and punctuation marks, half-punctuation marks, and symbols were excluded. Figure 1 shows an example of transforming documents into language networks.

If a document is split into two or more networks by using this method, we only used a large connected subnetwork as an analysis target in this paper. The reason is that these obtained networks consist of one large connected subnetwork and several small subnetworks or isolated vertices in many cases, and the small subnetworks and isolated vertices are not essential for capturing the characteristics of the co-occurrence relations of words in sentences which are the subject of our analysis.

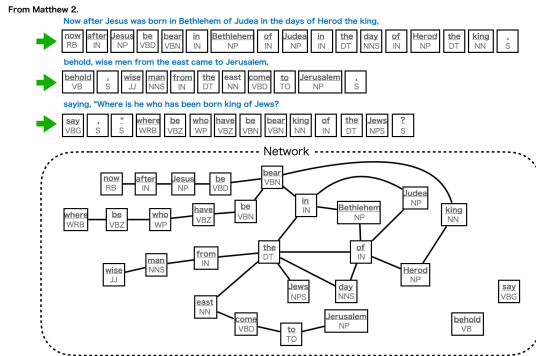


Figure 1: An example of transforming documents into language networks. Squares and rectangles are vertices. The upper part of the vertex expresses a prototype of the word, and the lower part of vertex expresses the PoS. For example, DT stands for determiner and VB stands for verb.

4. Statistical measures of network structures

We examined five statistical measures of network structures: the degree centrality, the betweenness centrality, the closeness centrality, the clustering coefficient and the square clustering coefficient. Let the number of vertices be N , the degree of the i th vertex be k_i , and the (i, j) th element of the adjacency matrix of the network be a_{ij} . If an edge between the i th vertex and the j th vertex exists, $a_{ij} = a_{ji} = 1$, otherwise, $a_{ij} = a_{ji} = 0$.

4.1 Degree centrality

The degree centrality (DC) quantifies the importance of vertices by focusing on their degrees. DC of the i th vertex,

$DC(i)$, is defined as

$$DC(i) = \sum_{j=1}^N a_{ij}. \quad (1)$$

4.2 Betweenness centrality

The betweenness centrality (BC) quantifies the importance of vertices by focusing on their frequency of appearance on the shortest paths in the network. BC of the i th vertex, $BC(i)$, is defined as

$$BC(i) = \sum_{j=1, j \neq i}^N \sum_{l=j+1, l \neq i}^N \frac{P_{jl}(i)}{P_{jl}}, \quad (2)$$

where j and l are start and end vertices of the shortest path, P_{jl} is the total number of shortest paths between j and l , and $P_{jl}(i)$ is the total number of shortest paths between j and l through the vertex i .

4.3 Closeness centrality

The closeness centrality (CC) quantifies the importance of vertices by focusing on whether they are close to other vertices in the network. CC of the i th vertex, $CC(i)$, is defined as

$$CC(i) = \frac{1}{N} \sum_{j=1}^N d_{ij}, \quad (3)$$

where d_{ij} is the shortest distance between the i th vertex and the j th vertex.

4.4 Clustering coefficient

The clustering coefficient ($C^{(3)}$) is the probability that two vertices that are neighbors of a given vertex are neighbors of each other. The clustering coefficient of the i th vertex, $C^{(3)}(i)$, is defined as

$$C^{(3)}(i) = \frac{2}{k_i(k_i - 1)} \sum_{j=1}^N a_{ij} \left(\sum_{l=j+1}^N a_{il}a_{jl} \right). \quad (4)$$

4.5 Square clustering coefficient

The square clustering coefficient ($C^{(4)}$) is the quotient between the number of squares and the total number of possible squares. For a given i th vertex and $\mathbf{V} = \{v_1, v_2, \dots, v_{k_i}\}$ which is the set of vertices of neighbors of i th vertex, the square clustering coefficient of the i th vertex, $C^{(4)}(i)$, is defined as

$$C^{(4)}(i) = \frac{\sum_{m=1}^{k_i} \sum_{n=m+1}^{k_i} q_{iv_m v_n}}{\sum_{m=1}^{k_i} \sum_{n=m+1}^{k_i} \{(k_{v_m} - \eta_{iv_m v_n})(k_{v_n} - \eta_{iv_m v_n}) + q_{iv_m v_n}\}}, \quad (5)$$

where q_{ivmv_n} is the number of squares actually constructed using vertices (i, v_m, v_n) , η_{ivmv_n} is defined as $\eta_{ivmv_n} = 1 + q_{ivmv_n} + \theta_{v_mv_n}$. If neighbors v_m and v_n are connected with each other, $\theta_{v_mv_n} = 1$, otherwise, $\theta_{v_mv_n} = 0$.

5. Experiments on the frequency of words and the PoS

First, we investigated the relationships between each of the five statistical measures of language networks described in Sec. 4 and the frequency of words. More specifically, we calculated the correlation coefficients and the correlation diagrams between the frequency of words and the five statistical measures to investigate their relationships.

Next, we investigated which characteristics of documents affect $C^{(3)}(i)$ and $C^{(4)}(i)$ by focusing on the PoS. First, we calculated the frequency of the PoS in all documents and networks. Then, we calculated and compared $C^{(3)}$ and $C^{(4)}$ for each PoS to investigate the differences of the PoSs which affect language network structure.

6. Results

Figure 2 shows the correlation coefficients between the frequency of words and the five statistical measures.

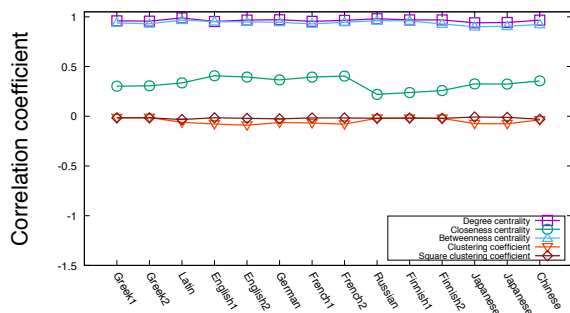


Figure 2: Correlation coefficients between the frequency of words and the five statistical measures.

Figure 2 shows that the linear correlation coefficients between the frequency of words and the two statistical measures—DC and BC—are very high. These results indicate that the frequency of words may affect the structural characteristics, namely DC and BC. On the other hand, the linear correlation coefficient between the frequency of words and CC is low.

To clarify these relations, we investigated correlation diagrams between the frequency of words and CC. Figure 3 is a result for English as an example. Figures 3(a) shows the correlation diagram between the frequency of words and CC, and Fig. 3(b) show the same results but in the semi-logarithmic scale. In Fig. 3(a), we found that the frequency of words and CC have a nonlinear correlation. Furthermore, in Fig. 3(b), a linear relationship appears in the semi-logarithmic scale, where their correlation coefficient in the semi-logarithmic scale is 0.72. These results suggest that there is a nonlinear correlation between the frequency of words and CC for

English. This nonlinear correlation can be found in other languages.

On the other hand, Fig. 2 shows that no correlation exists between the frequency of words and each of two statistical measures— $C^{(3)}$ and $C^{(4)}$. These results suggest that the frequency of words affects language network structure, namely DC, BC, and CC but, does not affect language network structure, namely $C^{(3)}$ and $C^{(4)}$.

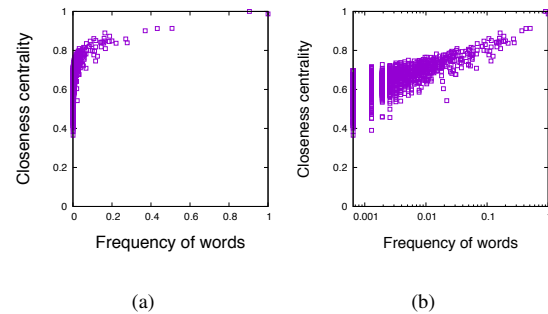


Figure 3: Correlation diagrams of the frequency of words and closeness centrality (English (KJV)) in (a) a linear scale and (b) the semi-logarithmic scale.

Next, we show the analysis results, focusing on the PoS. In Figs. 4–7, the horizontal axis is the type of PoSs and the vertical axis is the type of languages. Colors represent influence of the PoS. Here, the influence of the PoS is defined as the average value of statistics (Fig. 4 is the frequency of words, Fig. 5 is the frequency of vertices, Fig. 6 is the frequency of $C^{(3)}$ and Fig. 7 is the frequency of $C^{(4)}$) for each PoS. Also, the values of the influence for each PoS are normalized by the maximum value by languages. If a PoS is highly influential on its value, the color is close to red in Figs. 4–7. Besides, if no PoS exist in a language, they are shown in white.

Figures 4 and 5 show the results of the influence of the frequency of PoSs in all documents and networks. Figure 4 suggests that in the documents of many languages used in this paper, although nouns and verbs account for a high frequency of documents, other PoSs—such as adjective, pronoun, preposition, particle, also account for a high frequency of documents. However, Fig. 5 shows that in the network of many languages used in this paper, only nouns and verbs account for a very high frequency of language networks. These results suggest that nouns and verbs largely affect language network structure, among various PoSs that are frequently used in the documents.

Figures 6 and 7 show $C^{(3)}$ and $C^{(4)}$ for each PoS in all languages. Figure 6 suggests that PoSs that affect $C^{(3)}$ depend on languages. However, values of $C^{(3)}$ for nouns and verbs are high for almost all languages. Therefore, the nouns and verbs may strongly affect the characteristics of language network structure because frequencies of nouns and verbs are high from Fig. 5.

Figure 7 shows that adjectives and adverbs have higher val-

ues of $C^{(4)}$ in Latin, English, German, French, and Russian which belong to the Indo-European language family. However, the adjectives and adverbs does not strongly affect the characteristics of language network structure because frequencies of adjectives and adverbs are low from Fig. 5. These results suggest that $C^{(3)}$ and $C^{(4)}$ may be influenced by the PoS.

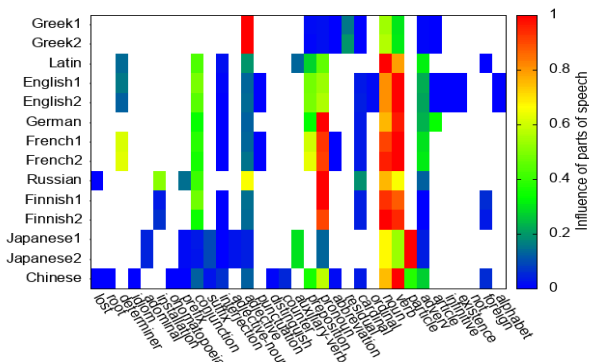


Figure 4: Influence of the frequency of words for each PoS in the documents.

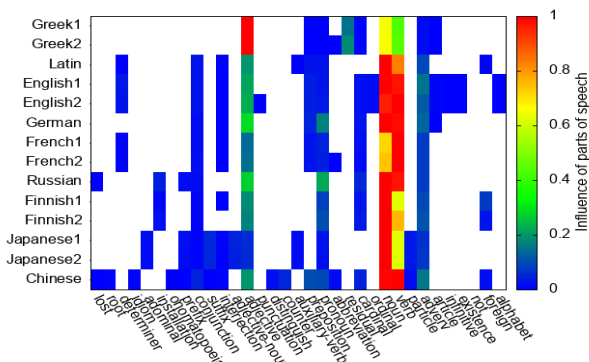


Figure 5: Influence of the frequency of vertices for each PoS in the language network.

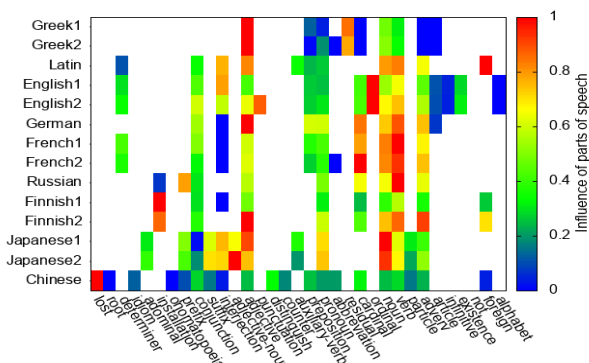


Figure 6: Influence of the frequency of the clustering coefficient for each PoS.

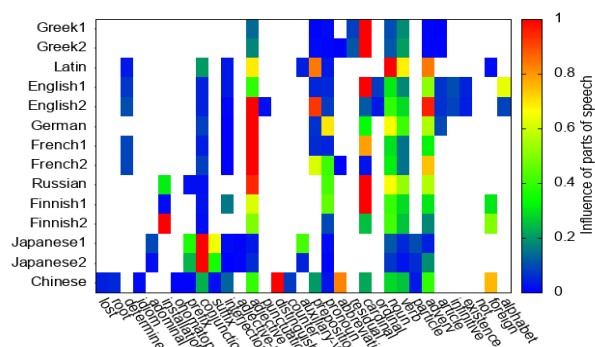


Figure 7: Influence of the frequency of the square clustering coefficient for each PoS.

7. Conclusions

In this paper, we investigated the relationships between the frequency of words and the five statistical measures of language networks—the degree centrality, the betweenness centrality, the closeness centrality, the clustering coefficient, and the square clustering coefficient. As a result, it is shown that the frequency of words affects the degree centrality, the betweenness centrality, and the closeness centrality, but does not affect the clustering coefficient and the square clustering coefficient. We also investigated the clustering coefficient and the square clustering coefficient, focusing on the PoS. As a result, the clustering coefficient and the square clustering coefficient may be influenced by the PoS.

The results of this paper suggest that in complex network theory, the clustering coefficient and the square clustering coefficient may capture the characteristics of PoSs. In the future, we should also investigate whether the method for creating language networks affects the results obtained in this paper.

Acknowledgment

This research is partially supported by the JSPS Grant-in-Aids for Scientific Research (Nos. JP17K00348, JP18K12701, JP18K18125 and JP20H00596).

References

- [1] R. F. i Cancho *et al.*, *Biological Sciences*, **268**, 1482, 2261–2265, 2001.
- [2] Y. Shimada *et al.*, *EPL*, **127**, 5, 56003, 2019.
- [3] K. Magishi *et al.*, Tech. Rep. of *IEICE*, **119**, 471, 77–82, 2020 (in Japanese).
- [4] K. Magishi *et al.*, Tech. Rep. of *IEICE*, **120**, 55, 27–32, 2020 (in Japanese).
- [5] K. Magishi *et al.*, Society Conf. *IEICE*, N-1-14, 2020 (in Japanese).
- [6] L. C. Freeman, *Social Networks*, **1**, 3, 215–239, 1978.
- [7] A. Bavelas, *J. Acoust. Soc. AM.*, **22**, 6, 725–730, 1950.
- [8] D. J. Watts *et al.*, *Nature*, **393**, 6684, 440–442, 1998.
- [9] P. G. Lind *et al.*, *PRE*, **72**, 5, 056127, 2005.
- [10] <https://www.bible.or.jp/> (2020-03-24)
- [11] <https://www.academic-bible.com/> (2020-04-13)
- [12] <http://www.cerbafaso.org/> (2020-04-16)
- [13] <http://www.nic.funet.fi/> (2020-04-20)
- [14] <https://www.biblegateway.com/> (2020-04-26)
- [15] <https://www.bebliä.com/> (2020-04-28)
- [16] <http://taku910.github.io/mecab/> (2020-01-04)
- [17] <https://www.cis.uni-muenchen.de/> (2020-04-27)

Influence of Contact Characteristics to Information Diffusion on Temporal Network

Luyan Xu^{1,†}, Kazuya Sawada¹, Yutaka Shimada² and Tohru Ikeguchi¹

¹Tokyo University of Science
 6-3-1 Nijjuku, Katsushika, Tokyo 125-8585, Japan
 Phone/FAX:+81-03-5876-1717
[†]E-mail: jyoroen@hisenkei.net

² Saitama University
 255 Shimo-okubo, Sakura, Saitama city, Saitama 338-8570, Japan

Abstract

Many researches on information diffusion have been reported. However, these researches mainly use static networks for analysis. In the real world, there are not only static networks, but also temporal networks (TNs) whose structure changes dynamically. We have already reported that the contact density is related to the information diffusion on the TNs. However, contact characteristics other than the contact density could also affect the information diffusion. Therefore, this paper investigated the relationship between the diffusion rate and the contact order on information diffusion on TNs while changing the contact density. Furthermore, we investigated the influence of the vertices with high centralities on the diffusion rate of information by exchanging contacts. As a result, we found that contact order strongly affects the diffusion rate of information. Moreover, we also found that information spreads more easily when the percentage of the contacts of vertices with high centralities is high.

1. Introduction

The relationships of people can be described as networks, where people are represented as vertices, and the social ties between two people are represented as edges. These network structures are static because the social ties between two people rarely change in the short term. Here, the social ties can be friendships on the internet or in the real world. In the real world, not only the social ties but also the contacts between two people can be represented as the edges of the networks. When we create the networks from the contacts of people, these network structures change temporally because the contacts between two people can change temporally. Therefore, the networks whose structures change temporally are referred to as temporal networks (TNs).

Many researches on information diffusion with the former static networks have been reported [1–5]. On the other hand, the information diffusion on the latter TNs are not well discussed. Therefore, we have already investigated final diffusion rates of information on the TNs of contacts between people using the information diffusion model that we pro-

posed [6] previously. As a result, we found that the contact density is related to the information diffusion. However, the diffusion rate of information diffusion are different, even when the contact density are almost the same [6]. These results indicate that other characteristics of TNs than the contact density can also affect the information diffusion. Then, we hypothesize that if contacts occurred in a different order, the diffusion rate of information can be different.

In this paper, to test the hypothesis, we investigate the relationship between the diffusion rate and the contact order on information diffusion on TNs while changing the contact density, by using the proposed model of information diffusion. Furthermore, we investigate the influence of the vertices with high centralities on the diffusion rate of information by exchanging contacts.

2. Contact data sets and temporal networks

In this section, we describe the data sets used in this paper. We create two TNs from two face-to-face contact data sets obtained at a high school and a hospital in France [7–9]. Let us describe the (i, j) th component of the adjacency matrix on the TNs at time $t[s]$ as $A_{ij}(t)$. If there is a contact between two vertices (people) i and j at time t , $A_{ij}(t) = A_{ji}(t) = 1$, otherwise $A_{ij}(t) = A_{ji}(t) = 0$. The data sets at the high school and the hospital consist of face-to-face contact records for 126 people during 4 days and for 75 people during 5 days, respectively. They are recorded at intervals of 20 seconds. In these data sets, the average numbers of contacts per 20 seconds (contact density ρ) at the high school and the hospital are 2.10 and 1.87.

3. Proposed model of information diffusion

In this section, we describe the model of information diffusion [6] we used in this paper. We define the interest level of vertex i for information as $x_i(t)$. Then, the dynamics of the

level of interest of vertex i is described as follows:

$$x_i(t+20) = g_i(\tau, t)x_i(t) + \alpha H(x_i(t) - \theta_i) \sum_{j \in G_i(t)} F(x_j(t) - \theta_j), \quad (1)$$

$$g_i(\tau, t) = \begin{cases} \tau & (t - t_i^* \equiv 0 \pmod{3600}), \\ 1 & (\text{otherwise}), \end{cases} \quad (2)$$

$$H(x) = \begin{cases} 0 & (x \geq 0), \\ 1 & (x < 0). \end{cases} \quad (3)$$

Here, $F(x)$ is the step function, t_i^* is the time when vertex i first receives the information such as e-mail, rumor, etc. τ ($0 < \tau < 1$) is a decay coefficient of the interest level for the information. Furthermore, $G_i(t)$ is the set of adjacent vertices of vertex i at time t , θ_i is the threshold of vertex i , and α is the influence of information. By $g_i(\tau, t)$, the interest level of vertex i for information decays every 3,600 seconds. By $H(x)$, vertex i receives information from adjacent vertices only if it has not reached its threshold θ_i .

4. Numerical experiments

In this section, we investigated the relationship between the diffusion rate and the contact order on information diffusion on TNs while changing the contact density. Furthermore, we investigated the influence of the vertices with high centralities on the diffusion rate of information by exchanging contacts.

4.1 The relationship of the contact order and information diffusion

4.1.1 Experimental conditions

Before the experiment, we divided the time series of contacts into windows with a width of $d[s]$, then we moved the contacts so that each window contains two contacts while keeping the temporal order. The contact density per 20 seconds is described as Eq. (4). From Eq. (4), when d is small, the contact density ρ is large. The relationship of d and ρ is shown in Fig. 1.

$$\rho = 40/d \quad (4)$$

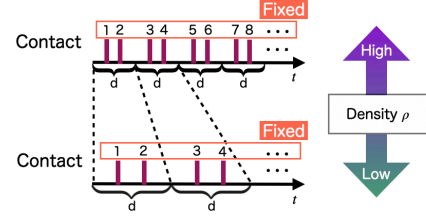


Figure 1: The relationship between d and ρ . The purple bars represent the contacts and the numerals above the bars represent the order of the contacts.

Then, we generated TNs from contact data sets with different ρ by changing d in the range of $[20, 60]$. To investigate the relationship between the diffusion rate and the contact order, we then generated 100 random shuffled (RS) TNs where the contact order is randomly shuffled. The results of the RSTNs are averaged over 100 trials.

In the experiments, we set α to 0.04, τ to 0.97, and threshold θ_i to 0.5 for all vertices. Furthermore, we spread the information by setting the interest level of the first vertex contacted to unity. The first vertex contacted refers to vertex i when the first contact occurred is (i, j) . Besides, we set the interest level of other vertices to 0. In this paper, we define the number of vertices whose interest level for information once reached their thresholds by time t as $n(t)$. The diffusion rate at time t is defined by Eq. (5).

$$s(t) = \frac{n(t)}{N}, \quad (5)$$

where N is the total number of people. Then, the final diffusion rate of information is defined by Eq. (6).

$$S = s(T), \quad (6)$$

where T is the time when the last contact is observed.

4.1.2 Results

We investigated the relationship between d and S on the TNs whose contacts are divided into windows. The results are shown in Fig. 2. Figure 2 shows that the final diffusion rates S decreases when the width of windows d increases in both cases of the high school and the hospital. It suggests that information spreads more easily, when the contact density ρ is high. In addition, Fig. 2 also shows that the grey area accounts for about more than 60% on the high school when $20 \leq d \leq 50$ and more than 90% on the hospital of the range $[0, 1]$ of S . These results suggest that the contact order strongly affects the information diffusion on TNs. Moreover, it is also necessary to discuss how to change the contact order to raise or lower S .

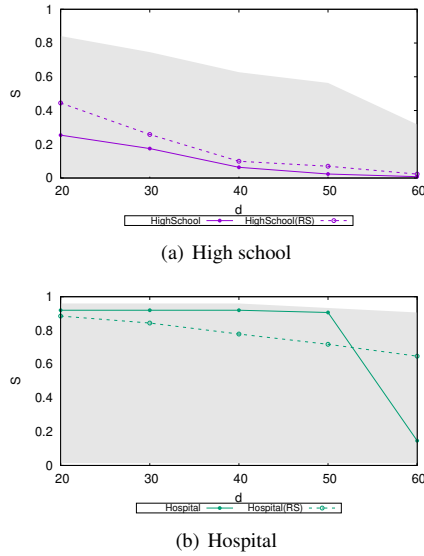


Figure 2: The relationship between d and S on the TNs ($\alpha = 0.04$, $\tau = 0.97$). The dashed lines and the grey areas show the average values of S and the ranges between the maximum and the minimum values of S obtained from 100 RSTNs.

4.2 The relationship of the order of the contacts with high centralities and information diffusion

Next, we investigated how to change the contact order to raise the final diffusion rate of information S . Generally, the vertices (people) with high centrality play an important role in information diffusion. If the contacts of the influential vertices occurred more frequently, the final diffusion rate S increases. Therefore, we move the contacts of the vertices with high centralities from the second half to the first half of the contact data to increase the percentage of the contacts of the vertices with high centralities in the first half of the data. We then investigated the influence of these vertices on the final diffusion rate of information diffusion S .

4.2.1 Experimental conditions

In this paper, we define vertices with high eigenvector centralities [10] as influential vertices in information diffusion. The eigenvector centrality is an indicator of the property that a vertex plays an important role in the networks, if adjacent vertices of the vertex has high degree. Therefore, it is considered to be appropriate to measure the influence of the vertex. In addition, we generated two networks from all contacts in the high school and the hospital, and calculated the eigenvector centrality of each vertex.

In the experiments, we set $0.005 \leq \alpha \leq 0.05$ and $\tau = 0.97$. Furthermore, we set the threshold θ_i to 0.5 for all vertices. Then, we defined the contacts of the vertices with high eigenvector centralities as high-centrality contacts. The

eigenvector centrality of contacts ξ_{ij} is denoted by Eq. (7).

$$\xi_{ij} = \max\{E_i, E_j\}. \quad (7)$$

Here, E_i indicates the eigenvector centrality of vertex i . Next, we define M as the total number of contacts. The values of ξ_{ij} for the high school and the hospital against observation time t are shown in Fig. 3.

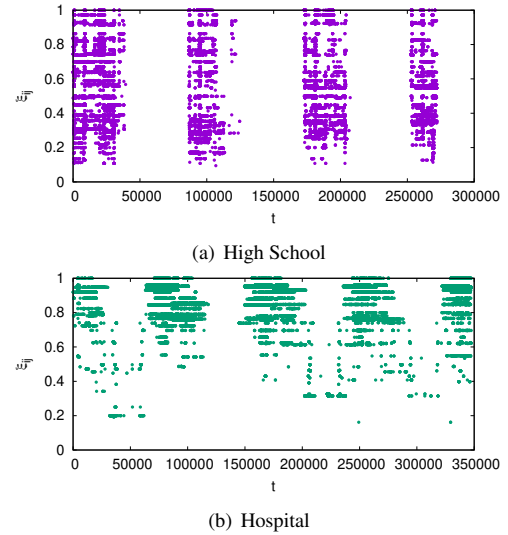


Figure 3: ξ_{ij} against observation time t . Here, contacts occurred in 20 seconds are plotted in the same time t , because the contacts were observed at intervals of 20 seconds. The blank areas mean that no contact was observed.

From Fig. 3, the contacts with high ξ_{ij} exist all over the range of time in both of the high school and the hospital. Furthermore, the hospital has a higher percentage of high-centrality contacts than the high school. We moved the high-centrality contacts in the second half into the first half of the contact data sets to increase the percentage of high-centrality contacts. In particular, we changed the orders of contacts before the experiments as follows:

- Step 1** Choose $M/10$ contacts randomly from the time range of $(0, T/2)$
- Step 2** Sort all contacts in the time range of $[T/2, T]$ by ξ_{ij} . Note that we sorted the contacts randomly if their values of ξ_{ij} are the same.
- Step 3** Choose $M/10$ contacts from the time range of $[T/2, T]$ in the sort order of Step 2. Note that we only choose half of the contacts of the same vertex. The reason is that if we chose all contacts of the same vertex, the vertex (person) would not contact with other vertices (people) in the time range of $[T/2, T]$.
- Step 4** Exchange the contacts chosen in Step 1 and Step 3.

Figure 4 shows the above procedure (Steps 1–4). Using the above procedure (Steps 1–4), we generated 100 TNs and investigated the averaged diffusion rates.

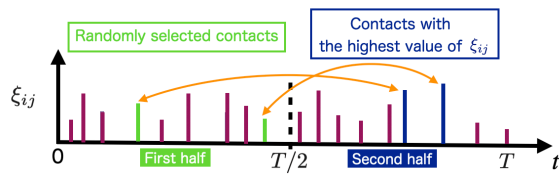


Figure 4: Schematic diagram of exchanging contacts. Green bars represent the contacts randomly chosen from the first half of the data. Blue bars represent the contacts with the highest values of ξ_{ij} in the second half of the data. The length of bars represents the value of ξ_{ij} of contacts.

4.2.2 Results

We investigated the relationship between α and S on the original TNs and the reconstructed TNs by exchanging contacts. The results are shown in Fig. 5.

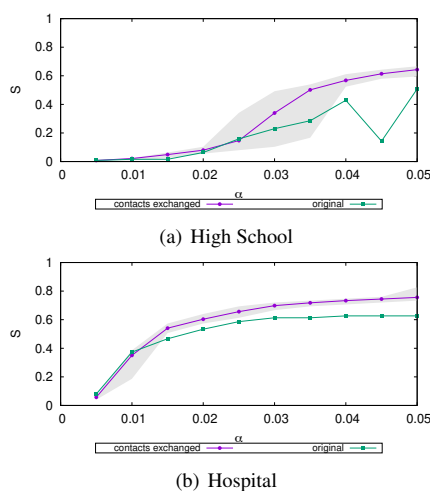


Figure 5: Relationship between α and S on the TNs. The dashed lines and the grey areas show the average values of S and the ranges between the maximum and the minimum values of S in the cases of the reconstructed TNs.

From Fig. 5, we found that average values of S in the cases of the reconstructed TNs are higher than that of the original TN when $\alpha \geq 0.03$ in the high school. In addition, we found that S of the reconstructed TNs are higher than that of the original TN when $\alpha \geq 0.015$ in the hospital. Furthermore, the values of S of the original TN are included in the grey area where $0.02 \leq \alpha \leq 0.035$ in the case of high school, and $0.005 \leq \alpha \leq 0.01$ in the case of hospital. These results suggest that information is likely to spread more widely when the contacts of vertices with high eigenvector centralities occur more frequently than usual in the first half of the contact data sets. In other words, the contacts of high eigenvector centrality vertices play an important role in the early phase of the information diffusion.

5. Conclusions

In this paper, we investigated the relationship between the diffusion rate and the contact order on information diffusion on the TNs while changing the contact density, by applying the proposed model of information diffusion. As a result, we found that information is likely to spread more widely when the contact density is high, and the contact order strongly affects the information diffusion. Furthermore, we investigated the influence of the vertices with high eigenvector centralities on the diffusion rate of information by exchanging contacts. As a result, we found that information spreads more easily when the percentage of the contacts of vertices with high centralities is high. From these results, it suggests that we can activate information diffusion by keeping the original contacts and changing their occurrence order.

Acknowledgment

This research is partially supported by the JSPS Grant-in-Aids for Scientific Research (Nos. JP15KT0112, JP17K00348, JP18K18125 and JP20H00596).

References

- [1] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ network,” *Nature*, Vol. 393, No. 6684, pp. 440–442, 1998.
- [2] D. Centola, “The spread of behavior in an online social network experiment,” *Science*, Vol. 329, No. 5596, pp. 1194–1197, 2010.
- [3] A. Tanaka et al., “Analysis on information diffusion of a mathematical model with dynamical sending probability,” *Proc. of NOLTA2013*, pp. 298–301, 2013.
- [4] L. Xu et al., “Information diffusion model with network centrality and community structure,” *IEICE Technical Report*, Vol. 119, No. 471, pp. 83–88, 2020 (in Japanese).
- [5] L. Xu et al., “Analysis of influence of network structure on information diffusion,” *IEICE Technical Report*, Vol. 120, No. 55, pp. 23–26, 2020.
- [6] L. Xu et al., “Analysis of information diffusion on temporal networks,” *IEICE Society Conf.*, N-1-11, 2020.
- [7] <http://www.sociopatterns.org/> (Accessed: 2021/01/07)
- [8] P. Vanhems et al., “Estimating potential infection transmission routes in hospital wards using wearable proximity sensors,” *PLOS ONE*, Vol. 8, No. 9, e73970, 2013.
- [9] J. Fournet and A. Barrat, “Contact patterns among high school students,” *PLOS ONE*, Vol. 9, No. 9, e107878, 2014.
- [10] P. Bonacich, “Factoring and weighting approaches to status scores and clique identification,” *Journal of Mathematical Sociology*, Vol. 2, No. 1, pp. 113–120, 1972.

A Method for Searching Edge Candidates of TSP

*Masayoshi Tosa *Honami Tsushima *Nanfu Miya *Tohru Ikeguchi

*Tokyo University of Science

Abstract

Traveling Salesman Problem is one of the \mathcal{NP} -hard combinatorial optimization problems. Thus, it is important to develop approximation algorithms for finding near-optimal solutions. The Lin-Kernighan-Helsgaun-heuristic (LKH) is one of the most powerful approximation algorithms, which can obtain good approximate solutions quickly. In the case of using LKH, LKH uses a set of edges presumed to be included in the optimal solution. Namely, the performance of LKH depends on the edge candidates. In this paper, we propose a method for generating good edge candidates, which has almost the same performance as the reduction of the computational time. Numerical experiments show that the proposed method can reduce the computation time while maintaining the solution performance of LKH.

1. Introduction

The traveling salesman problem (TSP) is defined as follows: given a set of n cities V and cost $d(i, j)$ between two cities i , and j , then find the shortest tour that visits only once all given cities. Thus, TSP is the problem of finding a bijection $\rho : V \rightarrow V$ that minimizes the objective function defined by

$$\sum_{v \in V} d(v, \rho(v)). \quad (1)$$

where $\rho(v)$ corresponds to the city visited after the city v .

The reason why TSPs are frequently solved is that it can be applied for solving a wide range of real-world problems, such as VLSI design [1]. However, TSP is one of the \mathcal{NP} -hard problems, then it is important to develop not only an exact algorithm but also an approximate algorithm.

The Lin-Kernighan-Helsgaun-heuristic (LKH) [2] is one of the high-performance approximate solution methods for solving TSP. It has been reported that LKH can find good approximate solutions in a short time frame by using a set of cities and edges that are estimated to be included in the optimal solution. In this paper, these edges are called edge candidates.

It is true that it is effective to find good edge candidates, however, if the number of cities increases, it takes a long time to obtain good edge candidates. In addition, the computational time and solution performance of LKH highly depend on the edge candidates. Therefore, it is an important issue to develop an efficient method of generating good edge candidates.

In Ref.[3], a method for searching edge candidates has been proposed (the conventional method). In the conventional method, one can get multiple tours generated by Partial Optimization Metaheuristic Under Special Intensification Conditions (POPMUSIC) [4], which can construct a tour shortly. All edges in obtained tours by POPMUSIC are regarded as edge

candidates. However, the computational time of the conventional method increases if the performance of edge candidates would be improved.

This paper proposes a method for obtaining good edge candidates by improving the conventional method by changing an algorithm in the POPMUSIC. Numerical experiments show that the proposed method improves the performance of edge candidates while reducing computational time.

2. The Conventional Method

In the conventional method, 20 tours are constructed using POPMUSIC, and all edges in the constructed tours are regarded as edge candidates. The process of generating edge candidates is shown in Fig. 1.

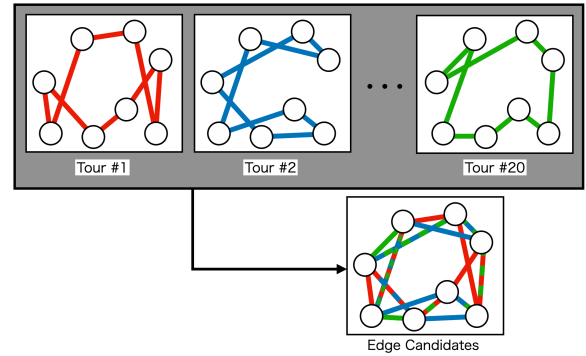
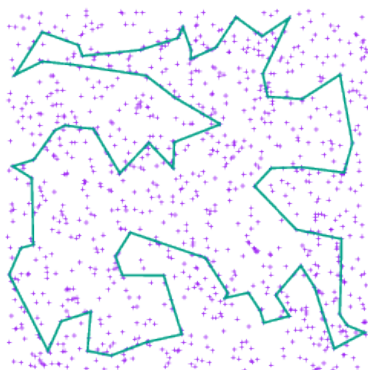


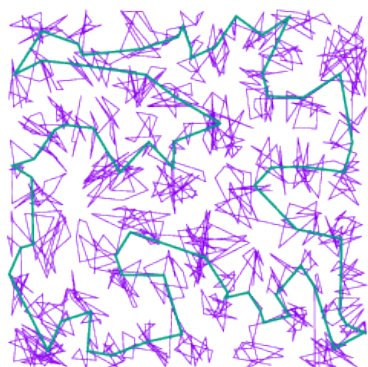
Figure 1: Process of generating edge candidates.

The algorithm of POPMUSIC is shown as follows:

1. Select n_a cities randomly from V to make the representative city set V' .
2. Construct a random tour T_s for n_a cities in the representative city set V' .
3. Optimize T_s using the Lin-Kernighan-heuristic[5](LK) (a tour colored in green of Fig. 2(a)).
4. Let the tour $T = T_s$.
5. For T , insert $v \notin V'$ immediately after $u \in V'$ with the smallest $d(u, v)$ (Fig. 2(b)).
6. For each city $v \in V'$ (Fig. 2(c)).
 - (a) For T_s , let w be 2 cities away from v .
 - (b) Let T' be a subtour that starts at the city visited immediately before v and ends at w in T .
 - (c) Fix the start and end points of T' and optimize it using the 2-opt method (Fig.3).
 - (d) Update T , following the subtour T' optimized in Step (c).
7. Terminate this algorithm (Fig.2(d)).



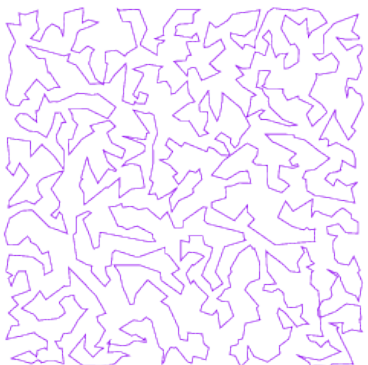
(a) Representative tour T_s (Green line) improved by LK. Purple dots represent all cities V .



(b) Insert not representative cities ($\notin U$) for T (purple line).



(c) Optimize partial tour.



(d) End of optimization.

Figure 2: A process of POPMUSIC. In this example, the instance E1k.0[6] is used.

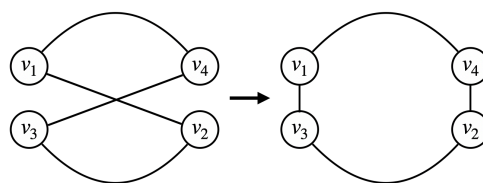


Figure 3: An example of 2-opt. If $d(v_1, v_3) + d(v_2, v_4) < d(v_1, v_2) + d(v_3, v_4)$, edges (v_1, v_2) and (v_3, v_4) are removed and cities v_1 and v_3 and cities v_2 and v_4 are connected.

In Step 3, POPMUSIC optimizes T_s using the LK. Then, in the following, we show the algorithm of LK. A procedure of the improvement of the solution is shown in Fig.4.

1. Construct a random tour T .
2. Let the city set $U = V$.
3. Select a city $v_1 \in U$ and an edge (v_1, v_2) .
4. Let the tour $T' = T$ and $i = 1, g = 0$.
5. Set $i = i + 1$ and select v_{2i-1} , and v_{2i} that satisfy the following conditions (a)~(c). If there are no such cities, proceed to Step 9.
 - (a) No overlap with $v_1 \sim v_{2i-2}$.
 - (b) T' contains the edge (v_{2i-1}, v_{2i}) .
 - (c) $d(v_{2i-1}, v_{2i-2}) < d(v_{2i-1}, v_{2i})$
6. Delete the edges (v_1, v_{2i-2}) and (v_{2i-1}, v_{2i}) of T' and add (v_1, v_{2i}) and (v_{2i-1}, v_{2i-2}) .
7. Let $g = g + d(v_1, v_{2i-2}) + d(v_{2i-1}, v_{2i}) - d(v_{2i-1}, v_{2i-2}) - d(v_{2i}, v_1)$.
8. If $g \leq 0$, return to Step 5. If $g > 0$, Let $T = T'$ and return to Step 2.
9. Remove v_1 from U and return to Step 3. If $U = \{\phi\}$, terminate the algorithm with T as the final solution.

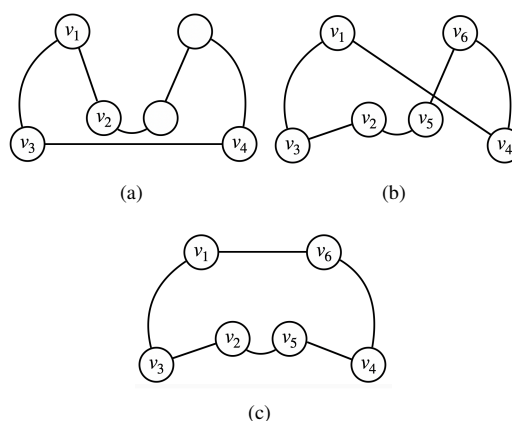


Figure 4: Improvement of the solution by LK. (a) Step 5 when $i = 1$. (b) Step 5 when $i = 2$. (c) Step 8 ($g > 0$).

The computational time of POPMUSIC is shorter than other algorithms such as LK and 2-opt. When the number of cities is n , the algorithmic complexity of each method is $O(n^{2.78})$ for the LK and $O(n^{2.29})$ for the 2-opt [3]. On the other hand, POPMUSIC is $O(n^{1.56})$ when $n_a = n^{0.56}$. In the following, we discuss the derivation of the algorithmic complexity of POPMUSIC.

POPMUSIC divides a large problem into small problems. By applying the LK and the 2-opt to the divided small problems, the algorithmic complexity of the POPMUSIC is reduced. Now, we consider the algorithmic complexity of Steps 3 and 6 because these steps could be bottlenecks.

Let the number of representative cities $n_a = n^h$ ($0 < h < 1$). In Step 3, POPMUSIC optimizes the representative cities using the LK. The number of representative cities is n^h , and the algorithmic complexity of the LK method is $O(n^{2.78})$. Therefore, the algorithmic complexity of Step 3 is $O(n^{2.78h})$.

In Step 6, POPMUSIC optimizes the subtour T' of the tour T using the 2-opt. The expected value of the number of cities in the subtour T' is $2n^{1-h}$, and the algorithmic complexity of the 2-opt is $O(n^{2.29})$. Thus, the algorithmic complexity for optimizing one subtour is $O(n^{2.29(1-h)})$. Since POPMUSIC needs to perform this operation for the number of representative cities n^h , the algorithmic complexity of Step 6 is $O(n^{2.29-1.29h})$.

Therefore, the total algorithmic complexity of POPMUSIC is $O(n^{2.78h + n^{1+h} + n^{2.29-1.29h}})$. Finding h that minimizes this algorithmic complexity is equivalent to finding h that minimizes $\max(2.78h, 1+h, 2.29-1.29h)$. The relationship between h and $2.78h$, $1+h$, and $2.29-1.29h$ is shown in Fig. 5. When $h = 0.56$, $\max(2.78h, 1+h, 2.29-1.29h)$ is minimized and the algorithmic complexity is $O(n^{1.56})$.

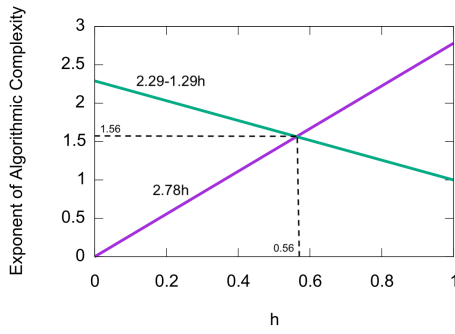


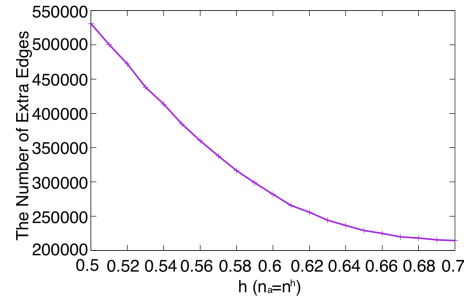
Figure 5: Relationship between the value of h and the algorithmic complexity of POPMUSIC. The algorithmic complexity of Step 3 is $O(n^{2.78h})$, and Step 6 is $O(n^{2.29-1.29h})$.

3. Proposed Method

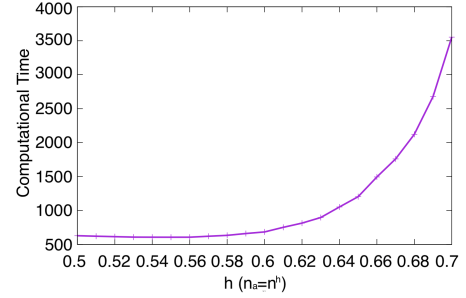
In Fig. 6(a), we show the relationship between n_a and the number of edges which are not the optimal solution but included in the edge candidates by the conventional method. We used pla85900 from the benchmark instance collection TSPLIB[1]. If the number of edges not included in the optimal solution is small, the edge candidates are good. From Fig. 6(a), it is necessary to increase the number of representative cities to obtain good edge candidates.

In Fig. 6(b), we show the relationship between the value of h and the computational time required to generate edge candidates. We used the same instance as Fig. 6(a). From Fig. 6(b), it is necessary to reduce the number of representative cities to reduce the computational time.

These results show that a large number of representative cities requires a lot of computational time, while a small number of representative cities results in a large number of extra edges. Therefore, there is a trade-off between the computational time and the performance of the edge candidates.



(a)



(b)

Figure 6: Relationships between the value of h and (a) the number of extra edges, and (b) the computational time.

In this paper, we propose a method that uses the 2-opt instead of the LK used in Step 3 in POPMUSIC. The number of representative cities n_a can be increased by replacing the LK with the 2-opt.

By replacing the LK with the 2-opt, for optimizing the representative city in Step 3, the algorithmic complexity of Step 3 becomes $n^{2.29h}$. Therefore, the total algorithmic complexity of Steps 3 and 6 is now $O(n^{2.29h + n^{2.29-1.29h}})$. Finding h that minimizes $O(n^{2.29h + n^{2.29-1.29h}})$ is equivalent to finding the h that minimizes $\max(2.29h, 2.29-1.29h)$. The relationship between h and $2.29h$, $2.29-1.29h$ is shown in Fig. 7. When $h = 0.65$, $\max(2.29h, 2.29-1.29h)$ is minimized and the total algorithmic complexity of Steps 3 and 6 is $O(n^{1.49})$.

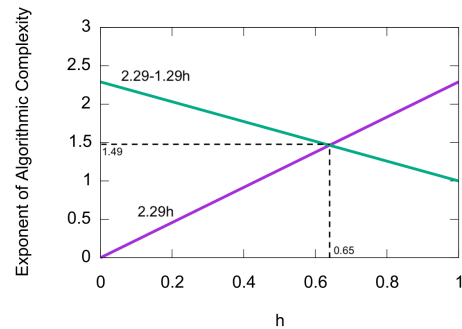


Figure 7: Relationship between the value of h and the algorithmic complexity of the proposed method. The algorithmic complexity of Step 3 is $O(n^{2.29h})$, and Step 6 is $O(n^{2.29-1.29h})$.

4. Numerical Experiments

To compare the performance of the conventional method and the proposed method, we used the gap and the computational

time. In this experiment, we used the benchmark instance collection DIMACS Challenge [6]. Most of the optimal solution of the DIMACS challenge instances are not obtained. Then, the gap is defined as follow:

$$\text{Gap}[\%] = \frac{(\text{Obtained solution}) - (\text{Known best solution})}{(\text{Known best solution})} \times 100. \quad (2)$$

Table 1 shows results of the numerical experiments. The first column represents the instance name, and the second column represents the number of cities. The third and fourth columns provide the gaps of the conventional and proposed methods. The fifth column represents the ratios of the conventional and proposed methods. Then, the sixth to the eighth columns represent the calculation time of both methods and their ratios. Table 1 shows no significant difference between the gaps in these methods. However, the computational time of the proposed method is reduced. These results indicate that the proposed method generates edge candidates with almost the same performance but reduces the computational time.

Table 1: Results of performance.

Instance	The number of cities	Gap(%)			Computational time(s)		
		Conv. (A)	Prop. (B)	Ratio (B/A)	Conv. (C)	Prop. (D)	Ratio (D/C)
E10k.0	10,000	0.081	0.077	0.951	69.66	49.29	0.708
E31k.0	31,623	0.13	0.13	1.000	392.60	289.42	0.737
E31k.1	31,623	0.77	0.76	0.987	395.09	288.56	0.730
E100k.0	100,000	0.82	0.84	1.024	2184.67	1580.14	0.723
E100k.1	100,000	0.17	0.19	1.118	2175.38	1571.78	0.723
E316k.0	316,228	0.19	0.21	1.105	11398.96	8541.49	0.749

5. Discussion

We investigated the detail of the computational time. Table 2 represents the time of generating edge candidates and calculating instances by LKH. For small-size instances, the proposed method takes less time than the conventional method both for generating edge candidates and conducting LKH. On the other hand, for large-size instances, the proposed method is inferior to generate edge candidates. However, the proposed method significantly reduces the computational time in LKH. As a result, the total computational time of the proposed method is shorter than that of the conventional method. These results show that the edge candidates generated by the proposed method can reduce the computational time of LKH.

Table 2: Results of computational time.

instance	The number of Cities	Generating time of edge candidate (s)			Computational time of LKH(s)		
		Conv. (A)	Prop. (B)	Ratio (B/A)	Conv. (C)	Prop. (D)	Ratio (D/C)
E10k.0	10,000	43.94	26.72	0.598	25.72	22.57	0.869
E31k.0	31,623	201.76	149.62	0.742	190.84	139.80	0.733
E31k.1	31,623	201.49	148.76	0.738	193.60	139.80	0.724
E100k.0	100,000	844.62	837.11	0.991	1340.05	743.03	0.554
E100k.1	100,000	837.27	827.13	0.988	1338.11	744.65	0.556
E316k.0	316,228	3575.71	4734.92	1.324	7823.25	3806.57	0.487

Next, we also investigated the number of edge candidates and the number of edge candidates included in the optimal solution. We used the benchmark instance collection TSPLIB [1]. If the number of the edge candidates is small and the edge candidates include many optimal edges, we can regard the edge candidates as having a good performance. The results are shown in Table 3. The proposed method maintains the number of edges in the optimal solution. However, the method reduces the total number of edges for the large-size instances compared

to the conventional method. It indicates that for large-size instances, the proposed method adequately removes extra edges that should not be included in the optimal solution.

Table 3: Evaluation of edge candidates.

Instance	The number of cities	Total number of edge candidates			The number of edges included in the optimum		
		Conv. (A)	Prop. (B)	Ratio (B/A)	Conv. (C)	Prop. (D)	Ratio (D/C)
pr2392	2,392	7,831	9,096	1.162	2,388	2,385	0.999
d15112	15,112	59,038	63,450	1.075	15,038	15,087	1.003
pla33810	33,810	146,022	133,363	0.913	33,748	33,717	0.999
pla85900	85,900	446,016	346,183	0.776	85,807	85,770	1.000

6. Conclusion

In this paper, we propose a method for generating good edge candidates for obtaining approximate solutions. Numerical experiments show that the proposed method can reduce the computational time compared to the conventional method.

The conventional and the proposed methods use 20 tours to generate edge candidates. The performance of the edge candidates is expected to depend on the number of tours used. It is an important future work to investigate an optimal value of the number of tours.

This research is partially supported by the JSPS Grant-in-Aids for Scientific Research (No. JP15KT0112, JP17K00348, and JP20H00596).

References

- [1] G. Reinelt, *ORSA J. on Comp.*, **3**, 4, 376–384, 1991.
- [2] K. Helsgaun, *EJOR*, **126**, 1, 106–130, 2000.
- [3] É. D. Taillard, and K. Helsgaun, *EJOR*, **272**, 2, 420–429, 2019.
- [4] É. D. Taillard, and S. Voss, *Essays and surveys in meta-heuristics*, 613–629, 2002.
- [5] S. Lin, and B. W. Kernighan, *Oper. Res.*, **21**, 2, 498–516, 1973.
- [6] D. S. Johnson, <http://archive.dimacs.rutgers.edu/Challenges/TSP/>, 2001.



Application of Recurrence Quantification Analysis to Hypertension Photoplethysmograms

Nina Sviridova and Tohru Ikeguchi

Faculty of Engineering, Tokyo University of Science
6-3-1 Nijjuku, Katsushika, Tokyo 125-8585, Japan
Email: nina.svr@rs.tus.ac.jp, tohru@rs.tus.ac.jp

Abstract– Hypertension is one of the widespread cardiovascular system disorders, which may cause various health complications. Hypertension can be diagnosed by the blood pressure (BP) measurement; however, conventional cuff-based devices cannot be used for continuous health monitoring. In recent studies, photoplethysmogram (PPG) was applied for cuff-less BP estimation. This PPG-based technique has great potential for application in the wearable devices for continuous real-time health monitoring and can be used for early identification of hypertension condition. Unlike the case of the clinical settings, real-time health monitoring by the PPG recorded by the wearable devices may require operating with the short PPG recordings. The PPG dynamics is highly complex, and the methods of nonlinear time series analysis can be efficiently used for its analysis. This data-driven study aimed to investigate the possibility of a recurrence quantification analysis application for the detection of hypertension conditions in short-recorded PPG signals. The results demonstrated that the mean diagonal line length, maximal diagonal line length, and the trapping time calculated from the recurrence plot can be potentially useful for differentiation between normal, prehypertension, stage 1 and 2 hypertension cases.

1. Introduction

Human photoplethysmogram (PPG) is one of the noninvasively measured biological signals widely applied to health monitoring. It is well known that PPG carries rich information on the performance of the cardiovascular system [1]. The heart rates and oxygen estimation have been main PPG applications for decades, and recent studies demonstrated that PPG can also be used for cuff-less blood pressure (BP) estimation [2]. High blood pressure or hypertension is the globally spread cardiovascular system disorder, which depends on its category may be potentially harmful to health, as well as lead to or worsen other cardiovascular system diseases. Timely identification of hypertension conditions is an important issue. Continuous real-time health monitoring by wearable devices, such as smartwatches, could be its possible solution. However, the number of issues naturally arises when the PPG measured by the wearable device is used. One of them is that unlike the PPG measured in clinical settings, where the measurement protocol and

devices are unified, the wearable devices, that are the most promising in terms of health monitoring, have a wide range of device setups and the types of the PPG used. To address this issue one may need to rely on the essential characteristics of the PPG signal dynamics. The most commonly used NIR and green light PPGs are known to have chaotic dynamics [3-5]; therefore many studies have been conducted towards the identification of various physiological and mental health disorders by the nonlinear time series analysis of the PPG signals [3, 6, 7], that can reveal its underlying dynamical characteristics. Recurrence quantification analysis (RQA) [8] is one of these methods [9].

Another important issue is that real-time health monitoring by signals continuously recorded during daily activities may require to operate with short recordings, as the PPG signal may be corrupted by the motion artifacts, presence of measurement noise, etc.

Taking the above-mentioned issues into consideration, this data-driven study aims to investigate whether the RQA is effective for the identification of the differences in the dynamics of the short recorded PPGs obtained from hypertension condition subjects in comparison with healthy subjects' PPGs.

2. Data

In this study, the short-recorded photoplethysmogram database for blood pressure monitoring [10] was utilized. Each dataset was recorded for 2.1s with sampling frequency 1 kHz. Detailed experimental protocol, database description, its availability and data quality check information can be found in [11].

For this study data of subjects with normal, prehypertension (preH), stage 1 hypertension (S1H), stage 2 hypertension (S2H) records, and with no records of diabetes, cerebral infarction, and cerebrovascular disease were selected.

It is well-known that the PPG data may significantly change with age [12]. To reduce influence of the subjects' age, among the 219 subjects (age range: 21-86) datasets provided in the database, the data corresponding to the subjects over 45 years old were chosen for this study. Used data age range, the total number of datasets (N), average BP for each PPG group are summarized in table 1.

Table 1. Summary of hypertension groups records

Hypertension	N	Age	Average BP, mmHg
Normal	49	45-75	110/64
Prehypertension	73	45-84	130/72
Stage 1 hypertension	31	46-77	148/82
Stage 2 hypertension	20	45-80	168/89

A typical example of the raw data from normal and S2H group is shown in Fig. 1.

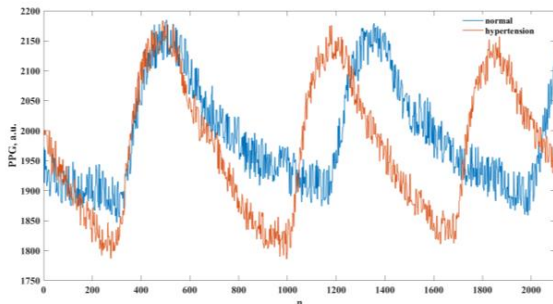


Fig. 1. A typical example of unprocessed PPG signal for normal (blue line) and stage 2 hypertension (red line) subject.

3. Results

3.1. Recurrence plot

The recurrence plot was calculated for each dataset. Based on the previous studies [4, 5], the applied embedding dimension was chosen equal to 4. Fig. 2 demonstrates an example of the obtained recurrence plots corresponding to the normal, preH, S1H, and S2H data.

3.2. Recurrence Quantification Analysis (RQA)

The RQA was applied to the calculated recurrence plots to get the quantitative characteristics of PPGs underlying dynamics. The RQA was conducted by the MATLAB CRP toolbox [8]. The following indexes were calculated: recurrence rate (RR), determinism (DET), mean diagonal line lengths (L), maximal diagonal line lengths (L_{max}), laminarity (LAM), trapping time (TT), and maximal vertical line lengths (V_{max}). The mean values of these indexes for each group are shown in table 2. Time series which DET values were below 0.9 were omitted from the study due to the possibility of significant noise contamination.

First, a one-factor ANOVA test was applied to each index to test the null hypothesis of equal means for four hypertension groups. The null hypothesis was confirmed for the RR and DET indexes; therefore, the difference in the mean values of these indexes is not significant; for other indexes the null hypothesis was rejected.

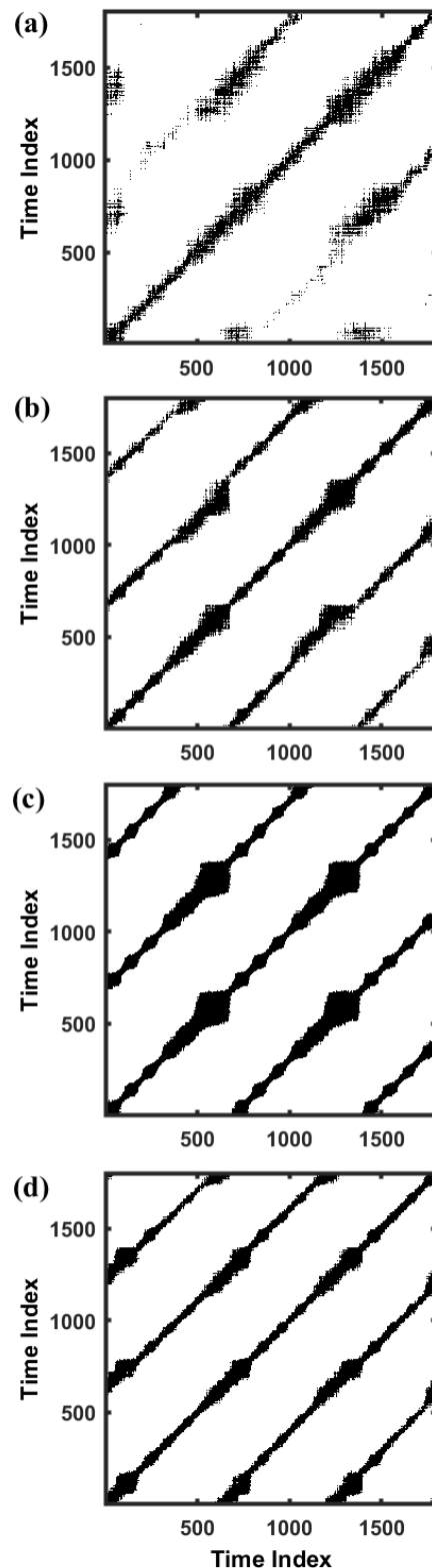


Fig. 2. An example of recurrence plot for the PPG signal corresponding to the (a) normal, (b) prehypertension, (c) stage 1 hypertension and (d) stage 2 hypertension subject.

Table 2. Mean values of calculated RQA indexes

	Normal	PreH	S1H	S2H
RR	0.15	0.15	0.15	0.16
DET	0.98	0.98	0.98	0.99
L	27.44	25.67	25.09	28.46
L _{max}	1600.78	1684.16	1742.39	1799.00
ENTR	2.98	2.94	2.93	3.03
LAM	0.99	0.99	0.99	1.00
TT	32.28	31.19	30.72	34.05
V _{max}	32.28	31.19	30.72	34.05

T-test was applied to the rest of the RQA indexes to test the null hypothesis on the equal means between pairs of different hypertension groups. The results demonstrated that none of the applied indexes could differentiate between S1H and S2H groups. For the following indexes and pairs of groups the null hypothesis was rejected with significance level 0.05:

- normal vs. preH: L, L_{max}, ENTR, LAM, TT, V_{max},
- normal vs. S1H: L, L_{max}, TT,
- normal vs. S2H: L, L_{max}, ENTR, LAM, TT,
- preH vs. S1H: L, L_{max}, ENTR, LAM, TT,
- preH vs. S2H: L, L_{max}, ENTR, LAM, TT.

4. Discussion

The aim of this study was to investigate whether the RQA indexes can be applied to differentiate the short-recorded PPG data obtained from subjects with normal, prehypertension, stage 1 and stage 2 hypertension. As seen from the typical examples of time series (Fig. 1) and different hypertension groups recurrence plots (Fig. 2), no significant qualitative difference can be observed. However, RQA results demonstrated that significant differences can be found in characteristics of the PPGs corresponding to normal, prehypertension, and hypertension data. The L, L_{max}, ENTR, LAM, and TT RQA indexes were found to be applicable to identify differences in the dynamics of studied data. However, no statistical differences were observed for the RR and DET indexes; and none of the applied indexes was able to differentiate between stage 1 and 2 hypertension groups.

It worths noticing that most of the PPG datasets for all groups are deterministic as follows from the results of DET calculation.

5. Conclusion

Results of this data-driven study demonstrated that the RQA measures, namely, mean diagonal line lengths, maximal diagonal line length, and the trapping time can potentially be applied for identification of the PPGs corresponding to the subjects with hypertension disorder.

This result can contribute to the area of real time health monitoring and early identification of hypertension.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number 19K14589 and 17K00348.

References

- [1] Allen, J. "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, 2007; 28: 1-39.
- [2] D. Fujita, A. Suzuki, K. Ryu, "PPG-Based Systolic Blood Pressure Estimation Method Using PLS and Level-Crossing Feature," *Appl. Sci.*, vol. 9, pp. 304 1-11, 2019.
- [3] I. Tsuda, "Chaotic pulsation in human capillary vessels and its dependence on mental and physical conditions," *International journal of bifurcation and chaos*, vol. 2(2), pp. 313–324, 1992.
- [4] N. Sviridova, K. Sakai, "Human photoplethysmogram: new insight into chaotic characteristics," *Chaos, solitons and fractals*, vol.77, pp. 53–63, 2015.
- [5] N. Sviridova, T. Zhao, K. Aihara, A. Nakano, K. Nakamura, "Photoplethysmogram at green light: Where does chaos arise from?" *Chaos, solitons and fractals*, vol.116, pp. 157–165, 2018.
- [6] T. D. Pham, T.C. Thang, M. Oyama-Higa, M. Sugiyama, "Mental-disorder detection using chaos and nonlinear dynamical analysis of photoplethysmographic signal," *Chaos, solitons & fractals*, vol. 51, pp. 64–74, 2013.
- [7] N. Sviridova, K. Sakai, "Application of photoplethysmogram for detecting physiological effects of tractor noise," *Engineering in agriculture, environment and food*, vol. 8(4), pp. 313–317, 2015.
- [8] N. Marwan, M. C. Romano, M. Thiel, J. Kurths, "Recurrence plots for analysis of complex systems," *Physics reports*, vol. 438, pp. 237–329, 2007.
- [9] N. Sviridova, "Detection of Preprocessing-Induced Changes in Chaotic Characteristics of Human Photoplethysmogram". *Proceedings of Int'l Symposium on Nonlinear Theory and its Applications 2019*, pp. 540-543, 2019.
- [10] Y. Liang, G. Liu, Z. Chen, M. Elgendi, PPG-BD database. Figshare. Dataset. <https://doi.org/10.6084/m9.figshare.5459299.v3>, 2018.
- [11] Y. Liang, Z. Chen, G. Liu, M. Elgendi, "A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China," *Scientific data*, vol. 5, pp. 180020 1-7, 2018.
- [12] J. Allen, A. Murray, "Age-related changes in peripheral pulse timing characteristics at the ears, fingers and toes," *J Hum Hypertens*, vol. 16, pp.11–17, 2002.

査読なし国内学会

ロジスティック写像のカオス応答とノイジーな周期解の判別

Discrimination between chaotic response and noisy periodic response of the logistic map

金丸 志生[†]

Nina Sviridova^{††}

島田 裕^{†††}

池口 徹^{†,††}

[†] 東京理科大学大学院 工学研究科

^{††} 東京理科大学 工学部

^{†††} 埼玉大学大学院 理工学研究科

1 はじめに

リカレンスプロット (RP) [1] は、観測された時系列の振る舞いを 2 次元の白黒画像へと可視化する手法であり、その画像に現れるプロットパターンから、時系列の特性を定性的に把握することができる。ただし、RP 上の黒画素の割合は、設定した閾値によって異なるため、閾値は重要なパラメータである。我々は既に閾値を変化させたときの RP の変化を定量的に調べることで、非線形時系列の特性を捉える手法を提案している [2]。提案手法では、まず、閾値の異なる複数の RP を作成する。次に、それらに対して再帰定量化解析 (RQA) [3] の一つである DET を算出して RP の変化を定量化する。閾値を変化させることで得られる DET の推移傾向を、時系列データの特性として捉える手法が提案手法である。本稿では、離散力学系から生じる決定論的カオスとノイジーな周期解の特性を、提案手法により判別することを目的とする。ただし本稿では、ノイジーな周期解を、安定な周期解にダイナミカルノイズが加わった応答と定義する。具体的には、決定論的非線形力学系であるロジスティック写像 [4] と、これにダイナミカルノイズが印加されたノイジーロジスティック写像 [5] から生成される時系列を対象に、提案手法の有効性を検討する。

2 提案手法

まず、解析対象とする時系列データはいずれも 1 次元の時系列信号のみを観測できると仮定するため、RP の作成に必要な状態空間を再構成する手法として遅延座標系への変換手法を用いる [6, 7]。次に、再構成状態空間内の 2 点間距離情報をもとに RP を作成する [1]。このとき、閾値は再構成状態空間内の 2 点間距離の最小値から最大値までの値に、 $0 \leq \theta \leq 1$ の割合を掛けて変化させ、2 点間距離が閾値以下の場合のみ、対応する画素を黒くする。その後、各閾値で作成した RP を、決定論性指標である DET [3] により定量化する。DET は、 $(45^\circ$ の対角線に平行な斜め線を構成する黒画素の総数)/(黒画素の総数) と定義される。すなわち、DET は RP 上に現れる斜め線の割合を表す。最後に、閾値と DET の関係から時系列の特性を判別する。

3 数値実験

本稿では、ロジスティック写像 [4] から生成される決定論的カオス応答と、ノイジーロジスティック写像 [5] から生成されるノイジーな周期解応答を対象に、提案手法による時系列の特性判別が可能であることを調査した。ここで、ノイジーロジスティック写像とは、 $x_{t+1} = ax_t(1 - x_t) + \sigma(x_t)\varepsilon_t$ で定義された、ロジスティック写像に右辺第二項の分散不均一性のダイナミカルノイズ項が加わったモデルである。ただし、 x_t は離散時刻 t における状態値、 a は分岐パラメータ、 ε_t は $(-0.5, 0.5)$ の一様分布に従う確率変数であり、 $\sigma(x_t)$ は $\sigma(x_t) = 0.5 \times \min\{ax_t(1 - x_t), 1 - ax_t(1 - x_t)\}$ と定義される。決定論的カオス応答は、 $a = 3.57, 3.69, 3.694, 3.6916, 4$ をロジスティック写像に与えて生成した。対してノイジーな周期解応答は、ロジスティック写像において 2^k ($k = 1, \dots, 7$) 周期解となるパラ

メータ $a = 3.4, 3.45, 3.55, 3.565, 3.569, 3.5698, 3.5699$ をノイジーロジスティック写像に与えて生成した。各時系列長は過渡状態として 100 点省いた後の 1,000 点とし、再構成状態空間の次元数は 3、遅れ時間は 1 とした。初期値と乱数シードを変更して 100 本の時系列を生成して得た、100 回平均の DET の推移傾向を図 1 に示す。横軸は対数軸で閾値を定めるパラメータ θ 、縦軸は DET を示す。図 1

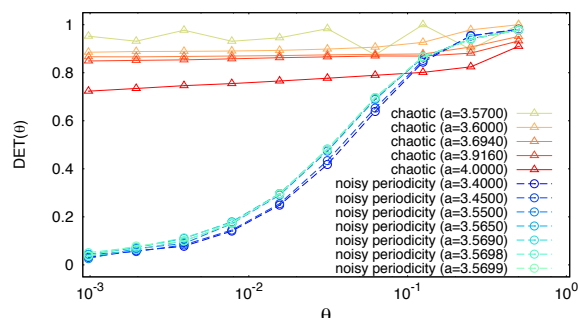


図 1: パラメータ θ に対する DET の推移。

より、カオスパラメータを与えたロジスティック写像の応答に対する DET の推移傾向 (実線三角印) と、周期パラメータを与えたノイジーロジスティック写像の応答に対する DET の推移傾向 (破線丸印) は明らかに異なることが分かった。決定論的カオス応答であれば、閾値によらず高い決定論性を有するのに対して、ノイジーな周期解応答では、閾値が大きくなるにつれて、DET の値が増加する傾向が得られた。従って、提案手法によって、離散力学系から生じる決定論的カオスとノイジーな周期解応答の判別が可能であることが明らかとなった。

4 まとめ

本稿では、RP の閾値を変化させることで得られる決定論性指標 DET の推移傾向から、決定論的カオスと周期解にダイナミカルノイズが加わったノイジーな周期解の特性判別が可能であることを調査した。結果より、DET の推移傾向がそれぞれ異なるため、提案手法は、離散力学系で見られる決定論的カオス応答とノイジーな周期解応答の判別に有効であることが分かった。なお、本研究の一部は JSPS 科研費 (Nos. JP17K00348, JP18K18125, JP19K14589, JP20H00596) の援助を受けて行われた。

参考文献

- [1] J. P. Zbilut *et al.*, Phys. Lett. A, **171**, 3–4, 199–203, 1992.
- [2] 金丸 志生 他, 信学技報, **120**, 330, 7–12, 2021.
- [3] N. Marwan *et al.*, Phys. Rep., **438**, 5–6, 237–329, 2007.
- [4] R. M. May, Nature, **261**, 5560, 459–467, 1976.
- [5] R. A. Meyers, “Complex systems in finance and econometrics,” Springer Science & Business Media, 2010.
- [6] F. Takens, Lecture Notes in Math., **898**, 366–381, 1981.
- [7] T. Sauer *et al.*, J. Stat. Phys., **65**, 3–4, 579–616, 1991.

因果性検出手法を用いたニューロン間の結合性推定

Estimating connectivity between neurons using causality detection methods

澤田和弥†

Kazuya Sawada

島田裕††

Yutaka Shimada

池口徹†, †††

Tohru Ikeguchi

東京理科大学大学院 工学研究科†

埼玉大学大学院 理工学研究科††

東京理科大学 工学部†††

1 はじめに

観測技術の向上により、様々な時系列データを観測できるようになっている。我々が観測する時系列データには、サンプリング間隔が一定な時系列以外にも、イベント発生時にのみ観測可能な点過程データも数多くある。例えば、地震、経済、神経系において観測されるデータは点過程の典型である。観測された点過程データ間の因果性を検出することが可能になれば、これらの複雑な振る舞いをする現象の理解に役立つだけでなく、予測・診断・効果的な制御などの工学的応用が可能となる。本稿では、神経細胞の数理モデルに対して、すでに我々が提案したスパイク列への因果性検出手法 [1] を用いた場合と発火率時系列に convergent cross mapping (CCM) [2] を用いた場合との因果性検出性能の比較を行う。

2 スパイク列および発火率時系列生成方法

本稿では、式 (1) に示す gap junction で結合する Izhikevich ニューロンモデルでスパイク列を生成した [3, 4].

$$\begin{cases} \dot{v}_i = 0.04v_i^2 + 5v_i + 140 - u_i + I_i + \sum_{j \neq i} w_{ij}(v_j - v_i) \\ \dot{u}_i = a_i(b_i v_i - u_i) \end{cases} \quad (1)$$

ここで、 $v_i \geq 30$ のとき、 $v_i \leftarrow c_i$, $u_i \leftarrow u_i + d_i$ のようにリセットされる。式 (1) において、 v_i , u_i はニューロン i の膜電位、回復変数である。また、 a_i, b_i, c_i, d_i, I_i はニューロン i のパラメータである。式 (1) を Runge-Kutta 法で解き、その際の刻み幅は 0.01 とした。本稿では、ニューロン数を $N = 2$ とし、パラメータは $a_1 = 0.2, a_2 = 0.201, b_1 = b_2 = 2, c_1 = c_2 = -56, d_1 = d_2 = -16, I_1 = I_2 = -99$ とした。また、結合強度 w_{ij} は (i) 両方向結合 ($w_{12} = 0.1, w_{21} = 0.05$), (ii) 片方向結合 ($w_{12} = 0, w_{21} = 0.05$), (iii) 結合なし ($w_{12} = w_{21} = 0$) の 3 つの場合で数値実験を行った。ここで、 v_i, u_i をリセットした時間にスパイク列を発生させた。本稿では、100 ステップで 1 秒とし、発火率は 100 秒ごとに発生したスパイク数の割合とした。

3 スパイク列に対する因果性解析手法

本稿では、発火率時系列に対し、CCM を適用した。CCM は、観測時系列に対して状態空間を再構成し [5, 6], 再構成状態空間の近傍情報に基づいて in-sample に相互予測することで、因果性を検出する手法である [2]。また、状態空間はスパイク列の発火間隔時系列から再構成可能なことも示されている [7]。そこで我々は、スパイク列の発火間隔時系列に対して発火時間を考慮して CCM を適用する手法を提案した [1]。提案手法では、発火時間が近いイベントの ISI から生成した遅延ベクトルを対応させることにより、予測の際の時間対応を取っている点が通常の CCM とは異なる。CCM および、CCM をスパイク列に拡張した提案手法では、影響を受けている時系列を用いて、影響を与えている時系列を予測する場合、予測に使用するデータ数の増加に伴い、予測精度が上昇・収束する。我々はこの

性質に基づき、文献 [2] と同様に相互予測精度から因果性を検出する。

4 結果

発火率時系列に対して CCM を適用した場合と、ISI 時系列に対して提案手法を適用した場合の予測に用いたデータ長と予測精度の関係を図 1 に示す。下の横軸は発火率

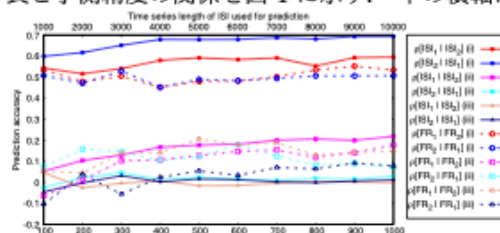


図 1 予測に用いたデータ長と予測精度の関係

時系列の時系列長、上の横軸は ISI 時系列の時系列長であり、縦軸は予測精度である。ここで、1 秒あたりの平均発火率はニューロン 1 が 0.082, ニューロン 2 が 0.088 である。したがって、発火率の時系列長が 1,000 の場合、スパイク数はニューロン 1 が 8,200, ニューロン 2 は 8,800 程度であるため、厳密には横軸の上下は一致していないが、取束傾向は変化しないと考えられる。また、 $\rho[ISI_i | ISI_j](k)$, $\rho[FR_i | FR_j](k)$ は、結合形態が k ($k = i, ii, iii$) の場合において、ニューロン j の ISI (ISI_j) もしくは発火率時系列 (FR_j) を用いてニューロン i の ISI (ISI_i) もしくは発火率時系列 (FR_i) を予測した結果を示す。図 1 から、発火率時系列を用いると、(i) の両方向結合の場合、予測精度は比較的高い値となっているが、(ii) の片方向結合、(iii) の結合なしの場合では、予測精度は時系列長によらず低い値で横ばいである。一方、ISI 時系列に提案手法を用いると、(i) の場合、(ii) の $\rho[ISI_i | ISI_j]$ (ii) の場合は、時系列長の増加に伴って予測精度は上昇し、(ii) の $\rho[ISI_i | ISI_j]$ (ii), (iii) の場合では予測精度は 0 に収束している。したがって、ISI を用いる手法では特に片方向結合による因果性、因果性なしの場合で有効であると考えられる。

5 まとめ

本稿では、発火率時系列に CCM を適用した場合と ISI 時系列に提案手法を適用した場合における因果性検出性能の比較・検討を行った。結果として、ISI を用いる提案手法は発火率時系列に CCM を適用するよりも優れていることが示された。なお、本研究の一部は JSPS 科研費 (JP15KT0112, JP17K00348, JP18K18125, JP20H00596) の援助を受けて行われた。

参考文献

- [1] 澤田 和弥 他, 信学ソ大, N-1-12, 2019.
- [2] G. Sugihara et al., *Science*, **338**, 6106, 496–500, 2012.
- [3] E. M. Izhikevich, *IEEE TNN*, **14**, 6, 1569–1572, 2003.
- [4] 内木 楓 他, 信学論 A, **J100-A**, 5, 195–204, 2017.
- [5] F. Takens, *Lecture Notes in Math.*, **898**, 366–381, 1981.
- [6] T. Sauer et al., *J. Stat. Phys.*, **65**, 3–4, 579–616, 1991.
- [7] T. Sauer, *Phys. Rev. Lett.*, **72**, 24, 3811–3814, 1994.

リカレンスプロット密度の変化による非線形時系列の特性抽出

金丸 志生[†] Nina Sviridova^{††} 島田 裕^{†††} 池口 徹^{†,††}

[†] 東京理科大学 大学院工学研究科 経営工学専攻 〒125-8585 東京都葛飾区新宿 6-3-1

^{††} 東京理科大学 工学部 情報工学科

^{†††} 埼玉大学大学院 理工学研究科 数理電子情報部門 〒338-8570 埼玉県さいたま市桜区下大久保 255

E-mail: †shikimaru@hisenkei.net

あらまし リカレンスプロットは力学系を定性的に把握するための有効な非線形時系列解析手法の一つである。また、再帰定量化解析手法を用いることでリカレンスプロットを定量化することができる。本稿では、リカレンスプロットとその拡張手法である同方向性リカレンスプロット、同方向的近傍プロットを使用し、リカレンスプロット密度を変化させることで得られる定量化指標の推移傾向から、非線形時系列の特性を抽出する手法を提案している。実験結果より、決定論的カオス時系列と周期応答にダイナミカルノイズが印加された時系列の特性判別が可能であることが示された。加えて、決定論的カオス時系列と非整数ブラウン運動のような有色ノイズ時系列の判別も可能であることが示された。キーワード リカレンスプロット、同方向性リカレンスプロット、同方向的近傍プロット、リカレンスプロット密度、再帰定量化解析、非線形時系列の特性抽出

Extraction of Property for Nonlinear Time Series by Changing Density of Recurrence Plots

Shiki KANAMARU[†], Nina SVIRIDOVA^{††}, Yutaka SHIMADA^{†††}, and Tohru IKEGUCHI^{†,††}

[†] Department of Management Science, Graduate School of Engineering, Tokyo University of Science
Nijuku 6-3-1, Katsushika-ku, Tokyo, 125-8585, Japan

^{††} Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science

^{†††} Department of Information and Computer Sciences, Graduate School of Science and Engineering,
Saitama University

255 Shimo-Okubo, Sakura-ku, Saitama City, Saitama 338-8570, Japan

E-mail: †shikimaru@hisenkei.net

Abstract A recurrence plot is one of the most effective nonlinear time series analysis methods for qualitatively understanding a nonlinear dynamical system. Furthermore, recurrence quantification analysis methods can be used to quantitatively evaluate the recurrence plot. In this report, we extract the properties of nonlinear time series from changes in quantification indices obtained by varying plot density on the recurrence plot and its extension methods, iso-directional recurrence plot and iso-directional neighbors plot. The experimental results suggest that the proposed method can discriminate the properties between deterministic chaotic time series and noisy periodic time series, and that it can also discriminate deterministic chaotic time series and the time series with colored noise such as a fractional Brownian motion.

Key words Recurrence Plot, Iso-Directional Recurrence Plot, Iso-Directional Neighbors Plot, Density of RP, Recurrence Quantification Analysis, and Extraction of Property for Nonlinear Time Series

1. はじめに

時系列解析に関する研究の目標の1つは、あるシステムから得られる複雑な時系列の予測と制御である [1], [2]。高い予測精度や制御効率を実現するためには、システムの特徴を捉え

たモデルの構築が求められ、良いモデリングを行うには、対象システムの構造の理解が重要となる。本稿では、アトラクタの構造を視覚化するために、リカレンスプロット (recurrence plot, RP) [3], [4] と呼ばれる手法を用いる。RP は、決定論的非線形力学系のアトラクタ上の各点間の相関関係を、時間発展情報を

保持しながら2次元画像へと視覚化する手法である。この2次元画像は、両軸が元のアトラクタの時間インデックスに対応した正方形の画像である。画素 (i, j) は、時刻 i と j の状態値が互いに近傍に存在する場合は1、存在しない場合は0とする。これにより、0と1で構成される2次元の2値画像が作成され、この画像のことをRPと呼ぶ。RPは非定常性の検出にも有効であるとされている[3]。また、RPにはアトラクタに関するトポロジカルな情報が含まれると考えられている[5]。RP上ではいくつかの代表的なパターン[6]が存在し、それらから時系列の背後に潜む力学系の特性を把握することができる。これらの理由から、RPは現在、最も有力な非線形時系列解析手法の1つとなっている。

また、RPの拡張手法も数多く提案されている[6]。例えば、カオス時系列解析において、アトラクタ上の点の近傍関係だけでなく、軌道ベクトルの情報を用いることも重要である。そこで、アトラクタ上の点がどの方向に遷移するかという軌道ベクトルの情報を考慮したRPの拡張手法として、同方向性リカレンスプロット(iso-directional recurrence plot, IDR) [7]が提案されている。RPがアトラクタ上の各点間の近傍関係が認められる場合にプロットする手法であるのに対して、IDRは、アトラクタ上の各点間の推移ベクトルの差から同方向性が認められる場合にプロットする手法である。さらに、RPとIDRを用いて、近傍かつ同方向である関係を可視化する同方向的近傍プロット(iso-directional neighbors plot, IDNP) [7]も提案されている。確率論的な時系列から作成したIDNP上には、近傍かつ同方向の関係となるプロット数が少なくなると予想される。したがって、IDNPは時系列の決定論性検出に効果的であると考えられる[7]。

しかしながら、RPやIDR、IDNPを適用する際には、慎重に設定しなければならないパラメータが存在する。RPの場合は、アトラクタ上の任意の2点間の距離に対して時間インデックスに対応する画素にプロットするか否かを判断するための閾値が必要となる。IDRの場合は、アトラクタ上の任意の2点から得られる推移ベクトル間の差のノルムに対する閾値が必要である。IDNPの場合は、RPとIDR上のプロットに応じてプロット数が決まるため、直接的な閾値は存在しない。ただし、IDNPを作成するうえで、RPとIDRの閾値は重要なパラメータである。これらの閾値の設定を誤った場合、時系列の特徴を捉えることができない場合がある。そのため、RPの閾値の設定方法についての研究が行われている[4],[8]~[11]。一方で、RPの最良な閾値は対象構造に強く依存することも示されている[6],[12]。また、各画素が0と1で構成される2次元画像ではなく、2点間距離を各画素に与えたりカレンス行列による解析手法も提案されている[13]。これにより、閾値の設定をする必要がなくなるため、あるアトラクタに対して一意にリカレンス行列を作成することができる。しかし、リカレンス行列による定量化は、RPを定量化することよりも難しいため、2値化したRPを使用することが一般的とされている[5]。

そこで我々は、閾値による2値化処理を施しつつも、ある1つの閾値を採用するのではなく、閾値を変化させて複数のRP、

IDR、IDNPを作成することで、時系列の特性を抽出する手法を提案している[14]~[17]。これらの手法では、閾値変化に伴う2次元画像上にプロットされる点の密度(RP密度)とRP、IDR、IDNP上に現れるパターンの関係に着目している。ただし、RP、IDR、IDNPは時系列を2次元画像へと変換することで定性的に時系列を解析する手法であるので、再帰定量化解析(recurrence quantification analysis, RQA) [6],[18],[19]と呼ばれる手法を用いて、これら2次元画像上に現れるパターンを定量化する。RQAの中でも、RP上に現れる斜め線の割合を定量化したDET [6],[18]と呼ばれる指標は、時系列の決定論性の指標として知られている。また、文献[7]では、RPとIDRのプロット数から決まる決定論性の指標として、IDNPのRPに対する生起確率 R を提案している。これら2つの決定論性の指標を用いて、RP密度の変化に対するDETと R の推移傾向が得られる。この推移傾向を時系列の特性として捉える手法が、我々が提案している手法[14]~[17]である。

しかしながら、提案手法について、ランダム力学系のような決定論と確率論が混在する系に対する調査は不十分である。そこで本稿では、提案手法を適用することで、決定論的カオス時系列と周期応答にダイナミカルノイズが印加された時系列(周期+ノイズ)について、それぞれで異なる特性が得られることを報告する。これまでに、これら2つのタイプの時系列の判別には、カオスの性質の一つである短期予測可能性と長期予測不可能性を利用した判別手法が提案されている[20]~[22]。これらの手法では、非線形予測手法の予測精度が予測ステップ数に対してどのように変化するかを調べることで、時系列の背後に存在するシステムの振る舞いを推定する。ただし、予測ステップ数とは、時刻 t までの時系列から s 点先の $t+s$ の状態値を予測する際の点数 s に対応する。例えば、予測ステップ数の増加に対して予測精度が落ちる場合は対象がカオス時系列、予測精度が予測ステップ数に依存せず一定であれば周期+ノイズ時系列であると判断する手法である[20]~[22]。一方で、この手法では非整数ブラウン運動(fractional Brownian motion, fBm) [23]のような有色ノイズを識別することは難しいと考えられる[2]。これはfBmがカオスと同様に短期予測可能性を有するためである。また、強い周期性を持つデータの場合も判別がうまく行われないことがある[24]。これに加えて、カオス時系列とfBm時系列の判別には、予測時間に対する初期変位の変化が、カオスであれば指数関数的に拡大されるのに対して、fBmであれば時間の冪に従うことを利用した手法が提案されている[25]。

そこで本稿では提案手法によるカオス時系列と周期+ノイズ時系列の判別に加えて、カオス時系列とfBmのような有色ノイズ時系列の判別も可能であることについても合わせて報告する。

2. 解析手法

本稿では、解析対象の時系列が1次元であることを仮定するため、まずは時間遅れ座標への変換によって状態空間を再構成する[26],[27]。再構成状態空間の次元数を m 、時間遅れを τ とし、観測時系列 $x_t (t = 1, \dots, n)$ を式(1)のように再構成状態

空間ベクトルへと変換する.

$$\mathbf{X}_t = (x_t, x_{t+\tau}, \dots, x_{t+(m-1)\tau}) \quad (1)$$

このとき, $t = 1, \dots, N$, $N = n - (m-1)\tau$ となる. 式 (1) より得られる再構成状態空間を用いて, 次の手順で $N \times N$ の 2 次元画像 RP [3],[4] を作成する. 任意の 2 点 $(\mathbf{X}_i, \mathbf{X}_j)$ のユークリッド距離 $D_{i,j}$ を式 (2) で定義する.

$$D_{i,j} = |\mathbf{X}_i - \mathbf{X}_j| \quad (2)$$

RP の第 (i, j) 画素値 $RP_{i,j}$ は, 式 (2) を用いて,

$$RP_{i,j} = \begin{cases} 1 & (\text{if } D_{i,j} \leq \theta_r) \\ 0 & (\text{otherwise}) \end{cases} \quad (3)$$

と与えられる. 式 (3) の θ_r が 2 点間の近傍性を定める閾値である. ここで, RP 上の画素値が 1 の点の組 (i, j) の集合を M_{RP} とする.

次に, IDRP [7] について説明する. RP が 2 点間の近傍関係を可視化する手法であるのに対して, IDRP は 2 点間の軌道ベクトルの差のノルムをもとに同方向性を可視化する手法である. すなわち, 式 (2) を式 (4) に代えることで, $(N-T) \times (N-T)$ の 2 次元画像 IDRP を作成できる.

$$DV_{i,j} = |(\mathbf{X}_{i+T} - \mathbf{X}_i) - (\mathbf{X}_{j+T} - \mathbf{X}_j)| \quad (4)$$

ここで, T は軌道ベクトルの推移時間長である. IDRP の第 (i, j) 画素値 $IDRP_{i,j}$ は, 式 (4) を用いて,

$$IDRP_{i,j} = \begin{cases} 1 & (\text{if } DV_{i,j} \leq \theta_d) \\ 0 & (\text{otherwise}) \end{cases} \quad (5)$$

と与えられる. 式 (5) の θ_d が 2 点間の同方向性を定める閾値である. ここで, IDRP 上の画素値が 1 の点の組 (i, j) の集合を M_{IDRP} とする.

IDRP では任意の 2 点 $\mathbf{X}_i, \mathbf{X}_j$ の時間発展の推移が近ければプロットされるため, それらが互いに近傍であるかどうかは考慮されない. そこで, RP と IDRP を組み合わせることによって, 同方向かつ近傍となる関係を可視化する IDNP [7] を作成することができる. IDNP 上の画素値 1 の点の組 (i, j) の集合 M_{IDNP} は, 集合 M_{RP} と集合 M_{IDRP} の共通集合として式 (6) のように定義される.

$$M_{\text{IDNP}} = M_{\text{RP}} \cap M_{\text{IDRP}} \quad (6)$$

したがって, IDNP の第 (i, j) 画素値 $IDNP_{i,j}$ は,

$$IDNP_{i,j} = RP_{i,j} \times IDRP_{i,j} \quad (7)$$

と定まる. つまり, 任意の 2 点 $\mathbf{X}_i, \mathbf{X}_j$ 間の距離が θ_r 以内であり, かつ, 推移ベクトル間の差のノルムが θ_d 以内の同方向性を有する関係のみが IDNP 上にプロットされる. また, IDNP のサイズは, IDRP と同様に $(N-T) \times (N-T)$ となる.

3. 定量化指標

時系列から得られる RP, IDRP, IDNP を定量化するために,

本稿では 2 つの定量化指標を用いる. まず注意点として, RP の対角線上の画素は必ず 1 になるため, 以下で述べる定量化指標においては考慮しないことが一般的である [6]. また, この対角線は LOI (line of identify) と呼ばれる.

本稿で扱う定量化指標の 1 つは, DET (determinism) [6],[18] と呼ばれる決定論性を評価する指標である. DET は, RP 上にプロットされる点の総数に対して, RP 上に出現する傾き 45 度の斜め線 (LOI に平行な斜め線) を構成する点の割合を定量化した指標である. LOI に平行な斜め線は, $\mathbf{X}_i, \mathbf{X}_j$ が次の時刻でも連続して近傍に留まる場合に現れる. これは, 近傍関係にある 2 点が次の時刻においても近傍関係を保ち続けるという決定論的なシステムの性質を RP 上に反映している. 解析対象時系列が決定論的である場合に DET は高い値を示す傾向にある [6],[18]. 斜め線を定量化するにあたって, RP 上で現れる長さ l の斜め線の本数 $P(l)$ を, 式 (8) で定義する.

$$P(l) = \sum_{i=1}^{N-l+1} \sum_{j=i+1}^{N-l+1} (1 - RP_{i-1,j-1})(1 - RP_{i+l,j+l}) \prod_{c=0}^{l-1} RP_{i+c,j+c} \quad (8)$$

ここで, RP は LOI について対称な画像となり, $RP_{i,j} = RP_{j,i}$ と冗長性が高いため, 式 (8) では $i < j$ の範囲で $P(l)$ を算出している. また, 式 (8) を用いて RP 上の斜め線を数えるためには, $RP_{0,j} = 0$ ($j = 1, \dots, N+1$), $RP_{i,N+1} = 0$ ($i = 1, \dots, N$) と定義して, RP の縁 ($RP_{1,j}$, または, $RP_{i,N}$) に存在するプロットによって構成される斜め線をカウントできるようにする必要がある. 次に, $P(l)$ を用いて DET を式 (9) で算出する.

$$\text{DET} = \frac{\sum_{l \geq l_{\min}}^{N-1} IP(l)}{\sum_{l \geq 1}^{N-1} IP(l)} \quad (9)$$

ここで, l_{\min} は斜め線の最小の長さを表し, 本稿では $l_{\min} = 2$ とする. また, 式 (9) 中の \sum の上限を $N-1$ とするのは LOI を考慮しないためである. DET は RP のみならず, IDRP, IDNP から作成した画像からも算出することが可能である. ただし, IDRP, IDNP から DET を算出する際は, 式 (8), (9) 中の N を $N-T$ とする. 各手法から求められる DET をそれぞれ, DET_{RP} , DET_{IDRP} , DET_{IDNP} とする.

もう 1 つの定量化指標は, 生起確率 R [7] と呼ばれる指標である. R は RP と IDNP を用いて算出される. まず, IDNP は, 式 (6) に示した通り, RP と IDRP の共通集合である. このときの同方向的近傍数 M_{IDNP} の RP の近傍点数 M_{RP} に対する生起確率 R を式 (10) で定義する.

$$R = \frac{|M_{\text{IDNP}}|}{|M_{\text{RP}}|} \quad (10)$$

ここで, $|M_{\text{RP}}|$ 及び $|M_{\text{IDNP}}|$ は, それぞれ集合 M_{RP} 及び M_{IDNP} の要素数を表す. R は, M_{RP} と M_{IDRP} によって決まる定量化指標である. DET と同様に, 解析対象時系列が決定論的である場合に, R は高い値を示す傾向にある [7]. これは, 決定論的なシステムを対象とした場合, 近接する集合は類似した方向に遷移して, 遷移後も近傍関係を保ちやすいという性質が, 式 (10) の分

母と分子の点数の割合として反映されているからである。

4. RP 密度の操作方法

RP と IDRP では閾値 θ_r と θ_d を設定する必要がある。RP の閾値については、アトラクタの 2 点間距離の平均または最大距離の 10% [4]、アトラクタの直径の数 % [8]、総画素数の 1% [9] 程度がプロットされるように設定する方法、システムに応じた設定方法 [10],[11] などが提案されている。しかし、RP の閾値選択は対象とするシステムに強く依存するため、任意の時系列に対して最適な閾値を一意に定めることは難しい [6],[12]。

そこで本稿では、最適な閾値を設定するのではなく、閾値を変化させて RP 密度の異なる様々な画像を 1 つの時系列に対して作成する。

RP 密度の変更方法について、RP を例に説明する。RP では、アトラクタ内の全ての 2 点間距離 $D_{i,j}$ を算出し、 $D_{i,j}$ が閾値 θ_r 以下であれば第 (i, j) 画素に点がプロットされる。本稿では、まず、 $D_{i,j}$ ($i \neq j$) のうちの最短距離 D_{\min} と最長距離 D_{\max} を求める。次に、2 点間距離が存在する範囲 $[D_{\min}, D_{\max}]$ の幅を $D_r = D_{\max} - D_{\min}$ として、

$$\theta_r = D_r \theta + D_{\min} \quad (11)$$

によって θ_r を与える。ここで、 θ は RP 密度を操作するパラメータであり、 $0 \leq \theta \leq 1$ の範囲を持つ。例えば、 $\theta = 0$ の場合は、 $D_{i,j} = D_{\min}$ を満たす RP 上の (i, j) 画素のみに点がプロットされる。反対に $\theta = 1$ の場合は、全ての 2 点間距離が可視化対象となるため、全ての画素が 1 となる。また、IDRP では $DV_{i,j}$ について RP の $D_{i,j}$ と同様に、 $DV_r = DV_{\max} - DV_{\min}$ を求め、 θ を操作して多様な θ_d を表現する。本手法では、このパラメータ θ を変化させることによって、RP 密度が異なる画像を作成する。ただし、本稿では、RP と IDRP の両方に共通の θ を与えて、同じ RP 密度を持つ RP と IDRP から IDNP を作成する。本稿では、 $\theta = 2^{-k}$ ($k = 1, 2, \dots, 10$) とした。

5. 数理モデル

本稿では、決定論的カオス時系列、周期応答にダイナミカルノイズが印加された時系列 (周期 + ノイズ)、非整数ブラウン運動 (fBm) を解析対象の時系列とする。

まず、カオス時系列と周期 + ノイズ時系列の生成に使用した確率 Lorenz モデル (stochastic Lorenz model, SLM) [28] を式 (12) に示す。

$$\begin{cases} \dot{x} = s(-x + y) + \sigma\omega_1 \\ \dot{y} = rx - y - xz + \sigma\omega_2 \\ \dot{z} = -bz + xy + \sigma\omega_3 \end{cases} \quad (12)$$

ここで、 x, y, z は状態変数、 s, r, b はパラメータ、 σ はノイズ強度、 $\omega_i(t)$ ($i = 1, 2, 3$) は独立なガウス増分を持つ標準 Wiener 過程に従う確率変数である。ただし、 $\sigma = 0$ のときは決定論的な Lorenz 方程式 [29] となる。本稿では、 $s = 10, b = 8/3$ として、カオス応答を $r = 28$ 、16 周期応答を $r = 215.45$ により生成し

た。また、周期 + ノイズ時系列を生成するためのノイズ強度 σ は $\sigma = 0.5, 1, 2, 5, 10, 20$ の 6 種類を用いた。これらの数値計算は、ステップ幅が $h = 10^{-4}$ の 4 次のルンゲ・クッタ法により行なった [30]。解析対象とする時系列は、過渡状態として 10^6 ステップを省いた後の 2×10^6 ステップの応答から、500 ステップ毎に 1 点をリサンプリングした 4,000 点の時系列長を持つ。

次に、fBm [23] 時系列 $B_H(t)$ は、増分 $B_H(t) - B_H(s)$ の平均が 0、増分の分散が $\sigma_f^2 |t - s|^{2H}$ と時間の冪乗で変化する確率過程より生成した。本稿では、分散 $\sigma_f^2 = 1$ 、ハースト指数 H を $H = 0.2, 0.5, 0.8$ の 3 種類を用いて fBm の時系列を生成した。ただし、 $H = 0.5$ のときの $B_H(t)$ はブラウン運動となる。fBm 生成のアルゴリズムは、ランダム中点変位法を修正した方法 [31] とした。また、解析対象とする時系列の長さは 4,000 点とした。

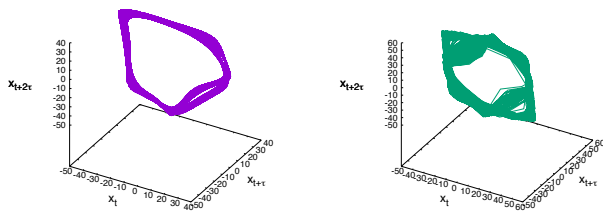
6. 数値実験

数値実験では、数理モデルに依存しない共通の実験条件として、1 つのシステムに対しパラメータを固定し、初期値と乱数シードを変えた 100 本の時系列を生成した。したがって、100 本のそれぞれの時系列に提案手法を適用して得られた、平均の定量化指標 (DET と R) の推移傾向を結果として示す。また、IDRP 適用時に設定する推移時間長 T は時間遅れ τ と等しくした ($T = \tau$)。

カオス時系列と周期 + ノイズ時系列は、それぞれのパラメータ r 、ノイズ強度 σ を与えた SLM から生成した第一変数 x の時系列を対象とした。再構成状態空間の次元数 m は $m = 7$ とした。時間遅れ τ は、解析対象時系列の自己相関関数が最初に $1 - 1/e$ となる時刻 C_i の値とした ($\tau = C_i$) [32]。また、非整数ブラウン運動 (fBm) から生成した時系列は、 $m = 3, \tau = C_i$ として状態空間を再構成した。

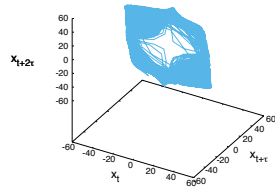
まず、パラメータ r 、ノイズ強度 σ を与えたときの SLM から生成される第一変数 x の時系列を、 $m = 7, \tau = C_i$ で再構成した状態空間の一例を図 1 に示す。図 1(a) はノイズ強度が $\sigma = 0$ であるため、決定論的な Lorenz 方程式から得られる 16 周期応答の時系列 x に対する再構成アトラクタを示している。一方、図 1(h) も $\sigma = 0$ の決定論的な応答であるが、パラメータは $r = 28$ なのでその応答はカオス応答となる。図 1 において注目すべき点は、 $\sigma > 0$ の場合に、周期応答の構造は崩れ (図 1(b)~(g))、カオス応答の再構成アトラクタの構造と類似の構造となることである (図 1(h))。このような特徴があるために、カオス時系列と周期応答にダイナミカルノイズが印加された時系列 (周期 + ノイズ) の判別を考える必要がある。

次に、カオス時系列 (SLM の $r = 28, \sigma = 0$)、周期 + ノイズ時系列 (SLM の $r = 215.45, \sigma > 0$)、fBm 時系列 ($H = 0.2, 0.5, 0.8$) それぞれに対して RP 密度を変更して得られる、RP, IDR, IDNP の定量化指標 (DET と R) の推移傾向を図 2 に示す。図 2 では、横軸を RP 密度を操作するパラメータ θ として片対数グラフで表している。縦軸は定量化指標を表しており、図 2(a) は RP より算出された DET_{RP} 、図 2(b) は IDRP より算出された DET_{IDRP} 、図 2(c) は IDNP より算出された DET_{IDNP} 、図 2(d) は生起確率 R である。図 2(a)~(d) において、凡例は共通で

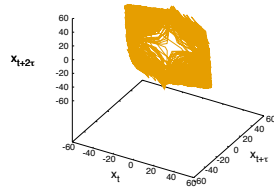


(a) $r = 215.45, \sigma = 0, \tau = 2$

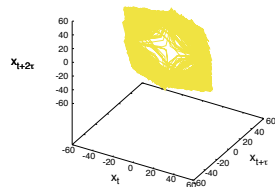
(b) $r = 215.45, \sigma = 0.5, \tau = 2$



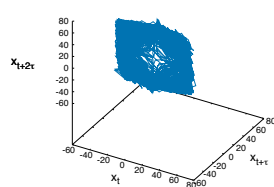
(c) $r = 215.45, \sigma = 1, \tau = 2$



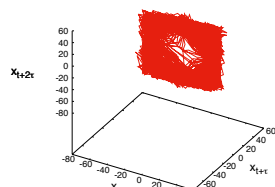
(d) $r = 215.45, \sigma = 2, \tau = 2$



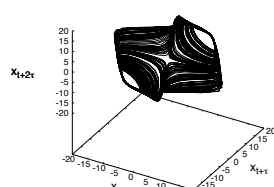
(e) $r = 215.45, \sigma = 5, \tau = 2$



(f) $r = 215.45, \sigma = 10, \tau = 2$



(g) $r = 215.45, \sigma = 20, \tau = 2$

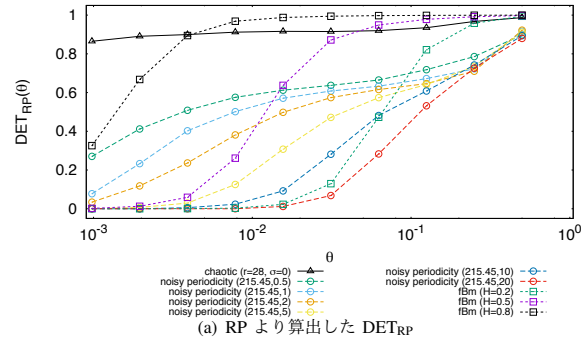


(h) $r = 28, \sigma = 0, \tau = 4$

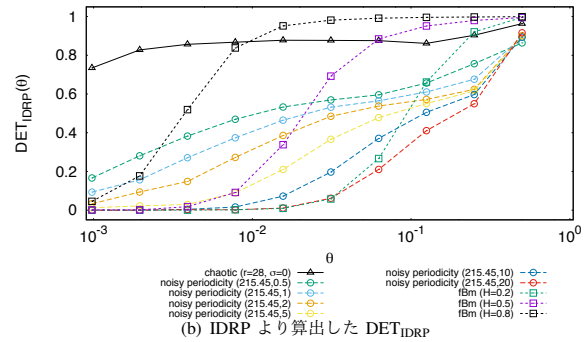
図1: パラメータ r とノイズ強度 σ を与えたときの SLM から生成される第一変数 x の時系列の再構成状態空間の一例。状態空間の再構成に使用するパラメータは、 $m = 7, \tau = C_i$ 。ただし、 C_i はそれぞれの時系列の自己相関関数が最初に $1 - 1/e$ になったときの時刻を示す。

あり、黒色の実線で結ばれた Δ 印は、カオス時系列 (SLM の $r = 28, \sigma = 0$)、破線で結ばれた \circ 印は、周期 + ノイズ時系列 (SLM の $r = 215.45, \sigma > 0$)、破線で結ばれた \square 印は、fBm 時系列 ($H = 0.2, 0.5, 0.8$) をそれぞれ示している。色の違いは、ノイズ強度 σ 、または、ハースト指数 H の違いを示している。

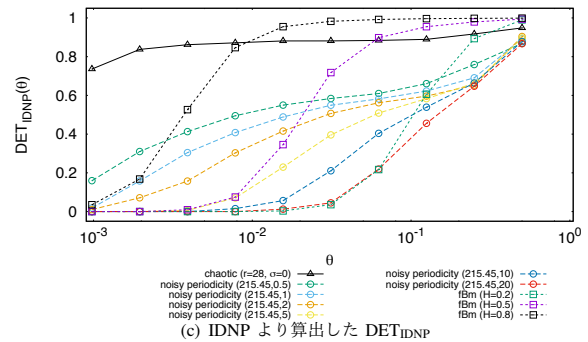
本稿において判別したい時系列の組は、(I) 決定論的カオス



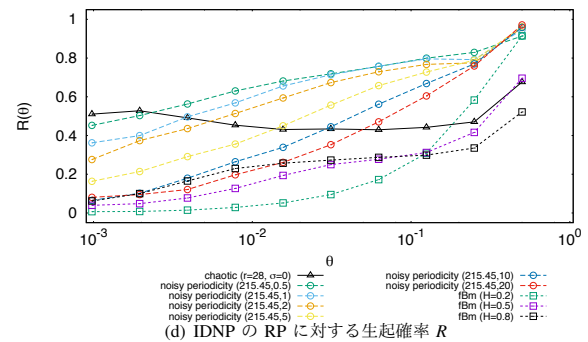
(a) RP より算出した DET_{RP}



(b) IDRP より算出した DET_{IDRP}



(c) IDNP より算出した DET_{IDNP}



(d) IDNP の RP に対する生起確率 R

図2: 提案手法をシステムが異なる時系列に適用した結果。横軸は RP 密度を操作するパラメータ θ 、縦軸は定量化指標であり、(a) は RP より算出された DET_{RP} 、(b) は IDRP より算出された DET_{IDRP} 、(c) は IDNP より算出された DET_{IDNP} 、(d) は生起確率 R をそれぞれ示している。黒色の実線で結ばれた Δ 印は、カオス時系列 (SLM の $r = 28, \sigma = 0$)、破線で結ばれた \circ 印は、周期 + ノイズ時系列 (SLM の $r = 215.45, \sigma > 0$)、破線で結ばれた \square 印は、fBm 時系列 ($H = 0.2, 0.5, 0.8$) をそれぞれ示している。線の色はノイズ強度 σ 、または、ハースト指数 H の違いを示している。

時系列と周期応答にダイナミカルノイズが印加された時系列と (II) 決定論的カオス時系列と fBm のような有色ノイズ時系列である。まず、図 2(a)~(c) で示した、各 θ の RP, IDRP, IDNP から算出された DET の推移傾向について、(I) カオス (黒色の実線で結んだ Δ 印) と周期 + ノイズ (破線で結んだ \circ 印) それぞれの結果を比較する。カオスの DET_{RP} , DET_{IDRP} , DET_{IDNP} の推移傾向は、いずれの θ においても高い値を保ち、また、 θ の変化による DET の増減は小さいことが分かる。一方で、周期 + ノイズの DET の推移傾向は、 θ が小さいときは DET も小さく、 θ が大きくなるにつれて、DET も増加していくことが分かる。これらの推移傾向の違いから、(I) カオス時系列と周期 + ノイズ時系列の判別は可能であると考えられる。

次に、(II) カオス (黒色の実線で結んだ Δ 印) と fBm (破線で結んだ \square 印) の判別は、図 2(d) の生起確率 R の推移傾向の違いより、可能であると考えられる。カオスであれば θ が小さい場合でも R は 0.5 付近となり、決定論性が認められる。一方で、fBm であればいずれのハースト指数 H においても、RP 密度が 10% 程度までは R は 0.3 程度にしか増加しない。したがって R の推移傾向によって、(II) カオス時系列と fBm 時系列の判別は可能であると考えられる。しかしながら、(I) の判別において有効であった図 2(a)~(c) の DET の推移傾向による (II) の判別は難しいと考えられる。特にハースト指数が $H = 0.8$ と大きい場合 (黒色の破線で結んだ \square 印)、決定論的カオス時系列よりも高い DET となる。これらの結果は、これまで使用されてきた RP 密度 [4], [8]~[11] では決定論性を誤検出する危険性があること、提案手法を用いることでこの危険性を回避できることを示すものである。

7. ま と め

本稿では、我々が提案しているリカレンスプロット密度の変化による時系列の特性抽出手法 [14]~[17] を用いて、周期応答にダイナミカルノイズが印加された時系列と、非整数ブラウン運動のような有色ノイズ時系列 [23] が、決定論的カオス時系列と判別可能であるかを調査した。提案手法では、まず、1次元の時系列を時間遅れ座標を用いて状態空間を再構成した。次に、一つの再構成状態空間について、リカレンスプロット (RP) [3], [4], 同方向性リカレンスプロット (IDRP) [7], 同方向的近傍プロット (IDNP) [7] の RP 密度を変化させて適用し、複数の画像を作成した。RP 密度が低い画像から高い画像へと変化していく際の DET [6], [18] と生起確率 R [7] を算出し、それらの推移傾向から時系列の背後に潜むシステムの判別を行った。数値計算の結果から、DET の推移傾向より、決定論的カオス時系列と周期ノイズ時系列の判別が可能となることが明らかになった。加えて、決定論的カオス時系列と非整数ブラウン運動の時系列の判別は、 R の推移傾向から可能となることが示された。一方で、これまでに扱われてきた閾値設定方法 [4], [8]~[11] により定まるような RP 密度においては、決定論的カオスと非整数ブラウン運動は判別を誤る危険性があること、従って、本手法のように RP 密度が低い範囲から大きい範囲まで変化させる手法が必要であることが示された。

なお、本研究の一部は JSPS 科研費 (No. 19K14589, 20H00596, 17K00348, 18K18125) の援助を受けて行われた。

文 献

- [1] 赤池 弘次, 中川 東一郎, サイエンス社, 1972.
- [2] 池口 徹, 山田 泰司, 小室 元政, 合原 一幸 編, 産業図書, 2000.
- [3] J.-P. Eckmann, S. Oliffson Kamphorst, and D. Ruelle, *Europhysics Letters*, Vol. 4, pp. 937–977, 1987.
- [4] J. P. Zbilut and C. L. Webber Jr., *Physics Letters A*, Vol. 171, Nos. 3–4, pp. 199–203, 1992.
- [5] M. Thiel, M. C. Romano, and J. Kurths, *Physics Letters A*, Vol. 330, No. 5, pp. 343–349, 2004.
- [6] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, *Physics Reports*, Vol. 438, Nos. 5–6, pp. 237–329, 2007.
- [7] 實来 俊介, 山田 泰司, 合原 一幸, *電気学会論文誌. C*, Vol. 122, No. 1, pp. 141–147, 2002.
- [8] G. B. Mindlin and R. Gilmore, *Physica D: Nonlinear Phenomena*, Vol. 58, Nos. 1–4, pp. 229–242, 1992.
- [9] J. P. Zbilut, J.-M. Zaldívar-Comenges, and F. Strozzi, *Physics Letters A*, Vol. 297, Nos. 3–4, pp. 173–181, 2002.
- [10] M. Thiel, M. C. Romano, J. Kurths, R. Meucci, E. Allaria, and F. T. Arcechi, *Physica D: Nonlinear Phenomena*, Vol. 171, No. 3, pp. 138–152, 2002.
- [11] L. Matassini, H. Kantz, J. Holyst, and R. Hegger, *Physical Review E*, Vol. 65, No. 2, 021102, 2002.
- [12] N. Marwan, *International Journal of Bifurcation and Chaos*, Vol. 21, No. 4, pp. 1003–1017, 2011.
- [13] G. McGuire, N. B. Azar, and M. Shelhamer, *Physics Letters A*, Vol. 237, Nos. 1–2, pp. 43–47, 1997.
- [14] 小川 徳紀, 金丸 志生, 山本 紘平, 島田 裕, 池口 徹, 電子情報通信学会 2019 年総大会講演論文集, N-1-33, 2019.
- [15] 金丸 志生, 島田 裕, 池口 徹, 電子情報通信学会 技術研究報告, Vol. 119, No. 19, NLP2019-5, pp. 23–28, 2019.
- [16] 金丸 志生, 島田 裕, 池口 徹, 2019 年電子情報通信学会ソサイエティ大会, N-1-13, 2019.
- [17] S. Kanamaru, Y. Shimada, and T. Ikeguchi, *Proceedings of the 2019 International Symposium on Nonlinear Theory and its Applications (NOLTA2019)*, pp. 536–539, 2019.
- [18] L. L. Trulla, A. Giuliani, J.P. Zbilut, and C.L. Webber Jr., *Physics Letters A*, Vol. 223, No. 4, pp. 255–260, 1996.
- [19] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, *Physical Review E*, Vol. 66, No. 2, 026702, 2002.
- [20] W. M. Schaffer and M. Kot, *Journal of Theoretical Biology*, Vol. 112, No. 2, pp. 403–427, 1985.
- [21] L. F. Olsen and W. M. Schaffer, *Science*, Vol. 249, No. 4968, pp. 499–504, 1990.
- [22] G. Sugihara and R. M. May, *Nature*, Vol. 344, No. 6268, pp. 734–741, 1990.
- [23] B. B. Mandelbrot and J. W. Van Ness, *Society for Industrial and Applied Mathematics Review*, Vol. 10, No. 4, pp. 422–437, 1968.
- [24] B. Cazelles and R. H. Ferriere, *Nature*, Vol. 355, No. 6355, pp. 25–26, 1992.
- [25] A. A. Tsonis and J. B. Elsner, *Nature*, Vol. 358, No. 6383, pp. 217–220, 1992.
- [26] F. Takens, *Lecture Notes in Mathematics*, Vol. 898, pp. 366–381, 1981.
- [27] T. Sauer, J. A. Yorke, and M. Casdagli, *Journal of Statistical Physics*, Vol. 65, Nos. 3–4, pp. 579–616, 1991.
- [28] I. Bashkirtseva, L. Ryashko and P. Stikhin, *International Journal of Bifurcation and Chaos*, Vol. 23, No. 5, 1350092, 2013.
- [29] E. N. Lorenz, *Journal of the Atmospheric Sciences*, Vol. 20, No. 2, pp. 130–141, 1963.
- [30] I. Bashkirtseva and L. Ryashko, *Physical Review E*, Vol. 79, No. 4, 041106, 2009.
- [31] 成田 清正, “例題で学べる 確率モデル”, 共立出版, 2010.
- [32] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, *Physica D: Nonlinear Phenomena*, Vol. 65, Nos. 1–2, pp. 117–134, 1993.

発火間隔時系列を用いたニューロンへの共通入力の再構成

三浦 英[†] 池口 徹[†]

[†] 東京理科大学 工学部 情報工学科
〒125-8585 東京都葛飾区新宿 6-3-1
E-mail: †ei.miura@hisenkei.net

あらまし ニューロンの出力信号の観測に比べて、ニューロンへの入力信号の観測は容易ではない。そのため、ニューロンの出力信号のみを用いて、ニューロンへの入力信号を再構成する手法が提案されている。本稿では、複数のニューロンより出力されたスパイク列から得られる発火間隔時系列に対して、非線形時系列解析手法の一つであるリカレンスプロットを用いたニューロンへの共通入力の再構成手法を提案している。その際、ニューロンの数理モデルとして、Leaky Integrated and Fire モデルを用いて提案手法の性能評価を行った。結果として、発火間隔時系列を用いた複数のニューロンへの共通入力の再構成を行う提案手法が有効であることが示された。

キーワード LIF モデル, リカレンスプロット, ニューロン

Reconstruction of Input Signal Using Common Interspike Interval Time Series

Ei MIURA[†] and Tohru IKEGUCHI[†]

[†] Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science
E-mail: †ei.miura@hisenkei.net

Abstract It is not easy to observe the input signals of neurons compared to the output signals of neurons. For this reason, several methods have been proposed to reconstruct the input signals of neurons using only the output signal of the corresponding neurons. In this paper, we propose a method for reconstructing common input signals of neurons using recurrence plots, a nonlinear time series analysis method, using inter spike interval time series observed from output spike trains from multiple neurons. We use the leaky integrated-and-fire model as the mathematical neuron model to evaluate the performance of the proposed method. As a result, we show that the proposed method using inter spike interval time series is effective to reconstruct the common input signals of multiple neurons.

Key words Leaky Integrated-and-Fire model, Recurrence plot, Neuron

1. はじめに

人間の脳内には、ニューロンと呼ばれる細胞が多数存在している。ニューロンは、ある閾値を超える入力を与えられると、スパイクを出力することで、断続的に情報を伝達する。一般的に、ニューロンの出力信号を観測することに比べて、ニューロンへの入力信号を観測することは容易ではないとされている。そのため、ニューロンの出力信号から入力信号を再構成する手法が提案されている。具体的には、ニューロンの出力信号であるスパイク列から、時間窓に対する発火数の割合を算出することで得られる発火率時系列を用いて、個体差のある複数のニューロンへの共通入力を再構成する手法である [1], [2]。この従来手法を用いることで、共通入力を再構成できることが示されている。しかし、従来手法では、発火率時系列を用いているため、ニューロンの発火の詳細な情報が欠落している可能性がある。

そこで本稿では、複数のニューロンより出力されたスパイク列から得られる直接的な情報である発火間隔時系列を用いることで、共通入力を再構成する手法を提案する。従来手法では、複数のリカレンスプロット (以下, RP) [3] の和集合を取ることで作成する重畳リカレンスプロット (以下, 重畳 RP) [1], [2] を用いている。重畳 RP を作成する際には、一辺の長さが異なる RP を重畳することが困難であるため、全ての RP の一辺の長さを一定にする必要がある。ここで、RP の一辺の長さは、ニューロンの発火数に比例する。そこで本稿では、発火数が一定となる 2 つの設定を導入することで、共通入力の再構成を行った。1 つ目の設定は、Leaky Integrated-and-Fire モデル (以下, LIF モデル) [4] の漏れ τ に対し、発火数が一定となるようにバイアス項 B_{τ} を調整した場合である。2 つ目の設定は、発火数が一定となる範囲で、LIF モデルの漏れ τ を微小に変化させた場合である。これらの 2 つの設定を導入することで調査を行った結

果、複数のニューロンの個体差として、LIF モデルの漏れのパラメータ τ を用いた場合でも、発火数が一定となるような条件で共通入力のリ構成を行えることが示唆された。

2. ニューロンモデルと発火間隔時系列

ニューロンは閾値 Θ を超える入力を与えられると発火し、スパイク列を出力する性質を持っている。LIF モデルは、これらの性質を模倣する単純なニューロンモデルであり、式 (1) で定義される。

$$\tau \frac{dV(t)}{dt} = -V(t) + S(t) \quad (1)$$

ここで、 $V(t)$ は膜電位、 $S(t)$ は入力電流を表す。なお、膜電位 $V(t)$ は閾値 Θ に達したとき式 (2) のようにリセットされる。

$$V(t) = 0 \quad (V(t) \geq \Theta) \quad (2)$$

また、 $1/\tau$ は漏れの割合を示すパラメータであり、 $1/\tau \rightarrow 0$ のとき、通常の積分発火モデルとなり、 $1/\tau$ が大きくなると、漏れの割合が大きくなる。

ニューロンは、入力信号により膜電位 $V(t)$ が上昇することで発火し、直後に $V(t)$ がリセットされることを繰り返す。 i 番目の発火時刻を $T(i)$ とすると、発火間隔 $t(i)$ は、式 (3) で求めることができる。

$$t(i) = T(i+1) - T(i) \quad (3)$$

式 (3) により、断続的に発火間隔を求めることで、発火間隔 (Inter spike intervals, ISI) 時系列が得られる。

3. 提案手法

本稿では、複数のニューロンより出力されたスパイク列から得られる発火間隔時系列を用いることで、共通入力を再構成する手法を提案する。提案手法の流れを図 1 に示す。ここでは、提案手法の流れについて説明する。

(1) 手順 1 では、LIF モデルの漏れ τ を個体差とした n 個のニューロンを用意し、共通入力にバイアス項を加えた時系列を入力することで、 n 本の発火間隔時系列を得る。

(2) 手順 2 では、 n 本の各発火間隔時系列に対して、アトラクタの再構成 [5]~[7] を行い、 n 個の RP を作成する。

(3) 手順 3 では、各ニューロンへの共通入力の情報を取り出すために、 n 個の RP から和集合を取ることで、重畳 RP を作成し、共通入力を再構成する [8]。

なお、各手順における具体的な手法については、以降の章で説明する。

4. 時系列解析手法

4.1 時間遅れ座標系への変換によるアトラクタの再構成

現実存在する m 次元のシステムから、複数の状態変数を同時に観測できる場合は稀である。また、多くの場合では元のシステムの次元数が不明である。このような場合、観測できた状態変数から元のシステムのアトラクタの再構成を行うことで、

アトラクタの特徴を再現できる。本稿では、アトラクタの再構成を行う手法として、時間遅れ座標系を用いた [5]~[7]。時間遅れ座標系への変換は、1つの状態変数から元の力学系の m 次元のシステムのアトラクタの軌道を再構成する際に用いられる手法である。具体的には、1変数の時系列信号 $y(t)$ から m 次元の再構成状態空間へ変換する場合、時間遅れ L を用いて式 (4) のように m 次元ベクトル $\mathbf{v}(t)$ を構成する。

$$\mathbf{v}(t) = (y(t), y(t+L), \dots, y(t+(m-1)L)) \quad (4)$$

また、時間遅れ座標系への変換によるアトラクタの再構成手法は、スパイク列に対しても拡張されている。発火間隔時系列 $t(i)$ を用いて、式 (5) のように m 次元ベクトル $\mathbf{u}(i)$ を構成することで、元の力学系のアトラクタを再構成できる [7]。

$$\mathbf{u}(i) = (t(i), t(i+1), \dots, t(i+(m-1))) \quad (5)$$

4.2 リカレンスプロット

RP は力学系のアトラクタの状態空間内における 2 点間の距離情報を可視化する手法である [3]。具体的には、アトラクタ上の点の数が N であるとき、一辺の長さが N となる二次元画像を用意する。式 (6) で $\mathbf{R}(i, j) = 1$ となるとき、第 (i, j) 画素をプロットすることで RP は作成される。

$$\mathbf{R}(i, j) = \begin{cases} 1 & (|\mathbf{u}(i) - \mathbf{u}(j)| < \theta_r) \\ 0 & (\text{それ以外}) \end{cases} \quad (6)$$

ここで、 θ_r は RP の閾値であり、 $\mathbf{u}(i)$ 、 $\mathbf{u}(j)$ は式 (5) により求めたアトラクタの状態空間内の 2 点を表す。また、 $|\mathbf{u}(i) - \mathbf{u}(j)|$ は $\mathbf{u}(i)$ と $\mathbf{u}(j)$ のユークリッド距離を表す。

4.3 重畳リカレンスプロット

重畳 RP は、2 つ以上の RP の成分の和集合をとる手法である。具体的には、 k 個の時系列から作成した k 個の RP を \mathbf{R}_s ($s = 1, 2, \dots, k$) とすると、式 (7) により、 \mathbf{R}_s から重畳 RP、 $\mathbf{R}_+(i, j)$ を作成できる。

$$\mathbf{R}_+(i, j) = \begin{cases} 1 & (\sum_{s=1}^k \mathbf{R}_s(i, j) \geq \theta_+) \\ 0 & (\text{それ以外}) \end{cases} \quad (7)$$

ここで、 θ_+ は重畳 RP の閾値 θ_+ である。式 (7) は RP 上の $\mathbf{R}_s(i, j) = 1$ となる点が θ_+ となるときに、重畳 RP 上の第 (i, j) 画素をプロットすることを意味している。なお、複数のニューロンから得られた出力信号の重畳 RP を求めることで、ニューロンへの共通の入力の情報を取り出すことができると考えられる [1], [2]。

4.4 RP からの時系列の再構成

RP の作成に用いた時系列の特徴が RP 上のプロットパターンとして可視化される。すなわち、RP から元の時系列を再構成することが可能である [8]。文献 [8] では、以下の 4 つの手順で時系列の再構成を行う。

(1) RP を隣接行列とみなし、RP 上の点 $\mathbf{R}(i, j) = 1$ である頂点 i と頂点 j の間に枝を張ることで無向グラフを作成する。

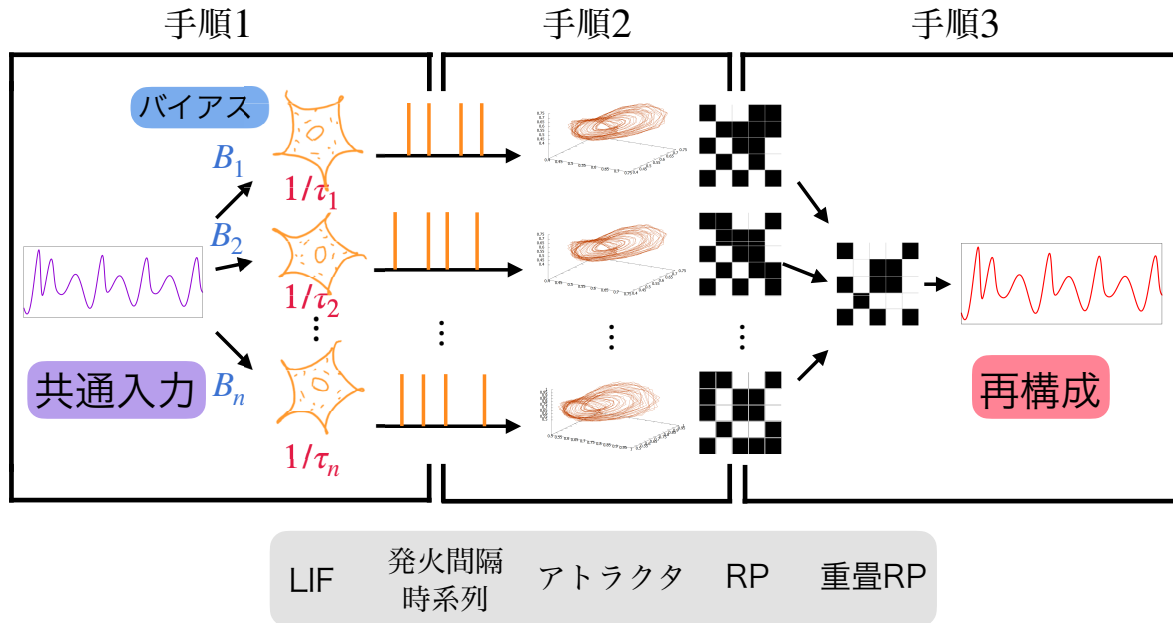


図 1: 提案手法の流れ

(2) 頂点 i と隣接する頂点の集合 G_i を $G_i = \{j | R(i, j) = 1\}$ とし, 頂点 i と頂点 j の枝の重み $w_{i,j}$ を式 (8) により定義する.

$$w_{i,j} = 1 - \frac{|G_i \cap G_j|}{|G_i \cup G_j|} \quad (8)$$

ここで, $|G_i|$ は集合 G_i の要素数, \cap は積集合, \cup は和集合である.

(3) 枝の重み $w_{i,j}$ で重み付けされた無向グラフの全ての頂点 i と頂点 j の間の最短距離 $l_{i,j}$ を計算し, 距離行列 $L = [l_{i,j}]$ を作成する.

(4) 距離行列 L に古典的多次元尺度法 [9] を適用し, 時系列を再構成する.

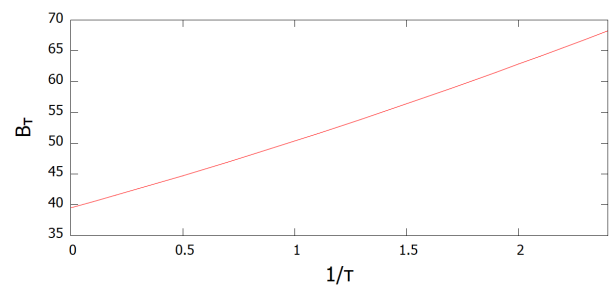
5. 数値実験

本稿では, 漏れのパラメータ τ を個体差とした LIF モデルから得られる発火間隔時系列を用いて, 複数のニューロンへの共通入力の再構成を行った. 具体的には, まず, 共通入力としたレスラー方程式 [10] の第 1 変数 $x(t)$ の時系列にバイアス項 B を加えた入力時系列 $S(t) = x(t) + B$ を作成し, LIF モデルの漏れ τ を個体差とした n 個のニューロンへ入力することで, n 本の発火間隔時系列を得る. ここで, レスラー方程式のパラメータは $a = 0.36, b = 0.4, c = 4.5$, LIF モデルの閾値 $\Theta = 20$ とした. 次に, 時間遅れ座標系への変換を用いて, n 本の各発火間隔時系列に対して, アトラクタの再構成を行い, n 個の RP を作成する. ここで, 次元数 $m = 5$, 遅れ時間 1 としてアトラクタの再構成を行った. また, RP の閾値 θ_r は RP のプロット割合が 10% となるように設定した. 最後に, 共通入力の情報を取り出すために, n 個の RP から和集合を取ることで, 重畳 RP を作成し, 共通入力を再構成する. ここで, 重畳 RP の閾値 $\theta_+ = n/2$ とした.

なお, LIF モデルの漏れ τ を変化させると発火数が変化する

ため, アトラクタのプロット数が変化し, RP の重畳が困難になる場合がある. そこで本稿では, 発火数が一定となる 2 つの設定を導入することで数値実験を行った.

1 つ目の設定では, LIF モデルの漏れ τ に対し, 発火数が一定となるようにバイアス項 B_τ を調整する. 具体的には, 漏れ $1/\tau = 0.0, 0.1, \dots, 2.4$ とした 25 個の LIF モデルへ共通入力したレスラー方程式の第 1 変数 $x(t)$ にバイアス項 B_τ を加えた時系列 $S_\tau(t) = x(t) + B_\tau$ を入力し, 得られる発火間隔時系列を用いて共通入力の再構成を行った. ここで, バイアス項 B_τ は, 漏れ τ として LIF モデルの発火数が 500 となるように設定した. 漏れ τ に対して設定したバイアス項 B_τ を図 2 に示す.

図 2: 漏れ τ に対し, 発火数が一定となるように設定したバイアス項 B_τ

2 つ目の設定では, 発火数が一定となる範囲で, LIF モデルの漏れ τ を微小に変化させる. 具体的には, LIF モデルの発火数が一定となる漏れ τ の範囲 $0.996 \leq 1/\tau \leq 1.00188$ で変化させた 50 個の LIF モデルへ共通入力したレスラー方程式の第 1 変数の時系列 $x(t)$ に, バイアス項 $B = 50.35$ を加えた入力時系列 $S(t) = x(t) + B$ を入力し, 得られる発火間隔時系列を用いて共通入力の再構成を行った. ここで, バイアス項 $B = 50.35$ は, 漏れ τ の範囲 $0.996 \leq 1/\tau \leq 1.00188$ における LIF モデル

の発火数が 500 となるように設定した値である。

なお、共通入力の再構成ができていないかを確認するため、共通入力と再構成時系列の標準化を行った 2 つの時系列の重ね書きを行い、共通入力と再構成時系列標準化を行った 2 つの時系列から抽出した極値の時系列間の相関係数を算出した。さらに、共通入力としたレスラー方程式の第 1 変数 $x(t)$ の時系列から RP を作成し、作成した RP と重畳 RP のプロットパターンを比較した。このとき、レスラー方程式の第 1 変数 $x(t)$ の時系列の 1/20 でサブサンプルした時系列から RP を作成した。

6. 実験結果

共通入力としたレスラー方程式の第 1 変数 $x(t)$ の時系列の RP と提案手法により作成した重畳 RP を図 3, 図 4 に示す。また、これらの RP から時系列を再構成することで得た共通入力時系列と再構成時系列を重ね書きした結果を図 5, 図 6 に示す。ここで、共通入力を紫色の時系列で、再構成時系列を赤色の時系列で示した。

図 3 及び図 4 から、発火数が一定となる 2 つの設定のどちらの場合でも、共通入力としたレスラー方程式の第 1 変数 $x(t)$ の時系列の RP と重畳 RP のプロットパターンが類似することがわかる。また、図 5 及び図 6 から、発火数が一定となる 2 つの設定のどちらの場合でも、共通入力としたレスラー方程式の第 1 変数 $x(t)$ の時系列と再構成時系列の概形が類似することもわかる。さらに、共通入力と再構成時系列の標準化を行った後、極値を抽出することで得た 2 つの時系列間の相関図を図 7 及び図 8 に示す。これらの相関係数は、発火数が一定となる 2 つの設定のどちらの場合でも 0.98 となった。以上の結果から、発火数が一定となる設定下においては、高精度に共通入力の再構成を行うことができることがわかる。

7. まとめ

本稿では、ニューロンより出力されたスパイク列から得られる発火間隔時系列に対して、非線形時系列解析手法の一つである RP を用いた共通入力の再構成手法を提案し、性能評価を行った。その結果、複数のニューロンの個体差として、LIF モデルの漏れのパラメータ τ を用いた場合、発火数が一定となる設定を導入することで共通入力の再構成を精度よく行えることが示された。

本稿では、発火数が一定となる設定で共通入力の再構成を行ったが、発火数が異なる設定での共通入力の再構成を行うことについての検討も必要である。また、ニューロンの数理モデルとして、単純な LIF モデルを用いたが、より複雑なニューロンモデルを用いた場合に、発火間隔時系列を用いた共通入力の再構成について検討する必要がある。

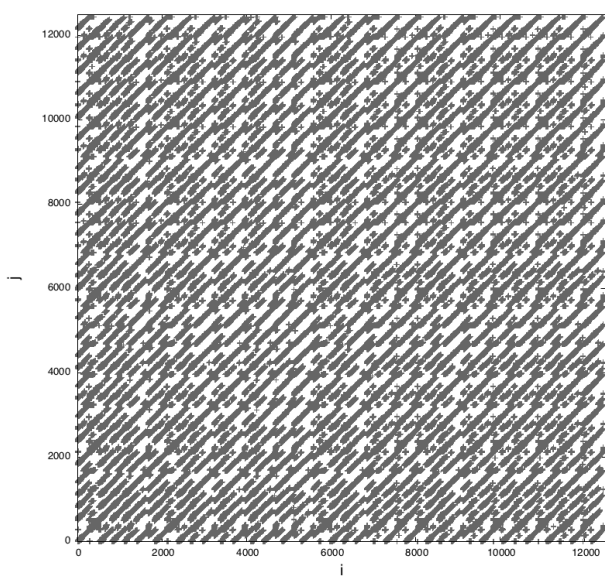
本研究の一部は、JSPS 科研費 (No. 17K00348, 20H00596) の援助を受けて行われた。

文 献

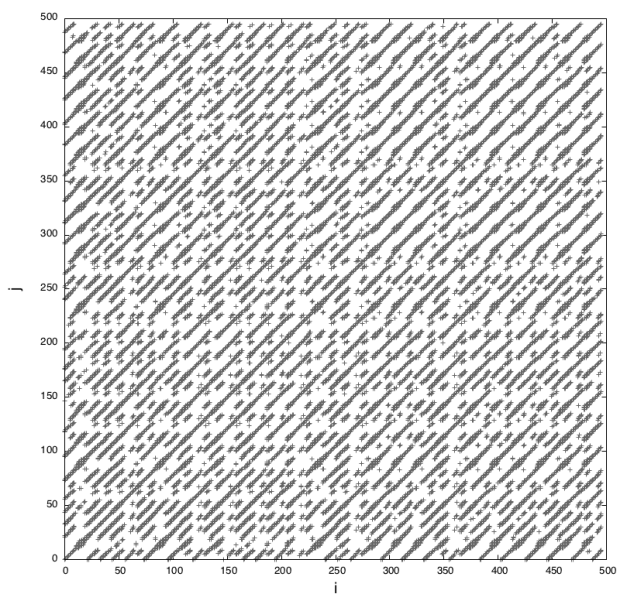
- [1] 野村亮太, 池口徹, “重畳リカレンスプロットを用いた共通入力再構成,” 電子情報通信学会 2019 年総合大会, N-1-24, 2019.
- [2] 野村亮太, 池口徹, “重畳リカレンスプロットを用いた共通入力

の再構成に関する一検討,” 電子情報通信学会技術通信研究報告, Vol. 119, No. 19, pp. 29–34, 2019.

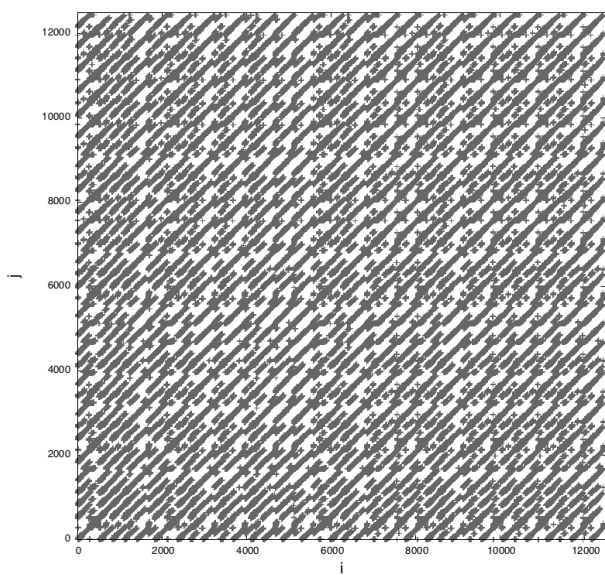
- [3] J.-P. Eckmann, S. Oliffson Kamphorst and D. Ruelle, “Recurrence plots of dynamical systems,” *Europhysics Letters*, Vol. 4, No. 9, pp. 973–977, 1987.
- [4] R. B. Stein, “Some models of neuronal variability,” *Biophysical Journal*, Vol. 7, No. 1, pp. 37–68, 1967.
- [5] Norman H. Packard, James P. Crutchfield, J. D. Farmer, and Robert S. Shaw, “Geometry from a time series,” *Physical Review Letters*, Vol. 45, No. 9, pp. 712, 1980.
- [6] F. Takens, “Detecting strange attractors in turbulence,” *Lecture Notes in Mathematics*, Vol. 898, pp. 366–381, 1981.
- [7] T. Sauer, “Reconstruction of Dynamical Systems from Interspike Intervals,” *Physical Review Letters*, Vol. 72, No. 24, pp. 3811–3814, 1994.
- [8] Y. Hirata, S. Horai, and K. Aihara, “Reproduction of distance matrices and original time series from recurrence plots and their applications,” *The European Physical Journal Special Topics*, Vol. 164, No. 1, pp. 13–22, 2008.
- [9] J. C. Gower, “Some distance properties of latent root and vector methods used in multivariate analysis,” *Biometrika*, Vol. 53, No. 3–4, pp. 325–338, 1966.
- [10] O. E. Rössler, “An equation for continuous chaos,” *Physics Letters A*, Vol. 57, No. 5, pp. 397–398, 1976.



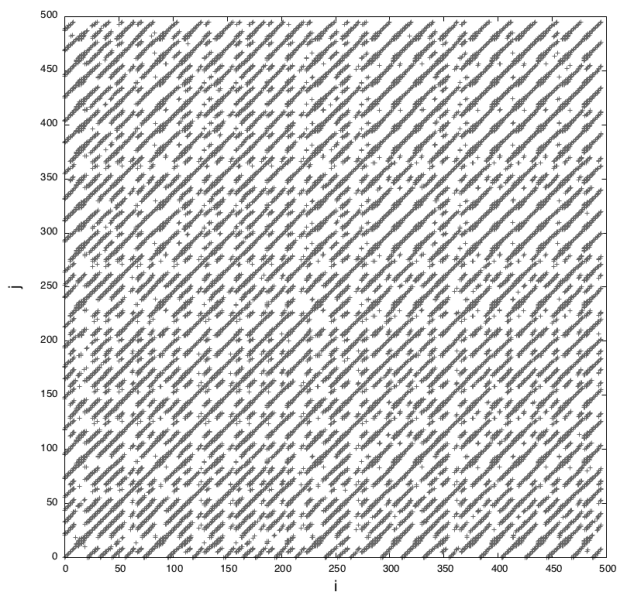
(a) 共通入力（レスラー方程式の第1変数）の RP



(b) 重畳 RP

図 3: LIF モデルの漏れ τ に対し、発火数が一定となるようにバイアス項 B_τ を調整した場合の RP の比較

(a) 共通入力（レスラー方程式の第1変数）の RP



(b) 重畳 RP

図 4: 発火数が一定となる範囲で、LIF モデルの漏れ τ を微小に変化させた場合の RP の比較

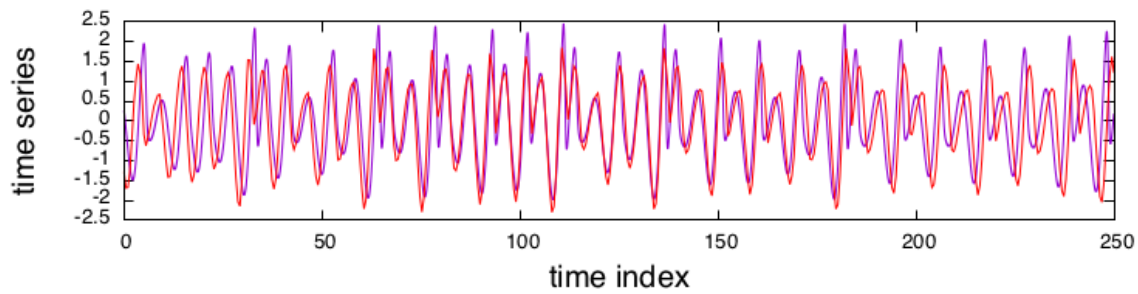


図 5: LIF モデルの漏れ τ に対し、発火数が一定となるようにバイアス項 B_τ を調整した場合の共通入力としたレスラー方程式の第 1 変数 $x(t)$ と再構成時系列の重ね書き

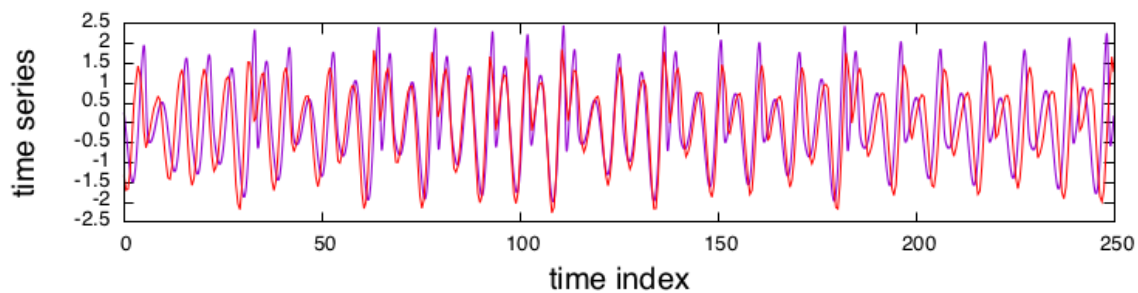


図 6: 発火数が一定となる範囲で、LIF モデルの漏れ τ を微小に変化させた場合の共通入力としたレスラー方程式の第 1 変数 $x(t)$ と再構成時系列の重ね書き

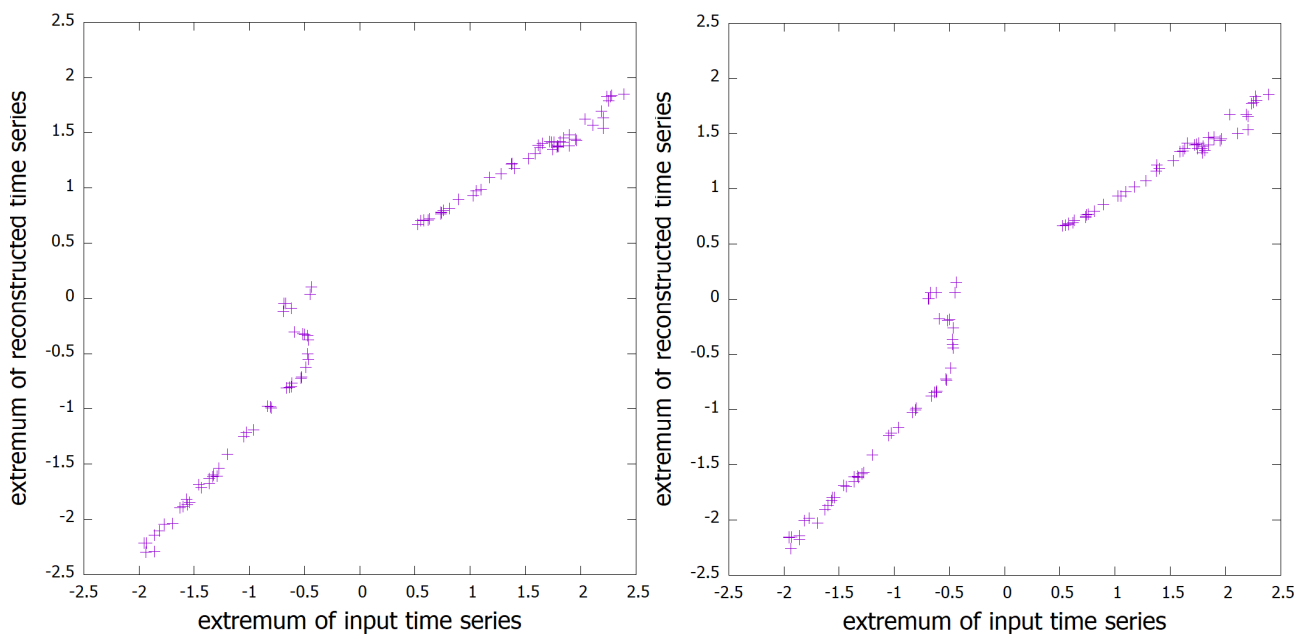


図 7: LIF モデルの漏れ τ に対し、発火数が一定となるようにバイアス項 B_τ を調整した場合の共通入力時系列と再構成時系列から抽出された極値間の相関図

図 8: 発火数が一定となる範囲で、LIF モデルの漏れ τ を微小に変化させた場合の共通入力時系列と再構成時系列から抽出された極値間の相関図

感染症数理モデルを用いた 感染者隔離効果の定量的評価

真鍋 歩未[†] 澤田 和弥^{††} 宮 南風^{††} 池口 徹^{†,††}

[†] 東京理科大学 工学部情報工学科

〒125-8585 東京都葛飾区新宿 6-3-1

^{††} 東京理科大学 大学院工学研究科 経営工学専攻

E-mail: [†]ayumi.manabe@hisenkei.net

あらまし 2020年、COVID-19が世界的に流行し、現在も感染拡大している。感染力が強く、無発症者が存在することで、無意識的に感染拡大を引き起こしたことが世界的大流行の一つの要因であると考えられている。COVID-19ウイルスに対するワクチン開発に時間がかかる中、感染拡大を抑制するために行われた主な対策の一つに、感染者の隔離がある。感染者の隔離により、どの程度、感染拡大を抑制できるのかを定量的に評価することは重要である。そこで本稿では、ネットワーク上での感染拡大抑制において感染者の隔離が有効であるかを確率的な感染症伝播モデルを用いて定量的に評価する。まず、周囲の感染状況によって回復確率を定め、感染者の隔離及び回復者のクラスを加えたr-SIARモデルを提案する。提案モデルはネットワーク上の頂点が確率的に状態遷移するモデルである。次に、r-SIARモデルを用いた感染症伝播の数値実験を行い、感染者を隔離した場合に最終感染率がどの程度減少するか調査した。その結果、ネットワーク構造と感染力によっては隔離という対策が効果的でないことがわかった。

キーワード 複雑ネットワーク、感染症伝播モデル、感染症拡散、COVID-19

Quantitative Evaluation of the Effects of Quarantine of Infected Person Using Mathematical Models of Infectious Diseases

Ayumi MANABE[†], Kazuya SAWADA^{††}, Nanfu MIYA^{††}, and Tohru IKEGUCHI^{†,††}

[†] Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science

^{††} Department of Management Science, Graduate School of Engineering, Tokyo University of Science

E-mail: [†]ayumi.manabe@hisenkei.net

Abstract In 2020, COVID-19 has become a global pandemic and is still spreading. One of the reason why COVID-19 spreads widely is that it has high infectivity and some patients infected uncsciously. While the development of a vaccine against COVID-19 virus would take much more time, one of the main strategies taken to control the spread of infection is the quarantine of infected persons. However, it is an important issue to quantitatively investigate to what extent the quarantine of infected patients can reduce the spread of infection. In this paper, we quantitatively evaluate the effectiveness of the quarantine to control the spread of infection, using a stochastic infection transmission model on networks. First, we propose an r-SIAR model that introduces the process of quarantine of infected persons and a class of recovering persons. The proposed r-SIAR model has stochastic dynamics of state transitions at each vertex on the network. Next, using the proposed model, we simulated the spread of infectious diseases and investigated how much the final infection rates are reduced when the infected patients were quarantined. We found that depending on the network structure and infectivity, the quarantine of infectious person is not effective.

Key words complex network, infection transmission model, spread of infectious disease, COVID-19

1. はじめに

2020年、COVID-19が世界的に流行し、現在も感染拡大している。COVID-19は感染力が強く、自覚症状のない感染者が多く存在する。これにより、無意識的な感染拡大を引き起こし、その結果、世界的大流行が発生したと考えられている。COVID-19に対するワクチン開発に時間がかかる一方で、感染拡大を抑制するために行われた対策の一つに、感染者の隔離がある。感染者の隔離によって感染拡大をどの程度抑制できるのかを定量的に評価することは重要である。

そこで本稿では、感染拡大を抑制するために、感染者を隔離することの有効性を感染症伝播モデルを用いて定量的に評価する。隔離の有効性を評価するために、周囲の感染状況によって回復確率を定め、感染者の隔離及び回復者のクラスを加えたr-SIARモデルを提案する。SI系モデル[1]は感染者数のダイナミクスを常微分方程式で記述するが、本稿で提案するr-SIARモデルはネットワークの頂点の状態が確率的に遷移するダイナミクスを有するとしている。提案モデルを用いて、感染症伝播の数値実験を行い、感染者の隔離をした場合と、隔離を行わなかった場合の最終感染率の差異を調査した。その際、UCM (Uncorrelated Configuration Model)[2]と、WSM (Watts-Strogatz Model)[3]をネットワークとして使用した。

2. 感染症伝播モデル

2.1 r-SIS モデル

文献[4]では、r-SISモデル(resource based Susceptible-Infectious-Susceptible Model)が用いられている。このモデルでは、死亡や誕生を考慮しないものとしている。SISモデルの状態遷移図を図1に示す。

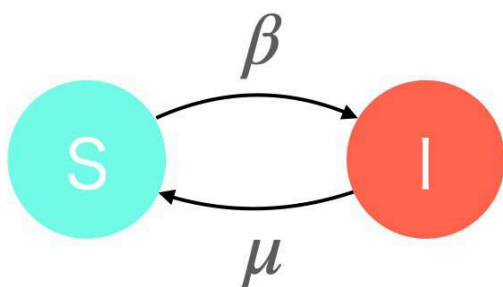


図1 SISモデルの状態遷移図

SISモデルでは、未感染者Sクラスから感染者Iクラスへの遷移が感染確率 β によって、また、IクラスからSクラスへの遷移が回復確率 μ によって発生する。本稿では、r-SISモデル[4]の振る舞いを離散的に再現し、これを改良したモデル(r-SIARモデル)を提案する。r-SISモデルにおいて、 $x_i(t)$ を頂点 i の時刻 t のときの状態値とする。頂点 i がSクラスするとき $x_i(t) = 0$ 、Iクラスするとき $x_i(t) = 1$ と定める。また、隣接行列の (i, j) 成分を Y_{ij} と表す。ここで、時刻 t でSクラス($x_i(t) = 0$)となる

頂点 i と隣接するIクラス($x_j(t) = 1$)である頂点 j の集合を $l_i(t)$ と表すと、 $l_i(t)$ は式(1)で表せる。

$$l_i(t) = \{j \mid x_i(t) = 0 \cap Y_{ij}x_j(t) = 1\} \quad (1)$$

すなわち、 $|l_i(t)|$ は時刻 t で $x_i(t) = 0$ となる頂点 i と隣接する $x_j(t) = 1$ である頂点 j の個数である。また、感染確率 β は、隣接するIクラスとの一回の接触でIクラスに遷移する確率を示すため、時刻 t で頂点 i がSクラス($x_i(t) = 0$)からIクラス($x_i(t+1) = 1$)に遷移する確率は $1 - (1 - \beta)^{|l_i(t)|}$ となる。

以上より、ネットワーク上の頂点 i の状態遷移のダイナミクスは式(2)となる。

$$x_i(t+1) = \begin{cases} x_i(t) + G(1 - (1 - \beta)^{|l_i(t)|}) & (x_i(t) = 0) \\ x_i(t) - G(\mu_i(t)) & (x_i(t) = 1) \end{cases} \quad (2)$$

ただし、 $G(p)$ は確率 p で1をとり、確率 $1-p$ で0をとる関数とする。一方、現実的には、回復確率は単純な確率に加えて治療や薬などの資源(リソース)にも依存すると考えられる。そこで、r-SISモデルでは、回復確率 μ を分配されるリソースの量に依存させている。時刻 t での頂点 i の回復確率 $\mu_i(t)$ を式(3)に示す。

$$\mu_i(t) = \mu_0 + \mu_r \frac{r_i(t)}{k_i} \quad (3)$$

ここで、 μ_0 は自己治癒確率を表す。本稿では、文献[4]と同様に、 $\mu_0 = 0$ とした。 k_i は頂点 i の次数、 μ_r は得たリソースの使用率、 $r_i(t)$ は時刻 t のとき、頂点 i が隣接するSクラスのノードから得るリソース量を表す。式(3)より、頂点 i の回復確率 $\mu_i(t)$ は得られるリソース量 $r_i(t)$ に比例する。また、リソースはSクラスの頂点からIクラスの頂点へ一方方向にのみ配分される。

各頂点の獲得リソース量は、周囲の頂点の感染状況、リソースの配分戦略によって異なる。例として、図2(a)のような状態を考える[4]。水色の頂点がSクラスの頂点、赤色の頂点がIクラスの頂点を表す。このとき、頂点3から頂点1, 2への感染力はそれぞれの隣接頂点のクラスによって決定する。図2(b)の頂点1では、隣接するSクラスの頂点が1つ、図2(c)の頂点2では3つある。つまり、頂点1よりも、頂点2の方が多くの頂点に疾病を伝染させる可能性が高い。このとき、頂点1をLI (Lower Infectiousness) 頂点、ノード2をHI (Higher Infectiousness) 頂点と呼ぶ。

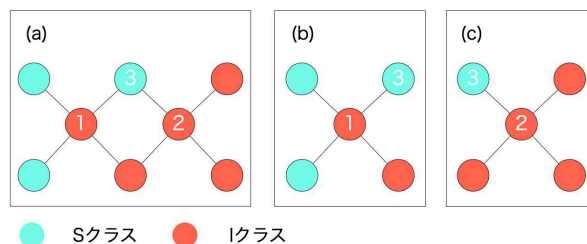


図2 (a) ネットワークの例、(b) 頂点1(周囲にSクラスの頂点が多い頂点: HI 頂点)とその隣接頂点、(c) 頂点2(周囲にIクラスの頂点が多い頂点: LI 頂点)とその隣接頂点

また、時刻 t のとき、 I クラスの頂点 i が持つ S クラスの頂点と接続する枝の本数を $m_i(t)$ とする。図 3 のようなネットワークでは、 $m_i(t) = 2$ (紫色)、 $m_k(t) = 3$ (緑色) となる。

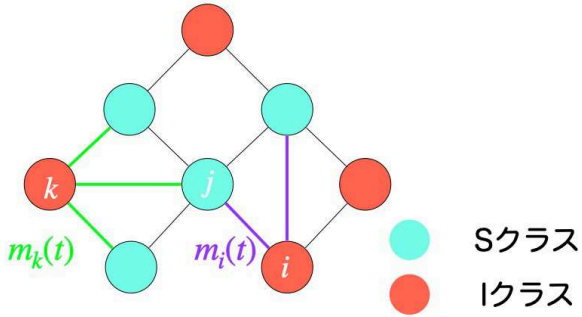


図 3 $m_i(t)$, $m_k(t)$ として数える枝

時刻 t で頂点 i が、隣接する S クラスの頂点 j から得られるリソース量を $r'_{j \rightarrow i}(t)$ として、式 (4) のように定義する。

$$r'_{j \rightarrow i}(t) = \frac{(m_i(t) + 1)^\alpha}{\sum_{q \in \{j(I)\}} (m_q(t) + 1)^\alpha} \quad (4)$$

ここで、 $\{j(I)\}$ は、頂点 j と隣接する I クラスの頂点の集合、 α はパラメータである。例えば、図 3 のようなネットワークにおいて、時刻 t で、 I クラスの頂点 i が S クラスの隣接頂点 j から得るリソース量 $r'_{j \rightarrow i}(t)$ は、 $m_i(t) = 2$, $m_k(t) = 3$ であるので、

$$r'_{j \rightarrow i}(t) = \frac{(m_i(t) + 1)^\alpha}{(m_i(t) + 1)^\alpha + (m_k(t) + 1)^\alpha} = \frac{3^\alpha}{3^\alpha + 4^\alpha}$$

となる。また、時刻 t で頂点 i が配分されるリソース量 $r_i(t)$ は式 (5) で表すことができる。

$$r_i(t) = \sum_{j \in \{i(S)\}} r'_{j \rightarrow i}(t) \quad (5)$$

ここで、 $\{i(S)\}$ は頂点 i と隣接する S クラスの頂点の集合である。式 (4) は、 $\alpha > 0$ のとき HI 頂点に対して優先的にリソース配分し、 $\alpha < 0$ のとき LI 頂点に対して優先的にリソース配分することを意味する。また、 $\alpha = 0$ のとき、各頂点には均等にリソースを配分することになる。

2.2 r-SIAR モデル

本稿では、感染者の隔離という対策の効果を定量的に評価するために、2.1 節の r-SIS モデルの I クラスを 2 つのクラスに分けたモデルを提案する。具体的には、r-SIS モデルの I クラスを、確率 δ で隔離された感染者クラス (I クラス) と、確率 $1 - \delta$ で隔離されなかった感染者クラス (A クラス) に分類する。提案モデルでは、隔離となった I クラス頂点をネットワークから外すこととする。また、感染症の種類によっては、人間は一度疾病に感染し、回復することで、免疫を得ることもある。免疫を獲得することで再感染せず、他者に疾病を感染させない状態となる。そこで、提案する r-SIAR モデルでは、このように再感染せず、他者へ疾病を感染させない回復者のクラス (R クラス) も導入した。ここで、 I クラスの頂点 i は隔離されてから時刻

u 経過後に回復者 R クラスに遷移し、隔離前のネットワーク上での枝を再度つなぎ直す。 A クラスの頂点は 2.1 節の r-SIS モデルと同様に、ネットワークから隔離されることはなく、獲得したリソースによる回復確率のもと、回復者 R クラスに遷移する。r-SIAR モデルの状態遷移図を図 4 に示す。

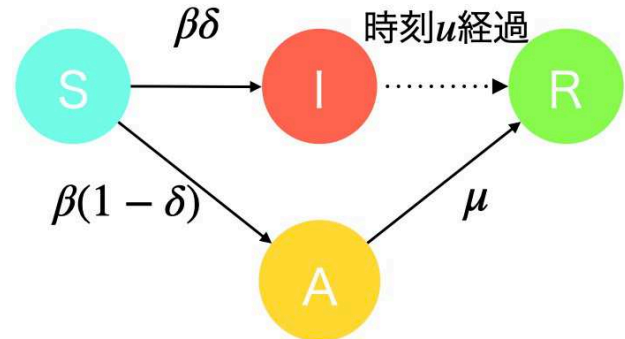


図 4 r-SIAR モデルの状態遷移図

次に、 $x_i(t)$ を頂点 i の時刻 t での状態値とし、頂点 i が S クラスのとき $x_i(t) = 0$, I クラスのとき $x_i(t) = 1$, A クラスのとき $x_i(t) = 2$, R クラスのとき $x_i(t) = 3$ とする。また、時刻 t のとき、隣接行列の (i, j) 成分を $Y_{ij}(t)$ と表す。ここで、時刻 t で S クラス ($x_i(t) = 0$) となる頂点 i と隣接する A クラス ($x_j(t) = 2$) である頂点 j の集合を $I'_i(t)$ と表すと、 $I'_i(t)$ は式 (6) で表せる。

$$I'_i(t) = \{j \mid x_i(t) = 0 \cap Y_{ij}(t)x_j(t) = 2\} \quad (6)$$

すなわち、 $|I'_i(t)|$ は、時刻 t で S クラス ($x_i(t) = 0$) となる頂点 i と隣接する A クラス ($x_j(t) = 2$) である頂点 j の個数を表す。また、感染確率 β は、隣接する A クラスとの一回の接触によって S クラスの頂点が I または A クラスに遷移する確率を示す。よって、 S クラス ($x_i(t) = 0$) から I クラス ($x_i(t+1) = 1$) または A クラス ($x_i(t+1) = 2$) に遷移する確率は $1 - (1 - \beta)^{|I'_i(t)|}$ となる。以上より、ネットワーク上の頂点 i の状態遷移のダイナミクスは式 (7) となる。

$$x_i(t+1) = \begin{cases} x_i(t) + G(1 - (1 - \beta)^{|I'_i(t)|})(G(1 - \delta) + 1) & (x_i(t) = 0) \\ x_i(t) + 2H(t - T_i - u) & (x_i(t) = 1) \\ x_i(t) + G(\mu_i(t)) & (x_i(t) = 2) \\ x_i(t) & (x_i(t) = 3) \end{cases} \quad (7)$$

ただし、関数 $H(x)$ はステップ関数である。ここで、式 (7) の T_i は頂点 i がネットワーク上で隔離された時刻とする。また、感染者クラスに遷移する頂点は、確率 δ によって I クラス、確率 $1 - \delta$ で A クラスに遷移する。 I クラスとなった頂点はネットワーク上で枝を外して隔離状態とし、一時的に孤立点とする。孤立点となった I クラスの頂点は、 u 時刻経過後、 R クラスへ遷移し、ネットワークに復帰する。一方、 A クラスの頂点はネットワーク上で隔離されずに、各時刻ごとに与えられる回復確率 $\mu_i(t)$ によって R クラスへ遷移する。ここで、回復確率 $\mu_i(t)$ は

式(3)~(5)と同様に定義する。ただし、r-SIAR モデルでは式(4)の $m_i(t)$ を、Aクラスの頂点 i が持つ S または R クラスの頂点と接続する枝の本数とする。

3. ネットワークモデル

本節では、感染症伝播の数値実験に使用した2つのネットワークモデルについて述べる。

3.1 UCM (Uncorrelated Configuration Model)

文献[4]では、次数分布が任意のべき則に従うネットワーク UCM (Uncorrelated Configuration Model) [2] が使用されている。UCMは、ネットワークの頂点数 N 、次数分布のべき指数 γ 、最小次数 k_{\min} を決め、最大次数が \sqrt{N} となるような制約のもと、次数列を作成する。作成した次数列をネットワークの各頂点に割り当て、ランダムに結ぶことで作成されるネットワークである。

3.2 WSM (Watts-Strogatz Model)

WSM[3]は、頂点数 N 、次数 k の規則的な円環状格子から、確率 p で枝をつなぎ変えて作成されるネットワークモデルである。 $p=0$ のときは枝をつなぎ変えが起きず、規則的な円環状格子ネットワークとなる。一方、 $p=0.01$ 前後ではスモールワールドネットワーク、 $p=1$ では全ての枝がつなぎ変えられたランダムネットワークになる。

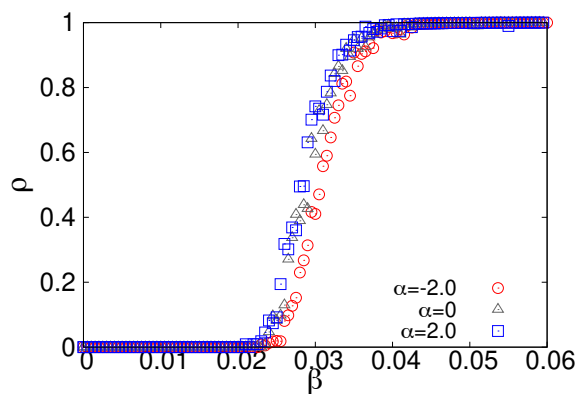
4. 数値実験

本稿で行った数値実験について述べる。本稿の目的は、ネットワーク上での感染拡大を抑制するために、感染者の隔離が有効であるかを定量的に評価することである。そこで、各感染症伝播モデルを用いて、一度の接触による感染確率 β と、最終感染率 ρ の関係を調査した。 $V = \{i \mid x_i(t) > 0\}$ としたとき、 V は r-SIS モデルでは I クラスの頂点集合、r-SIAR モデルでは I, A, R クラスの頂点集合を表す。最終感染率 ρ を式(8)のように定義する。

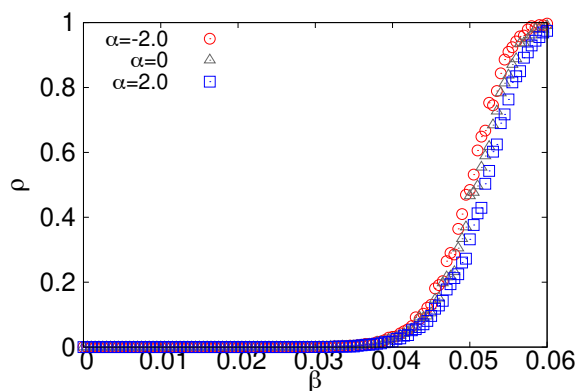
$$\rho = \frac{|V|}{N} \quad (8)$$

ネットワークは頂点数 $N = 5000$ 、最小次数 $k_{\min} = 3$ 、次数分布が $\gamma = 2.4$ のべき則に従う UCM と、頂点数 $N = 5000$ 、次数 $k = 8$ の規則的な円環格子状ネットワークを確率 $p = 0.01$ で枝をつなぎ変えた WSM を使用した。シミュレーション開始時には、全頂点の1%の50頂点を感染源としてランダムに配置した。また、r-SIAR モデルでは、感染源の50頂点に対し、確率 δ で I クラス、確率 $1-\delta$ で A クラスとした。リソースの使用率を $\mu_r = 0.60$ とし、リソース戦略パラメータ $\alpha = -2.0, 0, 2.0$ とした。

まず、r-SIS モデルを用いた場合の感染確率 β と最終感染率 ρ の関係を図5に示す。ここで、赤のプロットが $\alpha = -2.0$ 、灰色のプロットが $\alpha = 0$ 、青のプロットが $\alpha = 2.0$ のリソース配分戦略を表す。



(a) UCM

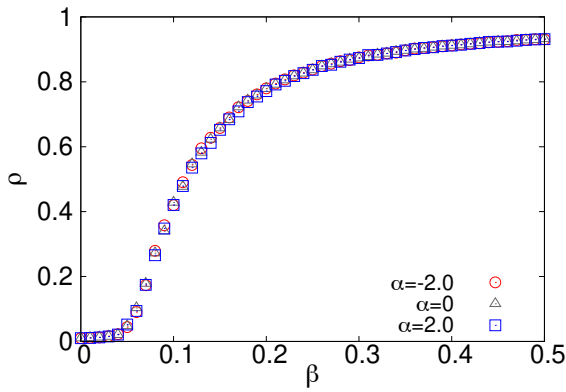


(b) WSM

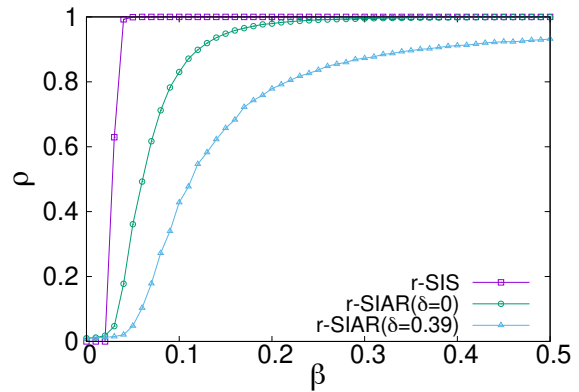
図5 感染確率 β と最終感染率 ρ の関係 (r-SIS モデル)

文献[4]では UCM を用いて、感染者数のダイナミクスを表す r-SIS モデルで数値実験が行われているが、文献[4]の図2(a)と図5(a)は、ほぼ同様の概形が得られた。ただし、文献[4]では $\beta \leq 0.015$ と $\beta > 0.015$ で有効な戦略の入れ替わりが起きているが、我々の数値実験では、どちらのネットワークモデルを使用しても各戦略間の最終感染率 ρ に対する有意な差はなく、戦略間の効果の入れ違いは起きなかった。

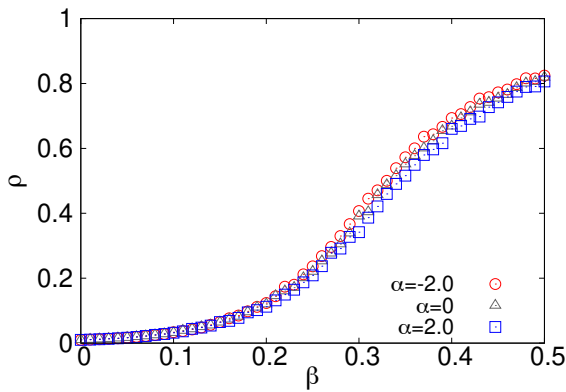
次に r-SIAR モデルを使用したときの感染確率 β と、最終感染率 ρ の関係を調査した。このとき、厚生労働省の新型コロナウイルス感染症の国内発生動向[5]より、7月15日までの東京都の陽性患者のうち、有症状者の割合を算出し、隔離確率を $\delta = 0.39$ と設定した。得られた結果を図6に示す。



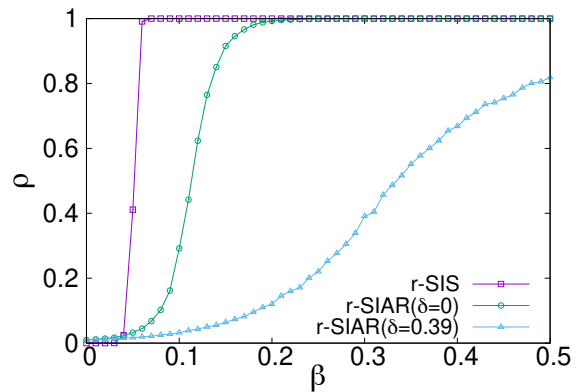
(a) UCM



(a) UCM



(b) WSM



(b) WSM

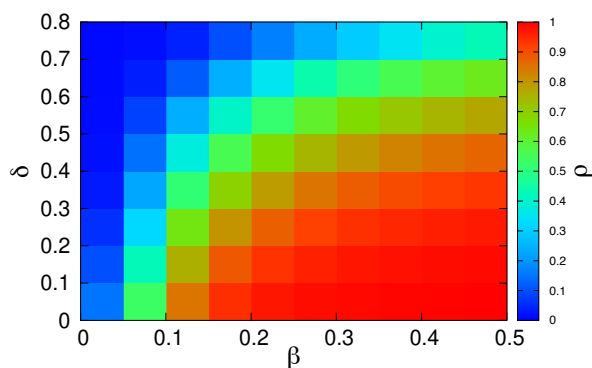
図6 感染確率 β と最終感染率 ρ の関係 (r-SIAR モデル)

図6から、再感染が起きないことと、ネットワークから約40% ($\delta = 0.39$) の感染者を隔離することによって、r-SISモデルよりも最終感染率 ρ を抑制している。また、図6(a), (b)から、どちらのネットワークモデルにおいても各戦略の違いが最終感染率に影響を与えていないこともわかる。したがって、これ以降は $\alpha = 0$ に固定して数値実験を行った。

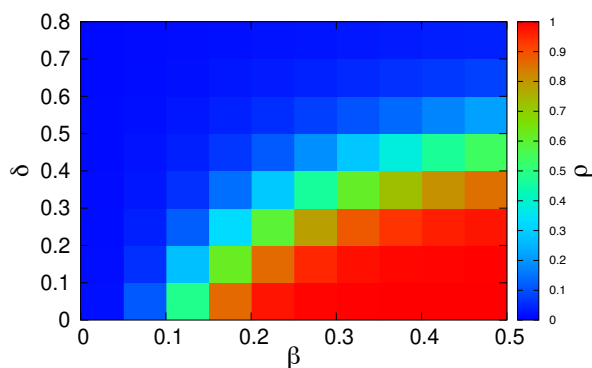
次に、再感染が起きないことと感染者隔離の効果をそれぞれ定量的に評価するために、r-SISモデルと、 $\delta = 0, 0.39$ としたr-SIARモデルについて、感染確率 β と最終感染率 ρ の関係を図7に示す。図7(a)を見ると、UCMにおいては、約40%の感染者の隔離という対策を行っても、 $0.1 \leq \beta \leq 0.2$ では約20%程度しか最終感染率を抑制できていないことがわかる。一方で、図7(b)より、WSMでは $0 < \beta < 0.3$ において、感染者を隔離しないときに比べ約50%以上の最終感染率 ρ を抑制できていることがわかる。

図7 r-SISモデル, r-SIARモデル ($\delta = 0, 0.39$) の感染確率 β と最終感染率 ρ の関係の比較

次に、r-SIARモデルについて、隔離確率 δ と感染確率 β を変化した場合に、最終感染率 ρ がどのように変化するかを調査した。結果を図8に示す。図8では、最終感染率 ρ ($0 \leq \rho \leq 1$) をカラーマップで示している。図8(a), (b)から、どちらのネットワークにおいても δ の値が大きいくほど、どの β に対しても感染を抑制できていることがわかる。しかし、図8(a)を見ると、UCMにおいては、 $\beta = 0.5$ のとき、6割以上の感染者を隔離しないと $\rho > 0.7$ となることがわかる。すなわち、感染力の強い疾病においては、感染者を半数以上隔離しても最終感染率は高いままであり、7~8割以上の感染者を隔離しなければ、感染の拡大を抑制することはできない。一方、図8(b)を見ると、WSMでは、UCMに比べて全体的に青色が占めている割合が高く、感染者の隔離が感染症の拡大抑制に効果的であることを示している。しかし、WSモデルでも、 $\beta = 0.5$ のとき半数以上の感染者を隔離しなければ $\rho \geq 0.8$ となる。これらの結果から、感染者を隔離する対策は、ネットワーク構造と、感染力の強さに依存して効果的となる場合もあるが、感染力が強い疾病 ($\beta \geq 0.4$) の場合、感染者の大多数を隔離しなければ、感染の拡散抑制に効果はない。



(a) UCM



(b) WSM

図8 感染確率 β と隔離確率 δ , 最終感染率 ρ の関係

5. ま と め

本稿では、確率的な感染症伝播モデルを用いて、感染者の隔離がネットワーク上での感染拡大抑制にどの程度有効な対策であるかを定量的に評価した。感染者の隔離が有効であるかを評価するために、r-SIS モデルに隔離の効果を導入した r-SIAR モデルを提案した。また、r-SIS モデルと r-SIAR モデルを用いて、感染確率と隔離確率を変化させたときの最終感染率を評価した。対象とするネットワークは UCM と WSM とした。

数値計算による検証の結果、感染力が高い感染症の場合は、感染者を 5 割以上隔離しなければ、8 割程度拡散してしまうことがわかった。この結果は、ネットワーク構造、感染力の強さによっては隔離という対策は効果的とならないことを示唆するものである。

今後は、どのようなネットワーク構造でより感染者の隔離が有効であるのかを調査することが必要である。また、COVID-19 の感染拡大傾向により適応したモデル化のために、再感染を考慮した感染症モデルや、隣接頂点だけでなく濃厚接触者という概念を含んだ感染症モデルの構築を検討する。

文 献

- [1] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the*

Royal Society of London. Series A, Containing papers of a mathematical and physical character, Vol. 115, No. 772, pp. 700–721, 1927.

- [2] Michele Catanzaro, Marián Boguná, and Romualdo Pastor-Satorras. Generation of uncorrelated random scale-free networks. *Physical Review E*, Vol. 71, No. 2, p. 027103, 2005.
- [3] Duncan J Watts and Steven H Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, Vol. 393, No. 6684, pp. 440–442, 1998.
- [4] Xiao-Long Chen, Shi-Min Cai, Ming Tang, Wei Wang, Tao Zhou, and Pak-Ming Hui. Controlling epidemic outbreak based on local dynamic infectiousness on complex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, Vol. 28, No. 12, p. 123105, 2018.
- [5] <https://www.mhlw.go.jp/content/10906000/000649533.pdf>, (最終閲覧日 2020/07/23).

大規模自転車再配置問題に対する 動的な重み係数調整を用いた解法の性能調査

對馬 帆南[†] 松浦 隆文^{††} 池口 徹^{†,††}

[†] 東京理科大学 大学院工学研究科 情報工学専攻 〒125-8585 東京都葛飾区新宿 6-3-1

^{††} 日本工業大学 先進工学部 情報メディア工学科 〒345-8501 埼玉県南埼玉郡宮代町学園台 4-1

^{†††} 東京理科大学 工学部 情報工学科 〒125-8585 東京都葛飾区新宿 6-3-1

E-mail: [†]honami@hisenkei.net, ^{††}matsuura@nit.ac.jp, ^{†††}tohru@rs.tus.ac.jp

あらまし 我々は既に、バイクシェアリングシステム (BSS) において、自転車回収車が短い時間で駐輪ポートの自転車台数を調整する経路を決定するために複数の回収車による自転車回収車再配置問題 (mBSSRP) と発見的解法を提案している。しかし、mBSSRP の制約条件が厳しいため、一部の問題例に対して実行可能解が得られない場合も存在する。そこで、mBSSRP の一部の制約を取り除き、その制約に対する違反を重み付きの違反量として mBSSRP の目的関数に加えたソフト制約付き自転車再配置問題 (mBSSRP-S) を提案している。さらに、mBSSRP-S を解くことで mBSSRP の優れた解を得るために、mBSSRP の実行可能解と実行不可能解を効率的に探索できる動的な重み係数調整法を提案している。数値実験により、ポート数が 50 程度の問題例に対して、mBSSRP の良好な近似解が得られることを報告しているが、実際に運営されている BSS のポート数は数百である。そこで本報告は、ポート数が多い大規模な BSS に対する提案解法の性能を調査する。数値実験の結果、大規模問題に対しても動的に重み係数を調整しながら mBSSRP-S を解く方法は有効であることを確認した。

キーワード 大規模バイクシェアリングシステム, 組合せ最適化問題, 自転車再配置問題, タブーサーチ

Dynamically changing penalties method for Large-Scale multiple-Vehicle Bike Sharing System Routing Problem

Honami TSUSHIMA[†], Takafumi MATSUURA^{††}, and Tohru IKEGUCHI^{†,††}

[†] Department of Information and Computer Technology, Graduate School of Engineering,
Tokyo University of Science

6-3-1 Nijuku, Katsushika-ku, Tokyo, 125-8585 Japan

^{††} Department of Information Technology and Media Design, Faculty of Advanced Engineering,
Nippon Institute of Technology

4-1 Gakuendai, Miyashiro-cho, Minami-Saitama-gun, Saitama, 345-8501 Japan

^{†††} Department of Information and Computer Technology, Faculty of Engineering,
Tokyo University of Science

6-3-1 Nijuku Katsushika-ku, Tokyo, 125-8585 Japan

E-mail: [†]honami@hisenkei.net, ^{††}matsuura@nit.ac.jp, ^{†††}tohru@rs.tus.ac.jp

Abstract In a bike sharing system (BSS), we proposed a multiple-Vehicle Bike Sharing System Routing Problem (mBSSRP) to determine short tours to adjust bicycles using vehicles in short time. However, it is difficult to construct feasible solutions for some instances due to strict constraints of mBSSRP. Therefore, we proposed a multiple-Vehicle Bike Sharing System Routing Problem with Soft Constraints (mBSSRP-S) that removes some constraints from mBSSRP and adds those violations to objective function of mBSSRP as penalties. Furthermore, we proposed a dynamic weight adjustment method to efficiently explore both feasible and infeasible solution space of mBSSRP by solving mBSSRP-S. Numerical experiments have shown that we obtained good solutions for instances with about 50 ports. However, the number of ports in real BSSs is more than hundreds. Thus, in this study, we investigate the performance for large-scale instances. As results of numerical experiments, we confirm that good solutions of mBSSRP are obtained for the large-scale instances.

Key words large-scale bike sharing system, combinatorial optimization problem, multiple-vehicle bike sharing system routing problem, tabu search

1. はじめに

近年、交通渋滞の緩和、CO₂削減、健康志向のために、バイクシェアリングシステム (Bike Sharing System, 以下、BSS) と

呼ばれる自転車を他の人と共有するサービスが、新しい交通システムとして世界中で導入されている。BSS では、街中の様々な場所に専用の駐輪ポートを設置する。各ポートの貸出可能な自転車台数と返却可能な空きスペースはリアルタイムに確認す

ることができ、簡易な手続きにより、自転車を利用することができる。BSSは、借りた自転車をどのポートへも返却できる柔軟性により、利用者が増加している。しかし、利用者が自由に出発地と目的地を決定できるため、自転車が一方ばかりに利用されてしまうと、自転車を借りることができないポート、自転車を返却できないポートが発生してしまう。このため、乗りたい場所で自転車を借り、好きな場所で返すことができるBSSの利点が失われてしまう。

この問題を解決するために、BSSでは、複数の自転車回収車を用いて駐輪ポートの自転車台数を調整している。我々は、複数の自転車回収車の効率的な自転車再配置経路を決定するために、複数の自転車回収車を用いた自転車再配置問題 (multiple-Vehicle Bike Sharing System Routing Problem, 以下、mBSSRP) を提案し、この問題に対する発見的解法を開発している [1]~[5]。提案解法は、短い時間で自転車回収車の巡回経路を構築することができたが、一部の問題例に対して実行可能な巡回経路を得ることができなかった。これは、mBSSRPは制約条件が厳しいため、制約を満たす巡回経路を構築できなかったことが原因の1つである。

そこで、mBSSRPの一部の制約を取り除き、その制約に対する違反を重み付きの違反量としてmBSSRPの目的関数に加えたソフトな制約付き自転車再配置問題 (multiple-Vehicle Bike Sharing System Routing Problem with Soft constraints, 以下、mBSSRP-S) を提案した [6]。mBSSRP-Sの解探索過程で得られる違反量が0の解はmBSSRPの実行可能解であり、違反量が0より大きい解はmBSSRPの実行不可能解である。我々は、mBSSRP-Sを解くことで、mBSSRPの良好な近似解を得られることを報告した。しかし、mBSSRPの良好な近似解を得るためには、mBSSRP-Sの目的関数に使用される重み係数を適切な値に設定することが不可欠であり、その設定は困難である。

そこで、mBSSRP-Sを解くことで優れたmBSSRPの解を得るために、mBSSRPの実行可能解と実行不可能解の両方を効率的に探索することができる動的な重み係数調整法も提案している [7]。数値実験を行った結果、提案解法は、ポート数が50程度の問題例に対してmBSSRPの良好な近似解が得られることを報告している。しかし、実際に運営されているBSSの駐輪ポート数は数百である。そこで本報告では、ポート数が数百ある大規模なmBSSRPに対する提案解法の性能調査を行う。数値実験の結果、大規模問題に対しても動的な重み係数調整法を用いてmBSSRP-Sを解く方法はmBSSRPの良い解を探索できることを確認した。

2. 自転車再配置問題

複数の自転車回収車を用いて各駐輪ポートの自転車台数を効率良く再調整する巡回経路を決定するために、multiple-Vehicle Bike Sharing System Routing Problem (mBSSRP) と multiple-Vehicle Bike Sharing System Routing Problem (mBSSRP-S) を提案している [1]。mBSSRPとmBSSRP-Sは、1つのデポと n 個の駐輪ポート、 m 台の自転車回収車が与えられる。各ポートは初期状態よりも自転車が超過しているポートまたは、不足しているポートの2種類に分けられる。回収車は超過しているポートで自転車を回収し、自転車が不足しているポートに自転車を補充する。mBSSRPとmBSSRP-Sの前提条件を以下に示す。

- (1) デポには予備の自転車が用意されているため、回収車はデポで自転車を積んで出発することが可能である。
- (2) 全ての回収車はデポを出発し、再びデポに戻らなければならない。
- (3) 再配置作業中に自転車台数は変化しない。
- (4) 各ポートには1台の回収車が1度しか訪問できない。

これらの前提条件のもとで、複数の自転車回収車を用いて再配置作業を行う。

2.1 multiple-Vehicle Bike Sharing System Routing Problem (mBSSRP)

mBSSRPの解は、前提条件に加えて以下に示す3つの制約条件を満たさなければいけない。

- (1) 制限時間制約: 回収車は制限時間内にデポに戻らなければならない
- (2) 積込み制約: 回収車は積載容量を超えて自転車を積むことができない
- (3) 補充制約: 1度の訪問で不足しているポートに必要な台数を補充しなければいけない

mBSSRPの目的は、前提条件と(1)~(3)の制約条件を満たす巡回経路中で、回収車の総移動時間が最小となる巡回経路を求めることである。

2.2 multiple-Vehicle Bike Sharing System Routing Problem with Soft Constraints (mBSSRP-S)

mBSSRP-Sでは、mBSSRPから(1)~(3)の制約条件を取り除き、それらの違反量を重み付きの係数として目的関数に加える。mBSSRP-Sの目的関数を以下に示す。

$$\text{総移動時間} + \alpha \times P_T + \beta \times (P_{b+} + P_{b-}) \quad (1)$$

式(1)において、総移動時間とは全ての回収車の移動時間の総和を示し、mBSSRPの目的関数である。 P_T は超過時間であり、制限時間を超えて作業を行った時間である。 P_{b+} と P_{b-} はそれぞれ積込み違反台数と補充違反台数を表す。積込み違反台数とは、自転車が超過しているポートで積載容量制約のために自転車を回収車に積込むことができなかった台数である。補充違反台数とは、自転車が不足しているポートにおいて回収車に積んでいる自転車台数が少ないために、自転車を補充することができなかった台数である。超過時間、積込み違反台数、補充違反台数が0の解は、mBSSRPの実行可能解である。 α 、 β はペナルティの重み係数である。

3. 提案解法の求解手順

mBSSRPとmBSSRP-Sを解くための提案解法の求解手順について説明する。まず始めに、構築法により初期解を生成する。mBSSRPを解く場合には最近挿入法、mBSSRP-Sの場合には最遠挿入法を基にした構築法により初期解を生成する。

次に、ルート内挿入法とルート内交換法を用いて1つの巡回経路内の改善操作を行う。ルート内挿入とは、ある巡回経路の3個以下の連続するポート列 $a'-b$ を同一巡回経路中の $c-c'$ に挿入する手法である(図1(a))。ルート内交換法とは、ある巡回経路の連続する3個以下のポート列 $a'-b$ と同一巡回経路中の連続す

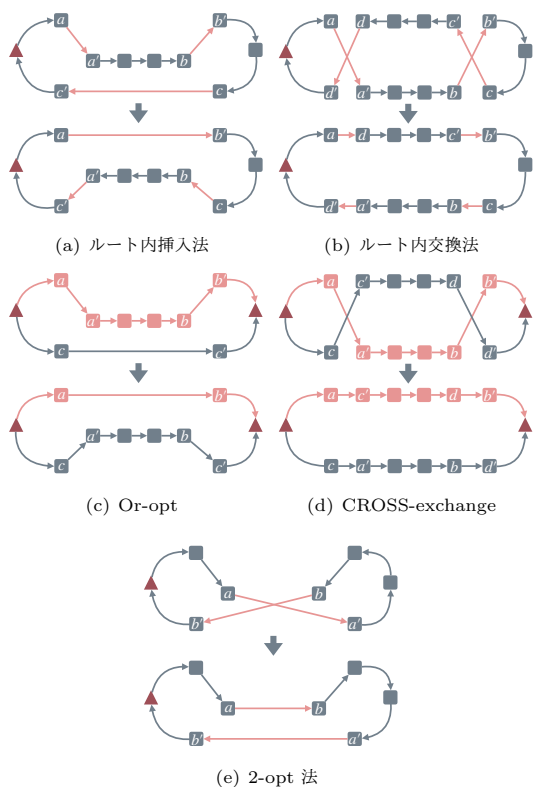


図 1 局所探索法の例。三角形はデポ、四角形はポートを示す。

る 3 個以下のポート列 $c' - d$ を交換する方法である (図 1(b)).

更に良い解を得るために、Or-opt [8] と CROSS-exchange [9] の実行をタブーサーチ [10]~[12] で制御する手法により解探索を行う。Or-opt とは、ある巡回路内の連続する 3 個以下のポート列 $a' - b$ を異なる巡回路 $c - c'$ に挿入する方法である (図 1(c)). CROSS-exchange とは、ある巡回路の 3 個以下の連続するポート列 $a' - b$ を異なる巡回路中の 3 個以下のポート列 $c' - d$ と交換する方法である (図 1(d)).

タブーサーチも局所探索と同様に、現在の解の近傍解に遷移する。局所探索では、現在の解よりも良い解にしか遷移することができないが、タブーサーチでは現在の近傍解の中で最も良い解に遷移する。したがって、現在の解よりも改悪な解に遷移する場合もある。タブーサーチは、周期的な解探索を防ぐために、過去の探索履歴をタブーリストに追加し、タブーリストに記載されている解へは一定期間遷移することができない。この期間のことをタブーニューアと呼ぶ。本研究では、タブーサーチが遷移する近傍を Or-opt または CROSS-exchange で構築する。本研究では、Or-opt または CROSS-exchange が実行された場合、交換される部分順列の始点の組み合わせをタブーリストに追加する。したがって、CROSS-exchange または Or-opt が実行された場合、交換される部分順列の始点の組み合わせ $a' - c'$ がタブーリストに保存され、これらを始点とする操作は一定期間禁止される。探索過程の中で mBSSRP の実行可能解が得られた場合 (超過時間、積込み違反台数、補充台数違反が 0 のとき)、2-opt 法を実行する。2-opt 法とは 1 つの巡回路内で枝 $a - a'$ と枝 $b - b'$ を繋ぎかえる方法である。あらかじめ決められた回数タブーサーチによる解遷移を繰り返し、解探索の過程で得られた最良解を出力する。その解に対して、ルート内挿入、ルート内交換を行い、1 つの巡回路内で改善操作を行う。

本研究では mBSSRP-S を解いて mBSSRP の良好な近似解を探索する場合、目的関数のペナルティの重み係数 α と β を探索状態に応じて動的に変化させる。 t 回目の解遷移における重み係数を $\alpha(t)$ 、積込み・補充制約に対する重み係数を $\beta(t)$ としたとき、 t 回目の解遷移における目的関数値は以下の式で表される。

$$\text{総移動時間} + \alpha(t) \times P_T(t) + \beta(t) \times (P_{b+}(t) + P_{b-}(t)) \quad (2)$$

式 (2) の t 回目の探索で得られた $[\alpha(t) \times P_T(t)]$ と $[\beta(t) \times (P_{b+}(t) + P_{b-}(t))]$ を比較し、式 (3), (4) の更新式を用いて係数調整を行う。

$$\alpha(t+1) = \begin{cases} \lambda\alpha(t) & \text{if } \alpha(t) \times P_T(t) > \beta(t)(P_{b+}(t) + P_{b-}(t)), \\ & \text{and } P_T(t) \neq 0, \\ \mu\alpha(t) & \text{if } \alpha(t) \times P_T(t) < \beta(t)(P_{b+}(t) + P_{b-}(t)), \\ & \text{and } P_T(t) \neq 0, \\ \alpha(t) & \text{if } \alpha(t) \times P_T(t) = \beta(t)(P_{b+}(t) + P_{b-}(t)), \\ & \text{and } P_T(t) \neq 0, \\ \alpha(t) & \text{if } P_T(t) = 0, \end{cases} \quad (3)$$

$$\beta(t+1) = \begin{cases} \lambda\beta(t) & \text{if } \alpha(t) \times P_T(t) < \beta(t)(P_{b+}(t) + P_{b-}(t)), \\ & \text{and } P_{b+}(t) + P_{b-}(t) \neq 0, \\ \mu\beta(t) & \text{if } \alpha(t) \times P_T(t) > \beta(t)(P_{b+}(t) + P_{b-}(t)), \\ & \text{and } P_{b+}(t) + P_{b-}(t) \neq 0, \\ \beta(t) & \text{if } \alpha(t) \times P_T(t) = \beta(t)(P_{b+}(t) + P_{b-}(t)), \\ & \text{and } P_{b+}(t) + P_{b-}(t) \neq 0, \\ \beta(t) & \text{if } P_{b+}(t) + P_{b-}(t) = 0. \end{cases} \quad (4)$$

ここで、 $\lambda (> 1)$ は重み係数の増加量を調整するパラメータ、 $\mu (0 < \mu < 1)$ は減少量を調整するパラメータである。

4. 数値実験

10km \times 10km のエリア内に、デポと 100 個のポートの分布が異なる問題例を 5 つ作成し性能評価を行う。過不足台数は $-5 \sim +5$ 台とし、最大積載容量 10 台の回収車 5 台が再配置作業を行う。回収車は 30[km/h] で移動し、自転車 1 台あたりの作業時間は 2 分、制限時間は 180 分とする。試行回数は 50 とし、タブーニューアは 70、タブーサーチの繰り返し回数 (イタレーション) は 1,000 とする。重み係数のパラメータ値は $\alpha(0) = 1$, $\beta(0) = 1$, $\lambda = 1.05$, $\mu = 0.7$ とした。

4.1 実験結果

表 1 に mBSSRP を解いた結果、表 2 に mBSSRP-S を解いて得られた mBSSRP の結果を示す。表 1, 2 は左から問題番号、目的関数の平均値、最良値、最悪値、計算時間、50 個の試行の中で得られた実行可能解の数を示す。表 1 より、問題 1, 2, 5 に対しては 50 試行の中で 1 つも実行可能解が得られなかったが、問題 3, 4 に対しては 50 試行の中で数回だけ実行可能解を得られたことがわかる。一方、mBSSRP-S を解いて、mBSSRP の実行可能解を得る方法では、全問題の全試行において実行可能解を得ることができている (表 2)。さらに、問題 3, 4 の最良解について比較すると、mBSSRP-S を解いたほうが短い巡回路を得ることができている。しかし、計算時間については mBSSRP を解くほうが短い結果となった。これは mBSSRP-S を解く場合、mBSSRP の実行不可能解も近傍解となるため、探索範囲が広がるからである。

次に、mBSSRP と mBSSRP-S の解探索の様子について考察する。図 2 は、問題 3 に対する各試行において、1000 イタレーション中に交換された部分列の先頭ポート a' と c' の組み合わせ

表 1 mBSSRP を解いた結果

No.	平均	最良	最悪	計算時間 [s]	得られた 実行可能解数
3	112,075.25	111,416	113,916	431.68	4
4	121,287.00	121,287	121,287	440.01	1

表 2 mBSSRP-S を解き得られた mBSSRP の結果

No.	平均	最良	最悪	計算時間 [s]	得られた 実行可能解数
1	119,663.24	116,138	125,689	539.44	50
2	118,280.04	113,323	125,884	532.35	50
3	113,082.80	108,516	119,708	549.36	50
4	120,097.32	116,042	127,243	555.68	50
5	122,788.34	119,708	128,506	536.73	50

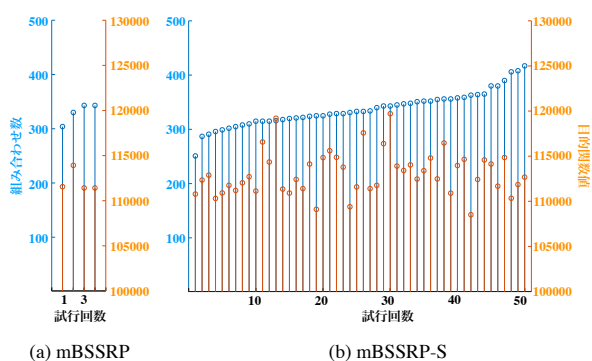


図 2 mBSSRP と mBSSRP-S を解いたとき、交換されたポートの組み合わせの種類数と目的関数値の関係

せの種類数と最良解を示す。なお、試行回数は組み合わせの種類数に従い昇順に並べ変えている。mBSSRP-S では mBSSRP の実行不可能領域へも遷移できるため、mBSSRP を解くよりも多彩な交換が行われ、それにより優れた解が得られたのかと予想される。しかし、図 2 より、交換されたポートの組み合わせ種類数と性能は関係性が見られなかった。

図 3 に、各試行の 1000 イタレーションの中で実行可能解を探索した回数と得られた最良解の関係を示す。試行回数は実行可能解と探索した回数に応じて昇順に並び変えている。mBSSRP の実行可能解をより多く見つけることで最良解の候補が増えるため、解探索性能が向上すると考えられるが、図 3 の結果からは、これらの間には関係がないことが分かる。

5. まとめ

本研究では、動的に重み係数を調整しながら mBSSRP-S を解き、mBSSRP の良好な近似解を求める方法が大規模な mBSSRP に対しても有効であるのか性能を調査した。数値実験の結果、mBSSRP-S を解き、mBSSRP の解を探索する方法は、mBSSRP を解くよりも短い経路を見つけられることを確認した。また、mBSSRP では実行可能解を見つけることができなかつた問題に対して、mBSSRP-S では mBSSRP の実行可能解を見つけることができた。しかし、mBSSRP-S では mBSSRP の実行不可能解も近傍解として扱われてしまうため、解探索の範囲が広がり、長い計算時間を要してしまう。従って、今後の

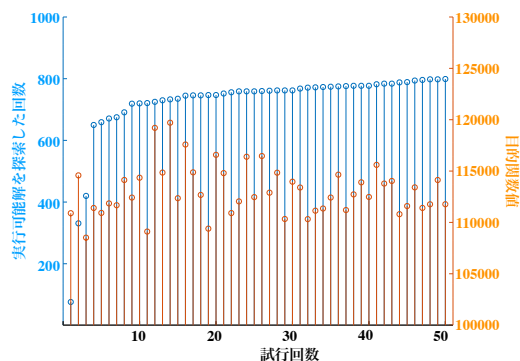


図 3 mBSSRP-S を解いたとき、実行可能解した回数と目的関数値の関係

課題としては、探索範囲を限定し計算時間を短縮することなどが考えられる。

また、mBSSRP-S を解くことで、良い解が得られる理由は mBSSRP の実行可能解と実行不可能解を探索するためだと考えられる。そこで、これを確認するために近傍操作の種類数と性能や実行可能解の探索回数と性能について調べたが、明らかな関係性を示してはなかった。提案解法の性能を向上させるためにも優れた性能が得られる理由について明らかにする必要はある。

謝辞 本研究は JSPS 科研費 19K04907, 15KT0112, 17K00348 の助成を受けたものである。

文 献

- [1] 對馬帆南, 松浦隆文, “バイクシェアリングシステムにおける複数の自転車回収車を用いた自転車再配置問題,” 2017 年電子情報通信学会 NOLTA ソサエティ大会, NLS-16, 2017.
- [2] H. Tsushima, T. Matsuura, and K. Jin'no, “Local Search Method for Multiple-Vehicle Bike Sharing System Routing Problem,” *Journal of Signal Processing*, Vol. 22, No. 4, pp. 157-160, 2018.
- [3] H. Tsushima, T. Matsuura, and T. Kimura, “Tabu Search Method for Multiple-Vehicle Bike Sharing System Routing Problem,” *Proc. of IEICE NOLTA 2018*, pp. 168-171, 2018.
- [4] 對馬帆南, 松浦隆文, 木村貴幸, “複数の自転車回収車を用いた自転車再配置問題に対するカオスサーチ法の提案,” *信学技報*, Vol. 118, No. 106, pp. 57-62, 2019.
- [5] Gurobi Optimizer: <http://www.gurobi.com>
- [6] 對馬帆南, 木村貴幸, 松浦隆文, “ソフトな制約付き自転車再配置問題に対する解法の研究,” *信学技報*, Vol. 119, No. 19, pp. 65-70, 2019.
- [7] 對馬帆南, 木村貴幸, 松浦隆文, “自転車再配置問題に対する実行可能解と不可能解の探索を動的に変化させる解法の提案,” *信学技報*, Vol. 119, No. 381, pp. 13-18, 2020.
- [8] Or. I, “Traveling salesman-type combinational problems and their relation to the logistics of blood banking,” *Xerox Univ. Microfilms* 1976.
- [9] E. Taillard, P. Badeau, M. Gendreau, F. Geurtin, and J.Y. Potvin, “A Tabu Search Heuristic for the Vehicle Routing Problem with Time Windows,” *Transportation Science*, Vol. 31, pp. 170-186, 1997.
- [10] F. Glover, “Tabu Search-Part I,” *ORSA Journal on Computing*, Vol. 1, No. 3, pp. 190-206, 1989.
- [11] F. Glover, “Tabu Search-Part II,” *ORSA Journal on Computing*, Vol. 2, No. 1, pp. 4-32, 1990.
- [12] F. Glover, “A user's guide to tabu search,” *Annals of Operations Research*, Vol. 41, pp. 3-28, 1993.

ネットワーク構造が情報拡散に与える影響の解析

ジョロエン[†] 澤田 和弥^{††} 島田 裕^{†††} 池口 徹^{†,††,†††}

[†] 東京理科大学 大学院工学研究科 情報工学専攻

〒125-8585 東京都葛飾区新宿 6-3-1

^{††} 東京理科大学 大学院工学研究科 経営工学専攻

^{†††} 埼玉大学 大学院理工学研究科 数理電子情報部門

〒338-8570 埼玉県さいたま市桜区下大久保 255

^{††††} 東京理科大学 工学部 情報工学科

E-mail: †jyoroen@hisenkei.net

あらまし 本稿では、固有ベクトル中心性を考慮した情報拡散の数理モデルを提案する。また、提案モデルを用いて、情報の影響力が情報拡散に与える効果とネットワーク構造の関係を調査する。まず、規則的なネットワークとランダムネットワーク上で情報の影響力を変化させた場合の最終的な情報の拡散率について調査した。次に、規則的なネットワークとランダムネットワーク上で情報の影響力を変化させた場合の平均拡散速度について調査した。その結果、情報の影響力が大きくなると、規則的なネットワークとランダムネットワーク上での最終的な情報の拡散率は規則的なネットワークの方が大きく、その差はピークに達してから小さい値に収束することがわかった。一方、情報がネットワーク全体まで広がる場合、規則的なネットワークとランダムネットワーク上での平均拡散速度はランダムネットワークの方が大きく、その差は情報の影響力が大きくなると共に拡大することがわかった。

キーワード 情報拡散, 複雑ネットワーク, 数理モデル, ネットワーク中心性

Analysis of influence of network structure on information diffusion

Luyan XU[†], Kazuya SAWADA^{††}, Yutaka SHIMADA^{†††}, and Tohru IKEGUCHI^{†,††,†††}

[†] Department of Information and Computer Technology, Graduate School of Engineering, Tokyo University of Science

Nijjuku 6-3-1, Katsushika-ku, Tokyo, 125-8585, Japan

^{††} Department of Management Science, Graduate School of Engineering, Tokyo University of Science

^{†††} Department of Information and Computer Sciences, Graduate School of Science and Engineering, Saitama University

255 Shimo-Okubo, Sakuraku, Saitama City, Saitama 338-8570, Japan

^{††††} Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science

E-mail: †jyoroen@hisenkei.net

Abstract In this paper, we propose an information diffusion model with eigenvector centrality. Using the proposed model, we investigated the relationship between the influence of information and the final diffusion rates on regular and random networks, when we change the influence of information. Next, we investigated the relationship between the influence of information and the average spreading speed on regular and random networks, when we change the influence of information. As a result, we found that the final diffusion rates on regular network are higher than those on random networks. When the influence of information increases, the difference of the final diffusion rates between regular and random networks increases to a peak value, and finally converges to a small value. On the other hand, when information spreads to the entire network, we found that the average spreading speed on random networks is faster than that on regular networks. In addition, the difference of the average spreading speed between regular and random networks increases with the influence of information.

Key words Information diffusion, Complex network, Mathematical model, Network centrality

1. Introduction

The social relationships of people can be described as networks, where people are represented as vertices, and their social ties are represented as edges. In recent years, there have been a lot of researches on information diffusion on such networks. For example, Centola reported that the information spreads more widely on a hexagonal lattice than on regular random networks by a social experiment [1]. To support the results of Centola's experiment, we have already proposed an information diffusion model in which the level of interest in information changes according to the eigenvector centrality of vertices. Then, we reported that the information spreads more easily on regular networks than on random networks through the proposed model [2].

In this paper, we investigated the relationship between the influence of information and the final diffusion rates on regular and random networks through the proposed model. Furthermore, we investigated the relationship between the influence of information and the average spreading speed on regular and random networks.

2. Proposed model

In this section, we describe the proposed model in which the level of interest in information changes according to the eigenvector centrality of vertices. Eigenvector centrality [3] is an indicator of the property that the more adjacent vertices of high degree has, the more likely the vertex plays an important role in the network. In other words, vertices with high eigenvector centralities are likely to play an important role in the networks.

We define the eigenvector centrality of vertex i as E_i . Then, the proposed model is shown by Eq. (1).

$$x_i(t+1) = e^{-1/\tau} x_i(t) + \sum_{j \in G_i} \alpha F(x_j(t) - \theta_j) \frac{E_j}{E_i}, \quad (1)$$

where $F(x)$ is a step function, $x_i(t)$ is the interest of information of the vertex i at time t , G_i is a set of indexes of adjacent vertices of the vertex i , θ_j is the threshold of the vertex j , τ is the decay parameter, and α is the parameter that determines the influence of information. In Eq. (1), $e^{-1/\tau}$ represents the decay of the level of interest in information over time. In Eq. (1), E_j/E_i indicates that the greater the eigenvector centrality of the information sender against the receiver, the higher level of interest in that information is.

3. Network models

In this section, we describe the networks which we used in this paper. As a regular network, we used the Hexagonal Lattice Network (HLN) [1] which is a torus network in

which each vertex has six neighbors as shown in Fig. 1(a). As a random network, we used random networks generated by rewiring edges in HLN in the same manner as the Watts and Strogatz model [4] as shown in Fig. 1(b). In addition, we define the rewiring probability as p , and the total number of vertices as N .

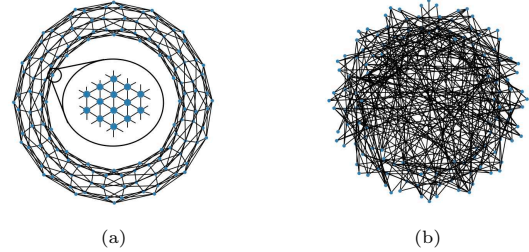


Figure 1 Examples of (a) the Hexagonal lattice network and (b) the random network ($N = 128$). The network in (b) is obtained by rewiring edges of the network in (a) with $p = 1$.

4. Numerical experiments

4.1 Experimental conditions

We set $0.05 \leq \alpha \leq 1$, and $e^{-1/\tau} = 0.8$. Under these conditions, we investigated the relationship between the influence of information and the final diffusion rates on regular and random networks through the proposed model. Furthermore, we investigated the relationship between the influence of information and the average spreading speed on regular and random networks. In the numerical experiments, we started the diffusion with a randomly choosing vertex i^* from the network and setting $x_{i^*}(0)$ which is larger than its threshold θ_{i^*} . The end timing of the diffusion T was defined as the previous timing when vertices that could spread information disappear. We define the number of vertices whose level of interest in information once reached their thresholds by time t as $n(t)$, the diffusion rate at time t as Eq. (2).

$$s(t) = \frac{n(t)}{N}. \quad (2)$$

Then, the final diffusion rate of information is defined by Eq. (3).

$$S = s(T). \quad (3)$$

The average spreading speed is defined by Eq. (4).

$$\bar{v} = \frac{S}{T}. \quad (4)$$

In this paper, we used HLN ($N = 512$) and their randomized networks. In addition, the threshold θ_i obeys a truncated normal distribution with an average $\mu = 0.7$ and a variance σ^2 which lies within the range $[0.4, 1.0]$. In this section, all results are averaged over 1,000 trials.

4.2 Results

First, we investigated the relationship between the influence of information α and the final diffusion rates S on regular and random networks when we applied the proposed model to HLN and random networks as shown in Fig. 2.

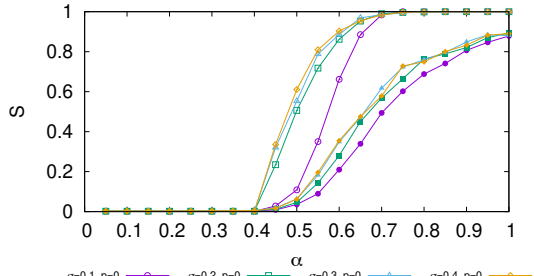


Figure 2 The relationship between the influence of information α and the final diffusion rates S when we applied the proposed model to HLN and random networks.

Figure 2 shows that the diffusion of information rarely occurs when the influence of information α is less than or equal to 0.4. On the other hand, when the value of α is greater than 0.4, the final diffusion rate S increases with the value of α on regular and random networks, and converges to unity on regular networks, when the value of α is greater than 0.7. In addition, the difference of S between regular and random networks changes significantly when $0.4 < \alpha \leq 1$. To calculate the difference of S between regular and random networks quantitatively, we investigated the relationship between the influence of information α and the difference of S on regular and random networks as shown in Fig. 3. In Fig. 3, we define the difference of S between regular and random networks as D_S .

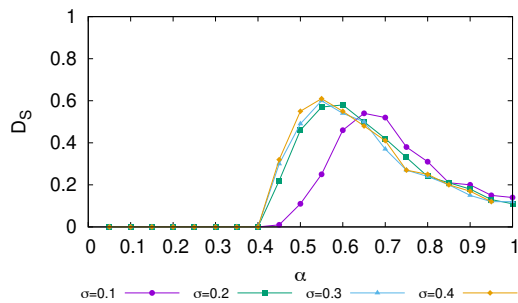


Figure 3 The difference of the final diffusion rates S between regular and random networks against the influence of information α .

Figure 3 shows that D_S increases to about 0.6 in the range of $0.4 \leq \alpha < 0.55$ and decreases to about 0.12 in the range of $0.55 \leq \alpha < 1$ when the standard deviation $\sigma = 0.2, 0.3$ and 0.4. On the other hand, D_S increases to 0.54 in the range of $0.4 \leq \alpha < 0.65$ and decreases to 0.12 in the range of $0.65 \leq \alpha \leq 1$, when $\sigma = 0.1$. These results show that

the greater the influence of information α is, the less the difference of the final diffusion rate S on regular and random networks, when $0.65 \leq \alpha \leq 1$.

Next, we investigated the average spreading speed \bar{v} on regular and random networks when the final diffusion rate $S = 1$. According to the frequency distribution normalized by the range $[0, 1]$ of the final diffusion rate S of random networks as shown in Fig. 4, we found that the cases that $S \in [0.9, 1]$ are less than 20 percents when $\alpha = 0.5$ and 0.55, about 20 ~ 40 percents when $\alpha = 0.6$. Figure 4 indicates that the cases that $S > 0.9$ are rare in random networks when α is less than 0.6. Then, we investigated the relationship between α and \bar{v} when $0.6 \leq \alpha \leq 1$ and $S = 1$. The results are shown in Fig. 5.

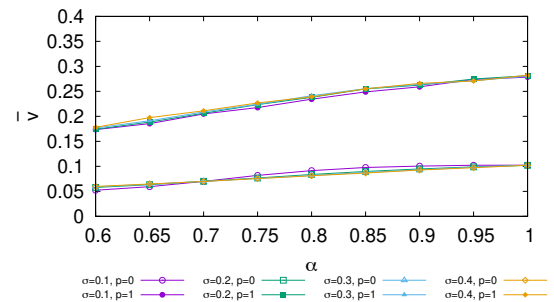


Figure 5 The relationship between the influence of information α and the average spreading speed \bar{v} when we applied the proposed model to HLN and random networks.

Figure 5 shows that the average spreading speed \bar{v} increases with the value of α in both regular and random networks. In addition, the values of \bar{v} on random networks are always greater than those on regular networks. Furthermore, the ratios of α and \bar{v} on random networks are greater than those on regular networks. To quantify the difference of \bar{v} on random and regular networks, we investigated the relationship between α and the difference of \bar{v} on random and regular networks as shown in Fig. 6, where we define the difference of \bar{v} between random and regular networks as $D_{\bar{v}}$.

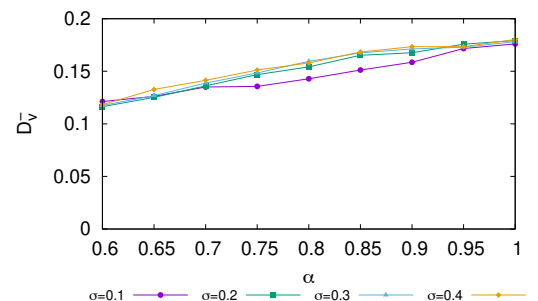


Figure 6 The difference of the average spreading speed \bar{v} between random and regular networks against the influence of information α

Figure 6 shows that $D_{\bar{v}}$ increases with the value of α in

the range of $0.6 \leq \alpha \leq 1$, when $S = 1$. In other words, the greater the influence of information α is, the faster the average spreading speed \bar{v} is. In addition, the spreading speed of information in the random networks is faster than regular networks, when $S = 1$.

5. Conclusion

In this paper, we investigated the relationship between the influence of information and the final diffusion rates on regular and random networks. Furthermore, we also investigated the relationship between the influence of information and the average spreading speed on regular and random networks.

As a result, we found that the information spreading on random networks tends to be stable and wide, almost as well as on regular networks, when the influence of information increases. On the other hands, when information spreads

to the entire network, we found that the average spreading speed of information on random networks tends to be faster than regular networks, when the influence of information increases.

Acknowledgments This research is partially supported by the JSPS Grant-in-Aids for Scientific Research (15KT0112, 17K00348, 18K18125).

References

- [1] D. Centola, "The spread of behavior in an online social network experiment," *Science*, Vol. 329, pp. 1194–1197, 2010.
- [2] L. Xu et al., "Information Diffusion Model with Network Centrality and Community Structure," *Technical report of IEICE*, Vol. 119, No. 471, pp. 83–88, 2020.
- [3] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification," *Journal of mathematical sociology*, Vol. 2, No. 1, pp. 113–120, 1972.
- [4] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, Vol. 393, pp. 440–442, 1998.

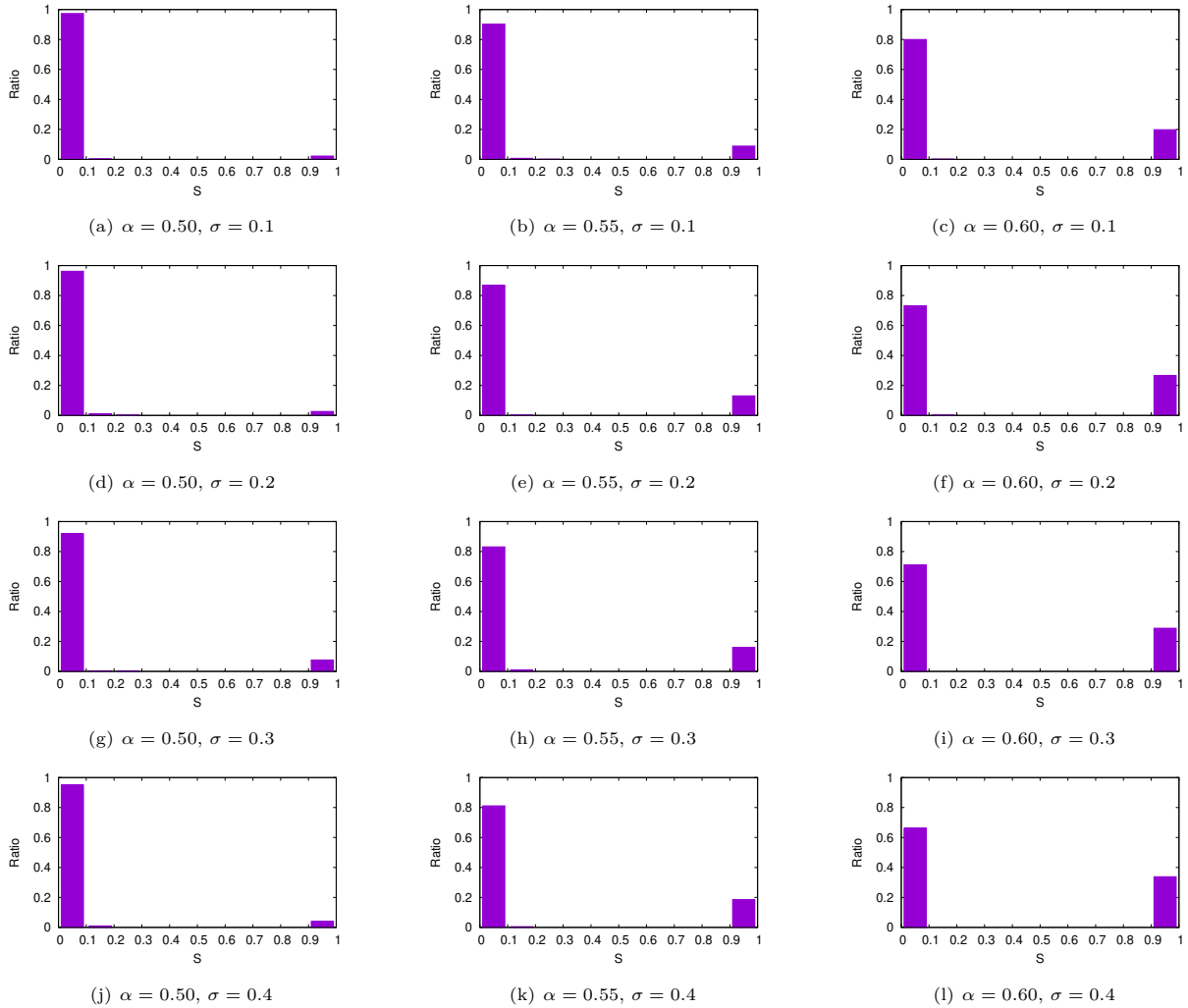


Figure 4 The frequency distributions of the final diffusion rates S calculated from 1,000 random networks.

新約聖書を対象とした言語ネットワークの解析

眞岸 祈平[†] 松本 朋子^{††} 島田 裕^{†††} 池口 徹^{†††}

[†] 東京理科大学 工学研究科 情報工学専攻

〒125-8585 東京都葛飾区新宿 6-3-1

^{††} 東京理科大学 理学部第一部 教養学科

〒162-0081 東京都新宿区神楽坂 1-3

^{†††} 埼玉大学 大学院理工学研究科 数理電子情報部門

〒338-8570 埼玉県さいたま市桜区下大久保 255

^{†††} 東京理科大学 工学部 情報工学科

〒125-8585 東京都葛飾区新宿 6-3-1

E-mail: †magikihe@hisenkei.net

あらまし 本稿では、文章を単語の共起に着目したネットワークとして表現することで、言語の文法構造がネットワークの構造的特徴にどのように現れるのかを調査した。その際、インド・ヨーロッパ語族を中心とする9種類の言語を対象とし、同じ内容の文章が異なる言語で表現されている新約聖書を使用した。また、使用する文章をネットワークに変換する手法を2種類導入し、それぞれから得られたネットワークに対して解析を行った。その結果、各言語は異なるネットワークの構造的特徴を有していること、異なる言語でもネットワークの構造的特徴が類似している言語があるということが明らかとなった。また、同一語族、語派に属していても、ネットワークの構造的特徴が類似している言語と類似していない言語があることも明らかとなった。

キーワード 言語ネットワーク, 新約聖書, 共起, 語族

Analysis of language networks for the New Testament

Kihei MAGISHI[†], Tomoko MATSUMOTO^{††}, Yutaka SHIMADA^{†††}, and Tohru IKEGUCHI^{†††}

[†] Department of Information and Computer Technology, Graduate School of Engineering, Tokyo University of Science 6-3-1 Nijuku, Katsushika-ku, Tokyo, 125-8585, Japan

^{††} Liberal Arts Department, Faculty of Science Division I, Tokyo University of Science 1-3 Kagurazaka, Shinjuku-ku, Tokyo, 162-0081, Japan

^{†††} Department of Information and Computer Sciences, Graduate School of Science and Engineering, Saitama University

255 Shimo-Okubo, Sakuraku, Saitama-shi, Saitama, 338-8570, Japan

^{†††} Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science 6-3-1 Nijuku, Katsushika-ku, Tokyo, 125-8585, Japan

E-mail: †magikihe@hisenkei.net

Abstract In this paper, we investigate differences in structural features of co-occurrence networks of words which are generated from documents written in different languages. For the analysis, we use the New Testament, in which the same texts are expressed in different languages. We investigate the structural features of the language network of the New Testament written in nine different languages, mainly Indo-European languages. In this paper, we use two types of methods for transforming networks from the New Testament texts. The results show that each language has its own unique network structure and some languages, which belong to different language families, have similar structural features of their networks. In addition, we show that some languages do not have similar structural features of the network, even though they belong to the same language family.

Key words Language network, New Testament, Co-occurrence, Language family

1. はじめに

近年では、インターネットや SNS などの普及により世界中の国や地域からデータが発信されており、それらに対するビッグデータ解析が重要視されている。これに伴って、多くのテキストデータから定型化されていない文字情報を解析し、有用な知見や特徴を抽出するテキストマイニングの研究が盛んに行われている。その一つとして、単語の共起をネットワーク上での連結性として捉え、複雑ネットワークの観点から文章を解析する研究が注目されている。例えば、文章中の単語の共起関係を表したネットワークの解析 [1]、コミュニティ分解の実験に言語ネットワークを使用した解析 [2]、小説を対象とした言語間のネットワーク解析 [3] などがある。これらの報告は、文章をネットワークとして捉え、クラスタ係数 [4]、平均頂点間距離 [4] などのネットワークの定量化指標を用いて、ネットワークの構造的特徴を定量的に調査している。しかし、ネットワーク間の構造的特徴の差異を定量化し、言語の特徴を比較することも重要である。

本稿では、言語による文法構造に着目し、ネットワーク理論の観点から言語の特徴を定量的に調査する。具体的には、9 種類の言語で書かれた同じ内容の文章を使用し、言語間の文法構造の違いに着目した解析を行う。本稿では、同じ内容が異なる言語で表現されている文章の例として新約聖書を解析対象とした。なお、本稿で使用した 9 種類の言語は、インド・ヨーロッパ語族を中心とするギリシャ語、ラテン語、英語、ドイツ語、フランス語、ロシア語、日本語、フィンランド語、中国語である。これは、世界の言語の中でインド・ヨーロッパ語族のユーザ数が最大であることに加え、新約聖書の始まりがギリシャ語であり、ヨーロッパ圏での言語の違いを調査することが先決であるからである。

はじめに、使用する新約聖書の文章に対して形態素解析を行い、単語の原型 (形態素) および品詞の情報を取得する。次に、得られた単語の情報を用いて、単語の共起に着目したネットワークを作成する。本稿では、クラスタ係数と平均頂点間距離を用いて、言語間のネットワークの構造的特徴の比較を行った。また、異なるネットワーク間の構造的特徴の差異の定量化が可能なスペクトルグラフ距離 [5] を用いて、言語間のネットワークの構造的特徴の差異について調査した。

その結果、同一言語で記述された文章は類似したネットワークの構造的特徴を有していること、インド・ヨーロッパ語族の中でも言語によって異なるネットワークの構造的特徴を示す言語があることが明らかとなった。また、異なる語族に分類されていても、ネットワークの構造的特徴が類似している言語が存在するということが明らかとなった。

2. 使用データ

本稿では、分析対象として新約聖書の福音書 [7]–[11] を使用した。具体的には、マタイによる福音書 (以下、マタイ伝)、マルコによる福音書 (以下、マルコ伝)、ルカによる福音書 (以下、ルカ伝)、ヨハネによる福音書 (以下、ヨハネ伝) の全 4 種類である。新約聖書の福音書を使用する理由は、文章の対応を取ることができることに加えて、文章の内容にストーリー性があるため、訳者の違いによる内容の差異が出にくいからである。

本稿で解析対象とした言語が属する語族、語派 [12]、語順 [13] を表 1 に、使用した新約聖書のデータを表 2 示す。表 2 における ID は、4 章以降の箇中で使用する。また、表 2 の括弧内の数値は第一版の作成年を表す。

表 1 本稿で解析対象とした言語の語族、語派、語順。

言語	語族	語派	語順
ギリシャ語	インド・ヨーロッパ語族	ギリシャ語派	SVO
ラテン語		イタリック語派	SOV
英語		ゲルマン語派	SVO
ドイツ語		ゲルマン語派	SVO
フランス語		イタリック語派	SVO
ロシア語		バルト語派	SVO
日本語	日本語族	日本語派	SOV
フィンランド語	ウラル語族	フィン語派	SVO
中国語	シナ・チベット語族	シナ語派	SVO

表 2 使用した新約聖書のデータ。

言語	ID	バージョン	作成年
ギリシャ語	GR1	NA28	2012(1898)
	GR2	UBS5	2014(1966)
ラテン語	La1	Bublia Sacra Vulgata	1955
英語	GB1	King James Version	1611
	GB2	English Standard Version	2001
ドイツ語	DE1	Lutherbible	1984
フランス語	FR1	Louis Segond	1910
	FR2	Segond 21	2007
ロシア語	RU1	Russian Synodal Version	1876
日本語	JP1	口語訳	1954
	JP2	新共同訳	1987
フィンランド語	FI1	Vanha Kirkkoraamattu	1642
	FI2	Uusi Kirkkoraamattu	1992
中国語	CN1	Chinese Standard Bible Simplified	2009

3. 言語ネットワーク作成手法

はじめに、文章を形態素解析し、単語の原型 (形態素)、品詞の情報を取得する。形態素解析ツールとして、日本語では mecab [14] を、他言語では tree-tagger [15] を使用した。分かち書きをしない日本語では、明確な単語の定義はないため、曖昧さや矛盾がなく、一義的にその単位を切り取ることが重要である。そのため、mecab には様々な辞書が用意されているが、本稿では unidic [16] を使用した。

分かち書きされた単語を一つの頂点とし、同一文章中で隣接する単語を枝で結ぶことで、重み無し無向ネットワークを作成した。ここで、単語の重複、同じ単語のペアのつながりの重複は認めず、句読点、半句読点、記号を除いた。また、句読点の前後の単語は枝で結ばないこととした。このネットワーク作成手法を手法 1 とする。手法 1 では、単語の活用を残したままネットワークを作成できるため、言語による活用の関係性をネットワークに反映することが可能である。手法 1 で作成された言語ネットワークの例を図 1 に示す。

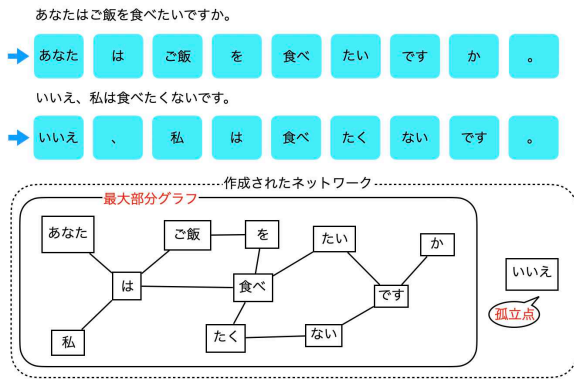


図1 手法1で作成された言語ネットワークの例。

次に、形態素解析で得られた単語の原型と品詞の組を一つの頂点とし、同一文章中で隣接する単語を枝で結ぶことで、重み無し無向ネットワークを作成した。このネットワーク作成手法を手法2とする。また、手法1と同様に、単語の重複、同じ単語のペアのつながりの重複は認めず、句読点、半句読点、記号を除いた。また、句読点の前後の単語は枝で結ばないこととした。手法2では、単語の品詞に着目した作成手法となるため、手法1よりも語順や文法構造をネットワークに反映することが可能である。手法2で作成された言語ネットワークの例を図2に示す。

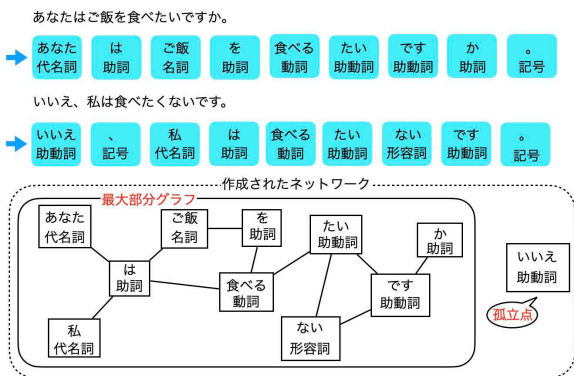


図2 手法2で作成された言語ネットワークの例。

なお、これらのネットワークの作成手法では、1つの文章が2つ以上のネットワークに分裂することも考えられる。しかし、このような場合の多くは、最大部分グラフと非常に小さいネットワークまたは孤立点で、全体のネットワークが構成されている。小さいネットワークと孤立点は、本稿の分析対象とする文章中の単語の共起関係の特徴を捉える上では本質的ではないと考えられる。したがって、本稿では文章中の最大部分グラフのみを抽出し、解析を行った。

4. 新約聖書から作成した言語ネットワークの解析

4.1 頂点数と枝数の関係

はじめに、新約聖書の文章で使用されている単語数について調査した結果を図3に示す。ここで、日本語における単語数とは文章を形態素解析した際に得られる、分かち書きされた単語の数である。

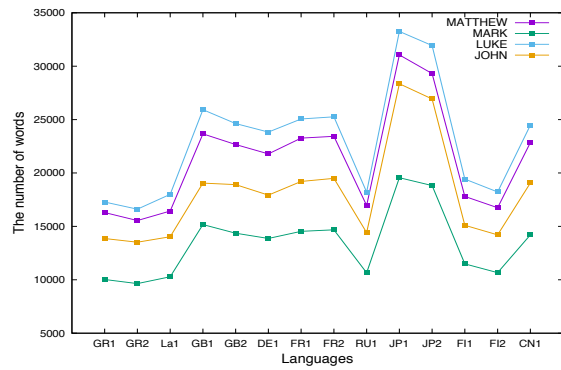


図3 マタイ伝 (MATTHEW)、マルコ伝 (MARK)、ルカ伝 (LUKE)、ヨハネ伝 (JOHN) の文章中で使用されている単語数。横軸は表2のIDである。

図3を見ると、同じ内容の文章を異なる言語で表現しているため、各福音書間での単語数は同じ傾向を示していることがわかる。また、同じ内容の文章を構成するのに、日本語は特に多くの単語数を必要とする言語であること、英語、ドイツ語、フランス語、中国語も比較的単語数を多く必要とする言語であることがわかる。一方、ラテン語、フィンランド語、ロシア語、ギリシャ語では、比較的少ない単語数で文章が構成されている。

語族、語派に着目すると、ヨーロッパ語族同士でも言語によって単語数が異なっていることがわかる。一方、ゲルマン派に属している英語、ドイツ語は単語数が類似していることがわかる。

次に、手法1、2で作成したネットワークについて、言語による頂点数、枝数の違いについて調査した。横軸を頂点数、縦軸を枝数とした結果を図4、5に示す。

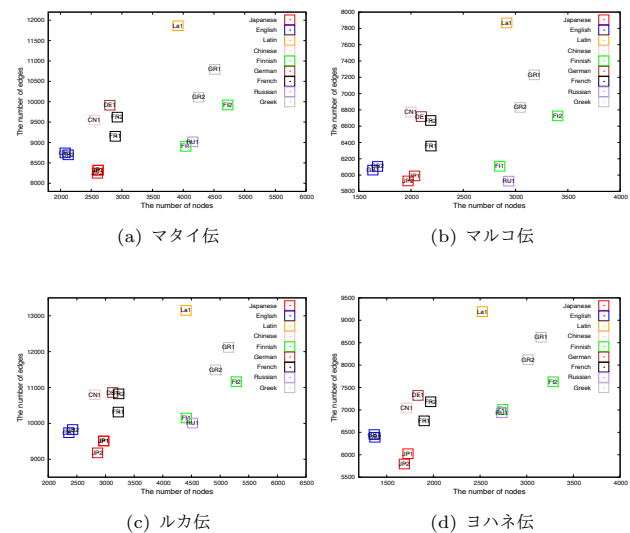


図4 各福音書を手法1により変換したネットワークの頂点数、枝数の関係。

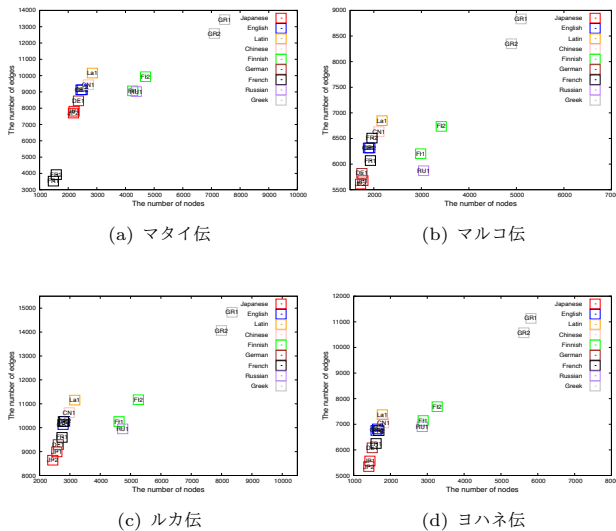
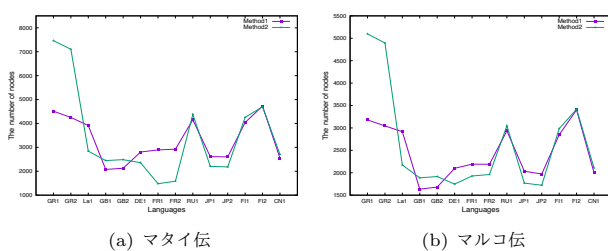


図5 各福音書を手法2により変換したネットワークの頂点数、枝数の関係。

図4, 5を見ると、日本語、英語、フランス語、中国語、ドイツ語は、比較的少ない頂点数、枝数でネットワークが構成されていることがわかる。一方、ギリシャ語、フィンランド語、ロシア語、ラテン語は頂点数、枝数を比較的多く必要とする言語であることがわかる。これらの結果は、図3の単語数の多さと逆の関係となっている。これは、日本語、英語、フランス語、中国語、ドイツ語が、同じ単語を繰り返し用いることで文章を表現している可能性を示唆するものである。また、頂点数、枝数の関係は同一語族であっても異なる傾向にある一方で、ゲルマン語派に属する英語とドイツ語には同様の傾向があることがわかる。

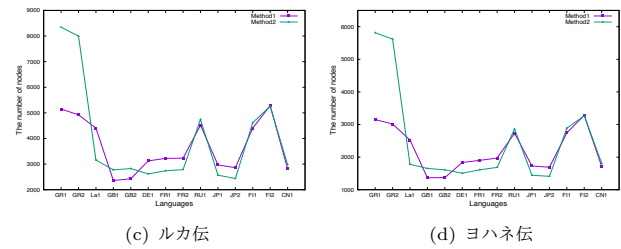
同言語に着目してみると、フィンランド語以外の同言語同士では頂点数、枝数はほとんど変化していないことがわかる。これは、同じ文章の異なる訳版であっても、訳し方に関わらず、言語の共起の特徴がネットワークの特徴として現れているためである。一方、難しい文法規則を使用しているとされるフィンランド語は、文脈による語尾の変化が多いとされている。この理由により、訳に依存してネットワークの頂点数、枝数に変化が起きていると考えられる。また、後述するネットワークの構造的特徴では、フィンランド語同士の文章は類似した構造を有していることを示している。

次に、ネットワーク作成手法の違いによる、頂点数の変化を表した結果を図6に示す。



(a) マタイ伝

(b) マルコ伝



(c) ルカ伝

(d) ヨハネ伝

図6 ネットワーク作成手法の違いによる頂点数の違い。

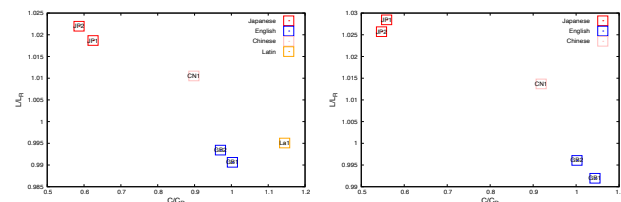
図6を見ると、日本語、ラテン語、ドイツ語、フランス語では手法1を用いた場合、頂点数が多くなっていることがわかる。これは、他の言語に比べて活用の種類や文章中で単語が活用される場合が多いためであると考えられる。また、イタリアック語派に属しているラテン語、フランス語で傾向が一致していることもわかる。一方、英語、ギリシャ語は、手法2を用いると頂点数が多くなるのがわかる。これは、同じ単語であっても異なる品詞となることが大きく影響していると考えられる。しかし、特に差が顕著に表れているギリシャ語では、一般的に受け入れられている、複雑な活用をする言語であるという評価とは逆の結果となった。また、フィンランド語、ロシア語は、ネットワーク作成手法による頂点数の違いは顕著に現れなかった。

4.2 クラスタ係数と平均頂点間距離の関係

次に、手法1, 2で作成したネットワークのクラスタ係数(以下, C) [4], 平均頂点間距離(以下, L) [4]を調査した。ここで, C, L はともに、元のネットワークをランダム化して得られたネットワークの C_R, L_R で正規化している。

ランダム化は、次数を保存した状態でランダム化を行う手法を使用した。本稿では、1つのランダムネットワークを作成するためのランダム化の試行回数を200回とし、作成されたランダムネットワークから C_R, L_R を導出した。

ただし、言語によっては、元のネットワークからランダム化して得られたネットワークが多くの場合で分裂してしまい、 L_R を計算できないことがある。これは、元のネットワークの言語による構造の特徴が、ランダム化の際に分裂しやすい構造を持つ可能性があると考えられる。そこで本稿では、連結なランダムネットワークが得られる言語に対してのみ $C/C_R, L/L_R$ を導出した。なお、連結なランダムネットワークが得られた言語は、図4, 5における頂点数、枝数が比較的少ない言語が多かった。横軸を C/C_R , 縦軸を L/L_R とした結果を図7, 8に示す。



(a) マタイ伝

(b) マルコ伝

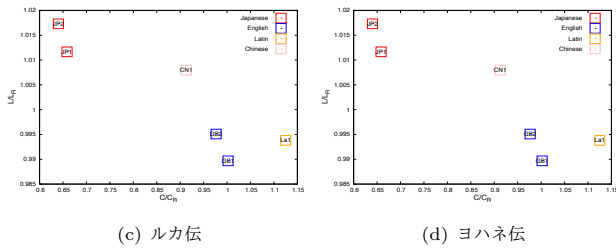


図 7 クラスタ係数, 平均頂点間距離の関係(手法 1).

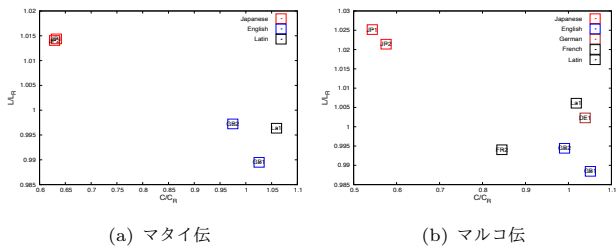


図 8 クラスタ係数, 平均頂点間距離の関係(手法 2).

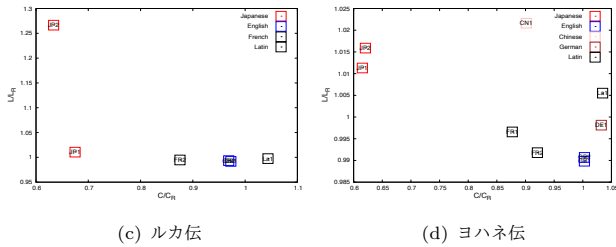


図 8 クラスタ係数, 平均頂点間距離の関係(手法 2).

図 7, 8 を見ると, 日本語は C が小さく, L が大きい傾向を示していることがわかる. このことから, 日本語は単語を使用するための語順の規則が少なく, 使用される単語同士の関連性が少ないことが, ネットワークの構造的特徴に影響していると考えられる. 一方, 英語, ラテン語, フランス語, ドイツ語は C が大きく, L が小さい傾向を示している. このことから, 英語, ラテン語, フランス語, ドイツ語は, 文章を構成するための文法規則が日本語より多く, 単語同士の関連性が強いと考えられる. また, ゲルマン語派に属する英語, ドイツ語, イタリア語派に属するラテン語, フランス語で同様の傾向にあることもわかる.

図 7, 8 のネットワーク作成手法の違いを比較しても, 作成されたネットワークの特徴量の言語間の違いはほとんど変わらない結果となった. このことから, 活用に着目した手法 1, 品詞に着目した手法 2 に関わらず, 作成されたネットワークの構造的特徴に, 言語の特徴が反映されていることが示唆される.

4.3 スペクトルグラフ距離に多次元尺度法を適用した相対的な位置関係

次に, 手法 1, 2 で作成したネットワークについて, 言語によるネットワーク間の構造的特徴の差異をスペクトルグラフ距離 [5] を用いて定量化した. また, 得られたスペクトルグラフ距離に対して, 多次元尺度法 [17] を適用することで, 距離情報をできるだけ保持した状態で, ネットワークの相対的な位置関係を可視化した. 言語間のネットワークの相対的な位置関係を

図 9, 10 に示す.

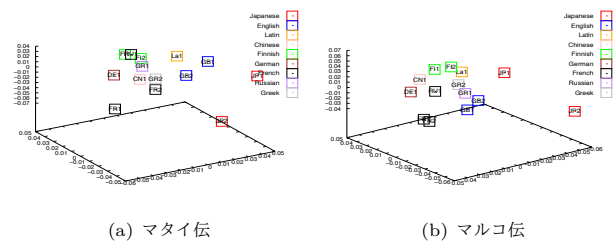


図 9 ネットワークの相対的な位置関係(手法 1).

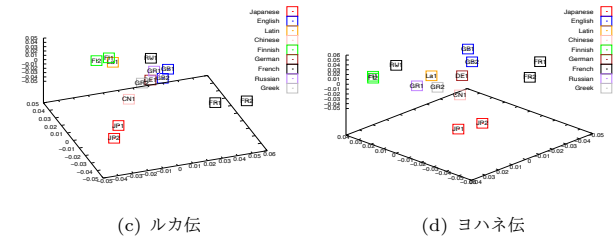


図 9 ネットワークの相対的な位置関係(手法 1).

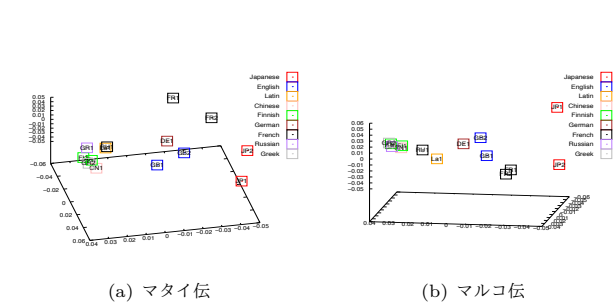


図 10 ネットワークの相対的な位置関係(手法 2).

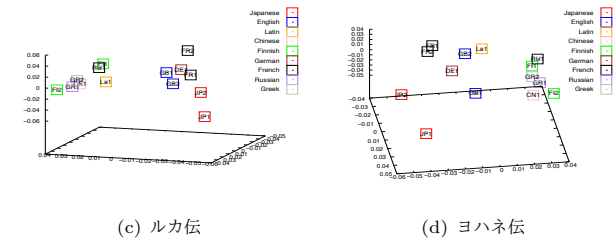


図 10 ネットワークの相対的な位置関係(手法 2).

図 9, 10 を見ると, どの言語でも同一言語同士で比較的類似したネットワークの構造的特徴を有していることがわかる. また, 日本語は, 他言語からの位置が離れているが, これは他言語にはない特徴的なネットワーク構造を日本語が有していることを示唆するものである. この結果は, 図 7, 8 で示した, 日本語は他言語に比べてクラスタ係数が小さく, 平均頂点間距離が大きい言語であるという結果と一致している.

一方, フランス語は, 文章およびネットワーク作成手法の違いによって, ネットワークの構造的特徴が他言語と類似する場合としない場合があることがわかる. フランス語同士は類似したネットワークの構造的特徴を有していることから, フランス語には単語の活用や語順などの文法規則以外にも, ネットワークの構造的特徴に影響を与える言語の特徴が存在している可能性が示唆される.

また, ギリシャ語, ロシア語, フィンランド語, ラテン語,

中国語は、比較的類似したネットワークの構造的特徴を有していることがわかる。しかし、ギリシャ語、ロシア語、ラテン語はともにインド・ヨーロッパ語族に属していることと一致した結果となったが、フィンランド語、中国語はともに異なる語族の言語であるため、語族分類とは反する結果となった。

ゲルマン語派に属する英語、ドイツ語は、手法1ではネットワークの構造的特徴が類似している場合としていない場合があるという結果となった。一方、品詞に着目した手法2では顕著に類似したネットワークの構造的特徴を有しているという結果となった。これらの言語は、頂点数、枝数の関係、クラスタ係数、平均頂点間距離の関係でも類似した言語であることがわかる。このことから、英語、ドイツ語は品詞に着目した語順などの文法規則が類似しているという点で、同じ語族に分類することができると思われる。

また、イタリック語派に分類されているラテン語、フランス語は、本稿の解析では異なるネットワークの構造的特徴を有しているという結果となった。この理由は前述した通り、フランス語にネットワークの構造的特徴に影響を与える言語の特徴が存在していることが原因であると考えられる。

表1の文法に着目すると、日本語は他言語と比較して、異なるネットワークの構造的特徴を持つという結果と一致したが、ラテン語では他言語と類似する結果となった。

4.4 ネットワークの特徴量と語族、語派との関係性

頂点数、枝数、クラスタ係数、平均頂点間距離の語族と語派の分類との関係性を、ネットワーク作成手法で比較した結果を表3に示す。ここで、インド・ヨーロッパ語族を語族1、日本語族を語族2、フィンランド語族を語族3、シナ・チベット語族を語族4と表した。また、ネットワーク作成手法1、2での頂点数をそれぞれ N_1 、 N_2 とした。表3では、手法の違いによる頂点数の比、クラスタ係数、平均頂点間距離が1を超えている場合を+、超えていない場合を-として示した。また、表3の*は値がないことを表す。

表3 ネットワークの特徴量と語族、語派との関係。

言語	語族	語派	語順	N_1/N_2	C/C_R 手法1	C/C_R 手法2	L/L_R 手法1	L/L_R 手法2
ギリシャ語	語族1	ギリシャ語派	SVO	+	*	*	*	*
ラテン語		イタリック語派	SOV	-	+	+	-	±
英語		ゲルマン語派	SVO	+	±	±	-	-
ドイツ語		ゲルマン語派	SVO	-	*	+	*	±
フランス語		イタリック語派	SVO	-	*	±	*	±
ロシア語		バルト語派	SVO	+	*	*	*	*
日本語	語族2	日本語派	SOV	-	-	-	+	+
フィンランド語	語族3	ウラル語派	SVO	+	*	*	*	*
中国語	語族4	シナ語派	SVO	+	-	-	+	+

表3を見ると、各言語でネットワークの構造的特徴が異なっていることがわかる。また、同語族、同語派であってもネットワークの構造的特徴が類似していないこともわかる。これらのことから、本稿で扱ったネットワーク特徴量以外にも、言語のネットワークの構造的特徴があることが示唆される。

5. まとめ

本稿では、異なる言語の文章を単語の共起ネットワークへ変換することで、言語が持つネットワークの構造的特徴の差異について調査した。その結果、インド・ヨーロッパ語族に属しているにも関わらず、他の語族の言語と類似したネットワークの

構造的特徴を有している言語がある一方で、一般的に知られている語族の分類と一致する言語もあることが明らかとなった。また、ゲルマン語派に属する英語、ドイツ語では、言語から作成されるネットワークの構造的特徴は類似しているという結果が顕著に現れた。

また、今回のネットワーク作成手法では活用と品詞に着目した2種類の手法を使用した。フランス語のように、この2つだけではまだ言語の特徴を捉えきれていない可能性があるということも明らかとなった。

本稿での解析結果では、言語学上で知られている言語と同様の絶対的分類はできなかったが、相対的な言語間の違いがあるということが明らかとなった。

本稿では、ネットワーク作成手法を変更することで、まだ知られていない文法規則を見つけられる可能性があることを示している。また、言語学で示されている語族や語派の分類に対して、ネットワークの観点から新たな言語学への視座を提案することができる可能性を示している。

謝辞 本研究の一部は、JSPS 科研費 (No. 15KT0112, 17K00348, 18K12701, 18K18125) の助成を受けて行われた。また、言語資料の収集に関してアドバイスをくださった Sviridova Nina 助教、毛福佳氏、ジョロエン氏に深く感謝申し上げます。

文 献

- [1] Ramon Ferrer i Cancho and Richard V. Solé, The small world of human language, *Proceedings of the Royal Society of London B: Biological Sciences*, Vol. 268, No. 1482, pp. 2261–2265, 2001.
- [2] Mark EJ Newman, Finding community structure in networks using the eigen-vectors of matrices, *Physical Review E*, Vol. 74, No. 3, 036104, 2006.
- [3] Yutaka Shimada, Mayumi Tatara, Kantaro Fujiwara, and Tohru Ikeguchi, Formation mechanisms of local structures in language networks, *Europhysics Letters*, Vol. 127, No. 5, 56003, 2019.
- [4] Duncan J. Watts and Steven H. Strogatz, Collective dynamics of small-world networks, *Nature*, Vol. 393, No. 6684, pp. 440–442, 1998.
- [5] Yutaka Shimada, Yoshito Hirata, Tohru Ikeguchi, and Kazuyuki Aihara, Graph distance for complex networks, *Scientific Reports*, Vol. 6, 34944, 2016.
- [6] https://www.bible.or.jp/read/vers_search.html (最終閲覧日: 2020年03月24日)
- [7] <https://www.academic-bible.com> (最終閲覧日: 2020年04月13日)
- [8] <http://www.cerbaso.org/> (最終閲覧日: 2020年04月16日)
- [9] <http://www.nic.funet.fi/> (最終閲覧日: 2020年04月20日)
- [10] <https://www.biblegateway.com> (最終閲覧日: 2020年04月26日)
- [11] <https://www.bebli.com/> (最終閲覧日: 2020年04月28日)
- [12] <https://www.ethnologue.com/> (最終閲覧日: 2020年05月12日)
- [13] <https://www.pnas.org/> (最終閲覧日: 2020年05月13日)
- [14] <http://taku910.github.io/mecab/> (最終閲覧日: 2020年01月04日)
- [15] <https://www.cis.uni-muenchen.de/schmid/tools/TreeTagger/> (最終閲覧日: 2020年04月27日)
- [16] <https://unidic.ninjal.ac.jp> (最終閲覧日: 2020年02月10日)
- [17] Michael AA Cox, and Trevor F. Cox. Multidimensional scaling. *Handbook of data visualization*, pp. 315–347, 2008.

決定論的非線形力学系からのマーク付点過程の生成と 力学系の特徴保存に関する検討

澤田 和弥[†] 島田 裕^{††} 池口 徹^{†,†††}

[†] 東京理科大学 大学院工学研究科 経営工学専攻
〒125-8585 東京都葛飾区新宿 6-3-1

^{††} 埼玉大学 大学院理工学研究科 数理電子情報部門
〒338-8570 埼玉県さいたま市桜区下大久保 255

^{†††} 東京理科大学 工学部 情報工学科
E-mail: †sawada@hisenkei.net

あらまし 様々な時系列データが観測されるようになってきている。また、時系列データにはマーク付き点過程として観測されるものも存在する。例えば、株式取引や地震データはマーク付き点過程の一種である。確率的なマーク付き点過程生成法は種々提案されているが、決定論的ダイナミクスに従うマーク付き点過程生成法は十分に検討されていない。決定論的ダイナミクスにしたがってマーク付き点過程を生成する手法の一つに、決定論的力学系から作成した観測時系列の極大値を抽出する手法がある。しかし、極大値を抽出することで元の力学系の特徴がどの程度保存されるのかについて検討することが重要である。そこで本稿では、極大値抽出により生成したマーク付き点過程が元の力学系の特徴を保持しているかについて検討した。具体的には、元の力学系のアトラクタとマーク付き点過程の再構成アトラクタでの2点間距離分布のKL情報量、JS情報量によって力学系の特徴を評価した。

キーワード 点過程, マーク付き点過程, KL情報量, JS情報量, ローレンツ方程式, Izhikevichモデル

On Investigation of Invariant Characteristics of Deterministic Nonlinear Dynamical Systems and Marked Point Processes

Kazuya SAWADA[†], Yutaka SHIMADA^{††}, and Tohru IKEGUCHI^{†,†††}

[†] Department of Management Science, Graduate school of Engineering, Tokyo University of Science
Nijuku 6-3-1, Katsushika-ku, Tokyo, 125-8585, Japan

^{††} Department of Information and Computer Sciences, Graduate School of Science and Engineering, Saitama University
255 Shimo-Okubo, Sakura-ku, Saitama-shi, Saitama 338-8570, Japan

^{†††} Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science
E-mail: †sawada@hisenkei.net

Abstract In recent years, various kinds of time series data have been observed. One of the kinds is a marked point process. For example, stock trading and seismic data are typical examples of marked point process. Although various stochastic methods for generating marked point processes have been proposed, a method for generating marked point processes that follows deterministic dynamics has not been sufficiently discussed. One of the methods for generating marked point processes according to deterministic dynamics is to extract maxima of observed time series produced from a nonlinear dynamical system. In this article, we investigate whether the marked point processes generated from the nonlinear deterministic dynamical system by the maximum extraction preserve the characteristics of the original dynamical system. To evaluate the characteristics of the dynamical system, we used KL and JS divergence of the distribution of distance between two points on attractors of the original dynamical system and the marked point process.

Key words point process, marked point process, KL divergence, JS divergence, Lorenz equation, Izhikevich neuron model

1. はじめに

様々な時系列データが高精度に観測されるようになっている。これらの時系列データの中には、イベントが発生したときのみデータが観測される点過程データも多く存在する。例えば、地震や株価変動、神経細胞の応答などはその典型である。神経細胞の応答はスパイク列と呼ばれ、神経細胞の振舞いを再現する様々な数理モデルが提案されている。また、点過程データは、例えば株価変動や地震などのデータのように付加情報を有する、マーク付き点過程データである場合も多い。マーク付き点過程データを生成するモデルとして、イベントタイミングやマーク値に確率的な振舞いを仮定した様々な数理モデルが提案されている [1]。

一方、株価変動はカオス性を有すること、従って、短期的な予測が可能であることを示唆する結果も報告されている [2]。即ち、決定論的非線形力学系に従うマーク付き点過程データの生成モデルを検討することは重要である。しかし、決定論的非線形力学系からマーク付き点過程データを生成する手法は十分に検討されていない。

そこで本稿では、連続力学系の極大値抽出によって生成したマーク付き点過程が元の力学系の特徴をどの程度保持しているのかについて調査する。具体的には、決定論的非線形力学系から生成された時系列データの極大値を抽出する [3] ことによって生成したマーク付き点過程から再構成したアトラクタ上の 2 点間距離分布と元の力学系のアトラクタ上の 2 点間距離分布を比較することで、2 点間距離分布の情報がどの程度保持されているのか評価した。2 つの 2 点間距離分布の比較には、Kullback-Leibler (KL) 情報量 [4] と Jensen-Shannon (JS) 情報量 [5] を用いた。また、本稿の数値実験には Lorenz 方程式 [6] と Izhikevich モデル [7] を使用した。

2. 極大値抽出によるマーク付点過程の生成

まず、本稿で用いた極大値抽出によるマーク付き点過程の生成 [3] について述べる。極大値を抽出することでマーク付き点過程を生成する様子を図 1 に示す。図 1 に示すように、任意の時系列信号の n 番目の極大値の時刻をイベント発生時刻 $t(n)$ と定義し、極大値をマーク情報 $M(n)$ と定義することで、マーク付き点過程を生成する。このとき n 番目のイベント発生間隔 (IEI) は、 $T(n) = t(n+1) - t(n)$ で定義できる。本稿では極大値抽出に

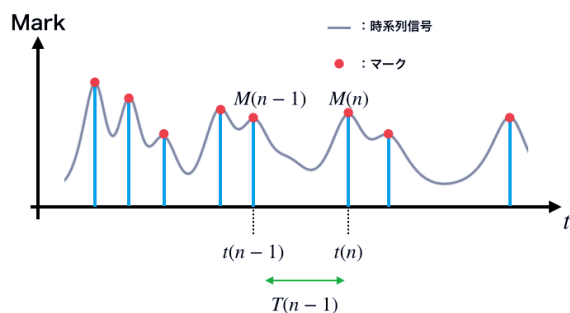


図 1 極大値抽出によるマーク付き点過程生成法

よるマーク付き点過程の対象として、Lorenz 方程式 [6] を用いた。Lorenz 方程式を式 (1) に示す。

$$\begin{cases} \frac{dx}{dt} = \sigma(-x + y) \\ \frac{dy}{dt} = -xz + rx - y \\ \frac{dz}{dt} = xy - bz \end{cases} \quad (1)$$

本稿では、Runge-Kutta 法により解を計算したが、その際の刻み幅は $h = 0.01$ とした。パラメータは $\sigma = 10$, $b = 8/3$ とし、 r は $24 \leq r \leq 100$ の範囲で 1 刻みで変化させた。ここで、カオス応答となる $r = 28$ の場合の z の時系列と、極大値抽出によるマーク付き点過程を図 2 に示す。図 2 を見ると、時系列信号とマーク付き点過程はカオス的な振舞いをしていることが確認できる。

次に、ニューロンモデルに対しても極大値を抽出することによってマーク付き点過程を生成することを考える。本稿では、Izhikevich ニューロンモデル [7] を用いてマーク付き点過程を生成する。Izhikevich ニューロンモデルを式 (2) に示す。

$$\begin{cases} \dot{v} = 0.04v^2 + 5v + 140 - u + I \\ \dot{u} = a(bv - u) \end{cases} \quad (2)$$

ここで、 $v \geq 30$ のとき、式 (3) で v, u はリセットされる。

$$\begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \quad (3)$$

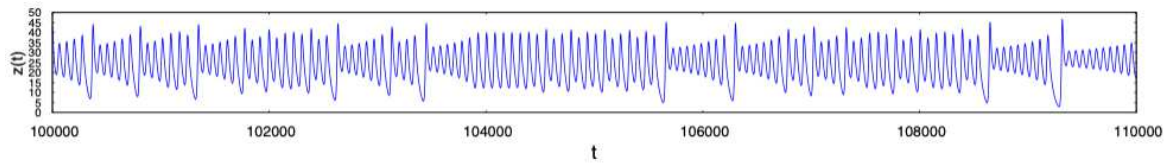
また、 v は膜電位、 u は回復変数、 I は外部入力、 a, b, c, d はパラメータである。解を計算する際の Runge-Kutta 法の刻み幅は $h = 0.01$ に設定した。本稿では、イベント発生タイミングを v, u が式 (3) でリセットされた時刻、マークの値をリセットされた時刻の u^2 としてマーク付き点過程を生成した。回復変数は発火時に極大値を取るため、極大値抽出していることに相当するが、マーク値の差をより大きくするためにマーク値を u^2 としている。ここで、カオス応答をする $a = 0.2, b = 2, c = -56, d = -16, I = -99$ のときの膜電位時系列とマーク付き点過程を図 3 に示す。図 3 から、膜電位時系列信号および、マーク付き点過程はカオス的な応答をしていることが確認できる。

3. 状態空間の再構成

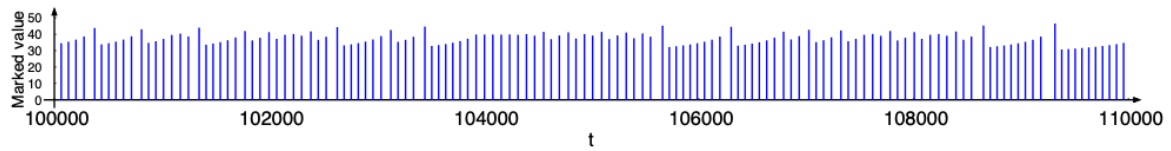
遅延座標系を用いた状態空間の再構成 [8] について述べる。時系列信号 $x(t)$ に対する遅延座標系を用いた状態空間の再構成は式 (4) で定義される。ここで、 τ は遅れ時間、 m は再構成次元を表す。

$$v(t) = (x(t), x(t + \tau), x(t + 2\tau), \dots, x(t + (m - 1)\tau)) \quad (4)$$

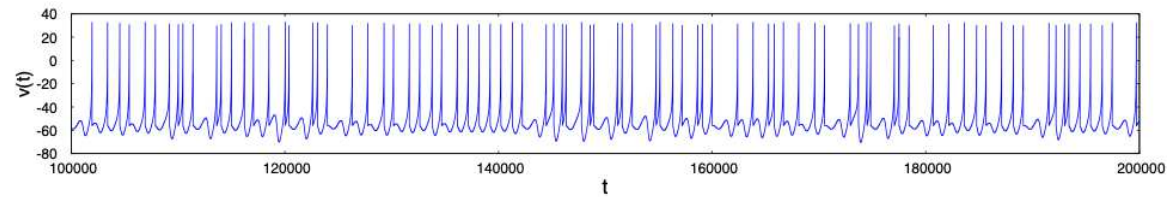
式 (4) にしたがって、十分大きい次元で遅延座標系による再構成を行うと、全変数が観測されていない場合でも元の力学系のダイナミクスを再構成できることが知られている [9]。本稿では、任意の力学系から極大値抽出によってマーク付き点過程を



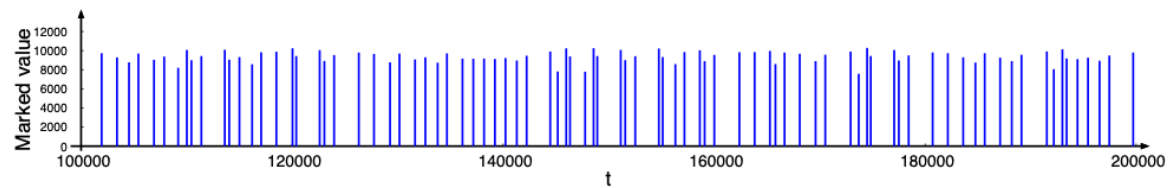
(a) 時系列信号



(b) マーク付き点過程

図2 $z(t)$ の時系列信号と極大値抽出により生成したマーク付き点過程

(a) 膜電位時系列信号



(b) マーク付き点過程

図3 膜電位時系列信号と生成したマーク付き点過程

生成しているが、極大値抽出はポアンカレ断面を取ることに対応する。つまり、極大値を抽出して生成したマーク付き点過程は元の力学系の情報を抽出していることになる。したがって、得られたマーク付き点過程を遅延座標系によって再構成することで、元の力学系のダイナミクスを捉えることが可能であると考えられる。また、イベント間隔からでも元の力学系を再構成可能であることもわかっている [10]。つまり、IEI を式 (4) で再構成することで元の力学系を再構成できる。式 (4) は観測時系列が 2 変数の場合に拡張しても再構成が可能である [11]。そこで、 n 番目の IEI を $T(n)$ 、 n 番目のマーク値を $M(n)$ とし、以下の式 (5) により再構成する。

$$\mathbf{V}(t) = (T(t), M(t), \dots, T(t + (m-1)\tau), M(t + (m-1)\tau)) \quad (5)$$

式 (5) において、 τ は遅れ時間で、 $2m$ が再構成次元となる。また、一般的には力学系の全変数を観測することは困難だが、本稿では数理モデルを用いているため、ローレンツ方程式の元の力学系として、 $x(t), y(t), z(t)$ の全変数を用いた 3 次元アトラクタを用いる。

4. 確率分布間の非類似度の定量化

マーク付き点過程の背後にある力学系の特徴がマーク付き点過程データにどの程度保持されているのかを定量的に評価する

ことを考える。我々はある力学系から極大値を抽出することによってマーク付き点過程を生成している。生成したマーク付き点過程から式 (4), (5) によって元の力学系を再構成できるとすれば、元の力学系の 2 点間距離分布と再構成した力学系の 2 点間距離分布は類似すると考えられる。したがって、2 点間距離分布の差を 4.1, 4.2 節で示す指標によって評価する。

4.1 Kullback-Leibler (KL) 情報量

まず、本稿では離散的な確率分布に対する Kullback-Leibler (KL) 情報量 [4] を用いる。KL 情報量は、任意の 2 つの確率分布がどの程度類似しているのかを定量化する指標である。確率分布 P の確率分布 Q に対する KL 情報量は以下の式 (6) で表される。

$$D_{KL}(P||Q) = \sum_{i=1}^n P(k_i) \log \left(\frac{P(k_i)}{Q(k_i)} \right) \quad (6)$$

ここで、 $P(k_i), Q(k_i)$ は確率分布において小さい方から i 番目の階級を取る確率、 n は全階級数を表す。また、任意の k_i において $Q(k_i) = 0$ の場合は計算が行えないため、本稿では $Q(k_i) = 0$ の場合は除いて計算を行なった。KL 情報量は非負性があり、通常は正の値を取るが、この処理によって、負の値も取ってしまう場合が存在する。また、KL 情報量には非対称性があり、一

一般的には $D_{KL}(P||Q) \neq D_{KL}(Q||P)$ であるため、結果の解釈が複雑となってしまふ。そこで本稿では、KL 情報量に比べて計算できない場合が少なく、2つの分布に対して一意に情報量が計算できる JS 情報量も用いた。

4.2 Jensen-Shannon (JS) 情報量

KL 情報量には非対称性があるが、任意の2つの分布に対して一意に情報量が決定する JS 情報量が提案されている [5]。確率分布 P と Q の JS 情報量の定義を式 (7) に示す。

$$D_{JS}(P, Q) = \frac{1}{2}(D_{KL}(P||M) + D_{KL}(Q||M)) \quad (7)$$

ただし、確率分布 M は確率分布 P と Q の平均の確率分布であり、 $M(k_i) = \frac{1}{2}(P(k_i) + Q(k_i))$ とする。KL 情報量では、任意の k_i において $Q(k_i) = 0$ が存在すると、計算が行えないために、負の情報量になってしまう場合があり、結果の信頼性が落ちてしまふ。しかし、JS 情報量は $M(k_i) = 0$ の場合、すなわち $P(k_i) = 0 \cap Q(k_i) = 0$ の場合のみ計算が行えなくなるため、KL 情報量に比べて信頼性の低い結果が出にくいと考えられる。また、JS 情報量には対称性があり、2つの分布から計算される情報量は一意に定まるため、結果の解釈が行いやすくなる。従って、本稿では KL 情報量に加えて JS 情報量も計算した。

5. 数値実験

本稿で行なった数値実験について述べる。本稿の目的は、極大値を抽出する変換によって、生成したマーク付き点過程データが元の力学系の空間的特徴をどの程度保持できているのかを定量的に評価することである。そこで、極大値を抽出することで生成したマーク付き点過程データと元の力学系のアトラクタの2点間距離分布の差を比較することで、元の力学系の情報がマーク付き点過程にどの程度保持されているかの調査を行なった。具体的には、元のアトラクタ上の2点間距離分布とマーク付き点過程の再構成アトラクタ上の2点間距離分布に対して KL 情報量と JS 情報量を計算した。その際、2点間距離は元のアトラクタと IEI、マーク値、IEI とマーク値の再構成アトラクタで範囲が異なり、単純に比較できないため、[0,1] の範囲で正規化した。また、マーク付き点過程の再構成次元は $m = 10$ 、遅れ時間は $\tau = 1$ とし、過渡状態として時系列から 100,000 点を除外した。2点間距離分布の計算に用いたアトラクタの点数は元の力学系、IEI、マーク値、IEI とマーク値から再構成したアトラクタの全てにおいて 10,000 点とした。また、確率分布は [0,1] の範囲で階級幅を 0.05、即ち階級数を $n = 20$ として求めた。

元のアトラクタ上の2点間距離分布と IEI、マーク値、IEI とマーク値から再構成したアトラクタ上の2点間距離分布の KL 情報量と JS 情報量を計算した結果を図4に示す。ここで、図4の横軸はローレンツ方程式のパラメータ r 、縦軸は KL 情報量、JS 情報量である。また、 P_0 は $x(t), y(t), z(t)$ を用いた3次元アトラクタにおける正規化2点間距離の確率分布である。 P_x, P_y, P_z はそれぞれ x, y, z の時系列の極大値を抽出して生成したマーク付き点過程から再構成したアトラクタにおける正規化2点間距離の確率分布である。

図4を見ると、KL 情報量は0から2程度の値を取っており、JS 情報量は概ね0から0.3程度の値を取っている。したがって、確率分布は比較的類似していると考えられる。また、 $r = 92, 93, 100$ のときの JS, KL 情報量は大きく変動しているが、カオス応答の場合とは異なり、このパラメータのときは周期軌道となっていることが原因だと考えられる。実際に $r = 92, 93, 100$ の周期軌道の場合と $r = 28$ のカオス軌道の場合のアトラクタを図5に示す。図5を見ると、 $r = 28$ の(a)はカオス的な軌道を描いているが、 $r = 92, 93, 100$ の(b)~(d)は周期軌道となっていることが確認できる。また、図4において、JS 情報量を比較すると、 $z(t)$ の時系列の極大値を抽出した場合に情報量の値が相対的に低くなっていることがわかる。これらの結果は、再構成次元 $m = 10$ 、遅れ時間 $\tau = 1$ とした場合であるため、他の条件でも同様の傾向となるか、またその原因についてはより詳細な調査が必要である。

次に、Izhikevich モデルから生成したマーク付き点過程に対しても Lorenz 方程式の場合と同様に JS 情報量を計算した。再構成次元 m 、遅れ時間 τ 、除外した過渡状態、用いた時系列長も同様である。また、本稿では、 a, b, c, d, I に関して5種類のパラメータを用いたが、それぞれ、(A) $a = 0.2, b = 2, c = -56, d = -16, I = -99$, (B) $a = 0.511, b = 0.2, c = -55, d = 2, I = 10$, (C) $a = 0.02, b = 0.498, c = -47.6, d = 2, I = 10$, (D) $a = 0.05, b = 0.2, c = -50.96, d = 2, I = 10$, (E) $a = 0.01, b = 0.61, c = -56, d = 2, I = 10$ である。ここで、(A), (B) はカオス応答するパラメータで、(C), (D), (E) は推定したりアプノフ指数が正の範囲であり、(C), (D) は Chattering と類似したカオスの発火パターンを示す場合、(E) は Regular Spiking と類似したカオスの発火パターンを示す場合である [12]。計算した JS 情報量の結果を表1に示す。表1を見ると、JS 情報量は0.1から

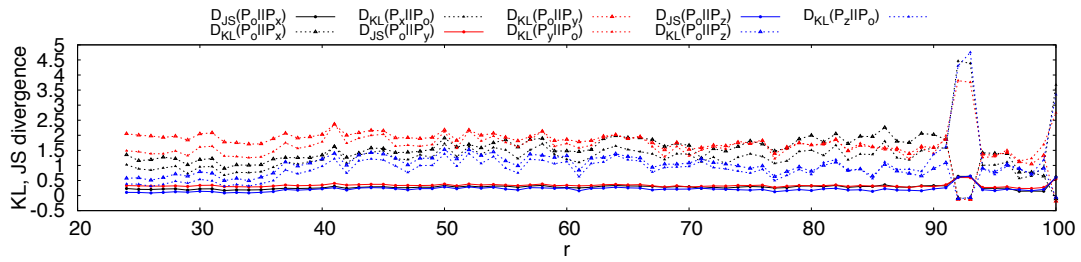
表1 Izhikevich モデルから生成したマーク付き点過程を用いて再構成したアトラクタの2点間距離分布と膜電位時系列から再構成したアトラクタの2点間距離分布間の JS 情報量

パラメータ	D_{JS}		
	マーク時系列	IEI 時系列	マークと IEI 時系列
(A)	0.465586	0.494907	0.459617
(B)	0.307387	0.310033	0.291579
(C)	0.143854	0.246331	0.165065
(D)	0.393453	0.376193	0.320988
(E)	0.412835	0.379912	0.290520

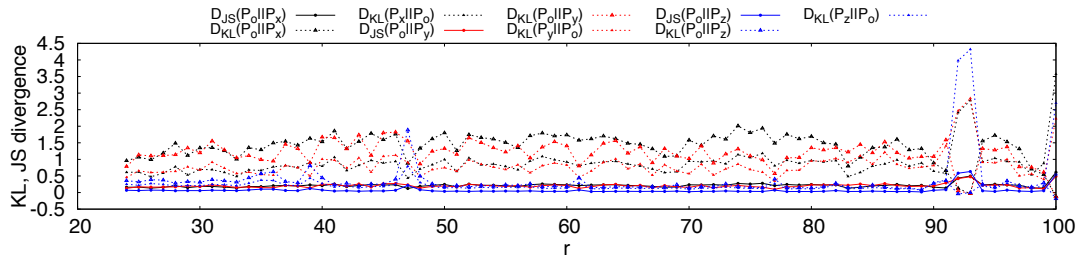
0.5の範囲の値を取っており、(C)のパラメータ以外では、マークと IEI 時系列から再構成した場合にもっとも低い値を取っている。

6. まとめ

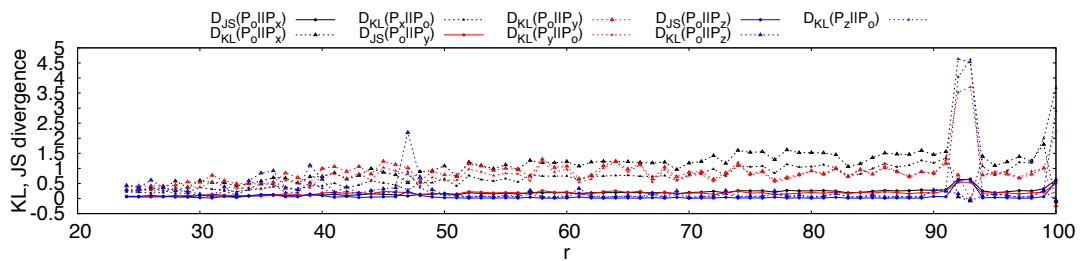
本稿では、決定論的非線形力学系を観測することで得た時系列信号の極大値を抽出して生成したマーク付き点過程データがどの程度元の力学系の空間的特徴を保持しているのかについて検討した。具体的には、元の力学系の2点間距離分布とマーク付き点過程の再構成状態空間の2点間距離分布に対して KL 情



(a) マーク時系列から再構成したアトラクタ上での 2 点間距離の確率分布との KL 情報量, JS 情報量



(b) IEI 時系列から再構成したアトラクタ上での 2 点間距離の確率分布との KL 情報量, JS 情報量



(c) マーク時系列と IEI 時系列から再構成したアトラクタ上での 2 点間距離の確率分布との KL 情報量, JS 情報量

図 4 r を変化させた場合の元のアトラクタ上の 2 点間距離の確率分布との KL 情報量, JS 情報量

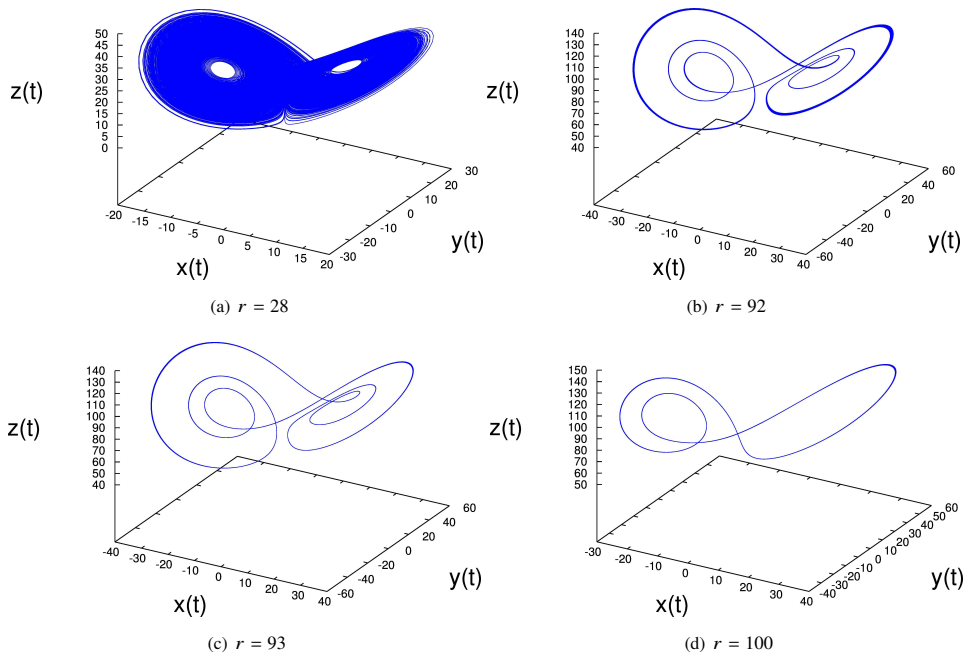


図 5 ローレンツ方程式のアトラクタ

報量, JS 情報量を計算することで, 変換前後の 2 点間距離分布の差を評価した. 結果として, 本稿の数値実験において得られ

た JS 情報量は概ね 0~0.5 程度であり, 2 点間距離分布という力学系の特徴はある程度保持できることが示唆された. 本稿の実

験では再構成の条件を固定したが、今後は条件の変更による詳細な調査が必要である。また、力学系の時間発展の特徴を評価するためには、リアプノフ指数や決定論性指標などを用いる必要もある。

謝辞 本研究の一部はJSPS 科研費 (No. 15KT0112, 17K00348, 18K18125) の援助を受けて行われた。

文 献

- [1] Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, Vol. 58, No. 1, pp. 83–90, 1971.
- [2] Yoshito Hirata and Kazuyuki Aihara. Timing matters in foreign exchange markets. *Physica A: Statistical Mechanics and its Applications*, Vol. 391, No. 3, pp. 760–766, 2012.
- [3] Naoki Yabuta and Tohru Ikeguchi. Prediction of high-dimensional multivariate information as an amplitude-event dynamical system. *Proceedings of NOLTA 2007*, pp. 188–191, 2007.
- [4] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, Vol. 22, No. 1, pp. 79–86, 1951.
- [5] Jianhua Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, Vol. 37, No. 1, pp. 145–151, 1991.
- [6] Edward Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, Vol. 20, No. 2, pp. 130–141, 1963.
- [7] Eugene M Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, Vol. 14, No. 6, pp. 1569–1572, 2003.
- [8] Norman Packard, James Crutchfield, Doyne Farmer, and Robert Shaw. Geometry from a time series. *Physical Review Letters*, Vol. 45, No. 9, pp. 712–716, 1980.
- [9] Floris Takens. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics*, pp. 366–381, 1981.
- [10] Tim Sauer. Reconstruction of dynamical systems from interspike intervals. *Physical Review Letters*, Vol. 72, No. 24, p. 3811, 1994.
- [11] Tim Sauer, James A Yorke, and Martin Casdagli. Embedology. *Journal of Statistical Physics*, Vol. 65, No. 3-4, pp. 579–616, 1991.
- [12] 内木楓, 島田裕, 藤原寛太郎, 池口徹. Izhikevich ニューロンモデルにおけるカオス応答とカオス同期. 電子情報通信学会論文誌 A, Vol. 100, No. 5, pp. 195–204, 2017.

コンタクトの特徴がテンポラルネットワーク上での情報拡散に与える影響の解析

Influence of Contact Properties to Information Diffusion on Temporal Networks

ジョ ロエン¹

澤田 和弥¹

島田 裕²

池口 徹^{1,3}

Luyan XU

Kazuya SAWADA

Yutaka SHIMADA

Tohru IKEGUCHI

東京理科大学 大学院工学研究科¹

埼玉大学 大学院理工学研究科²

東京理科大学 工学部³

1 はじめに

近年、人を頂点、人間関係を枝とした静的なネットワーク上での情報拡散に関する研究が盛んに行われている [1, 2]. 一方、現実世界には、静的なネットワークだけでなく、人と人のコンタクトを枝とするような構造が時間的に変化するテンポラルネットワーク (TN) も数多く存在する. 我々は既に、実世界の TN 上において、コンタクトの密度と情報拡散の間に関連性があることを報告している [3]. しかし、コンタクト密度が情報拡散にどの程度影響を与えるかについては明らかになっていない. また、両者のコンタクト回数と密度が等しい場合でも、コンタクトの順番が異なれば、拡散の効果も異なると考えられる. そこで本稿では、我々が提案した情報への関心度に着目した情報拡散モデル [3] を用いて、TN 上でのコンタクトの密度と順番が情報拡散に与える影響を調査する.

2 使用データとテンポラルネットワーク (TN)

本稿では、あるフランスの高校と病院で観測された、人と人のコンタクトのデータ [4-6] を用いた. 時刻 t での隣接行列 $A(t)$ の (i, j) 成分を $A_{ij}(t)$ とする. 時刻 t で実験参加者 i と j にコンタクトが存在する場合は、 $A_{ij}(t) = A_{ji}(t) = 1$, そうでない場合は、 $A_{ij}(t) = A_{ji}(t) = 0$ とすることで TN を作成した. 高校と病院でのデータはそれぞれ、126 名 (教員と学生) の 4 日間の対面接触記録と、75 名 (医療スタッフと患者) の 5 日間の対面接触記録である. なお、高校と病院でのコンタクトは 20 秒ごとに観測され、20 秒ごとの平均コンタクト数 (密度 ρ) はそれぞれ 2.09/[20 秒] と 1.87/[20 秒] である. 実験では、コンタクトの時系列を幅 $d[s]$ の窓に分割し、コンタクトの順番を保ちつつ各窓に 2 つのコンタクトが含まれるようにコンタクトを移動させたデータも作成し、比較対象としている. 20 秒あたりのコンタクト密度は $\rho = 40/d$ であり、 d が小さいほど ρ は大きくなる.

3 情報への関心度に着目した情報拡散モデル

時刻 t での頂点 i の情報への関心度を $x_i(t)$ とする. 関心度 $x_i(t)$ は、 $x_i(t+1) = g_i(\tau, t)x_i(t) + \alpha H(x_i(t) - \theta_i) \sum_{j \in G_i(t)} F(x_j(t) - \theta_j)$ に従って変化する. ここで、 t_i^* を頂点 i が最初に情報を受けとった時刻、 $\tau (\tau < 1)$ を情報への関心度の減衰係数とすると、 $g_i(\tau, t)$ は、 $t - t_i^* \equiv 0 \pmod{3600}$ のとき $g_i(\tau, t) = \tau$, それ以外のとき $g_i(\tau, t) = 1$ となる関数である. また、 $F(x)$ はステップ関数、 $H(x)$ は $x \geq 0$ のとき $H(x) = 0$, $x < 0$ のとき $H(x) = 1$ となる関数である. さらに、 $G_i(t)$ は時刻 t での頂点 i の隣接頂点の集合、 α は情報の影響力、 θ_i は頂点 i の閾値である.

4 数値実験

本実験では、 $\alpha = 0.05$, $\tau = 0.97$, $\theta_i = 0.5$ とした. 拡散開始の条件として、最初に発生したコンタクト (i, j) について

の頂点 i の関心度を、 $x_i(0) = 1$ とすることで情報を拡散させた. 本稿では、データの最終観測時刻までに少なくとも一度閾値に達した頂点の割合を最終拡散率 S とした. さらに、 d を $20 \leq d \leq 60$ の範囲で変化させることで、異なるコンタクト密度 ρ を有するデータから新たな TN を生成した. 高校と病院のデータについて、コンタクトが元の順番に従う場合と、コンタクトの順番をランダムにシャッフルした場合 (RS) の各 TN 上での S を図 1 に示す. 図 1 の破線は RS に対して求めた S の 100 回試行平均であり、灰色の範囲は S の最大値と最小値で囲まれた領域である.

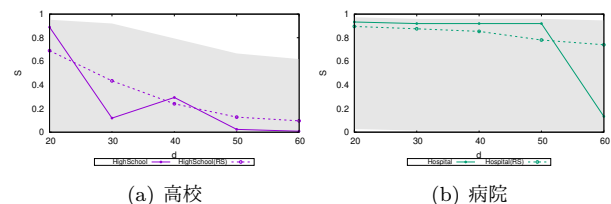


図 1: 時間間隔 d と最終拡散率 S の関係

図 1 から、 d が大きければ、 S が小さくなる傾向があることがわかる. この結果から、コンタクト密度 ρ が大きい場合、情報が広く拡散することがわかった. また、RS では、灰色の範囲が S の値の範囲 $[0, 1]$ の約 8 割以上を占めていることがわかる. この結果から、コンタクトの順番によって、情報がネットワークのほぼ全体まで広がる場合も、全く広がらない場合もあることが示唆された. これらの結果は、コンタクトの順番が情報拡散に大きな影響を与えていることを示唆するものである.

5 まとめ

本稿では、高校と病院でのコンタクトデータを用いて、コンタクトの密度と順番が情報拡散に与える影響について調査した. その結果、高校と病院のどちらでも、コンタクトの密度が大きい場合では、情報が広がりやすいことがわかった. また、コンタクトの順番が情報拡散に大きな影響を与えていることもわかった. この結果から、コンタクトの順番を適切に設定すれば、情報拡散を制御できる可能性がある. 今後の課題として、コンタクトの順番と情報拡散の関係を詳しく調査する必要がある.

なお、本研究の一部は JSPS 科研費 (Nos. JP15KT0112, JP17K00348, JP18K18125, JP20H00596) の援助を受けて行われた.

参考文献

- [1] D. J. Watts & S. H. Strogatz, *Nature*, **393**, 440-442, 1998.
- [2] D. Centola, *Science*, **329**, 1194-1197, 2010.
- [3] ジョ 他, 信学ソ大, N-1-11, 2020.
- [4] <http://www.sociopatterns.org/> (最終閲覧日: 2020 年 12 月 20 日)
- [5] P. Vanhems et al., *PLOS ONE*, **8**, e73970, 2013.
- [6] J. Fournet & A. Barrat, *PLOS ONE*, **9**, e107878, 2014.

複数の再構成アトラクタを用いたニューロンへの共通入力への推定

Estimation of Common Input Signals to Neuron Using Multiple Reconstructed Attractors

三浦 英¹

Ei Miura

澤田 和弥²

Kazuya Sawada

池口 徹^{1,2}

Tohru Ikeguchi

東京理科大学 工学部 情報工学科¹

東京理科大学 大学院工学研究科 経営工学専攻²

1 はじめに

ニューロンの出力信号を観測することに比べて、ニューロンへの入力信号を観測することは容易ではない。本稿では、複数のニューロンより得られる発火間隔 (ISI) 時系列から再構成したアトラクタ [1] を用いることで、ニューロンへの共通入力への推定を行う手法を提案する。

2 ニューロンモデルと入力信号

本稿では、ニューロンの数理モデルとして、Leaky Integrated and Fire (LIF) モデル [2] を用いる。LIF モデルのダイナミクスは、 $\tau \dot{V}(t) = -V(t) + S(t)$ と表現できる。ここで、 $V(t)$ は膜電位、 $S(t)$ は入力電流、 τ は漏れの割合を示すパラメータである。LIF モデルは、膜電位 $V(t)$ が閾値 Θ を超えると発火し、 $V(t)$ がリセットされる。ここで、 j 番目のニューロンの i 番目の発火時刻を $T_j(i)$ とすると、ISI 時系列は $t_j(i) = T_j(i+1) - T_j(i)$ である。本稿では、個体差のある複数のニューロンとして、 $\Theta = 20$, $1/\tau = 0.8, 0.9, \dots, 1.2$ とした 5 個の LIF モデルを用いた。さらに、ニューロンへの共通入力として、カオス応答を示すレスラー方程式 [3] の第 1 変数 $x(t)$ の時系列を用いて、 $S(t) = x(t) + 40$ とした。なお、レスラー方程式のパラメータは $a = 0.36, b = 0.4, c = 4.5$ とした。

3 ニューロンへの共通入力への推定手法

本稿では、以下に述べるように複数の再構成アトラクタの合成によりニューロンへの共通入力への推定を行う。そこで、再構成アトラクタの直径を揃えるため、以下の手順により処理を行う。(1) j 番目 ($j = 1, 2, \dots, 5$) のニューロンより得られた ISI 時系列 $t_j(i)$ を平均 0, 分散 1 として標準化した ISI 時系列 $t'_j(i)$ を得る。(2) 5 個の $t'_j(i)$ から m 次元遅延ベクトル $\mathbf{u}_j(i) = (t'_j(i), t'_j(i+1), \dots, t'_j(i+(m-1)L))$ を用いて [1], 5 つのアトラクタを再構成する。ここで、 L は遅れ時間である。(3) ベクトル $\mathbf{u}_j(i)$ を用いて再構成した 5 つのアトラクタから 1 つのアトラクタを合成する。具体的には、発火時刻 $T_j(i)$ とベクトル $\mathbf{u}_j(i)$ は対応しているので、5 つのベクトル $\mathbf{u}_j(i)$ を発火時刻 $T_j(i)$ の順番に並べ替え、アトラクタを合成した。図 1 は、 $j = 3, i = 1, 2$ とした場合の、ベクトル $\mathbf{u}_j(i)$ を対応する発火時刻 $T_j(i)$ の順番に並べ替える例である。(4) 合成したアトラクタのリカレンスプロット (RP) [4] を作成し、共通入力への推定した [5]。ここで、 $m = 5, L = 1$, RP の閾値はプロット割合が 10% となるように設定した。

$T_j(i)$	$\mathbf{u}_j(i)$	発火時刻 $T_j(i)$ の順に 並べ替える	$T_j(i)$	$\mathbf{u}_j(i)$
$T_3(1) = 3.2$	$\mathbf{u}_3(1)$		$T_2(1) = 2.8$	$\mathbf{u}_2(1)$
$T_1(2) = 5.3$	$\mathbf{u}_1(2)$	$T_1(1) = 3.2$	$\mathbf{u}_1(1)$	
$T_2(1) = 2.8$	$\mathbf{u}_2(1)$	$T_3(1) = 3.3$	$\mathbf{u}_3(1)$	
$T_2(2) = 5.5$	$\mathbf{u}_2(2)$	$T_1(2) = 5.3$	$\mathbf{u}_1(2)$	
$T_3(1) = 3.3$	$\mathbf{u}_3(1)$	$T_2(2) = 5.5$	$\mathbf{u}_2(2)$	
$T_3(2) = 7.4$	$\mathbf{u}_3(2)$	$T_3(2) = 7.4$	$\mathbf{u}_3(2)$	

図 1: 対応する発火時刻 $T_j(i)$ の順番でベクトル $\mathbf{u}_j(i)$ を並べ替える例

4 結果

共通入力の時系列 $x(t)$ から $m = 5, L = 15$ としてアトラクタの再構成を行い、そのアトラクタから作成した RP と合成したアトラクタの RP を図 2 に示す。ここで、図 2(左) は、

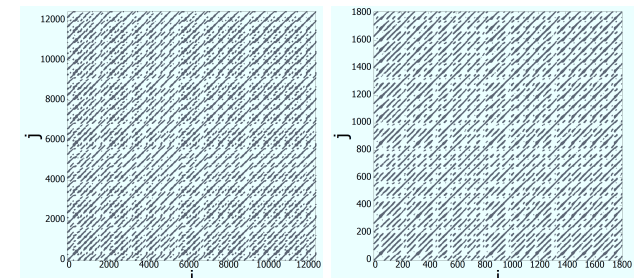


図 2: 共通入力 RP (左) と合成したアトラクタ RP (右)

共通入力の時系列 $x(t)$ を 1/20 でサブサンプルして作成した RP である。図 2 を見ると、共通入力 RP と合成したアトラクタ RP は類似することがわかる。また、共通入力の時系列 $x(t)$ (紫色) と図 2 (右) に示した RP から推定した入力時系列 $\hat{x}(t)$ (赤色) を図 3 に示す。ここで、共通入力の時系列

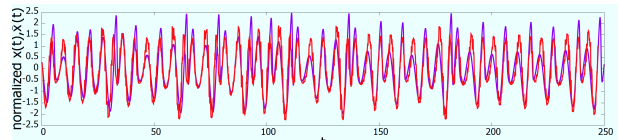


図 3: 共通入力時系列 $x(t)$ と推定した入力時系列 $\hat{x}(t)$

$x(t)$ と推定した入力時系列 $\hat{x}(t)$ は標準化を行った。図 3 を見ると、2 つの時系列の概形が類似していることがわかる。実際に $x(t)$ と $\hat{x}(t)$ の相関係数は、0.80 であり、入力信号を概ね推定できていることがわかる。

これらの結果は、個体差のある複数のニューロンより得られる ISI 時系列を用いて再構成したアトラクタを合成することで、共通入力への推定が可能であることを示すものである。

5 まとめ

本稿では、複数の ISI 時系列を用いてアトラクタを合成することで、個体差のある複数のニューロンへの共通入力への推定を行う手法を提案した。数値実験の結果、提案手法は個体差のある複数のニューロンに対する共通入力への推定において有効であることが示された。

なお、本研究の一部は、JSPS 科研費 (JP17K00348, JP20H00596) の援助を受けて行われた。

参考文献

- [1] T. Sauer, *Phys. Rev. Lett.*, **72**, 24, 3811–3814, 1994.
- [2] R. B. Stein, *Biophys. J.*, **7**, 1, 37–68, 1967.
- [3] O. E. Rössler, *Phys. Lett. A.*, **57**, 5, 397–398, 1976.
- [4] J. -P. Eckmann, *et al.*, *EPL*, **4**, 9, 973–977, 1987.
- [5] Y. Hirata, *et al.*, *Eur. Phys. J. Spec. Top.*, **164**, 1, 13–22, 2008.

マーク付点過程からの状態空間再構成について

On State Space Reconstruction from Marked Point Processes

澤田和弥¹
Kazuya Sawada

島田裕²
Yutaka Shimada

池口徹^{1,3}
Tohru Ikeguchi

東京理科大学大学院 工学研究科¹

埼玉大学大学院 理工学研究科²

東京理科大学 工学部³

1 はじめに

近年の観測技術の向上に伴い、様々な時系列データを高精度に観測できるようになっている。観測時系列データには、サンプリング間隔が一定ではない点過程データ、マーク付点過程データも存在する。典型的な点過程データとして、神経活動から得られるスパイク列が、マーク付点過程として地震・経済活動がある。もしこれらの観測データから因果関係を検出することができれば、様々な現象の理解に役立てることができる。例えば、多次元観測時系列のみからネットワーク構造を推定し [1]、システムを制御することなどへの応用が期待できる。

我々はすでに、非線形力学系理論に基づく、因果性検出法 (convergent cross mapping) [2] を拡張することで、点過程データ間の因果性を検出可能な手法を提案している [3, 4, 5]。提案手法は、再構成状態空間の情報に基づいて因果性を検出する。マークのない点過程については、すでにイベント間隔からの状態空間再構成手法などに関する議論がある [6]。一方で、マーク付点過程からの状態空間再構成については、観測時系列の極大値を抽出して得られる時系列には元の力学系の特徴が保持されることが文献 [7] で議論されている。しかし、マーク付点過程からの状態空間再構成は十分に検討されていない。そこで本稿では、マーク付点過程から状態空間を再構成する手法の妥当性の検討を行う。

2 解析手法と結果

本稿では、観測時系列の極大値を抽出することでマーク付点過程を生成する [8]。ここで、マーク付点過程から状態空間を再構成する方法について述べる。マーク付点過程の n 番目のイベント間隔 (IEI) を $T(n)$ 、マーク値を $M(n)$ とする。状態空間は、遅延座標系 $\mathbf{V}(n) = (M(n), T(n), \dots, M(n+(m-1)\tau), T(n+(m-1)\tau))$ を用いて再構成可能であると期待される [9]。ここで、 m は再構成次元、 τ は遅れ時間である。IEI とマーク値の時系列から状態空間を再構成することは、イベント間隔時系列から力学系を再構成可能なこと [6]、時系列の極大値には元の力学系の特徴が含まれていること [7]、多次元時系列から状態空間を再構成可能なこと [10] などから、埋め込みとなることが考えられる。

以下では、その妥当性を評価するために、Lorenz 方程式 [7] の全状態変数から構成したアトラクタ上での 2 点間距離分布と、極大値抽出で得たマーク付点過程から再構成した状態空間内のアトラクタ上の 2 点間距離分布間の類似度として、JS 情報量 [11] を算出する。2 点間距離を計算する際に使用したデータ長は 10,000 で、2 点

間距離は $[0, 1]$ の範囲で正規化した。また、2 点間距離分布のビン幅は 0.05 とし、再構成次元は $2 \leq m \leq 20$ の範囲で変化させ、遅れ時間は $\tau = 1$ とした。各変数 x, y, z に対する再構成次元と JS 情報量の関係を表 1 に示す。表 1 では、JS 情報量が最小となった値を赤字で

表 1 再構成次元と JS 情報量の関係

	再構成次元 m									
	2	4	6	8	10	12	14	16	18	20
x	0.062	0.018	0.026	0.062	0.098	0.118	0.151	0.184	0.208	0.238
y	0.095	0.054	0.041	0.056	0.084	0.121	0.151	0.185	0.215	0.249
z	0.201	0.166	0.110	0.071	0.050	0.042	0.042	0.044	0.050	0.059

示している。表 1 から、どの変数を用いてマーク付点過程を生成するかによって適切な再構成次元の値は異なるが、JS 情報量は 0.05 以下の小さな値を取っていることがわかる。これらの結果は、マーク付点過程を用いた時間遅れ座標系により、埋め込みとなる可能性を示唆するものである。ただし、2 点間距離分布が類似することは力学系が再構成可能な十分条件ではないため、今後はより詳細な調査が必要である。

3 まとめ

本稿では、マーク付点過程からイベント間隔 (IEI) とマーク値を用いて状態空間を再構成する手法の妥当性を検証した。具体的には、元の力学系のアトラクタ上の 2 点間距離分布とマーク付点過程の IEI とマークの値から再構成したアトラクタ上の 2 点間距離分布の間の類似度を JS 情報量で評価した。結果として、イベント間隔とマーク値を用いた状態空間の再構成が埋め込みとなる可能性が示唆された。なお、本研究の一部は JSPS 科研費 (Nos. JP15KT0112, JP17K00348, JP18K18125, JP20H00596) の援助を受けて行われた。

参考文献

- [1] K. Sawada et al., *NOLTA IEICE*, **11**, 4, 422–432, 2020.
- [2] G. Sugihara et al., *Science*, **338**, 6106, 496–500, 2012.
- [3] K. Sawada et al., *Proc. of ICMNS*, 80, 2019.
- [4] K. Sawada et al., *Proc. of NOLTA2019*, 591–594, 2019.
- [5] 澤田 和弥 他, 信学ソ大, N-1-6, 2020.
- [6] T. Sauer, *PRL*, **72**, 24, 3811–3814, 1994.
- [7] E. Lorenz, *J. Atmos. Sci.*, **20**, 2, 130–141, 1963.
- [8] N. Yabuta et al., *Proc. of NOLTA2007*, 188–191, 2007.
- [9] T. Sauer et al., *J. Stat. Phys.*, **65**, 3–4, 579–616, 1991.
- [10] E. Deyle et al., *PLoS ONE*, **6**, 3, e18295, 2011.
- [11] J. Lin, *IEEE Trans. Inf. Theory*, **37**, 1, 145–151, 1991.

単語の使用頻度と言語ネットワークの構造的特徴の関係

Relation between Frequency of Words and Language Network Structures

眞岸祈平[†]
Kihei MAGISHI

松本朋子[†]
Tomoko MATSUMOTO

島田裕^{††}
Yutaka SHIMADA

池口徹[†]
Tohru IKEGUCHI

[†] 東京理科大学

^{††} 埼玉大学

1 はじめに

我々は既に、複雑ネットワーク理論の観点から言語を解析することで、同一文書であっても記述する言語に依存して、ネットワークの構造的特徴が異なるものとなることを明らかにした [1-3]. しかし、文書のどのような特徴がネットワークの構造的特徴に反映するののかについては未解明である. そこで本稿では、文書の特徴の1つと考えられる単語の使用頻度とネットワーク構造の定量化指標との関係性について調査した. その結果、言語の種類に関わらず、単語の使用頻度と次数中心性・媒介中心性・近接中心性の間に相関があることが明らかとなったので報告する.

2 使用データと言語ネットワーク作成手法

本稿では、新約聖書のマタイによる福音書 [4-9] を対象とし、翻訳バージョンを含めた14種類のテキストデータを使用した. また、解析対象言語として、インド・ヨーロッパ語族に分類されるギリシャ語、ラテン語、英語、ドイツ語、フランス語、ロシア語、ウラル語族に分類されるフィンランド語、日本語族に分類される日本語、シナ・チベット語族に分類される中国語を用いた.

まず、形態素解析により文章を単語の原型(形態素)に分解し、その品詞を求めた. 形態素解析ツールとして日本語はmecab [10] を、他言語はtree-tagger [11] を使用した. 次に、単語の原型とその品詞の組を1つの頂点とし、文章中で隣接する頂点を接続することで、単語の共起に着目した重み無し無向ネットワークを作成した. その際、単語の重複、単語間のつながりの重複は認めず、句読点、半句読点、記号を除いた. なお、1つの文章から非連結なネットワークが得られた場合は、最大部分グラフを対象とした.

3 実験結果

まず、各言語のテキストデータの単語の使用頻度を算出した. 次に、5種類のネットワーク構造の定量化指標を求めた. 具体的には、次数中心性 [12]、媒介中心性 [12]、近接中心性 [13]、クラスタ係数 [14]、四角形クラスタ係数 [15] である. 最後に、各単語の使用頻度と各定量化指標の相関係数を算出することで、単語の使用頻度とネットワークの構造的特徴の関係性を調査した.

各単語の使用頻度とネットワーク指標(次数中心性、近接中心性、媒介中心性、クラスタ係数、四角形クラスタ係数)の相関係数を図1に示す. 図1の横軸は言語の種類、縦軸は相関係数を表している.

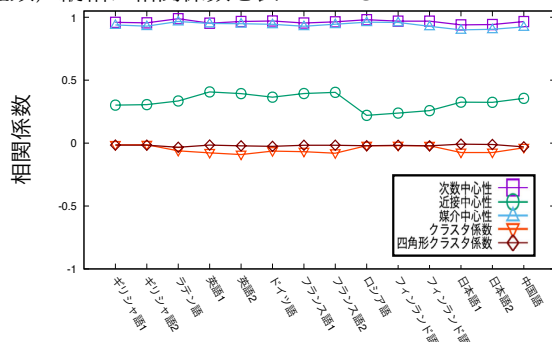


図1: 単語の使用頻度と各指標間の相関係数

図1を見ると、単語の使用頻度と次数中心性・媒介中心性の線形相関が非常に高いことが分かる. これらの結果は、単語の使用頻度がネットワークの構造的特徴に影響を及ぼす可能性があることを示唆するものである.

一方、単語の使用頻度と近接中心性の線形相関は低いことが分かる. しかし、英語について、単語の使用頻度と近接中心性の相関図を見ると、単語の使用頻度と近接中心性には非線形な相関があることが分かる(図2).

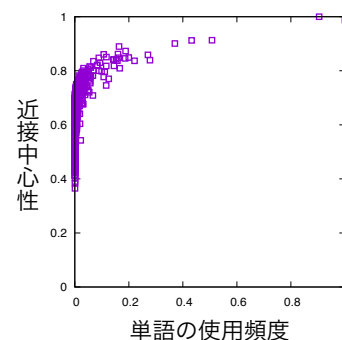


図2: 英語2に対する単語の使用頻度と近接中心性との相関図

また、単語の使用頻度とクラスタ係数・四角形クラスタ係数の間には相関はない.

4 まとめ

本稿では、文書の特徴の一つと考えられる単語の使用頻度とネットワーク構造の定量化指標との関係性について調査した. その結果、言語の種類や語族に関わらず、単語の使用頻度と次数中心性・媒介中心性との間には線形相関があること、近接中心性との間には非線形相関があることが明らかとなった. 本稿では、英語を対象とした結果のみ示したが、他言語においても同様な非線形相関が存在することは確認している. これらの結果の詳細については別途報告する予定である. 一方、クラスタ係数・四角形クラスタ係数との間には線形相関がないことも明らかとなった. 今後は、本稿の結果が言語ネットワーク作成手法に依存している可能性について調査する必要がある. また、単語の使用頻度以外の特徴とネットワーク構造の関係性を調査する必要がある.

謝辞 本研究の一部は、JSPS 科研費 (No. JP17K00348, JP18K12701, JP18K18125, JP20H00596) の助成を受けて行われた.

参考文献

- [1] 眞岸 祈平 他, 信学技報, **119**, 471, 77-82, 2020.
- [2] 眞岸 祈平 他, 信学技報, **120**, 55, 27-32, 2020.
- [3] 眞岸 祈平 他, 信学ソ大, N-1-14, 2020.
- [4] <https://www.bible.or.jp/> (2020-03-24 閲覧)
- [5] <https://www.academic-bible.com/> (2020-04-13 閲覧)
- [6] <http://www.cerbafo.org/> (2020-04-16 閲覧)
- [7] <http://www.nic.funet.fi/> (2020-04-20 閲覧)
- [8] <https://www.biblegateway.com/> (2020-04-26 閲覧)
- [9] <https://www.bebli.com/> (2020-04-28 閲覧)
- [10] <http://taku910.github.io/mecab/> (2020-09-16 閲覧)
- [11] <https://www.cis.uni-muenchen.de/> (2020-04-27 閲覧)
- [12] L. C. Freeman, *Social Networks*, **1**, 3, 215-239, 1978.
- [13] A. Bavelas, *J. Acoust. Soc. Am.*, **22**, 6, 725-730, 1950.
- [14] D. J. Watts, *et al.*, *Nature*, **393**, 440-442, 1998.
- [15] P. G. Lind, *et al.*, *PRE*, **72**, 5, 056127, 2005.

カオス時系列とノイズな準周期解の判別

Discrimination between Chaotic Time Series and Noisy Quasi-periodic Solutions

金丸 志生¹ Nina Sviridova² 島田 裕³ 池口 徹^{1,2}

東京理科大学大学院 工学研究科¹ 東京理科大学 工学部² 埼玉大学大学院 理工学研究科³

1 はじめに

非線形時系列解析手法の一つに同方向的近傍プロット (IDNP) [1] がある。IDNP は、アトラクタの 2 点間の近傍かつ同方向の関係を、2 次元の白黒画像で表現したものである。IDNP 上の黒画素の割合は、IDNP を構成するリカレンスプロット (RP) [2] と同方向性リカレンスプロット (IDRP) [1] の閾値の大きさに相当するパラメータ θ により定まる。我々は既に RP, IDRP の θ を徐々に大きくしたときの IDNP の変化を定量的に調べることで、非線形時系列の特徴を捉える手法を提案している [3-5]。これらの手法では、まず、RP と IDRP に対して、 θ の値を変化させることで複数の RP と IDRP を作成する。次に、同じ θ の値で作成された RP と IDRP の共通集合から、IDNP を作成する。最後に、IDNP と RP を用いて生起確率 R [1] を算出し、 θ に対する R の推移傾向を得る。文献 [3-5] では、この手順によって時系列データの特性を捉えることができることが示されているが、決定論的カオス時系列と準周期解に内部ノイズが印加された時系列の特性判別については未検討である。そこで本稿では、カオス解と準周期解に内部ノイズが印加された時系列の特性判別が提案手法により可能であるかを調査する。

2 解析手法

本稿では、1 次元の時系列データが観測されたと仮定する。この 1 次元の観測時系列データに対して、まずは、遅延座標系への埋め込み手法 [6, 7] を用いて状態空間を再構成する。次に、再構成状態空間から、2 点間距離の情報をもとに RP を作成する [2]。また、2 つの軌道ベクトル間の差の情報をもとに、同方向性を可視化する IDRP を作成する [1]。作成した RP と IDRP を用いることで、近傍かつ同方向の関係を可視化する IDNP を作成する [1]。ただし、RP と IDRP の閾値は、再構成状態空間内の任意の 2 点間についての情報 (RP であれば 2 点間距離, IDRP であれば軌道ベクトル間の差) の最小値から最大値までの値に、 $0 < \theta < 1$ の割合を掛け値とする。最後に、 θ の値を変化させることで作成した RP の総プロット数を分母、IDNP の総プロット数を分子としたときの生起確率 $R(\theta)$ を算出し、 θ の値の変化に対する $R(\theta)$ の推移傾向を得る。

3 数値実験

本稿では、Langford 方程式 [8] の状態変数 (x, y, z) にノイズ強度 σ の加法性の内部ノイズを印加した以下のモデルから生成される x を観測時系列とした。

$$\begin{cases} \dot{x} = (z - \beta)x - \omega y + \sigma \xi_1 \\ \dot{y} = \omega x + (z - \beta)y + \sigma \xi_2 \\ \dot{z} = \lambda + \alpha z - z^3/3 - (x^2 + y^2)(1 + \rho z) + \varepsilon z x^3 + \sigma \xi_3 \end{cases}$$

ここで、パラメータは $\alpha = 1$, $\beta = 0.7$, $\lambda = 0.6$, $\omega = 3.5$, $\rho = 0.25$ とし、 ε を変えてカオス解 ($\varepsilon = 0.25$) と準周期解 ($\varepsilon = 0$) を生成した。ただし、 $\xi_i(t)$ ($i = 1, 2, 3$) は独立なガウス増分を持つ標準 Wiener 過程に従う確率変数である。また、準周期解に内部ノイズが印加された時系列を生成するために、 $\sigma = 0.01, 0.02, 0.03, 0.04, 0.05$ の 5 種類を用いた。これらの数値計算はステップ幅が $h = 10^{-4}$ の 4 次のルンゲ・クッタ法により行い、時系列長は、過渡状態

として 10^6 ステップを省いた後の 2×10^6 ステップの応答から、500 ステップ毎に 1 点をサブサンプリングした 4,000 点とした。本稿では、一つの (ε, σ) の組に対して (x, y, z) の初期値を変更した 100 本の時系列を生成し、 $R(\theta)$ の平均を求めた。なお、 (x, y, z) の初期値は標準正規分布に従う確率変数とした。さらに、 θ を 2^{-i} ($i = 1, 2, \dots, 10$) とした。また、状態空間の再構成は次元数を $m = 7$ 、時間遅れ τ を観測時系列の自己相関関数が最初に $1 - 1/e$ になる時刻 [9] とした。

結果を図 1 に示す。図 1 の横軸はパラメータ θ 、縦軸は生起確率 $R(\theta)$ である。黒色の実線で結ばれた \triangle 印が、カオス時系列 ($\varepsilon = 0.25, \sigma = 0$) を、破線で結ばれた \circ 印が、準周期解に内部ノイズが印加された時系列 ($\varepsilon = 0, \sigma > 0$) を表している。線の色はノイズ強度 σ の違いを示している。図 1 において、カオス時系列は θ に依らず R が高い水準を

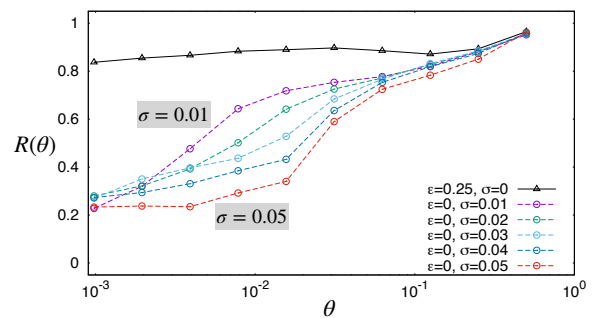


図 1: パラメータ θ に対する生起確率 $R(\theta)$ の推移。

保持し、その変動も小さい。一方、準周期解に内部ノイズが印加された時系列は、 θ を大きくすると R が増加する傾向がある。これらの推移傾向の違いから、提案手法によって、カオスと準周期解に内部ノイズが印加された時系列の判別が可能となることが示された。

4 まとめ

本稿では、パラメータ θ を変化させながら複数の RP と IDNP を作成することで得られる生起確率 $R(\theta)$ の推移傾向から、決定論的カオスと準周期解に内部ノイズが印加された時系列の判別が可能であることを調査した。その結果、これらの判別は可能であることが示された。なお、本研究の一部は JSPS 科研費 (Nos. JP17K00348, JP18K18125, JP19K14589, JP20H00596) の援助を受けて行われた。

参考文献

- [1] 寶来 俊介 他, 電学論 (C), **122**, 1, 141-147, 2002.
- [2] J. P. Zbilut *et al.*, Phys. Lett. A, **171**, 3-4, 199-203, 1992.
- [3] 小川 徳紀 他, 信学総大, N-1-33, 2019.
- [4] 金丸 志生 他, 信学技報, **119**, 19, 23-28, 2019.
- [5] S. Kanamaru *et al.*, Proc. of NOLTA, 536-539, 2019.
- [6] F. Takens, Lecture Notes in Math., **898**, 366-381, 1981.
- [7] T. Sauer *et al.*, J. Stat. Phys., **65**, 3-4, 579-616, 1991.
- [8] W. F. Langford, In: T. Küpper *et al.*, eds., "Numerical Methods for Bifurcation Problems," Int. Series of Num. Math., 285-295, 1984.
- [9] M. T. Rosenstein *et al.*, Physica D, **65**, 1-2, 117-134, 1993.

自転車再配置問題の実データに対する タブーサーチ法を用いた解法の性能調査

Performance evaluation of tabu search method for solving real data of multiple-vehicle bike sharing system routing problems

對馬 帆南¹ 松浦 隆文² 池口 徹¹
Honami Tsushima Takafumi Matsuura Tohru Ikeguchi

東京理科大学¹
Tokyo University of Science

日本工業大学²
Nippon Institute of Technology

1 まえがき

我々は既に Bike Sharing System (BSS) で生じる自転車の過不足を調整するために、複数の自転車回収車を用いた自転車再配置問題 (mBSSRP) [1] とその近似解を短時間で発見できる解法を提案している [2, 3]. これまでは、ランダムに駐輪ポートを配置した問題例を用いて提案解法の性能評価を行ってきた。本稿では、提案解法で使用している Or-opt を拡張した手法を導入し、実データ [4] に対する性能について報告する。

2 複数の自転車回収車を用いた自転車再配置問題

mBSSRP では、デポと n 個のポート集合 $V = \{0, 1, \dots, n\}$ ($\{0\}$ はデポ), 各ポートの 2 点間距離, ポート i に補充または回収しなければいけない自転車台数 q_i が与えられる。mBSSRP の目的は、最大積載容量 Q の回収車 M 台が自転車を回収しなければいけないポート ($q_i > 0$) で自転車を回収し、それらの自転車を不足しているポート ($q_i < 0$) に補充するとき、回収車の総移動距離が最小となる巡回経路を求めることである。mBSSRP では回収車は、自転車を積んでデポを出発することが可能であり、デポを出発し、再配置後に再びデポに戻る。自転車の再配置中にポートの自転車台数は変化しない。また、ポートには 1 台の回収車が 1 度のみ訪問可能である。mBSSRP の制約条件は、(1) 決められた時間内で作業を終える時間制約, (2) 回収車の容量を超えて自転車を積み込めない積み込み制約, (3) 1 度の訪問で回収と補充を行う補充制約である。

3 mBSSRP に対する解法

mBSSRP の良好な近似解を得るために、現在の解から次の解に遷移する近傍解候補を Or-opt と CROSS-exchange で生成する。Or-opt とは、ある巡回路の連続する 3 ポート以下の部分順列を他の巡回路のポート間に挿入する操作である [5].

これまで Or-opt で挿入できる連続ポート数 κ は 3 に設定していたが、文献 [6] では、 κ の値を 3 よりも大きくすることで、対称巡回セールスマン問題の短い巡回路を得られることが報告されている。そこで提案解法でも、 κ の値を 3 よりも大きく設定する。なお、CROSS-exchange は、従来通りの設定とした。また提案解法では、タブーサーチ法を用いて局所解に陥ることを防いでいる。

4 数値実験・結果

ポート数 $n = 90$ である実データ Ciudad de Mexico に対し、回収車の最大積載容量 $Q = 17$, 回収車台数 $M = 20$, 試行回数を 10 回, タブー期間を 250, タブーサーチ法の繰り返し回数を 1,000 と設定して数値実験を行なった。Or-opt で挿入できる連続ポート数は $\kappa = 3, 5, 10, 15, 20, 25$ と設定した。表 1 に実験結果を示す。表 1 より、提案解法は実データに対しても巡回経路を構築することができることが分かる。さらに、Or-opt での挿入連続ポート数 κ の増加に伴って、短い巡回経路を得られることが確認できた。

表 1 κ を変えたときの誤差率と計算時間

κ	誤差率 [%]		計算時間 [s]
	平均	最良	
3	11.28	9.33	84.83
5	10.64	7.39	92.92
10	8.96	7.73	105.28
15	8.84	8.15	120.00
20	8.10	6.67	134.07
25	7.97	5.78	138.77

5 まとめ

本稿では、実データに対する提案解法の性能評価を行なった。数値実験の結果、提案解法は実データに対しても実行可能解が得られることを確認した。さらに Or-opt 挿入で可能な連続ポート数 κ の増加により、短い巡回経路を得られることがわかった。なお、本研究は JSPS 科研費 (No. JP19K04907, JP17K00348, JP20H000596) の援助を受けて行われた。

参考文献

- [1] 對馬帆南ら, 2017 年 電子情報通信学会 NOLTA サイエティ大会, A-16, 2017.
- [2] H. Tsushima, et al., *Journal of Signal Processing*, **22**(4), pp. 157–160, 2018.
- [3] H. Tsushima et al., submitted, 2020.
- [4] D. Mauro, et al., *Omega*, **45**, pp. 7–19, 2014.
- [5] I. Or, *PhD thesis*, Northwestern University, Evanston, 1976.
- [6] G. Babin, et al., *Journal of the Operational Research Society*, **58**(3), pp. 402–407, 2007.

TSP の近似解法に与える枝候補作成手法について

On a Method for Searching Good Edge Candidate Sets of TSP

土佐 真義[†]

宮 南風[†]

對馬 帆南[†]

池口 徹[†]

[†] 東京理科大学大学院 工学研究科

1 はじめに

巡回セールスマン問題 (Traveling Salesman Problem, 以下, TSP) は組合せ最適化問題の 1 つである. TSP は NP 困難なクラスに属する問題であり, 厳密解法だけではなく近似解法の開発も重要である.

TSP の高性能な近似解法の 1 つに Lin-Kernighan-Helsgaun-heuristic [1] (以下, LKH 法) がある. LKH 法は, 都市集合と最適解に含まれると推測される枝の集合 (以下, 枝候補) も入力とすることで, 短い時間で良好な近似解を求めることができる. しかし, 枝候補を作成する手法は, 近似解法と同様に都市数の増加に伴って計算時間も増加する. すなわち, 効率的な枝候補の作成法も重要である.

文献 [2] では, 枝候補の作成手法を提案している (以下, 従来手法). 従来手法では, 高速な巡回路作成手法である Partial OPTimization Metaheuristic Under Special Intensification Conditions [3] を用いて複数の巡回路を用意する. 得られた巡回路に一度以上含まれる全ての枝を枝候補とする.

本稿では, 従来手法の計算時間削減を目的とした枝候補作成手法を提案する. 数値実験の結果, 提案手法は従来手法と比較して, 同等の解探索性能を有しながらも計算時間を削減できることがわかった.

2 従来手法

従来手法 [2] で用いられている巡回路作成アルゴリズムを以下に示す. なお, 2 つの都市 i, j 間の距離を d_{ij} とする.

- 問題都市集合 V から n_a 個の都市をランダムに選択し, 代表都市集合 V' とする.
- 代表都市集合 V' に含まれる n_a 都市に対してランダムな巡回路 T_s を作成する.
- 巡回路 T_s を Lin-Kernighan-heuristic [4] (以下, LK 法) を用いて改善する.
- 巡回路 $T = T_s$ とする.
- T について, $v \notin V'$ を d_{uv} ($u \in V'$) が最も小さくなる u の直後に挿入する.
- 各都市 $v \in V'$ について, 以下の作業を行う.
 - T_s について, v の 2 つ先に訪れる都市を w とする.
 - 巡回路 T において, v の直前に訪れる都市から始まり w で終わる部分巡回路を T' とする.
 - 部分巡回路 T' の始点と終点を固定し, 2-opt 法を用いて改善操作を行う.

- 手順 (c) で得られた部分巡回路 T' を用いて T を更新する.

上記の手順で巡回路を複数個作成し, それらに含まれる全ての枝を枝候補とする.

3 提案手法

従来手法では, 大規模な問題を小問題に分割することで高速な巡回路の作成を実現している. しかし, 従来手法では, 代表都市数 n_a を大きく設定すると計算時間が増加し, 少なく設定すると余分な枝が多く含まれてしまう. そこで本稿では, 手順 3 の代表都市の改善操作における LK 法の代わりに 2-opt 法を用いる手法を提案する. 2-opt 法を用いることで, 代表都市数 n_a を増やすことができる. これにより, 近似解法の計算時間を削減することができる.

従来手法と提案手法で得られた枝候補を LKH 法 [1] に与えた際の解の誤差率と総計算時間, ならびにそれぞれの枝候補に含まれる総枝本数と最適解に含まれる枝本数を比較した. 対象としたインスタンスはベンチマーク問題集 TSPLIB [5] のものである. 誤差率は, $\{(\text{巡回路長}) - (\text{最適巡回路長})\} / (\text{最適巡回路長}) \times 100(\%)$ で定義した. 結果を表 1 に示す. 表 1 をみると, 誤差率は従来手法が優れているものの, 提案手法は計算時間を削減できていることがわかる. これは, 提案手法では, 枝候補に含まれる最適解の枝本数の比率を 1 に保ったまま, 枝候補に含まれる総枝本数を削減できていることが要因である.

4 まとめ

本稿では, 近似解法に与える枝候補作成手法を提案した. 数値実験の結果, 提案手法を用いると, 従来手法よりも計算時間を削減できることを明らかにした.

本研究の一部は, JSPS 科研費 (No. 15KT0112, 17K00348, 20H00596) の助成を受けて行われた.

参考文献

- [1] K. Helsgaun, *EJOR*, **126**, 1, 106–130, 2000.
- [2] É. D. Taillard, and K. Helsgaun, *EJOR*, **272**, 2, 420–429, 2019.
- [3] É. D. Taillard, and S. Voss, *Essays and surveys in metaheuristics*, 613–629, 2002.
- [4] S. Lin, and B. W. Kernighan, *Oper. Res.*, **21**, 2, 498–516, 1973.
- [5] G. Reinelt, *ORSA J. on Comp.*, **3**, 4, 376–384, 1991.

表 1 従来手法と提案手法の性能の比較

問題 [5]	誤差率 (%)			計算時間 (秒)			枝候補に含まれる総枝本数			最適解に含まれる枝本数		
	従来手法 (A)	提案手法 (B)	比率 (B/A)	従来手法 (C)	提案手法 (D)	比率 (D/C)	従来手法 (E)	提案手法 (F)	比率 (F/E)	従来手法 (G)	提案手法 (H)	比率 (H/G)
pr2392	0.376	0.322	0.856	7.30	4.73	0.648	7,831	9,096	1.162	2,388	2,835	0.999
usa13509	0.380	0.475	1.250	125.35	90.16	0.719	51,114	55,215	1.080	13,462	13,463	1.000
d15112	0.076	0.092	1.211	168.28	135.24	0.804	59,038	63,450	1.075	15,038	15,087	1.003
pla33810	0.568	0.865	1.523	434.93	374.65	0.861	146,022	133,363	0.913	33,748	33,717	0.999
pla85900	0.229	0.299	1.305	1589.15	1288.62	0.811	446,016	346,183	0.776	85,807	85,770	1.000

文書内容の違いが言語ネットワークの構造に与える影響

Impact of the differences in document contents on the structure of language networks

眞岸祈平[†]
Kihei MAGISHI

松本朋子[†]
Tomoko MATSUMOTO

島田裕^{††}
Yutaka SHIMADA

池口徹[†]
Tohru IKEGUCHI

東京理科大学[†]

埼玉大学^{††}

1 はじめに

複雑ネットワークの観点から言語解析を行うことで、異なる言語で記述された文書から得られる言語ネットワークの構造的特徴の差異が、各言語の文法構造上の特徴を反映することを示す結果が報告されている [1,2]。しかし、言語ネットワークの作成に用いた文書内容が言語間で同一ではないため、言語ネットワークの構造的特徴の差異が、文法構造ではなく文書内容によって生じた可能性が残されている。

本稿では、同一内容の文書が複数の異なる言語で翻訳されている新約聖書を解析対象として、言語ネットワークの構造的特徴の差異を調査した。具体的には、各言語で書かれた新約聖書から重み無し無向ネットワークを作成し、クラスタ係数 [4]、平均頂点間距離 [4]、スペクトルグラフ距離 [5] を用いることで、文書内容の同一性がネットワークの構造的特徴の差異に与える影響の有無を調査したので報告する。

2 使用データ

本稿では、分析対象として新約聖書の福音書 [6-11] の翻訳を使用した。具体的には、マタイによる福音書 (以下、マタイ伝)、マルコによる福音書 (以下、マルコ伝)、ルカによる福音書 (以下、ルカ伝)、ヨハネによる福音書 (以下、ヨハネ伝) の全 4 種類である。なお、各福音書の文書の内容は異なることに注意する。また、解析対象として、印・欧語族に分類される希語、羅語、英語、独語、仏語、露語、日本語族に分類される日本語、烏拉語族に分類される芬語、支那・西蔵語族に分類される中国語を用いた。

3 言語ネットワーク作成方法

まず、形態素解析により文章を単語の原型 (形態素) に分解し、その品詞を求める。形態素解析ツールとして、日本語は mecab [12] を、他の言語は tree-tagger [13] を使用した。次に、単語の原型とその品詞の組を 1 つの頂点とし、文章中で隣接する場合に枝で結ぶことで、重み無し無向ネットワークを作成した。その際、単語の重複、単語間のつながりの重複は認めず、句読点、半句読点、記号を除いた。なお、1 つの文章から非連結なネットワークが得られた場合は、最大部分グラフを用いた。

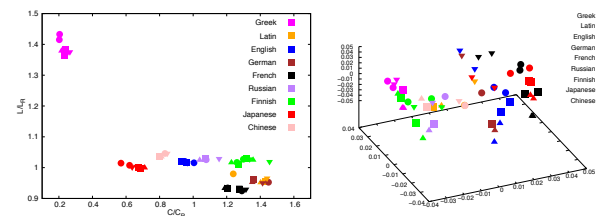
4 実験

本稿では、同一の福音書から得られたネットワークの構造的特徴と異なる福音書から得られたネットワークの構造的特徴を比較することで、文書内容の同一性がネットワークの構造的特徴に与える影響を調査する。まず、各福音書から得られたネットワークの構造的特徴を、クラスタ係数 (C) [4] と平均頂点間距離 (L) [4] を用いて算出した。その際、元のネットワークをランダム化したネットワークのクラスタ係数 (C_R) と平均頂点間距離 (L_R) を用いて C と L の値を正規化した [14]。次に、ネットワークの構造的特徴の類似性を、スペクトルグラフ距離 [5] を用いて調査した。その際、スペクトルグラフ距

離に多次元尺度法を適用することで、各ネットワークの類似性を可視化した。

5 結果

C/C_R と L/L_R の関係、ネットワークの類似性の結果を図 1 に示す。なお、図 1 では、□ がマタイ伝、○ がマルコ伝、△ がルカ伝、▽ がヨハネ伝である。また、図 1(a) の横軸は C/C_R 、縦軸は L/L_R である。



(a) C/C_R , L/L_R の関係 (b) ネットワークの類似性
図 1: ネットワーク解析の結果

図 1(a) を見ると、文書内容が同一でない場合でも、同一言語で書かれた福音書から得られたネットワークは類似した構造的特徴を有することがわかる。一方、図 1(b) を見ると、英語、仏語、日本語のように、文書内容が同一でないネットワークの構造的特徴が類似しない言語もあることがわかる。

6 まとめ

本稿では、言語間で文書の内容が同一であるか否かで、言語ネットワークの構造的特徴に差異が出現するかを調査した。その結果、文献 [1,2] などで用いられてきた、クラスタ係数、平均頂点間距離を用いた場合には、文書内容が同一でなくても結果が変化しないことが示唆された。一方、スペクトルグラフ距離を用いる場合は、ネットワーク構造に文書内容の差異に関する情報も含まれていることが示唆された。これらのことから、ネットワークの観点から言語解析を行う場合、文書内容の同一性は重要な要素であり、この意味でも、新約聖書を用いた解析は必要であると考えられる。

謝辞 本研究の一部は、JSPS 科研費 (No. 15KT0112, 17K00348, 18K12701, 18K18125) の助成を受けて行われた。

参考文献

- [1] Y. Shimada, et al., *EPL*, **127**, 5, 56003, 2019.
- [2] C. I. R. Ferrer, et al., *Biol. Sci.*, **268**, 1482, 2261, 2001.
- [3] 眞岸 祈平 他, 信学技報, **120**, 55, 27-32, 2020.
- [4] D. J. Watts, et al., *Nature*, **393**, 440-442, 1998.
- [5] Y. Shimada, et al., *Sci. Rep.*, **6**, 34944, 2016.
- [6] <https://www.bible.or.jp/> (2020-03-24 閲覧)
- [7] <https://www.academic-bible.com> (2020-04-13 閲覧)
- [8] <http://www.cerbafaso.org/> (2020-04-16 閲覧)
- [9] <http://www.nic.funet.fi/> (2020-04-20 閲覧)
- [10] <https://www.biblegateway.com> (2020-04-26 閲覧)
- [11] <https://www.bebliia.com/> (2020-04-28 閲覧)
- [12] <http://taku910.github.io/mecab/> (2020-01-04 閲覧)
- [13] <https://www.cis.uni-muenchen.de/> (2020-04-27 閲覧)
- [14] R. Milo, et al., *Science*, **303**, 1538-1542, 2004.

テンポラルネットワーク上での情報拡散の解析

Analysis of Information Diffusion on Temporal Networks

ジョ ロエン¹

Luyan XU

東京理科大学 大学院工学研究科¹

澤田 和弥¹

Kazuya SAWADA

埼玉大学 大学院理工学研究科²

島田 裕²

Yutaka SHIMADA

池口 徹^{1,3}

Tohru IKEGUCHI

東京理科大学 工学部³

1 はじめに

近年、複雑ネットワーク上での情報拡散に関する研究が盛んに行われている [1, 2]. 我々は既に、静的なネットワーク構造が情報拡散に与える影響について調査した結果を報告している [3, 4]. 一方、現実世界には静的なネットワークだけでなく、構造が動的に変化するテンポラルネットワークも数多く存在する. そこで本稿では、時間と共に構造が変化するテンポラルネットワークに対して、情報への関心度に着目した情報拡散モデル [3] を適用し、テンポラルネットワーク上での情報拡散に関する調査を行ったので報告する.

2 コンタクトネットワーク

本稿では、フランスの高校と病院で観測された、人と人のコンタクトのデータ [5-7] をテンポラルネットワークに変換し、2つのテンポラルネットワーク上での最終的な情報拡散率を調査した. 高校でのデータは、126名(教員と学生)の4日間の対面接触記録である. 病院でのデータは、75名(医療スタッフと患者)の5日間の対面接触記録である. また、高校と病院でのコンタクトは20秒ごとに観測される. 時刻 t で実験協力者 i と j にコンタクトが存在する場合、時刻 t でのテンポラルネットワークの隣接行列 $A(t)$ の (i, j) 成分は $A_{ij}(t) = A_{ji}(t) = 1$, そうでない場合は、 $A_{ij}(t) = A_{ji}(t) = 0$ となる.

3 情報拡散モデルを用いた数値実験

高校と病院のテンポラルネットワーク上で、情報拡散の数値実験を行い、情報の最終拡散率 S を調査した. 時刻 t での頂点 i の情報への関心度を $x_i(t)$ とする. 関心度 $x_i(t)$ は式 (1) に従って変化する.

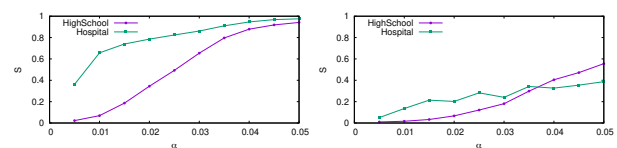
$$x_i(t+1) = g_i(\tau, t)x_i(t) + \alpha H(x_i(t) - \theta_i) \sum_{j \in G_i(t)} F(x_j(t) - \theta_j) \quad (1)$$

ここで、 t_i^* を頂点 i が最初に情報を受けとった時刻、 τ ($\tau < 1$) を情報への関心度の減衰係数とすると、 $g_i(\tau, t)$ は、 $t - t_i^* \equiv 0 \pmod{3600}$ のとき $g_i(\tau, t) = \tau$, それ以外のとき $g_i(\tau, t) = 1$ となる. また、 $F(x)$ はステップ関数、 $H(x)$ は $x \geq 0$ のとき $H(x) = 0$, $x < 0$ のとき $H(x) = 1$ となる関数である. さらに、 $G_i(t)$ は時刻 t での頂点 i の隣接頂点の集合、 θ_i は頂点 i の閾値、 α は情報の影響力である. 頂点 j の閾値 θ_j は、範囲が $[0, 1]$ の平均 $\mu = 0.5$, 標準偏差 $\sigma = 0.2$ の切断正規分布とした.

拡散開始の条件として、最初に接触した2つの頂点のうち、一方の関心度を1にすることで情報を拡散させた. $g_i(\tau, t)$ により、3600秒ごとに頂点 i の情報への関心度が減衰する. また、 $H(x)$ により、頂点 i は閾値に達していない場合のみ隣接頂点からの情報を受信する. 本実験では、情報の影響力を

$0.005 \leq \alpha \leq 0.05$ とし、 $\tau = 0.95, 0.99$ とした. なお、本稿での結果は、200回の試行平均である.

情報の影響力 α を変化させ、高校と病院のテンポラルネットワーク上での最終拡散率 S を調査した結果を図1に示す. 最終拡散率 S は、最後のコンタクトが終わるまでに、一度閾値に達した頂点数の全頂点に対する割合である.



(a) 減衰係数 $\tau = 0.99$

(b) 減衰係数 $\tau = 0.95$

図1: 高校と病院のテンポラルネットワーク上での情報の影響力 α と最終拡散率 S の関係

図1(a)より、減衰係数 $\tau = 0.99$ のとき、高校に比べて、病院の最終拡散率 S が大きいことがわかる. また、情報の影響力 α が大きくなるにつれ、病院に比べて、高校での最終拡散率 S が顕著に変化することがわかった. 次に、図1(b)より、減衰係数 $\tau = 0.95$ のとき、 $\alpha < 0.035$ の場合では、高校より病院の S が大きいことがわかった. 一方、 $\alpha \geq 0.035$ の場合では、病院より高校の S が大きいことがわかった. これは、病院より高校の S が α に強く依存することを示すものである. また、図1(a), (b)の結果を比較すると、減衰係数 τ による最終拡散率 S の変化は、高校に比べて、病院において顕著であることが示唆される.

4 まとめ

本稿では、高校と病院でのコンタクトデータ [5] をテンポラルネットワークに変換した. それらのネットワーク上で情報拡散の数値実験を行い、減衰係数 τ を変化させた際の情報の影響力 α と情報の最終拡散率 S の関係について調査した. その結果、高校での情報の最終拡散率は病院に比べて、情報の影響力 α に強く依存することがわかった. それに対して、病院での情報の最終拡散率は、高校に比べて、情報への関心度の減衰率 τ に強く依存することがわかった. なお、本研究の一部は JSPS 科研費 (No. 15KT0112, 17K00348, 18K18125) の援助を受けて行われた.

参考文献

- [1] D. J. Watts & S. H. Strogatz, *Nature*, **393**, 440-442, 1998.
- [2] D. Centola, *Science*, **329**, 1194-1197, 2010.
- [3] A. Tanaka et al., *Proc. of NOLTA2013*, 298-301, 2013.
- [4] ジョ, 他, 信学技報, **119**, 471, 83-88, 2020.
- [5] <http://www.sociopatterns.org/> (最終閲覧日: 2020年06月15日)
- [6] P. Vanhems et al., *PLOS ONE*, **8**, e73970, 2013.
- [7] J. Fournet & A. Barrat, *PLOS ONE*, **9**, e107878, 2014.

興奮性/抑制性ニューロン比率がシナプス伝達情報量に与える影響

Effects of excitatory/inhibitory neuron ratio on synaptic information entropy

宮南風¹ 島田裕² 藤原寛太郎³ 池口徹^{1,4}
 Nanfu Miya Yutaka Shimada Kantaro Fujiwara Tohru Ikeguchi

東京理科大学 大学院工学研究科 経営工学専攻¹ 埼玉大学 大学院理工学研究科 数理電子情報部門²
 東京大学国際高等研究所 ニューロインテリジェンス国際研究機構³ 東京理科大学 工学部 情報工学科⁴

1 はじめに

生理学的知見によれば、大脳皮質では興奮性ニューロンと抑制性ニューロンは約 4:1 の比率 (以下、生理学的比率) で存在するといわれている [1]。しかし、なぜこのような比率となるのか、またこの比率が変化した場合、脳内の神経活動にどのような影響があるのかについては明らかになっていない。

我々は既に、STDP 学習則 [2] に従うニューラルネットワークでは、生理学的比率において全ニューロンの平均発火率が最も上昇することを数値実験により示した [3]。本稿では、生理学的比率の必然性をニューロン間の伝達情報量の側面から解析する。情報伝達の因果性解析で用いられる Delayed Transfer Entropy [5] (以下、DTE) を拡張した指標を提案し、ニューロン比率の変化に対する伝達情報量の変化を調査した。

2 評価指標

文献 [5] の評価指標 $D_p[i \rightarrow j]$, $D_c[i \rightarrow j]$ を拡張する。まず、ニューロン i と j の間のシナプス伝達遅延を d と仮定した場合のシナプス伝達情報量 $DTE_{i \rightarrow j}(d)$ を算出する。さらに、 d を変化させた際の $DTE_{i \rightarrow j}(d)$ の最大値 $D_p[i \rightarrow j] = \max_{1 \leq d \leq d_{\max}} \{DTE_{i \rightarrow j}(d)\}$ と、DTE の信頼度を

$$\text{表す } D_c[i \rightarrow j] = \frac{\sum_{d=d_p-\tau/2}^{d_p+\tau/2} DTE_{i \rightarrow j}(d)}{\sum_{d=1}^{d_{\max}} DTE_{i \rightarrow j}(d)}$$

を用いてシナプスの結びつきの強さを評価する [5]。ただし、 d_{\max} は最大遅延時間であり、本稿では $d_{\max} = 30$ とした。 d_p は $D_p[i \rightarrow j]$ をとる遅延値 d である。また、 $D_c[i \rightarrow j]$ のパラメータ τ は信頼区間であり、本稿では $\tau = 5$ とした。 d の値によって $DTE_{i \rightarrow j}(d)$ が大きく変動し、かつピーク値の高いシナプスはその遅延による結びつきが強いと推定できる。

文献 [5] の手法では、伝達遅延を伴う情報量が正しく推定できないとき、 $D_p[i \rightarrow j]$ が高いが、 $D_c[i \rightarrow j]$ が低いという傾向が現れる。したがって、このような場合には、伝達情報量の値は低く算出されることが望ましい。そこで本稿では、伝達強度

を評価する指標として $TR[i \rightarrow j] = D_p[i \rightarrow j] \cdot D_c[i \rightarrow j]$ を提案する。提案指標を用いると、 $D_p[i \rightarrow j]$ と $D_c[i \rightarrow j]$ の両者が高い場合のみ $TR[i \rightarrow j]$ の値が高くなる。例として、

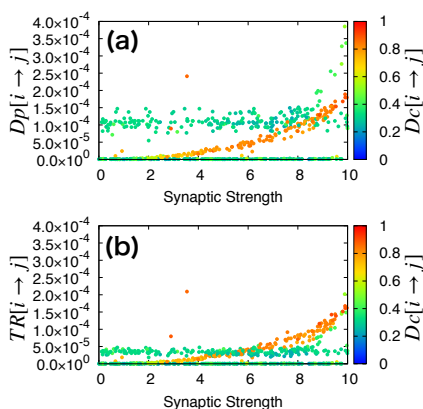


図 1. 結合強度と各情報量の関係

頂点数 100、ニューロンの結合確率 10%、興奮性/抑制性ニューロン比率 3:2 のネットワークで指標を算出した結果を図 1 に示す。本実験では、ニューロン間の結合強度は [0,10] の範囲での一様分布に従うとした。縦軸が従来法 (図 1(a)) と提案法 (図 1(b)) の情報量、横軸は結合強度、カラーバーは $D_c[i \rightarrow j]$ の値を表す。

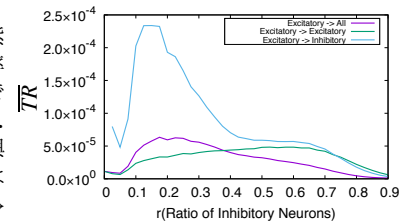
図 1(a) では、 $D_p[i \rightarrow j]$ が高く $D_c[i \rightarrow j]$ が低いシナプスが存在することがわかる。一方、図 1(b) では、それらのシナプスの情報量が全体的に低いことがわかる。また、 $D_c[i \rightarrow j]$ の高いシナプスの分布は双方の図で大きく変わらない。よって、 $TR[i \rightarrow j]$ を用いることにより、伝達情報量を正しく推定できていることが確認できる。

3 実験

全ニューロンに対する抑制性ニューロン比率の割合を、 $0 \leq r \leq 0.9$ の範囲で設定する。各 r の値に対して文献 [3] と同様のニューラルネットワークを用意し、STDP 学習を行う。学習終了後、十分に経過した時刻における、各興奮性シナプスの伝達強度を算出した。

シナプスごとに得られた伝達強度 $TR[i \rightarrow j]$ の結果を図 2 に示す。図 2 の縦軸 \overline{TR} は興奮性シナプス全体の $TR[i \rightarrow j]$ の平均値を表す。

図 2 より、 \overline{TR} (紫色) は生理学的比率である $r \approx 0.2$ で最も高い値をとった。また、興奮性 → 興奮性結合 (緑色) では $r \approx 0.7$ 、興奮性 → 抑制性結合 (水色) では $r \approx 0.1$ で高い値

図 2. ニューロン比率と \overline{TR} の関係

となった。 r が高いとき、興奮性 → 興奮性結合の学習機会は減少し、長期抑圧 (LTD) が少なくなる [3]。その結果、シナプス結合強度は全体的に増強され、伝達強度も高い値となる。その一方で、過度な抑制性ニューロンの存在は複数のニューロンへの情報伝達を直接的に障害する。ゆえに、興奮性シナプス全体では、 $r \approx 0.2$ で \overline{TR} が最大になると考えられる。

4 まとめ

本稿では、興奮性/抑制性ニューロン比率を変化させた場合のニューロン間のシナプス伝達情報量を調査した。まず、DTE を拡張した指標を提案し、その有用性を示した。さらに、提案指標を適用した結果、生理学的比率付近で興奮性シナプスの有する平均情報量が向上することを明らかにした。この結果は、ネットワーク全体の複数のニューロンに情報が伝播し、これにより発火が促進されることを示唆するものである。なお、本研究の一部は JSPS 科研費 (No.15KT0112, 17K00348, 18K18125) の援助を受けて行われた。

参考文献

- [1] De Felipe J. and Farinas I., *Prog. Neurobiol.*, Vol. 6, No. 39, pp. 563-607, 1992.
- [2] Bi G. Q. and Poo M. M., *J. Neurosci.*, Vol. 18, No. 8, pp. 10464-10472, 1998.
- [3] 宮南風 他, 信学総大, N-1-25, 2019.
- [4] 宮南風 他, 信学ソ大, N-1-10, 2019.
- [5] Ito, S. et al., *PLoS ONE*, Vol. 11, No. 6, e27431, 2011.

マーク付き点過程データに対する因果性解析

Causality Analysis for Marked Point Process

澤田 和弥¹

島田 裕²

池口 徹^{1,3}

Kazuya Sawada

Yutaka Shimada

Tohru Ikeguchi

東京理科大学 大学院工学研究科 経営工学専攻¹ 埼玉大学 大学院理工学研究科 数理電子情報部門²
東京理科大学 工学部 情報工学科³

1 はじめに

地震や神経活動, 経済活動は点過程データとして観測される. 我々は既に, 一定のサンプリング間隔で観測された時系列データ間の因果性を検出可能な Convergent Cross Mapping (CCM) [1] を拡張することで, 神経細胞の数理モデルから得られる発火時間間隔時系列に対して因果性を検出する手法 (ISICCM 法) を提案し, 点過程データに対しても高精度な因果検定が可能となることを示した [2, 3]. 一方, マーク付き点過程に対しても ISICCM 法が有効となるかを検証することも重要である. そこで本稿では, 数理モデルから生成したマーク付き点過程データに対して ISICCM 法により因果性が検出可能であることを調査したので報告する.

2 解析手法とデータ生成方法

本稿では, マーク付き点過程間の因果性解析に ISICCM 法 [2, 3] を拡張した手法を用いる. ISICCM 法は ISI 時系列から状態空間を再構成し, 発火時間による時間対応を取ることで ISI 時系列を予測する方法である. 本稿では, 状態空間を再構成する際, イベント間隔 (IEI) 時系列とマーク時系列の 2 次元時系列から 10 次元の状態空間を再構成し, ISICCM 法と同様の手順で IEI 時系列とマーク時系列をそれぞれ独立に予測した. また, マーク付き点過程は, 式 (1) に示すレスラー方程式とローレンツ方程式の結合系 [4] を基に両方向に結合させて得られた時系列データの極大値とその発生時刻として生成した [5].

$$\begin{cases} \dot{x}_1 = -\alpha(x_2 + x_3) + C_1 y_1 \\ \dot{x}_2 = \alpha(x_1 + 0.2x_2) \\ \dot{x}_3 = \alpha(0.2 + x_3(x_1 - 5.7)) \\ \dot{y}_1 = 10(-y_1 + y_2) \\ \dot{y}_2 = 28y_1 - y_2 - y_1 y_3 + C_2 x_2^2 \\ \dot{y}_3 = y_1 y_2 - 8/3 y_3 \end{cases} \quad (1)$$

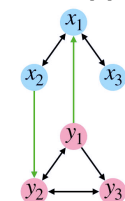
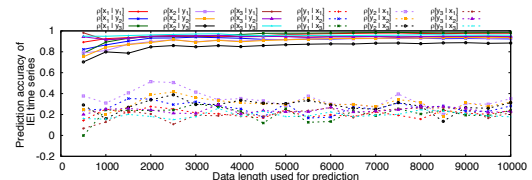


図 1 変数間の因果関係

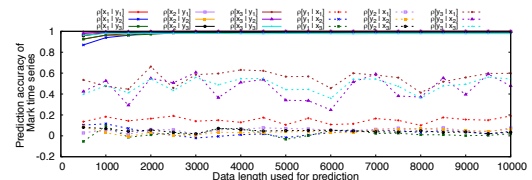
本稿では, $\alpha = 3$, $C_1 = 1$, $C_2 = 8$ とした. また, 式 (1) の変数間の因果グラフを図 1 に示す. 図 1 は, ローレンツ系とレスラー系が緑色の矢印で相互に影響を与えていることを示す.

3 数値実験

IEI 時系列とマーク時系列から生成した 10 次元再構成状態空間を用いて, IEI 時系列とマーク時系列を予測した結果を図 2 に示す. 横軸は予測に用いたデータ長, 縦軸は予測精度 (相関係数) を表す. また, 図 2 の $\rho[x_i|y_j]$ は y_j を用いて x_i を予測した結果, $\rho[y_i|x_j]$ は x_j を用いて y_i を予測した結果である. 図 2 を見ると, y の情報を用いて x を予測した全ての場合において, 時系列長が



(a) IEI 時系列を予測した場合



(b) マーク時系列を予測した場合

図 2 予測に用いたデータ長と予測精度の関係

500 のとき 0.7 以上の予測精度を示し, 時系列長 10,000 のとき 0.9 以上の高い予測精度に収束している. 時系列長を増加させた場合に, 予測精度が高い値に収束することは予測方向と逆の方向の因果性を検出していることを意味するため [1], x から y への因果性を検出できている. しかし, x の情報を用いて y を予測した場合は, 予測精度は 0 から 0.6 程度で推移している. ここで, レスラー方程式とローレンツ方程式は両方向に結合しているが, $C_1 = 1$, $C_2 = 8$ と x から y への因果性に比べて y から x への因果性が弱いいため, 予測精度が低下したと考えられる.

4 まとめ

本稿では, 既に我々が提案した点過程データ間の因果性検出手法 (ISICCM 法) [2, 3] をマーク付き点過程に拡張し, 適用した結果を報告した. 具体的には, IEI 時系列とマーク時系列の 2 次元時系列から状態空間を再構成した点が拡張部分である. その結果, マーク付き点過程データに対しても拡張した ISICCM 法は有効である可能性が示唆されたが, 結合強度に偏りが存在する場合は正しく因果性を検出できない場合もあることがわかった. なお, 本研究の一部は JSPS 科研費 (No. 15KT0112, 17K00348, 18K18125) の援助を受けて行われた.

参考文献

- [1] G. Sugihara et al., *Science*, **338**(6106): 496–500, 2012.
- [2] 澤田 和弥 他, 信学ソ大, N-1-12, 2019.
- [3] 澤田 和弥 他, 信学総大, N-1-9, 2020.
- [4] M. Quyen et al., *Physica D*, **127**(3–4): 250–266, 1999.
- [5] N. Yabuta and T. Ikeguchi, *Proc. of NOLTA*, 188–191, 2007.

自転車再配置問題に対する Or-opt と CROSS-exchange の近傍解を削減した解法の性能調査

Performance of Reduction of Or-opt and CROSS-exchange neighborhoods
for multiple-Vehicle Bike Sharing System Routing Problem

對馬 帆南¹
Honami Tsushima

松浦 隆文²
Takafumi Matsuura

池口 徹¹
Tohru Ikeguchi

東京理科大学¹
Tokyo University of Science

日本工業大学²
Nippon Institute of Technology

1 はじめに

バイクシェアリングシステム (BSS) で生じる自転車の過不足台数を調整するために、複数の自転車回収車を用いた自転車再配置問題 (mBSSRP) を提案している [1]。しかし、mBSSRP は制約条件が厳しいため、提案解法は一部の問題例に対して実行可能解を得ることができない場合があった。そこで、mBSSRP の一部の制約を取り除き、その制約に対する重み付きの違反量として mBSSRP の目的関数に加えたソフトな制約付き mBSSRP (mBSSRP-S) を提案した [2]。違反量が 0 の mBSSRP-S の解は mBSSRP の実行可能解となる。数値実験の結果、mBSSRP-S を解いて mBSSRP の実行可能解を得ることで mBSSRP を解くよりも優れた解を得られるが、計算時間は増加する [3]。そこで本稿では、計算時間を短縮するために近傍数を削減した解法を提案し、その性能を報告する。

2 複数の自転車回収車を用いた自転車再配置問題

mBSSRP と mBSSRP-S では、1 つのデポと自転車が超過・不足している 2 種類のポートが n 個与えられる。mBSSRP の目的は、容量制約のある複数の回収車が超過ポートにて自転車を回収し、不足ポートへ補充するとき、1) 時間制約、2) 積み込み制約、3) 補充制約を満たす巡回路の中で回収車の総移動時間が最短となる巡回路を求めることである。

一方、mBSSRP-S は上記の 3 つの制約を取り除き、これらの制約に対する違反量を目的関数に加えた問題である。mBSSRP-S の目的関数は、 $T(t) + \alpha(t) \times E(t) + \beta(t) \times P(t)$ である。 $T(t)$ は t イタレーション目の回収車の総移動時間であり、これは mBSSRP の目的関数である。 $E(t)$ は t イタレーション目の巡回路において回収車が制限時間を超えて作業を行った時間、 $P(t)$ は、 t イタレーション目に回収・補充できなかった自転車台数を示す。 $\alpha(t)$ と $\beta(t)$ は t イタレーション目の重み係数である。

3 近傍数を削減した解法

mBSSRP の良好な近似解を得るために、タブーサーチを用いた手法を提案している [4]。この解法ではタブーサーチが次の解へ遷移するための近傍を Or-opt と CROSS-exchange で生成している。Or-opt とは、ある巡回路の部分順列 $i-j$ を他の巡回路中に挿入する操作である (図 1 (a))。1 つの挿入操作に対し部分順列を正順、または逆順にして挿入する 2 通りがある。CROSS-exchange は、ある巡回路の部分順列 $i-j$ と他の巡回路の部分順列 $k-l$ を交換する操作である (図 1 (b))。CROSS-exchange も部分順列を正順と逆順で交換できるため 1 つの交換操作に対し 4 通りの解が存在する。

提案解法で実行される挿入・交換操作の割合を調べた結果、正順の部分順列を挿入する Or-opt が約 65%、正順の部分順列を交換する CROSS-exchange が約 25% 実行されており、部分順列を逆順にして挿入・交換する操作はほとんど行われていないことがわかった。そこで、部分順列を正順で挿入・交換する場合のみに限定し、計算時間の短縮を試みる。

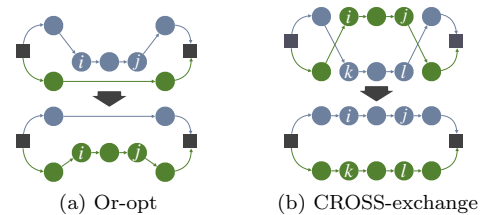


図 1 近傍操作

4 数値実験・結果

ポート数が 100 箇所の問題を 3 つ作成し、性能評価を行う。過不足台数は $-5 \sim +5$ 台とし、最大積載容量 10 台の回収車 5 台が作業を行う。回収車は 30[km/h] で移動し、1 台あたりの作業時間は 2 分、制限時間は 180 分とする。タブー期間は 50、イタレーションは 1,000、重み係数のパラメータ値は $\alpha(0) = 1$, $\beta(0) = 1$, $\lambda = 1.07$, $\mu = 0.30$ とした。

表 1 に mBSSRP-S を解き得られた mBSSRP の結果を示す。表 1 より、近傍数を削減したことで、問題 1, 2 ではわずかに性能が劣るが計算時間は 50% 以上短縮することができた。問題 3 については計算時間を削減しながら、良い性能を得ることができた。

表 1 mBSSRP-S を解き得られた mBSSRP の結果
(a) 全ての近傍解を用いた場合

No.	平均	最良	最悪	計算時間 [s]
1	119,966.02	116,229	125,588	461.10
2	118,408.50	113,405	123,847	457.67
3	113,967.06	110,167	119,364	449.33

(b) 近傍解を限定した場合

No.	平均	最良	最悪	計算時間 [s]
1	121,039.18	116,338	125,900	217.13
2	118,623.86	113,954	124,618	212.58
3	113,808.58	109,525	118,751	209.01

5 まとめ

本報告では、提案解法の性能を維持したまま、計算時間を短縮するために遷移頻度の高い 2 つの近傍探索に絞った解法を提案し、その性能を調査した。数値実験の結果、同等の性能を得ながら、計算時間を 50% 以上削減することに成功した。なお、本研究は JSPS 科研費 (No. JP19K04907, 15KT0112, 17K00348) の援助を受けて行われた。

参考文献

- [1] 對馬帆南ら, 2017 年 電子情報通信学会 NOLTA ソサイエティ大会, A-16, 2017.
- [2] 對馬帆南ら, 信学技報, **119**(19), 65-70, 2019.
- [3] 對馬帆南ら, 信学技報, **119**(381), 13-18, 2020.
- [4] H. Tsushima et al., submitted, 2020.

その他発表会

自転車再配置問題に対する計算時間を短縮するための探索戦略

對馬帆南

東京理科大学 博士1年 池口研究室

1 はじめに

バイクシェアリングシステム (BSS) で生じる自転車の過不足台数を調整するために、複数の自転車回収車を用いた自転車再配置問題 (mBSSRP) を提案している [1]. しかし, mBSSRP は制約条件が厳しいため, 提案解法は一部の問題例に対して実行可能解を得ることができない問題があった. この問題を解決するため, mBSSRP の一部の制約を取り除き, それらの制約に対する違反量を重み付きのペナルティとして mBSSRP の目的関数に加えたソフトな制約付き mBSSRP (mBSSRP-S) を提案した [2]. mBSSRP-S では mBSSRP の実行不可能解へも遷移しながら違反量が 0 となる mBSSRP の解に遷移することを目指す. 数値実験の結果, mBSSRP-S を解いて mBSSRP の実行可能解を得る戦略は mBSSRP を解いた場合よりも優れた解を得られるが, 計算時間は増加する [3]. 本発表では, 性能を維持しながら計算時間を削減する探索戦略について報告する.

2 複数の自転車回収車を用いた自転車再配置問題

mBSSRP と mBSSRP-S では, 1 つのデポと自転車が超過・不足している 2 種類のポートが n 個与えられる. mBSSRP の目的は, 容量制約のある複数の回収車が超過ポートにて自転車を回収し, 不足ポートへ補充するとき, 1) 時間制約, 2) 積込み制約, 3) 補充制約を満たす巡回路の中で回収車の総移動時間が最短となる巡回路を求めることである.

一方, mBSSRP-S は上記の 3 つの制約を取り除き, これらの制約に対する違反量を目的関数に加えた問題である. mBSSRP-S の目的関数は, $T(t) + \alpha(t) \times E(t) + \beta(t) \times P(t)$ である. $T(t)$ は t イタレーション目の回収車の総移動時間であり, これは mBSSRP の目的関数である. $E(t)$ は t イタレーション目の巡回路において回収車が制限時間を超過して作業を行った時間, $P(t)$ は, t イタレーション目に回収・補充できなかった自転車台数を示す. $\alpha(t)$ と $\beta(t)$ は t イタレーション目の重み係数である.

3 計算時間を短縮するための探索戦略

mBSSRP の良好な近似解を得るために, タブーサーチを用いた手法を提案している [4]. この解法ではタブーサーチで次の解へ遷移するための近傍解を Or-opt と CROSS-exchange で生成している. Or-opt とは, ある巡回路の部分順列を他の巡回路中に挿入する操作である. 1 つの挿入操作に対し部分順列を正順, または逆順にして挿入する 2 通りがある. CROSS-exchange は, ある巡回路の部分順列と他の巡回路の部分順列を交換する操作である. CROSS-exchange も部分順列を正順と逆順で交換できるため 1 つの交換操作に対し 4 通りの解が存在する.

提案解法で実行される挿入・交換操作の割合を調べた

結果, 正順の部分順列を挿入する Or-opt が約 66%, 正順の部分順列を交換する CROSS-exchange が約 20% 実行されており, 部分順列を逆順にして挿入・交換する操作はほとんど行われていなかった. また, mBSSRP-S を解いたときの解探索の様子を調査すると, 1 度 mBSSRP の実行可能解を探索すると連続して実行可能解を探索することもわかった. これら 2 つの傾向から, mBSSRP-S を解く際, 実行可能解を得る前と実行可能解を得た後で探索戦略を変える必要がある. 実行可能解を得る前までの戦略として, 全ての近傍解を探索した場合と最も遷移度の高い部分順列を正順で挿入・交換する場合にのみ限定した場合の比較実験を行う.

4 数値実験・結果

ポート数が 100 箇所の問題を 10 個作成し, 性能評価を行う. 過不足台数は $-5 \sim +5$ 台とし, 最大積載容量 10 台の回収車 5 台が作業を行う. 回収車は $30[\text{km/h}]$ で移動し, 1 台あたりの作業時間は 2 分, 制限時間は 180 分とする. タブー期間は 50, イタレーションは 1,000, 重み係数のパラメータ値は $\alpha(0) = 1, \beta(0) = 1, \lambda = 1.07, \mu = 0.30$ とした.

表 1 に mBSSRP-S を解いて初めて mBSSRP の実行可能解が得られたときのイタレーション, 計算時間, 50 個の初期解から得られた実行可能解の数を示す. 表 1 より, 全ての近傍解を探索した場合は近傍解を限定した場合に比べると計算時間は劣ってしまうが, 全ての試行で実行可能解を得られていること, 短いイタレーションで得られていることから, 実行可能解を探索する前の探索戦略としては, 全ての近傍解を探索することが有効である.

表 1 mBSSRP-S を解き得られた mBSSRP の結果

	全ての近傍解を探索した場合	近傍解を限定した場合
イタレーション	219.63	263.77
計算時間 [s]	103.13	55.89
得られた実行可能解の数	50	45

5 まとめ

本報告では, mBSSRP-S を解く際に, 性能を維持しながら, 計算時間を短縮するために実行可能解を得る前と実行可能解を得た後で探索戦略を変更する解法を提案した. 実行可能解を得る前の探索戦略としては全ての近傍解を探索することが有効である. 今後は, 実行可能解を得た後の探索戦略について考える必要がある.

参考文献

- [1] 對馬帆南ら, 2017 年 電子情報通信学会 NOLTA ソサイエティ大会, A-16, 2017.
- [2] 對馬帆南ら, 信学技報, **119**(19), 65–70, 2019.
- [3] 對馬帆南ら, 信学技報, **119**(381), 13–18, 2020.
- [4] H. Tsushima et al., submitted, 2020.

リカレンスプロット密度を変化させることによる 決定論的カオスと非線形確率時系列の特性抽出

金丸志生

東京理科大学 修士2年 池口研究室

1 はじめに

非線形時系列解析手法の一つにリカレンスプロット (RP) [1] がある. RP は時系列データの振る舞いを 2 次元の白黒画像で表現したものである. RP には画像中の黒画素の割合 (以下, RP 密度と呼ぶ) を定めるパラメータ θ がある. 我々は RP 密度を徐々に大きくしたときの RP の変化を定量的に調べることで, 非線形時系列の特性を捉える手法を提案した [2-4]. これらの手法では, まず, RP やその拡張手法 [5] に対して, RP 密度の異なる複数の画像を作成する. 次に, それらに対して再帰定量化解析 (RQA) [6] の一つである DET と, 生起確率 R を算出して RP の変化を定量化する. 文献 [2-4] では, RP 密度を変化させることで得られる DET と R の推移傾向から, 時系列データの特性を捉えられることが示されている.

しかしながらこれらの先行研究において, 決定論的ダイナミクスと確率論的ノイズが混在するランダム力学系を対象とした調査は行われていない. そこで本稿では, 決定論的カオスと周期解にノイズを含む時系列の特性を捉えることを目的とする. 具体的には, 決定論ダイナミクスであるロジスティック写像 [7] に確率論的ノイズが混在したノイジーロジスティック写像 [8,9] から得られる時系列を対象に, 提案手法の有効性を調査する.

2 提案手法

本稿で解析対象とする時系列データはいずれも 1 次元のみ観測できることを仮定する. そのため RP の作成に必要な状態空間を再構成する手法として遅延座標系への埋め込み手法を用いる [10,11]. 本稿では, 再構成して得られた状態空間から, 2 点間距離情報をもとに RP を作成する [1]. また, 2 つの軌道ベクトルの差のユークリッドノルムより同方向性を可視化する IDRP を作成し, RP と IDRP を用いて, 近傍かつ同方向の関係性を可視化する IDNP を作成する [5].

RP 密度は, 再構成状態空間内の任意の 2 点間についての情報 (RP であれば 2 点間距離, IDRP であれば軌道ベクトルの差のユークリッドノルム) の最小値から最大値までの値に, $0 \leq \theta \leq 1$ の割合を掛けて調節する. DET は, 分母に黒画素の総数, 分子に 45° の斜め線を構成する黒画素の総数とした, RP 上に現れる斜め線の割合を表す. 生起確率 R は, 分母に RP の総プロット数, 分子に IDNP の総プロット数とした, IDNP の同方向的近傍点数の RP の近傍点数に対する割合を表す. DET と R はいずれも時系列が決定論的であると値は高くなると考えられる.

3 対象時系列

本稿で対象とするランダム力学系であるノイジーロジスティック写像 [8,9] を式 (1) に示す.

$$x_{t+1} = ax_t(1 - x_t) + \sigma(x_{t-1})\varepsilon_t \quad (1)$$

ただし, x_t は離散時刻 t のときの状態値, a は分岐パラメータ, ε_t は $(-0.5, 0.5)$ の一様分布に従う確率変数であり, $\sigma(x)$ は $\sigma(x_{t-1}) = 0.5\min(x_t, 1 - x_t)$ と定義される.

4 数値実験

RP 密度を大きくしたときの DET と R の値の推移から, 決定論的カオスと周期解にダイナミカルノイズを含む時系列の特性を捉えられるかを調査した. 数値実験には, ロジスティック写像において 2, 4, 8, 16, 32, 64, 128 周期応答, カオス応答になるパラメータ値 $a = 3, 3.45, 3.55, 3.565, 3.569, 3.5698, 3.5699, 4$ を用いた. 各時系列長は過渡状態として 500 点省いた後の 1,000 点とし, 再構成状態空間の次元は 3, 遅れ時間は 1 とした. 同方向性を算出するのに必要な軌道ベクトルは, 1 ステップ後の点までのベクトルとした.

ロジスティック写像とノイジーロジスティック写像について各 a を与え, 初期値と乱数シードを変更して 100 本の時系列を対象に, 提案手法を適用して得られた指標の推移傾向を図 1 に示す. 横軸は対数軸で RP 密度を定めるパラメータ θ , 縦軸は各指標を示す.

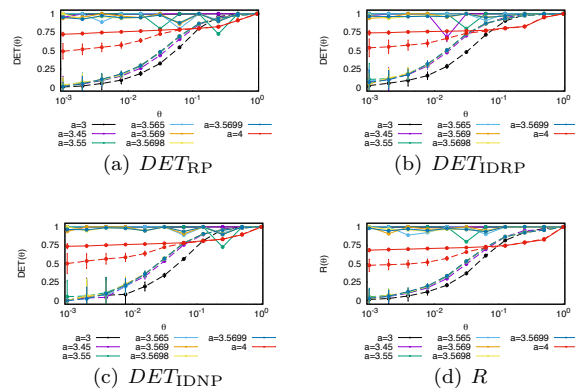


図 1. 1 つのパラメータ a に対して, 提案手法を初期値と乱数シードを変えて 100 回適用した結果. 実線はロジスティック写像, 破線はノイジーロジスティック写像の応答を対象としたときの各指標を表し, プロットは平均値, エラーバーは標準偏差を表す. 各図は, 閾値 θ を変化させて作成した (a) RP, (b) IDRP, (c) IDNP の DET と, (d) 生起確率 R の推移傾向を示す.

図 1 に示したいずれの指標においても, $a = 4$ のロジスティック写像のカオス応答に対する推移傾向 (赤い実線) と, その他の周期応答となる a でのノイジーロジスティック写像の応答に対する推移傾向 (赤以外の破線) は異なる. 従って, カオス応答と周期解にダイナミカルノイズを含む応答の判別が可能となることが示唆された.

5 まとめ

本稿では RP 密度を変化させることで得られる指標の推移傾向から, カオス時系列と周期解にダイナミカルノイズを含む時系列の特性抽出が可能であるかを調査した. 結果から, 閾値が小さい範囲での各指標の値は, 時系列がカオス応答であれば高く, 周期解にダイナミカルノイズを含む応答であれば低くなるという応答に応じて異なる特性が得られた. 一方で, RP 適用時に経験的に扱われる RP 密度である 0.1 付近や大きい範囲では, これらの見分けがつかないため, RP 密度を変化させることによって得られる推移傾向を時系列の特性とすることの有効性が示唆された.

参考文献

- [1] J.-P. Eckmann *et al.*, *Europhys. Lett.*, Vol. 4, pp. 937-977, 1987.
- [2] 小川 徳紀 他, 電子情報通信学会 2019 年総大会講演論文集, N-1-33, 2019.
- [3] 金丸 志生 他, 電子情報通信学会 技術研究報告, Vol. 119, No. 19, NLP2019-5, pp. 23-28, 2019 年 5 月.
- [4] S. Kanamaru *et al.*, *Proceedings of the 2019 International Symposium on Nonlinear Theory and its Applications (NOLTA2019)*, pp. 536-539, 2019.
- [5] 寶来 俊介 他, *電気学会論文誌. C*, Vol. 122, No. 1, pp. 141-147, 2002.
- [6] N. Marwan *et al.*, *Physics Reports*, Vol. 438, Nos. 5-6, pp. 237-329, 2007.
- [7] R. M. May, *Nature*, Vol. 261, No. 5560, pp. 459-467, 1976.
- [8] M. Robert A., *Springer Science & Business Media*, 2010.
- [9] K. Mohadeseh *et al.*, *International Journal of e-Navigation and Maritime Economy*, Vol. 6, pp. 53-63, 2017.
- [10] F. Takens, by DA Rand and L.-S. Young *Springer*, Berlin, Vol. 898, pp. 366-381, 1981.
- [11] T. Sauer *et al.*, *Journal of Statistical Physics*, Vol. 65, Nos. 3-4, pp. 579-616, 1991.

非線形力学系理論に基づいたマーク付点過程間の因果性検出

澤田和弥

東京理科大学 修士2年 池口研究室

1 はじめに

地震や神経活動, 経済活動は点過程データとして観測される. 我々は既に, 一定のサンプリング間隔で観測された時系列データ間の因果性を検出可能な Convergent Cross Mapping (CCM) [1] を拡張することで, 神経細胞の数理モデルから得られる発火間隔時系列に対して因果性を検出する手法 (ISICCM 法) を提案し, 点過程データに対しても高精度な因果検定が可能となることを示した [2, 3]. 一方, マーク付点過程に対しても ISICCM 法が有効となるかを検証することも重要である. そこで本発表では, 数理モデルから生成したマーク付点過程データに対して ISICCM 法により因果性が検出可能であることを調査したので報告する.

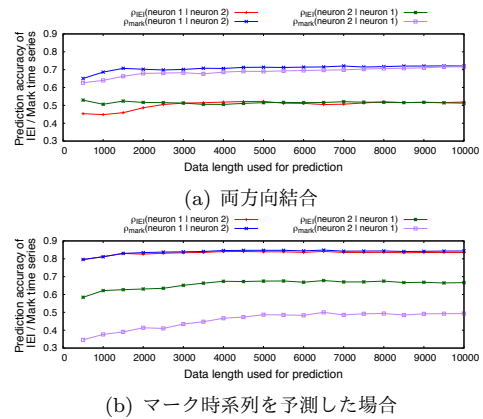
2 解析手法とデータ生成方法

本発表では, マーク付点過程間の因果性解析に ISICCM 法 [2, 3] を拡張した手法を用いる. ISICCM 法は発火間隔 (ISI) 時系列から状態空間を再構成し, 発火時間による時間対応を取ることで ISI 時系列を予測する方法である. 本発表では, 状態空間を再構成する際, イベント間隔 (IEI) 時系列とマーク時系列の 2 次元時系列から 10 次元の状態空間を再構成し, ISICCM 法と同様の手順で IEI 時系列とマーク時系列をそれぞれ独立に予測した. その際, 過渡状態として 100,000 点を除外し, 時系列長は 500 から 500 刻みで 10,000 まで変化させた. また, マーク付点過程は, gap junction を伴う Izhikevich ニューロンモデル: $\dot{v}_i = 0.04v_i^2 + 5v_i + 140 - u_i + I_i + \sum_{j=1, j \neq i}^N w_{ij}(v_j - v_i)$, $\dot{u}_i = a_i(b_i v_i - u_i)$ [4, 5] から得られた回復変数の時系列の極大値を抽出することで生成した [6]. ここで, $v_i \geq 30$ のとき, $v_i \leftarrow c_i$, $u_i \leftarrow u_i + d_i$ にリセットされる. ただし, 今回は $N = 2$ とした. また, パラメータは, $a_i = 0.03$, $b_i = 0.3$, $c_i = -55$, $d_i = 2$, $I_i = 10$ とし, 両方向結合の場合は $w_{12} = w_{21} = 0.05$ とし, 片方向結合の場合は $w_{12} = 0$, $w_{21} = 0.05$ とした.

3 数値実験

IEI 時系列とマーク時系列から生成した 10 次元再構成状態空間を用いて, IEI 時系列とマーク時系列を予測した結果を図 1 に示す. 図 1 の (a) は両方向結合の場合, 図 1 の (b) は片方向結合の場合の結果である. 横軸は予測に用いたデータ長, 縦軸は予測精度 (相関係数) を表す. また, $\rho_{IEI}(\text{neuron } i | \text{neuron } j)$ はニューロン j のマーク付点過程を用いてニューロン i のマーク付点過程の IEI を予測した結果を表し, $\rho_{\text{mark}}(\text{neuron } i | \text{neuron } j)$ はニューロン j のマーク付点過程を用いてニューロン i のマーク付点過程のマーク値を予測した結果を表す.

図 1 の (a) を見ると, ニューロン 1 のマーク付点過程を用いて, ニューロン 2 のマーク付点過程の IEI を予測



(b) マーク時系列を予測した場合

図 1 予測に用いたデータ長と拡張した ISICCM 法による予測精度の関係

した場合を除いて, 予測に用いるデータ長を増加させると予測精度は上昇し, 0.5 から 0.7 程度の比較的高い値に収束している. 図 2 の (b) を見ると, ニューロン 2 のマーク付点過程からニューロン 1 のマーク付点過程を予測した場合, データ長の増加に伴い予測精度は 0.8 以上の高い値に収束している. この結果は, ニューロン 1 からニューロン 2 への因果性を正しく検出できていることを示唆する. また, ニューロン 1 のマーク付点過程からニューロン 2 のマーク付点過程を予測した場合は, 本来であれば予測精度は上昇せず小さい値に収束するはずだが, 0.5 から 0.6 程度の値に上昇している. この結果は, 同一のパラメータを使用しているために, 力学系の特徴が似ていることが原因だと考えられる.

4 まとめ

本発表では, 既に我々が提案した点過程データ間の因果性検出手法 (ISICCM 法) [2, 3] をマーク付点過程に拡張し, 適用した結果を報告した. 具体的には, IEI 時系列とマーク時系列の 2 次元時系列から状態空間を再構成した点が拡張部分である. その結果, マーク付点過程データに対しても拡張した ISICCM 法は有効である可能性が示唆されたが, 力学系の特徴が似ている場合は, 因果性を誤検出してしまう場合があることがわかった. 今後は, サロゲートデータ [7] を使用して, 予測精度の上昇および収束が因果性によるものであるかの検定を行う必要がある. また, イベントタイミングの差を考慮することで, より高精度な予測器を生成することも重要であると考えられる.

参考文献

- [1] G. Sugihara et al., *Science*, **338**, 6106, 496–500, 2012.
- [2] 澤田 和弥 他, 信学ソ大, N-1-12, 2019.
- [3] 澤田 和弥 他, 信学総大, N-1-9, 2020.
- [4] E. M. Izhikevich, *IEEE TNN*, **14**, 6, 1569–1572, 2003.
- [5] 内木 楓 他, 信学論 A, **J100-A**, 5, 195–204, 2017.
- [6] N. Yabuta and T. Ikeguchi, *Proc. of NOLTA*, 188–191, 2007.
- [7] M. Thiel et al., *EPL*, **75**, 4, 535–541, 2006.

マーク付点過程に対する Visibility Graph の有効性の検討

毛福佳

東京理科大学 修士2年 池口研究室

1 はじめに

我々は、楽曲データをマーク付点過程データとして扱い、ネットワークに変換することで、複雑ネットワーク理論の観点から楽曲解析を行ってきた [1][2]. その際、Visibility Graph [3] (以下、VG 法) と呼ばれる時系列データをネットワークに変換する手法を使用した. しかし、VG 法は時系列データをネットワークに変換する手法であるため、マーク付点過程に適用するとき、作成したネットワークが元のマーク付点過程の特徴を反映できていない可能性が残されている. そこで、本研究では力学系非線型常微分方程式から作成したマーク付点過程に対して VG 法を適用し、ネットワークを作成する. 次に、ネットワーク間距離 [4] を用いて作成したネットワーク間の構造的特徴の差異を解析することで、マーク付点過程に対する VG 法の有効性を調査した結果を報告する.

2 データ

本研究では、マーク付点過程を作成するため、非線型常微分方程式であるローレンツ方程式 [5] とレスラー方程式 [6] 式を用いた. ルンゲクッタ法を用いて各方程式から得られた時系列データに対して、極大値を抽出し、マーク付点過程データを作成した. ローレンツ方程式のパラメータを $\sigma = 10$, $b = 3/8$, $r = 28$ に設定し、レスラー方程式のパラメータを $a = 0.2$, $b = 0.2$, $c = 5.7$ に設定した.

3 実験手順

まず、ローレンツ方程式から得られた時系列を用いてマーク付点過程を作成する. 次に、得られたマーク付点過程を図 1(a) に示すように複数の時間窓に分割し、マーク付点過程 w_1, w_2, \dots, w_n を作成する. その際、全てのマーク付点過程のイベント数が 200 個以上存在するように時間窓の幅を設定する. 次に、マーク付点過程 w_1, w_2, \dots, w_n に対して VG 法を適用し、図 1(b) のようにネットワーク G_1, G_2, \dots, G_n を作成する. 最後に、ネットワーク G_1 と他のネットワーク G_2, G_3, \dots, G_n のネットワーク間距離 $d(G_1, G_2), d(G_1, G_3), \dots, d(G_1, G_n)$ を算出する. なお、同じ力学系非線型常微分方程式であるローレンツ方程式から得られたマーク付点過程 w_1, w_2, \dots, w_n の構造が類似するため、ネットワーク G_1, G_2, \dots, G_n の構造も類似し、 $d(G_1, G_2), d(G_1, G_3), \dots, d(G_1, G_n)$ の値が小さいと考えられる. また、ネットワーク間距離の値が十分小さいことを確認するため、異なる力学系であるレスラー方程式から同じ手法で複数のネットワークを作成し、ローレンツ方程式とレスラー方程式から作成した各ネットワークのネットワーク間距離を比較することで、ネットワーク間距離の値が有意であることを確認する.

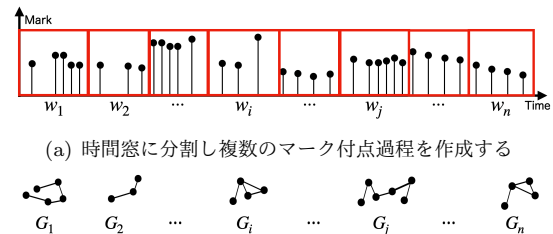


図 1 ローレンツ方程式からネットワークを作成する例.

4 結果

得られたマーク付点過程に対して VG 法を適用し、ネットワーク間距離を算出した結果を図 2 に示す. ここで、横軸は作成したネットワークのインデックス、縦軸はネットワーク間距離の値である. また、青線は同じローレンツ方程式から作成したネットワークのネットワーク間距離、赤線はローレンツ方程式から作成したネットワークとレスラー方程式から作成した各ネットワークのネットワーク間距離の値である.

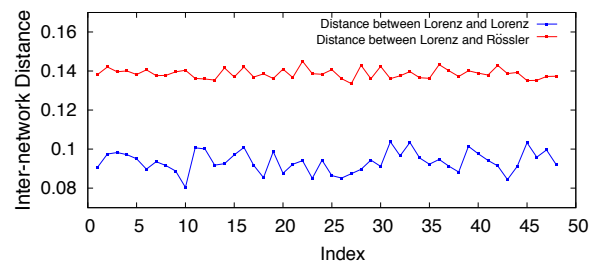


図 2 ネットワーク間距離を算出した結果

図 2 から、同じ力学系であるローレンツ方程式から作成したネットワークのネットワーク間距離の値が小さいため、各ネットワークの構造的特徴が類似することが分かる. これは、VG 法を用いることでマーク付点過程の特徴が反映されていることを示唆する.

5 まとめ

本報告では、VG 法をマーク付点過程に適用するとき、作成したネットワークが元のマーク付点過程の特徴を反映できるかの検証を行った. 結果として、マーク付点過程に対する VG 法が有効であることが示唆された.

参考文献

- [1] 毛福佳, 他, 信学技報, Vol. 119, No. 19, 47–52, 2019.
- [2] F. Mao *et al.*, NOLTA2019, 587–590, 2019.
- [3] L. Lacasa *et al.*, PNAS, Vol. 105, No. 13, 4972–4975, 2008.
- [4] Y. Shimada *et al.*, Scientific Reports, Vol. 6, 34944, 2016.
- [5] E. N. Lorenz, JAS, Vol. 20, No. 2, 130–141, 1963.
- [6] O.E.Rössler, Physics Letters A, Vol. 57, No. 5, 1976.

ニューロン比率と物理的距離がシナプスの結合性へ与える影響

宮南風

東京理科大学 修士2年 池口研究室

1 はじめに

生理学的知見によれば、大脳皮質では興奮性ニューロンと抑制性ニューロンは約 4:1 の比率 (以下、生理学的比率) で存在するといわれている [1]。しかし、なぜ生理学的比率が存在するのかについては明らかになっていない。これまでの研究で、生理学的比率に何らかの最適性を持つことをニューロンモデルを用いた数値実験で示した [2, 3, 4]。

さらに、ダウン症などの脳機能障害は生理学的比率に従っていないことが一因であると考えられており [5, 6]、ダウン症患者と健常者において、計測する領域間の物理的距離に応じた fMRI 信号の相関関係が異なることがわかっている [7]。

本発表では、前述の知見を検証する目的の下、ニューロン比率ごとのシナプス伝達情報量と物理的距離との関係性について数理モデルを用いて調査する。具体的には、まず異なるニューロン比率で生成された各ニューラルネットワーク上のシナプスのうち、情報量が上位のものについて着目する。それらのシナプス伝達遅延値の分布を算出し、ニューロン比率ごとに比較した。

2 シナプス伝達情報量

シナプスの情報量として、文献 [4] の伝達情報量 $TR[i \rightarrow j]$ を用いる。これは、シナプスを挟む 2 つのニューロン間のスパイク列から算出する Delayed Transfer Entropy を拡張したものである。伝達情報量 $TR[i \rightarrow j]$ が高いとき、ニューロン i から j への情報が多く流れていることを意味する。

3 数値実験

全ニューロンに対する抑制性ニューロン比率の割合を r とする。 $r = 0, 0.2$ (生理学的比率), $0.4, 0.6$ に対する文献 [4] と同様のニューラルネットワークを用意し、STDP 学習 [8] を行う。十分に学習が収束した ($t = 3,600$ [s]) 時刻以降において、学習を終了させる。さらに、一定期間 ($t = 5,400 \sim 7,200$ [s]) のスパイク列を取得し、伝達情報量 TR をシナプスごとに算出した。最後に、全てのシナプスのうち、伝達情報量が上位 20 % のシナプスを抽出し、物理的距離との関係性を調査した。また、ここでの物理的距離は数理モデル上のシナプス伝達遅延値 δ とした。 δ が小さい場合、物理的距離は近いとみなし、大きい場合は遠いとみなす。

4 実験結果

伝達情報量 TR が上位 20 % のシナプスの伝達遅延値 δ の頻度分布を図 1 に示す。生理学的比率である $r = 0.2$ では頻度がほぼ同等であるのに対し、それ以外の比率では分布に偏りがあることがわかる。特に抑制性ニューロ

ンが存在しない $r = 0$ のとき、遅延値 δ が小さいシナプスの占める割合が高く、大きいシナプスの割合が低い。fMRI 信号を解析した実験より、生理学的比率に従わないダウン症患者の脳内では、物理的距離が短距離の領域間の信号間の相関は高く長距離の領域間の相関は低いことがわかっている [7]。fMRI 信号が得られる領域の組と数理モデル上のニューロン間のシナプスは一対一には対応しないものの、これらの知見と一致していることがわかる。

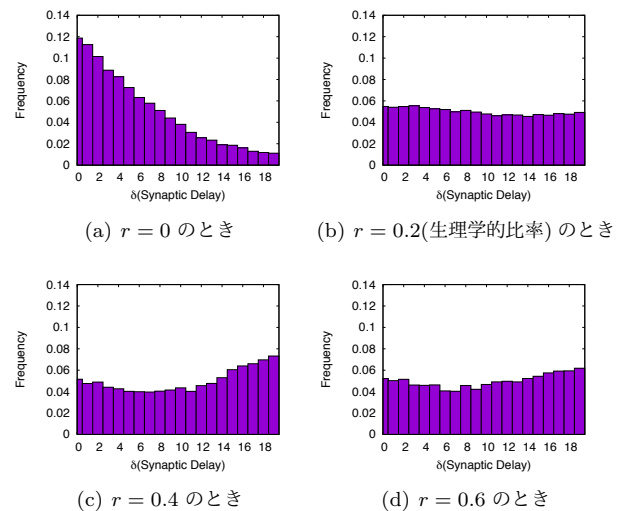


図 1 上位 10% の伝達情報量 TR を有するシナプスの伝達遅延 δ の分布

5 まとめ

本発表では、異なるニューロン比率でのニューラルネットワーク上で算出されたシナプス伝達情報量について、上位のシナプスのみを抽出し、評価を行った。その結果、シナプス伝達遅延値の頻度分布は生理学的比率でない場合に偏りが生じた。これは、fMRI 信号を解析した実験から得られた知見を支持する結果となった。今後は、ニューロンモデルより生成された神経活動データから Balloon モデル [9][10] を用いて、fMRI 信号に変換し、ニューロン比率ごとの相関関係を比較する。

参考文献

- [1] De Felipe J., et al., *Prog. Neurobiol.*, **6**(39), 563–607, 1992.
- [2] 宮南風 他, 信学総大, N-1-25, 2019.
- [3] 宮南風 他, 信学ソ大, N-1-10, 2019.
- [4] 宮南風 他, 信学ソ大, N-1-10, 2020.
- [5] P.V. Belichenko et al., *J. Comp. Neurol.*, **480**(3), 281–298, 2004.
- [6] J.E. Hanson et al., *J. Physiol.*, **579**(1), 53–67, 2007.
- [7] Jeffrey S. Anderson et al., *Neuroimage Clin.*, **2**, 703–715, 2013.
- [8] Bi G. Q. et al. *J. Neurosci.*, **18**(8), 10464–10472, 1998.
- [9] Buxton R. et al., *Magn Reson Med* **39**(6), 855–864, 1998.
- [10] Friston, K. J. et al., *NeuroImage*, **12**(4), 466–477, 2000.

テンポラルネットワークの構造が情報拡散に与える影響の解析

ジョロエン

東京理科大学 修士1年 池口研究室

1 はじめに

近年、複雑ネットワーク上での情報拡散に関する研究が盛んに行われている [1,2]. 我々は既に、情報拡散の数理モデルを提案し、静的なネットワーク構造が情報拡散に与える影響について調査した結果を報告している [3,4]. 一方、現実世界には人間関係を枝とするような静的なネットワークだけでなく、人と人のコンタクトを枝とするような時間的に構造が変化するテンポラルネットワークも数多く存在する. そこで本発表では、実世界のコンタクトデータから作成したテンポラルネットワークに対して、情報への関心度に着目した情報拡散モデル [3] を用いた数値実験によって、テンポラルネットワークの構造が情報拡散に与える影響について調査を行った結果を報告する.

2 用いたテンポラルネットワーク

本発表では、フランスの高校と病院で観測された、人同士のコンタクトのデータ [5-7] をテンポラルネットワークに変換し、2つのテンポラルネットワーク上で情報拡散の数値実験を行った. 高校と病院でのデータはそれぞれ、126名(教員と学生)の4日間の対面接触記録と75名(医療スタッフと患者)の5日間の対面接触記録である. また、高校と病院でのコンタクトは20秒ごとに観測され、20秒あたりの平均コンタクト数はそれぞれ2.09と1.80である. 時刻 t で実験協力者 i と j にコンタクトが存在する場合、時刻 t でのテンポラルネットワークの隣接行列 $\mathbf{A}(t)$ の (i, j) 成分は $A_{ij}(t) = A_{ji}(t) = 1$, コンタクトが存在しない場合は、 $A_{ij}(t) = A_{ji}(t) = 0$ とする.

3 情報への関心度に着目した情報拡散モデル

本発表では、情報への関心度に着目した情報拡散の数理モデルを用いて数値実験を行った. 時刻 t での頂点 i の情報への関心度を $x_i(t)$ とする. 関心度 $x_i(t)$ は式(1)に従って変化する.

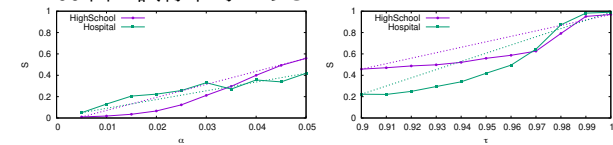
$$x_i(t+1) = g_i(\tau, t)x_i(t) + \alpha H(x_i(t) - \theta_i) \sum_{j \in G_i(t)} F(x_j(t) - \theta_j) \quad (1)$$

ここで、 t_i^* を頂点 i が最初に情報を受信した時刻、 $\tau(0 \leq \tau \leq 1)$ を情報への関心度の減衰係数とすると、 $g_i(\tau, t)$ は、 $t - t_i^* \equiv 0 \pmod{3600}$ の場合に $g_i(\tau, t) = \tau$, それ以外の場合に $g_i(\tau, t) = 1$ とする. 従って、頂点 i は初めて情報を受信してから、3600秒ごとに情報への関心度が減衰する. また、 $F(x)$ はステップ関数、 $H(x)$ は $x \geq 0$ の場合 $H(x) = 0$, $x < 0$ の場合 $H(x) = 1$ となる関数である. さらに、 $G_i(t)$ は時刻 t での頂点 i の隣接頂点の集合、 θ_i は頂点 i の閾値、 α は情報の影響力である. 従って、頂点 i は自身の閾値 θ_i に達するまで隣接頂点からの

情報を受信し続け、閾値に達すると情報を送信する. なお、各頂点の閾値は、範囲が $[0, 1]$ で平均 $\mu = 0.5$, 標準偏差 $\sigma = 0.2$ の切断正規分布とした.

4 数値実験

高校と病院のテンポラルネットワーク上で、情報拡散の数値実験を行い、情報の最終拡散率 S について調査した. 最終拡散率 S は、最後のコンタクトが終わるまでに、一度閾値に達した頂点数の全頂点に対する割合である. また、拡散開始の条件として、最初に接触した2つの頂点のうち、一方の関心度を1にすることで情報を拡散させた. 減衰係数 τ を $\tau = 0.95$ に固定し、情報の影響力 α を $0.005 \leq \alpha \leq 0.05$ の範囲で変化させた場合の S の結果を図1(a)に示す. また、 $\alpha = 0.05$ に固定し、 τ を $0.9 \leq \tau \leq 1$ の範囲で変化させた場合の S の結果を図1(b)に示す. 図1における実線と破線はそれぞれ最終拡散率とその増加傾向を示す. なお、本発表での結果は全て200回の試行平均である.



(a) α と S の関係 ($\tau = 0.95$) (b) τ と S の関係 ($\alpha = 0.05$)

図1: 高校と病院のテンポラルネットワーク上での情報の最終拡散率 S

図1(a)から、高校での最終拡散率 S の増加傾向は病院より大きいため、高校での S は、病院に比べて情報の影響力 α に強く依存することがわかった. 図1(b)から、病院での最終拡散率 S の増加傾向は高校より大きいため、病院での S は、高校に比べて減衰係数 τ に強く依存することがわかった. これらの傾向は、単位時間あたりの平均コンタクト数が、病院に比べて高校の方が大きいためだと考えられる.

5 まとめ

本発表では、情報拡散モデルを用いて、高校と病院のテンポラルネットワークの構造が情報拡散に与える影響について調査した. その結果、コンタクトの時間的密度が大きい場合の最終拡散率は、情報の影響力に比較的強く依存することが示唆された、一方、コンタクトの時間的密度が小さい場合の最終拡散率は、関心度の減衰に比較的強く依存することが示唆された. 今後は、この仮説を検証するために、コンタクトの時間的密度を変化させたサロゲートデータを用いた検定を行う必要がある.

参考文献

- [1] D. J. Watts & S. H. Strogatz, *Nature*, **393**, 440–442, 1998.
- [2] D. Centola, *Science*, **329**, 1194–1197, 2010.
- [3] A. Tanaka et al., *Proc. of NOLTA2013*, 298–301, 2013.
- [4] ジョ, 他, *信学技報*, **119**, 471, 83–88, 2020.
- [5] <http://www.sociopatterns.org/> (最終閲覧日: 2020年09月25日)
- [6] P. Vanhems et al., *PLOS ONE*, **8**, e73970, 2013.
- [7] J. Fournet & A. Barrat, *PLOS ONE*, **9**, e107878, 2014.

TSP の枝候補作成に用いる巡回路誤差率の比較

土佐真義

東京理科大学 修士1年 池口研究室

1 はじめに

巡回セールスマン問題 (Traveling Salesman Problem, 以下 TSP) は組合せ最適化問題の1つである。TSP は NP 困難であり、最適解に近い解を求めるアルゴリズム (以下、近似解法) が必要とされている。

近似解法で TSP を解くにあたり、最適解に含まれていると推測される枝 (以下、枝候補) をあらかじめ用意し、枝候補に限定して探索を行うことで、短時間で解を見つけることができる。しかし、都市数が大きな問題では枝候補を定める時間も長くなる。このことは、近似解法だけでなく、枝候補を抽出するアルゴリズムも重要であることを意味する。

先行研究 [1] では、枝候補の抽出手法を提案している。高速な巡回路生成手法である Partial Optimization Metaheuristic Under Special Intensification Conditions (以下、従来手法) を用いていくつかの巡回路を用意する。そして、各巡回路に一度以上含まれる枝を枝候補とする。

これまでの研究で、従来手法の一部を変更した手法 (以下、提案手法) を含む枝候補抽出手法を提案した。提案手法は都市数の多い問題に対して枝候補の中の最適解に含まれている枝の本数を維持しつつ、従来手法よりも枝候補に含まれる枝の総本数を減少させた。しかし、どのような過程でこれらの枝候補が得られたのかは明らかになっていない。そこで、枝候補の抽出の際に用いる巡回路の誤差率について従来手法と提案手法の比較を行う。

2 従来手法

枝候補探索の従来手法 [1] のアルゴリズムを以下に示す。

1. 問題都市集合 V から n_a 個の都市をランダムに選択し、代表都市集合 V' とする。
2. 代表都市集合 V' に含まれる n_a 都市に対してランダムな巡回路 T_s を生成する。
3. Lin-Kernighan-heuristic[2] (以下、LK 法) を用いて T_s を最適化する。
4. 巡回路 $T = T_s$ とする。
5. T について、 $v \notin V'$ を $d_{u,v}$ ($u \in V'$) が最も小さくなる u の直後に挿入する。
6. 各都市 $v \in V'$ について、以下の作業を行う。
 - (a) T_s について、 v の次の次に訪れる都市を w とおく。
 - (b) 巡回路 T において v の直前に訪れる都市から始まり w で終わる部分巡回路を T' とする。
 - (c) 部分巡回路 T' の始点と終点を固定し、2-opt 法を用いて最適化する。
 - (d) 手順 (c) で最適化された部分巡回路 T' を用いて T を更新する。

3 提案手法

従来手法では、大規模な問題を小問題に分割することで高速な巡回路の生成を実現している。しかし、従来手法は代表都市の個数を多く設定すると計算時間がかかり、少なく設定すると余分な枝が多く含まれてしまう。そこで、提案手法は代表都市数を従来手法より増やしつつ、手順 3 の代表都市の最適化に LK 法の代わりに 2-opt 法を用いる。2-opt 法を用いることで、代表都市の増加に伴う計算時間の増加を抑えることがで

きる。これにより計算時間を維持しつつ余分な枝の本数を抑えることができる。

4 数値実験

それぞれの手法について枝候補を作成し、枝候補に含まれる枝の本数と最適解に含まれる枝の本数を調査した。具体的には、初期解を 20 個生成し、それらに含まれる全ての枝を枝候補とする。対象とした問題は TSPLIB[3] の問題 pr2392, d15112, pla33810, pla85900 である。結果を表 1 に示す。枝候補に含まれる枝の本数が少ないほど、近似解法の計算時間を短縮することができる。また、枝候補が最適解に含まれる枝を多く含んでいるほど、最適解に近い解を得られる可能性が上がる。表 1 から、提案手法は従来手法と比較して最適解に含まれる辺の数は維持されているものの、都市数の大きい問題について枝の総本数を削減できていることがわかる。このことから、都市数の多い問題に対して提案手法は最適解に含まれない余分な枝を適切に取り除いていることがわかる。

表 1 枝候補の性能の評価

問題	都市数	枝候補の総本数		最適解の枝本数	
		従来手法	提案手法	従来手法	提案手法
pr2392	2,392	7,831	9,096	2,388	2,385
d15112	15,112	59,038	63,450	15,038	15,087
pla33810	33,810	146,022	133,363	33,748	33,717
pla85900	85,900	446,016	346,183	85,807	85,770

次に、従来手法と提案手法において、枝候補の作成に用いた巡回路を単体で見た際の誤差率を比較した。誤差率は $\frac{(\text{巡回路長}) - (\text{最適解の巡回路長})}{(\text{最適解の巡回路長})}$ で定義される。巡回路の誤差率の平均値を表 2 に示す。全ての問題において、従来手法が提案手法より平均誤差率が低いことがわかる。そのため、枝候補の抽出において必ずしも誤差率の低い巡回路を作成する必要はないことがわかった。

表 2 各手法の巡回路誤差率の平均値 (単位: %)

問題	従来手法	提案手法
pr2392	21.7368	24.4435
d15112	21.0794	28.5610
pla33810	34.8559	46.8461
pla85900	43.3363	57.2635

5 まとめ

提案手法を用いた枝候補抽出手法は従来手法のそれと比較して、枝候補に含まれる枝の本数を削減しつつ最適解に含まれる枝を維持できる。本発表では、その原因を調査するために枝候補の作成に用いる巡回路を単体で確認した際の誤差率を算出した。その結果、巡回路の誤差率は枝候補の枝の本数の削減には直接的に影響していないことがわかった。

参考文献

- [1] Taillard Éric D, and Helsgaun Keld, *European Journal of Operational Research*, **272**, 2, 420–429, 2019.
- [2] Lin, Shen and Kernighan, Brian W, *Operations research*, **21**, 2, 498–516, 1973.
- [3] G. Reinelt, *ORSA Journal on Computing*, **3**, 4, 376–384, 1991.

単語の使用頻度が言語ネットワークの構造的特徴に与える影響

眞岸 祈平

東京理科大学 修士1年 池口研究室

1 はじめに

我々は、複雑ネットワークの観点から言語解析を行うことで、異なる言語で記述された文書から得られる言語ネットワークの構造的特徴が、各言語で異なることを示してきた [1, 2]. しかし、文書をネットワークに変換した際に、文書のどのような特徴がネットワークの構造的特徴に影響しているかは調査していない。

本発表では、文書の特徴の1つと考えられる単語の使用頻度とネットワーク構造の定量化指標との関係性について調査した。結果として、言語の種類に関わらず単語の使用頻度がネットワークの構造的特徴に影響していること、単語の使用頻度以外にもネットワークの構造的特徴に影響する文書の特徴が存在することが示唆された。

2 使用データ

本発表では、新約聖書のマタイによる福音書 [4]–[9] を解析対象とし、翻訳バージョンを含めた14種類のテキストデータを使用した。また、解析対象言語として、印・欧語族に分類される希語、羅語、英語、独語、仏語、露語、日本語族に分類される日本語、烏拉語族に分類される芬語、支那・西藏語族に分類される中国語を用いた。

3 言語ネットワーク作成手法

まず、形態素解析により文章を単語の原型(形態素)に分解し、その品詞を求める。ここで、形態素解析ツールとして日本語はmecab[10]を、他言語はtree-tagger[11]を使用した。次に、単語の原型とその品詞の組を1つの頂点とし、文章中で隣接する場合に頂点間を枝で結ぶことで、重み無し無向ネットワークを作成した。その際、単語の重複、単語間のつながりの重複は認めず、句読点、半句読点、記号を除いた。なお、1つの文章から非連結なネットワークが得られた場合は、最大部分グラフを用いた。

4 実験

本発表では、単語の使用頻度と4種類のネットワーク構造の定量化指標との関係性を調査した。具体的にはまず、各言語でのテキストデータの単語の使用頻度を算出する。次に、各言語のテキストデータから作成した14種類の言語ネットワークに対して、次数中心性 [12]、近接中心性 [13]、媒介中心性 [12]、クラスタ係数 [14] を算出する。最後に、単語の使用頻度と各指標の相関係数を算出することで、単語の使用頻度がネットワークの構造的特徴に影響を与えているのかを調査した。

5 結果

単語の使用頻度と次数中心性、近接中心性、媒介中心性、クラスタ係数の相関を図1に示す。なお、図1の横軸は言語の種類を、縦軸は相関係数を表している。

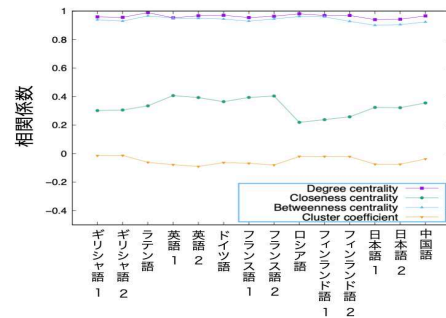


図1 各言語の単語の使用頻度と各指標の相関係数

図1を見ると、単語の使用頻度と次数中心性の相関が非常に高いことがわかる。この結果から、文書内で単語が使用される毎にネットワーク内で対応する頂点の次数が増加し、ネットワークの構造的特徴に影響を与えていることが考えられる。また、単語の使用頻度と媒介中心性も相関が非常に高いことがわかる。この結果から、文書内で多く使用される単語は、ネットワーク内の他頂点への媒介役として重要な役割を果たすという意味で、ネットワークの構造的特徴に影響を与えていることが考えられる。一方、近接中心性、クラスタ係数は相関が非常に低いことがわかる。この結果から、単語の使用頻度以外の文書の特徴が近接中心性、クラスタ係数の各指標に影響を与えていることが示唆される。

6 まとめ

本発表では、文書の特徴とネットワーク構造の定量化指標との関係性について調査した。結果として、言語の種類に関わらず、単語の使用頻度がネットワーク構造の定量化指標である次数中心性と媒介中心性に影響を与えていることが示唆された。一方、近接中心性とクラスタ係数の結果から、単語の使用頻度以外にもネットワーク構造に影響を与える文書の特徴が存在することも示唆された。

参考文献

- [1] 眞岸 祈平 他, 信学技報, **120**, 55, 27–32, 2020.
- [2] 眞岸 祈平 他, 信学ソ会, N-1-14, 2020.
- [3] D. J. Watts, et al., *Nature*, **393**, 440–442, 1998.
- [4] <https://www.bible.or.jp/> (2020-03-24 閲覧)
- [5] <https://www.academic-bible.com> (2020-04-13 閲覧)
- [6] <http://www.cerbaso.org/> (2020-04-16 閲覧)
- [7] <http://www.nic.funet.fi/> (2020-04-20 閲覧)
- [8] <https://www.biblegateway.com> (2020-04-26 閲覧)
- [9] <https://www.bebli.com/> (2020-04-28 閲覧)
- [10] <http://taku910.github.io/mecab/> (2020-09-16 閲覧)
- [11] <https://www.cis.uni-muenchen.de/> (2020-04-27 閲覧)
- [12] L. C. Freeman, *Social Networks*, **1**, 3, 215–239, 1978.
- [13] A. Bavelas, *ASA*, **22**, 6, 725–730, 1950.
- [14] D. J. Watts, et al., *Nature*, **393**, 440–442, 1998.

マスクの着用が生体信号に与える影響の調査

岡安高輝

東京理科大学 学部 4 年 池口研究室

1 はじめに

現在、COVID-19 の感染症対策として外出時にはマスクを着用することが推奨されている [1]. 本研究では、マスクの着用の有無が生体信号にどのような影響を与えるのかを明らかにすることを目的とする. そこで我々は心臓血管系の関連情報を持つ photoplethysmogram (以下 PPG) データ [2] を使用して調査する. また、マスクを着用せずに安静にした状態で測定された PPG データはカオスダイナミクスに起因して生じていると強く示唆されている [3]. 本稿では、マスクを着用する状態と着用しない状態での PPG データを実際に計測し、比較した結果を報告する.

2 PPG データ

PPG データは末梢血管の血流量の変化を表した時系列データである. PPG データを計測する様子は図 1 の通りで、指の一方から LED を投光し相対する受光素子により透過光を測定することによって非侵襲に得られる. 本実験では、マスクの着用の有無を対照実験の条件として、室温 26°C の静かな室内にてリラックスした状態で 1 回 3 分間のデータを 1 人あたり 3 回計測した. 計測機器のサンプリングレートは 409.6[Hz] である. 計測した PPG データは、マスクの着用なしが 2 人、マスクの着用ありが 4 人の計 18 回分である.



図 1 PPG データを計測する様子

3 アトラクタの再構成

アトラクタの再構成を行うことでデータの特徴を定性的に捉えることができる. 1 変数の時系列データから高次元空間のアトラクタを再構成するために、時間遅れ座標系への変換を行う. この変換は、1 変数の時系列データを $y(t)$ 、時間遅れの大きさを τ として m 次元の再構成状態空間において、 m 次元ベクトル式 (1) を構成する [4]. 次元は $m = 3$ として、 τ にはフーリエ変換して得られたスペクトラムの主要な周波数の周期を 4 で割った値を採用する [5].

$$v(t) = (y(t), y(t + \tau), \dots, y(t + (m - 1)\tau)) \quad (1)$$

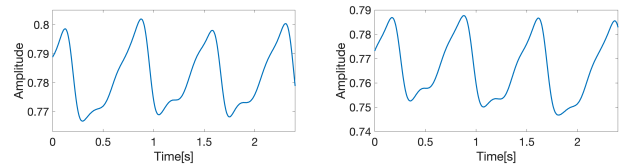
4 結果

計測した PPG データのうち、マスクの着用の有無が異なる PPG データの調査を行った結果の一例を示す. 時系列波形の一部を図 2 に示す. また PPG データをフーリエ変換して得られたスペクトラムを図 3 に示す. 基本統計量や図 3 より得られた τ は表 1 に示す.

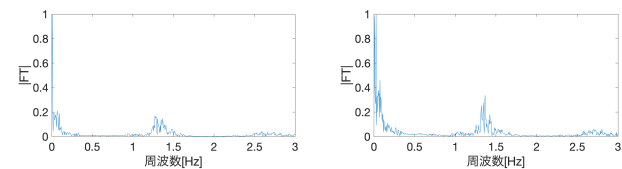
表 1 PPG データの基本統計量と τ を計算した結果

	データ数 [ステップ]	平均値	標準偏差	主要な周波数 [Hz]	τ [s]	τ [ステップ]
マスクなし	75,683	0.78233	0.02466	1.367	$\frac{1}{4 \times 1.272}$	81
マスクあり	77,009	0.77055	0.02461	1.272	$\frac{1}{4 \times 1.367}$	75

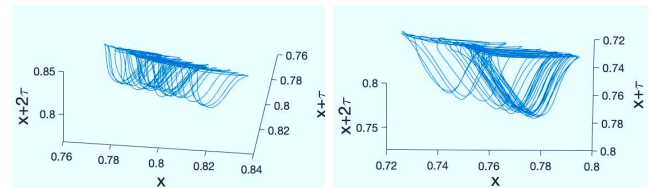
さらに PPG データを時間遅れ座標系へ変換してアトラクタの再構成を行った結果を図 4 に示す. ただし、見やすいようデータ数は 10,000 としている. 図 4 から、マスクの有無に関わらず生成されたアトラクタはストレンジアトラクタであることが確認できたが、マスクの着用の有無による違いは、定性的には捉えられなかった.



(a) マスクなし (b) マスクあり
図 2 時系列波形の例



(a) マスクなし (b) マスクあり
図 3 フーリエ変換で得られたスペクトラムの例



(a) マスクなし (b) マスクあり
図 4 再構成したアトラクタの例

5 まとめ

本稿では、リラックスした状態でマスクを着用する場合としない場合について PPG データを計測してアトラクタの再構成を行った. マスクの着用の有無いずれの場合もストレンジアトラクタを描けたが、定性的な違いはなかった. 今後は、他の条件下での PPG データの計測を行う. 加えて、リアプノフ指数、非線形決定論的予測、ポアンカレ写像、Wayland テスト、サロゲート法 [3] 等の非線形時系列解析手法を PPG データに適用して、マスクの着用の有無が生体信号にどのような影響を与えるのかを定量的に明らかにする.

参考文献

- [1] 新型コロナウイルス感染症対策専門家会議, “新型コロナウイルス感染症対策の状況分析・提言 (2020 年 5 月 4 日),” <https://www.mhlw.go.jp/content/10900000/000629000.pdf>, (Accessed on 09/28/2020).
- [2] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiological Measurement*, Vol. 28, No. 3, pp. R1–R39, feb 2007.
- [3] N. Sviridova *et al.*, “Human photoplethysmogram: new insight into chaotic characteristics,” *Chaos, Solitons & Fractals*, Vol. 77, pp. 53–63, 2015.
- [4] 池口 徹 他, “カオス時系列解析の基礎と応用,” 合原 一幸 編, 産業図書, 2000.
- [5] K. Sakai, “Nonlinear Dynamics and Chaos in Agricultural Systems,” *Developments in Agricultural Engineering. Elsevier Science*, 2001.

決定論的非線形力学系から生成するマーク付点過程と力学系の特徴保存についての調査

田所知也

東京理科大学 学部4年 池口研究室

1 はじめに

マーク付点過程を生成する手法として、観測時系列の極大値を抽出する手法がある。しかし、極大値抽出によって生成したマーク付点過程が元の時系列及び力学系の特徴をどの程度保存出来ているかについては詳細な調査が行われていない。そこで、本発表ではローレンツ方程式のパラメータを変化させた際に、極大値抽出により生成したマーク付点過程が元の力学系の特徴を保存出来ているかをリカレンスプロットの定量化指標 DET と L_{\max} を用いて評価した。

2 リカレンスプロット

2.1 定義

リカレンスプロット (RP) とは非線形時系列に対する解析手法の一つである。具体的には、状態空間内の2点間距離が閾値 θ より小さければ点をプロットし、そうでなければプロットしないという処理を行い、データの特徴を視覚化する方法である。 i 番目の状態ベクトルを \mathbf{X}_i とし、 i 番目、 j 番目の状態ベクトル間のユークリッド距離を $d(\mathbf{X}_i, \mathbf{X}_j)$ と表すと、RP の (i, j) 成分 $R(i, j)$ は式 (1) のように定義される [1]。

$$R(i, j) = \begin{cases} 1, & \text{if } d(\mathbf{X}_i, \mathbf{X}_j) < \theta \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

2.2 定量化指標

RP の点のパターンを定量化する指標として、プロットの斜め線を用いた DET と L_{\max} がある。本発表では、この2つの指標により力学系の特徴を評価する。まず、斜め線の長さが l の集合を $D(l)$ とすると、斜め線をなす点の割合 DET は $\text{DET} = \frac{\sum_{l \geq 2} |D(l)|}{\sum_{l \geq 1} |D(l)|}$ である。次に、RP の斜め線の最大長 L_{\max} は $L_{\max} = \max\{l | D(l) \neq \emptyset\}$ である。ここで、 \emptyset は空集合を表す [2]。

3 数値実験

ローレンツ方程式 [3] のパラメータ r を刻み幅 1 ずつ $0 \leq r \leq 400$ の範囲で変化させ、得られた解から RP を作成し、定量化指標 DET と L_{\max} を求めた。解を求める際にはルンゲ=クッタ法を用い、刻み幅は $\frac{8}{3}$ とした。パラメータ σ, b はそれぞれ $\sigma = 10, b = \frac{8}{3}$ 、用いたデータ数は $n = 5000$ とし、 $t = 1000$ までを過渡状態として除いて計算した。DET と L_{\max} は、ローレンツ方程式の各変数の時系列に対して、遅延座標系 [4] による再構成状態空間から RP を作成して求めた。その際、遅れ時間を $\tau = 10$ 、再構成次元を $m = 10$ とした。また、ローレンツ方程式の各変数の時系列から極大値を抽出して得たマーク付点過程のマーク時系列、イベント間隔時

系列、及びその両方に対して遅延座標系による再構成状態空間から RP を作成し、DET と L_{\max} を求めた。図 1 に各変数の時系列とマーク付点過程から得られた DET と $\frac{1}{L_{\max}}$ を示す。

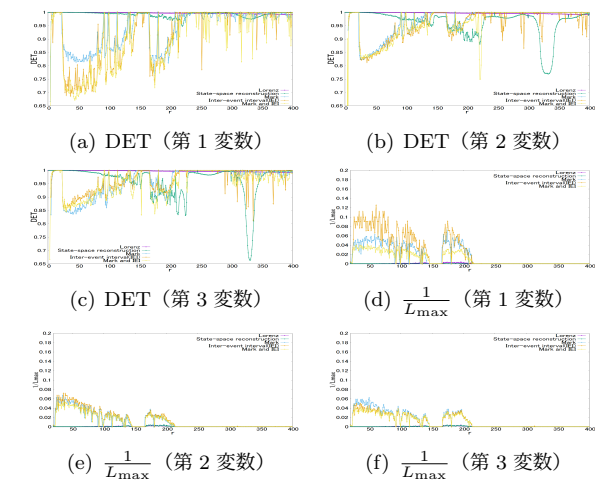


図 1: ローレンツ方程式のパラメータ r と DET, $\frac{1}{L_{\max}}$ の関係

ローレンツ方程式のパラメータ r については様々な調査がされており、例えば $24.06 < r < 313$ ではストレンジアトラクタとなることが知られている。実際、図 1(d), (e), (f) を見ると $25 \leq r \leq 215$ で $\frac{1}{L_{\max}}$ が大きくなっている。また、 $r = 150$ 付近及び $r > 214.4$ では周期的な振る舞いをする事が知られており、 $r = 150$ 付近及び $r > 215$ では $\frac{1}{L_{\max}}$ が小さくなっている。しかし、DET からは元の力学系の特徴を確認できなかった

4 まとめ

本発表では、観測時系列の極大値抽出により生成したマーク付点過程が元の力学系の特徴を保存できているかを調査した。具体的には、ローレンツ方程式のパラメータ r を変化させた際に RP の定量化指標である DET と L_{\max} が元の力学系に追従するかを評価した。結果として、定量化指標 L_{\max} では元の力学系の特徴が確認できた。先行研究により、ローレンツ方程式のパラメータ r の変化による力学系の変化が確認されており、今後は既知のパラメータ設定の周辺における調査をより細かい刻み幅で行いたい。

参考文献

- [1] J.-P. Eckmann et al., *EPL*, **4**, 9, 973–977, 1987.
- [2] 平田祥人, 数理解析研究所備講録, **1768**, 150–162, 2011.
- [3] E. N. Lorenz, *Journal of Atmospheric Sciences*, **20**, 130–141, 1963.
- [4] F. Takens, *Lecture Notes in Mathematics*, 366–381, 1981.

感染症数理モデルを用いた「感染者隔離」の定量的評価

真鍋歩未

東京理科大学 学部4年 池口研究室

1 はじめに

昨今における、COVID-19 の世界的な大流行は、感染力が強く、無症状者が存在することによる無意識的な感染拡大が一つの要因だと考えられている。ウイルスに対するワクチン開発に時間がかかる中、感染拡大を防止するために、海外渡航の制限や、感染者の隔離などの対策がとられている。しかし、その中でも感染者の隔離がどの程度感染拡大を防止できるのかは十分に調査されていない。そこで、各頂点の周囲の感染状況によって回復確率が定められる r-SIS モデル [1] に、隔離感染者を考慮したモデルを提案し、感染伝播の数値実験を行う。感染者の隔離の有無が最終感染率に与える影響を調査した。

2 r-SIAR モデル

SIR モデルなどの一般的な感染症伝播モデルは感染者数のダイナミクスを表すモデルであるが、本発表では、ネットワークの各頂点に状態遷移を表すダイナミクスを与えるモデルを使用した。具体的には、回復確率 μ が各頂点の周囲の感染状況によって定められる r-SIS モデル [1] を、回復者クラス及び隔離感染者クラスを加えるよう r-SIAR モデルへ拡張した。 $x_i(t)$ は時刻 t における状態値であり、 $x_i(t) = 0, 1, 2, 3$ でそれぞれ S クラス, I クラス, A クラス, R クラスであることを表す。各頂点の状態遷移のダイナミクスを式 (1), 状態遷移図を図 1 に示す。

$$x_i(t+1) = \begin{cases} x_i(t) + G(1 - (1 - \beta)^{|l_i(t)|})(G(1 - \delta) + 1) & (x_i(t) = 0) \\ x_i(t) + 2H(t - T_i - u) & (x_i(t) = 1) \\ x_i(t) + G(\mu_i(t)) & (x_i(t) = 2) \\ x_i(t) & (x_i(t) = 3) \end{cases} \quad (1)$$

ただし、 T_i は頂点 i がネットワークから隔離された時刻を表し、関数 $G(p)$ は確率 p で 1, 確率 $1-p$ で 0 をとり、関数 $H(x)$ は $x \geq 0$ のとき 1, $x < 0$ のとき 0 をとる階段関数とする。時刻 t における隣接行列の (i, j) 成分を $A_{i,j}(t)$ とすると、 $l_i(t) = \{j | x_j(t) = 0, A_{i,j}(t)x_j(t) = 2\}$ と定義する。未感染者 S クラスの頂点 i は感染確率 $1 - (1 - \beta)^{|l_i(t)|}$ によって感染者クラスへ遷移する。また、回復確率 $\mu_i(t)$ は、隣接頂点の感染状況、リソース配分戦略によって定められる [1]。確率 δ で I クラスに遷移すると、頂点 i はネットワークから枝を外す。さらに、時刻 u が経過すると、R クラスへ遷移しネットワークに復帰する。

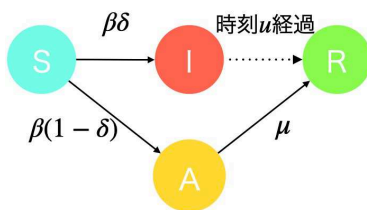


図 1 r-SIAR モデルの状態遷移図

3 数値実験

頂点数 $N = 5000$, 最小次数 $k_{\min} = 3$ であるべき指数 $\gamma = 2.4$ の次数分布に従う UCM(Unrelated Configuration Model)[2] を使用して、感染者隔離の有無と、感染確率 β , 最終感染率 ρ の関係を調査した。最終感染率 ρ は I クラス, A クラス, R クラスの頂点が全頂点に占める割合を示す。r-SIS モデル [1] と、r-SIAR モデルの隔離確率 $\delta = 0$ 及び 0.39 において感染確率 β と最終感染率 ρ を調査した結果を図 2 に示す。回復者クラスが存在することや、感染者の隔離を行うことで最終感染率 ρ を抑制できている。また、感染確率 β と最終感染率 ρ , 隔離確率 δ の関係を調査した結果を図 3 に示す。感染確率 $\beta \geq 0.3$ では感染者を 5 割以上隔離しないかぎり、最終感染率は 8 割を超えた。したがって、ネットワークの構造や、感染力の強さによっては感染者の隔離という策があまり効果的でないことが考えられる。

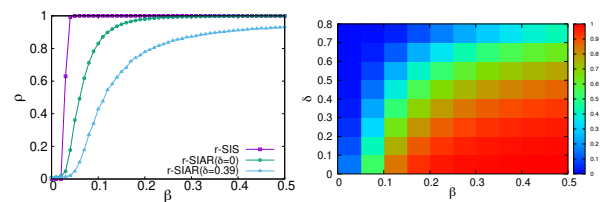


図 2 β と ρ の関係 図 3 β と δ, ρ の関係

4 まとめ

本発表では、確率的な感染症伝播モデルを用いて、感染者隔離のネットワーク上の感染拡大防止への有効性を定量的に評価した。具体的には、r-SIS モデル [1] に、隔離感染者クラス、回復者クラスを追加した r-SIAR モデルを提案し、感染確率、隔離確率と最終感染率の関係を調査した。結果として、感染力が高い感染症である場合は、感染者を 5 割以上隔離しなければ、感染症は 8 割ほど拡散してしまうことがわかった。よって、ネットワークの構造、感染力の強さにより感染者の隔離は効果的でないことを示唆した。今後は、感染者の隔離の有無とネットワークの構造的特徴の関連性の調査が必要である。また、COVID-19 により適応したモデル化のために、再感染の起きうる感染症モデルや、隣接頂点だけでなく濃厚接触者という概念を含んだ感染症モデルの構築も検討する。

参考文献

- [1] Chen *et al*, *Chaos: An Interdisced. J. of Nonlinear Sci.*, **28**(12), 123105, 2018.
- [2] Catanzaro *et al*, *PRE*, **71**(2), 027103, 2005.

複雑ネットワーク理論を用いた文章構造の時代的变化の解析

相原里江

東京理科大学 学部4年 池口研究室

1 はじめに

現在に至るまで、文章をネットワークに変換することで文章構造の特徴について解析する先行研究は多くなされてきた [1]. しかし、時代変遷に伴った文章構造や言語ネットワークの特徴の変化についての議論は不十分である. 本研究では、異なる時代の日本語の新聞記事をネットワークに変換することで、複雑ネットワーク理論の観点から時代変遷による文章構造の変化を調査した.

2 使用データ

本研究では解析対象として、神戸大学経済経営研究所の新聞記事文庫 [2] の異なる時代の新聞記事を使用した. 具体的には、各年代の文章が豊富にあること、書き手による文章の差が小さいこと、時代背景を反映していることなどの特徴を持つ東京朝日新聞を使用した.

3 ネットワークの作成手法

本研究では単語の共起に注目し、言語ネットワークを作成した. 具体的には、まず形態素解析ツールであるmecab[3]を用いて、文章を単語の原型に分解し、その品詞を得る. 次に得られた単語の原型と品詞を一つの頂点として扱い、文章中で隣接する場合に枝を付与し重み無しの無向ネットワークを作成した. その際単語の重複は認めず、句読点等の記号は削除した. また、一つの文章から複数のネットワークが得られた場合は、最大部分グラフのみを採用することとした.

4 実験手順

複雑ネットワークの観点から時代変遷に伴う文章構造の変化を調査するため、時代ごとに文章のネットワーク構造を比較した. 具体的には、新聞記事を言語ネットワークに変換し、得られたネットワークの特徴量である次数分布、平均頂点間距離 [4], クラスタ係数 [4] を計算した. その際、得られたネットワークをランダム化したネットワークのクラスタ係数と平均頂点間距離を用いて正規化処理を行った. 異なる時代の新聞記事から得られたネットワークの特徴量の比較をすることで、時代変遷に伴う文章構造の変化の傾向について調査した.

5 実験結果

1912年から1942年の新聞記事から得られた言語ネットワークの次数分布を計算した結果を図1に示す. また、1912年から1943年の新聞記事から得られた言語ネットワークの平均頂点間距離を図2、クラスタ係数を図3に示す. 図1から、徐々に次数分布が上に遷移しているように見えるが、1912年は114件、1922年は241件、1932年は109件、1942年は32件の新聞記事について次数分布をまとめたものであるため、記事数が次数分布に影響

している可能性があると考えられる. また、図2と図3から平均頂点間距離とクラスタ係数は年代間で差が存在することがわかるが、時代変遷に伴う値の変化傾向が見られない結果となった.

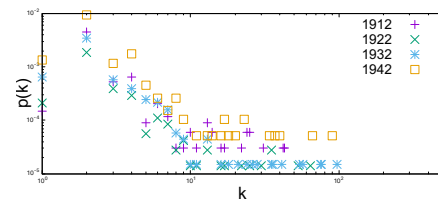


図1 次数分布

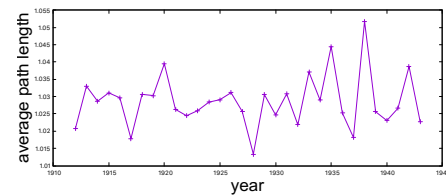


図2 平均頂点間距離

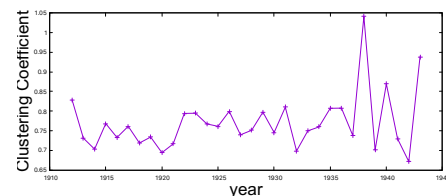


図3 クラスタ係数

6 まとめ

本研究では、時代変遷による文章構造の変化を調査するために、異なる時代の新聞記事をネットワークに変換し、特徴量を計算することで時代間の比較を行った. その結果、特徴量は年代間で差があったが、時代変遷に伴う文章構造の変化の傾向は見られなかった. 原因として、新聞の記事の文章量が全体的に少ないこと、記事ごとの文章量に差があることが考えられる. そのため、今後の課題として、特徴量を品詞ごとに計算し時代間の比較を行うこと、記事数を更に増やすことなどが挙げられる.

参考文献

- [1] Y. Shimada, *et al.*, EPL, 127, 5, 56003, 2019.
- [2] <http://www.lib.kobe-u.ac.jp/sinbun/> (2020-09-25 閲覧)
- [3] <https://taku910.github.io/mecab/> (2020-09-25 閲覧)
- [4] D. J. Watts, *et al.*, Nature, 393, 6684, 440–442, 1998.

ブートストラップリサンプリング法による時系列データの予測

大矢早紀

東京理科大学 学部4年 池口研究室

1 はじめに

現実世界には株価のデータや、気温のデータなどの様々な時系列データが存在している。これらの時系列データの大半は非線形な動きをしており、数値の予測を行うことが難しい。このような非線形時系列に対して、予測を行う手法がいくつか提案されている。原木らは先行研究 [1] で数値予測手法としてブートストラップリサンプリング法 (以下、ブートストラップ法) を提案し、池田写像に対して数値予測を行った結果を報告している。ブートストラップ法は従来の手法よりも、ノイズのある時系列に対して精度の高い数値予測を行うことができる。本発表では、ブートストラップ法の予測精度を評価するため、エノン写像を用いて従来の予測法との比較を行う。具体的には、予測ステップ数を変化させた場合の実データと予測データとの誤差を導出することによって、従来の予測手法とブートストラップ法の予測精度の比較を行う。

2 従来の予測手法

本発表では、従来の予測手法として、ローレンツの類推法 [2]、加重平均予測法 [3]、負の指数を用いた加重平均予測法、ヤコビ行列推定法 [4] の4つの予測手法を用いる。

3 ブートストラップ法

アトラクタ上の一点を $v(t)$ 、予測する点を $\hat{v}(t+1)$ とする。そのとき、ブートストラップ法によるヤコビ行列の推定の手順は以下のように示せる。

1. 学習データにおけるヤコビ行列 $G(t)$ を推定しておく。
2. $v(t)$ の近傍点を M 個探索する。ここでは $M = 5$ とする。
3. 探索した近傍点の情報を B 回リサンプリングする。すなわち、どの近傍点を用いてヤコビ行列を推定するかをランダムに選びなおし、 B 個のヤコビ行列を求める。なお、ここでは $B = 200$ とする。
4. 求めた B 個のヤコビ行列を推定し、平均をとる。

4 使用データ

カオス的な振る舞いを行う時系列データとして、エノン写像を使用した。エノン写像を式 (1)、式 (2) に示す。

$$x(t+1) = 1 - ax^2(t) + y(t) \quad (1)$$

$$y(t+1) = bx(t) \quad (2)$$

本発表ではノイズなしの時系列データとしてパラメータ a, b を $a = 1.4, b = 0.3$ としたエノン写像を用いた。また、ノイズ付きの時系列データとして、パラメータを $a = 1.25, b = 0.3$ とし、変数 x, y に $[-0.1, 0.1]$ のランダムノイズを付与したエノン写像を用いた。

5 数値実験

各予測手法を用いてエノン写像の数値予測を行い、予測ステップ数 s を変化させた場合の誤差を比較した。誤差は二乗平均平方根誤差 E を用いた。式 (3) に示す。

$$E = \frac{\sqrt{\langle (x(t) - \hat{x}(t))^2 \rangle}}{\sigma} \quad (3)$$

ここで、 $x(t), \hat{x}(t)$ はそれぞれ時系列データの実データと予測データであり、 σ は実データの標準偏差である。ノイズなしとノイズ付きの場合のエノン写像の数値予測を行った結果を図 1 に示す。

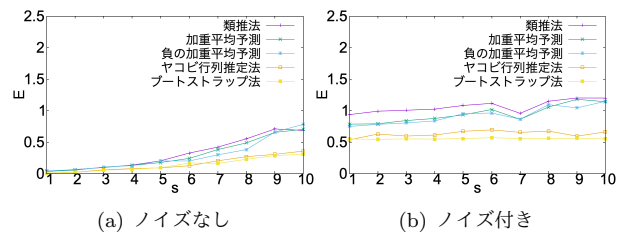


図 1: 予測ステップを変化させたときのエノン写像の予測誤差 E

類推法、加重平均予測、負の加重平均予測は、ヤコビ行列推定法とブートストラップ法より誤差が大きいのが分かった。図 1(a) において、予測ステップ数 s が大きくなるにつれて誤差 E の値が大きくなるのが分かった。それに対して、図 1(b) では誤差 E にあまり変化が見られなかった。これらの結果から、ブートストラップ法と従来のヤコビ行列推定法の精度の差にあまり大きな違いがないことが分かった。ヤコビ行列の推定は近傍点数 M とリサンプリング数 B の値が影響するため、誤差を小さくする条件を満たす適切な値を設定できていないと考えられる。

6 まとめ

本発表では、ブートストラップ法の有用性を示すため、エノン写像を用いて従来の予測手法との比較を行った。その結果、ブートストラップ法によるヤコビ行列推定の精度の高さを明確に確認することができなかった。今後の課題として、誤差を小さくするため、近傍点数 M や、リサンプリング回数 B を変化させたときの誤差を調査し、ヤコビ行列を正確に推定することが必要である。

参考文献

- [1] D. Haraki et al., *Phys.*, **75**, 5, 056212, 2007.
- [2] E. N. Lorenz, *J. Atmos. Sci.*, **26**, 4, 636–646, 1967.
- [3] T. Ikeguchi and K. Aihara, *Phys.*, **55**, 3, 2530, 1997.
- [4] Eckmann, et al., *Phys.*, **34**, 6, 4971, 1986.

スペクトルグラフ距離を用いた WS モデルの構造の解析

澤田幸輝

東京理科大学 学部4年 池口研究室

1 はじめに

複雑ネットワークとは、現実中存在するつながりや現象を頂点と枝で表したものである。複雑ネットワークを解析することにより、対象の特性を明らかにできる可能性がある。ネットワークを解析する指標は、ネットワーク内の狭さを定量化するクラスター係数、ネットワーク内のつながりの近さを定量化する平均頂点間距離などがある。しかし、これらの指標はネットワーク構造の差異を定量化することはできていない。

これに対してスペクトルグラフ距離 [1] は、2つのネットワーク間の距離を求めることができる。ネットワークの特性を反映したラプラシアン行列を利用することで、ネットワーク構造の差異を定量化することができる。本発表では、スペクトルグラフ距離を用いて、WS モデル [2] 同士の距離を計測した。

2 スペクトルグラフ距離

スペクトルグラフ距離はラプラシアン行列から得られる固有ベクトルを利用する。ここで、ラプラシアン行列 \mathbf{L} は、 \mathbf{A} を $n \times n$ の無向ネットワークの隣接行列、 \mathbf{D} を対角行列で $\text{diag}(k_1, \dots, k_n)$ として、 $\mathbf{L} = \mathbf{D} - \mathbf{A}$ と定義される。2つのネットワーク $G^{(i)}$, $G^{(j)}$ の距離を計算するスペクトルグラフ距離の式は以下で定義される。

$$d(G^{(i)}, G^{(j)}) \equiv \frac{1}{m_{ij} - 1} \sum_{r=2}^{m_{ij}} d'(\rho_r^{(i)}, \rho_r^{(j)})$$

2つのネットワークのラプラシアン行列の固有ベクトルの要素の分布 $\rho_r^{(i)}, \rho_r^{(j)}$ 間の距離 $d'(\rho_r^{(i)}, \rho_r^{(j)})$ の計算し、累積していくことで定義されている。 m_{ij} は $1 < m_{ij} \leq \min(n^{(i)}, n^{(j)})$ とし、固有ベクトルの個数が違う場合でも比較することができる。

3 WS モデル

WS モデル [2] は、1998 年に発表された論文でスモールワールド現象を説明する際に提案されたネットワークである。WS モデルは、以下の手順で作成することができる。

1. n 個の頂点を円環上に並び、各頂点から隣接する k 個の頂点で接続し、格子グラフの作成をする。
2. それぞれの枝について繋ぎ変え確率 p で枝の片方の頂点を切り離し、別の頂点に接続する。

このモデルは、 $p = 0$ のとき円環上の格子ネットワークとなり、 $p = 1$ のときランダムネットワークとなる。また、 $p = 0.01$ 付近のとき、スモールワールドネットワークとなることが知られている。

4 数値実験

スペクトルグラフ距離がネットワークの構造の差異を反映できているかを確認を行うために、WS モデルを用いて数値実験を行った。以下に数値実験の手順を示す。

1. 繋ぎ変え確率 p_0 の WS モデル G_{p_0} を作成する。

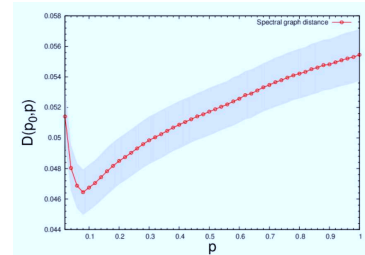
2. 繋ぎ変え確率 p の WS モデルを N 個作成し、それぞれを $G_p^{(1)}, G_p^{(2)}, \dots, G_p^{(N)}$ とおく。

3. $D(p_0, p) = \frac{1}{N} \sum_{i=1}^N d(G_{p_0}, G_p^{(i)})$ を計算する。

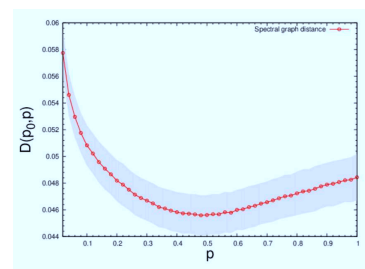
4. 手順 2, 3 を $p = 0.02$ から $p = 1$ まで 0.02 刻みで行う。

ただし、WS モデルの頂点数は 500、次数は 10 とした。また、 $N = 5000$ とした。

$p_0 = 0.1, 0.4$ で計算した結果を図 1 に示す。この図より、どちらの p_0 においても $p_0 = p$ 付近で最も $D(p_0, p)$ が小さくなっていることが確認できる。加えて、WS モデルは p が大きくなるにつれてランダムネットワークに近づくため、ネットワークとしての特徴が少なくなる。そのため、繋ぎ変え確率の大きい WS モデル同士では、繋ぎ変え確率が異なっても類似したネットワークとなる。図 1(b) は図 1(a) と比較して広範囲で $D(p_0, p)$ が小さくなっていることがわかる。このことからスペクトルグラフ距離は、WS モデル間の構造の差異を定量化することができると考えられる。



(a) $p_0 = 0.1$



(b) $p_0 = 0.4$

図 1 数値計算を行った結果

5 まとめ

本発表では、WS モデルを用いてスペクトルグラフ距離がネットワーク構造の差異を定量化できるかについて調査を行った。数値実験により、スペクトルグラフ距離は WS モデルのネットワーク構造の差異を定量化できることがわかった。

参考文献

- [1] Y. Shimada, *et al.*, *Sci. Rep.*, **6**, 34944, 2016.
- [2] D. J. Watts, *et al.*, *Nature*, **393**, 440–442, 1998.

ニューロンへの入力時系列の再構成

三浦英

東京理科大学 学部4年 池口研究室

1 はじめに

人間の脳内には、ニューロンと呼ばれる神経細胞が数多く存在している。ニューロンは、ある閾値を超える入力を与えられると、スパイク列を出力することで、断続的に情報を伝達する。一般的に、ニューロンの出力信号を観測することは容易であるが、ニューロンへの入力時系列を観測することは困難である。先行研究 [1] では、複数のニューロンから出力されるスパイク列から得られる発火間隔時系列を用いて、ノイズ量が大きい場合でもニューロンへの入力時系列を再構成できることを示している。なお、ニューロンモデルとして、Leaky Integrate-and-Fire モデル (以下、LIF モデル)[2] が使用されており、ニューロン数を 10 個と設定している。本発表では、ニューロン数を変更した場合に、再構成時系列の精度が、ノイズの大きさによって変化するかを検証した。

2 ニューロンへの入力時系列の再構成手法 [1]

まず、個体差のある複数のニューロンに対し、レスラー方程式の第 1 変数 $x(t)$ を入力することで得られる発火間隔時系列から複数のリカレンスプロット [5, 6] (以下、RP) を作成する。ここで、ニューロンの個体差はノイズを用いて表現した。次に、複数のニューロンへの共通入力の情報を取り出すために、複数の RP から和集合を取ることで、重畳 RP を作成する。最後に、重畳 RP から重み付きネットワークを作成し、ダイクストラ法を用いることで、距離行列を求め、多次元尺度法を適用することで、入力時系列を再構成する。[7, 8, 9]

3 実験

本発表では、ニューロン数を変更した場合に、ノイズ量の大きさによって再構成された入力時系列の精度が変化するかを調査するために、入力時系列と再構成時系列の相関係数を算出することで定量的に評価した。具体的には、ノイズ量を正規分布に従う平均 0 のガウスノイズを用いて、SN 比 p で定義し、 $p = -15, 0, 15, 20$ 、ニューロン数 $n = 1, 10, 50, 100$ と変更した場合の相関係数を求めた。なお、入力時系列と再構成時系列ではデータ数が異なるため、相関係数の計算には時系列の極値を用いた。

4 結果

入力時系列と再構成時系列を重ね書きした結果を図 1 に、ノイズ量を変更した場合の各ニューロン数における相関係数を図 2 に示す。ここで、時系列の振幅を合わせるため、入力時系列と再構成時系列を標準化した。

図 1, 2 を見ると、重ね書きの概形が類似したこと、相関係数が高いことから、ニューロン数を変更した場合で

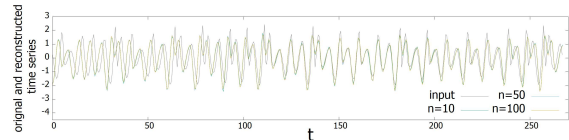


図 1 入力時系列と再構成時系列の重ね書き

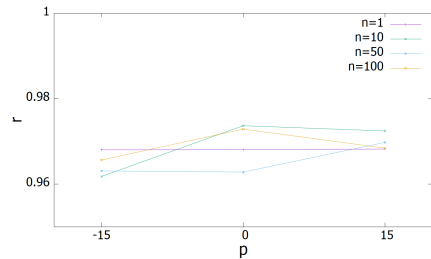


図 2 ノイズ量を変えた場合の相関係数

も入力時系列を再構成できていることがわかる。また、ノイズの SN 比 $p = 20$ とした場合では、ニューロン数 $n = 100$ と大きな場合でも再構成できなかった。これらの結果から、ニューロン数を増やした場合でも、再構成できるノイズ量の大きさは変化しないことがわかる。また、ニューロン数の違いによる、入力時系列と再構成時系列の相関係数に比例関係があるかについては傾向が見られなかった。

5 まとめ

本発表では、先行研究 [1] の手法を用いて、ニューロン数を変更した場合に、再構成時系列の精度がノイズの大きさによって、変化するかを検証した。結果として、ニューロン数を変更した場合でも、入力時系列を再構成できるノイズ量の大きさに変化がないことがわかった。

参考文献

- [1] 唐澤恵理花, 東京理科大学 経営工学科 卒業論文, 2020.
- [2] R. B. Stein, *Biophys. J.*, **7**, 1, 37–68, 1967.
- [3] E. M. Izhikevich, *IEEE Trans NN*, **14**, 6, 1569–1572, 2003.
- [4] O. E. Rössler, *Phys. Lett.*, **57A**, 5, 397–398, 1976.
- [5] J. -P. Eckmann, *et al.*, *EPL*, **4**, 9, 973–977, 1987.
- [6] 藤原寛太郎 他, *神経回路学会誌*, **21**, 2, 79–86, 2014.
- [7] Y. Hirata, *et al.*, *Eurphys J Spec Top*, **164**, 1, 13–22, 2008.
- [8] 野村亮太 他, *信学総大* 2019, N-1-24, 2019.
- [9] 野村亮太 他, *信学技報*, **119**, 19, 29–34, 2019.

数理モデルを用いたニューロン比率不均衡による脳機能障害の原因解明

宮南風†

島田裕††

藤原寛太郎†††

池口徹††††

東京理科大学 大学院工学研究科 経営工学専攻†
 東京大学国際高等研究所 ニューロインテリジェンス国際研究機構†††

埼玉大学 大学院理工学研究科 数理電子情報部門††
 東京理科大学 工学部 情報工学科††††

1 はじめに

ヒトの大脳皮質において、興奮性ニューロンと抑制性ニューロンは約 4:1 の比率 (以下、生理学的比率) で存在するといわれている [1]。一方、ダウン症のラットを用いた実験では NMDA 受容体の活性化の低下や GABA_A 拮抗薬による長期増強の回復が確認されており [2, 3]、生理学的比率に従っていないことも示唆されている。さらに、fMRI を用いたヒトを対象とした実験では、ダウン症患者と健常者にて、計測する領域間の物理的距離に応じた fMRI 信号同士の相関関係が異なっており、ダウン症における脳機能障害はニューロン比率不均衡が原因とも考えられる [4]。

本研究では、ダウン症患者の脳機能障害の原因解明を目的として、ニューロンの膜電位数理モデルを結合させ、両ニューロンの存在比率を意図的に変化させたネットワーク上でのニューロンの活動を調査する。さらに、ニューロンモデルより生成された活動時系列を fMRI 信号に変換する数理モデルを用いてダウン症患者の fMRI 解析実験 [4] の再現を行う。

2 Izhikevich ニューロンモデル [5]

式 (1) に示すニューロンの膜電位数理モデルを用いる。

$$\begin{cases} \dot{v}_i(t) = 0.04v_i(t)^2 + 5v_i(t) + 140 - u_i(t) + I_i(t) \\ \dot{u}_i(t) = a(bv_i(t) - u_i(t)) \end{cases} \quad (1)$$

if $v_i(t) \geq 30[\text{mV}]$, $v_i(t) \leftarrow c$, $u_i(t) \leftarrow u_i(t) + d$

t は時刻, $v_i(t)$ は時刻 t におけるニューロン i の膜電位, $u_i(t)$ は時刻 t におけるニューロン i の膜電位の回復変数, $I_i(t)$ は時刻 t におけるニューロン i への入力電流である。 a, b, c, d はパラメータであり、興奮性ニューロンは Regular Spiking ($a = 0.02, b = 0.2, c = -65, d = 8$), 抑制性ニューロンは Fast Spiking ($a = 0.1, b = 0.25, c = -65, d = 2$) を用いた。このとき、 $I_i(t) = I_{\text{ex}} + \sum_{j \in G_i} w_{ij} H(v_j(t) - \delta_{ij}) - 30$ とする。 I_{ex} は

外部入力電流, G_i はニューロン i にシナプス結合するニューロンの集合, δ_{ij} はニューロン j からニューロン i への軸索伝達遅延である。また、 $H(x)$ はステップ関数である。

3 Balloon モデル [6, 7]

式 (2) に示す Balloon モデルは、神経活動を入力として fMRI 信号時系列を再現する数理モデルである。

$$\begin{cases} \dot{s}_i(t) = \epsilon u_i(t) - \frac{s_i(t)}{\tau_s} - \frac{f_i(t) - 1}{\tau_f} \\ \dot{f}_i(t) = s_i(t) \\ \dot{z}_i(t) = \frac{1}{\tau_0} (f_i(t) - z_i^\alpha) \\ \tau_0 \dot{q}_i(t) = f_i(t) \frac{1 - (1 - E_0) \frac{f_i(t)}{E_0}}{E_0} - q_i(t) z_i(t) \frac{1}{\alpha} - 1 \\ y_i(t) = V_0 (k_1 (1 - q_i(t)) + k_2 (1 - \frac{q_i(t)}{z_i(t)}) + k_3 (1 - z_i(t))) \end{cases} \quad (2)$$

$s_i(t)$ は血管拡張性の信号, $f_i(t)$ は血流の流入量, $z_i(t)$ は血流の体積, $q_i(t)$ はデオキシヘモグロビン含有量, $y_i(t)$ は BOLD 反応 (fMRI 信号) 時系列, $u_i(t)$ はニューロンの活動信号である。また、 $\epsilon = 0.001, \tau_s = 20, \tau_f = 20, \tau_0 = 25, \alpha = 0.33, E_0 = 0.34, V_0 = 1, k_1 = 7E_0, k_2 = 2, k_3 = 2E_0 - 0.2$ はそれぞれパラメータである。本発表では、ニューロンの活動信号として $I_i(t)$ を用いる ($u_i(t) = I_i(t)$)。

4 数値実験

全ニューロン数を 1000 個、抑制性ニューロン比率 r に設定したニューラルネットワーク [8] を用意する。 r はネットワーク上の抑制性ニューロン数を全ニューロン数で除したものと定義する。このとき、興奮性シナプスの学習には STDP 学習則 [9] を用い、抑制性シナプスは固定値とした [10]。

初期状態からの学習が十分に収束した時間 ($t = 5,400 \sim 7,200[\text{s}]$) にて、 $y_i(t)$ を算出する。得られた時系列を興奮性 \rightarrow

興奮性シナプスを挟むニューロン間で相関係数を計算した。それらの値が 0.2 以上のシナプスを抽出し、そのシナプスの有する伝達遅延値 δ_{ij} の頻度分布を比較した。 δ_{ij} が小さい場合、信号の計測領域間の物理的距離は近い、大きい場合は遠いとみなす。このように対応づけることで、fMRI 解析実験の再現を試みる。

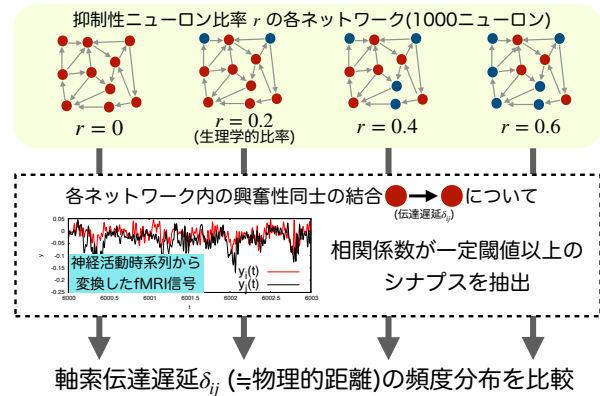
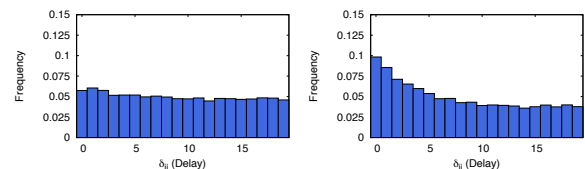


図 1: 数値実験の概要

5 実験結果

fMRI 信号時系列 $y_i(t)$ の相関を有するニューロン間の軸索伝達遅延値 δ_{ij} の頻度分布を図 2 に示す。生理学的比率である $r = 0.2$ では δ_{ij} によらず同頻度なのに対し、非生理学的比率である $r = 0$ のとき、 δ_{ij} が小さい頻度が高く、大きい頻度が低い。fMRI 信号を解析した実験より、生理学的比率に従わないダウン症患者の脳内では、物理的距離が短距離の領域間の信号間の相関は高く長距離の領域間の相関は低いことがわかっている [4]。以上より、fMRI 信号が得られる領域の組と数理モデル上のニューロン間のシナプス是一对一には対応しないものの、これらの知見と一致していることがわかる。



(a) $r = 0.2$ (生理学的比率) (b) $r = 0$ (非生理学的比率)

図 2: $y_i(t)$ の相関を有するシナプスの δ_{ij} の分布 (抜粋)

6 まとめ

本発表では、ニューロンの数理モデルより生成された神経活動から fMRI 信号に変換し、信号間で相関があるシナプスの遅延値分布をニューロン比率ごとに比較した。その結果、ダウン症患者の fMRI 解析実験から得られた知見を再現することに成功した。ニューロン比率の不均衡は、ネットワーク全体の情報過多を招き、領域間が短距離の伝達を強め、長距離の伝達を弱めることが示唆される。

参考文献

- [1] De Felipe J., et al., *Prog. Neurobiol.*, **6**(39), 563–607, 1992.
- [2] P.V. Belichenko et al., *J. Comp. Neurol.*, **480**(3), 281–298, 2004.
- [3] J.E. Hanson et al., *J. Physiol.*, **579**(1), 53–67, 2007.
- [4] Jeffrey S. Anderson et al., *NeuroImage Clin.*, **2**, 703–715, 2013.
- [5] Izhikevich E. M., *IEEE Trans Neural Netw.*, **6**(14), 1569–1572, 2003.
- [6] Buxton. R. et al., *Magn Reson Med* **39**(6), 855–864, 1998.
- [7] Friston, K. J. et al., *NeuroImage*, **12**(4), 466–477, 2000.
- [8] 宮南風 他, 信学ソ大, N-1-10, 2020.
- [9] Bi G. Q. and Poo M. M., *J. Neurosci.*, **18**(8), 10464–10472, 1998.
- [10] Izhikevich E. M. *Neural Computation*, **2**(18), 245–282, 2006.

発火間隔時系列を用いたニューロンへの共通入力への再構成

三浦英¹池口徹²東京理科大学 学部 4 年 池口研究室¹東京理科大学 工学部 情報工学科²

1 はじめに

人間の脳内には、ニューロンと呼ばれる神経細胞が数多く存在している。ニューロンは、ある閾値を超える入力を与えられると、スパイクを出力することで、断続的に情報を伝達する。一般的に、ニューロンの出力信号を観測することに比べて、ニューロンへの入力信号を観測することは困難である。本稿では、出力されたスパイク列から得られる発火間隔時系列を用いることで、共通入力を再構成する手法を提案する。ここで、ニューロンの数理モデルとして、Leaky Integrated and Fire モデル(以下、LIF モデル)を用いる。本稿では、複数のニューロンの個体差として、LIF モデルの漏れのパラメータ τ を用いて、共通入力の再構成を行った結果を報告する。

2 Leaky Integrated and Fire モデル [1]

LIF モデルは、式 (1) で定義される。

$$\tau \frac{dV(t)}{dt} = -V(t) + S(t) \quad (1)$$

ここで、 $V(t)$ は膜電位、 $S(t)$ は入力電流を表し、膜電位 $V(t)$ はある閾値 Θ に達したとき、 $V(t) \leftarrow 0$ とリセットされる。また、 τ は漏れの割合を示すパラメータであり、 τ が小さくなると、漏れの割合が大きくなる。本稿では、閾値 $\Theta = 20$ とした。

3 ニューロンへの共通入力の再構成

まず、共通入力としたレスラー方程式の第 1 変数 $x(t)$ の時系列にバイアス項を加えた時系列を作成し、LIF モデルの漏れ τ を個体差とした n 個のニューロンへ入力し、 n 本の発火間隔時系列を得る。次に、時間遅れ座標系への変換を用いて、 n 本の得られた発火間隔時系列から n 個のアトラクタの再構成 [3] を行い、 n 個のリカレンスプロット [4, 5] (以下、RP) を作成する。さらに、共通入力の情報を取り出すために、 n 個の RP から和集合を取ることで、重畳 RP を作成する。最後に、論文 [6] の手法を用いて、重畳 RP から共通入力を再構成する [7, 8]。

4 実験

LIF モデルの漏れ τ を変化させると、発火数が変化し、RP の重畳が困難になる場合があるため、発火数が一定となる条件にする必要がある。そこで本稿では、まず、各漏れ τ に対し、発火数が一定となるようにバイアス項を調整した場合に共通入力の再構成を行った。次に、発火数が変化しない範囲で、LIF モデルの漏れ τ を微小に変化させ、共通入力の再構成を行った。なお、共通入力の再構成ができていないかを確認するため、共通入力と再構成時系列の 2 つの時系列の標準化を行い、時系列から極値を抽出し、その相関係数を算出した。

5 結果

まず、漏れ $1/\tau = 0.0, 0.1, \dots, 2.4$ に対し、発火数が 500 で一定となるようにバイアス項を調整した場合に共通入力の再構成を行った結果を図 1 に示す。ここでは、ニューロン数を 25 とした。このとき、相関係数は 0.98 となった。

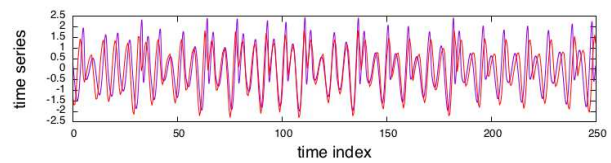


図 1 漏れ τ ごとにバイアス項を調整した場合の 2 つの時系列波形

次に、発火数が変化しない範囲 ($0.996 \leq 1/\tau \leq 1.00188$) で、LIF モデルの漏れ τ を微小に変化させ、共通入力の再構成を行った結果を図 2 に示す。ここでは、ニューロン数を 50、発火数が 500 となるようにバイアス項を 50.35 とした。このとき、相関係数は 0.98 となった。

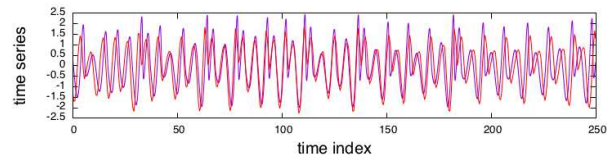


図 2 漏れ τ を微小に変化させた場合の 2 つの時系列波形

図 1, 2 を見ると、2 つの時系列波形の概形が類似していること、相関係数が高いことがわかる。この結果は、発火数が一定となるような条件では、共通入力の再構成を行えることを示すものである。

6 まとめ

本稿では、発火間隔時系列を用いてニューロンへの共通入力の再構成を行った。その結果、複数のニューロンの個体差として、LIF モデルの漏れのパラメータ τ を用いた場合でも、発火数が一定となるような条件で共通入力の再構成を行えることが示唆された。

参考文献

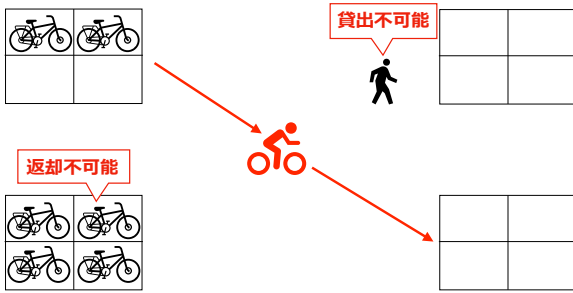
- [1] R. B. Stein, *Biophys. J.*, **7**, 1, 37–68, 1967.
- [2] O. E. Rössler, *Phys. Lett A.*, **57**, 5, 397–398, 1976.
- [3] T. Sauer, *Phys. Rev. Lett.*, **72**, 24, 3811–3814, 1994.
- [4] J. -P. Eckmann, *et al.*, *EPL*, **4**, 9, 973–977, 1987.
- [5] 藤原, 他, *神経回路学会誌*, **21**, 2, 79–86, 2014.
- [6] Y. Hirata, *et al.*, *Eurphys J Spec Top*, **164**, 1, 13–22, 2008.
- [7] 野村, 他, *信学総大*, N-1-24, 2019.
- [8] 野村, 他, *信学技報*, **119**, 19, 29–34, 2019.

非線形ワークショップ

自転車再配置問題の中規模サイズ実データに対するタブーサーチ法を用いた解法の性能調査

東京理科大学 池口研究室 D1 對馬帆南

1. バイクシェアリングシステム

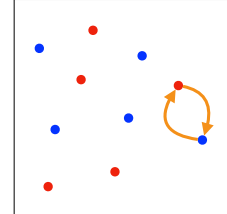


特徴	自由に乗り降りが可能
問題点	貸出・返却不可能なポートが発生
解決策	自転車回収車を用いて再配置作業が必要

2. 今回の発表内容

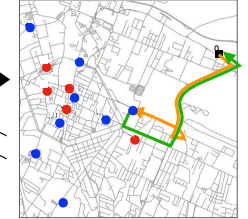
2.1 実データを使用 [2]

ランダムに駐輪ポートを配置



2点間のコストは等しい

Reggio Emiliaのデータ [1]

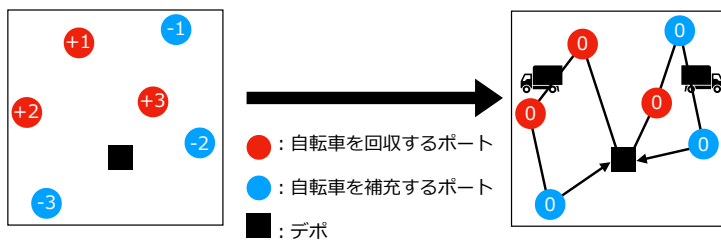


2点間のコストは等しくない

2.2 Or-optを拡張

挿入できるポート数 K を3に設定していたが、3よりも大きい値に設定する

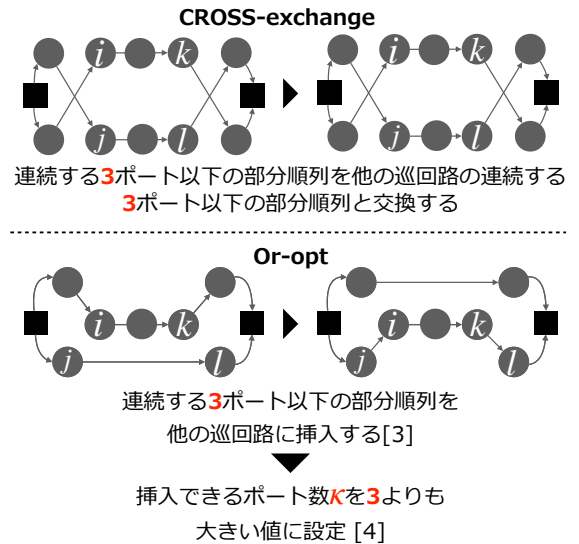
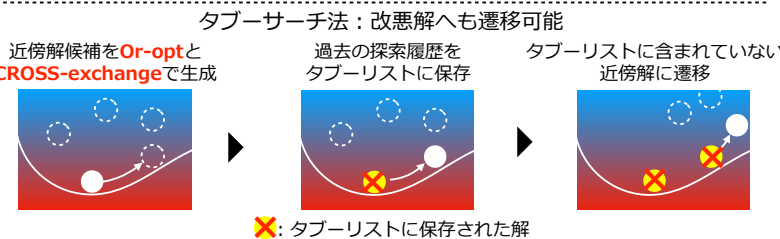
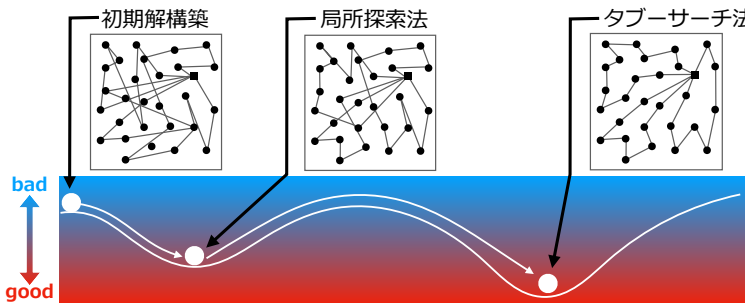
3. 複数の自転車回収車を用いた自転車再配置問題 [1]



回収車の移動時間を最小化

- 条件
- 再配置作業中に自転車台数は変化しない
 - 全ての回収車は必ずデポから出発し、再びデポに戻る
 - 回収車は自転車を積んでデポを出発することが可能
 - 1度の訪問で自転車台数を調整する
 - 回収車の積載容量を超えて自転車を積み込めない
 - 回収車は制限時間内に再びデポに戻ってくる

4. 複数の自転車回収車を用いた自転車再配置問題に対するタブーサーチ法を用いた解法



5. 数値実験・結果

表1: データ情報と実験条件

ポート数(Ciudad de Mexico)	90
各ポートの過不足台数	-17 ~ +17
回収車台数	20
回収車の最大積載容量	17
試行回数	10
タブー期間	250
タブーサーチの繰り返し回数	1,000
K	3, 5, 10, 15, 20, 25

表2: 実験結果

K	誤差率 [%]		計算時間 [s]
	平均	最良	
3	11.28	9.33	84.83
5	10.64	7.39	92.92
10	8.96	7.73	105.28
15	8.84	8.15	120.00
20	8.10	6.67	134.07
25	7.97	5.78	138.77

K の増加に伴い短い巡回路を得られた

参考文献

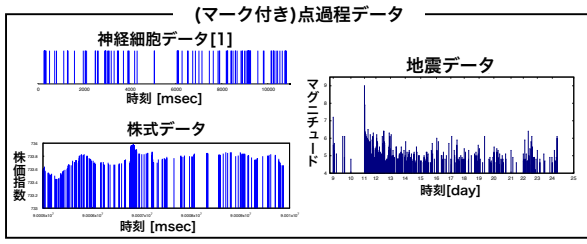
[1] 對馬帆南, 松浦隆文, Proc.of IEICE NOLTA ソサエティー 2017, NLS-16, 2017.
 [2] D. Mauro, et al., Omega, 45, pp. 7-19, 2014.
 [3] I. Or, PhD thesis, Northwestern University, Evanston, 1976.
 [4] G. Babin, et al., Journal of the Operational Research Society, 58(3), pp. 402-407, 2007.

点過程データ間の因果性検出手法の検討

東京理科大学 池口研究室 修士2年 澤田 和弥



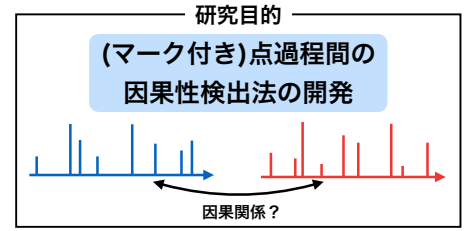
1. 研究動機と目的



時系列データに対する因果性検出手法

- Granger causality [2]
- Transfer entropy [3]
- Convergent cross mapping [4]

(マーク付き)点過程間の因果性検出手法はほとんどない



2. 因果性検出手法 (提案法)

(a) スパイク列 (例: ISI_2 を用いた ISI_1 の予測)

(i) 発火時間が近いスパイク (ISI) の参照
 スパイク列1: 1 2 3 4 5 6
 スパイク列2: 1 2 3 4 5 6 7 8 9 10

(ii) 状態空間の再構成[5] & 近傍を探す
 $ISI_2(n)$ vs $ISI_1(n)$

(iii) 重み w_1, w_3, w_8 の計算

(iv) 近傍ベクトルに対応するスパイクに発火時間が最も近い ISI を参照
 $ISI_1(1)$ $ISI_1(2)$ $ISI_1(5)$

(v) $ISI_1(4)$ の予測
 $\hat{ISI}_1(4) = w_1 ISI_1(1) + w_3 ISI_1(2) + w_8 ISI_1(5)$

(b) マーク付き点過程 (例: マーク付点過程2の情報を使ってマーク付点過程1を予測)

(i) タイミングに近いイベントを参照
 マーク付点過程1: $M_1(n)$ (1-6)
 マーク付点過程2: $M_2(n)$ (1-10)

(ii) 状態空間の再構成 & 近傍を探す
 $M_2(n)$ vs $M_1(n)$

(iii) 重み w_1, w_3, w_8 の計算

(iv) 近傍ベクトルに対応するイベントとタイミングに近いイベントを参照する
 $T_1(1)$ $T_1(2)$ $T_1(5)$

(v) IEIとマークの予測
 $\hat{T}_1(4) = w_1 T_1(1) + w_3 T_1(2) + w_8 T_1(5)$
 $\hat{M}_1(4) = w_1 M_1(1) + w_3 M_1(2) + w_8 M_1(5)$

3. データ生成モデル

(a) スパイク列 (gap junctionで結合したIzhikevichモデル)[6, 7]

$$\begin{cases} \dot{v}_i = 0.04v_i^2 + 5v_i + 140 - u_i + I_i + \sum_{j \neq i} w_{ij}(v_j - v_i) \\ \dot{u}_i = a_i(b_i v_i - u_i) \end{cases} \quad \begin{cases} v_i \geq 30 \text{ のとき} \\ v_i \leftarrow c_i \\ u_i \leftarrow u_i + d_i \end{cases}$$

(b) マーク付き点過程 (結合ローレンツ方程式[8])

$$\begin{cases} \dot{x}_1 = \sigma_1(y_1 - x_1) + C_{12}(x_2 - x_1) \\ \dot{y}_1 = r_1 x_1 - y_1 - x_1 z_1 \\ \dot{z}_1 = -b_1 z_1 + x_1 y_1 \\ \dot{x}_2 = \sigma_2(y_2 - x_2) + C_{21}(x_1 - x_2) \\ \dot{y}_2 = r_2 x_2 - y_2 - x_2 z_2 \\ \dot{z}_2 = -b_2 z_2 + x_2 y_2 \end{cases}$$

4. 結果

(a) スパイク列

発火率同じ (6ニューロン) vs 発火率異なる (6ニューロン)

予測精度 vs 予測に使ったISI時系列の長さ

大津法[9]による因果性検出結果のネットワーク図

(b) マーク付き点過程

両方向 vs 片方向(1→2) vs 結合なし

結合パラメータ $C_{12} = C_{21} = 1$, $C_{12} = 0, C_{21} = 1$, $C_{12} = C_{21} = 0$

因果性あり vs なし

提案手法の問題点

イベント数が大きく異なる場合

多 vs 少

窓を設定してイベント数を揃える?

5. まとめ

イベント間隔からの状態空間再構成 → イベント間隔の時間対応 → 提案手法により点過程間の因果性を検出可能

6. 今後(博士課程)の課題

- イベント数が異なるような点過程データに対して、点過程の変更を最小限にとどめ、イベントの挿入などを行う方法の検討
- マーク付点過程からの埋め込みの理論的保証
- 観測データに欠損があるような状況での因果性検出・推定
- 複数の因果性検出手法の検討及び統合の基盤 (非線形力学系理論+情報理論的なアプローチ)
- 実データ(神経スパイク列・経済データ・地震データなど)の解析とその結果得られた知見から手法の改善を検討
- 複雑ネットワーク上での因果性検出
- 因果性検出手法を用いた複雑ネットワーク上での情報流推定

参考文献

[1] S. Funahashi et al., *Cerebral Cortex*, 10, 6, 535–551, 2000.
 [2] C. W. J. Granger, *Econometrica*, 37, 3, 424–438, 1969.
 [3] T. Schreiber, *PRL*, 85, 2, 461–464, 2000.
 [4] G. Sugihara et al., *Science*, 338, 6106, 496–500, 2012.
 [5] T. Sauer, *PRL*, 72, 24, 3811–3814, 1994.
 [6] E. M. Izhikevich, *IEEE Trans. on NN*, 14, 6, 1569–1572, 2003.
 [7] 内木, 他, 信学論, J100-A, 5, 195–204, 2017.
 [8] V. S. Anishchenko et al., *Phys. Rev. E*, 57, 1, 316–322, 1998.
 [9] N. Otsu, *IEEE Trans. on systems, man, and cybernetics*, Vol. 9, No. 1, pp.62–66, 1979.

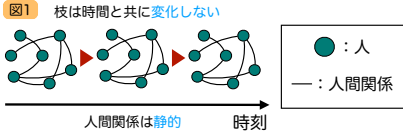
コンタクトの順番が情報拡散に与える影響の解析

東京理科大学 池口研究室 M1 ジョ ロエン



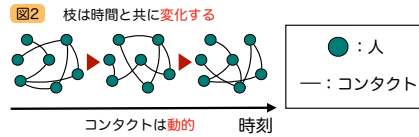
1 研究背景と動機

● 静的なネットワーク上での情報拡散



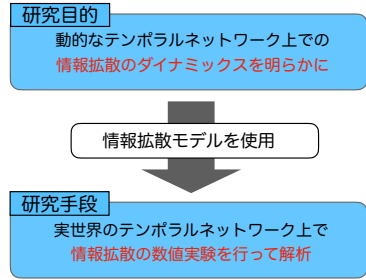
数値実験と社会実験は近年盛んに行われている[1-4]
数値実験で社会実験の結果を再現できている[3, 4, P1, P2]

● テンポラルネットワーク上での情報拡散



感染症拡散の研究では用いられているが、
それと類似した情報拡散の研究では用いられていない

2 研究目的



3 使用データ[5-7]

- (1) 高校での接触記録 (参加者: 126名, 平均コンタクト数(密度): 2.09/[20sec])
- (2) 病院での接触記録 (参加者: 75名, 平均コンタクト数(密度): 1.87/[20sec])

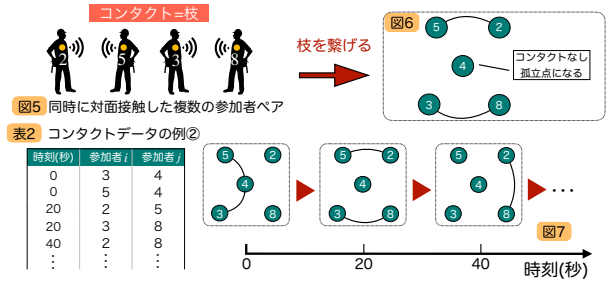
コンタクトは20秒ごとに一回観測される

● コンタクトデータの取り方



図3 コンタクトセンサー [5] 図4 対面接触した参加者

4 テンポラルネットワークの作成手法

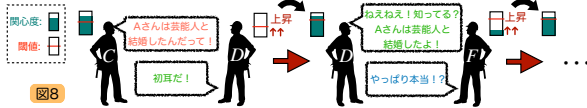


5 情報への関心度に着目した情報拡散モデル[4, P3]

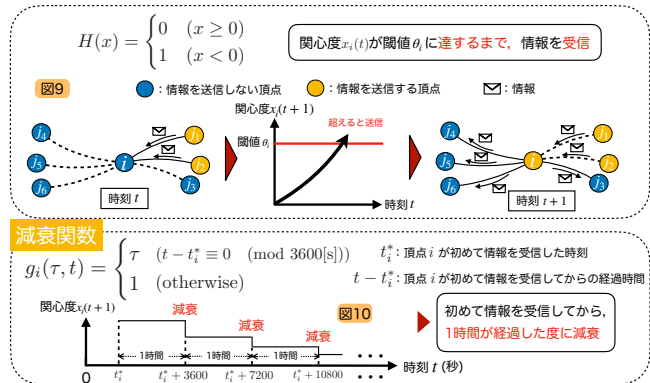
関心度 減衰関数 隣接頂点からの情報入力

$$x_i(t+1) = g_i(\tau, t) x_i(t) + \alpha H(x_i(t) - \theta_i) \sum_{j \in G_i(t)} F(x_j(t) - \theta_j)$$

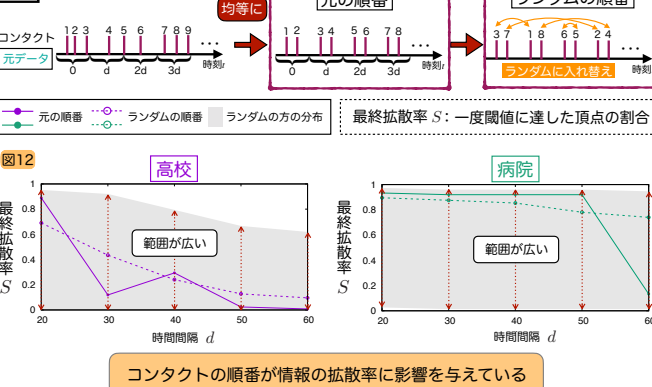
τ : 減衰係数 ($0 < \tau < 1$) α : 情報の影響力 $F(x)$: ステップ関数
 $G_i(t)$: 時刻 t での頂点 i の隣接頂点の集合 θ_j : 頂点 j の関心度の閾値 ($\theta_j \in [0, 1]$)



興味を持つ人から情報を受信し、興味を持ったら情報を送信する $x_i(t) > \theta_i$

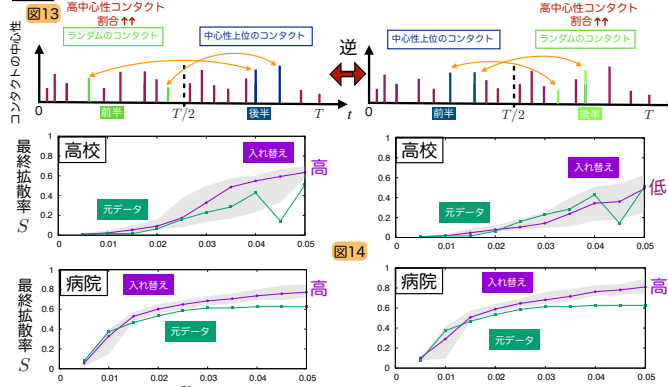


6 数値実験 A



コンタクトの順番が情報の拡散率に影響を与えている

7 数値実験 B



高中心性のコンタクトの割合大 → 情報が広く拡散する傾向

8 まとめ

実世界のコンタクトデータにおけるコンタクトの順番を変化させ、
情報の最終的な拡散率を調査した

高中心性のコンタクトの短期的な割合が
情報拡散に影響を与えていることが示唆

参考文献

[1] D. J. Watts & S. H. Strogatz, *Nature*, 393, 1440-1442, 1998.
 [2] D. Centola, *Science*, 329, 1194-1197, 2010.
 [3] 石黒, 他, 信学ソス, N-1-14, 2017.
 [4] A. Tanaka et al., Proc. of NOLTA2013, 298-301, 2013.
 [5] http://www.sociopatterns.org/ (最終閲覧日: 2020/11/26).
 [6] P. Vanhems et al., *PLOS ONE*, 8, e73970, 2013.
 [7] J. Fournet & A. Barrat, *PLOS ONE*, 9, e107878, 2014.

発表文献

[P1] ジョ, 他, 信学技報, 119, 471, 83-88, 2020.
 [P2] L. Xu et al., *IEICE Technical Report*, 120, 55, 23-26, 2020.
 [P3] ジョ, 他, 信学ソス, N-1-11, 2020.

TSPの近似解法に与える枝候補作成手法の検討

東京理科大学 池口研究室 土佐真義



1. TSP

全ての都市を訪れて最初に訪れた都市に戻る
最短の巡回路を求める組合せ最適化問題

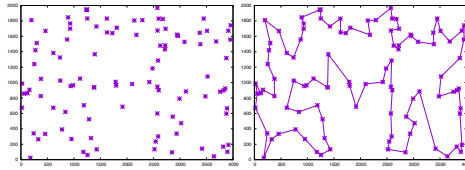


図1: TSPの問題例
(kroA100[1])

図2: kroA100の最適解

TSPを解くアルゴリズムは重要

2. 研究背景

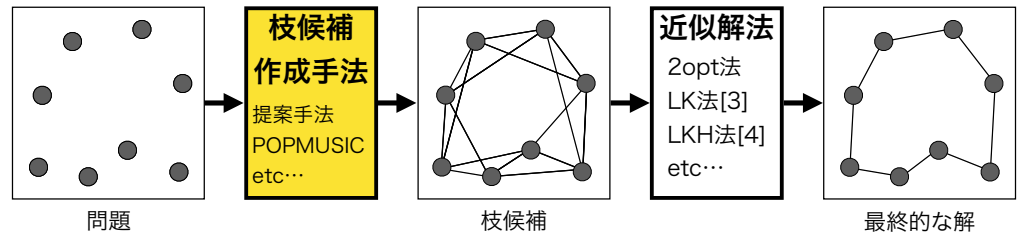
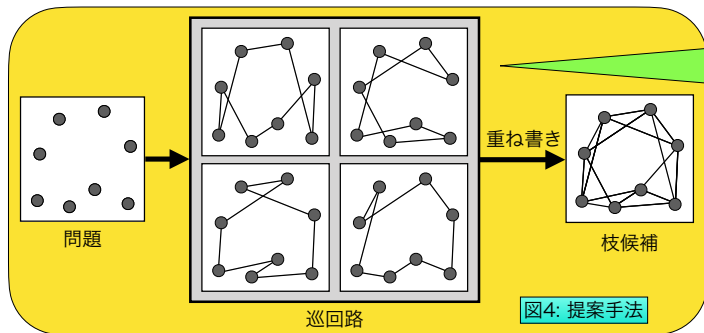


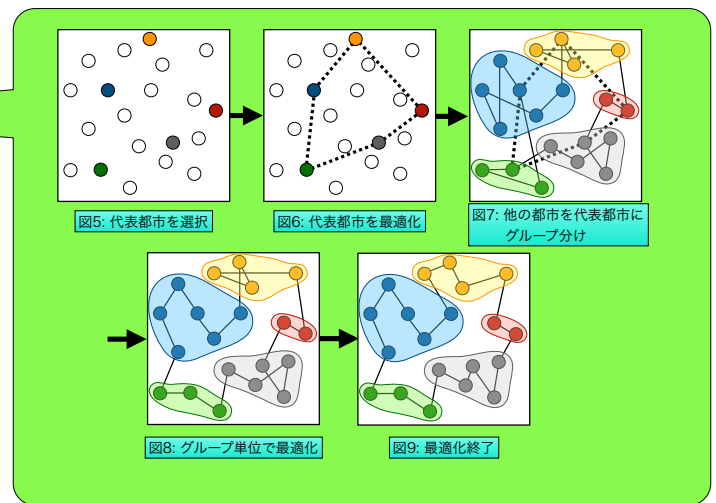
図3: 枝候補と近似解法の関係

枝候補は近似解法による最終的な解や計算時間に大きく影響

3. 提案手法



重ね書きに用いる巡回路の個数による枝候補の性能変化を調査



4. 数値実験

・重ね書きに用いる巡回路数と計算時間の関係

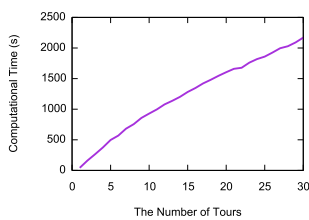


図10: 巡回路数と計算時間
(100,000都市問題)

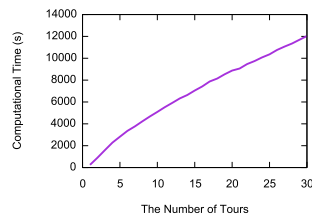


図11: 巡回路数と計算時間
(316,227都市問題)

計算時間は線形に増加

・重ね書きに用いる巡回路数と誤差率の関係

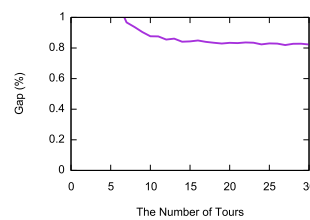


図12: 巡回路数と誤差率
(100,000都市問題)

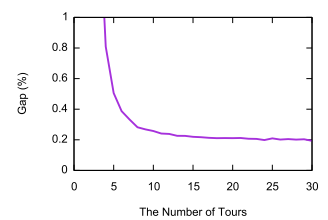


図13: 巡回路数と誤差率
(316,227都市問題)

ある程度巡回路数を増やせば誤差率に変化はない

・計算時間の詳細

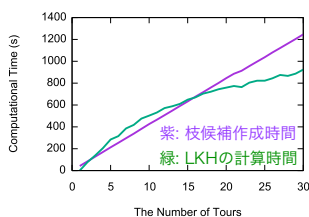


図14: 巡回路数と計算時間
(100,000都市問題)

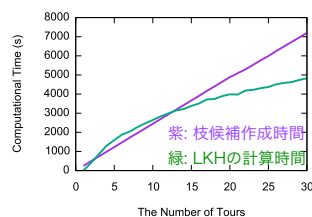


図15: 巡回路数と計算時間
(316,227都市問題)

LKHの計算時間は対数関数的に増加

・巡回路の個数と枝候補の枝本数

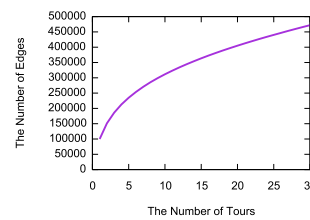


図16: 巡回路数と誤差率
(100,000都市問題)

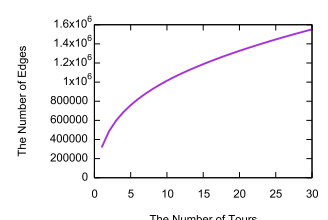


図17: 巡回路数と誤差率
(316,227都市問題)

LKHの計算時間と同様に対数関数的に増加

6. まとめ

- ・TSPの枝候補は近似解法の性能に直結
- ・枝候補の作成に用いる巡回路の個数も重要

参考文献

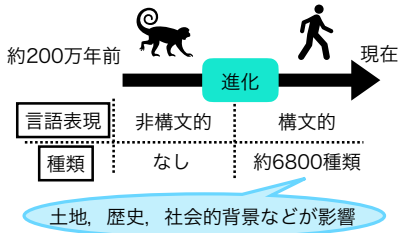
- [1] Gerhard Reinelt, *ORSA Journal on Computing*, 3, 4, 376–384, 1991.
- [2] Eric D Taillard, Keld Helsgaun, *European Journal of Operational Research*, 272, 2, 420–429, 2019.
- [3] Lin Shen, Kernighan Brian W, *Operations Research*, 21, 2, 498–516, 1973.
- [4] Keld Helsgaun, *European Journal of Operational Research*, 126, 1, 106–130, 2000.
- [5] David S Johnson, Lyle A McGeoch, Cesar Rego, Fred Glover, <http://archive.dimacs.rutgers.edu/Challenges/TSP/>, 2001.

新約聖書を対象とした言語ネットワークの解析

東京理科大学 池口研究室 M1 眞岸祈平



1. 研究背景と目的



言語の共通点

- ◆限られた単語から無限の文を作成できる
- ◆文生成の過程が迅速で頑健

単語のつながりが要因?

複雑ネットワーク理論の観点から言語を解析

先行研究

- スモールワールド性, スケールフリー性[1]
文を迅速に生成できる要因を示唆
- 共通の形成メカニズム[2]
各言語で局所的にネットワーク構造が類似
解析対象: 小説, コーパス
- 問題点① 文書内容が統一されていない
定性的な言語分類とは異なる[3]
同じ語族であってもネットワーク構造が異なる
異なる6言語を解析
- 問題点② 言語の種類が増やせない

本研究

新約聖書を用いて言語解析

- ① 文書内容の統一可能
- ② 言語の種類が豊富

目的

- ◆各言語の文章構造を定量的に解析
- ◆文書内容統一の必要性の検討
- ◆単語の使用頻度の影響度の調査

2. 使用データ

新約聖書

- ◆マタイによる福音書 ◆マルコによる福音書
- ◆ルカによる福音書 ◆ヨハネによる福音書

表1

言語	語族	バージョン	作成年
ギリシャ語1	インド・ヨーロッパ語族	NA28	2012(1898)
ギリシャ語2		UBS5	2014(1966)
ラテン語		Bublia Sacra Vulgata	1955
英語1		King James Version	1611
英語2		English Standard Version	2001
ドイツ語		Lutherbible	1984
フランス語1		Louis Segond	1910
フランス語2		Second 21	2007
ロシア語		Russian Synodal Version	1876
日本語1		日本語族	口語訳
日本語2	新共同訳		1987
フィンランド語1	ウラル語族	Vanha Kirkkoraamattu	1642
フィンランド語2		Uusi Kirkkoraamattu	1992
中国語	シナ・チベット語族	Chinese Standard Bible	2009

3. 言語ネットワーク作成手法

イエスがヘロデ王の代に、ユダヤのベツレヘムで

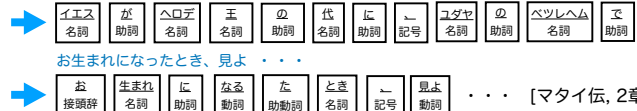
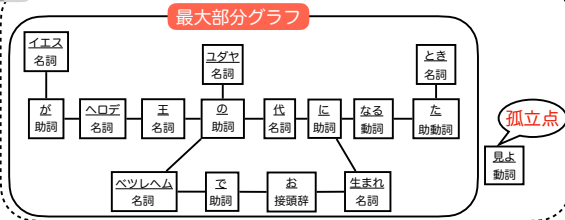


図1 重み無し無向ネットワーク



単語の共起に着目

単語の重複
つながりの重複
句読点
半句読点
記号

4. 実験と結果

(i) 単語数

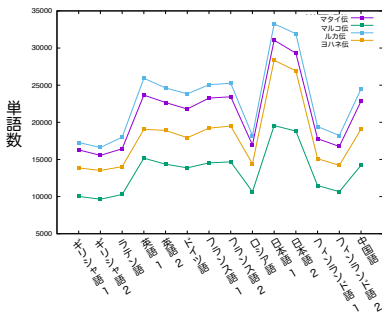


図2: 文章中の単語数

(ii) 頂点と枝数

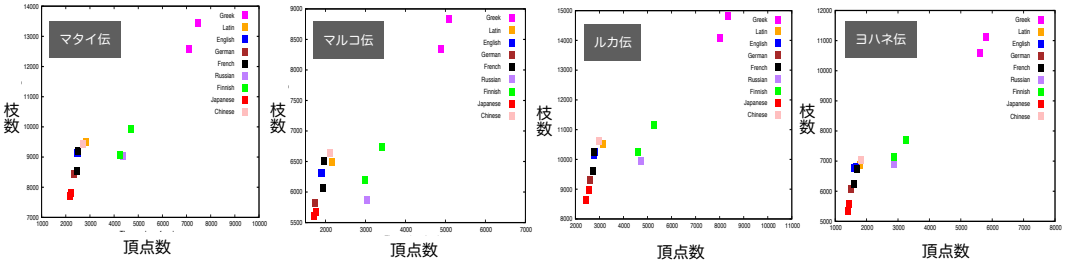


図3: 頂点数と枝数

◆文書内容に関わらず言語ごとに傾向は類似

(iii) クラスタ係数と平均頂点間距離

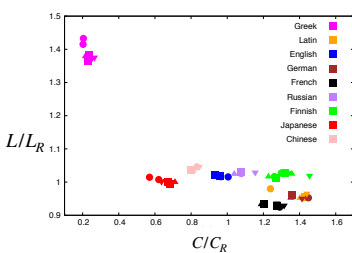


図4: クラスタ係数Cと平均頂点間距離L

(iv) ネットワーク構造の類似性

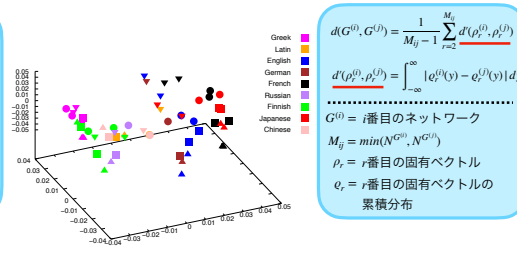


図5: スペクトルグラフ距離d(G⁽ⁱ⁾, G^(j))

(v) 単語の使用頻度の影響度

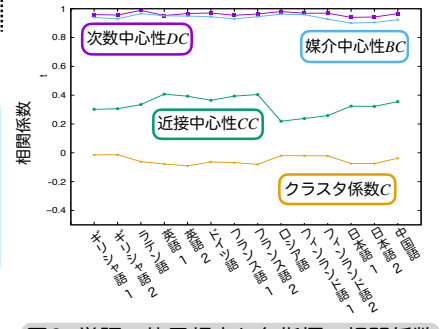


図6: 単語の使用頻度と各指標の相関係数

- ◆言語の文章構造の特徴を定量化可能
- ◆各言語の文章構造は特徴的

◆文書内容が異なると結果が異なる可能性

◆単語の使用頻度は中心性に影響

6. まとめ

- ◆言語の文章構造を定量化可能
- ◆各言語特有の文章構造が存在
- ◆文書内容の統一は重要な要素
- ◆単語の使用頻度は中心性に影響

課題

- ◆品詞ごとの構造解析
- ◆時代間比較

参考文献

- [1] Ramon Ferrer I Cancho et al., *biological science*, **268**, 1482, 2265, 2001.
- [2] Yutaka Shimada et al., *EPL*, **127**, 5, 56003, 2019.
- [3] Gao Yuyang et al., *Physica A*, **393**, 579–589, 2014.
- [4] Duncan J Watts et al., *Nature*, **393**, 440–442, 1998.
- [5] Yutaka Shimada, et al., *Scientific Reports*, **6**, 34944, 2016.

発表文献

- [P1] 眞岸 祈平 他, 信学技報, **119**, 471, 77–82, 2020.
- [P2] 眞岸 祈平 他, 信学技報, **120**, 55, 27–32, 2020.
- [P3] 眞岸 祈平 他, 信学ソ会, N-1-14, 2020.

スペクトルグラフ距離における累積寄与率と構造的特徴

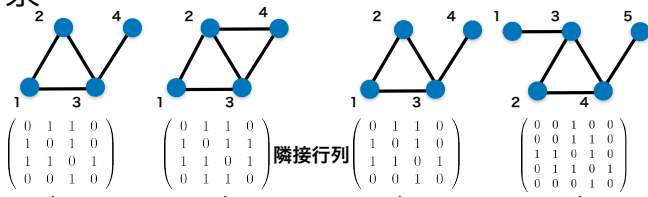
東京理科大学 工学部 情報工学科 B4 澤田 幸輝 池口研究室



1. 背景

同じ頂点数のネットワーク

異なる頂点数のネットワーク



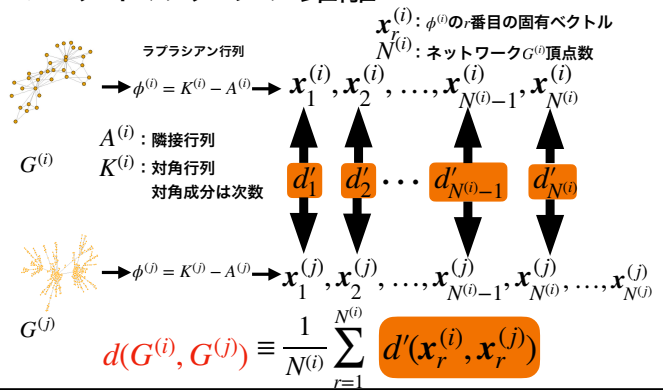
ハミング距離

ハミング距離

スペクトルグラフ距離

- 異なる頂点数のネットワークにも利用可能
- 有向ネットワークに利用可能

2. スペクトルグラフ距離



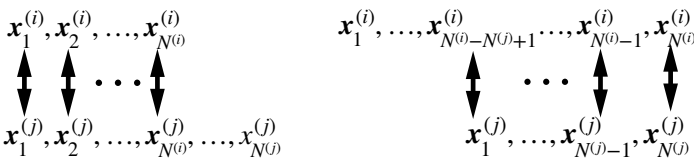
3. 提案手法

$x_r^{(i)}$: 各固有値に対応する固有ベクトル

$N^{(i)}$: 頂点数

従来手法

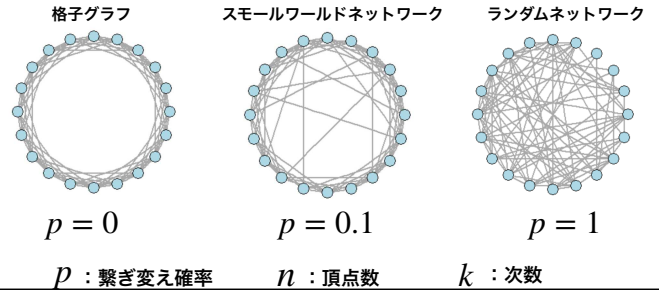
提案手法



- 従来: 寄与率の小さい固有ベクトルから使用
- 提案: 寄与率の大きい固有ベクトルから使用

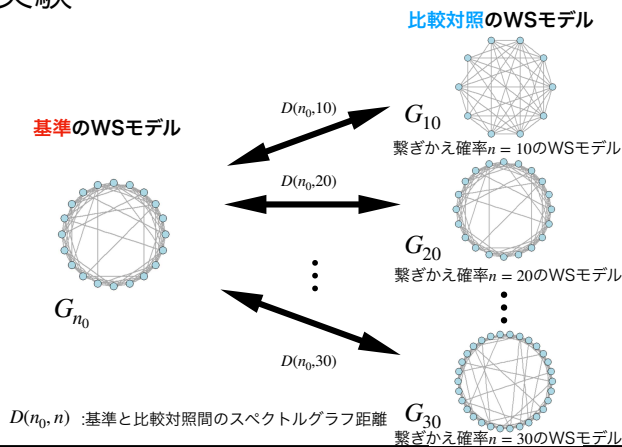
4. 使用ネットワーク

Watts-Strogatzモデル

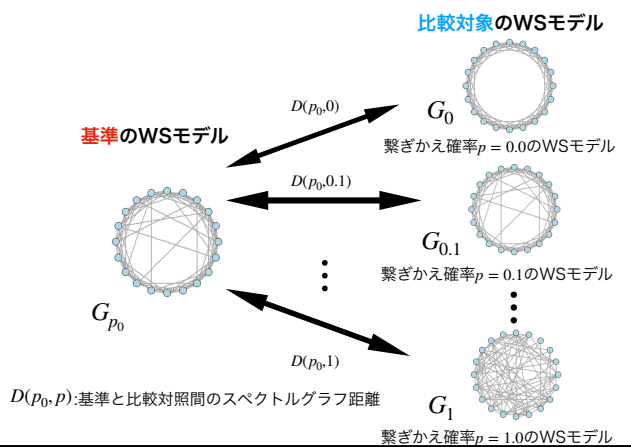


5. 数値実験

5.1 頂点数とスペクトルグラフ距離

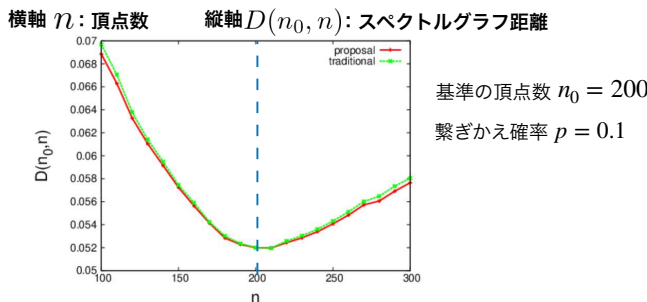


5.2 繋ぎかえ確率とスペクトルグラフ距離



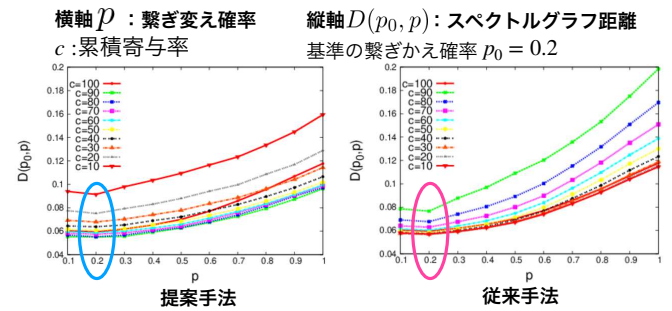
6. 結果

6.1 頂点数とスペクトルグラフ距離



構造の差異を定量化
提案手法と従来手法で傾向は変わらない

6.2 繋ぎかえ確率とスペクトルグラフ距離



提案手法: 寄与率の高い固有ベクトルのみで説明できる
従来手法: 寄与率の低い固有ベクトルのみで説明できる

7. まとめ

- 従来: 寄与率の小さい固有ベクトルから使用
 - 提案: 寄与率の大きい固有ベクトルから使用
- 寄与率の大小にかかわらず、
固有ベクトルが同等のネットワークの特性を反映

参考文献

[1] Yutaka Shimada, Yoshito Hirata, Tohru, Ikeguchi and Kazuyuki Aihara, "Graph distance for complex networks," *Scientific reports*, Vol. 6, No. 1, pp. 1–6, 2016.
[2] Duncan J Watts and Steven H Strogatz, "Collective dynamics of 'small-world' networks," *nature*, Vol. 393, No. 6684, pp. 440–442, 1998.

学位論文

修士論文要旨

主査: 池口 徹 教授
副査: 立川 智章 准教授
副査: 鮎川 矩義 助教

論文提出者 経営工学専攻 金丸 志生

論文題目 A Method for Extracting Nonlinear Time Series Property

by Changing Threshold of Recurrence Plot

リカレンスプロットの閾値変化による非線形時系列の特性抽出法

本論文では、観測時系列の特徴を2値の2次元画像へと可視化するリカレンスプロット (RP) を用いた非線形時系列の新しい特性抽出手法を提案している。RP は、2つの対応する時刻の状態間の距離が閾値より近ければ、対応する画素に点を打ち、遠ければ点を打たないと定義される。RP はシンプルな定義であるにも関わらず、時系列データの背後に潜む力学系の様々な情報を定性的に把握することができるため、有力な非線形時系列解析手法の一つとなっている。

また、RP で用いられる2点間距離情報は、カオス時系列解析において基礎となる情報である。その他にも重要な情報として、軌道ベクトルの情報が挙げられるため、本論文では、軌道ベクトルの情報を考慮した RP の拡張手法である同方向性リカレンスプロット (IDRP) も用いる。加えて、RP と IDRP から、同方向的近傍プロット (IDNP) を作成することができる。IDNP は、2つの対応する時刻の状態が近傍かつ同方向性の関係である場合のときのみ、対応する画素に点を打つ手法である。特に IDNP は時系列の決定論性検出に有効な手法である。また、IDNP は直接的な閾値を持たず、RP と IDRP のそれぞれの閾値に依存する。

一方で、RP、IDRP の閾値は慎重に設定する必要がある。これは閾値の設定を誤ると時系列の特徴をうまく捉えられない場合があるためである。しかしながら、時系列に応じた最適な閾値は、力学系の構造に強く依存することから、一意に定めるのは難しい。

そこで本論文では、RP と IDRP の閾値を変化させて、閾値の変化により RP、IDRP、IDNP 上のプロットパターンがどのように推移するのかを調査し、その推移傾向を時系列の特性として捉える手法を提案する。このとき、再帰定量化解析手法 (RQA) を用いることで、閾値と RP 上のプロットパターンの関係を定量的に捉えることができる。

まず、提案手法の妥当性を調査するために、非線形力学系から観測される決定論的カオス時系列と確率論的な時系列それぞれに提案手法を適用し、これらの特性判別の有効性を調査した。次に、振る舞いが似るとされる決定論的カオス時系列と周期応答にダイナミカルノイズが加わった時系列に提案手法を適用し、これらの特性判別の有効性を調査した。数値実験結果より、決定論的カオス時系列と周期応答にダイナミカルノイズが加わった時系列の特性判別が可能であることが示された。加えて、提案手法によって、決定論的カオス時系列と非整数ブラウン運動のような有色ノイズ時系列の判別も可能であることが示された。また、既存の RP の閾値設定方法による特性の誤検出の危険性が明らかとなり、提案手法による閾値変化に伴うプロットパターンの推移傾向を調査することの必要性が示唆された。

修士論文要旨

主査: 池口 徹 教授
副査: 立川 智章 准教授
副査: 鮎川 矩義 助教

論文提出者 経営工学専攻 金丸 志生

論文題目 A Method for Extracting Nonlinear Time Series Property

by Changing Threshold of Recurrence Plot

リカレンスプロットの閾値変化による非線形時系列の特性抽出法

In this thesis, using a recurrence plot (RP) which visualizes the features of an observed time series into a binary two-dimensional image, we proposed a new method for extracting properties of nonlinear time series. RP can be defined as follows: if the distance between two states at two corresponding times is closer than a threshold, a point is plotted on the corresponding pixel; if it is farther, no point is plotted. Despite its simple definition, RP is one of the most powerful nonlinear time series analysis methods because it can qualitatively grasp difference between various dynamical systems behind the time series.

The distance information between two points used in RP is the basic information in chaotic time series analysis. In addition, the information of the trajectory vectors is also important. Therefore, in this thesis, we also use the information of the trajectory vectors in the iso-directional recurrence plot (IDRP), which is an extended method of RP. Besides, from the RP and the IDRP, we can create an isotropic neighborhood plot (IDNP). IDNP is a method of plotting the corresponding pixel only when the states of the two corresponding times are near and in the same direction. In particular, IDNP is an effective method to detect determinism of time series. Moreover, IDNP does not have a direct threshold but depends on the threshold of RP and IDRP.

On the other hand, the thresholds of RP and IDRP need to be set carefully. This is because if the thresholds are set incorrectly, the features of the time series may not be captured well. However, it is difficult to uniquely determine the best threshold value for each time series because it strongly depends on the structure of the target system. In this thesis, we investigate the relationship between the threshold values and the plot patterns on RP, IDRP, and IDNP. Then, we proposed a method for capturing the time series properties by observing the transition of the recurrence quantification analysis (RQA) statistics. Using RQA, the relationship between the thresholds and the plot patterns on the RPs, IDRPs and IDNPs can be captured quantitatively.

First, to investigate the validity of the proposed method, we applied the proposed method to deterministic chaotic time series observed from nonlinear dynamical systems and stochastic time series. Next, we applied the proposed method to the deterministic chaotic time series and the noisy periodic time series. Then, we investigated the effectiveness of the discrimination of these properties. The numerical experiments showed that the proposed method could discriminate the deterministic chaotic time series and the noisy periodic time series. Besides, it is shown that the proposed method can discriminate between deterministic chaotic time series and colored noise time series, such as the fractional Brownian motion. In addition, the danger of false detection of the properties by the existing threshold setting method of RP was revealed, which implies that it is essential to apply the proposed method which investigates the trend of the plot patterns with the change change of the threshold of RP.

修士論文要旨

主査: 池口 徹 教授
副査: 立川 智章 准教授
副査: 鮎川 矩義 助教

論文提出者 経営工学専攻 澤田 和弥

論文題目 Detecting Causality for Point Processes Based on Nonlinear Dynamical System Theory
非線形力学系理論に基づいた点過程データに対する因果性検出

観測技術の向上により、様々な分野で高精度に多次元の時系列データを観測可能となった。我々は、サンプリング間隔が一定な時系列データ以外にもサンプリング間隔が一定ではない多次元の点過程データも数多く観測することができる。典型的な点過程データとしては、神経活動により得られるスパイク列や、地震活動、経済活動から得られるマーク付点過程データが存在する。スパイク列はニューロンの発火時間のみの点過程データであるが、マーク付点過程は、イベント発生時刻に何らかの情報が含まれているデータである。

しかし、このように多次元の点過程データを観測できるようになったにも関わらず、点過程生成の背後にある事象間の因果関係が明らかになっていない場合が多い。得られた時系列データ、点過程データから、事象間の関連性、すなわち因果関係を明らかにすることができれば、自然現象や経済活動など我々の社会に遍在する種々の複雑現象を理解する一助となるだけでなく、予測・診断・制御などの種々の工学的応用が可能となる。

これまで、時系列データ間の因果性を検出する手法として、線形回帰に基づく Granger causality や情報理論に基づく transfer entropy などが広く用いられている。しかし、これらの手法は決定論的非線形力学系から得られる時系列データに対しては、真の因果関係と疑似相関を区別できないことなど、様々な問題を有しているのも事実である。一方、非線形力学系理論に基づく因果性検出手法として convergent cross mapping (CCM) が提案されており、決定論的非線形力学系に対する因果性検出において有効であることが示されている。しかし、非線形力学系理論の立場からの点過程データに対する解析手法は十分に整備されていないのが現状であり、点過程データに対する因果性検出手法はほとんどない。

そこで、本論文では、CCM を拡張することで、スパイク列のようなマークのない点過程データ、マーク付点過程データに対する因果性検出手法を提案する。CCM を用いて因果性を検出する際には、観測時系列から状態空間を再構成する必要があるが、スパイク列に対して遅延座標系を用いることで、状態空間を再構成可能なことはすでに議論されているが、マーク付点過程からの状態空間再構成は議論されていない。本論文では、マーク付点過程に対して遅延座標系を用いて状態空間を再構成する方法を提案し、その妥当性を数値実験により検証した。結果として、遅延座標系を用いることで、マーク付点過程から状態空間を再構成できることが示唆された。また、数理モデルから生成したスパイク列、マーク付点過程に対して、提案手法を適用することで、因果性検出が可能となることがわかった。

修士論文要旨

主査: 池口 徹 教授
副査: 立川 智章 准教授
副査: 鮎川 矩義 助教

論文提出者 経営工学専攻 澤田 和弥

論文題目 Detecting Causality for Point Processes Based on Nonlinear Dynamical System Theory
非線形力学系理論に基づいた点過程データに対する因果性検出

In recent years, improvements in observation technologies have made it possible to observe multi-dimensional time series data with high precision in many fields. In addition, we can also observe not only time series data with regular sampling intervals but also point process data with irregular sampling intervals. Typical examples of the multi-dimensional point process data and marked point process data include spike trains obtained from neural, seismic and economic activities. Spike trains are point process data only with the firing timing of neurons, while marked point processes are data with some information and the event timing.

In this way, although we can observe multi-dimensional point process data, causal relationships between the events behind the generation of point processes are not yet clear in many cases. If we can clarify the causal relationships between events from observed multi-dimensional point process data, it cannot only help us understand various complex phenomena that are ubiquitous in our society, such as natural phenomena and economic activities, but also enable various engineering applications such as prediction, diagnosis, and effective control.

Various methods have been proposed to detect causality between time series data, such as the Granger causality based on linear regression and transfer entropy based on the information theory. However, these methods have various problems, such as the inability to distinguish actual causality from spurious causality for time series data obtained from nonlinear deterministic dynamical systems. Therefore, convergent cross mapping (CCM), a causality detection method based on nonlinear dynamical systems theory, has been proposed. It is effective in causality detection for nonlinear deterministic dynamical systems. However, the analysis method for point process data based on the nonlinear dynamical system theory has not been sufficiently developed yet, and there are few causality detection methods for point process data.

Therefore, in this thesis, we propose a method for detecting causality for spike trains and marked point processes by modifying CCM. To detect causality using CCM, it is necessary to reconstruct the state space from the observed time series. The state space reconstruction from interspike intervals of spike train has already been discussed. However, the state space reconstruction from marked point processes has not been fully explored. Therefore, in this thesis, first we propose a method to reconstruct the state space from the marked point process using a delay coordinate system; then, we verified the validity of the proposed method by numerical experiments. As a result, it is suggested that the state space can be reconstructed from the marked point process using the delayed coordinate system. Next, we apply the proposed method of causality detection to the spike trains and marked point processes generated from the mathematical model. As a result, the proposed method is effective for detecting causality for spike trains and marked point processes.

修士論文要旨

主査: 池口 徹 教授
副査: 立川 智章 准教授
副査: 鮎川 矩義 助教

論文提出者 経営工学専攻 モウ フクカ

論文題目 Effectiveness of Visibility Graph for Marked Point Processes and its Application to Analyzing Structural Features of Musical Composition
マーク付点過程に対する Visibility Graph の有効性の検討と楽曲構造解析への応用

複雑ネットワーク理論において、時系列データとネットワークの関係を解析する研究が盛んに行われている。例えば、2006年にJ. Zhangらにより時系列をネットワークに変換する手法が提案されている。また、2012年に島田らによりネットワークを時系列に変換する手法が提案されている。これらの手法により、時系列データとネットワークを関連づけて解析することの有効性が証明されている。しかし、現実世界においては、連続的に観測して得られた時系列データよりもむしろ、離散的なイベント系列に何らかの情報が付与されたマーク付点過程データのほうが多く存在する。このように記録されるデータの例として、地震の発生タイミング、各種金融市場の値動きなどがある。しかし、複雑ネットワーク理論において、これらのマーク付点過程とネットワークの関係については十分に議論されていない。そこで、本論文では、マーク付点過程とネットワークの關係に注目する。

本論文では、まずマーク付点過程に対して、Visibility Graph (VG) 法を適用するときの有効性を検討する。VG法は時系列データをネットワークに変換する手法の一つであるが、時系列データの値と時間情報を両方使用するため、マーク付点過程にも適用できると考えられる。そこで、マーク付点過程に対してVG法の有効性を調査するため、非線形常微分方程式から作成したマーク付点過程に対してVG法を適用し、ネットワークを作成する。具体的には、ネットワーク間距離を用いて算出した構造的特徴の差異を解析することで、マーク付点過程に対するVG法の有効性を調査した。その結果、マーク付点過程に対してVG法が有効であることが示された。

また、本論文ではマーク付点過程をネットワークに変換する応用として、楽曲データをマーク付点過程データとして扱い、ネットワークに変換することで複雑ネットワーク理論の観点から解析を行う。その際、VG法を拡張した新たなネットワーク変換手法を提案する。拡張したVG法は音符の高さだけではなく、音符の継続時間も考慮しているため、従来のVG法よりも楽曲データの解析に適していることが明らかとなった。提案手法を用いて得られたネットワークに対し、ネットワーク間距離を定量化する手法を用いて、楽曲構造の差異を解析する。更に、楽曲の特徴を定量化するため、リカレンスプロットとその定量化手法を用いた楽曲の決定論性を計算する手法を提案し、決定論性の差異という観点から楽曲解析を行う。結果として、決定論性は楽曲の構造を定量化できる指標となり得ることが示された。

修士論文要旨

主査: 池口 徹 教授
副査: 立川 智章 准教授
副査: 鮎川 矩義 助教

論文提出者 経営工学専攻 モウ フクカ

論文題目 Effectiveness of Visibility Graph for Marked Point Processes and its Application to Analyzing Structural Features of Musical Composition
マーク付点過程に対する Visibility Graph の有効性の検討と楽曲構造解析への応用

The relationship between time series and networks has been widely discussed in the theory of complex networks. For example, J. Zhang, M. Small *et al.* proposed a method to transform time series into networks in 2006, and Y. Shimada, T. Ikeguchi *et al.* proposed a method to transform networks into time series in 2012. Both of these studies have already been proved to be effective to analyze the relationship between time series and networks. Most commonly, time series data are a sequence taken at successive equally spaced points in time. However, there are different types of data that can only be observed at irregular intervals in the real world, which are called marked point process data. For example, occurrence timing in seismic activities data, changes in the stock price in financial markets, and many other data are described by the marked point process data. Therefore, it is also important to study the relationship between marked point processes and networks which has rarely been discussed in the theory of complex networks.

In this thesis, we first discuss the possibility to transform marked point processes into networks by using the visibility graph (VG) method. Despite the VG method being one of the methods to transform time series into networks, it is considered that the VG method can also be applied to marked point processes because the VG method uses both values and time information. To investigate the effectiveness of the VG method for marked point processes, we apply the VG method to the marked point processes created from the nonlinear ordinary differential equations and investigate the difference in structural features between networks created by using the inter-network distance. As a result, we show that the VG method can reflect the characteristics of the marked point processes properly.

Also, as an application of transforming marked point processes into networks, we treat musical compositions data as marked point processes and transform them into networks that preserve structural features of the corresponding marked point processes. We can analyze the structures of musical compositions through the corresponding marked point processes and networks. Then, we propose a two-dimensional extended marked point process for musical compositions and a new method of transforming the extended marked point processes to networks by using the VG method. We show that the proposed method is more suitable for analyzing musical composition data than the conventional VG method because the proposed method treats not only the heights of notes but also the duration of notes. Further, we investigate how to extract features of musical compositions by focusing on the temporal structure changes. In particular, we calculate determinism of the musical composition data, which can represent the regularity of musical composition. As a result, we show that determinism can be a good indicator for quantifying the regularity of musical composition.

修士論文要旨

主査: 池口 徹 教授
副査: 立川 智章 准教授
副査: 鮎川 矩義 助教

論文提出者 経営工学専攻 宮 南風

論文題目 A Study on Emerging of Brain Disorders Induced by Neuron Ratio Imbalance
Using a Mathematical Model of Spiking Neural Networks
数理モデルを用いたニューロン比率不均衡が導く脳機能障害の発現に関する研究

ヒトの脳内には千億個以上のニューロンが存在する。ニューロンは信号を受け取り、内部状態がある一定値を超えて発火することで結合する他のニューロンに信号を送る。また、ニューロンには他のニューロンの発火を促進する興奮性ニューロンと過度な発火を制御する抑制性ニューロンが存在する。これらの均衡は同期発火の促進と過剰発火を制御していると考えられている。生理学実験により得られた知見から、ヒトの大脳皮質には興奮性ニューロンが80%、抑制性ニューロンが20%存在することが知られている(生理学的比率)。しかし、なぜこの比率なのか、また、この比率が変化した場合、脳の情報処理にどのような影響を与えるのかは明らかにされていない。一方、興奮性/抑制性ニューロン比率の不均衡が自閉スペクトラム症、ダウン症、統合失調症、アルツハイマー病など、様々な疾患で観測されており、この不均衡性がこれらの脳機能障害を誘引していることが示唆されている。

本論文では、まずスパイクングニューラルネットワークの数理モデルを用いて、興奮性/抑制性ニューロン比率がネットワークに与える影響の定量的評価を行う。特に、実際の脳内で観測されているスパイクタイミング依存可塑性(STDP)のシナプス学習を適用することにより、ニューロンの長期的な活動や構造形成への影響を調査する。その際、神経科学分野における基本的な手法に加え、情報理論、複雑ネットワーク理論を用いた定量的解析を行う。

結果として、第一に、ニューロンの活動度を調査したところ、ネットワーク上の平均発火率、同期度は生理学的比率で最大化することが明らかとなった。また、シナプスの学習過程を調査した結果、STDPによる長期抑圧と抑制性ニューロンによる直接的な抑圧のトレードオフ関係が存在することを確認した。この結果は、2種類の抑圧による影響が最小化された比率が生理学的比率に対応することを示唆するものである。第二に、情報理論を用いた解析として、Delayed Transfer Entropyを拡張した指標を用いてシナプスの情報流を定量化した。その結果、シナプス単体での局所的な情報量は、生理学的比率で必ずしも最大化されないが、複数のシナプスにまたがる大域的な情報量は最大化されることが明らかとなった。第三に、複雑ネットワーク理論に基づいた有向重み付きネットワークとしての解析を行った。その結果、STDP学習により形成されるフィードフォワード構造がニューロンの活動に与える影響は微小であることを明らかにした。また、各ニューロンの出シナプス強度の相関が高い場合に発火率は上昇することも示された。

これらの結果を受けて、本論文では、ニューラルネットワークで得られる神経活動データから、fMRI信号に変換する数理モデルを用いて、ダウン症患者の脳ネットワークの異常性の知見再現を試みた。その結果、物理的な距離の違いによる活動相関の分布を再現することに成功した。

以上より、興奮性/抑制性ニューロン比率の不均衡性は脳機能の活性低下を引き起こすことを数値実験から明らかにした。これらの結果は、数理モデルを用いた調査により、各疾患の障害の原因解明を可能とすること、さらには、これらの疾患の治療法の開発に貢献するものである。

修士論文要旨

主査: 池口 徹 教授
副査: 立川 智章 准教授
副査: 鮎川 矩義 助教

論文提出者 経営工学専攻 宮 南風

論文題目 A Study on Emerging of Brain Disorders Induced by Neuron Ratio Imbalance
Using a Mathematical Model of Spiking Neural Networks
数理モデルを用いたニューロン比率不均衡が導く脳機能障害の発現に関する研究

There are more than 100 billion neurons in the human brain. A neuron receives signals and sends signals to other connected neurons by firing when its internal state exceeds a threshold value. Also, excitatory neurons promote the firing of other neurons and inhibitory neurons that depress excessive firing. We consider that balancing these two types of neurons can contribute to the control of synchronized firings and excessive firings. Physiological experiments have shown that the cerebral cortex contains 80 percent excitatory neurons and 20 percent inhibitory neurons (the physiological ratio). However, it has not been clarified why this ratio exists and how this ratio affects information processing in the brain. In addition, excitatory/inhibitory imbalance has been observed in various diseases such as autism spectrum disorder, Down syndrome, schizophrenia, and Alzheimer's disease. It has been suggested that this neural imbalance induces brain impairment. In this thesis, we use a mathematical model of spiking neural networks to evaluate the effect of neuron ratio on the brain functions quantitatively. In particular, we apply the synaptic learning of Spike timing-dependent plasticity (STDP), which has been observed in the real brain, to investigate the effects on activities in the long-term and structure formation of neural networks. We use information theory and complex network theory for quantitative evaluation in addition to basic neuroscience methods.

As a first result, from the investigation of neural activity, we clarified that the average firing rate and the degree of synchronization in the network are maximum at the physiological ratio. We also investigated the learning process of synapses and suggested a trade-off relation between long-term depression by STDP and direct depression by inhibitory neurons. Results show that the physiological ratio corresponds to the ratio in which the two types of suppression was minimum. Secondly, we quantified synaptic information flow for analysis based on information theory using an extended Delayed Transfer Entropy. The results showed that the global information flow across multiple synapses was maximum though the local information at single synapses was not maximum by the physiological ratio. Thirdly, we analyzed the results as a directed weighted network based on the complex network theory. Then, we clarified that the feed-forward structure formed by STDP learning had only a small effect on the neural activity. We also found that the firing rate increased when the correlation between the output synaptic strengths of each neuron was high. Based on these results, in this thesis, we attempt to reproduce the abnormalities in the brain networks of patients with Down syndrome using a mathematical model that converts neural activity data obtained by neural networks into fMRI signals. As a result, we succeeded in reproducing the distribution of activity correlations due to differences in physical distance. We conclude that the ratio of excitatory/inhibitory neuron imbalance depressed in functional brain activities from numerical experiments. These results allow elucidating the cause of disorders and contribute to the development of a therapeutic method by using the mathematical model.

マスクの着用の有無を対照条件として測定した 光電脈波に対する非線形時系列解析

岡安高輝 (池口徹 教授, Nina Sviridova 助教)

1 はじめに

現在, 日本政府は新型コロナウイルス (COVID-19) 感染症対策として外出時にはマスクを着用することを推奨している [1]. しかし, マスクの着用の有無が身体にどのような影響を及ぼすかは十分な調査が行われていない. 一方, 光電脈波 (photoplethysmogram, 以下 PPG) データは非侵襲に測定可能な心臓血管系の関連情報を持つデータとして知られている [2]. また, マスクを着用せずに安静にした状態で測定された PPG データは, カオス的振る舞いを示すことも報告されている. [3, 4]. そこで本論文では, 非線形力学系理論の立場から, マスクの着用の有無が身体にどのような影響を与えるのかを, PPG データに対するカオス時系列解析により明らかにする. 具体的には, カオス力学系の特徴の一つである初期値鋭敏依存性を定量化する指標であるリアプノフ指数を算出し, マスクを着用した場合としない場合で, PPG データのリアプノフ指数に有意差があるのかを調査した.

2 実験

2.1 Photoplethysmogram データ

PPG データは, 末梢血管の血流量の変化を表した時系列データである. 血管の血流量は定期的に変化することを利用して, 指先の一方から LED を投光し相対する受光素子により透過光の変化量を測定する. PPG データからは, 心拍数, 血中酸素飽和度 (以下 SpO_2) や血圧の検出, 自律神経機能の評価などの心血管系に関する情報を得ることができる [2]. また, PPG データは非侵襲な計測方法であり計測機器が安価という特徴を持つため, 医療機器の他にスマートウォッチなどのウェアラブルデバイスにも広く用いられている.

2.2 実験条件

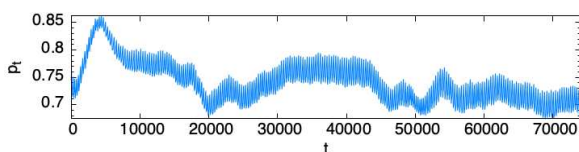
実験では, パルスオキシメータ (東京デバイセズ株式会社) [5] を用いて PPG データを非侵襲に計測した. 計測機器のサンプリング周波数 $f_s = 409.6$ [Hz] である. 被験者は, 東京理科大学の学生のうち, 21~28 歳の健康な男女 14 名である. 「東京理科大学における人を対象とする医学系研究に関する倫理規程」 [6] に従い, 予め実験手順を被験者に対して説明し, 同意を得た後実験を行った. 室温 $22 \sim 26^\circ\text{C}$ の静かな室内にて, リラックスした状態で 1 回 3 分間の PPG データを 1 人あたり 2 回もしくは 3 回計測した. この実験をマスクを着用した場合としない場合で実施した. また, 実験を実施する前に, 血圧計と SpO_2 が計測できるパルスオキシメータを用いて, 血圧と SpO_2 , 心拍数を記録した. 計測した PPG データは, マスクの着用ありが 13 人の 37 回分, マスクの着用なしが 12 人の 33 回分である.

2.3 実験結果

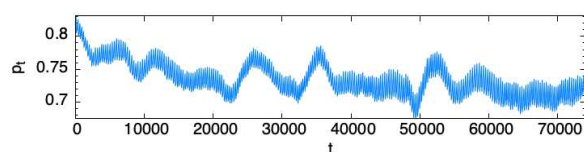
マスクあり, なしの条件における被験者 4 の 3 回目の PPG データを対象に解析した結果を示す. 表 1 は, 実験実施時の血圧と心拍数, SpO_2 の記録と, 計測した PPG データの基礎統計量 (データ数, 最大値, 最小値, 平均, 標準偏差) である. また, 計測した PPG データ p_t の時系列波形を図 1 に示す.

表 1: 実験情報と計測した PPG データの基礎統計量.

条件	日時	室温	最高血圧	最低血圧	心拍数	SpO_2	データ数	最大値	最小値	平均値	標準偏差
マスクあり	9月16日 13:37	25[°C]	106[mmHg]	76[mmHg]	81[bpm]	98[%]	79,946	0.862	0.674	0.737	0.037
マスクなし	9月16日 13:22	26[°C]	106[mmHg]	76[mmHg]	81[bpm]	98[%]	74,032	0.831	0.674	0.734	0.026



(a)



(b)

図 1: 計測した PPG データ p_t の時系列波形. (a) マスクあり, (b) マスクなし.

3 解析手順

3.1 トレンドの除去

まず、実験によって得られた PPG データから、被験者が実験中に計測機器を動かしたことなどで生じた観測ノイズを取り除くことを目的として、 t ステップ毎に前後 30 点の移動平均 $T_t = \frac{1}{61} \sum_{k=t-30}^{t+30} p_k$ をトレンドとして算出する。解析では、トレンドを除去した時系列 $x_t = p_{t+30} - T_{t+30}$ を対象とする。

3.2 状態空間の再構成

現実存在するシステムから、全ての状態変数を観測できる例は稀である。本論文においても、1次元の PPG データが計測されるのみで、元のシステムの全ての変数を観測することはできない。そこで、1次元の時系列データ $x_t (t = 1, 2, 3, \dots, n)$ を時間遅れ座標系へと変換して状態空間を再構成する [7, 8]。この変換は、再構成状態空間の次元数を m 、遅れ時間を τ として、 $\mathbf{X}_t = (x_t, x_{t+\tau}, \dots, x_{t+(m-1)\tau})$ を構成する。このとき、 $t = 1, 2, 3, \dots, N$ 、 $N = n - (m-1)\tau$ となる。本論文では、各 PPG データに対して、False nearest neighbors (以下 FNN) 法 [9] を適用して再構成次元数 m を推定する。また、自己相関関数の値が最初に $1/e$ 以下になる時刻を遅れ時間 τ とした [10]。

3.3 False nearest neighbors 法による再構成次元数 m の検討

FNN 法 [9] は、状態空間を再構成するときに必要な条件としての次元数 m を定めるために用いられる手法である。これは、再構成状態空間の次元数 m を大きくしたときに、誤り近傍点数が 0 に収束する m を採用する手法である。再構成状態空間の次元を m から $m+1$ としたときに、 m 次元では点 \mathbf{X}_i の近傍であった点 \mathbf{X}_j が $m+1$ 次元では近傍でなくなったとき、点 \mathbf{X}_j は点 \mathbf{X}_i の誤り近傍点という。FNN 法によって推定される次元数 m が埋め込みとなる数学的な保証はないが、実際の解析では、より低次元の再構成状態空間を用いることが望ましいため、次元数 m を決める際の一つの基準として用いられている。

FNN 法の手順を示す。まず、 m 次元再構成状態空間における点 \mathbf{X}_i とその最近傍点 \mathbf{X}_j の 2 点間距離は、

$$D_m(i, j) = \sqrt{\sum_{k=0}^{m-1} (x_{i+k\tau} - x_{j+k\tau})^2}$$

である。次に、 m 次元と $m+1$ 次元の再構成状態空間における点 \mathbf{X}_i と点 \mathbf{X}_j の 2 点間距離についての相対距離は $R_L = \sqrt{\frac{D_{m+1}(i, j)^2 - D_m(i, j)^2}{D_m(i, j)^2}} = \frac{|x_{i+m\tau} - x_{j+m\tau}|}{D_m(i, j)}$ と表す。このとき、 R_L が閾値 θ より大きい場合、点 \mathbf{X}_j は点 \mathbf{X}_i の誤り近傍点であると定める。本論文では $\theta = 10$ とした [9]。最後に、誤り近傍点を再構成状態空間内の全ての点に対して算出する。 $R_L > \theta$ である点の総数を R_m とおけば、誤り近傍率は $\frac{R_m}{n - (m-1)\tau}$ となる。

3.4 最大リアプノフ指数の算出

カオス力学系の最も重要な性質の一つに初期値鋭敏依存性がある。カオス力学系において、ある初期値とその初期値に微小な変化を与えた第 2 の初期値を与えると、これらの 2 つの初期値から始まる解軌道の挙動が全く異なるという性質である。このような性質を軌道不安定性とも呼ばれる。軌道不安定性はリアプノフ指数によって定量化される。最も大きいリアプノフ指数 (largest Lyapunov exponent, 以下 LLE) が正のとき、その力学系はカオス応答を示す。本論文では、LLE を文献 [11] に基づいて求める。文献 [11] のアルゴリズムは、利用可能なすべてのデータを用いるため、小さなデータセットに対しても LLE の推定精度が高く、加えて 1次元のデータからも LLE の推定が可能という特徴を持つ。

具体的な LLE 算出の手順を示す。まず、アトラクタ上の点 \mathbf{X}_j と点 $\mathbf{X}_{\hat{j}}$ の微小変位として異なる軌道上の点 $\mathbf{X}_{\hat{j}}$ を求める。ただし、点 $\mathbf{X}_{\hat{j}}$ は $|j - \hat{j}| > 205$ ステップ (PPG データの平均周期の半分程度) を満たす点 $\mathbf{X}_{\hat{j}}$ の近傍点とした。次に、 \mathbf{X}_j と $\mathbf{X}_{\hat{j}}$ 間の i ステップ経過後の拡大距離 $d_j(i) = \|\mathbf{X}_{j+i} - \mathbf{X}_{\hat{j}+i}\|$ を求める。この

ような $d_j(i)$ を全ての j について求め、 i ステップ経過後の平均拡大距離を $d(i) = \frac{1}{N-i} \sum_{j=1}^{N-i} \|\mathbf{X}_{j+i} - \mathbf{X}_{\hat{j}+i}\|$

とする。一般的に、カオス力学系において近接する軌道は指数関数的な速さで離れていく。すなわち、時系列のサンプリング周期を $T_s (= 1/f_s)$ とすると、 i ステップ経過後、つまり $iT_s [s]$ 後の平均拡大距離は $d(i) \sim e^{\lambda i}$ と表せる。すなわち、 $\ln d(i) \sim \lambda i$ である。したがって、経過時間 $iT_s [s]$ と微小変位の平均拡大距離 $d(i)$ の片対数グラフ上に直線部分が表れたとき、その傾き λ が LLE の推定値である。

4 解析結果

計測した PPG データからトレンドを除去した時系列波形を図 2 に示す。次に、3次元の再構成状態空間に変換した結果を図 3 に示す。遅れ時間 τ は自己相関関数の値が最初に $1/e$ 以下になる時刻としたところ、マスクありのとき $\tau = 24$ 、マスクなしのとき $\tau = 22$ であった。

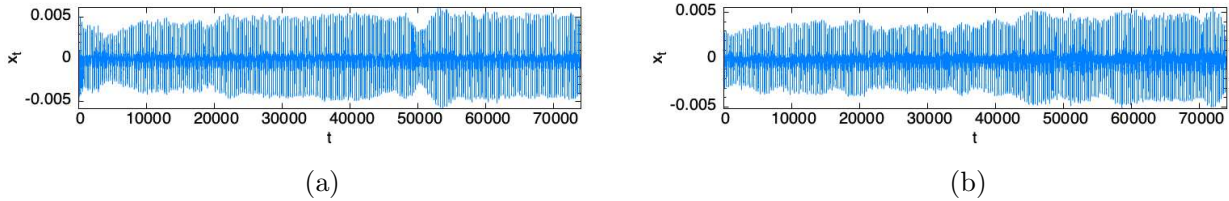


図 2: 計測した PPG データ p_t のトレンドを除去した PPG データ x_t の時系列波形. (a) マスクあり, (b) マスクなし.

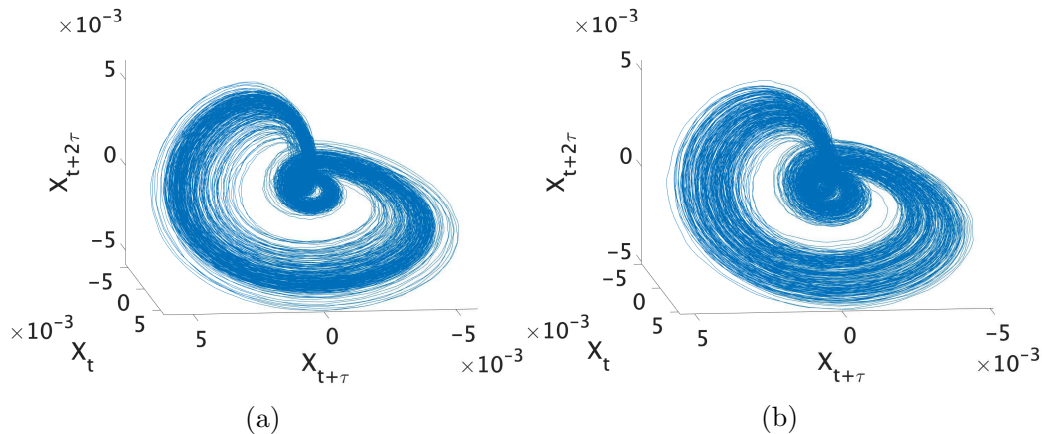


図 3: 再構成次元数 $m = 3$ として状態空間を再構成した結果. (a) マスクあり ($\tau = 24$), (b) マスクなし ($\tau = 22$).

図 3(a) と (b) を比較すると、マスク着用の有無によって 3次元の再構成状態空間に定性的な違いは見られなかった。

次に、FNN 法を適用した結果を図 4 に示す。図 4(a) は横軸が次元 m 、縦軸が誤り近傍率である。図 4(a) よ

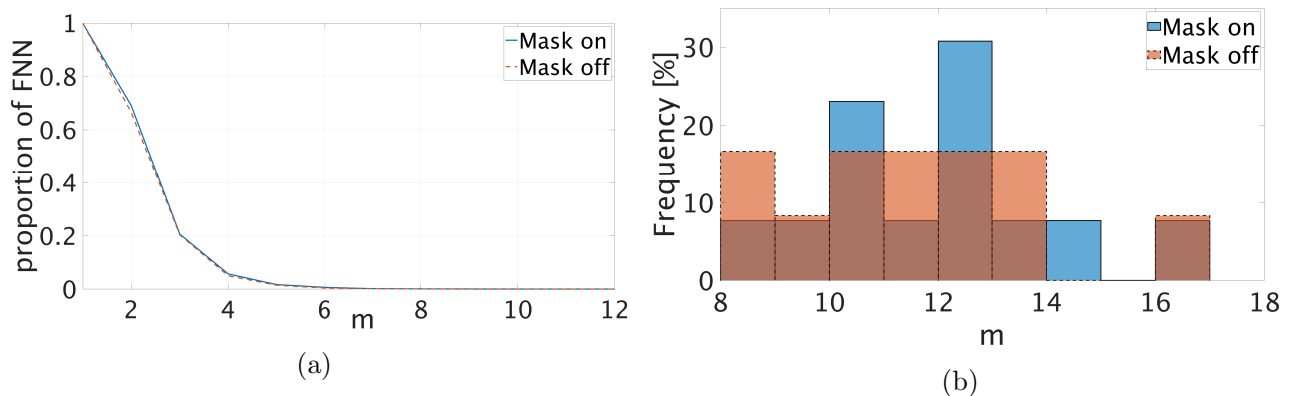


図 4: (a) ある PPG データに FNN 法を適用した結果例. 実線がマスクあり, 破線がマスクなしの結果である. (b) 全 PPG データに FNN を適用して得られた再構成次元 m の度数分布. 実線で囲んだデータがマスクあり, 破線で囲んだデータがマスクなしの結果である。

り、マスクありとマスクなしで誤り近傍率が 0 に収束する次元数はいずれも $m = 11$ であることが分かる。図 4(b) は全 PPG データに対して FNN 法適用することで求めた次元数 m の分布である。実線で囲まれたデータがマスクあり, 破線で囲まれたデータがマスクなしのときの結果である。横軸は再構成次元 m 、縦軸は度数 [%] である。階級幅は 1 とした。データによって再構成次元 m は異なる値をとることが分かる。

続いて、推定された m と τ を用いて状態空間を再構成し、LLE を算出した。図 5(a) に経過時間 iT_s [s] と微小変位の平均拡大距離の自然対数 $\ln d(i)$ の関係を示す。実線でマスクありのときの解析結果と傾き一定の区間における近似直線を描いた。破線でマスクなしのときの解析結果と傾き一定の区間における近似直線を描

いた。 $\ln d(i)$ はアトラクタサイズに達すると収束する。LLE はマスクありのとき 0.7185, マスクなしのとき 0.6716 であった。同様の手順で全 PPG データの LLE を算出し, 各被験者に対するマスクありの場合とマスクなしの場合で平均を取って図示した結果が図 5(b) である。実線に囲まれたデータがマスクあり, 破線に囲まれたデータがマスクなしのときの結果である。横軸は LLE の階級, 縦軸は度数 [%] である。階級幅は 0.1 とした。図 5(b) より, 全 PPG データの LLE が正であり軌道不安定性を有することが分かった。

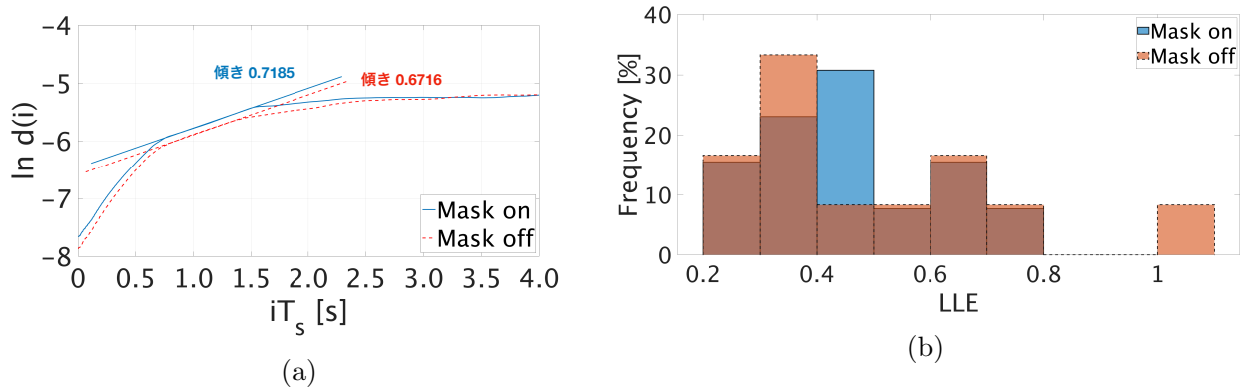


図 5: (a) 経過時間 iT_s [s] と微小変位の平均拡大距離の自然対数 $\ln d(i)$ の関係。実線がマスクあり, 破線がマスクなしのときの結果。(b) 全 PPG データの LLE を算出した結果の度数分布。実線で囲んだデータがマスクあり, 破線で囲んだデータがマスクなし。

最後に, 全 PPG データから求めた再構成次元 m と LLE について, t 検定を行い, マスクの着用の有無によって有意差があるか調査した。いずれも有意水準 $\alpha = 0.05$ で両側検定とした。再構成次元 m については, 帰無仮説 H_0 : 「マスクの着用の有無により m に差はない」, 対立仮説 H_1 : 「マスクの着用の有無により m に差がある」と設定した。LLE については, 帰無仮説 H_0 : 「マスクの着用の有無により LLE に差はない」, 対立仮説 H_1 : 「マスクの着用の有無により LLE に差がある」と設定した。結果, いずれの帰無仮説も棄却されなかった。すなわち, マスクの着用の有無によって PPG データの再構成次元 m と LLE に差が生じるとは言えない。

5 まとめ

本論文では, マスクの着用が身体にどのような影響を及ぼすのかを PPG データに対して非線形時系列解析を適用することで調査した。まず, PPG データを計測し, FNN 法により再構成次元 m を推定し, LLE を算出して PPG データの軌道不安定性を定量的に評価した。その結果, マスクの着用の有無に関わらず PPG データは軌道不安定性を有すること, マスクの着用の有無による LLE に有意差はないことが明らかになった。以上の結果は, マスク着用による身体への影響は存在しない可能性を示唆するものである。今後の課題として, LLE 以外の指標を用いて PPG データを解析することで, マスク着用の有無が身体に与える影響について調査を行う。

参考文献

- [1] 新型コロナウイルス感染症対策本部, https://corona.go.jp/expert-meeting/pdf/kihon_h.20210107.pdf, (Accessed on 01/13/2021).
- [2] J. Allen, *Physiol. Meas.*, Vol. 28, No. 3, pp. R1–R39, 2007.
- [3] T. Sumida *et al.*, *Int. J. Bifurcat. Chaos.*, Vol. 10, pp. 2245–2255, 2000.
- [4] N. Sviridova *et al.*, *Chaos Solitons Fractals*, Vol. 77, pp. 53–63, 2015.
- [5] 東京デバイス株式会社, <https://tokyodevices.com/items/188>, (Accessed on 01/13/2021).
- [6] 東京理科大学における人を対象とする医学系研究に関する倫理規程, 平成 27 年 7 月 16 日施行, https://letus.ed.tus.ac.jp/pluginfile.php/745226/mod_resource/content/1/2020_%E4%BA%BA%E3%82%92%E5%AF%BE%E8%B1%A1%E3%81%A8%E3%81%99%E3%82%8B%E5%8C%BB%E5%AD%A6%E7%B3%BB%E7%A0%94%E7%A9%B6_%E8%AC%9B%E7%BF%92%E4%BC%9A%20%E8%B3%87%E6%96%99.pdf, (Accessed on 01/13/2021).
- [7] N. H. Packard *et al.*, *Phys. Rev. Lett.*, Vol. 45, No. 9, pp. 3712–3716, 1980.
- [8] F. Takens *et al.*, *In Dynamical systems and turbulence*, Vol. 45, No. 9, pp. 366–381, 1981.
- [9] M. B. Kennel *et al.*, *Phys. Rev. A*, Vol. 45, No. 6, pp. 3403–3411, 1992.
- [10] A. M. Albano *et al.*, *Phys. Rev. A*, Vol. 38, No. 6, pp. 3017–3026, 1988.
- [11] M. T. Rosenstein *et al.*, *Physica D*, Vol. 65, Nos. 1–2, pp. 117–134, 1993.

隔離を導入した感染症数理モデルを用いた感染拡大抑制策の定量的評価

真鍋 歩未 (池口 徹 教授, Nina Sviridova 助教)

1 はじめに

2020年, 新型コロナウイルス感染症 (COVID-19) が世界的に流行し, 現在も感染拡大が続いている. COVID-19 は感染力が強く, 自覚症状のない感染者が多く存在する. これにより, 無意識的な感染拡大を引き起こし, 世界的大流行が発生したと考えられている [1]. COVID-19 を含む種々の感染症は, ワクチンが開発される, あるいは治療法が確立されるまでの間, 感染者の隔離によって感染拡大の抑制を試みる. その際, 隔離の効果を定量的に評価することも重要である [2]. しかし, 現実社会において隔離の効果を定量化することは困難である. そこで本論文では, 感染症数理モデルを用いて, 感染者の隔離の有効性を定量的に評価することを目的とする. 具体的には, 周囲の状況によって感染状態の遷移を行い, 感染者の一部をネットワークから隔離させるモデルを提案する. 提案モデルでは, ネットワークの頂点の感染状態がネットワーク構造に依存して変化する. 2種類の複雑ネットワークモデル上で, 提案モデルによる数値実験を行い, 各ネットワーク上での隔離と最終感染率の関係を調査した.

2 感染症数理モデル

2.1 従来モデル

代表的な感染症数理モデルの一つに, Kermack らが提案した, SIR (Susceptible–Infected–Recovered) モデル [3] がある. SIR モデルは, 人々の集団を, 疾病への感染状態によって, S クラス, I クラス, R クラスに分類する. S クラスは未感染者であり, 疾病に感染する可能性がある状態である. I クラスは感染者であり, 疾病に感染している状態である. R クラスは回復者であり, 疾病から回復した状態で, 他者に疾病を感染させず, 自身も再感染することはない状態である. SIR モデルでは, 各クラスの人数の遷移が確率的に行われることで, 感染症の伝播を表現する. 一方, ネットワークの各頂点を人とみなし, 感染状態のクラスを与えるモデルとして, r-SIS (resource based Susceptible–Infected–Susceptible) モデルが提案されている [4]. r-SIS モデルは, 感染確率 β で S クラスから I クラスへ遷移し, 回復確率 μ で I クラスから S クラスへの遷移する. また, 各頂点の周囲の感染状況によって感染確率および回復確率が頂点ごとに変動する [4].

2.2 提案モデル – r-SIAR モデル –

本論文では, 感染者の隔離という対策の効果を定量的に評価するために, 従来モデルの r-SIS モデル [4] の I クラスを2つのクラスに分けた r-SIAR モデルを提案する. 具体的には, r-SIS モデルの I クラスを隔離率 δ で隔離された感染者 (I クラス) と, 確率 $1 - \delta$ で隔離されなかった感染者 (A クラス) とする. r-SIAR モデルでは, 隔離となった I クラスの頂点をネットワークから外すこととする. ここで, 隔離率 δ は, 現実社会において感染者を陽性と確認した確率とも捉えることもできる. また, 感染症の種類によっては, 人間は一度疾病に感染し, 回復することで, 免疫を得ることもある. 免疫を獲得することで再感染せず, 他者に疾病を感染させない状態となる. そこで, 提案モデルでは, このように再感染せず, 他者へ疾病を感染させない回復者のクラス (R クラス) も導入した. ここで, I クラスの頂点 i は隔離されてから時刻 u 経過後に回復者 R クラスに遷移し, 隔離前のネットワーク上での枝を再度つなぎ直す. A クラスの頂点は r-SIS モデル [4] と同様に, ネットワークから隔離されることはなく, 獲得したリソースによる回復確率のもと, 回復者 R クラスに遷移する. r-SIAR モデルの状態遷移図を図1に示す. $x_i(t)$ を頂点 i の時刻 t での状態値とし, 頂点 i が S クラスのとき $x_i(t) = 0$, I クラスのとき $x_i(t) = 1$, A クラスのとき $x_i(t) = 2$, R クラスのとき $x_i(t) = 3$ とする. また, 時刻 t のとき, 隣接行列の (i, j) 成分を $Y_{ij}(t)$ と表す. ここで, 時刻 t で S クラス ($x_i(t) = 0$) となる頂点 i と隣接する A クラス ($x_j(t) = 2$) である頂点 j の集合 E_i^t は, $E_i^t = \{j \mid x_i(t) = 0 \cap Y_{ij}(t) x_j(t) = 2\}$ で表せる. すなわち, $|E_i^t|$ は, 時刻 t で S クラス ($x_i(t) = 0$) となる頂点 i と隣接する A クラス

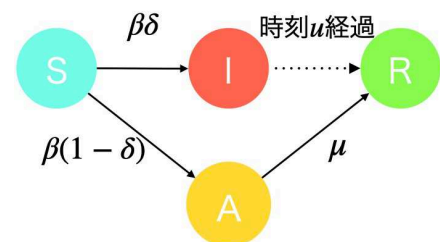


図 1: r-SIAR モデルの状態遷移図

$(x_j(t) = 2)$ である頂点 j の個数を表す. また, 感染確率 β は, 隣接する A クラスとの一回の接触によって S クラスの頂点が I または A クラスに遷移する確率を示す. よって, 時刻 t において, S クラスの頂点 $i(x_i(t) = 0)$ が, 時刻 $t+1$ において I クラス ($x_i(t+1) = 1$) または A クラス ($x_i(t+1) = 2$) に遷移する確率は $1 - (1 - \beta)^{|E_i^t|}$ となる. 以上より, ネットワーク上の頂点 i の状態遷移のダイナミクスは式 (1) となる.

$$x_i(t+1) = \begin{cases} x_i(t) + G(1 - (1 - \beta)^{|E_i^t|})(G(1 - \delta) + 1) & (x_i(t) = 0) \\ x_i(t) + 2H(t - T_i - u) & (x_i(t) = 1) \\ x_i(t) + G(\mu_i(t)) & (x_i(t) = 2) \\ x_i(t) & (x_i(t) = 3) \end{cases} \quad (1)$$

ただし, 関数 $G(p)$ は確率 p で 1, 確率 $1-p$ で 0 をとる関数, 関数 $H(x)$ はステップ関数である. 一方, 現実的には, 回復確率は単純な確率に加えて治療や薬などの資源 (リソース) にも依存すると考えられる. そこで, r-SIAR モデルでは, r-SIS モデル [4] に倣い, 回復確率 μ を分配されるリソースの量に依存させている. 時刻 t での頂点 i の回復確率 $\mu_i(t)$ を $\mu_i(t) = \mu_0 + \mu_r r_i(t)/k_i$ と定義する. ここで, μ_0 は自己治癒確率を表す. ただし, 本論文では文献 [4] と同様に, $\mu_0 = 0$ とした. k_i は頂点 i の次数, μ_r は得たリソースの使用率, $r_i(t)$ は時刻 t のとき, 頂点 i が隣接する S クラスの頂点から得るリソース量を表す. つまり, 頂点 i の回復確率 $\mu_i(t)$ は得られるリソース量 $r_i(t)$ に比例する. また, リソースは S クラスおよび R クラスの頂点から A クラスの頂点へ配分される. また, 時刻 t のとき, A クラスの頂点 i が持つ S クラスまたは R クラスの頂点と接続する枝の本数を $m_i(t)$ とする. 時刻 t で頂点 i が, 隣接する S クラスまたは R クラスの頂点 j から得られるリソース量を $r'_{j \rightarrow i}(t)$ として, $r'_{j \rightarrow i}(t) = (m_i(t) + 1)^\alpha / \sum_{q \in \Theta_j^t} (m_q(t) + 1)^\alpha$

と定義する. ここで, 集合 Θ_j^t は, $\Theta_j^t = \{i \mid x_j(t) = 0 \cup x_j(t) = 3 \cap Y_{ij}(t)x_i(t) = 2\}$ で表し, 時刻 t で S または R クラスの頂点 j と隣接する A クラスの頂点 i の集合, α はパラメータである. $\alpha > 0$ のとき, 周囲に S クラスおよび R クラスの多い頂点に対して優先的にリソース配分し, $\alpha < 0$ のとき, 周囲に S クラスおよび R クラスの少ない頂点に対して優先的にリソース配分することを意味する. また, $\alpha = 0$ のとき, 各頂点には均等にリソースを配分することになる. ここで, 集合 Ω_i^t を $\Omega_i^t = \{j \mid x_i(t) = 2 \cap (Y_{ij}(t)(x_j(t) + 1) = 1 \cup Y_{ij}(t)x_j(t) = 3)\}$ と表し, 時刻 t で A クラスの頂点 i と隣接する S クラス, R クラスの頂点 j の集合とすると, 時刻 t で頂点 i が配分されるリソース量 $r_i(t)$ は $r'_{j \rightarrow i}(t)$ を用いて, $r_i(t) = \sum_{j \in \Omega_i^t} r'_{j \rightarrow i}(t)$ と定義する.

3 ネットワークモデル

3.1 Uncorrelated Configuration Model

Uncorrelated Configuration Model (以下, UCM)[5] は, スケールフリー性を持つネットワークモデルである. ここで, スケールフリー性とは, ネットワークの次数分布がべき則に従うことを指す. UCM は, 任意のべき指数 γ を与え, 以下の手順で作成される.

1. 頂点数 N , 最小次数 k_{\min} , べき指数 γ を指定する.
2. 次数分布に従い, 最小次数が k_{\min} , 最大次数が \sqrt{N} となるような次数順列を作成し, 各頂点に割り当てる.
3. 各頂点に割り当てられた次数から, 高次数の頂点から順に, 他の頂点と枝を接続する. ただし, 自己ループ, 多重辺は作成しない.

3.2 Watts-Strogatz Model

Watts-Strogatz モデル (以下, WSM) は, スモールワールド性を有するネットワークモデルである [6]. ここで, スモールワールド性とは, 現実世界では「友人の友人も自分の友人」であるような世間の狭さを表す性質である. 複雑ネットワーク理論上では, ネットワークのクラスター係数が高く, 同時に平均頂点間距離が小さいときにスモールワールド性を有すると定義される. WSM は, 以下の手順で作成される.

1. 頂点数 N , 平均次数 \bar{k} の規則的な円環状格子を作成する.
2. 存在する枝全てについて, 確率 p でつなぎ変える. ただし, 枝をつなぎ変える際は枝の両端の頂点のうち, 一方の頂点を等確率で切り離し, ランダムに選んだ頂点に枝を接続する. また, 自己ループ, 多重辺は作成しない.

この手順で作成されたWSMは、 $p = 0$ のとき、規則的な円環状格子ネットワーク、 $p = 0.01$ 前後ではスモールワールドネットワーク、 $p = 1$ のとき、全ての枝がつなぎ変えられたランダムネットワークとなる。

4 数値実験

感染者隔離の定量的評価を行うために、提案モデル(r-SIARモデル)とr-SISモデルを用いて、感染確率 β と最終感染率 ρ の関係を調査した。まず、最終時刻 $t'(=200)$ における $V = \{i \mid x_i(t') > 0\}$ とすると、 V はr-SISモデルではIクラスの頂点集合、r-SIARモデルではI、A、Rクラスの頂点集合を表す。これより、最終感染率 ρ を $\rho = |V|/N$ と定義する。感染伝播を行うネットワークとして、頂点数 $N = 5000$ 、最小次数 $k_{\min} = 3$ 、次数分布が $\gamma = 2.28$ のべき則に従うUCMと、頂点数 $N = 5000$ 、平均次数 $k = 8$ 、枝つなぎ変え確率 $p = 0.01$ のWSMを使用した。実験開始時において、全頂点の1%の頂点を感染源としてランダムに配置した。また、r-SIARモデルでは、感染源の頂点に対し、確率 δ でIクラス、確率 $1 - \delta$ でAクラスとした。隔離後回復に要する時間を $u = 14$ 、リソースの使用率を $\mu_r = 0.60$ とし、リソース戦略パラメータ $\alpha = -2.0, 0, 2.0$ とした。また、厚生労働省の新型コロナウイルス感染症の国内発生動向 [7] より、7月15日までの東京都の陽性患者のうち、有症状者の割合を算出し、隔離率を $\delta = 0.39$ と固定した。まず、r-SIARモデルのリソース戦略の違いによる感染確率 β と、最終感染率 ρ の関係を調査した結果を図2に示す。

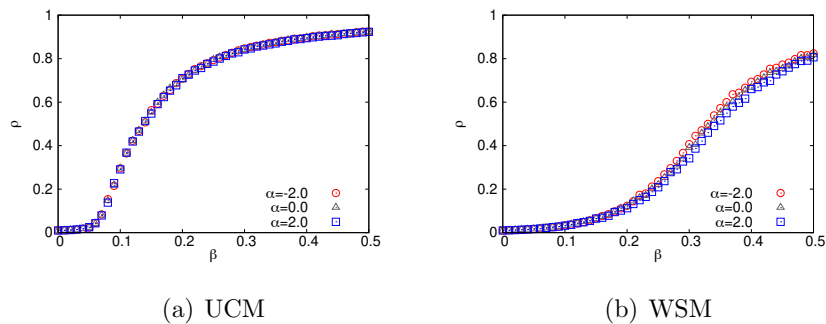


図 2: リソース戦略の違いによる感染確率 β と最終感染率 ρ の関係

図2から、リソース戦略パラメータ α の値によらず、ネットワークから約40% ($\delta = 0.39$)の感染者を隔離することで、疾病の感染力が弱い際は、UCMに比べ、WSMにおいてネットワーク全体への疾病の拡散を抑制できていることがわかる。しかし、疾病の感染力が強い際は、UCM, WSM共に、ネットワーク全体に疾病が拡散する。また、各リソース戦略の違いが最終感染率に影響を与えていないことから、これ以降は $\alpha = 0$ に固定して数値実験を行った。

次に、再感染が起きないことと感染者隔離の効果とをそれぞれ定量的に評価するために、r-SISモデルと、 $\delta = 0, 0.39$ としたr-SIARモデルについて、感染確率 β と最終感染率 ρ の関係を図3に示す。

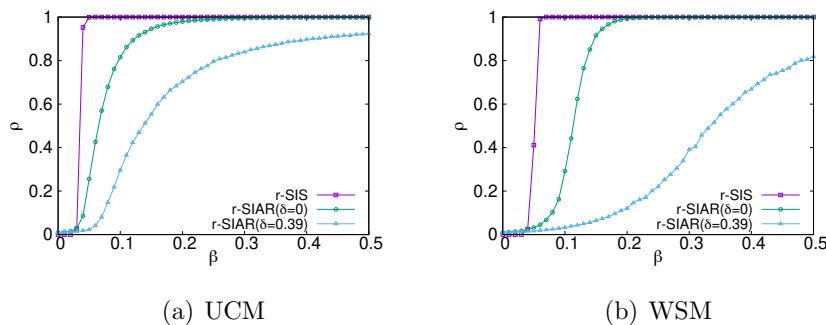


図 3: r-SISモデル、r-SIARモデル($\delta = 0, 0.39$)の感染確率 β と最終感染率 ρ の関係の比較

図3(a)を見ると、UCMにおいては、約40%の感染者の隔離という対策を行っても、 $0.1 < \beta \leq 0.2$ では約30%程度しか最終感染率を抑制できていないことがわかる。一方、図3(b)より、WSMでは $0 < \beta < 0.3$ において、感染者を隔離しないときに比べ約50%以上の最終感染率 ρ を抑制できていることがわかる。

最後に、r-SIARモデルについて、隔離率 δ と感染確率 β を変化させた場合に、最終感染率 ρ

がどのように変化するかを調査した．結果を図4に示す．図4では，最終感染率 ρ ($0 \leq \rho \leq 1$) をカラーマップで示している．

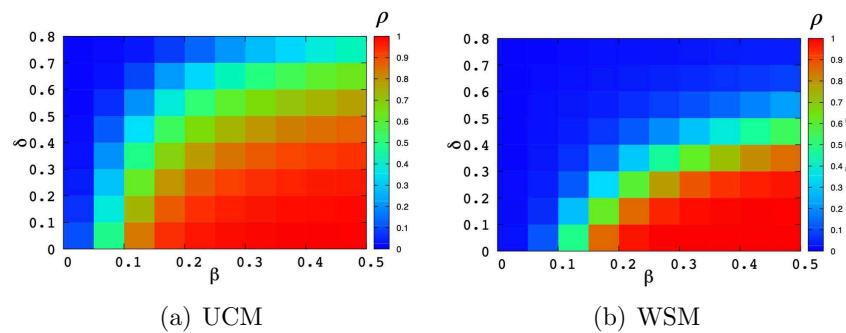


図4: 感染確率 β と隔離率 δ ，最終感染率 ρ の関係

図4 (a), (b) から，どちらのネットワークにおいても δ の値が大きいくほど，どの β に対しても感染を抑制できていることがわかる．しかし，図4 (a)を見ると，UCMにおいては， $\beta = 0.5$ のとき，6割以上の感染者を隔離しないと，最終感染率は7割を超える．すなわち，感染力の強い疾病においては，感染者を半数以上隔離しても最終感染率は高いままであり，7～8割以上の感染者を隔離しなければ，感染の拡大を抑制することはできない．一方，図4(b)を見ると，WSMでは，UCMに比べて全体的に青色が占めている割合が高く，感染者の隔離が感染症の拡大抑制に効果的であることを示している．しかし，WSMにおいても， $\beta = 0.5$ のとき半数以上の感染者を隔離しなければ，最終感染率は6割を超える．これらの結果から，感染者を隔離する対策は，ネットワーク構造と，感染力の強さに依存して効果的となる場合もあるが，感染力が強い疾病 ($\beta \geq 0.4$) の場合，感染者の大多数を隔離しなければ，感染拡大の抑制に効果はない．

5 まとめ

本論文では，感染者の隔離がネットワーク上での感染拡大抑制にどの程度有効な対策であるかを定量的に評価した．感染症数理モデルの一つである r-SIS モデルに，隔離を考慮した r-SIAR モデルを提案した．また，r-SIS モデルと r-SIAR モデルを用いて，感染確率と隔離率を変化させたときの最終感染率を評価した．対象とするネットワークは UCM と WSM とした．数値実験による検証の結果，感染力が高い感染症の場合は，感染者を5割以上隔離しなければ，7割程度拡散してしまうことがわかった．この結果は，ネットワーク構造，感染力の強さによっては隔離という対策は効果的な戦略とならないことを示すものである．

参考文献

- [1] Yan Bai, Lingsheng Yao, Tao Wei, Fei Tian, Dong-Yan Jin, Lijuan Chen and Meiyun Wang, “Presumed Asymptomatic Carrier Transmission of COVID-19,” *JAMA*, Vol. 323, No. 14, pp.1406–1407, 2020.
- [2] WHO, “Clinical Management of COVID-19,” <https://www.who.int/publications/i/item/clinical-management-of-covid-19> (最終閲覧日：2020/12/18)
- [3] Kermack, William Ogilvy and McKendrick, Anderson G, “A Contribution to the Mathematical Theory of Epidemics,” *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, Vol. 115, No. 772, pp. 700–721, 1927.
- [4] Xiao-Long Chen, Shi-Min Cai, Ming Tang, Wei Wang, Tao Zhou and Pak-Ming Hui, “Controlling Epidemic Outbreak Based on Local Dynamic Infectiousness on Complex Networks,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, Vol. 28, No. 12, p. 123105, 2018.
- [5] Michele Catanzaro, Marián Boguná and Romualdo Pastor-Satorras, “Generation of Uncorrelated Random Scale-free Networks,” *Physical Review E*, Vol. 71, No. 2, p. 027103, 2005.
- [6] Watts, Duncan J and Strogatz, Steven H, “Collective Dynamics of ‘small-world’ Networks,” *Nature*, Vol. 393, No. 6684, pp. 440–442, 1998.
- [7] 厚生労働省, “新型コロナウイルス感染症の国内発生動向,” <https://www.mhlw.go.jp/content/10906000/000649533.pdf> (最終閲覧日：2020/07/23)

発表文献

- [P1] 真鍋歩未, 澤田和弥, 宮南風, 池口徹, “感染症数理モデルを用いた感染者隔離効果の定量的評価,” 電子情報通信学会 技術研究報告, Vol. 120, No. 214, pp. 15–20, 2020年10月.

複雑ネットワーク理論を用いた文章構造の時代的变化の解析

相原里江 (池口徹 教授 , Nina Sviridova 助教)

1 はじめに

現在では多くのテキストデータが電子的に配信され、インターネットから容易に手に入れられるようになってきている。それに伴って、文章をネットワークに変換することで文章構造の特徴について解析する先行研究は多くなされている。例えば、Canchoらは文章を単語の共起に基づいてネットワークを作成し解析を行っていたり [1]、島田らの研究では様々な小説をネットワーク化し、その特徴の定量的な評価を行っていた [2]。しかし、これらは時代的な観点から解析を行っていなかったり、歴史の中の二点間での比較に留まっており、時代変遷に伴った文章構造や言語ネットワークの特徴の変化についての議論は不十分である。そこで本論文では、解析対象を日本語の新聞記事とし、戦前と戦後の新聞記事をネットワークに変換することで、複雑ネットワーク理論の観点から時代変遷による文章構造の変化を調査した。具体的には、戦前と戦後の新聞記事を対象に作成した言語ネットワークに対して、ネットワークの特徴量であるクラスタ係数 [3]、平均頂点間距離 [3]、次数を用いて調査した。これに加えて、出現頻度が最も高い単語の各ネットワーク特徴量に対して判別分析を行った。その結果、時代変遷に伴う文章構造の変化の傾向は見られなかったが、戦前と戦後による文章を分類できることが明らかになった。

2 使用データ

本論文では解析対象として、神戸大学経済経営研究所の新聞記事文庫 [4]、朝日新聞社の聞蔵Ⅱ [5] から異なる時代の新聞記事のテキストデータをダウンロードして使用した。その際、各年代の記事が豊富に存在すること、書き手による文章の差が小さいこと、時代背景を反映していることなどの特徴を持つ戦前の東京朝日新聞および戦後の朝日新聞を使用した。具体的には、1912年から1942年の東京朝日新聞と、1990年、2000年、2010年、2019年の朝日新聞である。

3 ネットワークの作成手法

本論文では単語の共起に注目し、文章を重み無し無向ネットワークに変換した。具体的には、まず形態素解析ツールであるmecab [6] を用いて、文章を単語の原型に分解し、その品詞を得る。次に得られた単語の原型と品詞を一つの頂点として扱い、文章中で隣接する場合に枝を付与することで重み無し無向ネットワークを作成した。その際単語の重複は認めず、句読点等の記号は削除し単語のみをネットワークの頂点とした。また、一つの文章から複数のネットワークが得られた場合は、得られたネットワークのうち最大のもののみを使用することとした。具体例を図1に示す。

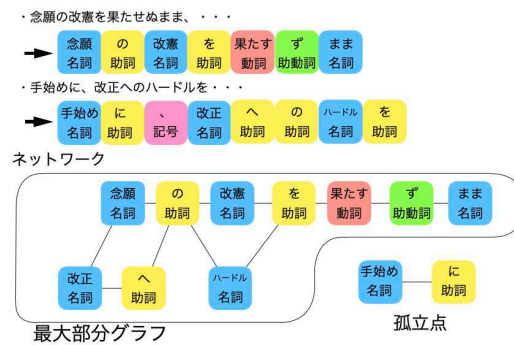


図 1: ネットワーク作成例

4 実験手順

複雑ネットワークの観点から時代変遷に伴う文章構造の変化を調査するため、時代ごとに文章のネットワーク構造を比較した。具体的には、新聞記事を言語ネットワークに変換し、得られたネットワークの特徴量である次数分布、平均頂点間距離、クラスタ係数を計算した。その際、得られたネットワークをランダム化したネットワークのクラスタ係数と平均頂点間距離を用いて正規化処理を行った。異なる時代の新聞記事から得られたネットワークの特徴量の比較をすることで、時代変遷に伴う文章構造の変化の傾向について調査

した。次に文章中で使用されている単語の中で頻出の単語を調査した。頻出の単語について、同様にクラスタ係数、次数、平均頂点間距離を計算し、それぞれランダム化したネットワークの指標を用いて正規化処理を行った。

5 結果

戦前の1912年から1942年、戦後の1990年、2000年、2010年、2019年の新聞記事から得られた言語ネットワークの特徴量を図2に示す。ここで横軸は年、縦軸はネットワーク特徴量を正規化した値となっている。

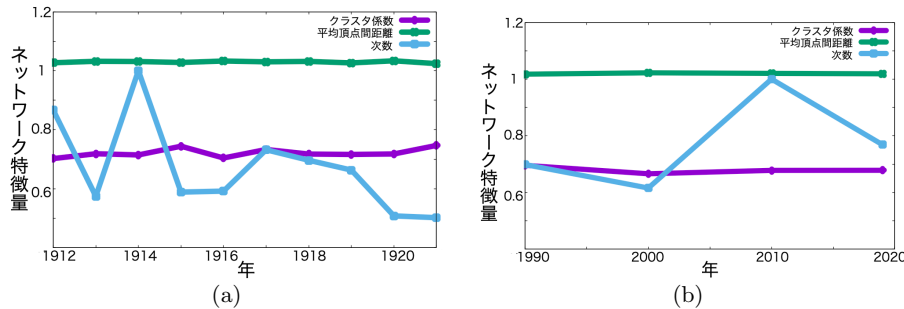


図 2: (a) 戦前と (b) 戦後の新聞記事のネットワーク特徴量

図2から、時代変遷に伴う値の変化傾向が見られない結果となった。原因として、一つの記事に対して一つのネットワークを作成しているため、文章量が多くなくネットワークの規模が小さいこと、記事ごとの文章量に差があることが考えられる。それらを改善するために、単語単位で特徴量の比較を行った。具体的に扱う単語として、時代に関係無くどの記事にも普遍的に使用されていることが望ましいため、単語の使用頻度についてまず調査した。戦前の1912年から1921年、戦後の1990年、2000年、2010年、2019年の新聞記事について、多くの記事で登場している上位10個の動詞と、それらがそれぞれ記事内で使用された頻度を調査した結果を図3、図4に示す。ここで図3(a)、図4(a)の横軸は年、縦軸は対応する単語が登場した記事数その年の全ての記事数に占める割合である。また、図3(b)、図4(b)の横軸は年、縦軸は対応する動詞がその年の記事内の全ての動詞に占める割合である。

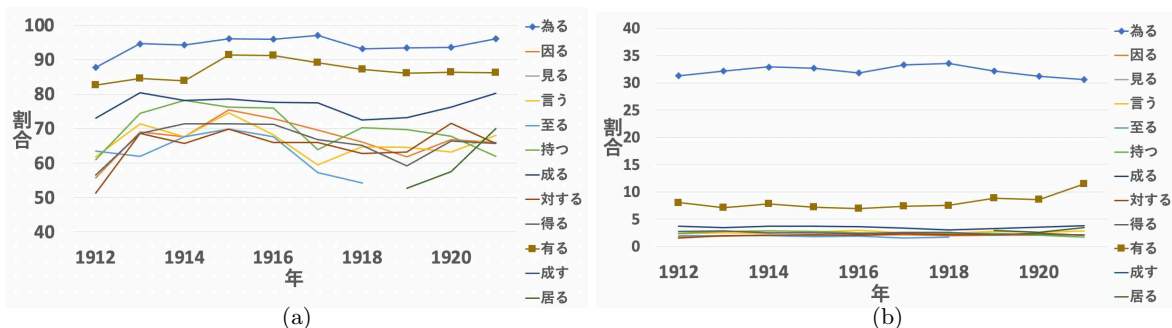


図 3: 戦前の記事に対して (a) 各動詞の出現した記事の割合と (b) 各動詞の記事内の割合を調査した結果

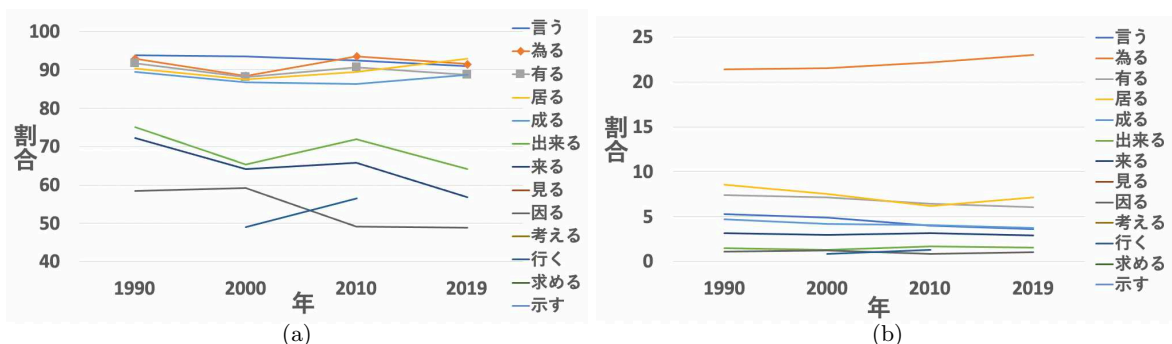


図 4: 戦後の記事に対して (a) 各動詞の出現した記事の割合と (b) 各動詞の記事内の割合を調査した結果

図3、図4を見ると動詞においては「為る」、「有る」が最もよく使用されていることがわかる。そのため、「為る」と「有る」のネットワーク特徴量であるクラスタ係数、平均頂点間距離、次数を調査・比較した。その際、特徴量の尺度を等しくするため、それぞれ平均値を引き標準偏差で割ることで、それぞれの分布と範囲を揃えるように正規化を行った。「為る」の結果を図5、「有る」の結果を図6に示す。ここで横軸は年、縦軸はそれぞれのネットワーク特徴量の値である。

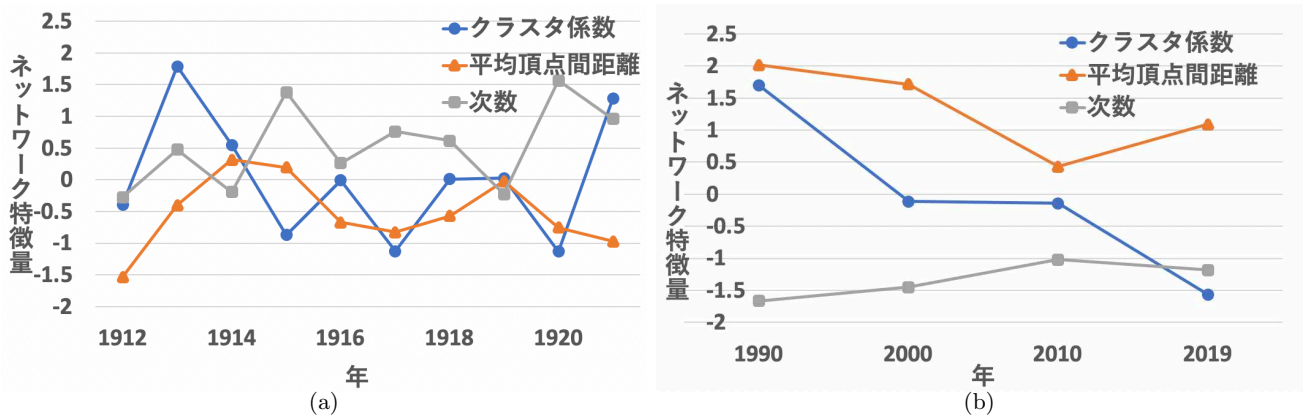


図 5: (a) 戦前と (b) 戦後の「為る」のネットワーク特徴量

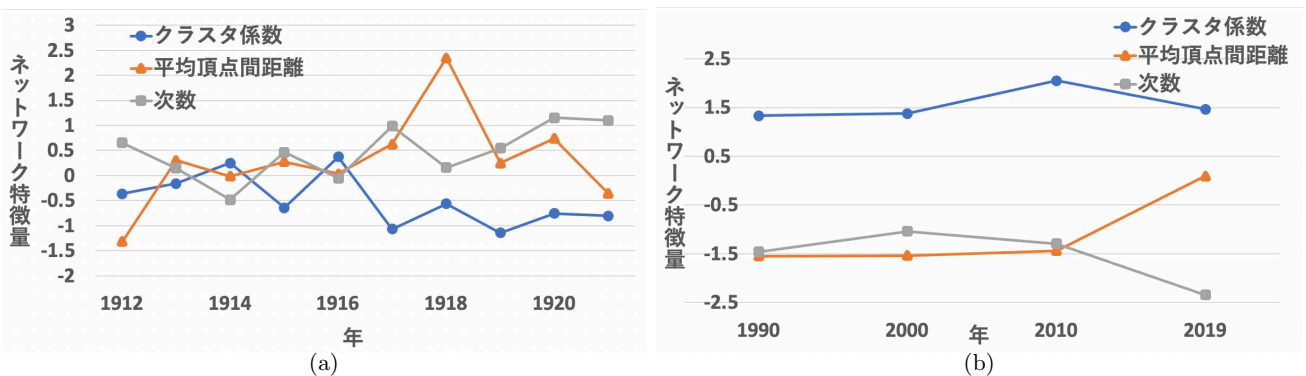


図 6: (a) 戦前と (b) 戦後の「有る」のネットワーク特徴量

更にこれらの結果に対し、戦前と戦後のプロットの重心の垂直二等分線を引くことで判別分析を行った。クラスタ係数と平均頂点間距離の関係を図 7, クラスタ係数と次数の関係を図 8, 平均頂点間距離と次数の関係を図 9 に示す。ここで、数字はデータの年を表す。また、青のプロットが戦前、オレンジのプロットが戦後、グレーのプロットが戦前の結果の重心、黄色のプロットが戦後の結果の重心、赤の直線が二つの重心の垂直二等分線である。

図 7 から図 9 の結果を見ると、判別分析により戦前と戦後で分類されており、ネットワーク構造に差があることがわかる。図 8, 図 9 を見ると、「為る」, 「有る」共に、次数は戦前に比べ戦後の方が小さい。図 8(b) では戦後の方がクラスタ係数は大きい値をとっており、次数と同様にネットワーク構造に影響しているが、図 8(a) ではクラスタ係数がネットワーク構造にほとんど影響していないことがわかる。

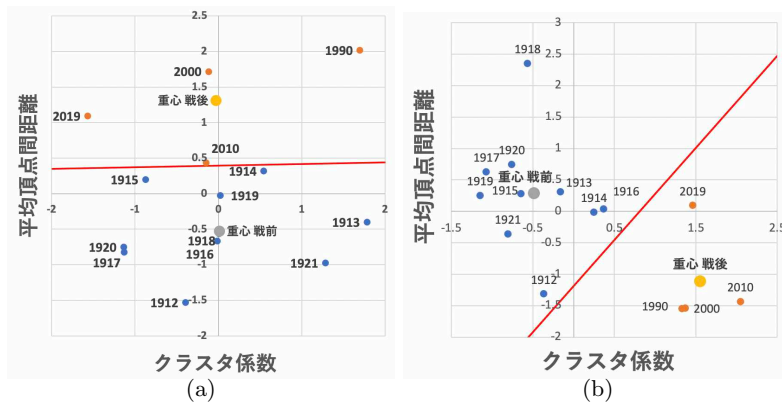


図 7: (a) 「為る」と (b) 「有る」のクラスタ係数と平均頂点間距離の関係

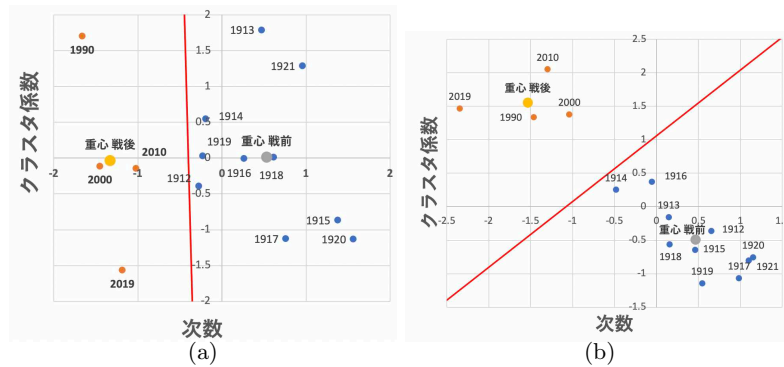


図 8: (a)「為る」と (b)「有る」のクラスタ係数と次数の関係

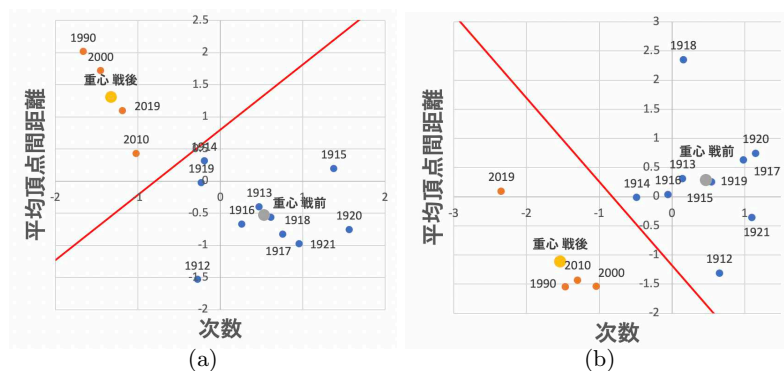


図 9: (a)「為る」と (b)「有る」の平均頂点間距離と次数の関係

図 9 を見ると、平均頂点間距離と次数どちらも同様にネットワーク構造に影響しているが、図 9(a) では戦後の方が平均頂点間距離が大きい値をとっている一方で、図 9(b) では戦前の方が平均頂点間距離が大きい値となっており、二つの動詞で共通する特徴は見られない。結果として、図 8、図 9 から次数が戦前と戦後の新聞記事の特徴を最も反映できていることが分かった。

6 まとめ

本論文では、新聞記事を対象に言語ネットワークを作成し、複雑ネットワークの観点から見た時代間での日本語構造の差を調査した。その結果、文章全体でのネットワーク構造においては傾向がよく見られなかったが、文章中でよく使用されている単語に着目し、その特徴量を比較することで戦前と戦後の新聞記事のネットワークを分類することができた。更に、単語の次数は戦前と戦後の新聞記事の特徴を反映できることが示唆された。

参考文献

- [1] R. F. Cancho, R. V. Solé, “The small world of human language”, *Proceedings of the Royal Society of London B: Biological Sciences*, 268, 1482, 2261–2265, 2001
- [2] Y. Shimada, M. Tatara, K. Fujiwara and T. Ikeguchi, “Formation mechanisms of local structures in language networks”, *EPL*, Vol. 127, No. 5, 2019.
- [3] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks”, *Nature*, Vol. 393, No. 6684, pp. 440–442, 1998.
- [4] 神戸大学経済経営研究所, 新聞記事文庫, <http://www.lib.kobe-u.ac.jp/sinbun/> (2020-09-25 閲覧)
- [5] 朝日新聞社, 聞蔵 II, <http://database.asahi.com/library2/main/top.php> (2020-09-25 閲覧)
- [6] 形態素解析ツール Mecab, <https://taku910.github.io/mecab/> (2020-09-25 閲覧)

ブートストラップリサンプリング法によるリアプノフ指数の推定

大矢早紀 (池口徹 教授, Nina Sviridova 助教)

1 はじめに

現実世界には様々な時系列データが存在しており, これらの多くは複雑な非線形な動きをしているため数値予測を行うことが困難である. 特にカオス性を持つ時系列データは, 初期値の微小な変位が指数関数的に拡大するという初期値鋭敏性を持つため, 長期的な推定は困難である. しかし, 決定論的ダイナミクスを有するため, 良いモデルを作成することができれば, 短期的な推定は可能である. 予測手法の一つに, ヤコビ行列推定法 [1] がある. ヤコビ行列推定法は, 微小な変位間の遷移を定める, ヤコビ行列を推定することにより予測を実現する. ヤコビ行列推定法による推定は, 十分に望ましい近傍点を参照すれば精度の高い推定を行うことができるが, 時系列データにノイズが入ると, 適切でない偽近傍点を参照してしまうため, 推定精度が下がってしまう. そこで原木ら [2] はブートストラップリサンプリング法 (以下, ブートストラップ法) を用いた推定手法を提案し, 池田写像 [4] に対して数値予測を行った. ブートストラップ法では, ヤコビ行列を一意に求めるのではなく, 近傍点を複数回参照し直すことによって, ノイズの影響による推定精度の減少を抑えることができる. 本稿では, ヤコビ行列推定法とブートストラップ法をリアプノフ指数の推定に適用し, その性能を比較する. その際, エノン写像 [3] と池田写像を用いて比較を行ったところ, 池田写像において, 二つの手法によるリアプノフ指数の推定に差が見られた.

2 非線形力学系とヤコビ行列

ある非線形写像があるとき, 時間が経過するごとにその写像がどのような値のヤコビ行列をとるのか, 理論的に求めることができる. $\mathbf{x}(t)$ を \mathbf{R}^k での時間 t における状態とし, \mathbf{f} は k 次元の非線形写像としたとき, 式 (1) の力学系を考える.

$$\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t)) \quad (1)$$

ここで, $\mathbf{x}(t)$ における微小な変位を $\Delta\mathbf{x}(t)$ とすると,

$$\mathbf{x}(t+1) + \Delta\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t) + \Delta\mathbf{x}(t)) \quad (2)$$

を得る. ここで, 式 (2) の右辺をテイラー展開をすると, ヤコビ行列 \mathbf{J} を用いて, 式 (3) のように近似することができる.

$$\mathbf{f}(\mathbf{x}(t) + \Delta\mathbf{x}(t)) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{J}(\mathbf{x}(t))\Delta\mathbf{x}(t) \quad (3)$$

式 (3) を式 (2) に代入すると, 式 (4) となる.

$$\mathbf{x}(t+1) + \Delta\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{J}(\mathbf{x}(t))\Delta\mathbf{x}(t) \quad (4)$$

ここで式 (1) を用いると, 状態値 $\mathbf{x}(t)$ における \mathbf{f} のヤコビ行列を $\mathbf{J}(t)$ として, 式 (4) は式 (5) に変換できる.

$$\Delta\mathbf{x}(t+1) = \mathbf{J}(t)\Delta\mathbf{x}(t) \quad (5)$$

よって, \mathbf{f} の第 i 成分を f_i , $\mathbf{x}(t)$ の第 j 成分を x_j をすると, ヤコビ行列 $\mathbf{J}(t)$ は, 式 (6) で表すことができる.

$$\mathbf{J}(t) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_k} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \frac{\partial f_k}{\partial x_2} & \cdots & \frac{\partial f_k}{\partial x_k} \end{pmatrix} \quad (6)$$

3 リアプノフ指数の導出方法

式 (1) の力学系を N 回反復するとき、ヤコビ行列 $\mathbf{J}(t)$ の N 回積の行列を $\mathbf{M}(\mathbf{x}(0), N) = \prod_{t=0}^{N-1} \mathbf{J}(t)$ と定義する。また、 $\mathbf{J}(t)$ の正定値行列を $\{\mathbf{M}(\mathbf{x}(0), N)\}^\dagger \{\mathbf{M}(\mathbf{x}(0), N)\}$ とし、その固有値を $\sigma_i(N)$ としたとき、リアプノフ指数は式 (7) で定義される。

$$\lambda_i = \lim_{N \rightarrow \infty} \frac{1}{N} \log \sigma_i(N) \quad (7)$$

ここで $i = 1, 2, \dots, m$ であり、 m は時系列データの次元である。しかし、この定義では、計算するのが困難である。本来、初期値の微小な変位の伸びの度合いを指しているリアプノフ指数が、負の値をとるとき、縮まってしまふからである。そのため、ヤコビ行列の QR 分解を行うことが必要となる。 $\mathbf{Q}(t)$ を直交行列、 $\mathbf{R}(t)$ を上三角行列とすると、 $\mathbf{J}(t)\mathbf{Q}(t) = \mathbf{Q}(t+1)\mathbf{R}(t+1)$ と分解できる。よって、式 (8) を得る。

$$\{\mathbf{M}(\mathbf{x}(0), N)\}^\dagger \{\mathbf{M}(\mathbf{x}(0), N)\} = \mathbf{Q}(2N) \prod_{k=1}^{2N} \mathbf{R}(k) \quad (8)$$

ただし、 $k \geq N$ のとき、 $\mathbf{J}^\dagger(N - (k+1))\mathbf{Q}(k) = \mathbf{Q}(k+1)\mathbf{R}(k+1)$ とする。式 (8) より、リアプノフ指数は式 (9) で求められる。

$$\lambda_i = \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{k=1}^{2N} \log |\mathbf{R}_{ii}(k)| \quad (9)$$

ただし、 $\mathbf{R}_{ii}(k)$ は、行列 $\mathbf{R}(k)$ の第 i 対角要素である。式 (9) を用いて、理論値のリアプノフ指数を求めることができる。本論文で用いるエノン写像や池田写像は 2 次元の写像であるため、 λ_1 と λ_2 の二つの値が導出されることになる。一般的に負のリアプノフ指数 λ_2 の推定は困難であり、最大リアプノフ指数 λ_1 の値が重要である。

4 推定手法

4.1 ヤコビ行列推定法 [1]

ヤコビ行列推定法によるヤコビ行列の推定の手順を以下に示す。アトラクタ上の一点 $\mathbf{x}(t)$ を考える。 $\mathbf{x}(k_i)$ は $\mathbf{x}(t)$ の i 番目の近傍点とする。ここで、 $i = 1, 2, \dots, M$ であり、 M を近傍点数とする。本稿では、アトラクタ上の 2 点間の最長距離を 1 としたときの割合 r を用いて、 $\mathbf{x}(t)$ を中心とした近傍半径 r 内に存在する点を、 $\mathbf{x}(t)$ の近傍点とみなす。変位ベクトル \mathbf{y}_i 、 \mathbf{z}_i をそれぞれ $\mathbf{y}_i = \mathbf{x}(k_i) - \mathbf{x}(t)$ 、 $\mathbf{z}_i = \mathbf{x}(k_i+1) - \mathbf{x}(t+1)$ とする。これらの変数の関係を、図 1 に示した。 \mathbf{W} を \mathbf{y}_i の分散行列、 \mathbf{C} を \mathbf{z}_i の共分散行列ヤコビ行列とすると、 $\mathbf{J}(t)$ は式 (10) で定義される。

$$\mathbf{J}(t) = \mathbf{C}\mathbf{W}^{-1} \quad (10)$$

\mathbf{W} 、 \mathbf{C} は $\mathbf{W}_{kl} = \frac{1}{M} \sum_{i=1}^M \mathbf{y}_{ik} \mathbf{y}_{il}$ 、 $\mathbf{C}_{kl} = \frac{1}{M} \sum_{i=1}^M \mathbf{z}_{ik} \mathbf{z}_{il}$ となる。ただし、 \mathbf{y}_{ik} 、 \mathbf{z}_{ik} は \mathbf{x}_i 、 \mathbf{y}_i の第 k 成分、 \mathbf{W}_{kl} 、 \mathbf{C}_{kl} は行列 \mathbf{W} 、 \mathbf{C} の (k, l) 成分である。

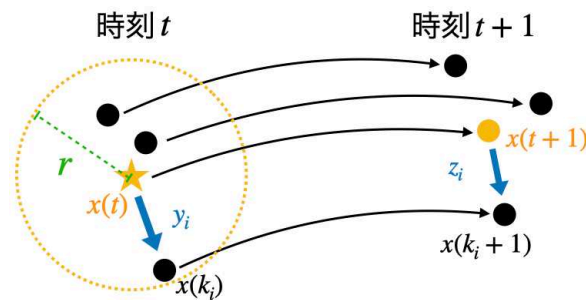


図 1: 時刻 t における点 $\mathbf{x}(t)$ とその近傍点の推移の関係

4.2 ブートストラップ法 [2]

ブートストラップ法は、近傍点を複数回選び直してヤコビ行列を求めることによって、ノイズの影響による推定精度の減少を抑える手法である。ブートストラップ法によるヤコビ行列の推定の手順を以下に示す。

1. アトラクタ上の一点 $\mathbf{x}(t)$ の近傍点を探索する。
2. どの近傍点を用いてヤコビ行列を推定するかをランダムにリサンプリングする。第 $b(b = 1, 2, \dots, B)$ 回目にリサンプリングした近傍により推定したヤコビ行列を $\mathbf{J}_b(t)$ とする。なお、本稿では $B = 200$ とする。
3. B 個のヤコビ行列の平均をとる。つまり、時刻 t におけるブートストラップ法によるヤコビ行列を、 $\mathbf{J}(t) = \frac{1}{B} \sum_{b=1}^B \mathbf{J}_b(t)$ とする。

5 観測ノイズ

本論文では、信号対雑音比 (以下, SN 比) により大きさを定めた観測ノイズを、エノン写像 [3] と池田写像 [4] に付加する。カオス性を持つようにパラメータを設定した。SN 比 $p[\text{dB}]$ は $p = 10 \log_{10}(\delta_s/\delta_n)$ となる。ここで、 δ_s は入力する時系列データの分散であり、 δ_n はノイズの分散である。本論文では平均が 0、分散が δ_n の正規分布に従うガウスノイズを生成した。

6 実験結果

両手法で推定したエノン写像のリアプノフ指数の相対誤差を、図 2 に示す。データ数 $N = 1,000, 10,000$ とした。近傍半径 r を $0.0005 \leq r \leq 0.1$ の範囲で 0.005 刻みで変化させ、SN 比は $10 \leq p \leq 40$ の範囲で 5 刻みで変化させた。

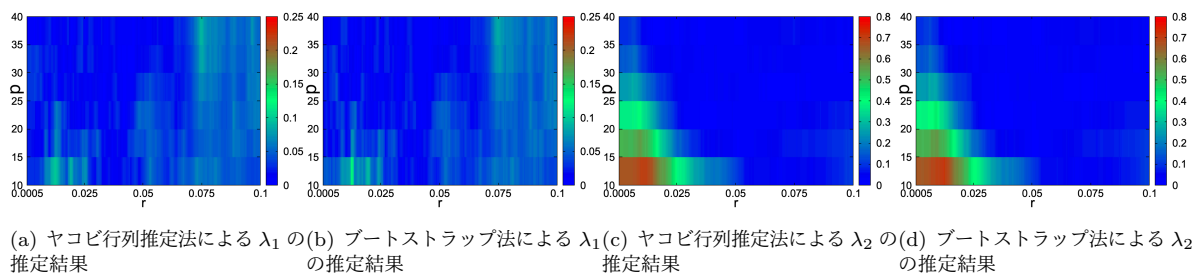


図 2: 推定したエノン写像の λ_i の相対誤差 ($N = 1,000$)

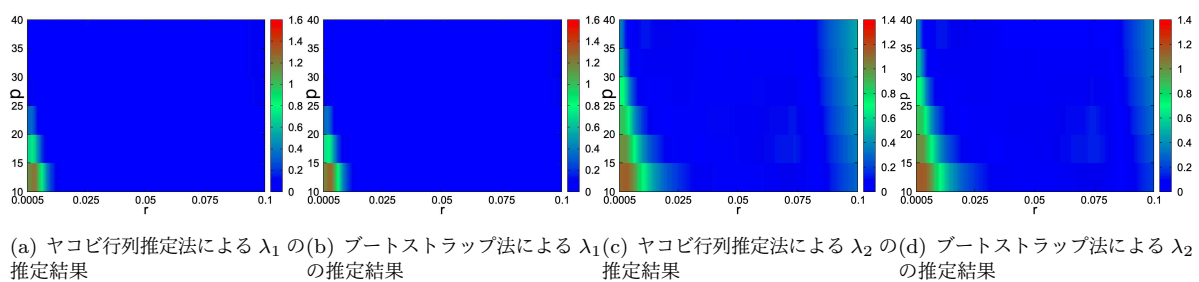


図 3: 推定したエノン写像の λ_i の相対誤差 ($N = 10,000$)

図 2(a), (b) を比較したとき、図が類似しているのがわかる。また図 2(c), (d) を、図 3(a), (b) を、図 3(c), (d) を比較したときも、それぞれ図が類似しているのがわかる。これらの結果から、エノン写像のリアプノフ指数の推定においては、両手法に差がほぼ見られないことがわかった。次に、池田写像に対して同様の実験を行った。

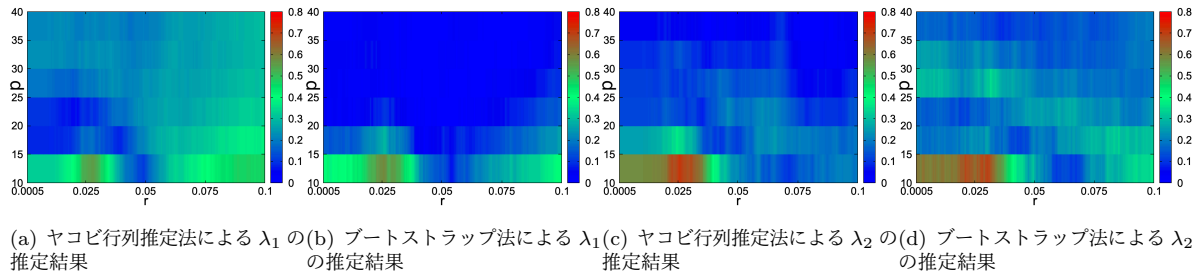


図 4: 推定した池田画像の λ_i の相対誤差 ($N = 1,000$)

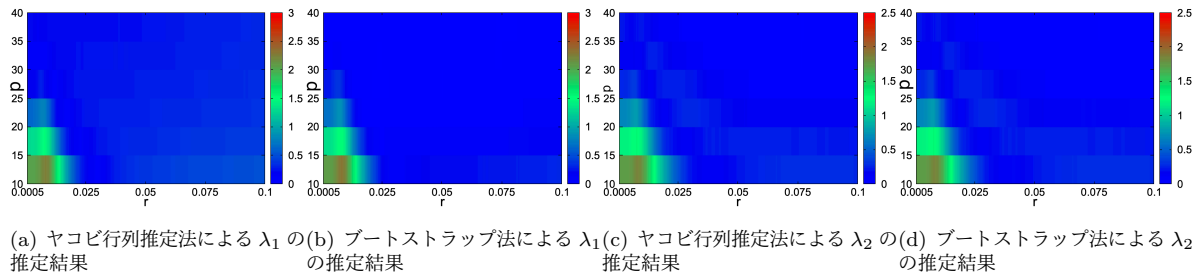


図 5: 推定した池田画像の λ_i の相対誤差 ($N = 10,000$)

図 4(a), (b) を比較すると、ヤコビ行列推定法の方が誤差が大きいのがわかる。また、図 4(c), (d) を比較すると、ブートストラップ法の方が誤差が大きい。一方、図 5(a) と図 5(b), 図 5(c) と図 5(d) を比較したときは、図 4 ほどの差が見られない。これらの結果から、池田画像においては、データ数 N が少ないとき、ブートストラップ法は λ_1 の推定精度が高く、ヤコビ行列推定法は λ_2 の推定精度が高いことがわかった。

7 まとめ

本稿では、エノン写像と池田写像に対してヤコビ行列推定法とブートストラップ法を適用し、リアプノフ指数の推定することで、両手法の性能の比較を行った。その結果、エノン写像においては両手法に差はほとんど見られなかった。一方、池田写像においてはデータ数が少ないとき、ブートストラップ法は特に重要な指標である λ_1 の推定精度が高く、ヤコビ行列推定法は λ_2 の推定精度が高いことがわかった。エノン写像と池田写像の形状の違いにより、池田写像は比較的ノイズの影響を受けやすいのだと考えられる。参考にできるデータ数が少ない場合に、ブートストラップ法はヤコビ行列推定法よりも、リアプノフ指数の推定に関して優れているということが示唆された。

参考文献

- [1] 池口徹, 山田泰司, 小室元政, “カオス時系列解析の基礎と応用,” 合原一幸編, 産業図書, 初版, 2000.
- [2] D. Haraki, T. Suzuki, H. Hashiguchi and T. Ikeguchi, “Bootstrap Nonlinear Prediction,” *Physical Review E*, Vol. 75, p. 056212, 2007.
- [3] M. Henon, “A two-dimensional mapping with a strange attractor,” *The Theory of Chaotic Attractors*, pp 94–102, 1976
- [4] K. Ikeda, H. Daido, and O. Akimoto. “Optical turbulence: chaotic behavior of transmitted light from a ring cavity,” *Physical Review Letters*, Vol 45, No 9, p 709, 1980

固有値の累積寄与率がスペクトルグラフ距離に与える影響

澤田幸輝 (池口徹 教授, Nina Sviridova 助教)

1 はじめに

ネットワークとは、現実中存在するつながりや現象を頂点と枝で表したものである。例えば、人を頂点、2者間の人間関係を枝とすることで社会ネットワークを作成できる。このようなネットワークを解析することで、実世界の様々な問題を解決することができる。複数のネットワークの違いを定量化することで、時間とともに構造が変化するテンポラルネットワークを一次元の時系列データとして扱うことができる。このようにして得られた時系列データに対して既存の時系列解析の手法を用いることで、テンポラルネットワークの特徴を解明できる可能性がある。本論文では、ネットワーク間の距離を測る指標の1つであるスペクトルグラフ距離 [1] ならびに、まず、有向拡張した有向スペクトルグラフ距離 [2] の性能の確認を行った。結果として、有向スペクトルグラフ距離は、有向ネットワーク構造の差異を定量化できることを確認した。また、スペクトルグラフ距離は、ネットワークの隣接行列から得られるラプラシアン行列の固有ベクトルを用いる。しかし、異なる頂点数のネットワーク間の距離を測る際には、頂点数の少ないネットワークに合わせて距離の計算を行うため、頂点数の多いネットワークの固有ベクトルはすべて使用することはできない。その際、論文 [1] では固有値の小さい固有ベクトルから順に計算に使用し、距離の計算を行う方法が紹介されている。しかし、一般的に寄与率の高い固有値、すなわち高い固有値に対応する固有ベクトルの方がその行列の特徴を説明できると考えられる。そのため、固有値の寄与率に着目し、使用する固有ベクトルによる距離の推移について調査を行った。結果として、すべての固有ベクトルが同様にネットワークの情報を有することがわかった。

2 使用ネットワーク

2.1 WS モデル

Watts-Strogaz モデル (以下 WS モデル) [3] はスモールワールド現象を説明する際に提案されたネットワークである。WS モデルは以下の手順で作成することができる。

1. n 個の頂点を円環上に並べ、各頂点から隣接する k 個の頂点で接続し、格子グラフの作成をする。
2. それぞれの枝について確率 p で枝の片方の頂点を切り離し、別の頂点に接続する。

このモデルは、図 1 のように $p = 0$ のとき円環上の格子ネットワークとなり、 $p = 1$ のときランダムネットワークとなる。また、 $p = 0.01$ 付近のとき、スモールワールドネットワークとなることが知られている。

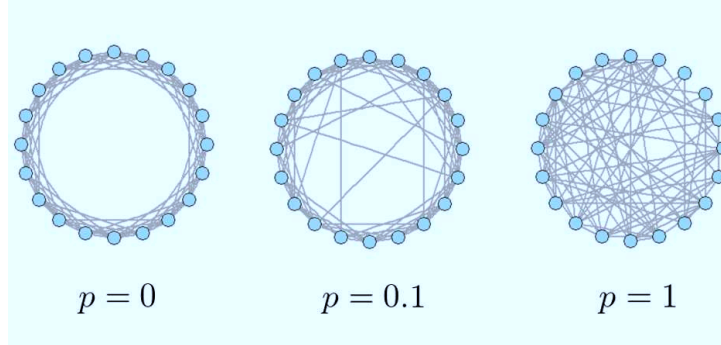


図 1: WS モデル

2.2 有向 WS モデル

この WS モデルを有向グラフに拡張した有向 WS モデル [2] は以下の手順で作成する。

1. WS モデルと同様に n 個の頂点を円環上に並べ、各頂点から隣接する k 個の頂点で接続し、格子グラフの作成を行う。
2. 頂点間の枝を 1 本から 2 本に増やし、それぞれの枝の向きを互いに異なるように決める。
3. それぞれの枝について確率 p_w で枝の片方の頂点を切り離し、別の頂点に接続する。

3 解析手法

3.1 スペクトルグラフ距離

本章では、スペクトルグラフ距離について説明を行う。ネットワーク G に対するラプラシアン行列 ϕ の定義を式 (1) に示す。

$$\phi \equiv K - A \quad (1)$$

A は $N \times N$ の無向ネットワークの隣接行列、 $K = \text{diag}(k_1, \dots, k_N)$ である。ただし、 k_i は頂点 i の次数である。

ϕ は固有値分解を用いて式 (2) のように書き換えることができる。

$$\phi = \mathbf{X}\Lambda(\mathbf{X})^T \quad (2)$$

$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$ である。ただし、各 λ は ϕ の固有値であり、 $\lambda_1 = 0 \leq \lambda_2 \leq \dots \leq \lambda_N$ とする。また、 \mathbf{x}_r を ϕ の r 番目の固有値 λ_r に対応する固有ベクトルとすると、 $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ となる。この固有ベクトルを用いて、2つのネットワーク $G^{(i)}$ 、 $G^{(j)}$ のスペクトルグラフ距離は式 (3) で定義される [1]。

$$d(G^{(i)}, G^{(j)}) \equiv \frac{1}{M_{ij} - 1} \sum_{r=2}^{M_{ij}} d'(\mathbf{x}_r^{(i)}, \mathbf{x}_r^{(j)}) \quad (3)$$

ここで、ネットワーク $G^{(i)}$ 、 $G^{(j)}$ の頂点数を $N^{(i)}$ 、 $N^{(j)}$ と定義すると、 $M_{ij} = \min(N^{(i)}, N^{(j)})$ である。また、 $d'(\mathbf{x}_r^{(i)}, \mathbf{x}_r^{(j)})$ は各ネットワークの r 番目の固有ベクトル間の距離である。 $d'(\mathbf{x}_r^{(i)}, \mathbf{x}_r^{(j)})$ は式 (4) で定義される。このとき、式 (4) は図 2 のように表せる。

$$d'(\mathbf{x}_r^{(i)}, \mathbf{x}_r^{(j)}) \equiv \int_{-\infty}^{\infty} |\varrho_r^{(i)}(y) - \varrho_r^{(j)}(y)| dy \quad (4)$$

ここで、 $\varrho_r(y)$ は式 (5) で定義されるベクトル \mathbf{x}_r の要素の累積分布関数であり、

$$\varrho_r(y) \equiv \frac{1}{N} \sum_{l=1}^N H(y - x'_{rl}) \quad (5)$$

となる。ここで、 $H(y)$ はステップ関数である。また、 x'_{rl} は最大値を 1、最小値を 0 となるように正規化したベクトル \mathbf{x}_r の l 番目の要素である。ただし、正規化は式 (6)

$$x'_{rl} \equiv \frac{x_{rl} - x_{r \min}}{x_{r \max} - x_{r \min}} \quad (6)$$

とする。ここで、 x_{rl} は固有ベクトル \mathbf{x}_r の l 番目の要素、 $x_{r \max} = \max(x_{r1}, x_{r2}, \dots, x_{rN})$ 、 $x_{r \min} = \min(x_{r1}, x_{r2}, \dots, x_{rN})$ である。

式 (5) は図 2 のように各固有値の固有ベクトルの累積分布となる。また、式 (4) は、2つのネットワークの固有ベクトルの累積分布から得られる斜線部の面積である。式 (4) は、各固有ベクトルの組みから求められ、その和の平均が式 (3) である。

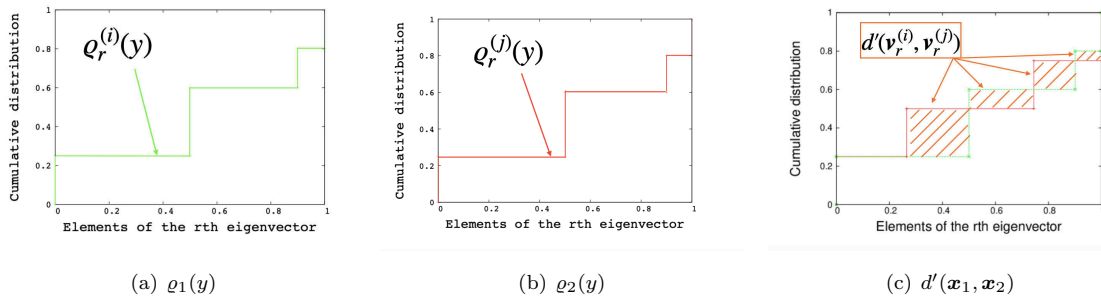


図 2: 式 (4) について

3.2 有向スペクトルグラフ距離

有向スペクトルグラフ距離 [2] では、ラプラシアン行列を用いて計算を行う。有向スペクトルグラフ距離のラプラシアン行列は式 (7) で定義される。

$$\phi_D \equiv A - K_D \quad (7)$$

ここで、 $K_D = \text{diag}(k_{D1}, \dots, k_{DN})$ である。ただし、 k_{Di} は頂点 i の入り次数である。ここで、 ϕ_D の特異値分解は式 (8) のように書き換えることができる。

$$\phi_D = \mathbf{U}\mathbf{S}(\mathbf{V})^T \quad (8)$$

ここで、 \mathbf{S} はラプラシアン行列 ϕ_D の r 番目の特異値 σ_r を用いた対角行列で $\mathbf{S} = \text{diag}(\sigma_1, \dots, \sigma_N)$ である。また、 $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_N)$ 、 $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_N)$ はそれぞれのラプラシアン行列の r 番目の特異値に対応する左特異ベクトル

ルと右特異ベクトルである。左特異ベクトル, 右特異ベクトルを用いて, 2つのネットワーク $G^{(i)}, G^{(j)}$ の有向スペクトルグラフ距離は式 (9), (10) で定義される。

$$d_u(G^{(i)}, G^{(j)}) \equiv \frac{1}{M_{ij} - 1} \sum_{r=2}^{M_{ij}} d'(\mathbf{u}_r^{(i)}, \mathbf{u}_r^{(j)}) \quad (9)$$

$$d_v(G^{(i)}, G^{(j)}) \equiv \frac{1}{M_{ij} - 1} \sum_{r=2}^{M_{ij}} d'(\mathbf{v}_r^{(i)}, \mathbf{v}_r^{(j)}) \quad (10)$$

M_{ij} は $\min(N^{(i)}, N^{(j)})$ である。有向スペクトルグラフ距離では, 左特異ベクトル, 右特異ベクトルのそれぞれでスペクトルグラフ距離 $d_u(G^{(i)}, G^{(j)})$, $d_v(G^{(i)}, G^{(j)})$ を求めることができる。

4 解析結果

4.1 有向スペクトルグラフ距離の性能調査

有向 WS モデルに対して有向スペクトルグラフ距離を適用し, 距離の推移について調査を行った。このとき, 繋ぎ変え確率を固定し, 頂点数を変化させた。

- 1 繋ぎ変え確率 p_0 の有向 WS モデル G_{p_0} を作成する。
- 2 繋ぎ変え確率 p の有向 WS モデルを Y 個作成し, それぞれを $G_p^{(1)}, G_p^{(2)}, \dots, G_p^{(Y)}$ とおく。
- 3 $D(p_0, p) = \frac{1}{Y} \sum_{i=1}^Y d(G_{p_0}, G_p^{(i)})$ を計算する。
- 4 $p = 0.02$ から $p = 1$ まで 0.02 刻みとして, 手順 2, 3 を繰り返す。

ただし, 有向 WS モデルの頂点数は 500, 次数は 10 とした。また, $Y = 5000$ とした。結果を図 2 に示す。どちらの図からも $p_0 = p$ となる付近で $D(p_0, p)$ が最も短くなっていることがみてとれる。また, $p_0 = 0.4$ のとき広範囲で $D(p_0, p)$ が短くなるという傾向もみてとれた。これは, 有向 WS モデルは p が大きくなるほどネットワークとしての特徴量が少なくなっていることが原因となっている。これらのことから, 有向スペクトルグラフ距離もネットワーク構造の差異を定量化できていることがわかる。

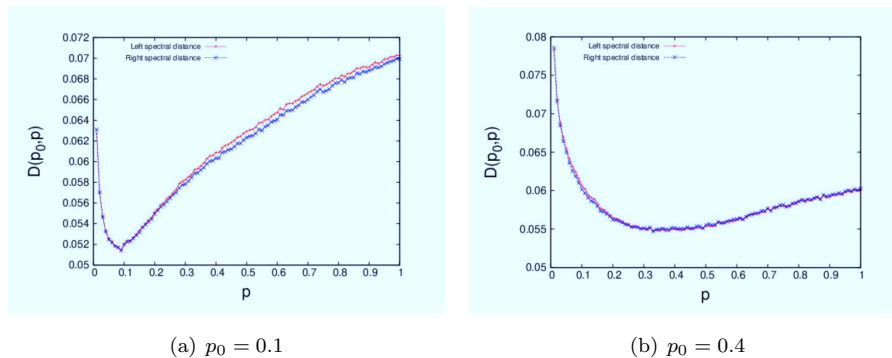


図 3: 有向 WS モデルの頂点数を固定し, 繋ぎ変え確率 p を変化させた結果

4.2 累積寄与率を用いて使用する固有値の数を変化させた際のスペクトルグラフ距離の推移

使用する固有ベクトルの個数によって, スペクトルグラフ距離にどのような変化がみられるのかを確認した。累積寄与率によって使用する固有ベクトルの数を決め, 計算を行った。累積寄与率 c_i に達するまでの使用した最小の固有ベクトルの数を m とする。ここで, 累積寄与率を加味したスペクトルグラフ距離の計算をする式を $d_m(G^{(i)}, G^{(j)}) = \frac{1}{m} \sum_{r=M_{ij}-m+1}^{M_{ij}} d'(\mathbf{x}_r^{(i)}, \mathbf{x}_r^{(j)})$ で定義する。このスペクトルグラフ距離では, 従来のものとは異なり, 固有値の大きい固有ベクトルから順に計算を行う。このスペクトルグラフ距離を用いて以下の数値実験を行った。

- 1 繋ぎ変え確率 p_0 の WS モデル G_{p_0} を作成する。
- 2 繋ぎ変え確率 p の WS モデルを N 個作成し, それぞれを $G_p^{(1)}, G_p^{(2)}, \dots, G_p^{(N)}$ とおく。
- 3 $D(p_0, p) = \frac{1}{N} \sum_{i=1}^N d_m(G_{p_0}, G_p^{(i)})$ を計算する。
- 4 $p = 0.1$ から 1 まで 0.1 刻みとして, 手順 2, 3, 4 を繰り返す。

数値実験に用いた WS モデルの頂点数 $n = 100$, 次数 $k = 4, 50$, $p_0 = 0.1$ とした。また, 累積寄与率 c を 20% から 100% まで 10% 刻みにした際の変化について調査を行った。結果を図 4 に示す。横軸は p , 縦軸は $D(p_0, p)$ を表している。結果から, $c = 100$ のとき, $p = 0.2$ で最も $D(p_0, p)$ が小さくなる。この傾向は c を変化させても変わらない。このことから, 寄与率の高い固有値のみでネットワークの特徴を説明できているといえる。

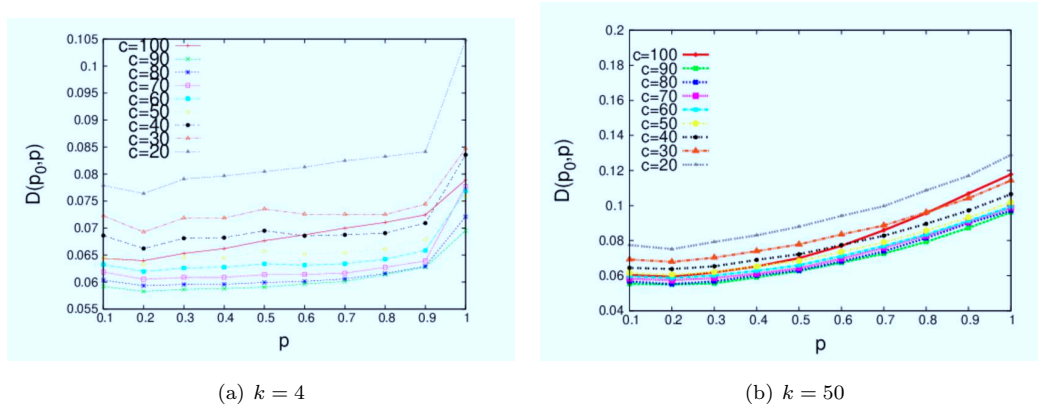


図 4: 繋ぎ変え確率を変化させた際のスペクトルグラフ距離の推移

また、同様の実験を小さな固有値から順に用いて計算を行った。結果を図5に示す。この場合も、図4と同様に、 c の値を変化させても傾向は変化していないことがわかる。このことから、WSモデルは、寄与率の高さによらず、すべての固有ベクトルが同等にネットワークの特徴を表現できていると考えられる。

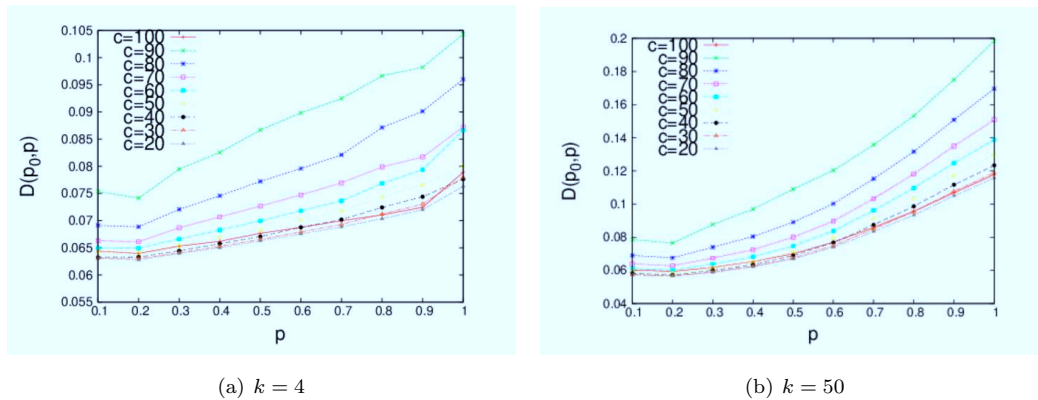


図 5: 小さい固有値から順に累積寄与率を計算した際のスペクトルグラフ距離の推移

5 まとめ

本論文では、有向スペクトルグラフ距離の性能確認を行った。結果として、有向スペクトルグラフ距離もスペクトルグラフ距離と同様にネットワーク構造の差異を定量化できることがわかった。また、累積寄与率を加味したスペクトルグラフ距離を用いて、数値実験を行った。結果として、寄与率の高さによらず、すべての固有ベクトルが同等にネットワークの特性を反映できていることがわかった。

参考文献

- [1] Yutaka Shimada, Yoshito Hirata, Tohru Ikeguchi and Kazuyuki Aihara, "Graph distance for complex networks," *Scientific reports*, Vol. 6, No. 1, pp. 1–6, 2016.
- [2] Dongheon Kim, Yutaka Shimada and Takaomi Shigehara, "Laplacian based Spectral Graph distance for Directed Complex Networks," *Proc. 2019 Int. Symposium. Nonlinear Theory and Its Applications*, pp. 595–598, 2019.
- [3] Duncan J Watts and Steven H Strogatz, "Collective dynamics of 'small-world' networks," *nature*, Vol. 393, No. 6684, pp. 440–442, 1998.

発火間隔時系列から再構成したアトラクタの合成によるニューロンへの共通入力への推定

三浦英 (池口徹 教授, Nina Sviridova 助教)

1 はじめに

人間の脳内には、ニューロンと呼ばれる細胞が多数存在している。ニューロンは、ある閾値を超える入力を与えられると、スパイクを出力することで、情報を他のニューロンに伝達する。脳内の情報表現は、スパイク発生間隔により実現されていると考えられている。すなわち、どのような入力により、どのような出力が得られるのかを解析することは、脳内情報処理原理の解明という立場から重要である。しかし、ニューロンの出力スパイク列を観測することに比べて、ニューロンへの入力信号を観測することは容易ではない。そのため、ニューロンの出力スパイク列から入力信号を推定する手法が提案されている [1, 2]。しかし、文献 [1, 2] の手法では、発火率時系列を用いているため、ニューロンの発火の詳細な情報が欠落している可能性がある。そこで本論文では、個体差のある複数のニューロンより得られる発火間隔時系列から再構成したアトラクタを合成することで、ニューロンへの入力信号を推定する手法を提案する。結果として、提案手法は個体差のある複数のニューロンに対する共通入力への推定において有効であることが示された。

2 ニューロンモデルと発火間隔時系列

本論文では、ニューロンの数理モデルとして、Leaky Integrate and Fire (LIF) モデル [3] を用いた。ニューロンは閾値 θ を超える入力を与えられると発火し、スパイク列を出力する性質を持っている。LIF モデルは、これらの性質を模倣する単純なニューロンモデルであり、式 (1) で定義される。

$$\tau \frac{dV(t)}{dt} = -V(t) + S(t) \quad (1)$$

ここで、 $V(t)$ は膜電位、 $S(t)$ は入力電流を表す。また、 τ は漏れの割合を示すパラメータであり、 $1/\tau \rightarrow 0$ のとき、通常の積分発火モデルとなり、 $1/\tau$ が大きくなると漏れの割合が大きくなる。

入力信号によりニューロンの膜電位 $V(t)$ が上昇することで、閾値 θ に達するとスパイクを出力（発火）し、直後に $V(t) \leftarrow 0$ とリセットされることを繰り返す。ここで、 i 番目の発火時刻を $T(i)$ とすると、発火間隔 (interspike intervals, ISI) 時系列 $t(i)$ は、式 (2) で求めることができる。

$$t(i) = T(i+1) - T(i) \quad (2)$$

3 提案手法

3.1 全体の手順

本論文では、個体差のある複数のニューロンより得られる発火間隔時系列から再構成したアトラクタを合成することで、共通入力への推定する手法を提案する。提案手法の手順を図 1 に示す。

1. 手順 1 では、LIF モデルの漏れ τ を個体差とした n 個のニューロンに対して、共通入力にバイアス項を加えた時系列を入力することで、 n 本の発火間隔時系列を得る。
2. 手順 2 では、 n 本の各発火間隔時系列に対して、アトラクタの再構成を行う [4]。
3. 手順 3 では、各ニューロンへの共通入力の情報を取り出すために、 n 個のアトラクタから 1 個のアトラクタを合成する。
4. 手順 4 では、手順 3 で合成したアトラクタのリカレンスプロット (RP) [5] を作成し、共通入力への推定する [6]。

3.2 時間遅れ座標系への変換によるアトラクタの再構成

現実存在する m 次元の力学系から、複数の状態変数を同時に観測できる場合は稀である。また、多くの場合では元の力学系の次元も不明である。このような場合、観測できた状態変数から元の力学系のアトラクタの再構成を行うことで、元の力学系の情報を再現できる。本論文では、アトラクタの再構成を行う手法として、時間遅れ座標系を用いた [4, 7, 8]。時間遅れ座標系への変換は、1つの状態変数から

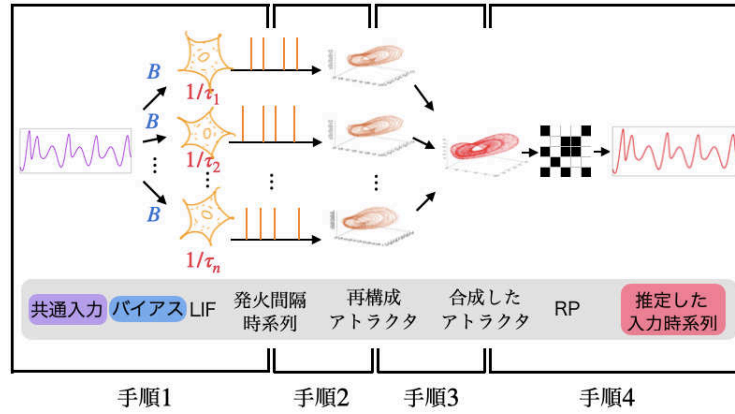


図 1: 提案手法の手順

元の力学系の m 次元のシステムのアトラクタの軌道を再構成する手法である．具体的には，1 変数の時系列信号 $y(t)$ から m 次元の再構成状態空間へ変換する場合，時間遅れ L を用いて式 (3) のように m 次元ベクトル $\mathbf{v}(t)$ を構成する．

$$\mathbf{v}(t) = (y(t), y(t+L), \dots, y(t+(m-1)L)) \quad (3)$$

また，時間遅れ座標系への変換によるアトラクタの再構成手法は，スパイク列に対して拡張されている．具体的には，発火間隔時系列 $t(i)$ を用いて，式 (4) のように m 次元ベクトル $\mathbf{u}(i)$ を構成することで，元の力学系のアトラクタを再構成できる [4]．

$$\mathbf{u}(i) = (t(i), t(i+1), \dots, t(i+(m-1))) \quad (4)$$

3.3 リカレンスプロット

RP は，力学系のアトラクタの状態空間内における 2 点間の距離情報を可視化する手法である [5]．具体的には，アトラクタ上の点の数が N であるとき，一辺の長さが N となる二次元画像を用意する．ここで，式 (5) で $\mathbf{R}(i, j) = 1$ となるとき，第 (i, j) 画素をプロットすることで RP は作成される．

$$\mathbf{R}(i, j) = \begin{cases} 1 & (|\mathbf{u}(i) - \mathbf{u}(j)| < \theta_r) \\ 0 & (\text{それ以外}) \end{cases} \quad (5)$$

なお， θ_r は RP の閾値であり， $\mathbf{u}(i)$ ， $\mathbf{u}(j)$ は式 (4) により求めたアトラクタの状態空間内の 2 点を表す．また， $|\mathbf{u}(i) - \mathbf{u}(j)|$ は $\mathbf{u}(i)$ と $\mathbf{u}(j)$ のユークリッド距離である．

3.4 RP からの時系列の推定

RP は，時系列の特徴を保持しているため，RP から元の時系列を推定することが可能である [6]．文献 [6] では，以下の 4 つの手順で時系列の推定が可能となることが示されている．

1. RP を隣接行列とみなし，RP 上の点 $\mathbf{R}(i, j) = 1$ である頂点 i と頂点 j の間に枝を張ることで無向グラフを作成する．
2. 頂点 i と隣接する頂点の集合 G_i を $G_i = \{j | \mathbf{R}(i, j) = 1\}$ とし，頂点 i と頂点 j の枝の重み $w_{i,j}$ を式 (6) により定義する．

$$w_{i,j} = 1 - \frac{|G_i \cap G_j|}{|G_i \cup G_j|} \quad (6)$$

ここで， $|G_i|$ は集合 G_i の要素数， \cap は積集合， \cup は和集合である．

3. 枝の重み $w_{i,j}$ で重み付けされた無向グラフの全ての頂点 i と頂点 j 間の最短距離 $l_{i,j}$ を計算し，距離行列 $\mathbf{L} = [l_{i,j}]$ を作成する．
4. 距離行列 \mathbf{L} に古典的多次元尺度法 [9] を適用し，時系列を推定する．

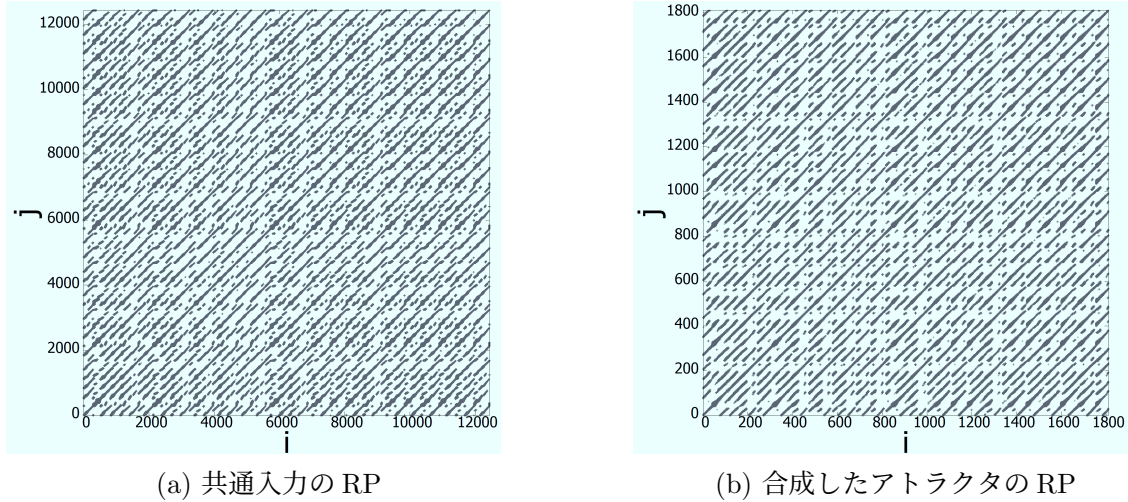


図 3: 共通入力の RP と合成したアトラクタの RP の比較

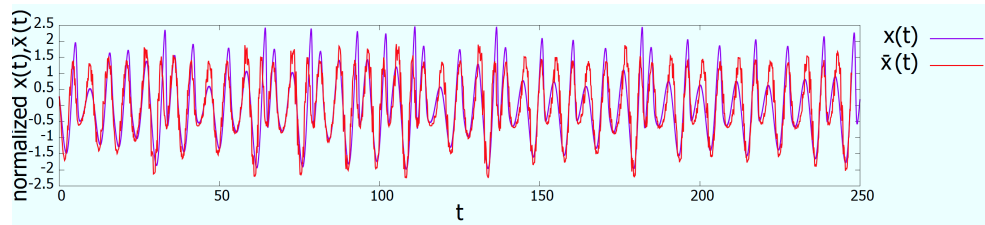


図 4: 共通入力時系列 $x(t)$ と推定した入力時系列 $\hat{x}(t)$

ことがわかる。これらの結果は、個体差のある複数のニューロンより得られる発火間隔時系列から再構成したアトラクタを合成することで、共通入力の推定が可能であることを示している。

5 まとめ

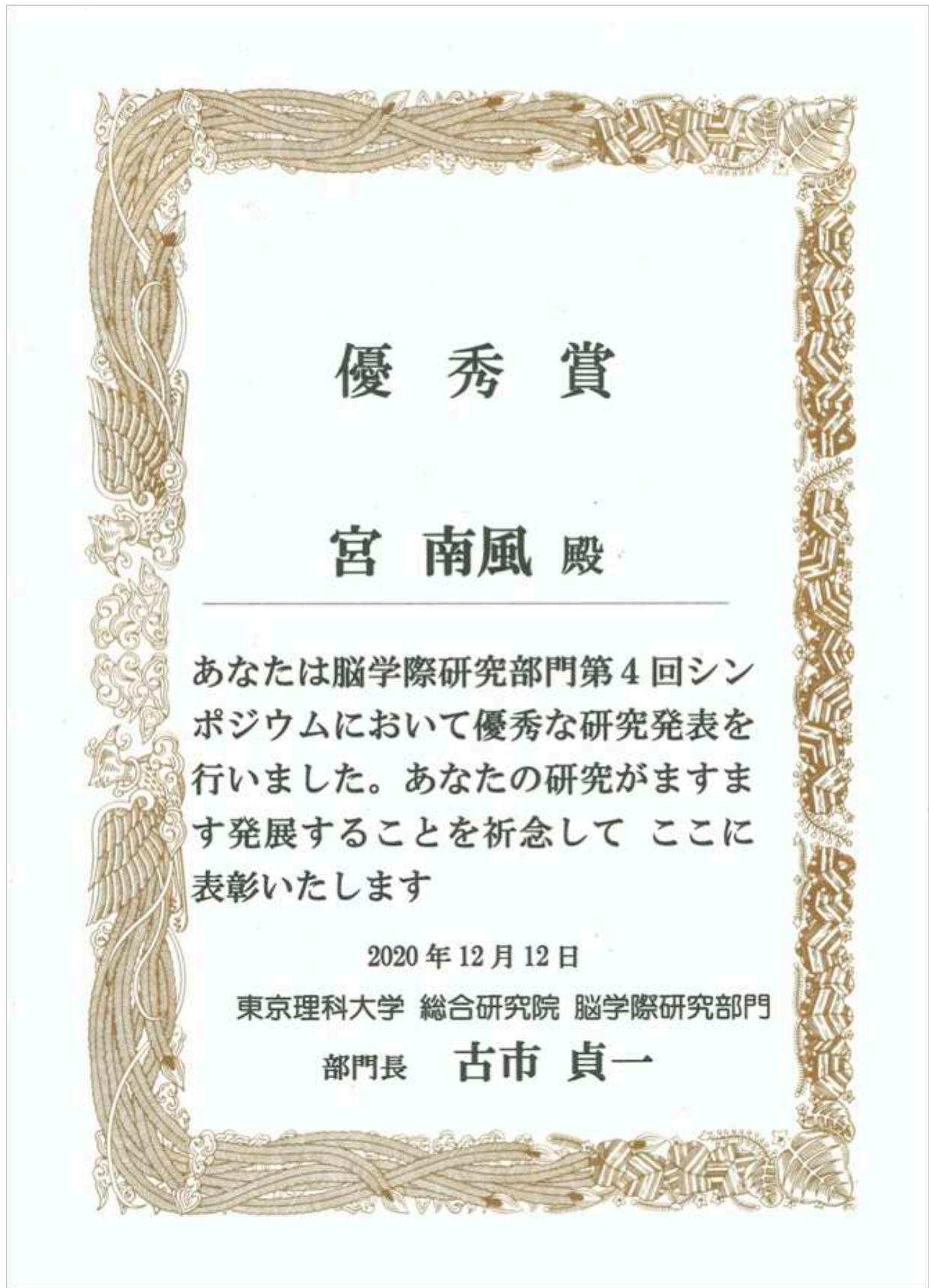
本稿では、個体差のある複数のニューロンより得られる発火間隔時系列から再構成したアトラクタを合成することで、共通入力を推定する手法を提案し、性能評価を行った。結果として、提案手法は個体差のある複数のニューロンに対する共通入力の推定において有効であることが示された。

参考文献

- [1] 野村亮太, 池口徹, “重畳リカレンスプロットを用いた共通入力の再構成,” 2019 電子情報通信学会 総合大会, N-1-24, 2019.
- [2] 野村亮太, 池口徹, “重畳リカレンスプロットを用いた共通入力の再構成に関する一検討,” 電子情報通信学会 技術通信研究報告, Vol. 119, No. 19, pp. 29–34, 2019.
- [3] R. B. Stein, “Some models of neuronal variability,” *Biophysical journal*, Vol. 7, No. 1, pp. 37–68, 1967.
- [4] T. Sauer, “Reconstruction of Dynamical Systems from Interspike Intervals,” *Physical Review Letters*, Vol. 72, No. 24, pp. 3811–3814, 1994.
- [5] J.-P. Eckmann, S. Oliffson Kamphorst and D. Ruelle, “Recurrence plots of dynamical systems,” *Europhysics Letters*, Vol. 4, No. 9, pp. 973–977, 1987.
- [6] Y. Hirata, S. Horai, and K. Aihara, “Reproduction of distance matrices and original time series from recurrence plots and their applications,” *The European Physical Journal Special Topics*, Vol. 164, No. 1, pp. 13–22, 2008.
- [7] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, “Geometry from a time series,” *Physical Review Letters*, Vol. 45, No. 9, pp. 712, 1980.
- [8] F. Takens, “Detecting strange attractors in turbulence,” *Lecture Notes in Mathematics*, Vol. 898, pp. 366–381, 1981.
- [9] J. C. Gower, “Some distance properties of latent root and vector methods used in multivariate analysis,” *Biometrika*, Vol. 53, No. 3–4, pp. 325–338, 1966.
- [10] O. E. Rössler, “An equation for continuous chaos,” *Physics Letters A*, Vol. 57, No. 5, pp. 397–398, 1976.

発表文献

- [P1] 三浦英, 池口徹, “発火間隔時系列を用いたニューロンへの共通入力の再構成,” 電子情報通信学会 技術研究報告, 2021 年 1 月発表予定.
- [P2] 三浦英, 澤田和弥, 池口徹, “複数の再構成アトラクタを用いたニューロンへの共通入力の推定,” 2021 電子情報通信学会 総合大会, 2021 年 3 月発表予定.



宮南風(M2), 東京理科大学より優秀賞受賞



金丸志生(M2), 通信処理学会よりNCSP'21
Student Paper Award受賞



澤田和弥(M2), 通信処理学会よりNCSP'21
Student Paper Award受賞