

# トポロジーの二重性を導入した Echo State Network の性能評価 Performance Evaluation of Echo State Network with Dual Complex Network Structure

北川 蓮也\*      信川 創\*

Renya Kitagawa      Sou Nobukawa

## 1 はじめに

Echo State Network(ESN) はリザーバーコンピューティングの一種であり、主に入力層、中間層(リザーバー層)、出力層の三層からなるニューラルネットワークである。リカレントニューラルネットワークでは、ネットワーク内の全てのシナプス結合を学習によって更新するが、リザーバーコンピューティングでは、入力層、リザーバー層のシナプス荷重は更新せず、出力層のシナプス荷重のみ更新を行う。入力層とリザーバー層の重みを更新しないことで、計算コストが大幅に削減できることがリザーバーのネットワークの利点である [1]。

ESN の性能に影響を及ぼすパラメータとして、スペクトル半径があげられる。リザーバー層のシナプス結合をあらわした行列をスペクトル半径によってスケーリングし、スペクトル半径が 1 付近で性能が最大化することが知られている [1]。さらにリザーバーのネットワークの構造も性能に影響を及ぼす重要な要素である。特に Kawai らの研究では、実際に解剖学的に脳にみられる神経回路網の構造としてスモールワールドネットワークを ESN に導入し、ネットワーク構造が計算性能に影響を及ぼすことを指摘した [2]。

一方、近年の神経生理学の進展により、シナプスの結合強度は対数正規分布を示し、少数の強いシナプス結合がスモールワールドネットワーク性を多数の弱いシナプス結合がランダムネットワーク性を示す二重のネットワークトポロジーを持つことが指摘されている [3][4][5]。そこで本研究では、ネットワークトポロジーが ESN の計算性能を左右するという観点から、リザーバーにこの二重性を導入し、シナプス荷重を対数正規乱数で与えることで、計算性能にどのような影響が生じるかの評価を行った。

## 2 手法

### 2.1 Echo State Network

入力層のノードの数を  $N_{in}$ 、リザーバーのノードの数を  $N$ 、出力層のノードの数を  $N_{out}$  と表す。時刻  $t$  における入力ベクトルを  $\mathbf{u}(t)$ 、リザーバーの状態ベクトルを  $\mathbf{x}(t)$ 、出力ベクトルを  $\mathbf{y}(t)$  とすると、それぞれ、

$$\mathbf{u}(t) = (u_1(t), \dots, u_{N_{in}}(t))^T \quad (1)$$

$$\mathbf{x}(t) = (x_1(t), \dots, x_N(t))^T \quad (2)$$

$$\mathbf{y}(t) = (y_1(t), \dots, y_{N_{out}}(t))^T \quad (3)$$

と表せる。また、入力層とリザーバー、リザーバー内のニューロン同士、リザーバーと出力層をつなぐシナプス結合をそれぞれ、

$$W_{in} = \{w_{ij}^{in}\} \quad (4)$$

$$W_0 = \{w_{ij}\} \quad (5)$$

$$W_{out} = \{w_{ij}^{out}\} \quad (6)$$

と表す。

本論文では Watts and Strogatz モデルに従い [6]、スモールワールドの隣接行列を  $W_0$  として構築した。1 次元のリング形

状パターン  $N$  のノードを用意し、各ノードに隣接する  $E$  個ノードに接続して通常のネットワークを構成する。その後再結線確率  $p$  によって繋ぎ変える。繋ぎ変え確率を変化させるネットワークの隣接行列を  $W_1$ 、ランダムネットワークの隣接行列を  $W_2$  とする。二重性を導入するためにこれら二つの隣接行列を足し合わせる。よって、

$$W_0 = W_1 + W_2 \quad (7)$$

と表現できる。

さらにスペクトル半径によって以下の式でスケーリングを行う。

$$W = \frac{W_0}{\rho(W_0)} \alpha \quad (8)$$

リザーバーの状態の時間発展は、リザーバーのノード  $i(i = 1, \dots, N)$  について、

$$x_i(t+1) = f\left(\sum_{i=1}^{N_{in}} w_{ij}^{in} u_i(t+1) + \sum_{i=1}^N w_{ij} x_j(t)\right) \quad (9)$$

となる。出力ノード  $k(k = 1, \dots, N_{out})$  については、

$$y_k(t+1) = f\left(\sum_{i=1}^N w_{ki}^{out} x_i(t+1)\right) \quad (10)$$

となる。ここで活性化関数  $f$  は  $\tanh$  である。

### 2.2 学習アルゴリズム

学習アルゴリズムはリッジ回帰を用いる。よって  $W_{out}$  は、以下の式に従い更新される。

$$W_{out} = DX^T(XX^T + \beta I)^{-1} \quad (11)$$

ここで、 $X$  は入力データ、 $D$  は教師信号、 $I$  は単位行列、 $\beta(> 0)$  は正則化項の大きさを調整するパラメータである。

### 2.3 性能評価

性能評価は正規化平方平均二乗誤差 ( $NRMSE$ ) によって行う。 $NRMSE$  は次の式で計算する。

$$NRMSE = \sqrt{\frac{\sum_{t=1}^T (y(t) - y_d(t))^2}{T\sigma^2(y_d)}} \quad (12)$$

ここで、 $y_d$  は教師信号であり、 $\sigma$  は標準偏差である。

## 3 実験環境

本実験では、リザーバーは 500 個のノード ( $N = 500$ ) を持つ。入力ノードと出力ノードの数はそれぞれ 10, 100 とした ( $N_{in} = 10, N_{out} = 100$ )。エッジについては、二重トポロジーのうち、ランダムネットワークは 6000 本、確率を変化させるネットワークは 18000 本のエッジを持つように構築した。

また、 $W_{in}$  と  $W_0$  の接続重みは、それぞれ  $[0, 1]$  の一様分布、対数正規乱数によりサンプリングした ( $\mu = \sigma^2 + \log 0.2, \sigma = 1.0$ )。二重性において、スモールワールドネットワークでは 1 未満の

\* 千葉工業大学情報工学科  
Department of Computer Science, Chiba Institute of Technology

値をシナプス荷重として採用し、確率を変化させるネットワークでは、1以上の値を採用した。式(11)の正規化パラメータ $\beta$ は0.2とした。再配線の確率 $p$ は、0, 0.01, 0.1, 1.0で実験した。スペクトル半径 $\alpha$ は、0.1から2.5の範囲で、0.1間隔で設定した。10回試行をおこない、標準誤差を求めた。

学習データには Mackey-Glass 方程式をもちいた。この方程式は

$$\frac{dx}{dt} = \frac{0.2x(t-\tau)}{1+x(t-\tau)^{10}} - 0.1x(t) \quad (13)$$

で与えられる。 $\tau$ は遅延を表す定数である。本実験では $\tau = 32$ と設定した。

#### 4 結果

本実験の実行結果を図1に示す。スペクトル半径 $\alpha$ が0.1から1.5ほどまでは、 $p = 1$ の時に最も性能が良く、それにつづき $p = 0.1, p = 0.001, p = 0$ の順となっている。一方でスペクトル半径 $\alpha$ が1.5から2.5ほどまでは、 $p = 0.1$ の時に最も性能が良く、それにつづき $p = 0, p = 0.001, p = 1$ の順となっている。これは、大きなスペクトル半径下においては、スモールワールドネットワークを持つ二重ネットワークトポロジーが性能劣化をより抑制できることを示している。

#### 5 考察

本研究では、ESNのリザーバーにスモールワールドネットワークとランダムネットワークが混在した二重トポロジーを導入し、スモールワールドネットワーク部分には1以上、ランダムネットワーク部分には1未満のシナプス荷重を対数正規乱数で与えることで、性能にどのような影響を与えるかを評価した。その結果、スペクトル半径 $\alpha$ が0.1から1.5ほどまでは、ランダムネットワークの単一トポロジーである( $p = 1$ )が最も性能が良く、それにつづき、シナプス結合がスモールワールド性を示すトポロジー( $p = 0.1$ )、弱めのスモールワールド性を持つトポロジー( $p = 0.01$ )、レギュラーネットワークトポロジー( $p = 0$ )の順となっている。一方でスペクトル半径 $\alpha$ が1.5から2.5ほどまでは、 $p = 0.1$ の時に最も性能が良いことがわかった。このようにスモールワールド性を持つネットワークにおけるスペクトル半径の増加に対しての性能劣化抑制に関しては、Kawaiらのスモールワールドネットワーク単体での結果と同じような結果が得られた[2]。

リザーバーにおいて、適度な同期性、すなわち、比較的小さな自由度のダイナミクスが時系列予測においては必要と考えられる。さらにスペクトル半径が大きくなるにつれ、リザーバーにおけるシナプス荷重が大きくなる。よって高いスペクトル半径が1.5から2.5のスモールワールドネットワーク( $p = 0.1$ )では、ネットワークにおけるクラスタ間の適度な相互作用が、リザーバーにおけるニューロン間の自由度を低下させ、性能が向上

することが考えられる。一方、格子状のネットワーク( $p = 0$ )では、クラスタは存在しているが、高い平均経路長に起因して、クラスタ間の相互作用が少なく、自由度が高くなり、性能が劣化すると考えられる。さらにランダムネットワーク( $p = 1$ )ではクラスタは存在せず、高い自由度により、性能が劣化すると考えられる。

実際にみられている脳のネットワーク構造は、スモールワールドネットワーク、ランダムネットワークだけでなく、リッチクラブ[7]、モジュール形成などがある[8]。今後の研究の方向性としては、これらを適用し、より脳に倣った構造にすることで、性能劣化の抑制が可能であると考えられる。

#### 6 おわりに

本研究では、ESNにおけるリザーバー(二重トポロジー)に対数正規乱数でシナプス荷重を与え、大きなスペクトル半径の時は、スモールワールド性を持つトポロジーが内部ニューロン間の自由度を低下させ、性能劣化を他のトポロジーと比べ抑制することがわかった。この結果は、二重トポロジーの導入によって、計算性能のスペクトル半径の増減による、性能劣化の抑制可能性があることを示唆している。よってリザーバーの構造を脳に近づけることによる、性能劣化抑制などの機能的意義を理解するための土台となったといえる。

#### 参考文献

- [1] Jaeger, H. (2001). The "echo state" approach to analyzing and training recurrent neural networks. Tech. rep., GMD Report 148, German National Research Center for Information Technology.
- [2] Yuji Kawai, Jihoon Park. & Minoru Asada. (2019) A small-world topology enhances the echo state property and signal propagation in reservoir computing. *Neural Networks* 112 15-23.
- [3] Sou Nobukawa, Haruhiko Nishimura, Teruya Yaminishi. (2019) Temporal-Scientific complexity of spiking patterns in spontaneous activity induced by a dual complex network structure. *Scientific Reports* 9.
- [4] Teramae, J.-N., Tsubo, Y. & Fukai, T. Optimal spike-based communication in excitable networks with strong-sparse and weak-dense links. *Scientific Reports* 2 (2012).
- [5] Watanabe, K., Teramae, J.-N. & Wakamiya, N. Inferred duality of synaptic connectivity in local cortical circuit with receptive field correlation. In *International Conference on Neural Information Processing*, 115-122 (Springer, 2016).
- [6] Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393(6684), 440-442. Xiaohu, L., Xiaoling, L., Jinhua, Z., Y
- [7] van den Heuvel, M.P. et al. (2012) High-cost, high-capacity backbone for global brain communication. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11372-11377
- [8] Kötter, R. et al. (2001) Connectional characteristics of areas in Walker's map of primate prefrontal cortex. *Neurocomputing* 38, 741-746

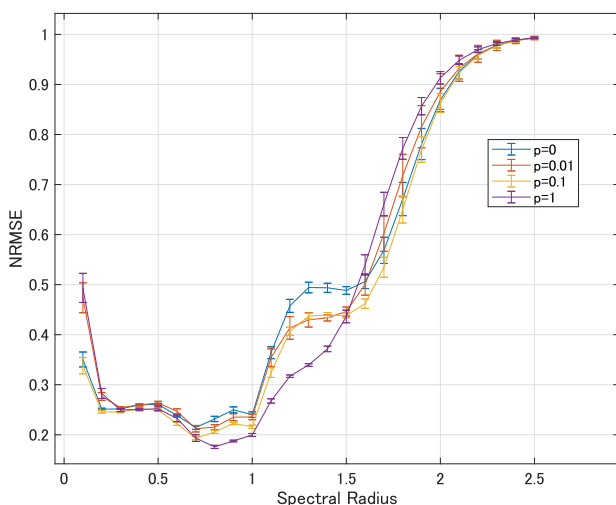


図1 各再結線確率による、予測性能(NRMSE)のスペクトル半径への依存性。