Interpretable Graph Reservoir Computing With the Temporal Pattern Attention

Xinyu Han and Yi Zhao

***Abstract*— Graph reservoir computing (GraphRC) gains increasing attention by virtue of its high training efficiency. However, since GraphRC is developed without knowledge of its internal mechanism, it cannot be fully trusted to deploy in practice. Although there are some existing approaches that can be extended to interpret GraphRC, the specific role played by each neuron (i.e., reservoir node) of GraphRC is far less explored. To address this issue, the latent short-term memory property of each reservoir node of GraphRC is qualitatively characterized to unravel its role in predicting the graph signal, thereby enabling an interpretable GraphRC. Specifically, we first deduce the equivalence between the GraphRC and conventional reservoir computing (RC). Then, the underlying memory properties of the GraphRC and its reservoir nodes can be characterized in theory by the multisource reachability among the reservoir nodes in the transformed RC. Moreover, the distinct temporal patterns hidden in reservoir nodes are identified, and then, an attention mechanism based on the identified temporal patterns is deployed in the GraphRC to improve its performance. In addition, the effectiveness of the interpretability for GraphRC and improved GraphRC is verified on the Lorenz-96 spatiotemporal dynamical system. The experimental results of the Lorenz-96 spatiotemporal chaotic system and three real-world traffic datasets demonstrate that the improved GraphRC is superior to original GraphRC and can achieve prediction performance comparable to the state-of-the-art baseline models, but with much less training cost.**

***Index Terms*— Graph reservoir computing (GraphRC), interpretability analysis, Kronecker product of graphs, spatiotemporal prediction, temporal pattern attention.**

I. Introduction

RECENTLY, conventional reservoir computing (RC) [1] has been applied to a wide range of practical scenarios, such as meteorological series prediction [2], multivariate time series classification [3], and quantum computing [4]. Since many real-world data are presented in the form of graph, conventional RC is improved to handle the graph signal, and then, graph reservoir computing (GraphRC) [5] is proposed. GraphRC serves as a pioneer type of graph neural network (GNN) [6] and becomes an attractive area due to its low computational cost. For example, GraphRC can be improved to tackle the typical graph classification problem with high training efficiency [7]. However, like the GNNs, GraphRC is associated with low transparency, which hampers its employment in practice. In particular, the complex topology structure of the graph signal imposes extra challenges on the interpretability analysis of GraphRC, compared with conventional RC. Therefore, interpreting GraphRC is a critical but quite challenging task.

Although no direct study on the interpretability of GraphRC is available, there are some existing interpretability methods for GNN [8], which are model-agnostic approaches and compatible with any GNN without any assumption on the learning task. These model-agnostic interpretability methods can be extended to explain GraphRC. GNNExplainer [9] provides an explanation for the output of a trained GNN. Specifically, the real-valued graph mask is introduced to identify the crucial node features and edges of the input graph. The identified node features and edges convey specific semantic meanings, thereby explaining how GNN generates the certain output. Similarly, PGExplainer [10] adopts a multilayer perceptron to learn the edge distribution and then identify important edges that serve as an explanation for GNN predictions. OrphicX [11] further constructs a generative model to produce an adjacency mask of the input graph, where such generated adjacency mask provides the explanation for the trained GNN. In GNN-LRP [12], GNN predictions are decomposed on the sequences of edges in the input graph with the higher order Taylor expansion such that the importance of the sequences of edges confers the interpretation for GNN. SubgraphX [13] is proposed to explain GNN with crucial subgraphs of the input graph, which are identified by Monte Carlo tree search [14]. Unlike the above five methods, Graphlime [15] applies an interpretable surrogate method (i.e., Hilbert–Schmidt independence criterion (HSIC) lasso) [16] to approximate the trained GNN outputs of the target nodes and its N-hop neighboring nodes. The explanation of the given GNN is then naturally obtained through the selected surrogate method. Note that the six methods provide an input-dependent explanation for GNN. In contrast, XGNN [17] gives an input-independent interpretation for GNN. In this method, the input graph that is constructed by the proposed graph generator, instead of the original input graph, is fed into the given trained GNN. The graph generator aims to figure out which graph structures lead to a certain GNN output, thereby bringing a general understanding of how GNN works. While these existing

Manuscript received 15 July 2021; revised 1 August 2022 and 6 November 2022; accepted 15 December 2022. Date of publication 29 December 2022; date of current version 9 July 2024. This work was supported by the Natural Science Foundation of Guangdong, China, under Grant 2020A15150108 and Grant 2021A1515011594. The work of Xinyu Han was supported by the China Scholarship Council under Grant 202006120443. *(Corresponding author: Yi Zhao.)*

The authors are with the School of Science, Harbin Institute of Technology, Shenzhen, Guangdong 518055, China (e-mail: hanxinyu@stu.hit.edu.cn; zhao.yi@hit.eud.cn).

Digital Object Identifier 10.1109/TNNLS.2022.3231620

Chaos ARTICLE scitation.org/journal/cha

A tighter generalization bound for reservoir computing

Cite as: Chaos **32**, 043115 (2022); doi: 10.1063/5.0082258
Submitted: 13 December 2021 · Accepted: 22 March 2022 ·
Published Online: 11 April 2022

View Online Export Citation CrossMark

Xinyu Han,[1] Yi Zhao,[1,a)] and Michael Small[2]

AFFILIATIONS

[1]Harbin Institute of Technology, Shenzhen, 518055 Guangdong, China
[2]Complex Systems Group, Department of Mathematics and Statistics, The University of Western Australia, Crawley, Western Australia 6009, Australia

[a)]**Author to whom correspondence should be addressed:** zhao.yi@hit.edu.cn

ABSTRACT

While reservoir computing (RC) has demonstrated astonishing performance in many practical scenarios, the understanding of its capability for generalization on previously unseen data is limited. To address this issue, we propose a novel generalization bound for RC based on the empirical Rademacher complexity under the probably approximately correct learning framework. Note that the generalization bound for the RC is derived in terms of the model hyperparameters. For this reason, it can explore the dependencies of the generalization bound for RC on its hyperparameters. Compared with the existing generalization bound, our generalization bound for RC is tighter, which is verified by numerical experiments. Furthermore, we study the generalization bound for the RC corresponding to different reservoir graphs, including directed acyclic graph (DAG) and Erdős–Rényi undirected random graph (ER graph). Specifically, the generalization bound for the RC whose reservoir graph is designated as a DAG can be refined by leveraging the structural property (i.e., the longest path length) of the DAG. Finally, both theoretical and experimental findings confirm that the generalization bound for the RC of a DAG is lower and less sensitive to the model hyperparameters than that for the RC of an ER graph.

Reservoir computing (RC) has been widely applied in practice and is gaining popularity in theory. Previous studies have established that RC enables to infer unseen parts of the model input that are never encountered in the training phase. However, a bottleneck for the inference of RC is the generalization issue that results in the significant improvement of the RC performance for known data but not for unseen data. As a result, it is essential to study the generalization bound for the RC trained on finite data, which characterizes a measure of how the RC performs with the unseen data. Moreover, the hyperparameter setting of the RC normally plays a key role in its performance. In this paper, we propose a method under the probably approximately correct (PAC) learning framework to estimate the theoretical generalization bound of RC with explicit dependencies on its hyperparameters. In addition, the reservoir graph type serves as a latent factor influencing the RC performance. Hence, the generalization bound of RC is also discussed theoretically in response to specific reservoir graphs, including directed acyclic graph (DAG) and Erdős–Rényi undirected random graph (ER graph). To verify the theoretical findings and effectiveness of the derived generalization bound for RC, a series of experiments are conducted on the chaotic time series benchmark.

I. INTRODUCTION

Recently, reservoir computing[1] (RC)—a special sub-class of conventional recurrent neural network (RNN)—has emerged as a versatile approach for processing sequential data by virtue of its low training costs, such as chaotic time series prediction[2] and signal separation.[3] RC has revealed to be able to construct abstract representations from the observed examples of model input,[4] thereby enabling RC to infer the existence and shape of attractors of model input not seen explicitly during the training phase.[5] Moreover, a pre-trained RC can be applied to predict the unseen dynamics of the model input under different control parameters.[6] It is worth emphasizing that RC presents a promising potential for modeling previously unseen data, where such potential is normally termed as the generalization capability. As a result, interpreting the underlying generalization capability of RC is conducive to reaching

Chaos ARTICLE scitation.org/journal/cha

Revisiting the memory capacity in reservoir computing of directed acyclic network

Cite as: Chaos **31**, 033106 (2021); doi: 10.1063/5.0040251
Submitted: 11 December 2020 · Accepted: 5 February 2021 ·
Published Online: 2 March 2021

View Online Export Citation CrossMark

Xinyu Han,[1] Yi Zhao,[1,a)] and Michael Small[2,3]

AFFILIATIONS

[1] Harbin Institute of Technology, Shenzhen, 518055 Guangdong, China
[2] Complex Systems Group, Department of Mathematics and Statistics, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
[3] Mineral Resources, Commonwealth Scientific and Industrial Research Organisation, 26 Dick Perry Ave., Kensington, WA 6151, Australia

[a)] **Author to whom correspondence should be addressed:** zhao.yi@hit.edu.cn

ABSTRACT

Reservoir computing (RC) is an attractive area of research by virtue of its potential for hardware implementation and low training cost. An intriguing research direction in this field is to interpret the underlying dynamics of an RC model by analyzing its short-term memory property, which can be quantified by the global index: memory capacity (MC). In this paper, the global MC of the RC whose reservoir network is specified as a directed acyclic network (DAN) is examined, and first we give that its global MC is theoretically bounded by the length of the longest path of the reservoir DAN. Since the global MC is technically influenced by the model hyperparameters, the dependency of the MC on the hyperparameters of this RC is then explored in detail. In the further study, we employ the improved conventional network embedding method (i.e., struc2vec) to mine the underlying memory community in the reservoir DAN, which can be regarded as the cluster of reservoir nodes with the same memory profile. Experimental results demonstrate that such a memory community structure can provide a concrete interpretation of the global MC of this RC. Finally, the clustered RC is proposed by exploiting the detected memory community structure of DAN, where its prediction performance is verified to be enhanced with lower training cost compared with other RC models on several chaotic time series benchmarks.

Reservoir computing (RC) has been extensively applied in many practical applications, especially in the modeling of the chaotic dynamical system. Previous researchers have established that the short-term memory is a crucial dynamic property of RC for its functioning in the prediction of chaotic systems. The memory capacity (MC) is proposed to measure the short-term memory property of RC quantitatively and the MC of typical RC of the undirected reservoir network has been systematically analyzed in previous studies. Moreover, some research studies have proved that the RC of the directed line network [i.e., a special case of the general directed acyclic network (DAN)] is often sufficient for obtaining high MC than that of the undirected network, which indicates the MC is sensitive to the reservoir network structure change. However, there is a lack of comprehensive studies on the MC of the RC of general DAN. Therefore, it is necessary to re-examine the short-term memory capacity of the RC of general DAN. In this paper, we give the theoretical bound of the MC and then characterize the effects of model hyperparameters on the MC. Since the MC is a global memory index, we further mine the memory community structure of the reservoir DAN by using the improved network embedding method, thereby gaining a more definitive depiction of the global MC from a local perspective. On this basis, the clustered RC is proposed to make an accurate prediction of some chaotic system benchmarks at a low training cost.

I. INTRODUCTION

Reservoir computing[1] is a typical recurrent neural network and is regarded as a unified computing framework since its core idea has been independently discovered both in the echo state network (ESN)[2] and the liquid state machine (LSM).[3] A number of

Neurocomputing 445 (2021) 134–148

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Reservoir computing dissection and visualization based on directed network embedding

Xinyu Han, Yi Zhao *

Harbin Institute of Technology, Shenzhen 518055, Guangdong, China

ARTICLE INFO

Article history:
Received 18 July 2020
Revised 20 January 2021
Accepted 12 February 2021
Available online 2 March 2021
Communicated by Zidong Wang

Keywords:
Reservoir computing
Directed network embedding
Memory community structure
Time series prediction

ABSTRACT

The reservoir computing (RC) has recently gained considerable attention in practice and many methods have been developed to study its internal mechanism. However, the specific role played by the reservoir nodes of RC in time series prediction is still to be defined. An interpretable RC model wherein its reservoir network is designated as the directed acyclic network (DAN) is proposed with focus on time series prediction in this paper. In virtue of asymmetric transitivity and hierarchical structure of DAN, we present a directed network embedding method to identify the latent memory property of each node in the DAN. Such memory property is utilized to characterize the roles played by the reservoir nodes on the prediction performance of the RC. Meanwhile, it can also be leveraged to identify the corresponding memory community of DAN. As a result, we demonstrate how the reservoir network structure takes effect on the prediction performance from the perspective of memory community. In addition, two novel hyperparameters with the deterministic meaning are introduced to quantify the influence of the model initialization on the reservoir input so as to facilitate further dissection of the interpretable RC. The experimental results indicate that tuning these hyperparameters, which is explicable in terms of the Taylor expansion of the activation function, serves as an essential step in achieving superior prediction performance. Finally, comparative experiments with some other RC models on various time series benchmarks are also conducted.

1. Introduction

Recently, reservoir computing has been adopted in many practical applications, such as speech recognition [1], taxi destination prediction [2] and image identification [3] by virtue of its advantages of hardware realization [4] and low training cost. As noted by Jaeger [5], RC can be regarded as a unified computing framework which includes echo state network (ESN) [6], liquid state machine (LSM) [7] and backpropagation decorrelation (BPDC) [8]. In addition, the reservoir computing framework is naturally suitable for temporal information processing, especially for time series prediction [9]. To date, various variants of RC have been proposed to improve prediction accuracy. Gallicchio et al. [10] introduce the deep architecture into the RC and propose a deep reservoir computing (Deep RC). Qiao et al. [11] provide a Growing ESN (GESN) for automatic design of the reservoir scale and topological structure, and demonstrate that GESN yields better performance than basic ESN in terms of training time and prediction accuracy. Moreover, several studies introduce the regularization methods into RC to avoid the over-fitting problem so as to obtain high-precision prediction models, such as Lasso Regression-based ESN [12] and Ridge Regression-based ESN [13]. More recently, Xu et al. [14] apply RC to predict the spatial–temporal meteorological data with successes. They employ the cubic spline method to address the interaction of spatial and temporal effects, and then present a Spatio-Temporal-Interpolated ESN.

Although there have been fruitful achievements on the improvement and application of reservoir computing, its interpretability analysis is still a nascent field and gains increasing attention recently. Furthermore, the echo state property (ESP) [6] and fading memory property (FMP) [15] are the two pivotal properties of RC, which play key roles in delineating RC dynamical mechanism. Lymburn et al. [16] give a comprehensive portrait of the ESP in RC by using the consistency concept and assess the model performance in a range of consistency regimes on a memory task. Grigoryeva et al. [17] conclude that RC is a universal uniform approximant in the context of discrete-time fading memory filters with uniformly bounded inputs. The more universality statements of RC with respect to the L^p criteria are presented by Gonon [18]

* Corresponding author.
E-mail address: zhao.yi@hit.edu.cn (Y. Zhao).

https://doi.org/10.1016/j.neucom.2021.02.029
0925-2312/

Identification of Dynamical Behavior of Pseudoperiodic Time Series by Network Community Structure

Xinyu Han, Yi Zhao, and Michael Small, *Senior Member, IEEE*

***Abstract*—Although some transformation mappings demonstrate that different time series result in networks with distinct topological properties, it remains unclear whether the community structure in networks is related to the dynamical characteristics of the original time series. In this brief, we capture the underlying deterministic dynamics of a time series with intrinsic community structure of the corresponding network on the basis of a novel transformation method. Specifically, our findings suggest that there exist strong associations between the community structure in networks and the dynamical regimes of the time series. Emphasis is put on identifying typical characteristics of chaotic time series in terms of the variation trend of the community structure in the transformed network, especially taking the unstable periodic orbits of the dynamical system into consideration. Notably, the community structure characterizes the initial sensitivity and ergodicity of chaotic time series well. Moreover, sparse chaotic attractors and filled-in chaotic attractors are distinguished by the network community structure.**

***Index Terms*—Community structure, pseudoperiodic time series, complex network, dynamical detection.**

I. Introduction

Recently, complex networks have developed greatly both in theory [1] and in application [2]–[4]. Along with the boom interdisciplinary research, complex networks are also applied to the realm of time series analysis, of which the identification of underlying dynamical behaviour of pseudoperiodic time series based on the complex networks has aroused a lot of focus [5]–[12]. For this reason, many algorithms have been proposed to transform time series to complex networks and explore the underlying dynamics of time series by network statistics. Zhang and Small [5] construct complex networks from pseudoperiodic time series and adopt each cycle of the time series as a network node. The degree distribution of complex networks is used to identify the dynamics of the corresponding time series. For example, they claim that chaotic time series correspond to the networks which exhibit the small-world and scale-free properties. Lacasa *et al.* [6] present a simple and fast algorithm, called the visibility algorithm, to transform time series into complex networks, where the "visibility" of data points determines the connection of network nodes. In addition, Xu *et al.* [7] use embedded points as the nodes in a complex network and construct the k-nearest neighbor network. As a result, they utilize the distribution of motifs instead of the degree to distinguish the different types of diverse continuous dynamics: periodic time series, chaotic time series, both with and without noise. When the weight and direction of the network are both taken into consideration, an algorithm with an ordinal pattern characterization [8] is then proposed to detect the dynamical transitions in the underlying system by global properties of networks. Remarkably, based on the methods founded on recurrence plots, higher-order statistical properties of time series are vividly depicted by recurrent networks [9]–[11]. In consideration of the equivalence of the transformation between time series and complex networks, Zhao *et al.* [12] propose a quasi-isometric map with the theoretical proof which preserves the geometrical characters of complex system during transformations. Recently, researchers have focused on the transformation between multilayer network and multivariate time series for their reciprocal characterisation [13]–[15].

The previous works establish the connections between some dynamical characteristics of time series and global statistics of the network. Nonetheless, not only should the global statistics of the networks be applied to distinguish the dynamical characteristics (e.g., chaotic or periodic dynamics) but also the local features of the networks should positively interpret the refinement states of the time series. Paradoxically, global statistics may become inadequate or even inapplicable to dynamical identification of the time series due to their limited ability to capture local and detailed characters hidden in time series. Besides, chaos that originates from the deterministic nonlinear systems, is the seemly random movement which heavily depends on several detailed features, such as sensitivity to initial values and exponent divergence of adjacent trajectories [16]. Hence, these works merely emphasize the qualitative analysis of the underlying dynamics of the time series from the network perspective. To the best of our knowledge there are few attempts that can exhibit the dynamical characters accurately, such as the strictly periodic character based on the network structural features. Accordingly, when we take into consideration some local features of network, it is presumably

Manuscript received December 29, 2018; revised February 21, 2019; accepted March 5, 2019. Date of publication March 8, 2019; date of current version November 1, 2019. This work was supported in part by the National Natural Science Foundation of China under Project 61573119, and in part by the Fundamental Research Project of Shenzhen under Project JCYJ20170307151312215 and Project JCYJ20150403161923533. This brief was recommended by Associate Editor X. Wu. *(Corresponding author: Yi Zhao.)*

X. Han and Y. Zhao are with the School of Science, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China (e-mail: hanxinyu@stu.hit.edu.cn; zhao.yi@hit.eud.cn).

M. Small is with the Department of Mathematics and Statistics, University of Western Australia, Crawley, WA 6009, Australia, and also with the Mineral Resources, CSIRO, Kensington, WA 6105, Australia (e-mail: michael.small@uwa.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSII.2019.2903936