

文章编号:1007-5321(2021)05-0010-04

DOI:10.13190/j.jbupt.2021-021

Dropout 回声状态网络的网络流量预测

牟晓惠, 李丽香

(北京邮电大学 网络空间安全学院, 北京 100876)

摘要: 提出了一种基于 Dropout 方法的回声状态网络(ESN). 将基于 Dropout 方法的 ESN(Dropout ESN)应用到实际的网络流量预测任务中, 设置储备池内神经元以不同的概率停止工作, 将经典的 ESN 和 Dropout ESN 进行了对比, 分析了 2 种网络对预测性能的影响, 将基于 Dropout 方法的 ESN 和其他网络的正规化方均根差进行对比分析. 仿真结果表明, Dropout ESN 对网络流量预测效果更优.

关 键 词: 机器学习; 回声状态网络; 网络流量预测; Dropout 方法

中图分类号: TP183

文献标志码: A

Network Traffic Prediction of Dropout Echo State Network

MU Xiao-hui, LI Li-xiang

(School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: An echo state network (ESN) based on Dropout method is proposed. The ESN based on Dropout method (Dropout ESN) is applied to the actual network traffic prediction task, in which the neurons in the reservoir are set to stop working with different probability. Dropout ESN is compared with the classical ESN to analyse the influence of the two networks on the prediction performance. In addition the normalized root mean square error of Dropout ESN and other models are compared and analyzed. Simulation results show that Dropout ESN has better prediction performance on network traffic than other ESN models.

Key words: machine learning; echo state network; network traffic prediction; Dropout method

互联网的高速发展改变了人们的生活方式. 对网络流量进行预测可以规避网络中容易出现的问题, 有效地优化和调整网络资源, 进一步保障重要节点的网络连接. 网络流量预测结果可以作为未来一段时间内流量走势的重要参考^[1]. 目前, 对网络流量的研究有很多^[2-3]. 很多学者提出了预测网络用来分析网络流量的特点. 基于线性时间序列流量预测网络可用于短期预测, 但这些简化了实际流量的假设条件, 未考虑流量的波动性等非线性变化, 相关的网络及应用有自回归网络^[4]、自回归滑动平均网络^[5]、自回归差分滑动平均网

络^[6-8]、自回归分数滑动平均网络^[9]等. 由于网络流量不是平稳变化, 而是存在很大的波动性. 因此, 接近真实情况的网络必定是非线性的, 如人工神经网络^[10]. 回声状态网络(ESN, echo state network)是一种递归神经网络, 其随机产生的储备池结构和权值会在一定程度上影响网络对网络流量数据非线性刻画. 通常情况下, 可以通过调整 ESN 储备池的谱半径来改进网络动力学特性以增强网络的性能, 也可以通过调整储备池中神经元与神经元之间的连接突触的权值来优化储备池的性能. 很多研究还从不同的角度改进了储备池的

收稿日期: 2021-03-11

基金项目: 国家重点研发计划项目(2020YFB1805402); 国家自然科学基金项目(62032002, 61972051)

作者简介: 牟晓惠(1991—), 女, 博士生.

通信作者: 李丽香(1978—), 女, 教授, 博士生导师, E-mail: lixiang@bupt.edu.cn.

结构,增加了储备池的动力学特性. Strauss 等^[11]在保证网络的回声状态特性的前提下降低了噪声的影响. Bianchi 等^[12]分析研究了储备池的动力学特性,构造了最优储备池结构. Ozturk 等^[13]对 ESN 储备池的动力学特性进行了衡量. Deng 等^[14]利用小世界网络和无标度网络的拓扑结构重新构造了储备池的结构,增强了回声状态的特性. 但是,这也增加了网络的层数和复杂度. 这些文献都没有很好地解决储备池中神经元的随机连接问题,改善网络的预测精度.

基于上述分析,为了满足网络流量预测的实时性要求,提高网络的预测性能,提出了一种基于 Dropout 方法的 ESN (Dropout ESN). 将探讨 ESN 的储备池结构的优化问题,把 Dropout 方法引入到 ESN 中,对储备池结构进行优化,并且用 Dropout ESN 预测网络流量序列.

1 Dropout ESN 描述

为了更好地描述所提的改进的 ESN,首先介绍 Dropout 方法的相关知识,然后介绍接受者操作特性 (ROC, receiver operating characteristic) 曲线的相关知识.

1.1 Dropout 方法的介绍

1) Dropout 方法的提出

在 2012 年,研究人员在论文中用到了 Dropout 方法,该方法可以防止过拟合^[10]. Dropout 方法在神经网络的训练过程中,按照一定的概率将神经元暂时从网络中丢弃. 通俗地讲,在前向传播的时候,让某个神经元的激活值以一定的概率停止工作,这样可以使网络泛化性更强.

2) Dropout 方法在神经网络中的使用

在训练阶段,在训练网络的每个神经元都要添加一道概率流程. 在代码层面上,网络让某个神经元以概率 p 停止工作,就是让它的激活函数值以概率 p 变为 0. 例如,某一层网络的神经元个数为 1 000,其激活函数输出值为 $y_1, y_2, \dots, y_{1000}$, Dropout 的概率 p 为 0.4,那么这一层神经元经过 Dropout 后,大约会有 400 个神经元的值被置为 0. 屏蔽掉某些神经元,使其激活值为 0 以后,还需要对向量 $y_1, y_2, \dots, y_{1000}$ 进行缩放,也就是乘以 $1/(1-p)$. 如果训练时经过置 0 后,没有对 $y_1, y_2, \dots, y_{1000}$ 进行缩放,那么在测试的时候,就需要对权重进行缩放.

在测试阶段预测时,每一个神经元的权重 W 要

乘以概率 p ,如图 1 所示,其中,圆表示神经元,箭头表示神经突触, W 表示突触权重矩阵.

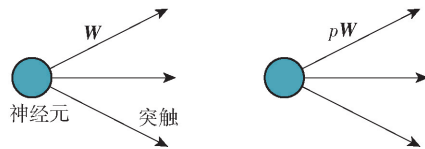


图 1 Dropout 方法的操作

1.2 ROC 曲线

ROC 曲线是机器学习中二分类常用的评价指标. 当测试集中的正负样本的分布变化的时候, ROC 曲线能够保持不变. 因为,在数据集中会存在类不平衡的情况,即负样本远远多于正样本(或者相反),在测试的数据中,随着时间的变化,正负样本的分布也可能不断变化.

1.3 网络总体框架

ESN 的核心结构是一个随机生成且保持不变的储备池. 其内部的动态储备池包含了大量稀疏连接的神经元,蕴含系统的运行状态,并具有短期记忆功能. 储备池的规模与神经元的个数有关. 储备池规模越大,神经元个数越多,越有助于 ESN 对给定动态系统的描述. 但是,储备池中神经元的随机连接会影响网络的预测性能.

为了解决上述问题,对经典的 ESN 进行了改进,在储备池中加入了 Dropout 方法,提出 Dropout ESN. 将 Dropout ESN 应用到实际的网络流量预测任务中,设置了 Dropout ESN 的储备池内神经元以不同的概率停止工作,将 Dropout ESN 的预测性能和其他神经网络进行了对比分析.

图 2 所示为 Dropout ESN,其中 W^{in} , W^{out} 和 W^{back} 分别为输入、输出和反馈连接的权重矩阵,虚线圆为储备池中被随机屏蔽掉的神经元. 与经典的 ESN 不同的是,该网络在储备池中随机屏蔽掉部分神经元,神经元以一定的概率停止工作.

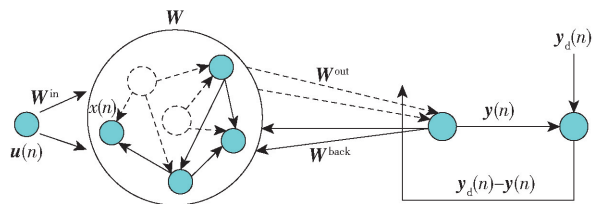


图 2 Dropout ESN 的结构

在采样时间 n , Dropout ESN 的更新方程和输出方程为

$$\mathbf{x}(n+1) = \tanh(\mathbf{W}^{\text{in}}\mathbf{u}(n+1) + \mathbf{W}^*\mathbf{x}(n) + \mathbf{v}(n)) \quad (1)$$

$$\mathbf{y}(n+1) = \mathbf{W}^{\text{out}}\mathbf{x}(n+1) \quad (2)$$

其中: $\mathbf{u}(n+1)$ 为输入向量; \mathbf{W}^* 为在训练时储备池的内部连接权重矩阵, $\mathbf{W}^* = p\mathbf{W}$, $\mathbf{v}(n)$ 为噪声矢量; $\tanh(\cdot)$ 为激活函数. 式(2)为单输出网络的输出公式,输出神经元的连接权重 \mathbf{W}^{out} 为 $l \times r$ 维矩阵.

1.4 稳定性分析

当使用 Dropout 方法使储备池中神经元的激活值以一定概率停止工作后,储备池依然具备回声状态特性. 当储备池中的内部连接权重矩阵 \mathbf{W}^* 的最大奇异值小于1时,即 $\sigma(\mathbf{W}^*) = \|\mathbf{W}^*\| < 1$,储备池具备回声状态特性. 该特性的存在性可以通过储备池的状态来验证^[15].

若 $\mathbf{x}(n+1)$ 和 $\mathbf{x}'(n+1)$ 是储备池的2个状态向量,则

$$\begin{aligned} & \|\mathbf{x}(n+1) - \mathbf{x}'(n+1)\|_2 = \\ & \|\tanh(\mathbf{W}^*\mathbf{x}(n) + \mathbf{W}^{\text{back}}\mathbf{y}(n) + \mathbf{v}(n)) - \\ & \tanh(\mathbf{W}^*\mathbf{x}'(n) + \mathbf{W}^{\text{back}}\mathbf{y}(n) + \mathbf{v}(n))\|_2 \leq \\ & \|\mathbf{W}^*\mathbf{x}(n) + \mathbf{W}^{\text{back}}\mathbf{y}(n) + \mathbf{v}(n) - \\ & (\mathbf{W}^*\mathbf{x}'(n) + \mathbf{W}^{\text{back}}\mathbf{y}(n) + \mathbf{v}(n))\|_2 = \\ & \|\mathbf{W}^*\mathbf{x}(n) - \mathbf{W}^*\mathbf{x}'(n)\|_2 \leq \\ & \|\mathbf{W}^*\|_2 \|\mathbf{x}(n) - \mathbf{x}'(n)\|_2 \end{aligned}$$

由此得出,当储备池中的权重矩阵 \mathbf{W}^* 的最大奇异值小于1时,储备池的神经元状态趋近稳定,使用 Dropout 方法使储备池中的神经元的激活值以一定概率停止工作后,储备池依然具备回声状态特性.

2 数值模拟仿真实验

2.1 实验设置

实验所需的仿真工具为 JetBrains PyCharm 2019.3,采用实际的网络流量数据进行仿真分析. 数据集来自英国学术骨干网^[16],共计19 888个数据点. 该网络流量数据的采样间隔较长、周期较长、随时间的变化较为平缓.

2.2 评价标准

在对流量未来的走势进行预测时,产生的数值必然与实际的数值存在误差,称之为预测误差. 为了评估预测性能,采用正规化方均根差(NRMSE, normalized root mean square error)来度量预测精度,其数学描述为

$$R = \left\{ \sum_{j=1}^k [d_j(m) - y_j(m)]^2 / k\sigma^2 \right\}^{1/2} \quad (3)$$

其中: k 为测试样本的长度, d_j 和 y_j 分别为在测试阶段实际的测试输出和期望输出(网络的预测值), m 为实验中数据的个数, σ^2 为期望输出的方差.

2.3 实验结果

将 Dropout ESN 应用于网络流量的预测,通过调整储备池内神经元的连接概率,从预测性能、NRMSE 预测精度和 ROC 曲线 3 个方面分析了 Dropout ESN 的预测能力.

1) 预测性能分析

图3所示为经典 ESN 和 Dropout ESN 对网络流量数据预测的对比,其中,Dropout ESN 将储备池中的神经元,按照0.5的概率暂时丢弃(在多数实验的情况下,概率取值为0.5). 由图3可知,在预测1 000步的时候,Dropout ESN 预测的曲线更接近于真实的网络流量样本数据. 因此,Dropout ESN 的预测效果优于经典 ESN.

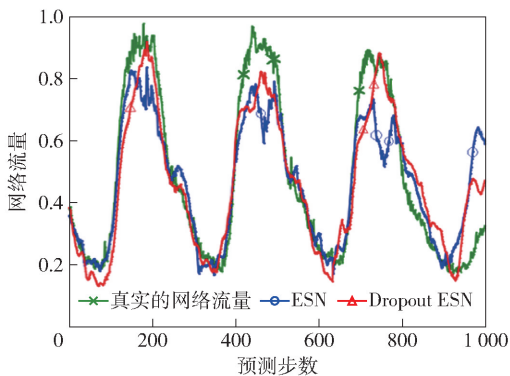


图3 网络流量数据预测对比图

2) NRMSE 预测精度分析

利用 NRMSE 指标来分析 Dropout ESN 在网络流量中的预测性能. 图4所示为预测步数与 NRMSE 的关系,对比了 Dropout ESN 和经典 ESN. 其中,Dropout ESN 储备池中的神经元以0.4的概率停止工作,储备池中的神经元是以0.6的概率进行连接的,而经典 ESN 储备池中的神经元是随机连接的. 由图4可知,在预测的前17步,Dropout ESN 的 NRMSE 高于经典 ESN;在17步以后,Dropout ESN 的 NRMSE 低于经典 ESN. 从总体趋势看,Dropout ESN 的 NRMSE 低于经典 ESN. 由此可以得出,Dropout ESN 的预测误差低于经典 ESN 的预测误差.

下面,将 Dropout ESN 与 BP 神经网络(BPNN, back propagation neural network)、支持向量回归(SVR, support vector regression)和小波分解结合回声状态网络(W-ESN, wavelet decomposition and echo

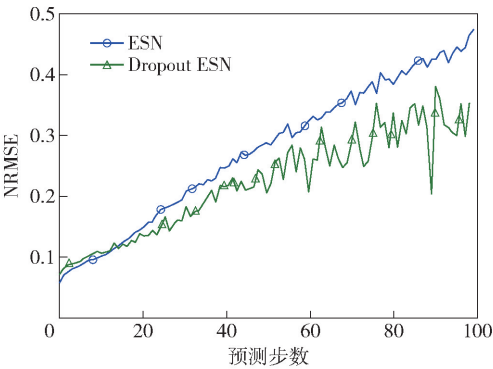


图 4 NRMSE 与预测步数的关系

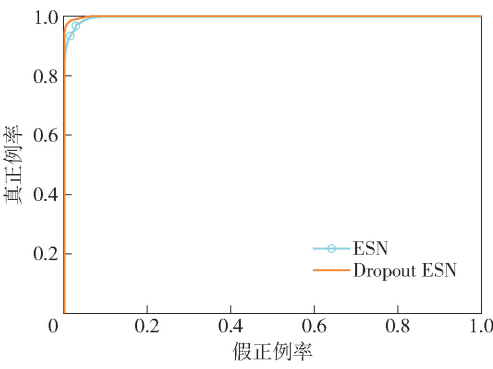


图 5 ROC 曲线

state networks)^[17]进行 NRMSE 的对比和分析,如表 1 所示.

表 1 Dropout ESN 与其他网络的 NRMSE 对比

网络	NRMSE 值
BPNN	0.149 5
SVR	0.157 4
W-ESN	0.071 3
Dropout ESN	0.059 1

由表 1 可知,Dropout ESN 对网络流量数据预测结果的 NRMSE 值最小,为 0.059 1. 由此可得,Dropout ESN 对网络流量的预测精度较其他三者更优.

3) ROC 曲线分析

在神经网络的预测中,将预测值是否在允许的误差范围内看成一个二分类情况. 下面,将利用二分类常用的评价指标 ROC 曲线分析 Dropout ESN 的预测性能. ROC 曲线越接近左上角,该网络的性能越好.

当 Dropout ESN 储备池中的神经元以概率 0.5 停止工作时,经典 ESN 和 Dropout ESN 的 ROC 曲线如图 5 所示. 这 2 条曲线都十分接近左上角. 但是,在纵坐标取值范围为 0.8 ~ 1.0 时,Dropout ESN 的 ROC 曲线比经典 ESN 的 ROC 曲线更接近于左上角,这说明相对于经典 ESN 而言,Dropout ESN 可以将数值进行更准确的分类,在预测时生成的数据跟样本数据更接近,预测精度也更高.

3 结束语

针对经典的 ESN 储备池中神经元连接结果的问题,提出了一种 Dropout ESN. 通过降低储备池内

神经元的连接度,优化了储备池结构. 将 Dropout ESN 应用到实际的网络流量预测任务中,调整 Dropout ESN 的储备池内参数,对储备池内神经元的连接度进行仿真,从 3 个方面对比分析了 Dropout ESN 对网络流量的预测能力. 实验结果表明了当 Dropout ESN 储备池中的神经元以一定的概率停止工作时,Dropout ESN 预测精度较 ESN 更高.

参考文献:

[1] 王瑞雪. 基于机器学习方法的网络流量感知与预测研究[D]. 无锡: 江南大学, 2013: 2-16.

[2] Zhou Dingding, Chen Songlin, Dong Shi. Network traffic prediction based on ARFIMA model[J]. International Journal of Computer Science Issues, 2012, 9(6): 84-87.

[3] Gao Feng. Network traffic prediction based on neural network[C]//2015 International Conference on Intelligent Transportation, Big Data and Smart City. Halong Bay: IEEE, 2016: 527-530.

[4] Hu Shenmin, Xu Weisheng, Wang Zhongjie, et al. Network traffic modeling and prediction based on fractional differencing and fuzzy-AR[J]. Computer Engineering and Applications, 2006, 42(19): 104-107.

[5] Periyamayagi S, Sumathy V. S-ARMA model for network traffic prediction in wireless sensor networks[J]. Journal of Theoretical and Applied Information Technology, 2014, 60(3): 524-530.

[6] Wang Jin. A process level network traffic prediction algorithm based on ARIMA model in smart substation[C]//2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC 2013). Kunming: IEEE, 2013: 1-5.

- nique in SIC and PIC receiver[C]//International Conference on Circuit, Power and Computing Technologies. Kollam: IEEE Press, 2017: 1-6.
- [6] Li Kai, Wang Jue, Xu Chen. Design of PDMA communication system based on OQAM-OFDM [C] // IEEE Advanced Information Management Communicates, Electronic and Automation Control Conference. Xi'an: IEEE Press, 2018: 2098-2104.
- [7] Peng Mingyao, Zeng Jie, Su Xin, et al. Power allocation in downlink PDMA systems[C]//IEEE Global Communications Conference. Abu Dhabi: IEEE Press, 2018: 1-6.
- [8] Liu Xiqing, Chen Hsiaohwa, Chen Shuyi, et al. Symbol cyclic-shift equalization algorithm: a CP-free OFDM/OFDMA system design[J]. IEEE Transactions on Vehicular Technology, 2017, 66(1): 282-294.
- [9] Bomfin R, Chafii M, Fettweis G. A novel iterative receiver design for CP-Free transmission under frequency-selective channels[J]. IEEE Communications Letters, 2020, 24(3): 525-529.
- [10] Zhang Jing, Wen Chaokai, Jin Shi, et al. Artificial intelligence-aided receiver for a CP-Free OFDM system: design, simulation, and experimental test[J]. IEEE Access, 2019, 7: 58901-58914.
- [11] Kim Dukhyun, Stuber G L. Residual ISI cancellation for OFDM with applications to HDTV broadcasting[J]. IEEE Journal on Selected Areas in Communications, 1998, 16(8): 1590-1599.
- [12] Andrea Goldsmith. Wireless communications[M]. London: Cambridge University Press, 2005: 214.
- [13] Wang Xianbin, Ho Paul, Wu Yiyang. Robust channel estimation and ISI cancellation for OFDM systems with suppressed features[J]. IEEE Journal on Selected Areas in Communications, 2005, 23(5): 963-972.

(上接第13页)

- [7] Marquez P, Pinos D, Juan I O. Performance comparison in network traffic prediction for polynomial regression to PIP versus ARIMA and MWM [C]//2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). Moscow: IEEE, 2018: 77-82.
- [8] Wei Yongtao, Wang Jinkuan, Wang Cuirong. Network traffic prediction based on wavelet transform and season ARIMA model [C]//International Symposium on Neural Networks. Berlin: Springer, 2011: 152-159.
- [9] Dethe C G, Wakde D G. On the prediction of packet process in network traffic using FARIMA time-series model [J]. Journal of the Indian Institute of Science, 2004, 84310(2): 31-39.
- [10] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [11] Strauss T, Wustlich W, Labahn R. Design strategies for weight matrices of echo statenetworks [J]. Neural Computation, 2012, 24(12): 3246-3276.
- [12] Bianchi F M, Livi L, Alippi C. Investigating echo-state networks dynamics by means of recurrence analysis [J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(2): 427-439.
- [13] Ozturk M C, Xu D, Principe J C. Analysis and design of echo state networks [J]. Neural Computation, 2007, 19(1): 111-138.
- [14] Deng Zhidong, Zhang Yi. Collective behavior of a small-world recurrent neural system with scale-free distribution [J]. IEEE Transactions on Neural Networks, 2007, 18(5): 1364-1375.
- [15] Jaeger H, Haas H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication[J]. Science, 2004, 304(5667): 78-80.
- [16] British Academic Backbone Network. Internet traffic time series datasets [2021-02-21]. <http://www3.dsi.nminho.pt/pcortez/series/>.
- [17] 吴佳东. 基于回声状态网络的网络流量预测研究[D]. 兰州: 兰州大学, 2016: 37.