

摘要

拓扑可迁移的复杂网络动力学方程学习方法

现实世界中存在着多种多样的复杂系统，如生态系统，社交网络等，通常利用微分方程对这些复杂网络上的动态进行建模来表达各种复杂现象。然而当前相当多的复杂系统的底层动态尚不清楚。因此，开发一种数据驱动的复杂网络动力学学习方法，理解复杂网络动力学底层原理，是对其实施干预和控制的第一步，在气候灾害预测、阻断传染病传播等挑战中具有相当重要的现实意义。随着深度学习研究的普及，处理图数据相关的深度学习方法得到充分发展，图神经网络以其卓越的性能成为当前研究热点，以图神经网络为基础，建模复杂网络上的动力学为当前复杂系统研究开辟一条新的研究方向。

基于以上出发点，本文的主要工作如下：

(1) 本文提出动力学预测任务应该满足的三个测试准则。通过分析发现当前复杂网络动力学建模相关研究暴露出的一个误解，即图神经网络中将节点特征嵌入潜在空间的思想，被许多复杂网络上的神经动力学模型自然采用。然而高维嵌入虽使模型很好的拟合了观测数据，但没有正确的学习到表达观测数据的动力学原理。现有的研究局限在动力学动态的短期预测任务中，在高维空间中拟合能够提供很好的预测结果，但在长期预测任务中，迅速崩溃。针对这些问题，本文提出了三个测试准则，并通过经验分析结果展示了基于嵌入的模型不能通过三个测试准则，并分析嵌入模型失败的原因，尤其以拓扑共轭的视角进行理论分析。

(2) 通过证明不合理的嵌入导致的这些挑战，本文提出了一种基于参数化两个向量场分量的简单无嵌入建模方法，即复杂网络神经动力学模型（Dy-Net Neural Dynamics）。通过实验验证了本文模型能够可靠的学习到不同拓扑结构上的一类动力学现象。因其直接在物理空间中建模观测数据，模型的学习参数与拓扑无关，天然具有拓扑迁移能力。模型分别捕获自动力学和交互动力学，通过引入符号回归能将两个向量场训练的模型参数分别转化为直观的数学表达式，增强模型的可解释性。相对于现有工作提出的基于数据驱动的复杂网络神经动力学模型具有明显的性能优势和可靠性。

关键词：

图神经网络，复杂网络动力学，神经动力学，拓扑迁移

Abstract

Topologically transferable method for learning complex network dynamics equations

There are a variety of complex systems in the real world, such as ecosystems, social networks, etc. The dynamics on these complex networks are usually modelled using differential equations to represent various complex phenomena. However, the underlying dynamics of a significant number of complex systems are currently unknown. Therefore, developing a data-driven learning method to understand the underlying principles of complex network dynamics is the first step towards implementing interventions and controls on them, and is of considerable practical importance in challenges such as climate disaster prediction and interrupting the spread of infectious diseases. With the popularity of deep learning research, deep learning methods related to processing graph data have been fully developed, and graph neural networks have become a current research hotspot with their excellent performance. Modelling the dynamics on complex networks based on graph neural networks opens up a new research direction for current complex systems research.

Based on the above issues, the main work of this paper is as follows:

(1) This paper finds that current research related to modelling the dynamics of complex networks exposes an under-recognised misconception that the idea of embedding node features into potential spaces in graph neural networks is naturally adopted by many neural dynamic models on complex networks. However, while the high-dimensional embedding allows the model to fit the observed data well, it does not properly learn the dynamics that express the observed data. Previous research has been limited to short-term prediction tasks of dynamical dynamics, where fitting in high-dimensional space provides good predictions, but rapidly breaks down in long-term prediction tasks. To address these issues, this paper proposes three test criteria that dynamical prediction tasks should satisfy, demonstrates through empirical analysis of the results that embedding-based models fail the three test criteria, and

analyses the reasons for the failure of embedded models, particularly in a topological conjugate perspective for theoretical analysis.

(2) By demonstrating that unreasonable embeddings lead to these challenges, this paper proposes a simple embedding-free modelling approach based on parameterising two vector field components, namely the complex network neural dynamics model (Dy-Net Neural Dynamics). It is experimentally verified that the model in this paper is able to reliably learn a class of dynamical phenomena on different topologies. As it models observations directly in physical space, the learning parameters of the model are topology-independent and naturally have topological migration capabilities. The models capture autodynamics and interaction dynamics separately, and the model parameters trained in each of the two vector fields can be transformed into intuitive mathematical expressions by introducing symbolic regression, enhancing the interpretability of the models. This provides significant performance advantages and reliability over previously proposed data-driven neurodynamic models for complex networks.

Keywords:

Graph Neural Networks, Complex Network Dynamics, Neural Dynamics, Topological Transfer

关于学位论文使用授权的声明

本人完全了解吉林大学有关保留、使用学位论文的规定，同意吉林大学保留或向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅；本人授权吉林大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或其他复制手段保存论文和汇编本学位论文。

论文级别： ☒ 硕士 ☐ 博士

学科专业： 计算机技术

论文题目： 拓扑可迁移的复杂网络动力学方程学习方法

作者签名： 刘冰

指导教师签名： 黄品

2022 年 5 月 22 日

目 录

目 录.....	I
第 1 章 绪 论.....	1
1.1 研究意义	1
1.2 研究现状	2
1.3 研究内容	5
1.4 本文组织结构	6
第 2 章 相关工作	8
2.1 图网络	8
2.1.1 基于频谱的图神经网络（Spectral-based GCN ^[20] ）	8
2.1.2 基于空间的图神经网络（Spatial-based GCN）	9
2.1.3 图关系归纳偏差	9
2.2 神经微分方程	10
2.2.1 神经常微分方程（Neural ODE）	10
2.2.2 GRAND 图神经扩散.....	12
2.2.3 复杂网络上的神经动力学	13
2.3 符号回归	14
2.3.1 偏微分方程的数据驱动发现	14

2.3.2 复杂网络动态推断	15
2.4 本章小结	16
第 3 章 拓扑可迁移的复杂网络动力学方程学习	17
3.1 问题定义	17
3.2 问题分析	18
3.2.1 基于编码器-解码器的神经网络动力学模型.....	18
3.2.2 神经动力学模型的三个准则	19
3.2.3 基于编码器-解码器的神经动力学模型违背准则经验分析	20
3.2.4 基于编码器-解码器的神经动力学模型违背准则原因剖析	22
3.3 算法设计	24
3.3.1 置换不变性聚合替代图拉普拉斯算子	24
3.3.2 复杂网络神经动力学模型（Dy-Net Neural Dynamics）	25
3.3.3 仿射正则化	26
3.3.4 损失预热策略设计	26
3.4 复杂网络神经动力学模型理论分析	27
3.5 本章小结	29
第 4 章 实验与分析	30
4.1 实验数据	30
4.1.1 复杂网络描述	30

4.1.2 动力学方程描述	31
4.1.3 仿真数据生成及数据集划分	32
4.2 对比算法	32
4.3 实验结果与分析	32
4.3.1 训练数据外泛化实验结果与分析	32
4.3.2 初始值迁移实验结果与分析	36
4.3.3 拓扑迁移实验结果与分析	37
4.3.4 动力学学习稳定性和收敛性实验	39
4.3.5 复杂网络上的动力学方程重建实验结果与分析	39
4.3.6 损失预热策略消融实验结果与分析	43
4.4 本章小结	45
第 5 章 总结与展望	46
5.1 工作总结	46
5.2 工作展望	47
参考文献.....	48
作者简介及科研成果	53
致 谢.....	54

第1章 绪论

1.1 研究意义

人类从未停止对复杂系统的研究与探索,在复杂网络这个工具尚未被人类掌握之前,人类就在探寻着控制这个世界的机制,例如通过对行星数据的观察与分析从而发现万有引力定律就是一个很好的证明。在探索系统动力学规律的征途中,微分方程的提出极大的增强了人类对连续时间系统的表达能力,优美的洛伦兹蝴蝶曲线作为发现的第一个混沌方程,打开了人类认识混沌世界的一扇大门。近年来,深度学习成为当前研究热点,深度学习从语音识别、文本理解和图像识别等数据丰富的应用中学习潜在表示,在预测性能方面取得重大成功。将网络和深度学习结合起来使用去理解复杂系统的动态演化,成为当前新的研究热点。

网络系统通常是由许多单元组成的复杂系统,其研究范围覆盖极广,包含生态系统、物理学、经济学、互联网和社会学等^{[1][2][3]}。这些单元通过特定的,可能变化的拓扑结构耦合在一起,网络动力学相关研究指网络系统中单元状态随时间变化的相关问题。现实世界的复杂系统,如大脑、生态系统、基因调控、社交网络生态网络^[4]等,通常被建模为受一些潜在非线性动力学支配的复杂网络上的动态变化。利用微分方程对这些复杂网络上的动态进行建模对于揭示自然界中的复杂系统演化动力学至关重要。尽管非线性动力系统理论已在不同领域得到广泛研究,包括应用数学、统计物理学、工程学和生态学等,但这些开发的模型通常基于对网络演化机制的清晰认识,因此通常被称为规则模型,如对疾病传播建模的 SIR 模型等。规则模型为了满足计算上的要求,通常抽象为简约的模型,假设条件较强,难以应对实际中存在的复杂现象。

已有基于神经网络的数据驱动方法对连续时间动力学数据进行建模,例如 Charles Murphy 等人提出的 GNN-RNN 模型框架^[5],通过对西班牙 COVID-19 的传播情况进行学习,取得了较好的预测结果。各种基于深度学习的模型被设计用于气候和天气预测,包括近地表空气温度预测、空气质量推断、降水预测等。天气由空间和时间数据组成,时空动力学模型可以对其进行建模和预测。DPGN^[6]设计为一种物理知识学习的架构来预测空气温度,其将可微物理方程与图神经网络

络框架相结合,并利用正则项约束将领域专家通过潜在空间中给出的隐含物理知识结合起来,使气候预测任务得到显著改进。

随着神经常微分方程的研究兴起,出现了将图神经网络与神经微分方程结合的相关研究,比如 NDCN^[7]是一个数据驱动模型,能够从给定的时空数据中预测复杂网络上动力学的演化趋势,但是经过本文分析,此类模型并非学习到真实的动力学演化规律,仅是拟合了当前观测数据,在长距离预测任务中即失效。但这仍是数据驱动连续神经网络模型在复杂网络动力学研究领域的一次突破,能够完成对任意间隔的连续时间序列数据建模,对热扩散、基因调控和生态系统等动力学进行短期预测,其对从复杂网络上的时间序列数据中学习动力学规则指引了一个方向。

鉴于现实世界的复杂性,仍然存在大量复杂网络,其底层动力学尚不清楚或可能无法通过显式数学函数建模。同时这些网络在演化过程中通常会产生海量数据,从观察时间序列数据中学习到此类动力系统的控制机制是理解复杂网络动力学特性并最终实施干预和控制的第一步。因此,现代数据驱动的方法在复杂网络上的动态学习任务中具有重要的研究意义。

1.2 研究现状

提出一种成功的数据驱动方法来对复杂网络上的连续时间动力学进行建模非常具有挑战性:现实世界网络中节点的交互结构复杂,节点和边的数量很繁杂,这被称为复杂系统的高维性。控制复杂网络中节点状态动态变化的规则在时间上是连续的并且是非线性的。系统内的结构和动态依赖性很难通过明确的数学表达式来进行建模。

早期动态反应扩散过程和元种群模型是对各种现象的标准建模方法^[8],在这些现象中,局部量如粒子根据物理定律扩散和相互作用。反应扩散过程能够用于模拟多种现象,如化学反应、种群进化、流行病传播和许多其他空间分布系统。具有异构拓扑结构且对节点占用数没有限制的网络上定义的两个基本反应扩散过程行为可以表达为 $B \rightarrow A$ 和 $A + B \rightarrow 2B$ 。元种群流行病模型即可以看作扩散过程的一个实例,其中粒子代表人们在不同地点之间移动,人群被分为多种类别以表达个体在建模疾病方面的状态,例如易感、感染、免疫等,反应过程说明了同

一地点的个人可能接触并根据情况改变他们状态的可能感染动态。基于 SI 模型，路由策略对基于交通流的流行病传播动力学的影响得到一些研究者的关注^[9]。这些方法通过抽象，将复杂的现象抽象为简单的易处理的表达式形式，进而对动态进行预测。

COVID-19 爆发以来，复杂动力学系统的预测与人类健康息息相关，预测传染病动力学的演变仍然是一个悬而未决的问题，传统规则模型只能提供部分答案如 SIR 模型，为了满足数学和计算上的易处理性，这些模型通常依赖于简化假设，从而限制了预测的定量准确性和可建模动态的复杂性。一类基于 GNN+RNN 的深度学习方法被提出^{[5][10]}，可以有效地从时间序列数据中学习到复杂网络动力学，此类模型采用图神经网络架构对动力学建模，假设要求很少，可以对任意网络结构上的动态进行模拟，甚至探索训练数据之外的动力学特性，在阻断社会网络传染病传播过程中有助于设计更好的公共卫生政策。基于 GNN 的时空序列预测方法在电力和工业数据预测任务中也得到应用，能够保证实时性和准确性^[11]。但此类数据驱动模型对数据量要求较高，当数据的拓扑结构或底层动力学规律发生变化时，模型无法进行迁移应用。

在典型的问题设置中，多元时间序列数据 $x(t_1), x(t_2), \dots, x(t_N)$ 是从网络动力系统中观察到的，目标是从中归纳推断动力学系统来建模 $x(t)$ 。当时间序列数据时间间隔均匀时，存在多种解决方案如文献[12][13]。其中 SINDy^[12]方法通过构建候选基函数库，通过差分等数值方法求得时间序列的导数，而后基于稀疏回归的方法进行基函数项的选择，求得一个简洁稀疏的偏微分方程表达式，具备可解释性的来对多种非线性动力学进行建模，但是其也存在固有问题，一般非线性动力系统只有很少的交互项组成，在采用回归方法求解系数时，组合复杂度随着构建基函数库中基项的数量增加而增加，在由数万个基项和交互项组成的大型网络上能否有效的学习连续时间动态在很大程度上仍然是未知的，同时基函数库的构建严重依赖于先验知识，假设条件较强。

近年来，随着捕获复杂网络系统结构和行为的经验数据的可用性大大增加，将稀疏回归方法推广到复杂网络数据上的两阶段方法^[14]被提出，该方法可以看作是 SINDy 方法在复杂网络上学习非线性动力学的应用推广，基于网络上节点的自动动力学和交互动力学解耦的通用表达式，分别构建自动动力学基项和交互动力

学基项,进而通过回归方法进行有效项的筛选。两阶段方法,通过归一化消除数据方差大带来的影响、引入修正的赤池信息量准则(wAIC)来指导稀疏项的选择,在网络动力学学习任务中取得较好的结果,这些策略在一定程度上克服了数据不完整,存在噪声等问题,但其同样存在 SINDy 面临的固有问题,并且对数据密度同样具有要求,时间间距太大的稀疏数据以及假设基库中存在缺失项等问题会导致出乎意料的结果。为有效解决时滞以及噪声等因素对复杂网络动态系统的同步控制,一种基于 LaSalle 型不变性原理的自适应控制器被提出,并借助 Lyapunov 稳定性理论给出复杂网络与外部系统同步的基本条件^[15]。

当前由于许多复杂系统观察到的时间序列数据观测时间间隔不均匀,基于 GNN-RNN 组合模型的深度学习方法对时间间隔具有较强的假设依赖,在面对非等间隔时间序列数据时束手无策。而两阶段方法等稀疏回归方法在求解导数时通过差分等数值方法进行求解,不均等间隔将会带来极大的计算误差,进而造成结果的不可靠性。因此学习复杂网络动态仍然是一个挑战。

随着 Ricky T. Q. Chen 等人基于残差网络提出的神经常微分方程模型^[16],将离散的神经网络层级连续化,使用神经网络参数化隐藏状态的导数,引发了对连续神经网络研究的热潮。数据驱动发现常微分方程(ODE)或偏微分方程(PDE)来捕获连续时间动态的相关研究不断涌现。已有将连续的神经网络和图神经网络相结合提出连续层级的图神经网络,如 GRAND^[17]、GDE^[18]等。通过将图神经网络层级连续化,自适应层级克服了当前图神经网络研究过程中存在的深层过平滑等问题。连续化的层级和物理意义中的时间概念相对应,自然的可以用于连续时间轨迹预测,例如 GDE 不仅可以解决图数据中节点分类任务,也可以用于解决交通流量轨迹预测等任务,天然的适用于任意时间间隔的任务中。将图神经网络和神经常微分模型结合,非常适合作为黑盒模型解决复杂网络动力学相关研究问题。

数据驱动的复杂网络动力学模型 NDCN^[7]的出现成功的解决了上述面临的复杂系统观测时间序列时间间隔不均等的问题。NDCN 提供了一个解决复杂网络动力学相关研究的框架,通过将原物理空间状态映射到高维空间中,在高维空间中进行图拉普拉斯扩散操作,最后映射回原空间进行网络动态的学习,其中的扩散操作可以自适应定义,将当前各种高效的图神经网络模型嵌入其中,以适应各

种复杂的动力学任务，获得更优结果。该模型也存在基于频谱空间的图神经网络的固有特性，即无法将拓扑结构和训练参数分离，扩散过程是在高维空间中进行的，表示图结构的邻接矩阵信息不可避免的影响了参数的学习过程，导致训练的模型只能在当前图结构上有效，已训练模型的泛化能力较差。

1.3 研究内容

在深度学习的相关研究中，通过证明能够成功解决任务的架构设计模式，通常会被迁移应用到解决相关任务中。编码器-解码器结构在几乎所有有竞争力的模型中都有使用，无论是多头注意力机制模型（transformer）架构^[19]还是图神经网络模型（graph neural network）^[20]。最近神经动力学模型^{[16][17]}成为热点研究，人们自然认为编码器-解码器结构迁移应用到神经动力学模型中也是有帮助的和必要的。本文在从时间序列数据中学习复杂网络动力学的背景下，挑战编码器-解码器对复杂网络动力学学习任务上的适用性假设，提出一种在原物理状态空间中学习的拓扑可迁移的复杂网络动力学学习方法。

最新研究利用编码器-解码器结构从时间序列数据中学习耦合动态^{[7][21][22][23][24]}，其中最具代表性的模型是 NDCN^[7]，如本文 3.2.1 节分析，NDCN 不是从原始观察数据 $\mathbf{x}$ 中学习流 Φ ，而是寻求同时学习编码器 f_e 和共轭流 Φ' 使得 $\Phi'(t_i, f_e(x_0)) \approx (f_e \cdot \Phi)(t_i, x_0)$ 。由于流 Φ' 在高维潜在空间中，可能具有比 Φ 更简单的形式，就像核方法允许将线性算法用于非线性问题一样^[25]。然而，这种看似无害的编码器-解码器结构的使用可能会学习到不正确的动力学模型，如同 3.2.3 节的经验结果验证。特别是，即使训练数据来自稳定观测流，NDCN 估计的流通常也是动态不稳定的（即最大 Lyapunov 指数为正值）。本文通过分析潜在的原因，得出编码器是不必要的结论，本文通过提出一种有效并且通用的无嵌入架构，直接在原始状态空间中对动态进行建模是可能的，也是更加可取的。同时由于学习到的是自动力学和交互动力学各自的部分，在原物理状态空间中进行信息传播，学习到的参数与图结构信息无关，在一个数据集上训练的模型参数，在满足动力学底层规则不变的情况下，能够直接迁移到其他初始化和拓扑结构数据中完成连续时间动态预测任务。

本文的主要贡献如下：

(1) 本文揭示了基于编码器-解码器的复杂网络神经动力学模型的严重局限性：多余的潜在维度会导致不正确的流。特别是，本文通过实验验证表明一些广泛使用的模型架构甚至不能产生有效的动力学系统。本文在多种动态和网络上进行实验展示了这种现象。

(2) 本文提出一种无编码器架构 Dy-Net Neural Dynamic (DNND) 模型，该模型能够从时间序列数据中学习复杂网络上的动力学。DNND 模型超越传统的拟合模型，可以可信地学习到复杂网络上的动力学特征，学习到的动力学特征可以迁移到不同拓扑结构的数据中。

(3) 本文提出一个损失预热策略来解决在不规则时间序列数据上训练神经动力学模型时的数值不稳定性。

(4) 本文通过分析非线性复杂动力系统的共同目标：稳定性和约束行为等一致性指标来改进现有的神经动力学模型的评估协议。

1.4 本文组织结构

本文首先阐述所做研究工作的意义与研究现状，而后介绍了与本文研究内容相关的已有工作，接着详细介绍了本文研究提出的拓扑可迁移的复杂网络动力学方程学习模型，并通过实验验证了本文提出方法的有效性，最后对本文的研究内容进行总结与展望。本文组织结构将展开如下：

第一章为绪论，简单介绍了复杂网络动力学研究的现实意义，总结了国内外发展研究现状，初步描述了本文所研究的内容和主要贡献。

第二章是相关工作，以本文的研究内容和研究方法所涉及的相关工作为切入点，首先描述了处理图数据相关的图神经网络算法，然后介绍了当前神经微分方程相关工作，以及将图神经网络和神经微分方程相结合的以处理图上时间序列数据的当前最新黑盒研究工作，最后简单介绍了基于符号回归的处理时间序列数据的具有可解释性的白盒研究工作。

第三章首先对复杂网络动力学学习问题进行定义，给出相应的符号说明；然后对当前最前沿的基于编码器-解码器的神经动力学模型方法进行简单阐述，提出三个神经动力学模型理论上应该满足的数学准则，并通过实验进行验证表明当前基于编码器-解码器的神经动力学模型在面临本文提出的准则存在的困境，然

后对当前模型违背相应准则的深层原因进行剖析；最后针对复杂网络上的连续时间序列数据学习任务提出本文的复杂网络神经动力学模型（**Dy-net Neural Dynamics**），并阐述了设计模型过程中提出的一些策略：置换不变性聚合替代图拉普拉斯算子、仿射正则化以及损失预热策略。

第四章通过充分的实验验证了本文提出的复杂网络神经动力学模型的有效性。包括三种动力学在五种复杂网络上的实验、三种动力学在拓扑结构更换时的泛化性实验、三种动力学在初始值更换时的泛化性实验、三种动力学学习稳定性和收敛性实验验证以及通过回归方法重建复杂网络动力学表达式的结果与分析。

第五章对本文的工作进行总结，并对未来的相关研究工作进行展望。

第2章 相关工作

2.1 图网络

近年来,深度学习极大的改变了许多机器学习领域的任务,从图像和视频的处理到语音识别和对自然语言的理解。当然这些任务中的数据通常在欧几里得空间中表示。然而当前更多的应用数据是从非欧几里得数据域产生的,并表示为对象之间具有复杂关系和相互依赖的图结构。图数据的复杂性对早期机器学习算法的挑战引发了当前对图数据深度学习方法的研究热潮。CNN 在计算机视觉领域取得的成功极大的鼓舞了卷积方法在图数据上的应用。图卷积神经网络主要有两个重要分支,基于频谱的方法和基于空间的方法。

2.1.1 基于频谱的图神经网络 (Spectral-based GCN^[20])

从图信号处理的角度来看,基于频谱的方法通过引入滤波器概念来定义图卷积操作,其中图卷积操作可以从为图信号去除噪声的角度来理解。图信号在处理过程中引入了傅里叶变换,该变换将信号从时域转换到频域分析,进而完成一些在时域空间中难以处理的操作,基于频谱的图卷积神经网络正是借鉴了这种信号变换的核心思想。

在基于频谱的图神经网络中,图通常定义为无向图,用正则化的图拉普拉斯矩阵来表示无向图是数学上鲁棒的一种表达方式。图拉普拉斯矩阵一般定义的形式为 $L = D - A$,其中, L 是拉普拉斯矩阵, D 是顶点的度矩阵, A 是图的邻接矩阵,在文献[20]中拉普拉斯矩阵的定义为 $L^{sys} = D^{-1/2}LD^{-1/2}$ 。正则化的图拉普拉斯矩阵具有一些特殊性质,比如满足实对称半正定的性质,基于该性质,可以分解正则化的拉普拉斯矩阵为 $L = U\Lambda U^T$, L 的特征向量构成矩阵 U , Λ 是对角矩阵。正则化图拉普拉斯矩阵的特征向量在数学上满足正交基的定义。在处理图信号的过程中,图信号可以解释为构成图中所有节点的特征向量。图的傅里叶变换就是将输入的图信号投影到正交空间,在新空间中完成图信号的滤波操作,最后通过傅里叶反变换映射回原空间。基于频谱的图卷积网络都遵循这个范式,不同谱图卷积的核心区别在于选择的滤波器的不同。

基于谱图卷积的图神经网络存在的一个关键问题是,该类方法需要将整个图结构信息加载到内存中从而执行图卷积操作,导致在处理大型图时效率较低。Hamilton 等人提出的 GraphSAGE^[26]展现了一种新的算法框架,该框架基于学习节点的信息是通过聚合当前节点的邻居节点特征而来的基本思想,并提出邻居节点采样方法来缓解大图卷积效率低下的问题。谱图卷积还面临其他限制,如对图数据的任何扰动都会导致节点特征发生变化、学习到的过滤器是域相关的意味着不能用于具有不同结构的图中、以及特征分解的复杂度较高。

2.1.2 基于空间的图神经网络 (Spatial-based GCN)

基于空间的图网络方法将把从邻域聚合特征信息看作图卷积操作,在节点层次执行图卷积操作时,图池化模块与图卷积层相互交错,从细粒度的图结构中提取出粗化的高级结构。该空间图卷积网络从传统卷积网络对图像的卷积操作中受到启发,与传统的图像卷积操作不同,图卷积网络通过对节点的空间关系定义图卷积操作。对于一般的图,基于空间的图卷积通过将邻域内的所有节点特征和当前中心节点表示聚合在一起用来更新当前中心节点的表示。空间图卷积运算本质上是沿网络结构进行节点信息的传递。

为使节点特征消息充分传播,一种通用的处理方式是将多个图卷积层叠加起来,不同的图卷积堆叠方式得到不同类型的空间图卷积网络,通常可以分为两种类型:基于循环的和基于组件的空间图神经网络。基于循环的图神经网络方法使用相同的图卷积层来更新隐藏层表示,基于组件的图神经网络方法则使用不同的图卷积层来更新隐藏层表示,使用固定数量的层在每一层中使用不同的权重来解决循环相互依赖关系。

基于空间的方法能够得到广泛研究与应用,得益于其具有的灵活性、高效率和通用性。典型模型有 Micheli A 提出的第一个针对基于空间的图神经网络 NN4G^[27]。受 NN4G 启发,在保持空间局部性的同时,Micheli A 等人又提出一种具有概率可解释性的图马尔可夫模型 CGMM^[28]。

2.1.3 图关系归纳偏差

近年来,创造性的机器学习架构通常遵循一种设计模式,即通过将基本模块

组合成更加复杂的结构，同时导致计算层次更深。这种按层组合起来的复杂结构引出了一种分层处理的特殊关系归纳偏差。其中的计算是分阶段进行的，通常会导致输入信号中信息之间越来越长的交互作用。图表示可以表达任意的成对关系结构，图上的计算提供了强大的关系归纳偏差。文献[29]提出了一种图网络(GN)框架，该框架定义了用于在图结构上进行关系推理的一类函数，该框架概括和扩展了各种图神经网络，并支持从简单的基础块构建复杂的网络结构。

图网络块作为图网络框架的主要计算单元，将图数据作为输入，而后在结构上执行计算，并返回图作为结果。实体由图的节点表示，关系由边表示，节点由特征向量表示，图级属性由全局属性表示。GN架构的块可通过定制，综合表达所需关系归纳偏差的新架构。每个块包含三个更新函数 ϕ^e, ϕ^v, ϕ^u ，分别更新边、节点和全局图表示；每个块还包含三个聚合函数 $\rho^{e \rightarrow v}, \rho^{e \rightarrow u}, \rho^{v \rightarrow u}$ ，每个 ρ 函数的输入是一个集合，简化表示为聚合信息的单个元素， ρ 函数的一个关键约束就是采用可变数量的参数，同时必须对其输入满足排列不变性，例如逐元素求和、平均值、最大值等。

2.2 神经微分方程

近年来，动力系统和深度学习的结合成为研究热点。特别是，神经微分方程(NDEs)表明神经网络和微分方程具有潜在的一致性。许多流行的神经网络架构，例如残差网络和循环网络，都是离散化的微分方程。神经微分方程通常适用于解决生成问题、动力学系统和时间序列拟合等任务。因此，无论是现代机器学习还是传统数学建模任务都对其很感兴趣。NDEs提供高容量的函数逼近、强大的模型先验空间、处理不规则数据的能力，以及在内存效率和丰富的可用理论等方面都具有优势。

2.2.1 神经常微分方程(Neural ODE)

Ricky T. Q. Chen等人提出了一个新的深度神经网络模型架构Neural ODE，没有通过显示指定出网络结构中隐藏层的离散序列，而是使用神经网络参数化隐藏状态的导数，网络的输出是通过使用黑盒的微分方程求解器得到。连续深度神经网络的一个优势就是内存占用成本恒定，因此可以明确的在数值精度和执行速

度之间取得平衡，这些性质文中通过连续深度的残差网络模型和连续时间隐变量模型进行展现。连续的归一化流是一种通过最大似然进行训练的生成模型，通过连续深度神经网络进行构建，从而不必对数据维度进行排序或分区操作。Neural ODE 可以通过任意 ODE 求解器进行计算，并在训练过程中对反向传播进行扩展，这种训练方式对神经网络的内部训练细节不做要求，从而允许 ODE 在大模型中进行端到端的训练。

目前为止，最常见的神经微分方程是 Neural ODE^[16]：

$$\frac{dy}{dt}(t) = f_{\theta}(t, y(t)) \dots \dots \dots (2.1)$$

$$y(0) = y_0 \dots \dots \dots (2.2)$$

其中 y_0 是任意维度的向量， θ 表示可学习参数的向量， f_{θ} 通常是一些标准的简单神经网络架构，例如前馈神经网络或者卷积网络。

假设 f_{θ} 是利普希茨连续函数，这对于神经网络通常来说都是正确的，Picard 存在定理保证了 Neural ODE 的存在性和唯一性。与非微分方程的模型相比，神经微分方程通常还需要额外的两个限制条件，首先是必须能够获得微分方程的数值解。其次必须能够通过微分方程进行反向传播，以获得其参数 θ 的梯度。文献[16]中已经提出了解决这些问题的方法，并将执行这些任务的代码标准化。

Neural ODE 具有内存效率优势，能得以恒定内存成本进行训练的原因是 ODE 在前向传播过程中不存储任何中间变量，仅通过对 ODE 求解器的输入计算标量损失梯度，在之前，内存限制是训练深度神经网络模型的主要瓶颈。Neural ODE 可以自适应计算，现在 ODE 求解器提供关于近似误差增长的保证，监控误差水平，按照要求的精度水平自适应的动态调整自身的评估方法，从而允许模型评估的成本依问题的复杂性而动态扩展，在训练完成后，根据预测精度要求完成对实时性和功耗的调整。Neural ODE 由于其状态连续的特点，天然的可以用于连续时间序列建模任务中，与需要离散化观察和固定间隔的传统递归神经网络不同，连续定义的动态可以自然地包含在任意时间到达的数据。连续的另一个优势是可以适用于可扩展和可逆的归一化流任务中，通过连续模型构建的一类新的可逆密度模型，避免了归一化流模型构建中的瓶颈，并且可以直接通过最大似然进行训练。

2.2.2 GRAND 图神经扩散

基于消息传递机制的图神经网络在图数据相关任务上取得很大成功，Neural ODE 的出现将离散神经网络推向连续的深层神经网络，结合二者的发展，将图上的深度学习作为一个连续的扩散过程进行处理是无可置疑的发展方向。文献 [17] 提出了图神经扩散模型 (GRAND)，将图上的深度学习作为一个连续的扩散过程进行处理，并将图神经网络视为 PDE 的离散化。GRAND 将层结构和拓扑对应于时间和空间算子的离散化选择，该模型导致了一类广泛的新型 GNN，这些 GNN 能够解决图表示学习模型的常见问题，例如深层 GNN 架构不可行、特征过度平滑^[30]和瓶颈问题^[31]。GRAND 模型成功的关键是相对于数据扰动带来的不稳定性，在隐式和显式离散化方案中都能得到解决。

扩散方程是一个偏微分方程，具有时间和空间分量。在图上任务设置中，前者是连续的，后者是离散的，因此，扩散算子继承了输入图数据中的邻接矩阵结构数据。GRAND 描述的图神经扩散模型是一种源自图扩散形式的新型 GNN 架构：

$$\mathbf{X}(T) = \mathbf{X}(0) + \int_0^T \frac{\partial \mathbf{X}(t)}{\partial t} dt \dots\dots\dots (2.3)$$

$$\mathbf{X}(0) = \phi(\mathbf{X}_{in}) \dots\dots\dots (2.4)$$

该架构可以实现可学习的编码器 ϕ 和解码器函数 ψ ，以及可学习的图扩散过程，

以生成节点嵌入 $\mathbf{Y} = \psi(\mathbf{X}(T))$ ， $\frac{\partial \mathbf{X}(t)}{\partial t}$ 由如下图扩散方程给出：

$$\frac{\partial \mathbf{X}(t)}{\partial t} = \text{div}[G(\mathbf{x}(t), t) \nabla \mathbf{x}(t)] \dots\dots\dots (2.5)$$

不同的 GRAND 架构相当于选择可学习的扩散函数 G 和式 2.5 的时间空间离散化。

时间离散化相当于数值积分方法的选择，可以使用固定的或者自适应步长，可以是显式的或隐式的方法。在这里，连续时间可以看作是离散层的连续化，每一层对应于求解器的一次迭代，使用自适应时间步长求解器时，层数是未定的，根据自适应计算方法自主确定。显式数值积分方法使用残差结构，这些结构来自于严格的数值稳定性的结果。隐式数值积分方法提供了一种权衡深度和宽度（指扩散核空间）的自然方式。空间离散化相当于修改给定的图，或者在没有给定图的情况下构建一个图，并且可以假设数据位于某个特征空间或连续流形上。

GRAND 通过将离散时间和空间算子相结合，提供了图学习架构的新蓝图，

适当的离散化和数值积分方法的选择能够设计出非常深的图神经网络,解决当前离散化模型架构存在的一些问题,并在流行的基准测试如节点分类任务中获得出色的性能。

2.2.3 复杂网络上的神经动力学

复杂网络上的连续时间动力学对于理解、预测和控制科学与工程中的复杂系统至关重要。高维系统结构的组合复杂性、难以捉摸的连续时间非线性动力学及其结构-动力学依赖性导致该任务非常具有挑战性。Chengxi Zang 等人建议结合神经微分方程与图神经网络,以数据驱动的方式学习复杂网络上的连续时间动态,提出了 NDCN^[7]模型,通过 GNN 对微分方程系统进行建模。该模型在前向传播的过程中并非通过传统的离散数量的图神经网络层进行映射,而是在连续时间内以数值积分方式整合 GNN 层,从而在图结构数据上捕获连续时间动态。NDCN 模型可以解释为连续时间的 GNN 模型或者图神经 ODE 模型,连续深度可以解释为连续的物理时间。

NDCN 首先将复杂网络上的节点特征映射到隐空间,在高维隐空间上完成动力学特征的传播,最后映射回原空间。表达式形式如下:

$$\frac{\arg \min}{W(t), \Theta(T)} \mathcal{L} = \int_0^T \mathcal{R}(X, G, W, t) dt + S(Y(X(T), \Theta)) \dots \dots \dots (2.6)$$

$$\text{subject to } X_h(t) = f_e(X(t), W_e), X(0) = X_0$$

$$\frac{dX_h(t)}{dt} = f(X_h, G, W_h, t)$$

$$X(t) = f_d(X_h(t), W_d) \dots \dots \dots (2.7)$$

其中 f_e 将 $X(t)$ 映射到隐空间 $X_h(t)$, f 是图扩散操作函数, f_d 将经过扩散操作后的隐空间状态变量 $X_h(t)$ 映射回原物理状态空间 $X(t)$, 文章通过消融实验证明了状态变量映射到隐空间能够增强模型的表达能力,本文对此有不同见解,详见第三章研究内容。

该模型可以在一个统一的框架内完成连续时间复杂网络上的动力学预测任务、结构化序列预测任务和半监督节点分类任务。同时文中通过实验验证了 NDCN 在捕获复杂网络系统结构和动力学的能力。其是第一个由微分方程系统建模的在复杂网络上学习连续时间动力学的模型,其在结构化序列预测任务中以更少的参数实现了比图神经网络模型更优的精度,其以连续时间动态的角度来完成

节点特征的传播并预测节点的未知标签,与众多图神经网络模型相比表现出非常有竞争力的性能。

2.3 符号回归

从数据中学习控制方程是不同科学和工程领域的核心挑战。数据复杂丰富,但模型往往难以学习,例如在气候预测、神经科学、生态学和流行病等领域。如何从数据中学习可解释的模型,对于充分利用现代社会环境能够采集到的各种数据,提供未来趋势预测等方面具有重要意义。

2.3.1 偏微分方程的数据驱动发现

基于神经网络的方法对充分挖掘数据中隐藏的规律进行建模在近数十年成为研究热点,但是不可解释性成为其发展过程中难以克服的困难。传统的基于符号回归的方法具有可解释性,能够提供可信的建模结果。Brunton S L 等人提出一种将传统的符号回归方法和机器学习与非线性动力学相结合的方法 SINDy^[12],该方法能够从嘈杂的测量数据中发现控制方程。SINDy 唯一假设只有少数几个重要的动力学单项式,学习到的方程在可能的函数空间中是稀疏的,这个假设在适当的基础上适用于许多现实世界中的物理系统。SINDy 使用稀疏回归来确定准确表示数据所需的动态控制方程中的单项式,通过调整稀疏回归的参数可以在准确性和模型复杂性之间取得平衡以避免学习到过度拟合的模型。

SINDy 考虑形式为 $\frac{dx(t)}{dt} = f(x(t))$ 的动力系统,其中 $x(t)$ 是系统在 t 时刻的状态, f 函数定义了一些物理上的约束。当前可以使用的数据是在不同时间点 $t_1, t_2, \dots, t_m$ 所感兴趣的系统状态 $x_1, x_2, \dots, x_n$, 并将其排列成 $n * m$ 维度的矩阵形式 X , 用 $\dot{X}$ 表示其导数形式, $\dot{X}$ 一般是通过差分方法等普通的数值方法获得。事实上,很多有趣的物理系统仅由很少的一些项构成,于是通过合理假定,等式右侧的函数 F 在某个函数空间中存在某种稀疏的表达方式能够充分表示当前测量到的数据。该函数空间 $\Theta(X)$ 一般情况下可以使用常见的函数构成,如多项式函数、三角函数等,这些候选的函数基库项可能并非是最简约的,但是并不影响其用来逼近目标函数。于是,搜索动力系统的数学表达式问题变成了标准的多元回归问

题。即 $\dot{X} = \Theta(X)\Xi + \eta Z$ ，其中 $\Xi = [\xi_1, \xi_2, \dots, \xi_n]$ 是系数矩阵， Z 中的元素表示的是满足独立同分布约束的高斯白噪声。可以采用 LASSO 回归等方法求得系数矩阵，SINDy 通过对数据进行多次最小二乘回归，每次回归之后将低于一定阈值的系数设定为 0，直到收敛停止得到系数矩阵。

SINDy 展示了其强大的表达数据的能力，可以在不假设控制方程形式的情况下从数据中识别非线性动力系统，该方法能够扩展到高维系统，如洛伦兹系统等。同时该方法在测量数据存在噪声的情况下仍然能够适用。

2.3.2 复杂网络动态推断

近年来，复杂网络系统的结构和数据的可用性增加，但从数据中揭示复杂网络系统中各个节点自身及其相互之间交互动态的方法仍然较少，基于图神经网络和神经 ODE 方法的复杂网络动力学发现方法依旧存在可解释性困境，不可迁移等固有难题，如何抽象、转化观测数据以进一步理解复杂系统的底层动力学机制，即自主地推理出隐含的决定复杂系统行为的动态方程，是学界亟需解决的一个关键科学问题。Gang Yan 等人最近提出了一种用于复杂网络动力学自主推理的两阶段推理方法，并通过在各种合成和真实网络上推断神经元、遗传、社会和耦合振荡器动力学证明了该方法的有效性，使用该方法发现的复杂网络上的动力学具有可解释性强，鲁棒性强等优势。

该方法可以看作是 SINDy 方法在复杂网络数据上的扩展，并进一步优化回归过程。具体来说，通过第一阶段全局粗粒度回归，对函数矩阵进行归一化，避免由于函数数量级差导致的误判；从冗余的函数库中发现与系统动力学最相关的多个函数，组成子函数空间。第一步选择的函数组成的常微分方程不具备生成能力，故进行第二步细粒度筛选，通过拓扑采样，利用信息不一致作为推断指标，以得到更加简洁的最优动力学形式，此时得到的动力学表达式具备生成能力。

该两阶段推理方法对不完整和带有噪声的数据具有较强的鲁棒性，即使数据是低分辨率、带有动态噪声的、连边丢失和虚假连边以及动态异质性的。但也存在假设较强，动力学学习结果对数据依赖性较强、对基函数空间的先验设定依赖较强等问题。作者将两阶段推理方法用于推断 H1N1 流感在全球航空公司网络上的早期传播动力学，该推断的动力学方程成功的捕获了 SARS 和 COVID-19 疾病

的传播。该方法为研究大量真实网络系统隐藏的微观机制提供了一条途径。

2.4 本章小结

本章首先介绍了近年来对图数据处理研究最经典的图神经网络相关方法,包括基于频谱和基于空间的图神经网络,以及提出的统一图神经网络算法的 GN 框架。进一步介绍了当前研究热点神经微分方程的典型算法如 Neural ODE , 以及将神经微分方程与图神经网络结合的图神经扩散 (GRAND) 和复杂网络上的神经动力学 (NDCN)。最后介绍了当前从时间序列数据中学习动力学方程的传统符号回归方法,包括学习一般的非线性方程的方法 SINDy 以及将其推广到复杂网络时间序列数据上的两阶段推理方法。

第3章 拓扑可迁移的复杂网络动力学方程学习

3.1 问题定义

定义一个具有 n 个节点的无向图，用一个对称邻接矩阵 A 表示， A 中的每个元素 A_{ij} 定义为：

$$A_{ij} = \begin{cases} 1, & \text{节点 } i \text{ 和节点 } j \text{ 具有公共边} \\ 0, & \text{其他} \end{cases} \quad (3.1)$$

本文假设的复杂网络没有自环，即对于任意的节点 i ，用于表示复杂网络的邻接矩阵 $A_{ii} = 0$ 。

定义 $\mathcal{T} \triangleq \mathbb{R}_+$ （或者更一般情况下定义为一个么半群），给网络中的每个节点 i 定义一个连续时间上的实值状态变量 $x_i: \mathcal{T} \rightarrow \mathbb{R}$ ，复杂网络动态系统由一个带有初始状态值的常微分方程定义：

$$\begin{cases} \dot{x}_i(t) = \Psi_i(x_1(t), x_2(t), \dots, x_n(t)) \\ x_0(t) = b_i \end{cases} \quad (3.2)$$

其中 $t \in \mathcal{T}$ 、 $1 \leq i \leq n$ 。本文采用 $\dot{x}$ 来表示 $\frac{dx}{dt}$ （例如：速度）， $\Psi_i: \mathbb{R}^n \rightarrow \mathbb{R}$ （速度算子），本文假设所有的算子都是平滑的并且满足全局利普希茨连续条件（globally Lipschitz），并且 Ψ_i 是非时变得。

定义 $\mathbf{x}$ 表示向量 $(x_1, x_2, \dots, x_n)^T$ ，演化算子 Ψ_i 映射到状态流 $\Phi: \mathcal{T} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ：

$$\Phi(t, \mathbf{x}(0)) = \mathbf{x}(t) \quad (3.3)$$

$$= \int_0^t \left(\Psi_1(\mathbf{x}(\tau)), \dots, \Psi_n(\mathbf{x}(\tau)) \right)^T d\tau \quad (3.4)$$

本文方法要解决的问题是从多元时间序列 $\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_N)$ 中恢复演化算子 Ψ_i ，时间序列不要求必须是等间隔的，满足 $0 \leq t_1 < t_2 < \dots < t_N$ 。

对于如式(3.2)的通用形式所示，从单个时间序列中学习动力学 Φ 是一个定义不明确的问题。当 $\Delta t_i \triangleq t_{i+1} - t_i$ 作为基本函数 $r(\mathbf{x}(t_i))$ 的输入值时， Φ 是一个特殊的流，在时间变化的情况下有无限多个特殊的流。因此，本文引入额外的归纳偏差。根据文献[7]假设引入一个所有状态变量共享的操作 $\Psi: \mathbb{R} \times \mathcal{X}_A \rightarrow \mathbb{R}$ ，即

$$\dot{x}_i(t) = \Psi(x_i(t), \{x_j(t): A_{ij} = 1\}) \dots \dots \dots (3.5)$$

其中 $\mathcal{X}_A \subset 2^R$ 。式(3.5)表示 $x_i(t)$ 不仅与其自身状态相关，并且与其邻居状态相关， Ψ 对当前节点 i 的邻居节点具有排列不变性。

3.2 问题分析

在本节中，首先介绍当前最新基于编码器-解码器的神经网络动力学模型。本文提出将标量状态变量 $\mathbf{x}(t)$ 编码为潜在向量 $\mathbf{h}(t)$ 会导致过度拟合和错误进化算子的问题，并通过一个简单的热动力学在复杂网络上的例子来演示问题，分析潜在的原因。

3.2.1 基于编码器-解码器的神经网络动力学模型

在最新复杂网络神经动力学模型（如 NDCN）、图神经网络^[20]以及一般神经 ODE^[32]中都遵循一种常见的编码器-解码器架构。在 NDCN 中（如图 3.1 所示），作者定义一个编码器 $f_E: R \rightarrow R^d$ 和一个解码器 $f_D: R^d \rightarrow R$ ， $d \in \mathbb{Z}_+$ 是隐空间维度。复杂网络中每个节点的初始状态 $x_i(0)$ 都被联合编码为向量 $h_i(0) = f_E(x_i(0))$ 。归纳为如下定义：

$$\dot{H}(t) = f_\theta(LH(t)) \dots \dots \dots (3.6)$$

其中 $H(t) = (h_1(t), h_2(t), \dots, h_n(t))^T$ ， $f_\theta: R^d \rightarrow R^d$ 是一个向量值函数。显然采用图拉普拉斯算子进行聚合是一个合适的选择，但是它限制了相邻节点之间的动态耦合。

训练 NDCN 模型涉及联合拟合 f_E ， f_θ ， f_D 以最小化观测数据的经验损失，该模型的学习过程与标准的监督学习设置本质上相同，尤其是对拟合流的动力学行为没有明确的要求时。编码器将原始变量 $\mathbf{x}$ 嵌入到高维潜在空间中，演化算子使用设定的线性耦合机制（图拉普拉斯操作）来近似顶点之间的未知动态相互作用。NDCN 存在几个问题，包括超出训练数据时泛化能力差（在图 3.1 中显示为 $x-t$ 曲线在 $t > 5$ 时偏离真实曲线）。

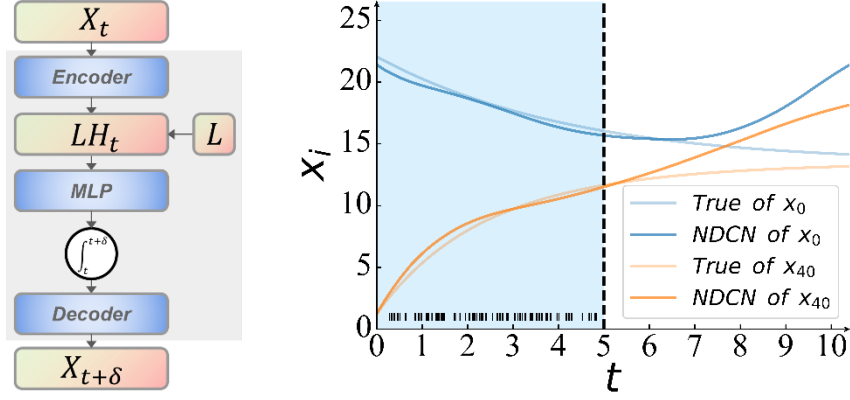


图 3.1 基于嵌入的复杂网络神经动力学模型 NDCN

3.2.2 神经动力学模型三个准则

根据通用逼近定理^[33], 可以用神经网络来拟合逼近任意时间序列。在 NDCN 中, 如果不断增加嵌入层维度 d , 总是能够拟合给定的观测时间序列数据。但是对于一个具有意义的拟合神经动力学的模型, 应该期望诱导流具有正确的动力学特征。因此, 本文提出三个用于评估神经动力学模型的准则。

3.2.2.1 动力学期长期预测准则

拟合神经动力学的模型, 通常仅有有限的观测时间点数据 $t_1, t_2, \dots, t_N$, 这些观测数据通常是在接近初始条件的短期瞬间收集起来的。

许多 NODE 相关研究表明, 使用来自与训练样本重叠的数据或短期外推数据作为测试样本, 这与标准的监督学习设置一致, 其中假设数据是独立同分布或者是接近独立同分布的。

然而在动力学建模任务中, 更重要的是拟合模型的长期行为。因此, 模型需要在长期范围内做出准确的预测, 才更具有实用价值。所以本文的第一个测试实验是估计模型是否可以在训练数据的观测时间点之外做出准确的长期预测。

3.2.2.2 动力学预测稳定性准则

一些研究学者认为在动态系统建模任务中, 能够预测正确的最终状态比中间状态值的预测更加有意义。与动力学的长期预测测试任务相关, 拟合模型需要与真实数据具有相同的固定点集合。两个对应的固定点应该具有相同的动态稳定性, 在本文中, 通过采用最大的 Lyapunov 指数^[34]来检查流的稳定性。

3.2.2.3 微分方程数值积分准则

最后一个准则是最重要却经常被忽视的：拟合的神经动力学模型是否为流引入满足数学原理的向量场。特别是，当模型引入一个状态转换函数，该函数是一个么半群动作。这意味着若 Φ 是一个流，那么：

$$\Phi(0, \mathbf{x}) = \mathbf{x} \dots \dots \dots (3.7)$$

$$\Phi(t_2, \Phi(t_1, \mathbf{x})) = \Phi(t_2 + t_1, \mathbf{x}) \dots \dots \dots (3.8)$$

对于任意的 $t_1, t_2 \in \mathcal{T}$ 。换句话说，动力学轨迹 $\mathbf{x}(t)$ 不能在任意时间 t 有分支。

3.2.3 基于编码器-解码器的神经动力学模型违背准则经验分析

本文通过使用 NDCN 模型进行示例演示基于嵌入的神经动力学模型存在的问题。在此示例中，NDCN 在短时域外推任务中表现出优于各种 RNN+GNN 基线的性能（详见文献[7]中表3）。

参照 NDCN 的设置，本文采用在 grid 网络上定义的热扩散方程来模拟生成时间序列数据。设置复杂网络的规模为 400 个节点，这些节点紧凑的排列在二维的正方形格子中。设 $\{v(k, l): 1 \leq k, l \leq 20\}$ 为顶点，两个顶点 $v(k, l)$ 和 $v(k', l')$ 相连当且仅当 $|k - k'| \leq 1, |l - l'| \leq 1$ 。换句话说，这是具有 2 维边界的 8 正则图变体。在该网络上，流由热扩散方程定义，对于每个节点 i ：

$$\dot{x}_i(t) = -\sum_{j=1}^N A_{ij} \alpha(x_i - x_j) \dots \dots \dots (3.9)$$

其中 α 是扩散常数。 $x_i(0)$ 设置为 $[0, 25]$ 中的随机值。在理论上，该式的耦合算子也可以通过图拉普拉斯算子表示。

从上述的热扩散流中，仿真生成了 80 个不规则时间间隔的样本数据，其中 $t \in [0, 5]$ 。使用 NDCN 作者提供的代码拟合 NDCN 模型（代码链接为：[GitHub - calvin-zcx/ndcn: Neural Dynamics on Complex Networks](https://github.com/calvin-zcx/ndcn: Neural Dynamics on Complex Networks)）。

3.2.3.1 基于编码器-解码器的动力学模型不能完成动力学的长期预测

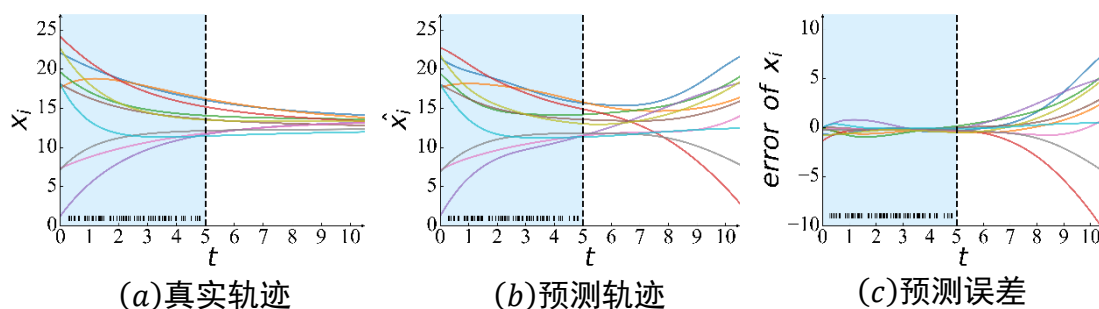


图 3.2 热扩散动力学在 grid 网络上随机 10 个节点的状态轨迹

图 3.2 显示了热扩散动力学在 grid 网络上随机 10 个节点的状态轨迹。其中图 3.2(a)是热扩散动力学在 grid 网络上 10 个节点的真实状态轨迹，可以看到各个节点收敛于平均温度。图 3.2(b)是 NDCN 模型使用虚线左侧的 80 个观测值进行训练（在轴须图中可见），预测的轨迹很好的拟合了训练观测数据，但是预测轨迹在训练观测数据之外（虚线右侧）迅速偏离。很明显，这些轨迹的动力学性质与热扩散动力学的性质不一致。图 3.2(c)显示预测误差随着远离训练观测数据而迅速增加。

通过热扩散动力学在 grid 网络上的示例可以明显得出经验性结论：基于编码器-解码器的动力学模型无法完成远离观测数据的预测任务，进一步表明基于编码器-解码器的动力学模型仅过度拟合了当前观测到的时间序列数据，进而不能完成动力学的长期预测。

3.2.3.2 基于编码器-解码器的动力学模型无法收敛到稳定状态

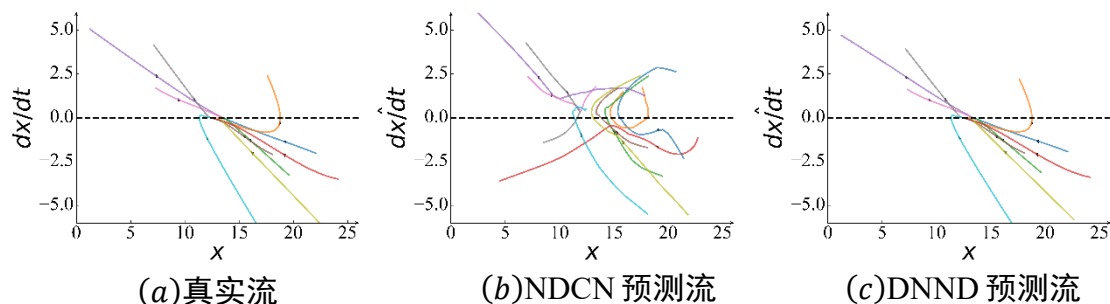


图 3.3 热流在 grid 网络上随机 10 个节点轨迹的二维投影

复杂网络上的热扩散动力学有一个简单的固定点：热稳态，即最终时刻每个节点 x_i 都收敛到复杂网络上所有节点的平均温度。图 3.3(b)显示 NDCN 模型未能捕获到达到热稳态的固定点，因为在拟合的 NDCN 模型中所有节点的状态变量

实际上随着时间的推移而发散。

本文计算了真实向量流和预测向量流的最大 Lyapunov 指数, 真实值为-35.83, 该数值表明向量流稳定。而预测流的值为 9.36, 正值表明该向量流不稳定。

通过绘制节点相位图的二维可视化揭示了相同的问题。如图 3.3(a)所示, 真实流轨迹收敛到系统的唯一固定点, 每个节点都具有平均热量。3.3(b)显示出 NDCN 预测流最初接近平均温度, 但最终偏离了正确的固定点。

3.2.3.3 基于编码器-解码器的动力学模型违反么半群动作准则

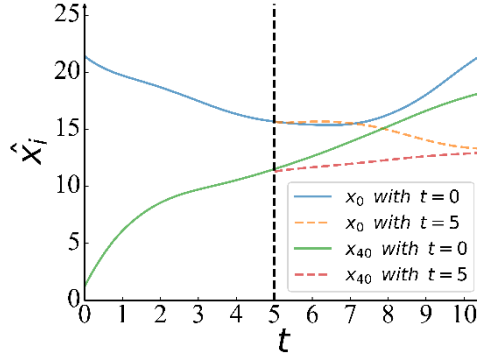


图 3.4 NDCN 预测流在重置初始条件下轨迹分支图

从 NDCN 预测的流 $\hat{\Phi}$ 中, 生成两个时间序列。一个 $t = 0$, 初始条件为 $x_i(0)$; 另一个 $t = 5$, 初始条件为 $\hat{\Phi}(5, x_i(0))$ 。如果 $\hat{\Phi}$ 是一个满足数学原理的流, 那么这两个时间序列应从 $t \geq 5$ 开始重合。图 3.4 显示了两个节点的时间序列, 从图中可以明显的看出重置初始条件导致了完全不同的时间序列, 图中表现为同一轨迹分支成两条轨迹, 表明 NDCN 没有将适当的动力学定义为 $\Phi(t_2, \Phi(t_1, x)) = \Phi(t_2 + t_1, x)$, 违反了么半群动作的要求。

3.2.4 基于编码器-解码器的神经动力学模型违背准则原因剖析

上面的三个例子表明, 神经动力学模型可以很好的拟合嵌入空间中的观察时间序列, 但无法产生正确的流。状态特征嵌入已用于从规则间隔的时间序列重建流^[35]。此外, 已经表明, p 维欧几里得空间中的任何同构映射 $h: \mathcal{X} \rightarrow \mathcal{X}, \mathcal{X} \subset \mathbb{R}^p$ 都可以通过在 $2p$ 维欧几里得空间上运行的神经 ODE $\hat{\Phi}: \mathcal{T} \times \mathbb{R}^{2p} \rightarrow \mathbb{R}^{2p}$ 来近似^[36]。但是基于嵌入的模型却无法恢复一个简单的热力学。本文认为有如下两点原因: 一是基于编码器-解码器的神经动力学模型中的编码器与解码器理论上应该互为

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/675210132041011114>