

## 基于图注意力网络的服务功能链路径优化研究

黄万伟<sup>\*①</sup> 李松<sup>①</sup> 张超钦<sup>②</sup> 王苏南<sup>③</sup> 张校辉<sup>④</sup>

<sup>①</sup>(郑州轻工业大学软件学院 郑州 450002)

<sup>②</sup>(郑州轻工业大学计算机与通信工程学院 郑州 450001)

<sup>③</sup>(深圳职业技术学院电子与通信工程学院 深圳 518005)

<sup>④</sup>(河南信安通信技术股份有限公司 郑州 450000)

**摘要:** 服务功能链(SFC)路径生成旨在为用户提供高速率、低延迟的多样化网络功能定制服务。现有的SFC路径选择算法大多都针对特定网络拓扑结构或使用单一评价指标,存在泛化性弱、评价标准单一的问题。针对此,该文提出基于图注意力网络(GAT)的SFC路径选择优化模型(SFC-GAT),其核心不再固定网络拓扑结构,而是将SFC路径选择问题建模为路径注意力问题。通过重新设计路径选择图和路径更新策略方式,增强模型泛化性;从时延和负载能力综合角度评价路径优化效果,解决路径性能单一评价问题。仿真实验结果表明:相比于约束下最短路径和最小过载路径选择策略,SFC-GAT至少能提高12%和7%选择路径综合性能。

**关键词:** 服务功能链;图注意力网络;注意力机制;路径选择

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2022)08-2833-09

DOI: [10.11999/JEIT210458](https://doi.org/10.11999/JEIT210458)

## Research on Optimization of Service Function Chain Path Based on Graph Attention Network

HUANG Wanwei<sup>①</sup> LI Song<sup>①</sup> ZHANG Chaoqin<sup>②</sup>  
WANG Sunan<sup>③</sup> ZHANG Xiaohui<sup>④</sup>

<sup>①</sup>(College of Software Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China)

<sup>②</sup>(College of Computer and Communication Engineering, Zhengzhou University of Light Industry,  
Zhengzhou 450001, China)

<sup>③</sup>(School of Electronic and Communication Engineering Shenzhen Polytechnic, Shenzhen 518005, China)

<sup>④</sup>(Henan Xin'an Communication Technology Co. Ltd, Zhengzhou 450000, China)

**Abstract:** Service Function Chain (SFC) path generation aims to provide users with diversified network function customization services with high speed and low delay. Most of the existing SFC path selection algorithms aim at specific network topology or use a single evaluation index, which has the problems of weak generalization and single evaluation standard. In view of this, an SFC path selection optimization model (SFC-GAT) based on Graph Attention network (GAT) is proposed. Its core is no longer to fix the network topology, but to model the SFC path selection problem as a path attention problem. The generalization of the model is enhanced by redesigning the path selection diagram and path update strategy; The effect of path optimization from the perspective of delay and load capacity is evaluated to solve the problem of single evaluation of path performance. The simulation results show that compared with the shortest path and minimum overload path selection strategy under constraints, SFC-GAT can improve the comprehensive performance of path selection by at least 12% and 7%.

**Key words:** Service Function Chain(SFC); Graph Attention network(GAT); Attention mechanism; Path selection

收稿日期: 2021-05-25; 改回日期: 2021-10-26; 网络出版: 2021-11-19

\*通信作者: 黄万伟 huangww79@163.com

基金项目: 国家自然科学基金(62072416, 62072414)

Foundation Items: The National Natural Science Foundation of China (62072416, 62072414)

## 1 引言

5G网络旨在为用户提供高速率、低延迟的多样化网络功能定制服务,这些网络功能服务通常以服务功能链<sup>[1]</sup>(Service Function Chain, SFC)的形式提供给用户。传统网络功能服务中网络功能(如防火墙、深度数据包检测、入侵检测等)都部署在专用硬件设备上,网络功能与专用硬件存在较强的耦合性,使SFC存在更新难、成本高和维护性差等问题。网络功能虚拟化<sup>[2]</sup>(Network Function Virtualization, NFV)和软件定义网络<sup>[3]</sup>(Software Defined Networking, SDN)技术将网络功能虚拟化,摆脱专用硬件设备的限制,以虚拟网络功能<sup>[4]</sup>(Virtualization Network Function, VNF)的形式运行在通用硬件设备平台上,通过SDN控制器对服务功能链请求(SFC Request, SFCR)进行统一管理,利用流量管控能力引导流量按序通过VNFs,从而生成SFC路径。

网络功能种类繁多,每类网络功能有多个虚拟网络功能实例<sup>[5]</sup>(VNF Instance, VNFI)。随着服务功能链长度的增加,符合SFCR的可选路径数目呈超指数级增长。考虑到实际网络服务请求中用户对网络功能的需求具有多样性与可变性,针对原需求的路径生成方案不能较好地适用于新需求路径生成,造成路径时延大、负载能力小,不具有泛化性。基于以上分析,如何在众多SFC路径中选择一条时延与负载能力综合评价较优的服务功能路径是当前SFC路径选择问题的研究重点。Liu等人<sup>[6]</sup>通过优化中间件放置位置,以端到端延迟作为评价标准,可以有效降低路径时延,但是在给定网络信息和策略规范的情况下,不能适用于节点变化的情况;Bremner-Barr等人<sup>[7]</sup>通过将来自不同服务请求中相同的虚拟网络功能进行合并,实现虚拟网络功率实例复用,减少功能实例个数以提高路径选择效率,但没考虑节点的负载能力,造成节点负载率过大。

随着人工智能技术的兴起,智能化SFC路径选择成为当前研究热点。Gu等人<sup>[8]</sup>设计了一个用模型辅助的深度强化学习框架实现SFC的业务调度,但模型训练消耗时间过长,且不能适用于动态变化的SFCR;Subramanya等人<sup>[9]</sup>采用神经网络的方法预测所需的虚拟网络功能实例数量来实现自动缩放,利用整数线性规划进行SFC路径选择从而达到了最小时延,但未考虑节点负载能力,造成路径负载过大,业务处理效率降低,影响服务质量(Quality of Service, QoS)。综上所述,目前传统的SFC路径选择算法和基于机器学习的智能路径选择主要存在两个问题:(1)泛化性弱,针对的是单一SFCR下的路径生成,不能很好地适应实时多变的SFCR。(2)评价机

制不完备,大多数路径选择都只以路径时延作为评价标准,未考虑节点负载能力,易造成路径拥堵。

针对上述问题,本文基于图注意力网络(Graph Attention network, GAT)<sup>[10]</sup>中注意力机制(attention mechanism)针对节点特征对边进行聚合的能力,提出一种服务功能链路径注意力模型(SFC-GAT)。将SFC路径图重构为带权有向图结构,功能节点与逻辑链路的属性进行向量化表示;以相邻的两列节点构成的全连接网络作为一个训练层,采用attention机制,得到节点之间注意力权重,对其进行加权求和后得到预测路径,通过与真实标签对比对SFC-GAT进行更新,使模型精度逐渐提高;在存在多条服务路径或SFCR变更的情况下,经过训练的SFC-GAT能优选时延与负载能力综合评价指数较好的服务功能路径生成方案。

本文主要贡献概括如下:

(1)分析传统SFC路径生成与智能化SFC路径生成策略中普遍存在泛化能力弱、评价机制不完备问题,基于图注意力网络的注意力机制能针对节点特征对边进行聚合的能力,提出了一种服务功能链路径注意力模型(Service Function Chain-Graph Attention network, SFC-GAT)。

(2)建立SFC路径选择图重构规则,增强对不同SFCR情况下网络拓扑的泛化能力;使用图注意力网络对SFC重构后路径选择图进行优化训练,得到路径时延和节点负载能力对选择路径评价质量的重要指数,从而提高路径对链路时延和负载能力的综合评价指数。

## 2 相关研究

### 2.1 服务功能链

服务功能链是指根据用户需求将VNF按网络流量通过的顺序所构成的一条顺序逻辑链<sup>[11]</sup>,在SDN\NFV环境下,一个网络功能节点可以部署多种不同的网络功能实例。图1为SDN\NFV环境下服务功能链拓扑结构图,在SDN\NFV环境下,服务功能链拓扑结构图是一个无向图,网络流量转发业务全部通过SDN控制器进行转发。SFC路径选择问题在结构上与传统网络路由选择问题类似,大量研究都是将端到端时延作为服务功能链路径选择质量的主要评价指标,忽视了其他指标对路径质量的影响。Gu等人<sup>[12]</sup>提出了基于分层图的SFC编排方案,以延迟作为分层图中链路权值,并基于模拟退火的批量处理算法寻找最短服务路径,但未考虑网络节点负载情况,链路资源利用率低。Clayman等人<sup>[13]</sup>提出了服务功能链路径选择问题的动态优化部署模型,分别设计了针对负载均衡、降低路径开销和保障QoS 3个目标的优化算法,但缺乏对多种类型参数的联合优化。

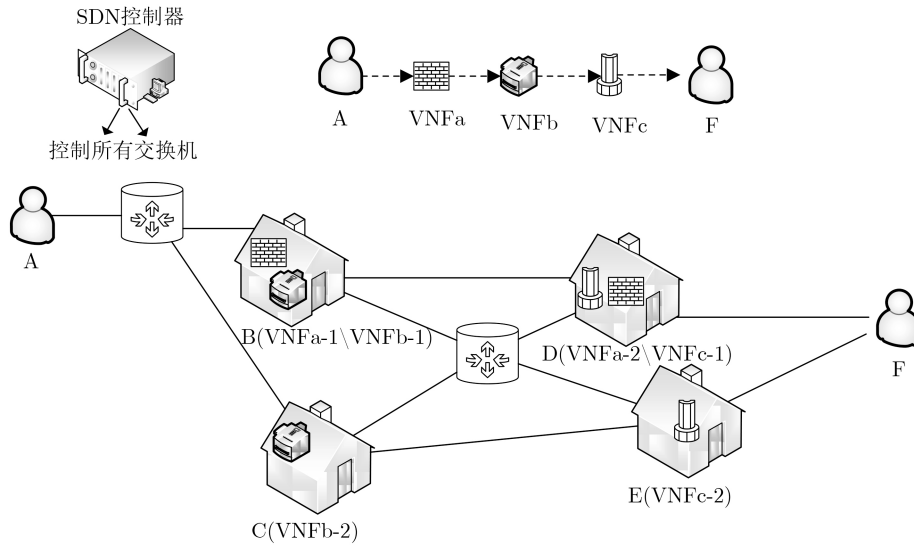


图1 SDN/NFV环境下服务功能链拓扑结构图

### 2.2 图注意力网络

2018年Veličković等人<sup>[14]</sup>提出了一种应用于图结构数据的图注意力网络。图注意力网络不同于先前一些基于谱域的图神经网络，它可以通过注意力机制对邻居节点做聚合操作，考虑各个邻居节点的相关度，实现对不同邻居节点权重的自适应分配，具有高效、可移植性等优势。在GAT图结构中，设图中任意节点 $v_i$ 在第 $l$ 层所对应的特征向量为 $h_i$ ,  $h_i \in R^{d(l)}$ ,  $R^{d(l)}$ 表示节点的特征长度，经过一个以注意力机制为核心的聚合操作后，输出的是每个节点新的特征向量 $h'_i$ ,  $h'_i \in R^{d(l+1)}$ ,  $R^{d(l+1)}$ 表示的是输出的特征向量的长度。这个聚合操作称为图注意力层(Graph Attention Layer, GAL)。图2为图注意力层操作图。其中，中心节点为 $v_i$ ，可设邻居节点 $v_j$ 到 $v_i$ 的权重系数为： $\alpha_{ij} = a(Wh_i, Wh_j)$ ,  $W \in R^{d(l+1) \times d(l)}$ 是该层节点特征变换的权重参数。 $a(\cdot)$ 是计算两个节点相关度的函数。经过计算，可以得到 $v_i$ 所有邻居节点对 $v_i$ 相关度大小，再经过softmax归一化处理，可以得到 $v_i$ 与所有相邻节点的注意力权重。SFC-GAT根据注意力权重和SFCR的约束，对路径上通过的节点进行加权求和，得到一条注意力值和最大的路径即为SFC-GAT选择的路径。

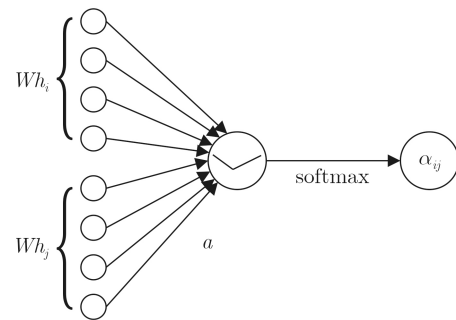


图2 图注意力层操作图

向图 $G = (u, V, E)$ ,  $u$ 表示SFCR，是选择SFC路径的约束条件。 $V$ 表示网络功能实例节点的集合， $E$ 表示虚拟链路的集合。

(2)基于SFCR对虚拟网络功能顺序的约束，相邻的两个功能节点在物理链路中并不一定是相邻关系，故在构建过程忽略物理链路中节点的位置关系，只考虑节点之间逻辑链路关系，采用虚线指向性箭头表示节点间的链接关系，以表示节点在物理链路中不确定的相邻关系。虚线权值代表该链路时延(传输时延+响应时延)。

(3)在SDN/NFV环境中，VNF往往是多实例进行部署，故将同一类型的VNF排在一列。节点的权值代表当前节点的负载能力，从而形成 $N$ 层的图结构。

SFC路径选择图的注意力机制规范定义：给定一个图3中第 $t$ 列的第 $x$ 个节点 $v_x^t$ 和第 $t+1$ 列中所有节点构成的 $\{v_1^{t+1}, v_2^{t+1}, \dots, v_y^{t+1} | v_x^t\} \in \Gamma_{v_x^t}$ ,  $f'(v_x^t, \Gamma_{v_x^t})$ 表示 $v_x^t$ 与 $\Gamma_{v_x^t}$ 中每个节点映射到相关性得分的函数

$$f' : \{v_x^t\} \times \Gamma_{v_x^t} \rightarrow [0, 1] \quad (1)$$

相关性得分表示该列节点相对重要性，满足

## 3 SFC-GAT模型

### 3.1 模型约束

本文旨在给定网络拓扑结构的情况下，根据SFCR将服务功能链拓扑结构图重构为带权有向的SFC路径选择图，并采用注意力机制从SFC路径选择图中找寻一条时延与负载能力综合评价优秀的服务路径，因此需要定义重构规则与图中注意力机制的规范。SFC路径选择图的重构规则如下：

(1)将服务功能链拓扑结构图进行重构为有

$$\sum_{i=1}^{\Gamma_{v_x^t}} f'(v_x^t, v_i^{t+1}) = 1 \quad (2)$$

### 3.2 SFC-GAT模型架构

SFC-GAT是结合SFC路径选择图与图注意力网络而构建的模型，旨在通过图注意力网络所具有的灵活、结构化的输入输出特点和根据节点特征对边进行聚合的能力，从而对SFC路径选择做出精准预测。SFC-GAT模型由图构建、信息传递和整合预测3部分组成，如图4所示。

(1)图构建：将服务功能链路拓扑结构图重构为带权有向图结构，将每个虚拟网络功能实例作为一个节点，按SFCR对节点进行排列，同一列为具有相同网络功能的虚拟网络功能节点，后一列是前一列网络流或业务的流向方向。若SFCR长度为 $N$ ，生成图则对应一个 $N \times N$ 的矩阵。

(2)信息传递：将相邻两列节点进行全连接，前一列中节点与后一列中任意节点都存在链接关系，在每两列之间利用GAT原理求得相邻节点之间的相关性得分 $\alpha_{ij} = a(Wv_i, Wv_j)$ ，从而在图中进行信息传递。

(3)整合预测：根据SFCR中对虚拟网络功能的顺序约束整合路径的相关注意力得分之和 $\sum_{i=1}^N V^t$ ，从中得到一个和最大且满足SFCR约束的节点集合，进行二分类预测。

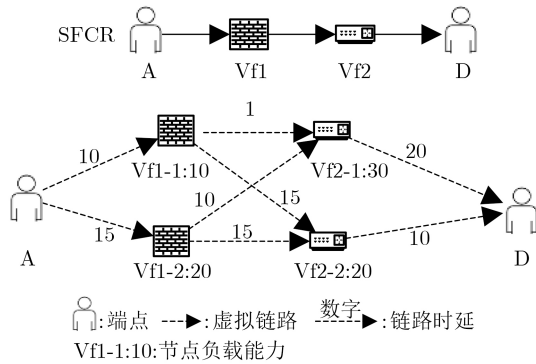


图3 重构后服务功能链路径选择图

### 3.3 路径性能评估

对于生成的SFC路径，本文将从平均路径时延、平均路径负载能力和路径综合评价指数对路径性能进行评价。对于平均路径时延，在SFC路径重构图 $G = (u, V, E)$ 中， $dealy(v_i, v_{i+1})$ 表示从节点 $v_i$ 到节点 $v_j$ 的路径总时延(传输时延+响应时延)，对于一条SFC路径而言，路径时延能够反映出该条路径对SFCR的响应速度，是路径性能的一个重要评价指标，式(3)为平均路径时延计算公式

$$Delay_{SFC} = \frac{\sum_{i=1}^{len(u)-1} dealy(v_i, v_{i+1})}{n} \quad (3)$$

其中， $n$ 为SFC路径实际通过的物理网络节点的数量， $len(u)$ 表示服务功能链的长度， $\sum_{i=1}^{len(u)-1} dealy(v_i, v_{i+1})$ 是该SFC路径的时延之和。

对于平均路径负载能力，在SFC路径选择图 $G = (u, V, E)$ 中， $load(v_i)$ 为当前该节点的负载能力，对于一条SFC路径而言，路径负载能力能够反映出该条路径对SFCR的处理速度，也是路径质量的一个重要评价指标，式(4)为平均路径负载能力计算公式

$$Load_{SFC} = \frac{\sum_{i=1}^n load(v_i)}{n} \quad (4)$$

其中， $n$ 为SFC路径实际通过的物理网络节点的数量， $\sum_{i=1}^n load(v_i)$ 是该条SFC所通过网络节点的负载能力之和。

本文对路径的综合评价指数定义为对平均路径时延的倒数与平均路径负载能力的加权求和，式(5)为路径综合评价指数计算公式

$$P_{SFC} = w_1 \frac{1}{Delay_{SFC}} + w_2 Load_{SFC} \quad (5)$$

其中， $w_1, w_2$ 为路径优化模型中计算路径优化效果的重要评价指标，代表模型路径优化方向，通过训练不断迭代更新。

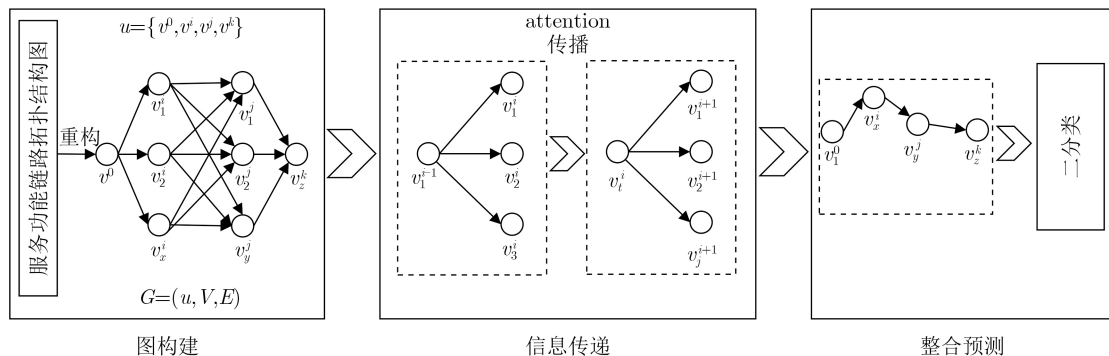


图4 SFC-GAT模型

## 4 SFC-GAT映射

### 4.1 图构建

在SDN/NFV的环境中，将服务功能链拓扑结构图重构为带权有向的SFC路径选择图 $G=(u, V, E)$ 。其中 $u=(v_a, v_b, \dots, v_m)$ 是服务功能链路径选择的约束条件， $v_a$ 表示SFCR的源节点， $v_m$ 表示目的节点，中间节点的顺序是SFC网络流或业务流通过网络功能的顺序； $V=(v_{t_1}^1, v_{t_2}^2, \dots, v_{t_n}^N)$ 表示网络功能实例节点的集合， $N$ 表示一共有 $n$ 种不同的网络功能， $t_n$ 表示该网络功能共有多少虚拟网络功能实例， $\text{load}(v_i^q)$ 表示第 $q$ 类网络功能的第 $r$ 个虚拟网络功能实例节点的剩余负载大小。 $E$ 表示虚拟链路的集合， $e_{i,j}^{x,y}=(v_i^x, v_j^y) \in E$ ，意为节点 $v_i^x$ 指向 $v_j^y$ 的边，其值 $\text{delay}(v_i^x, v_j^y)$ 为 $v_i^x$ 到 $v_j^y$ 节点的时延。

### 4.2 消息传递

在SFC-GAT模型消息传递模块中，本文引入了GAT，它将重构后的SFC路径选择图以每相邻两列作为一个图注意力层，从而拆分为若干个图注意力层堆叠。第1层输入的是源目标节点和第1列节点的特征向量集合 $\mathbf{v}_1^1 = \{v_1^1, v_2^1, \dots, v_{l_1}^1\} \in \Gamma_{v_1^1}$ ， $\mathbf{v}_1^1, \mathbf{v}_i^2 \in \mathbf{R}^{d(l)}$ 。

在第 $t$ 层图注意力层中，输入的是 $\mathbf{v}_x^t$ ，输出的是 $\mathbf{v}_y^{t+1}$ 。为了更好地获取节点之间的信息，本文先用一个矩阵 $\mathbf{W}^t \in \mathbf{R}^{d(l+1) \times d(l)}$ 完成对节点的特征转换，再利用图注意机制计算公式得到 $e_{i,j}^{t,t+1} = a(\mathbf{W}^t v_i^t, \mathbf{W}^t v_j^{t+1})$ ，表示 $v_j^{t+1}$ 节点对 $v_i^t$ 节点的相关程度，其中 $a(\cdot)$ 为一个共享的注意力计算函数。

本文规定每个节点的信息都可以传递到相邻节点，所以对每一对相邻节点都要计算注意力系数。对于第 $t$ 列的第 $i$ 个节点 $v_i^t$ 和第 $t+1$ 列中所有节点，将会把 $t+1$ 列中所有节点对 $v_i^t$ 的注意力做一个正则化处理，这里使用softmax函数

$$\alpha_{i,j}^{t,t+1} = \text{softmax}(e_{i,j}^{t,t+1}) = \frac{\exp(e_{i,j}^{t,t+1})}{\sum_{j=1}^n \exp(e_{i,j}^{t,t+1})} \quad (6)$$

在计算注意力系数时，为了避免某个注意力系数的值远大于其他值而造成不易训练以及为了增强模型的泛化拟合能力，对线性变化后的值加入非线性激活函数LeakyReLU(负数斜率为0.2)， $\alpha_{i,j}^{t,t+1}$ 可以重写为

$$\alpha_{i,j}^{t,t+1} = \frac{\exp(\text{LeakyReLU}(a(\mathbf{W}^t v_i^t || \mathbf{W}^t v_j^{t+1})))}{\sum_{j=1}^n \exp(\text{LeakyReLU}(a(\mathbf{W}^t v_i^t || \mathbf{W}^t v_j^{t+1})))} \quad (7)$$

### 4.3 整合预测

在消息传递过程中通过更新节点表示，根据每

一层节点计算的注意力系数，得到SFC选择路径序列图 $h^G$ 。得到 $h^G$ 后，根据4.2节归一化后得到的值，从中找到和最大的一条路径即为本次预测路径。

$$p_\theta = \text{Max} \left( \sum_{i=1}^n \alpha_{i,j}^{t,t+1} \right) \quad (8)$$

其中， $n$ 为服务功能链长度， $\alpha_{i,j}^{t,t+1}$ 为两点相关度大小。本文通过训练优化 $\theta$ 来最小化交叉熵损失，训练优化公式为 $\Phi = (p_\theta, L) = \text{mean}(\text{sum}(L - p_\theta)^2)$ 。其中 $p_\theta$ 为SFC-GAT模型预测的SFC路径， $L$ 是真实标签，是该SFC拓扑图下最优的SFC路径生成方案。GAT-SFC路径优化算法如表1所示。

## 5 仿真实验

### 5.1 参数设置

为了对SFC-GAT模型选择服务功能链路径的性能进行验证，该实验运行在i5-10600KF-CPU,16 GB-DDR4内存和两块GTX-1080 8GB显卡的设备上，基于Python3.5,tensorflow1.6-gpu进行仿真实验。使用Networkx按照Waxman-Salam模型随机生成网络拓扑，网络节点个数为10，每个网络节点能提供的负载能力为10 Mb/s，网络中可提供10种不同种类的网络功能，每类网络功能可选择实例数为1~3个。在该网络拓扑结构中各功能实例随机分布在所有节点中，任意相邻节点间的链路总时延在1~10 ms。

图5为Networkx所生成网络拓扑示意图，图中节点属性包括网络功能节点名称、所部署功能实例

表1 GAT-SFC路径优化算法

输入	重构后SFC路径选择图 $G=(u, V, E)$
输出	SFC预测路径
(1)	<b>Initialization</b>
(2)	lr = 0.005 //学习率
(3)	l2_coef = 0.0005 //权重衰减率
(4)	<b>Function</b> GAT-SFC( $G=(u, V, E)$ )
(5)	<b>Initialization</b> $W$ //根据 $G$ 中节点的特征维度初始化
(6)	<b>For each</b> $t \in \{1 \dots \text{len}(u) - 1\}$ //从第1层计算节点对下层节点的注意力
(7)	<b>For each</b> $i \in \{1 \dots \text{len}(t)\}$ //第 $t$ 层一共多少个实例节点
(8)	<b>For each</b> $j \in \{1 \dots \text{len}(t+1)\}$ // $t+1$ 层节点数量
(9)	Update( $\alpha_{i,j}^{t,t+1}$ ) //根据式(7)计算两个节点之间的注意力值
(10)	$p_\theta = \text{Max}(\sum_{i=1}^n \alpha_{i,j}^{t,t+1})$ //根据式(8)整合预测结果，得到预测路径
(11)	Loss = mean(sum( $L - p_\theta$ ) <sup>2</sup> ) //结果对比，计算损失值
(12)	$W = \text{training}(\text{loss}, \text{lr}, \text{l2\_coef})$ //更新权重
(13)	<b>End Function</b>

以及当前节点所剩负载大小，边的属性为链路时延。服务请求中的源节点和目标节点在网络中随机选择，服务功能链请求由2~10个互不相同的网络功能随机组成，每次服务请求中占网络节点负载为1 Mb/s，考虑到随着实验进行，网络节点会被重复利用，节点的负载能力会逐渐下降，因此规定一个服务请求在经过3次请求后会释放该次请求所占的网络节点的负载。为了提高算法评估的精确性，采用蒙特卡洛方法，在每个实验场景下都进行了100组仿真测试，取100组仿真测试的平均值作为测试结果。

### 5.2 路径优化效果分析

本文从SFC平均路径时延、平均路径负载能力和路径综合评价指数3方面对SFC路径选择优化效果进行对比，对比方法选择了约束最短路径算法(Constrained Shortest Path, CSP)和最小过载算法<sup>[15]</sup>(Least Busy Path, LBP)。其中，CSP是通过路径矩阵来计算每组实例节点的累计时延，使用回溯搜索符合SFCR的最短时延路径；LBP将每个功能可选的功能实例节点的剩余负载进行排序，并选择剩余负载最大的节点。为了验证本模型对SFC路径生

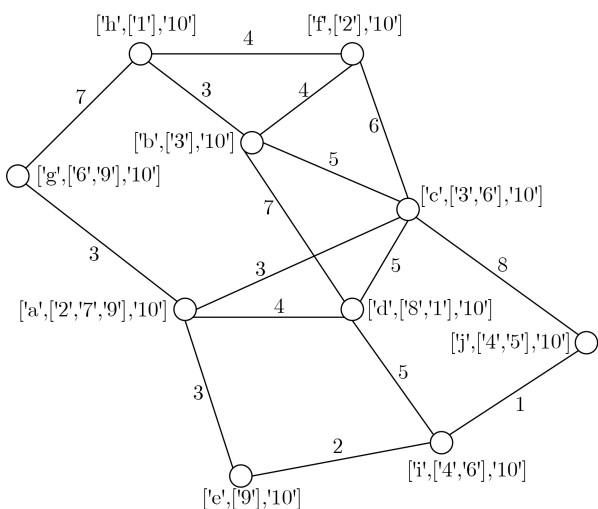


图5 Networkx生成网络拓扑示意图

成具有良好的泛化能力，分别设计了在两种情景下对生成路径的优化效果进行评估。其中：(1)SFCR不变时，网络中网络功能实例个数变化对SFC路径时延和节点负载能力的影响；(2)SFCR变化时，服务功能链长度对SFC路径时延和节点负载能力的影响。

#### 5.2.1 训练优化

为了避免模型在与对比算法开始对比时误差过大，需先对模型进行前期训练以提高模型的对优选路径的准确率。图6给出了在SFCR长度为5时以时延和负载能力为特征，SFC-GAT模型损失值和准确率优化训练过程，总训练次数为500次。从图6(a)可以看出，随着训练次数的增加，损失值呈逐渐降低的趋势；图6(b)可以看出，随着训练次数的增加，优选路径的准确率呈现逐渐增高的趋势，并在训练约470次之后模型预测准确率达到85%。

#### 5.2.2 SFCR不变时路径选择优化

本次实验旨在发现在给定SFCR情况下，SFC-GAT模型能否在给定的网络拓扑下寻找到较低时延且具有较高负载能力的SFC路径。图7(a)给出了在SFCR长度为5的情况下，路径时延随网络功能实例个数变化的仿真结果，仿真结果取100次随机实验的平均值。从图7(a)可以看出，由于CSP算法是根据满足SFCR下的最短时延路径，因此路径时延最低；经过训练后的SFC-GAT模型是将路径时延作为综合路径优化目标的一部分，生成路径的时延略高于CSP；而LBP是根据节点剩余负载选择，对路径时延并无要求，因此选择路径时延远大于其他两种。

图7(b)给出了与在SFCR长度为5的情况下，路径负载能力随网络功能实例个数变化的仿真结果，仿真结果取100次随机实验的平均值。从图7(b)可以看出，LBP是针对节点负载能力性能选择SFC，因此它对SFC节点负载能力优化效果最好；SFC-GAT将负载能力作为一个优化目标，得到的SFC

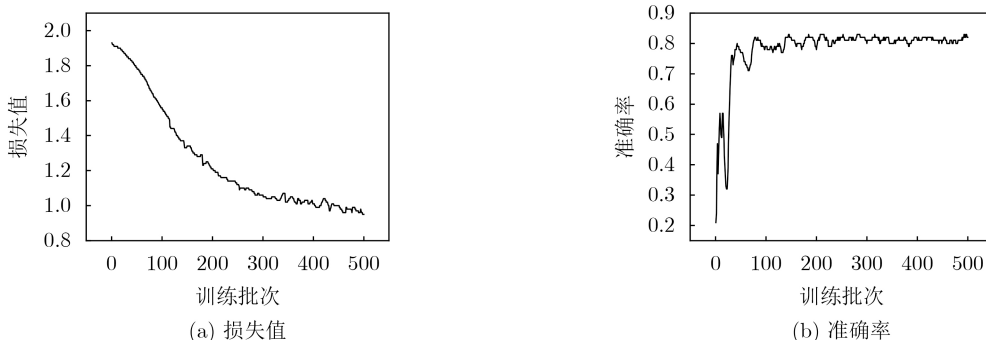


图6 SFC-GAT模型损失值和准确率优化训练过程

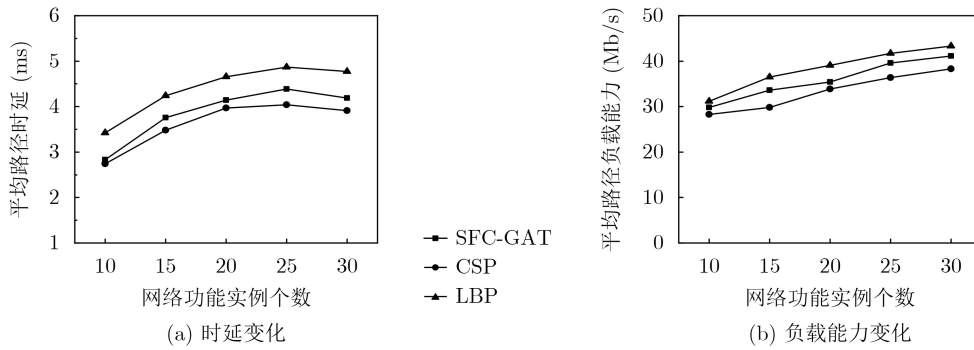


图7 SFCR不变情况下时延与负载能力变化

负载优化效果与LBP算法接近；而CSP在选择服务路径时没有考虑节点负载能力，因此路径负载能力最差。在实验过程中，每经过3次请求，节点会释放1次请求所占负载，因此随着功能实例个数的增加，节点负载释放的比例会增加，导致负载能力呈上升趋势，总体实验结果符合预测趋势。

在SFCR不变时，网络功能实例个数的增加表示符合SFCR的可选路径数目呈指数级增长。CSP和LBP路径选择算法在面对指数级数量的可选路径时，只针对单一评价进行选择路径，造成选择路径的时延与负载能力之间存在较大偏差，路径综合指数较低。本模型的路径优化算法以时延与节点负载能力作为预测路径的综合评价指标，模型在优化训练时也是以两个指标的综合值为优化方向。根据式(5)可以计算得到SFC-GAT选择的SFC路径综合评价指数相比CSP提高12.43%，比LBP提高7.59%。实验结果如图8所示。

### 5.2.3 SFCR变化时路径选择优化

真实网络场景中SFCR往往是动态变化的，因此设计在存在30个功能实例的情况下，SFC链路长度变化下各个算法在评价路径时延与节点负载能力方面的对比，仿真结果取100次随机实验的平均值。图9(a)为在SFCR变化时，SFC链路长度增加对路径时延产生的影响。从图9(a)可以看出，随着

SFC链路长度的增加，CSP以路径总时延大小为评价标准，选择的路径时延最小；LBP以总路径负载能力为评价指标，未考虑选择路径的时延，选择的路径时延最大；SFC-GAT路径优化算法在选择路径时以路径时延与负载能力为综合评价指标，故选择路径的平均时延比LBP算法低，比CSP算法高。

图9(b)为在SFCR变化时，SFC链路长度增加对路径负载能力影响的仿真结果，仿真结果取100次随机实验的平均值。从图9(b)可以看出，LBP以总路径剩余负载为评价指标，所选择SFC路径平均负载效果最好；CSP单以路径时延为评价指标，未考虑路径负载能力对路径效果的影响，所选择路径的平均负载能力效果最差；SFC-GAT路径优化算法以路径负载能力与时延为路径优化的综合评价，

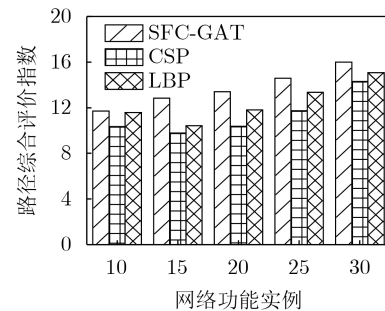


图8 SFCR不变情况SFC路径综合评价指数

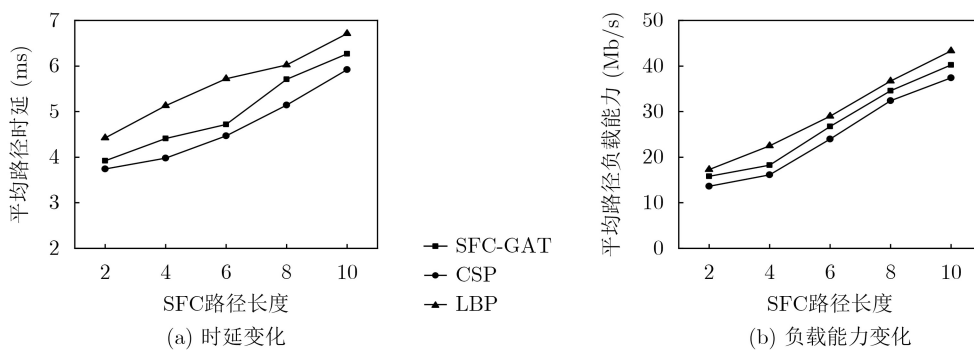


图9 SFCR变化情况下时延与负载能力变化

其选择路径的平均负载效果在LBP之下, SCP之上; 在实验过程中, 每经过3次请求, 节点会释放1次请求所占负载, 但SFC链路长度的增加会造成SFC通过网络功能节点数量增加, 导致整体负载能力呈上升趋势, 总体实验结果符合预测趋势。

随着SFCR请求长度的增加, 固定数量的网络功能实例导致可选择路径数目减少, 但可选路径长度增加会导致选择路径复杂度上涨。CSP通过计算路径的累计时延, 使用回溯搜索符合SFCR的最短时延路径, 忽略了负载能力对路径的影响; LBP按SFCR通过的节点顺序, 将节点按剩余负载大小排列, 并选择剩余负载大的节点构成最大负载能力路径, 缺少对时延的评价; SFC-GAT路径优化算法以时延和负载作为训练优化指标, 在对路径综合评价时, 优于单指标评价策略的CSP和LBP。根据式(5)可以计算得到SFC-GAT选择的SFC路径综合评价指数相比CSP提高14.58%, 比LBP提高7.69%。实验结果如图10所示。

## 6 结束语

针对服务功能链路径选择优化问题, 本文提出一种基于GAT的SFC路径优化模型SFC-GAT, 采用图注意力机制以时延与节点负载能力作为特征, 进行优化训练, 在给定网络拓扑结构下针对SFCR有无变化, 对网络功能实例个数和SFC链路长度对SFC路径选择优化影响进行仿真实验, 以提高SFC路径在时延与节点负载能力的综合评价指数。文中将SFC-GAT模型与CSP, LBP路径选择算法进行对比, 表明了SFC-GAT模型对比两种路径选择算法CSP, LBP至少能提升SFC路径综合性能指数12%和7%, 并且SFC-GAT模型通过对网络拓扑图进行重构, 能够实现不同网络拓扑结构的推广, 增强了模型泛化能力。后续工作将进一步考虑优化SFC-GAT模型路径选择优化算法设计结构, 提高模型预测准确率, 加入对链路抖动、节点失效概率等特征的关注, 提高模型综合评价指数的优化范围。

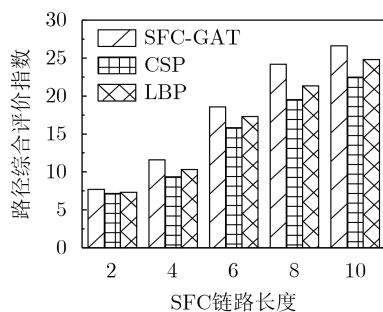


图 10 SFCR变化时SFC路径综合评价指数

## 参考文献

- [1] PANDEY S, HONG J W K, and YOO J H. Q-Learning based SFC deployment on edge computing environment[C]. The 21st Asia-Pacific Network Operations and Management Symposium, Daegu, Korea (South), 2020: 220–226.
- [2] HUFF A, VENÂNCIO G, GARCIA V F, *et al.* Building multi-domain service function chains based on multiple NFV orchestrators[C]. 2020 IEEE Conference on Network Function Virtualization and Software Defined Networks, Leganes, Spain, 2020: 19–24.
- [3] 卢旭, 易波, 王兴伟, 等. SDN/NFV下的5G网络资源切片管理机制[J]. 小型微型计算机系统, 2021, 42(5): 1082–1087. doi: 10.3969/j.issn.1000-1220.2021.05.032.  
LU Xu, YI Bo, WANG Xingwei, *et al.* 5G network resource slice management mechanism in software-defined networking/net-work function virtualization[J]. *Journal of Chinese Computer Systems*, 2021, 42(5): 1082–1087. doi: 10.3969/j.issn.1000-1220.2021.05.032.
- [4] Pham Chuan, Tran Nguyen H, Ren Shaolei, *et al.* Traffic-Aware and Energy-Efficient vNF Placement for Service Chaining: Joint Sampling and Matching Approach[J]. *IEEE Transactions on Services Computing*, 2020, 13(1): 172–185. doi: 10.1109/tsc.2017.2671867.
- [5] WANG Xiaolei, XIE Lijun, QIN Zhiqiang, *et al.* Clustering of virtual network function instances oriented to compatibility in 5G network[J]. *China Communications*, 2017, 14(12): 111–119. doi: 10.1109/cc.2017.8246341.
- [6] LIU Jiaqiang, LI Yong, ZHANG Ying, *et al.* Improve service chaining performance with optimized middlebox placement[J]. *IEEE Transactions on Services Computing*, 2017, 10(4): 560–573. doi: 10.1109/TSC.2015.2502252.
- [7] BREMLER-BARR A, HARCHOL Y, and HAY D. OpenBox: A software-defined framework for developing, deploying, and managing network functions[C]. 2016 ACM SIGCOMM Conference, Florianópolis, Brazil, 2016: 511–524.
- [8] GU Lin, ZENG Deze, LI Wei, *et al.* Intelligent VNF orchestration and flow scheduling via model-assisted deep reinforcement learning[J]. *IEEE Journal on Selected Areas in Communications*, 2020, 38(2): 279–291. doi: 10.1109/JSAC.2019.2959182.
- [9] SUBRAMANYA T, HARUTYUNYAN D, and RIGGIO R. Machine learning-driven service function chain placement and scaling in MEC-enabled 5G networks[J]. *Computer Networks*, 2020, 166: 106980. doi: 10.1016/j.comnet.



- 2019.106980.
- [10] CHOWDHURY P N, SHIVAKUMARA P, KANCHAN S, *et al.* Graph attention network for detecting license plates in crowded street scenes[J]. *Pattern Recognition Letters*, 2020, 140: 18–25. doi: [10.1016/j.patrec.2020.09.018](https://doi.org/10.1016/j.patrec.2020.09.018).
- [11] 唐伦, 曹睿, 廖皓, 等. 基于深度强化学习的服务功能链可靠部署算法[J]. *电子与信息学报*, 2020, 42(12): 2931–2938. doi: [10.11999/JEIT190969](https://doi.org/10.11999/JEIT190969).
- [12] GU Yunjie, DING Yuehang, HU Yuxiang. Joint optimization of delay guarantees and resource allocation for service function Chaining[J]. *IEICE Transactions on Information and Systems*, 2019, E102.D(12): 2611–2614.
- [13] CLAYMAN S, MAINI E, GALIS A, *et al.* The dynamic placement of virtual network functions[C]. 2014 IEEE Network Operations and Management Symposium, Krakow, Poland, 2014: 1–9.
- [14] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, *et al.* Graph attention networks[C]. The 6th International Conference on Learning Representations, Vancouver, Canada. 2018.
- [15] XIE Lijun, JIANG Yiming, WANG Binqiang, *et al.* An approach for network Function combination based on least busy placement algorithm[J]. *China Communications*, 2016, 13(S1): 167–176.
- 黄万伟：男，1979年生，副教授，研究方向为人机智能交互系统、宽带信息网络和大数据技术。
- 李松：男，1997年生，硕士生，研究方向为软件定义网络、服务功能链。
- 张超钦：男，1978年生，副教授，研究方向为计算机网络体系架构、软件定义网络。
- 王苏南：男，1984年生，副教授，研究方向为网络系统架构、数字信号处理、网络信息安全、数据流分析。
- 张校辉：男，1979年生，博士，研究方向为网络空间安全技术、大数据及人工智能技术的公共安全应用。

责任编辑：余蓉