



面向视频云微服务系统的智能运维技术

Intelligent Operation and Maintenance Technology for Video Cloud Microservice System

摘要: 提出了一种基于人工智能(AI)的保障视频云体验质量(QoE)的系统架构。该系统针对多个维度创建运维知识图谱,例如运行数据、运行环境、运维数据,以用于建模、感知、映射和分析。在对系统的微服务保障中,运用了图神经网络(GNN)等方法进行分类和预测。通过知识图谱和机器学习,该系统可实现实时监控、自愈恢复、智能预测和主动运维,从而实现QoE的智能保障。

关键词: 视频云系统;数据挖掘;知识图谱;图神经网络;智能运维

Abstract: Based on artificial intelligence (AI), a system architecture that guarantees video cloud quality of experience (QoE) is proposed. According to multiple dimensions, the system creates an operation and maintainance knowledge map for modeling, sensing, mapping, and analyzing, such as operating data, operating environment, and operation and maintenance data. In the microservice system, methods such as graph neural network (GNN) are introduced for classification and prediction. Based on the technology of knowledge graph and machine learning, the system can perform realtime monitoring, self-healing, intelligent predicting and active operation and maintainance to implement intelligent guarantee of QoE.

Keywords: video cloud system; data mining; knowledge graph; graph neural network; AIOps

徐代刚 / XU Daigang

姜磊 / JIANG Lei

梅君君 / MEI Junjun

(中兴通讯股份有限公司, 中国 深圳 518057)
(ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTETJ.202101014

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20210128.1824.002.html>

网络出版日期: 2021-01-29

收稿日期: 2020-11-15

随着5G时代的到来,增强型移动宽带(eMBB)、海量机器类通信(mMTC)、超可靠低时延通信(URLLC)3大应用场景,尤其是相关的音视频应用,得到了快速发展。此外,由于受到新冠肺炎疫情等不确定性因素的影响,企业或团体远程办公、远程会议的场景和需求日益增多,相应的通信系统也变得日益复杂。以视频会议为例,它需要支持双流会议、多流会议,除了要具备编解码、链路控制、会议控制等多种基本功能外,还要能够管理用户的安全接入、权限控制等。因此,一个好的视频系统,不仅要满足用户的业务需求,支撑音视频编解码、播放、合成、录制、扩展现实(XR)以及会议等多种业务,

还要支撑用户的质量需求,从听得清、看得清到听得懂、看得懂,直至听得真、看得真,以提升用户体验质量(QoE)。对于一个企业级视频系统而言,为了提供优良的用户感知,系统除了提供音视频编解码、XR渲染等技术外,还要能够支持多租户、高并发、大流量。整个系统应稳定可靠且具有高安全性,不仅能灵活控制终端的接入和退出,还可支持云边协同,使系统部件可弹性伸缩。

视频云系统是一个基于微服务架构的云化视频业务系统。它拥有良好的软件设计和硬件架构,支持多业务、多租户、大流量,并支持灵活控制、云边部署、协同协调,可以满足平滑扩容弹缩,并有严格的安全策略设

计,以保证终端用户接入和系统管理的安全性。从视频云系统的角度来看,QoE的保障不仅要求对具体业务指标,如丢包率、抖动和时延等,进行调参调模等,还要求服务端的软件系统具有较好的稳定性和连续性(如基于微服务的视频云系统)^[1]。因此,一旦微服务本身的运行状况出现劣化,上层业务的QoE将会受到影响。基于经验来看,当前期系统调试运行稳定后,中后期系统(如微服务模块)运行状况会出现瓶颈问题,从而导致QoE下降。在系统运行过程中,有两种情况会导致QoE指标下降,且最终影响用户感知:一种是软件代码质量问题,如代码漏洞、场景考虑不周全和压力不足等,或者是系统架构设计有问题,

如无法应对数据风暴；另一种是系统本身策略性问题，如对网络感知出现异常，在需要相应微服务模块进行弹性伸缩时，策略执行动作执行得不够迅速。

因此，我们有必要为视频云系统建立一个智能运维系统。这是因为智能运维系统不仅能够监控具体业务性能指标劣化，还能监控视频业务运行的软件系统劣化，并在此基础上协助分析、定位异常和系统瓶颈，甚至能够主动运维使系统自愈，以更好地提高用户感知。

对于早期的系统运维，运维人员和软件模块开发人员根据巡检和发生的告警，分析日志和代码，来定位和解决问题，其中大多属于事后分析和人肉运维。大量的人工参与，不仅耗时而且容易出现错误，因此，如何能减少人工操作并实现故障自愈走上运维的舞台。随着人力成本的不断增高和业务场景的日益复杂，自动化运维应运而生。自动化运维不仅能够端到端地解决重复、简单而又经验化的低阶运维工作，而且在提取相关经验形成知识库后，还能根据策略定义实施故障自愈。策略闭环和专家经验知识库是自动化运维的两大基石。对于复杂的视频云系统来说，故障自愈不仅是锦上添花，更是一个必备的保障功能，因为它把运维能力从低阶提升到了中阶。

但是，随着 5G 时代来临，由于业务、场景、数据急剧增加，系统架构变得更加复杂，自动化运维已经无法有效满足数据的海量性、流程的复杂性和应用的新颖性需求。因此，在算法、大数据和算力的支撑下，借助人工智能的运维逐渐成熟起来^[2-3]。通过数据挖掘和概率统计来分析海量数据，包括业务专家和流程专家标签异常数据，并通过机器学习（包括深度学习、强化学习）训

练学习异常来提炼规则并融入知识图谱，使自动化运维进一步迈向高阶的智能运维（AIOps）^[4]。

1 视频云系统云边部署架构

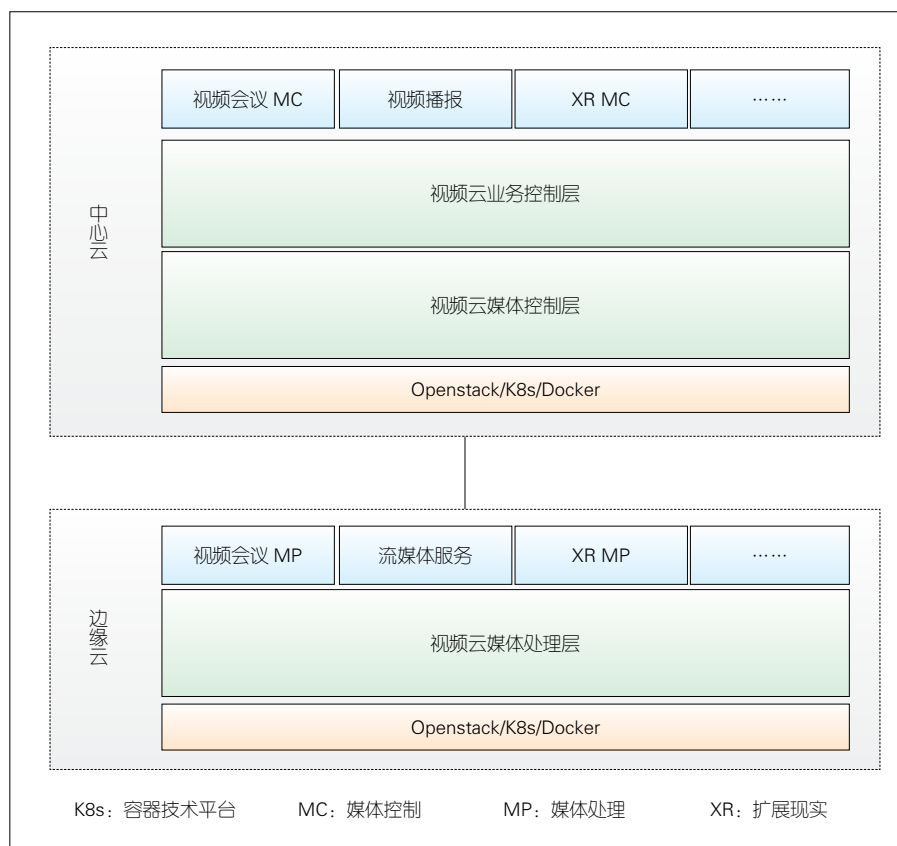
从业务角度来看，视频云系统主要提供媒体应用、媒体控制、媒体处理等服务。其中，媒体应用包括会议电视、音视频会议和直播等；媒体控制除了要对媒体应用的接入进行解析外，还需要进行业务调度和控制业务逻辑，如在视频会议中对不同用户的加入/退出、画面以及静音等的控制；媒体处理要进行音视频编解码、XR 识别等操作。

为了更快速地提供音视频服务，目前视频云系统服务端一般采用云边部署方式^[5]，即媒体面靠近用户边缘部署，控制面在云端中心部署，如图 1 所示。

中心云部署以业务控制和媒体控制为主，包括会议应用服务器、会议控制、资源控制、数据协作、消息群组、水印服务、实时录制控制。边缘云部署以媒体处理为主，其中媒体处理包括多流媒体处理、视频转码合成、实时媒体推流、实时媒体录制，可提供音视频算法库、转码和 QoE 通信引擎等。基于 Kubernetes（也称 K8s）的视频云系统支持微服务的弹性伸缩。

2 视频云系统智能运维架构

借助云边部署和云边协调，系统在中心管理控制并提供业务支持，同时计算下沉到边缘侧，能够更快速地响应和反馈，减少骨干传输网以及上层核心网的资源占用，可以很好地保障用户感知。然而，由于 QoE 的影响可能是来自边缘侧，也可能是来自中心云，甚至可能是两者交互后的结果，因此运维监



▲图 1 视频云系统云边部署示意图

管需要对云边进行统一运维。

如图 2 所示，左侧是视频云系统的中心云和边缘云，右侧是 AIOps 智能运维系统。中心云和边缘云都嵌入智能代理，分别通过代理模块将数据上报给运维系统，并先通过分析定位后再通过代理模块分别给云边下发指令进行修复自愈。整个系统实现云边运维闭环自愈。

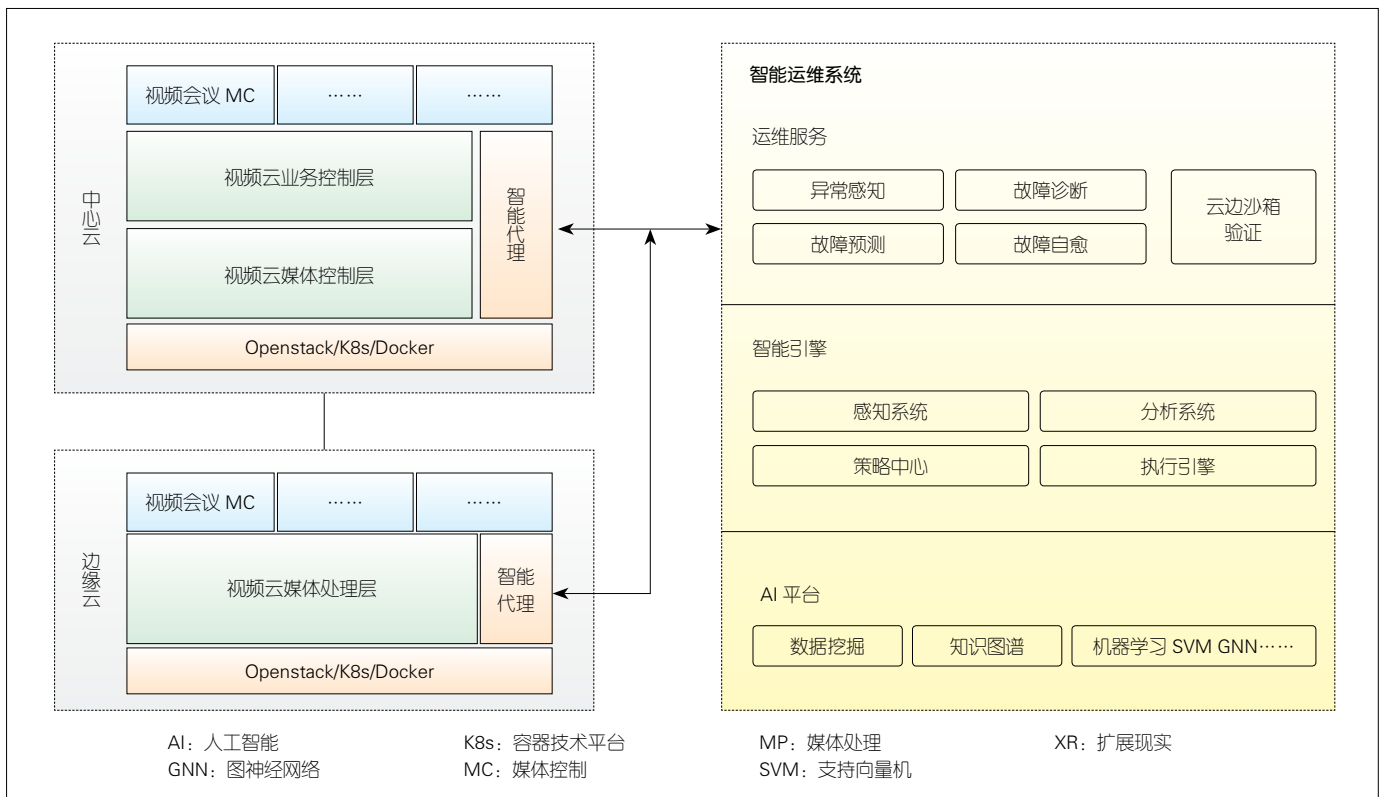
为了保障 QoE，智能代理对边缘云和中心云需要采集至少两方面数据。一个方面是运维数据，它不仅包括告警数据、性能数据、日志数据和拓扑资源数据等，还包括微服务的运行数据，如微服务的中央处理器（CPU）和内存等数据；另一个方面是业务数据，即那些会影响 QoE 的数据。运维数据和业务数据统称为感知数据。其中业务数据具体又分为 3 类：（1）网络数据，如收发端带宽、丢包率、抖动、时延等；（2）音视频参数，如帧率、

分辨率等；（3）直接体验 QoE 的终端数据，如手机型号、手机系统、会议室终端、机顶盒终端等。另外，由于视频云系统本身非常复杂，一些模块或者网络在调参（甚至硬件调整）后，需要重新模拟验证，因此一些视频云系统利用数字孪生进行仿真验证。智能代理模块也可以配置探针，采集相关影响 QoE 的孪生数据并将其上传到保障系统，以进行云边统一沙箱验证。

由图 2 右侧可以看出，运维保障系统大致包括 3 层：上层提供运维业务服务，中间框架层提供基本架构和运行支撑，底层是 AI 支撑。

上层不仅提供具体视频云系统的 QoE 运维，还提供异常感知、故障诊断、故障预测和故障自愈。异常感知是对感知数据进行异常检测，具体包括两种检测方法：一种是在数据直接异常时，有严重告警或者指标异常，例如抖动超过阈值、微服务 CPU 直接

超限等，这种异常一般可以通过阈值定义直接检测；另外一种综合判断，比如对某个时段，虽然没有指标超过阈值，但整体已经劣化。此时，可先通过人工直接标注是否异常，然后通过机器学习来学习和推理。故障诊断可对故障进行定界定位，比如，通过关联分析和根因分析，系统发现声音激励切换（VAS）模块导致终端会议在视频切换时出现卡顿等。同时，这些分析的结果将被直接注入知识图谱以供后续推理使用。故障预测可通过对历史数据进行机器学习来推断实时的运行情况，具体包括两种预测：一种是微观预测，例如根据历史数据学习来判断某微服务是否会出现内存异常；另一种是宏观预测，例如当视网络处理单元（NPU）内存消耗较高时，机器学习认为终端会议可能不会出现问题，但 XR 可能会出现问题。因此，把当前终端类型和 NPU 内存等维度输



▲图 2 视频云系统智能运维架构

入机器学习,可推理预测是否会出现劣化。故障自愈则是根据诊断定位故障原因,或者根据预测判断 QoE 是否发生劣化,并根据知识图谱学习的规则通过策略实施自愈措施,把恢复命令(如微服务弹缩或者微服务迁移指令)通过代理发送给相应系统。

中间的框架层包括感知系统、分析系统、策略中心和执行引擎。感知系统接收代理上传的数据(包括云和边的感知数据),并对这些数据进行分类、清洗和归一化。如果数据是非离散的,归一化可以按照高斯分布或者伯努力分布对其进行处理,例如高斯分布可按照马氏距离来处理。分析系统提供数据分析,如告警、性能以及日志等数据的关联分析和根因分析。策略中心定义相关策略和动作。例如,当 NPU 内存消耗达到 80% 且持续时间超过 60 s 时,系统就会执行弹缩扩容。再例如,如果 AI 预测 NPU 的内存会在视频终端用户数超过 10 个时就会冲高,那么制定策略会将其设定为当用户数超过 8 个时就弹缩(百分比、时间和数量等数据是非标准数据,在此仅做举例说明)。执行引擎保障系统自动化闭环执行,例如对数据清洗及归一化、再挖掘分析、预测、执行策略弹缩、发送指令、全程自动化。

底层 AI 支撑提供 AI 算法和训练推理框架。AI 算法包括数据挖掘和机器学习的算法,并提供知识图谱框架。其中机器学习除了涉及常规分类算法外,如支持向量机(SVM)、逻辑回归(LR)等,还用到了深度学习中的图神经网络(GNN)。知识图谱框架则提供知识图谱的创建、融合和推理等。相对于传统人工专家知识库来说,知识图谱具有更好的可视性、准确性和扩展性。

从异常感知、故障诊断到故障预测、故障自愈,是一个从被动运维到

主动运维的过程。其中,虽然故障自愈的场景不多,但是对于复杂的视频云系统来说,故障自愈是一个非常必要的功能。下面我们以后故障自愈的流程来详细说明系统的运作流程。

3 视频云系统故障自愈流程

一般来说,视频云故障自愈有如下 3 种情况:(1)当系统发生故障时,如检测到硬件温度过高,则进行微服务迁移,把有问题的硬件上的微服务迁移到备份机上;(2)当微服务性能关键性能指标(KPI)异常时,相关的微服务可能会进行弹性扩容或者提高服务等级;(3)根据业务指标超常进行微服务弹缩^[6]。

传统自动化自愈等做法就是设计门限阈值或者相关事件,制定相应策略进行迁移、扩容或者熔断等以保障自愈闭环。以微服务弹缩扩容来说,自动化运维最大的问题在于不确定性。这可能是因为指标(如 CPU、内存和带宽)抖动冲高但时间没有达标而不弹缩,也可能是无法一步弹缩到位,即会按照策略多次弹缩才能最终满足要求。还有一种可能就是策略设置错误,比如在一个周期内当最高值达到门限时就会弹缩。然而,由于内存瞬间冲高可能会回落,误弹缩的情况也会发生,一旦发生将浪费不必要的资源。

更高阶的智能化运维能够通过机器学习来寻找比较好的解决方案。基于机器学习的智能自愈,可建立在对历史数据进行机器学习、统计和分析诊断的基础上,生成相关学习模型和知识规则,在异常发生时,能够根据学习到的模型和规则,按照策略定义实施自愈手段,实现闭环自愈。此外,基于机器学习的智能运维,可以在预测服务或者系统必然劣化时,就提前采取措施实施主动运维,即把被动运

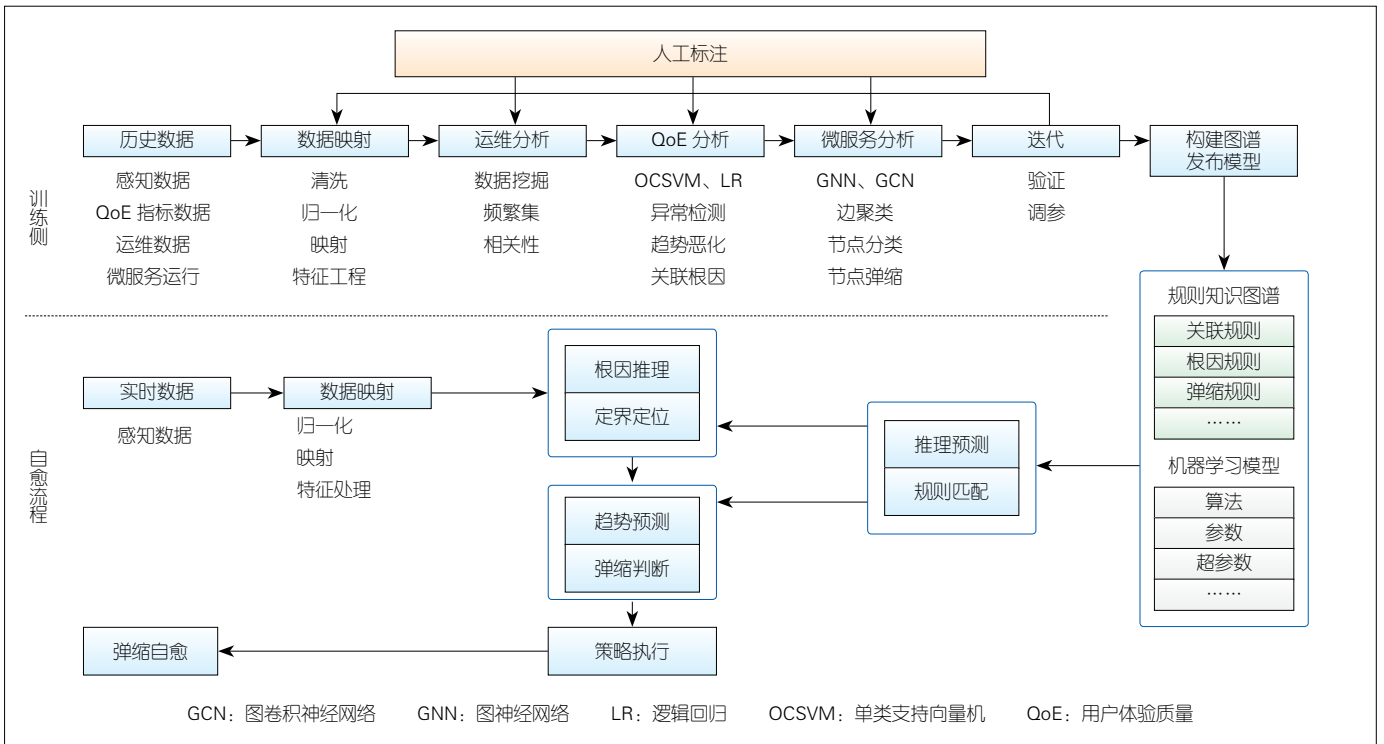
维和主动运维相结合,以保障更好的用户体验效果。

下面我们对视频云系统的故障自愈流程做一个总体描述。如图 3 所示,自愈流程分为上下两部分:上面是训练侧,下面是推理侧。

在训练侧,历史数据就是感知数据,包括 QoE 指标数据、运维数据和微服务运行数据。在采集数据后,系统会首先对数据进行清洗、归一化和映射处理,把运维数据进行频繁集挖掘并分析得到相关性,再通过机器学习对 QoE 进行分析,通过回归或分类来判断趋势是否会恶化(也可以定位相关恶化指标的关联根因)。接着,GNN 学习微服务运行趋势,得到边和节点的关系,以及是否需要弹缩的回归模型和分类模型,并进行验证。在这一阶段,如果效果不好就需要调参进行再次迭代。最终系统形成知识图谱和机器学习模型。值得注意的是,在训练侧,需要人工干预,即不仅在数据挖掘和机器学习阶段要标注,在验证阶段也要交互反馈。

在推理侧,当获得实时监控数据后,如果数据映射感知到异常,系统将进行定位诊断,即根据知识图谱得到问题根因模块,并根据机器学习模型判断是否可能会劣化、是否需要弹缩等。紧接着,系统将依据策略定义执行微服务弹性伸缩或者迁移等具体动作,最终达到自愈。

可以看出,在分析阶段进行数据挖掘后,系统采用传统的机器学习方法进行 QoE 分析,如使用单类支持向量机(OCSVM)进行异常检测,使用 LR 进行趋势是否恶化的分类判断等。由于数据挖掘技术已经比较成熟,比如工业界大多采用的频繁模式树(FP-Growth)和最大频繁项集(FP-MAX)等算法,加之传统机器学习和知识图谱构建也比较成熟,本文不再赘述。



▲图 3 视频云故障自愈流程示意图

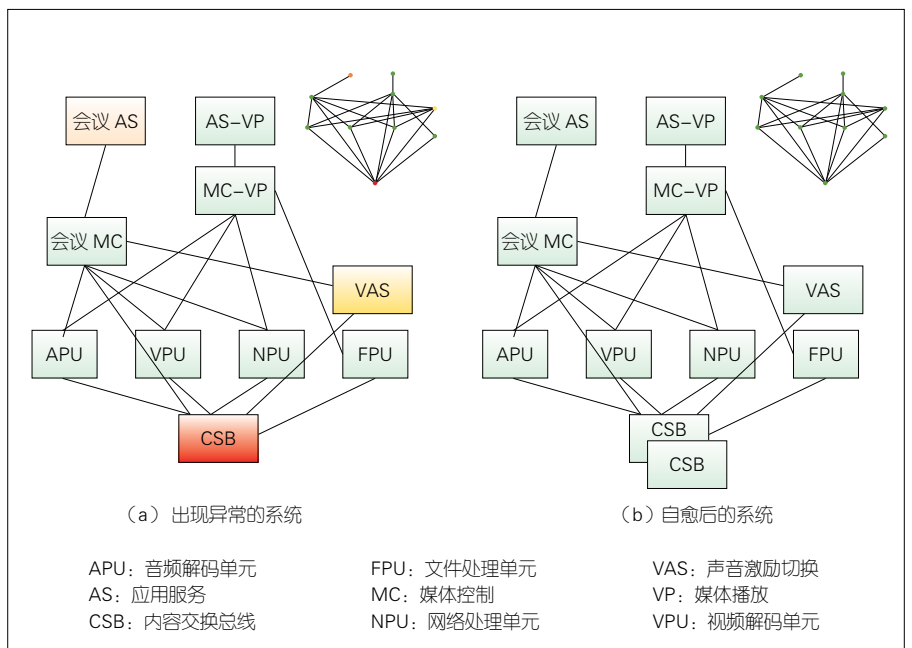
下面我们将重点介绍用于微服务判断和预测的图神经网络算法。

4 视频云系统智能运维模型算法研究

4.1 微服务系统故障分析

在讨论 GNN 算法之前，我们首先介绍基于微服务架构的视频云系统。在视频云系统中，不同微服务组件既有一定的独立性，又有一定的相关性和耦合性。它们的组织架构更像一张图，具体如图 4 所示。

视频云系统表示一个视频会议所需要的微服务局部子图。其中，应用服务（AS）为业务逻辑提供服务。媒体控制可调用不同的媒体处理单元。网络处理单元（NPU）负责收发和接收网络媒体包，同时保障媒体 QoS。内容交换总线（CSB）提供媒体数据的分发。在视频会议中，声音激励切换（VAS）可使发言者的画面被展示



▲图 4 视频云微服务部署示意（终端会议和媒体播放）

出来。这些组件与终端会议 AS、媒体播放应用服务（AS-VP）、音频解码单元（APU）、视频解码单元（VPU）均以微服务的形式部署。需要注意的是，图 4 仅是示意图，其中的 APU 和

VPU 等都以各自不同的微服务形式部署。而在实际部署时，为了实现更快的模块间交互，APU、NPU 和 CSB 可能会被部署在同一个微服务中。这是因为如果视频会议出现问题，就需要

快速定位是哪个服务出现了问题。例如，在图4(a)中，如果切换发言出现视频卡顿，终端会议AS将会出现延迟(图4中橘色模块)，同时VAS的日志将显示调用缓慢(图4中黄色模块)。一般来说，造成这种现象的原因可能是VAS、VPU或NPU出现了问题，也可能是CSB出现了问题(图4中红色模块)。此时，系统可借助模块间的快速交互，实现问题的准确定位：CSB内存超限影响VAS，导致切换激励出现延迟，进而导致AS出现卡顿。

深度神经网络(DNN)、卷积神经网络(CNN)、LR、SVM+梯度直方图(HOG)等，都是传统的机器学习和深度学习的方法。虽然它们在提取欧氏空间数据的特征方面取得巨大的成功，并在线性分类、图像分类、声音处理等相关应用上有着非常优秀的表现，但是在处理非欧式空间数据时(如社交网络、交通网络和化学分子式)，由于数据中每个节点的边或者邻域是不固定的，所以这些学习方法无法对此建立模型，即无法使用同样尺寸的卷积核来表达或者泛化。而GNN，如图卷积网络(GCN)，可同时结合图和卷积的特点，能够取得比较好的效果^[7]。例如，在参考文献[8]中，CHAI D.等利用GCN和多层全联接神经网络(MLP)模型，预测了共享单车流量问题；在参考文献[9]中，作者提出R-GCN模型，并分别将其运用到联系预测和实体分类两项任务上，在关系图的多个推理步骤中使用编码器模型来积累信息，显著改进了链路预测模型；在参考文献[10]中，YING R.等把GCN运用在推荐系统中，拼趣公司(Pinterest)在此论文基础上把GCN运用在商业系统中以推荐图片给不同用户。

如前所述，微服务方式部署实际

上是一个典型的图结构。图4中的节点就是各个微服务，边是各个服务之间的调用关系。例如，假如AS和VAS没有直接的调用关系，那么它们两个节点之间就没有相对应的边。因此，视频云运维系统在对图进行学习和预测时，不仅采用了传统的机器学习算法(例如OCSVM和LR等)，还结合了GCN等算法。需要注意的是，在真实视频云系统中，微服务数量有上千个，不同服务之间的关联更加复杂，微服务部署的复杂度远远超过图4所示的结构。因此，简单的图搜索和图嵌入都不能进行有效的推理和根因定位。

4.2 GNN

GNN是一个很宽泛的概念。一般来说，GNN就是图+神经网络。目前，与GNN相关的模型算法有GCN、图注意力网络(GAT)等。本文中，我们以GCN为例来做具体讨论。

一般来说，GNN可被用来处理3类任务：

- 节点层面任务。该任务主要包括对微服务节点进行分类和预测，例如判断这个节点和其他节点是否属于同一类，或者当该节点的业务数据有异常时，是否需要微服务弹缩或

迁移等。

- 边层面任务。该任务与微服务直接相关，比如微服务调用链、根因分析等。

- 图层面的任务。该任务可对整个模块或者子图进行分类预测，例如预测整个视频云是否正常。

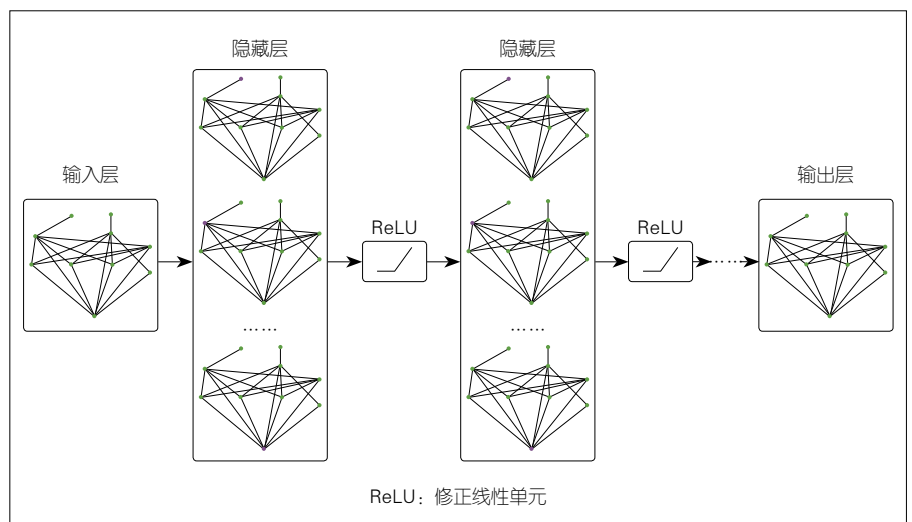
我们提出的视频云保障系统主要考虑节点层面任务和边层面任务。我们需要首先对微服务的组成架构进行图表示，然后再进行标注和训练以获得模型。

我们构建微服务图，其中GCN把整个图 G 、每个节点 V 、每条边 E 转化为稠密向量。当然，并不是每次都要把 G 、 V 、 E 进行向量化的，哪部分需要向量化取决于实际的应用场景。

以图4中的微服务为例，我们搭建了一个GCN网络，如图5所示。其中，左边为输入层，右边为输出层，中间有2个或3个隐藏层，并且每个隐藏层的激活函数均为修正线性单元(ReLU)。

以图4的子图数据为例，我们对其进行迭代传播训练。相关传播公式如公式(1)~(3)所示。

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}), \quad (1)$$



▲图5 神经网络

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}, \tag{2}$$

$$\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}, \tag{3}$$

其中， σ 是非线性激活函数（这里为 ReLU）， \mathbf{D} 是度矩阵， \mathbf{I} 是单位矩阵， \mathbf{W} 是卷积权重矩阵。传播过程中有 N 个节点（在图 4 中， $N=10$ ），每个节点都有自己的特征，包括 CPU 值、内存大小、业务指标、KPI 等 M 个维度的数据。这些节点的特征数据经过归一化处理首先组成一个 $N \times M$ 维的矩阵 \mathbf{H} ，然后各个节点之间的关系也会形成一个 $N \times N$ 维的矩阵 \mathbf{A} （也称为邻接矩阵）。这里， \mathbf{H} 和 \mathbf{A} 是我们模型的输入。通过 \mathbf{D} 和 \mathbf{A} 可计算并形成拉普拉斯矩阵。

以图 4 的子图为例，图 6 是邻接矩阵和和度矩阵的实际数据。通过这些数据计算得到的拉普拉斯矩阵，随后将进行公式（1）—（3）中的卷积运算。

整个学习过程为有监督学习。系统先根据业务异常对节点进行标签，然后把整个数据按 7 : 3 的比例拆分成训练数据和验证数据。损失函数可以是比较通用的交叉熵损失函数。通过训练学习，系统可以得到不同节点之间的分类、关联程度和节点是否异常的模型。

通过学习我们可以看出，由于受到相邻和其他节点的影响（相邻关系越近，影响越大），图中的每个节点都在时刻改变着自己的状态，直至平衡。这里，我们仍然以图 4 的子图为例，VAS 和 CSB 就是同一类节点，因此系统可以通过它们来确定关联关系和根因关系。

此外，有 3 点需要注意：（1）前文所述的传播表达方式是基于谱域方法而提出的。但是对整个网络来说，由于节点和边的关系非常复杂，除了

内存消耗巨大外，对 \mathbf{D} 和 \mathbf{A} 的求逆和行列式计算也将非常耗时，因此很多优化方法目前被引入进来，比如基于空域的方法和节点采样的方法等。

（2）和 DNN 近似无限表达能力不一样，GNN 的表达能力是比较受限的^[11]，当然，在云视频微服务系统中，GNN 的表达能力是完全能够覆盖远不足万计的微服务。（3）在一般的网络搭建中，GCN 后面会再设置一层 MLP 以协助业务判断。由于这也是通用模型，不是本文讨论的重点，故这里不再赘述。

5 视频云系统智能运维能力的提升效果

视频云微服务系统是基于 K8s 集群进行部署的，它主要通过配置运维策略、实时采集指标、可视化监控、故障工单等，来实现信息技术运维（ITOps）。AIOps 巧妙地将机器学习和知识图谱相结合，取得了更好的运维效果。下面我们以微服务弹缩自愈场景来进行对比说明。一般来说，微服务指标包括 CPU 值、内存、输入输出（IO）以及 Java 虚拟机（JVM）内存等。传统 ITOps 对 CPU 指标超限和

# 邻接矩阵										
#AS1	MC1	APU	VPU	NPU	VAS	FPU	CSB	AS2	MC2	
0	1	0	0	0	0	0	0	0	0	#AS1 会议 AS
1	0	1	1	1	1	0	1	0	0	#MC1 会议 MC
0	1	0	0	0	1	0	1	0	1	#APU
0	1	0	0	0	1	0	1	0	1	#VPU
0	1	0	0	0	1	0	1	0	1	#NPU
0	1	1	1	1	0	0	1	0	0	#VAS
0	0	0	0	0	0	0	1	0	1	#FPU
0	1	1	1	1	1	1	0	0	1	#CSB
0	0	0	0	0	0	0	0	0	1	#AS2 AS-VP
0	0	1	1	1	0	1	1	1	0	#MC2 MC-VP

# 度矩阵										
#AS1	MC1	APU	VPU	NPU	VAS	FPU	CSB	AS2	MC2	
1	0	0	0	0	0	0	0	0	0	#AS1 会议 AS
0	6	0	0	0	0	0	0	0	0	#MC1 会议 MC
0	0	4	0	0	0	0	0	0	0	#APU
0	0	0	4	0	0	0	0	0	0	#VPU
0	0	0	0	4	0	0	0	0	0	#NPU
0	0	0	0	0	5	0	0	0	0	#VAS
0	0	0	0	0	0	2	0	0	0	#FPU
0	0	0	0	0	0	0	7	0	0	#CSB
0	0	0	0	0	0	0	0	1	0	#AS2 AS-VP
0	0	0	0	0	0	0	0	0	6	#MC2 MC-VP

APU: 音频解码单元	FPU: 文件处理单元	VAS: 声音激励切换
AS: 应用服务	MC: 媒体控制	VP: 媒体播放
CSB: 内容交换总线	NPU: 网络处理单元	VPU: 视频解码单元

▲图 6 邻接矩阵和度矩阵数据

一问题能够处理得很好，但是对于内存超限却很难判断，这是因为内存会出现回收的情况。下面我们将对比测试这两种情况。

视频云的微服务蓝图主要包括：逻辑主机（Pod）数量共 1 个，CPU 为 2 核，内存为 2 GB。传统 ITOps 定义策略为：感知时间为 1 min，当采集间隔时间为 5 s，即 1 min 采样 12 次，弹缩消耗时间为 1 min。在感知时间内，采样中平均 CPU 消耗达 85% 时弹缩 1 个 Pod，内存消耗超过 80% 时弹缩 1 个 Pod。测试时，我们按最终弹缩 4 个 Pod 为例，具体测试效果如图 7 所示。

由图 7 可知，传统 ITOps 的弹缩是台阶式弹缩，即感知、弹缩、再感知、再弹缩，最终感知没有异常时，则弹缩到位。在一次感知后，AIOps 根据机器学习到的模型，直接进行一步到位的弹缩。图 7 的测试是在假设感知时间是一致的条件下进行的。而实际上，AIOps 的感知时间是按过去的时间窗口来预测的，它的感知时间会远远小于 1 min。但即便是对于简单的场景，AIOps 的效果也会远远好于基于策略的 ITOps。

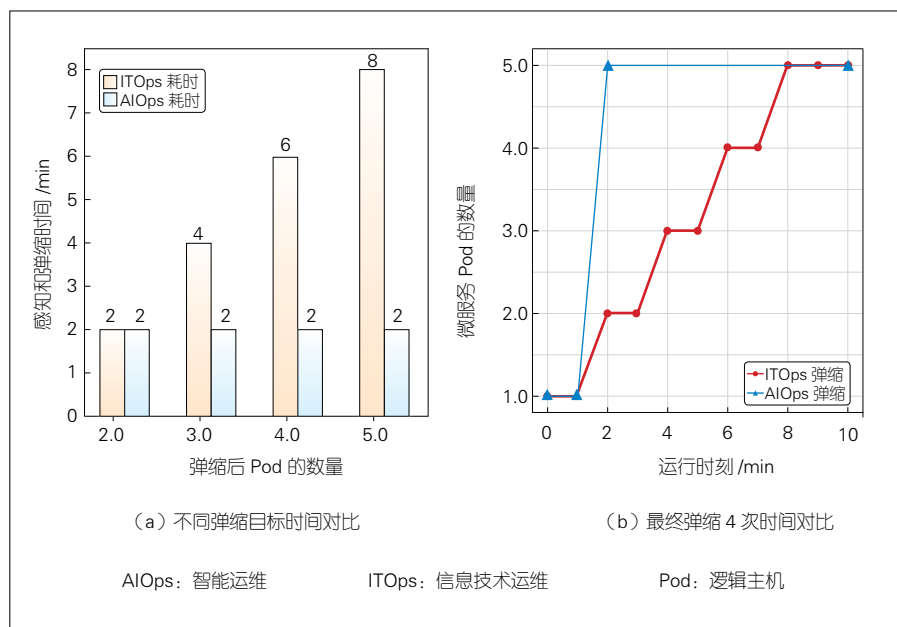
从分类是否正确的角度来看，即是否应该弹缩和是否不弹缩来看，ITOps 根据策略定义来分类，AIOps 根据机器学习来分类。这里，我们定义 TP 为应该弹缩，实际弹缩；FP 为不应该弹缩，实际弹缩；FN 为应该弹缩，实际未弹缩；TN 为不应该探索，实际未弹缩。测试显示，对于 ITOps 来说，TP=67，FP=7，FN=6，TN=20；对于 AIOps 来说，TP=65，FP=22，FN=8，TN=5。表 1 给出了在测试环境中 100 次弹缩是否准确的评价数据统计（统计公式见参考文献 [12]）。

由表 1 可知，AIOps 在分类方面的表现比完全根据策略定义的 ITOps 更好。

除了微服务弹缩，在整个系统运维保障能力方面，AIOps 比 ITOps 也有极大的提升，包括异常感知、故障诊断、故障自愈和故障预测等技术指标。以图 4 的终端会议和媒体播放局部场景为例，该场景共有 61 类不同告警。故障平均修复时间（MTTR）包括故障感知、故障定位、故障诊断和故障恢复。由表 2 可知，与 ITOps

相比，AIOps 至少提升了 60% 的故障修复效率。

除了在质量保障和效率方面的提升外，AIOps 在成本优化方面也有较大的提升。由于成本优化是一个比较复杂的课题，而本文的阐述重点是质量保障和故障自愈，因此，本文中我们仅以视频云微服务系统的 K8s 集群的成本，来做对比说明下。如表 3 所



▲图 7 AIOps 和传统 ITOps 自愈时间对比

▼表 1 AIOps 和传统 ITOps 分类评价

运维系统	准确率 /%	精确率 /%	召回率 /%
ITOps	70	84	89
AIOps ^[19]	87	90	92

AIOps: 智能运维 ITOps: 信息技术运维

▼表 2 AIOps 和传统 ITOps 故障处理效率对比

运维系统	故障感知时间 /min	故障定位时间 /min	故障诊断时间 /min	故障恢复时间 /min
ITOps	15 ~ 20	> 60	> 50	15
AIOps	< 5	< 15	< 20	15

AIOps: 智能运维 ITOps: 信息技术运维

▼表 3 AIOps 和传统 ITOps 对 K8s 集群的成本优化对比

运维系统	Pod 资源优化 /%	非忙时间资源缩减 /%	云存储成本	GPU 服务器	调整集群大小
ITOps	11	10	< 5%	< 5%	< 5%
AIOps	14	17	22%	21%	13%

AIOps: 智能运维 GPU: 图形处理器 ITOps: 信息技术运维 Pod: 逻辑主机

示, 通过策略和人工经验的 ITOps, 仅在 Pod 资源优化和非忙时间资源缩减方面有一定成效而 AIOps 在云存储成本、GPU 服务器和调整集群大小等方面表现更显著。这是因为 AIOps 通过历史数据的机器学习和数据挖掘, 在成本优化方面做得更科学、更深入, 而不是简单依靠人工经验等给出策略来控制成本。

6 结束语

在面向视频云系统的云边端复杂部署场景时, 系统的运维保障变得极为复杂。AIOps 技术能够有效地提升视频云的运维能力和 QoE 等级。与传统 ITOps 相比, AIOps 技术不仅能够节约运维成本、提升运维效率, 还能够实现更加精准的运维服务、提升用户体验。

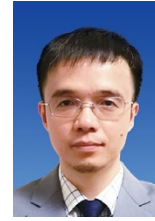
致谢

本研究得到中兴通讯股份有限公司胡锐、付迎春、周祥生、弄庆鹏等的帮助, 谨致谢意!

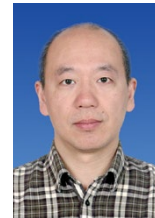
参考文献

- [1] ZHANG X, XIE L, GUO Z. Quality assessment and measurement for internet video streaming [J]. ZTE communications, 2019, 17(1): 12-17. DOI: 10.12142/ZTECOM.201901003
- [2] 董德尊, 欧阳硕. 分布式深度学习系统网络通信优化技术 [J]. 中兴通讯技术, 2020, 26(5): 2-7. DOI: 10.12142/ZTETJ.202005002
- [3] 李建飞, 曹畅, 李奥, 等. 算力网络中面向业务体验的算力建模 [J]. 中兴通讯技术, 2020, 26(5): 34-38. DOI: 10.12142/ZTETJ.202005007
- [4] AIOps 标准工作组. 企业级 AIOps 实施建议白皮书 [R]. 2018
- [5] 高志鹏, 尧聪聪, 肖楷乐. 移动边缘计算: 架构、应用和挑战 [J]. 中兴通讯技术, 2019, 25(3): 26-27. DOI: 10.12142/ZTETJ.201903004
- [6] 龚正, 吴治辉, 王伟, 等. Kubernetes 权威指南: 从 Docker 到 Kubernetes 实践全接触 [M]. 北京: 电子工业出版社, 2017: 93-171
- [7] WU Z H, PAN S R, CHEN F W, et al. A comprehensive survey on graph neural networks [EB/OL]. [2020-11-10]. <https://arxiv.org/abs/1901.00596>
- [8] CHAI D, WANG L Y, YANG Q. Bike flow prediction with multi-graph convolutional networks. [EB/OL]. [2020-11-10]. <https://arxiv.org/abs/1807.10934>
- [9] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks [C]//European Semantic Web Conference. Crete, Greece: Springer, 2018: 593-607
- [10] YING R, HE R Y, CHEN K F, et al. Graph convolutional neural networks for web-scale recommender systems [EB/OL]. [2020-11-10]. <https://arxiv.org/abs/1806.01973v1>
- [11] RYOMA S. A survey on the expressive power of graph neural networks [EB/OL]. [2020-11-10]. <https://arxiv.org/abs/2003.04078>
- [12] 李航. 统计学习方法 [M]. 北京: 清华大学出版社, 2012: 10-19

作者简介



徐代刚, 中兴通讯股份有限公司网络智能平台研发总工、OES 运营技术专家委员会主任及首席架构师; 研究方向为电信运营系统微服务架构和云化网络智能技术。



姜磊, 中兴通讯股份有限公司网络智能平台高级架构师、OES 运营技术专家委员会委员; 研究方向为电信网络智能运维、数据挖掘、机器学习和深度学习。



梅君君, 中兴通讯股份有限公司视频云平台研发总工、大视频技术专家委员会委员; 研究方向为视频能力基础设施云原生架构和低延时高性能实时音视频通信技术。