

## 智能自适应边缘系统：探索与挑战

王旭<sup>1</sup>, 陈南希<sup>1,2</sup>, 张柔佳<sup>1,3</sup>

(1. 中国科学院上海微系统与信息技术研究所, 上海 200050; 2. 中国科学院大学, 北京 100049;  
3. 华为技术有限公司上海研究所, 上海 201206)

**摘要:** 边缘智能已成为新一代物联网的发展趋势。边缘计算设备地理分布广, 设备种类多, 服务多样化, 时延敏感, 终端具备移动性。因此, 边缘系统需要提供灵活多样的、可重构可扩充的服务。通过将自适应思想融入边缘计算, 首先探索了智能自适应边缘系统应用需求, 分析并总结了现有自适应边缘系统基础框架, 并将深度学习、强化学习等人工智能技术应用于自适应边缘系统。然后, 介绍了如何在特定的应用领域设计专门的智能算法。最后, 探讨了该领域的发展潜力以及未来面临的挑战。

**关键词:** 边缘计算; 自适应系统; 物联网; MAPE-K 循环控制

**中图分类号:** TP393.0

**文献标识码:** A

**doi:** 10.11959/j.issn.2096-3750.2021.00210

## Intelligent adaptive edge systems: exploration and open issues

WANG Xu<sup>1</sup>, CHEN Nanxi<sup>1,2</sup>, ZHANG Roujia<sup>1,3</sup>

1. Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China  
2. University of Chinese Academy of Sciences, Beijing 100049, China  
3. Shanghai Research Institute, Huawei Technology Co., Ltd, Shanghai 201206, China

**Abstract:** Edge intelligence has emerged as a promising trend of the new generation of Internet of things. Edge computing devices are widely distributed, with various diverse end devices and services, delay sensitive, and serve mobile terminals. Therefore, the edge system needs to provide flexible, diverse, reconfigurable and scalable services. From the application fields of adaptive edge computing, the application requirements of intelligent adaptive edge systems were explored, the existing adaptive edge systems and their basic framework were analyzed and summarized, and the application of artificial intelligence technologies was discussed, such as deep learning and reinforcement learning. Then, how to design a special intelligent algorithm in specific application fields was introduced. Finally, the research status and future challenges in this field were discussed.

**Key words:** edge computing, adaptive system, Internet of things, MAPE-K control loop

### 1 引言

物联网、人工智能、5G 的全面发展将推动更多智能应用的产生, 同时也对底层计算与通信系统架构提出了新的挑战。目前, 边缘计算已成为智能应用架构技术的新趋势。如在智能交通系统中, 自动驾驶车辆可以受益于移动边缘计算提供的低时

延通信, 以协调其行为并实时做出决策<sup>[1]</sup>; 视频<sup>[2]</sup>、在线游戏等应用<sup>[3]</sup>也能够利用边缘服务获得高效的计算、通信服务, 从而提升用户的体验质量 (QoE, quality of experience)。处理数据爆炸和网络流量、对更智能的计算范式的需求以及可持续的能源消耗等, 是许多应用选择边缘计算的关键原因<sup>[4]</sup>。借助边缘计算, 通过在各种应用中采用不同的电子设备、

收稿日期: 2021-01-11; 修回日期: 2021-02-03

通信作者: 陈南希, nanxi.chen@mail.sim.ac.cn

基金项目: 上海市青年科技英才扬帆计划资助项目 (No.19YF1455900)

Foundation Item: Shanghai Sailing Program (No.19YF1455900)

传感器和执行器,来改善如智能家居、监视系统、环境保护、车辆交通和运输、医疗保健以及天气和水资源管理系统等的服务质量(QoS, quality of service),智能城市等概念将更接近现实<sup>[5]</sup>。

物联网节点由于受功耗、体积、成本等因素的约束,计算能力较弱。云服务器一般用于集中控制,负责处理全局性的任务。由于云服务器距离物联网节点较远,网络带宽有限,很难提供实时的数据处理服务。边缘计算使计算靠近数据源(如网络边缘),提供计算处理、存储、缓存等服务供数据使用者访问<sup>[6-7]</sup>。因此,边缘节点的设计必须满足如低时延、上下文感知和移动性支持等要求,并且所提供的服务必须可靠且可扩展<sup>[8]</sup>。然而,大多数服务依赖于在环境中动态移动的多个设备的协作与资源协调。在这种不确定且高度动态的环境中,无法预料的事件(如设备故障或网络不可用)将会影响整个系统。因此,边缘计算必须能够适应这些变化并自我组织、管理资源以保证正常运行并保持服务稳定。

本文从自适应边缘计算的适用领域出发,探索了智能自适应边缘系统的应用需求,并且根据经典自适应系统的设计原则,分析现有自适应边缘系统,总结自适应边缘系统的基础框架,并将深度学习、强化学习等人工智能技术应用到自适应边缘系统框架中。然后,介绍了如何在特定的应用领域设计专门的智能算法。最后,探讨了该领域的研究现状,并从协调与合作的角度讨论未来面临的挑战。

## 2 自适应边缘系统的特点

### 2.1 时延敏感

许多应用对时延很敏感,因此需要在网络边缘直接进行数据处理<sup>[9-12]</sup>。在同一应用中,不同的服务对时延的要求不同。如在自动驾驶汽车的车载应用中,可能同时存在地图定位、视频播放、障碍物检测 3 种服务。地图定位对时延的容忍度较高;视频播放则对时延要求较严格;车辆的运动速度很快,并且与人的生命健康直接相关,所以必须将障碍物检测的时延减小到最低程度。这就要求边缘服务器能够根据设备的需求,自适应地调整各类计算任务的优先级。如当同时有地图定位类和视频播放类的服务需求时,优先保障视频播放任务的服务需求;若出现行驶需求或遇到较差路况时,则需要提升障碍物检测服务在资源利用中的优先级。

### 2.2 网络系统开放

在边缘计算网络运行过程中,不断有新的设备加入,同时也有设备因为故障等原因退出,这导致其拓扑结构处于不断变化的状态<sup>[13-14]</sup>。此外,随着科技的发展,不断有新的设备被发明并且投入使用。如传统的收银服务仅支持现金收款,后来能够支持 POS(point of sales) 机刷卡支付、手机扫码支付,最新的服务还能够支持基于人脸识别、指纹/掌纹识别等技术的支付方式。因此,实现同一功能的服务还可能涉及多个不同类型的终端设备或终端应用<sup>[15-16]</sup>,所以边缘计算要能够根据系统中设备的种类、规模等的变化,不断调整其服务。

### 2.3 终端可移动

随着无线通信技术的迅速发展,越来越多的终端设备通过蜂窝网或 Wi-Fi 与边缘服务器通信,而移动设备与边缘服务器之间的网络拓扑结构随着时间的变化不断发生变化<sup>[16]</sup>。如某个设备前一时刻在这个位置,请求此处的边缘服务器进行计算,在边缘服务器完成计算准备反馈结果时,设备可能已经移动到另一个位置了。这就要求边缘服务器能够跟踪设备位置的变化,及时将计算结果反馈给设备<sup>[14]</sup>。同时,物联网中可移动的设备也需要能够在自身位置发生变化后,及时搜索网络环境,更新邻近的边缘服务器列表。

### 2.4 环境不断变化

物联网设备工作在真实的物理环境中,会受到多种外部环境因素的影响,如刮风、下雨、噪声、干扰、光照、温度、障碍物等<sup>[17]</sup>。其外部环境的变化不受软件系统的控制,所以无法对环境进行如调度、调优等操作,即便监测到环境发生变化,也会受传感器能力的限制难以及时处理<sup>[18]</sup>,这就要求边缘计算能够灵活地根据环境的变化调整自身的架构以及配置。另外,蜂窝网或 Wi-Fi 通信所依赖的无线信道也随时间变化而不断变化,所以网络的传输带宽也随之变化,因此边缘服务器要能够根据网络带宽的变化对自身进行调整<sup>[19]</sup>。

### 2.5 电力供应受限

由于传统的计算机系统通过交流电供电,所以不用考虑电源供电不足。而许多物联网设备、移动设备等采用电池供电,电池电量是有限的,在处理当前任务时,还要为后面可能的任务预留一定的电量<sup>[20]</sup>。这就要求物联网设备结合以往的经验、目前的任务状况以及当前的电池情况综合考虑,合理地

分配电池电量。近年来，一些边缘服务器开始使用可再生能源作为电力来源<sup>[17]</sup>，如风能、太阳能等。与火力发电等传统供电方式相比，可再生能源分布广泛，对环境无污染，但电力供应不稳定，容易受自然环境的影响。因此，当使用可再生能源的边缘服务器时，需要时刻监控自然环境的变化，对将来一段时间可能新充入的电量进行预测，并在此基础上，对计算任务进行合理调配<sup>[21]</sup>。

### 3 自适应边缘系统中的技术挑战与关键技术

自适应是指边缘计算系统在没有用户控制或仅由用户进行最小控制的情况下决定如何调整<sup>[1]</sup>。从体系结构的角度来看，自适应系统一般由两部分组成，包括管理系统和受管系统。受管系统是一组可配置的软/硬件资源，如边缘节点、网络带宽、组合服务等。管理系统也称为自适应逻辑，是一组软件模块，负责观察环境和管理受管系统，分析对适应更新的需求，计划此类适应操作以及控制适应操作的执行。智能自适应边缘系统框架如图 1 所示。

MAPE-K 控制框架是自适应系统中一种常用的控制方法<sup>[22]</sup>，主要由监测、分析、规划、执行、知识库这 5 个部分组成<sup>[18]</sup>。MAPE-K 控制框架是一种通用的方法，其思想广泛存在于各类自适应边缘计算场景中。本文将从监测、分析、规划、执行、知识库 5 个角度分析现有的自适应边缘系统，并将

深度学习、强化学习等人工智能技术应用于 MAPE-K 控制框架中。然后，介绍了一些其他自适应控制框架。

#### 3.1 自适应边缘系统中的技术挑战

自适应系统能够根据对环境和系统本身的感知来调整自身结构和相关配置，其关键是感知和适应。在边缘计算中，比较典型的感知有自我感知<sup>[23]</sup>和上下文感知<sup>[17]</sup>。其中，自我感知能够意识到关于系统资源的变化；上下文感知能够获取系统执行时计算环境和物理环境中出现的变化信息。一般而言，系统获取内部资源变化是相对容易的，可以通过日志、设置探针等手段获取内部资源变化。为了实现上下文感知，系统需要使用传感器收集环境的信息，再根据这些信息进行推理<sup>[24]</sup>。

当自适应系统感知到变化信息后，要对相关信息进行分析，确定异常的严重程度，并推测引起异常的原因，从而决定是否采取措施进行补偿。物联网系统中节点众多、功能复杂，根据少量观测到的异常值准确推测出引起异常的原因是较难实现的，近年来新兴的深度学习技术在这方面有很大的应用潜力。当完成异常信息分析之后，自适应系统还要规划如何补偿异常。物联网系统一般涉及多个功能模块，对系统进行完整的理论分析较困难。传统的做法依赖于以往的经验，并根据经验建立知识库，而近年来兴起的强化学习技术则能够自动生成最优的补偿策略。

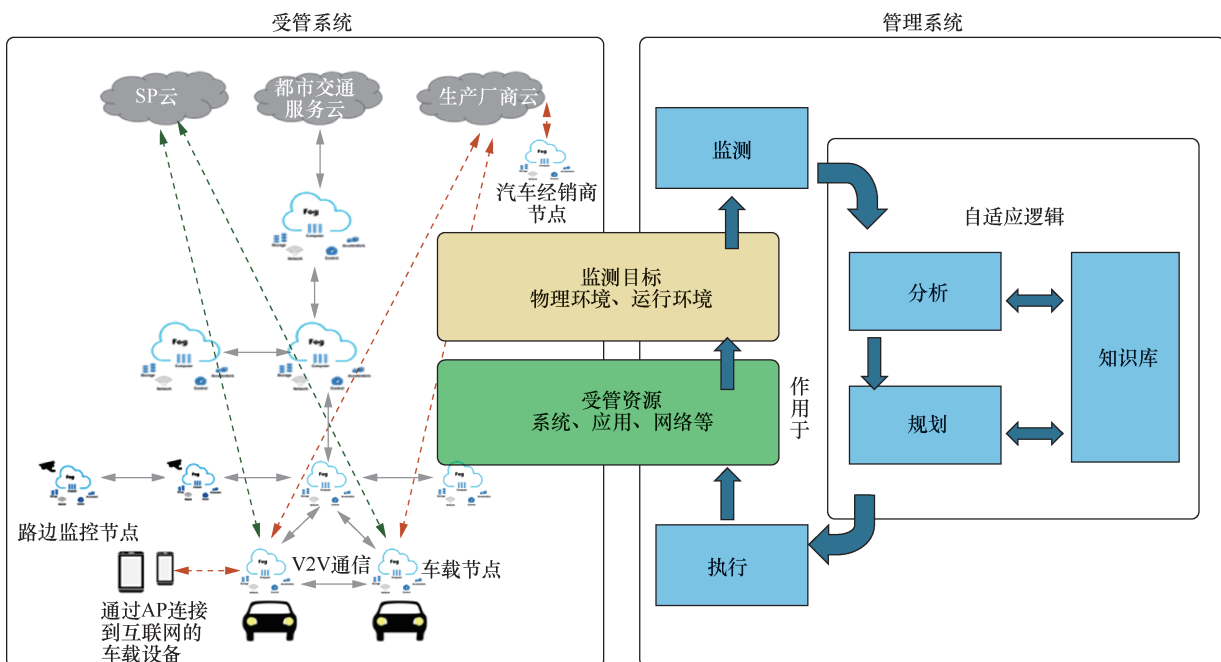


图 1 智能自适应边缘系统框架

## 3.2 自适应边缘系统中的关键技术

### 3.2.1 环境变化的监测

自适应系统中的监测首先要设定监测目标,当监测目标与预期状态偏离时,系统分析原因并寻找合适的方法进行补偿。

一般常见的监测目标来源于周围物理环境。对于智能家居<sup>[25]</sup>系统而言,监测目标可以是光照、室温、噪声等;对于智能医疗系统<sup>[26]</sup>而言,监测目标可以是病人的心跳、体温、血压等;对于智能停车系统<sup>[27]</sup>而言,监测目标可以是停车场的空位数量以及空位的分布等。环境监测可通过传感器搜集和控制目标相关的数据。如果有需要,还可以在监测过程中添加新的数据源。当传感器搜集到数据之后,将比对预设的期望值以确定监测目标是否出现偏离<sup>[28]</sup>。

计算任务本身的运行环境可以是自适应边缘系统的监测目标。为了满足计算任务的 QoS 与 QoE 需求,将监测可以分为 4 个层级<sup>[29]</sup>:系统级(如操作系统、虚拟机)、计算平台级、网络连接级、应用级。系统级的监控一般通过系统日志及相关驱动提供的监测接口实现;一般计算平台(如 Docker)自带监控自身状态的命令,利用第三方工具(如 cAdvisor、Prometheus、DUCP、Scout)提供额外的监控功能<sup>[29]</sup>。网络连接的主要监测目标包括吞吐量、时延、分组丢失率、抖动等<sup>[29]</sup>。设备之间的网络连接也是常见的监测目标,一般可以通过设备间周期性发送短小的数据分组来实现。应用级的监控一般通过在执行具体计算任务的相关服务中预先设置探针来实现<sup>[28]</sup>。如果服务开发者对与任务关系密切的环境因素进行预先分析并设置好监控端口,边缘服务器就能在资源短缺时更好地调度相关资源来满足需求。

### 3.2.2 基于深度学习的自适应分析

自适应分析过程将通过评估监测过程传递的数据的异常程度,来判断是否采取补偿行动。一般来说,如果异常只是一个小范围的随机波动,那么将不采取任何行动。如果数据的异常程度影响了自适应系统目标(如关键任务的运行时长超过其时延允许上限、系统资源短缺影响 QoS),则必须采取补偿措施。此时,还要进一步分析异常的特征,以此推测引起异常的可能原因。

简单的参数异常可通过传统的数值比对方法进行判断<sup>[18]</sup>。当监测目标种类较多或监测目标指向

的异常类型较多时,则需要更复杂的自适应分析方法来判断。近年来,许多研究工作开始尝试使用基于深度学习的自适应分析方法<sup>[22,30-31]</sup>。如在机场、港口等建有大型仓库的地方,需要按照消防安全的要求对货物分类存储。文献[30]提出利用射频识别(RFID, radio frequency identification)技术对货物的种类、数量识别后,形成该批货物的环境信息。然后利用深度学习技术对环境信息进行特征提取,并结合以往货物信息的知识库,提出货物存放方法的建议。仓库管理员对建议进行评估,并将评估结果反馈给知识库,从而实现知识库的动态更新。在物联网的网络安全方面,传统的基于统计分析的分布式拒绝服务(DDoS, distributed denial of service)攻击方法需要人工干预,所以无法处理加密的 DDoS 流量,基于深度学习的 DDoS 方法<sup>[32]</sup>就很好地解决了这个问题。另外,对于零日漏洞攻击来说,传统的基于 ECC-memory 的监测方法需要源代码,而边缘服务器经常无法做到这一点。使用深度学习之后,可以直接对内存进行分析,并标出可疑的地方<sup>[33]</sup>。

### 3.2.3 基于强化学习的规划更新

规划更新的目标是寻求补偿监测目标异常的方法。常见的解决方案是根据自适应分析得到的异常原因,在知识库查找可能的补偿方法。如果之前出现过类似的异常,但还没成功地补偿该异常,那么规划更新过程将继续尝试其他可能的解决方法。如果规划更新过程已经尝试了所有可能的解决方案但都不能解决该异常,那么该过程将决定是否触发异常警报,从而请求其他手段(如人工干预)来解决问题。

在物联网应用运行的过程中,周围的环境以及用户的目标都可能发生变化,并且这些变化可能完全超出系统设计时的预估,因此预先设定的补偿规则可能失效。因此,自适应边缘系统要进行在线规划更新。强化学习在理论上可以自动生成最优的更新策略,近年来在许多解决方案中得到了应用<sup>[17,22-23,34-35]</sup>。强化学习本质上是一种试错学习,系统需要不断与环境交互,在尝试中逐渐学会从环境状态到自适应策略的最优映射。如果环境发生了变化,系统要先付出一定的成本,然后才能逐渐调整系统的状态。但在一些应用领域(如自动驾驶),试错成本带来的损失不可估量。目前的解决方案一般面向具体自适应目标,如文献[17]提出根据预计的充电电量和将来的任务,利用强化学习技

术设计计算卸载策略; 文献[22]将知识库与强化学习结合来降低新策略的探索代价; 文献[34]提出每个节点利用强化学习从以往的经历中生成选择策略, 从而克服了单个节点缺乏全局信息的困难; 文献[35]提出利用强化学习优化网络连接。

### 3.2.4 策略执行与知识库的利用

策略执行过程具体执行规划更新过程生成的自适应策略, 再依据该策略调整系统资源或参数, 如调整开关状态<sup>[24]</sup>和系统配置<sup>[28]</sup>、修改视频分辨率<sup>[36]</sup>、调整运行时的资源分配情况等<sup>[37]</sup>。

在执行策略之后, MAPE-K 控制框架将继续监测控制目标, 并不断进行分析、规划、执行等步骤, 将相关结果记录在知识库中。随着 MAPE-K 控制框架的不断运行, 知识库中积累的经验越来越多, 后续出现问题可以更快地找到解决方案。

在自适应系统的知识库中, 一般记载了规划更新过程中面临的各类异常的解决方案、环境监测和自适应分析过程所需要的经验知识。如对于环境监测而言, 需要明确怎样的数据偏离程度可以认定为数据异常。换言之, 环境监测需要设置合适的数据异常阈值<sup>[24]</sup>。部分场景下可以通过理论分析得到最优阈值, 而实际中更多情况需要根据经验来设置最优阈值。利用知识库可以在缺乏经验的情况下, 协助自适应分析过程实现对该阈值的自动调优。如可以在开始时设置一个较小的阈值, 保证所有数据异常都能被捕捉。随着时间的推移, 知识库中存储的案例越来越多。监测器通过对以往案例的分析, 逐步增大阈值, 以降低虚警概率, 达到最优的阈值。文献[18]针对外界环境和设备自身状态进行建模, 并利用存储环境感知信息的知识库预存外界环境或自身状态变化时所对应的自适应策略。文献[22]的知识库则是先预存根据历史经验生成的一组策略, 从而形成一个初始的离线知识库。然后在运行过程中针对环境的变化尝试新的策略, 记录相关结果并更新知识库中的策略。文献[30]根据以往的货物信息生成初始知识库, 在运行自适应分析之后, 将人工的确认反馈发送给知识库, 从而实现知识库更新。

### 3.2.5 MAPE-K 框架的扩展

自适应系统的实现根据其具体自适应目标划分可分为自我配置、自我愈合、自我保护和自我优化<sup>[22]</sup>。在边缘计算中, 自我配置意味着系统能够自行配置, 不需要易于出错的手动安装和配置<sup>[22]</sup>。自我愈合包含在没有人工支持的情况下监测和修补

问题<sup>[18]</sup>。自我保护是指针对恶意攻击和恶意行为的自动防御<sup>[38]</sup>。在系统的运行阶段, 自我优化至关重要, 软件会根据环境的变化调整参数<sup>[10,16]</sup>, 从而使自身性能满足用户需求。根据不同的自适应需求, MAPE-K 控制框架中的模块可以有不同的侧重, 当前, 基于 MAPE-K 在不同需求下的扩展, 已出现具备目标驱动、质量驱动等不同方向的新型自适应框架<sup>[39]</sup>。但在边缘计算领域, 自适应系统研究大多停留在自我配置和自我优化方面。随着边缘计算应用的发展, 面向更多需求的自适应框架研究将是一个热门的研究方向。

## 4 自适应边缘计算应用实例

目前的研究已在多个应用领域探索了自适应边缘计算的应用价值, 如视频流系统<sup>[9]</sup>、视频监控与分析系统<sup>[10]</sup>、实时交通路线管理<sup>[2]</sup>等。本节将通过具体应用实例分析自适应边缘计算在实际应用中的作用, 并介绍如何针对特定的应用实例设计专门的智能算法。

### 4.1 移动视频

近年来, 视频类的网站和应用得到了更广泛的关注, 在手机端观看视频已经成为许多移动用户的日常习惯, 这给无线网络带来了海量的数据。据 Cisco 的调研报告, 2019 年视频数据占移动网络数据量的 72%, 视频类应用具有数据量大、对时延敏感等特点。而无线网络很容易受障碍物遮挡、墙面反射、多径干扰等因素的影响, 特别是在公共交通等使用场景下, 用户处于不断移动的状态, 可能需要从一个基站切换到另一个基站。为了提高视频传输的吞吐量, 降低传输时延, 文献[9,12,36]提出在基站上部署移动边缘服务器, 对用户使用过的视频数据进行边缘缓存, 从而提高其他同需求用户对视频的缓存速度。一个基站同时服务多个用户, 这些用户的无线信道质量各不相同, 因此需要不同的视频质量, 如 480P、720P、1 080P 等。低质量视频可以由高质量视频通过码率转化得到, 因此如果有一个视频的高质量版本和低质量版本, 只需要在边缘服务器缓存高质量版本的视频即可, 从而减少了对存储空间占用。但是, 从高质量视频转码到低质量视频需要消耗边缘服务器的计算能力, 并产生一定的处理时延。边缘服务器需要根据用户无线信道质量的变化不断调整边缘缓存, 文献[9]提出能够适应用户的信道质量变化的视频分辨率调整方法, 能

较好地解决此类问题。在使用边缘缓存之后,一般认为边缘缓存的命中率越高越好。文献[40]指出,由于用户与边缘服务器之间一般通过无线网络连接,将命中的缓存传输给用户也需要消耗一定的时间。如果用户与边缘服务器之间的传输带宽不足,那么降低边缘缓存反而会缩短下载视频所需的时间。也就是说,边缘缓存的最优容量应该与无线传输带宽相匹配。

## 4.2 智能家居

在基于边缘计算的智能家居系统<sup>[18,25]</sup>中,房间内安装了可以自动调节亮度的灯。为了节能环保,灯的亮度不宜调得太高,但亮度太低又会影响人们的正常生活。智能家居系统可以控制灯的亮度,使得室内的光照强度始终维持在一个适宜的范围(如 650~750 lux)内。与传统软件控制系统不同的是,智能家居系统是一个软硬件协同的系统。其中不仅有控制软件,还有灯、传感器、开关等硬件设备,这些硬件设备可能以一些软件无法直接感知的方式发生故障。如智能家居系统认为室内亮度太高,要将灯关闭。开关执行了这个命令,并向智能家居系统报告灯已经被关闭。但是,此时开关的硬件有可能发生了故障,开关中的电路实际并没有断开,这便是物联网中的软硬件一致性问题<sup>[18]</sup>。

为了确保硬件能够按照智能家居系统的指令完成相应的动作,仅依赖于硬件自身的报告是不够的,还要利用 MAPE-K 控制回路监测、修复硬件的不一致<sup>[18]</sup>。当开关报告已经将灯关闭后,智能家居系统要利用光线传感器监测室内的光线是否变暗,如果光照强度未发生变化,则说明硬件部件可能发生故障。此时智能家居系统应尝试修复该故障,如再次尝试关闭开关,或者关闭上一级开关。如果修复成功,则可以在知识库中添加记录,再遇到同样的问题就可以迅速解决。如果尝试后仍然无法修复故障,可以发出警告,请求人工干预来解决相关故障。

## 4.3 计算卸载

手机、便携式计算机等移动智能设备在日常生活中已经得到广泛使用。由于受体积、功耗、散热等因素的影响,移动智能设备在计算能力、存储空间、电池电量等方面与台式计算机等固定设备仍有一定差距。面对计算密集型或资源密集型任务,移动设备无法进行有效处理,需要将其上传到边缘服务器,由边缘服务器处理后再将结果反馈给移动设

备。移动设备通过蜂窝网或 Wi-Fi 等无线通信方式与边缘服务器通信,无线信道会随着时间的变化发生变化。这意味着将任务上传到边缘服务器所需的电量也会随之发生变化。当没有干扰时,移动设备可以迅速完成上传,因为此时分组丢失率很低。在有干扰的情况下,分组丢失率急剧上升,上传任务的速度随之下降。为了减小分组丢失率,移动设备只能通过提高发射功率,来保证边缘服务器在接收数据时有足够的信干噪比,这会导致通信消耗的电量大幅度增加。

在视频监控场景下,摄像头可以将采集的图像信息传输到边缘服务器进行处理,以减少摄像头的电池消耗<sup>[10]</sup>。如果网络连接带宽有限,摄像头要对采集的图像进行一定的简单处理再上传至边缘服务器。同时,为了精细化管理电池电量,还可以将视频监控所需的计算任务进一步分解,分成如目标检测与分割、压缩、特征提取、分类等阶段,每个阶段都制定从简单到复杂的多个算法。在计划进行计算卸载时,根据无线信道状态,卸载特定的一组算法组合至边缘服务器。其测试结果表明,优化后的计算卸载最多可以节省 4.3 倍电量<sup>[10]</sup>。

如果移动端采用可再生能源(如太阳能、风能)等非连续性能源供能<sup>[21]</sup>,既可以在电池电量不足时将计算任务上传到边缘服务器以节省电量,又可以通过自适应策略监控天气因素,从而预测未来的充电情况,合理地规划计算任务的工作量,以克服太阳能、风能来源不稳定的问题,延长边缘节点的使用寿命。另外,边缘服务器也可能采用可再生能源。为了能够向物联网设备提供稳定的服务,文献[41]提出利用网络切片技术将一部分位置相近的边缘服务器联合起来,然后根据可用能源、用户需求、QoS 要求等向各个设备动态分配计算资源。实验结果表明,即使每个边缘服务器只和最邻近的服务器节点合作,所有边缘服务器可以用于计算卸载的总计算资源也能翻倍。

如果多个设备同时通过无线网络向同一个边缘服务器进行计算卸载,这些设备的无线信号会相互干扰,从而显著降低设备与边缘服务器之间的传输带宽。文献[42]从博弈论的角度研究了多用户计算卸载问题,指出这些竞争的设备之间存在纳什均衡,并且可以通过分布式计算卸载算法达到纳什均衡。在可用的系统资源和计算任务量随时间变化而改变的情况下,以长时间的平均响应时间为优化目

标, 因为许多应用对于偶发的短时性能下降并不敏感, 只要长期的平均性能达到要求即可。文献[43]提出了一种分布式优化算法, 能够根据任务量的变化, 以长期的平均响应时间为目标自适应调整计算卸载策略, 并优化资源分配。随着物联网设备数量的不断增长, 经常出现大量设备需要进行计算卸载的情况。为了使边缘服务器能向更多的物联网设备提供服务, 文献[44]提出了一种计算卸载方法, 在满足每个设备 QoS 需求的前提下尽可能减少单个物联网设备对边缘服务器的占用。

## 5 自适应边缘计算的未來研究方向

### 5.1 分布式的自适应边缘计算

在物联网中, 终端设备可能分散部署在数千平方千米的范围内。这些地点的环境可能完全不同, 如室内与室外、城区与郊区等。边缘计算设备贴近终端部署, 因此也具备大范围分布的特性<sup>[14,45]</sup>。这意味着自适应边缘系统需要能有效处理大范围区域内各类复杂情况的适应性分析与决策, 尽可能避免由于某个终端或应用的自适应策略影响其他终端或应用的 QoS。

目前, 使用较广泛的是集中式自适应决策<sup>[23,46-49]</sup>, 即单个边缘节点仅考虑其所服务的终端或其托管的服务的自适应需求<sup>[30]</sup>, 或由计算资源较多的边缘节点(甚至是云计算中心)收集监测目标数据、生成适应补偿策略。但集中式的自适应控制难以避免因中心决策节点故障或失效造成的负面影响。此外, 决策节点所生成的适应补偿策略具有时效性。如果边缘服务器与决策节点的通信距离过长, 决策结果送达边缘节点时将有可能因环境的进一步变化而失效。半集中式<sup>[14]</sup>和分布式<sup>[34,50]</sup>的自适应控制是解决大范围分布需求的有效途径。但在分布式决策下, 如何避免终端之间、边缘节点之间的自适应补偿策略出现冲突仍然是一大挑战。目前, 关于分布式自适应边缘计算大多集中在计算卸载<sup>[40-44]</sup>, 其他应用领域中的分布式自适应边缘计算仍有待进一步研究。

### 5.2 可扩展可重构的自适应边缘计算

随着终端设备和边缘计算节点数量的不断增多, 自适应边缘系统需要具备高可扩展性, 以应对用户数量的扩充和需求的增加。另外, 物联网终端具有移动性, 用户对不同类型的服务需求也会随时间的变化发生变化, 所以一定区域内的终端节点数

量和对服务的需求也会发生变化。自适应边缘系统需要能够弹性灵活地对自身结构进行调整, 在满足用户需求和 QoS 的同时, 避免计算资源的闲置与浪费。特别是在部署新的服务或功能时, 自适应边缘系统需要能够满足其相应的 QoS 需求。这意味着与新服务、新功能相关的监测目标、自适应分析、规划更新、执行策略或知识库也需要相应地部署到系统中。在这种情况下, 需要对原来部署的自适应边缘系统进行重构, 以满足新的需求。

目前, 自适应边缘系统对终端设备的增加<sup>[25,51-53]</sup>和用户量的增长<sup>[13,37,54]</sup>讨论较多, 对边缘节点之间的协同与弹性调度<sup>[46,48]</sup>讨论较少, 尤其缺乏对自适应边缘系统重构的讨论。因此, 能够满足终端、边缘节点、用户需求等多层次弹性需求的自适应边缘系统有待进一步探索。

### 5.3 预测式适应的自适应边缘计算

反馈式适应是指系统出现异常后, 才根据异常采取一系列补救措施。预测式适应指系统根据监测目标的变化, 推理出未来可能出现的异常, 并提前采取措施预先补偿潜在的异常。

目前, 反馈式适应在自适应边缘系统中较常见<sup>[19-20,49,55-56]</sup>。这种模式能够迅速定位异常, 从而进行适应性补偿。由于自适应分析过程相对较简单, 所以对于因单一环境变化而产生的自适应更新需求有较好的效果, 如用户需求变化<sup>[22]</sup>、网络负载变化<sup>[36]</sup>、节点可用性<sup>[54]</sup>等。反馈式适应可能造成较长时间的进程等待与恢复, 因此无法在时延敏感型任务中使用。预测式适应可以提前采取措施, 因此更适合在时延敏感型任务中使用。值得注意的是, 如果使用了深度学习等人工智能技术进行自适应分析和规划, 由于计算量较大, 需要设计合理的算法以尽可能减少由计算引入的额外时延。目前, 基于预测式适应的自适应边缘系统大多集中在节能减排<sup>[17]</sup>、计算卸载<sup>[51]</sup>等业务上, 面向时延敏感型任务的预测式适应还需要进一步探讨。

### 5.4 面向 6G 的智能自适应边缘计算

6G 面向 10 年后的需求, 将对物联网、边缘计算、人工智能等技术进行深度融合。传统的通信网络主要完成了数据的传输, 6G 更多地强调智能。即在开始设计通信系统架构时期就考虑到智能应用的各种需求<sup>[57]</sup>。由于深度学习等人工智能技术对计算能力的强烈需求以及相对较高的能耗水平, 在终端设备上往往仅能够直接运行一部分小型的人

工智能程序,难以满足广泛的智能应用需求。因此,智能边缘计算将在 6G 中发挥重要作用。

许多人工智能算法假定其运行环境在一段时间内是固定不变的。由于 6G 终端设备具有移动性,所以系统的网络拓扑结构随时间的变化不断改变。而在不同地点所采集的数据也可能存在数据分布上的差异,从而使原本的人工智能算法无法正确完成新地点的数据分析任务。此外,无线信道具有随机衰落特性,这会影晌终端与边缘节点之间的通信连接,进而影响人工智能算法中同步通信和异步通信的平衡,并可能影响数据完整性。因此,边缘计算技术以及部署于边缘系统之中的人工智能算法需要能够根据周围环境的变化自适应地对自身结构和相关配置进行调整。目前,对通信系统中自适应边缘计算的讨论主要集中在计算卸载<sup>[40-44]</sup>,面向 6G 与人工智能应用的智能自适应边缘计算<sup>[57]</sup>还有待进一步研究。

## 6 结束语

物联网是一个网络开放、终端可移动、周围环境不断变化、对时延和功耗敏感的系统。因此,边缘系统需要提供灵活多样的、可重构可扩充的服务,实现智能自适应边缘计算。本文首先介绍了自适应边缘系统的特点,然后讨论了自适应边缘计算中的技术挑战与关键技术,接下来介绍了自适应边缘计算的一些典型应用,最后提出了自适应边缘计算的未采研究方向。随着 5G 应用的逐步落地和 6G 研究的兴起,自适应边缘计算在管理百万级的物联网设备和处理海量数据方面将发挥越来越重要的作用。

## 参考文献:

- [1] SASAKI K, SUZUKI N, MAKIDO S, et al. Vehicle control system coordinated between cloud and mobile edge computing[C]//Proceedings of 2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE). Piscataway: IEEE Press, 2016: 1122-1127.
- [2] BARTHÉLEMY J, VERSTAEVEL N, FOREHEAD H, et al. Edge-computing video analytics for real-time traffic monitoring in a smart city[J]. *Sensors*, 2019, 19(9): 2048.
- [3] OKAY F Y, OZDEMIR S. A fog computing based smart grid model[C]//Proceedings of 2016 International Symposium on Networks, Computers and Communications (ISNCC). Piscataway: IEEE Press, 2016: 1-6.
- [4] VARGHESE B, WANG N, BARBHUIYA S, et al. Challenges and opportunities in edge computing[C]//Proceedings of 2016 IEEE International Conference on Smart Cloud (SmartCloud). Piscataway: IEEE Press, 2016: 20-26.
- [5] BIBRI S E, KROGSTIE J. Smart sustainable cities of the future: an extensive interdisciplinary literature review[J]. *Sustainable Cities and Society*, 2017, 31: 183-212.
- [6] YANG Y. Multi-tier computing networks for intelligent IoT[J]. *Nature Electronics*, 2019, 2(1): 4-5.
- [7] YANG Y, LUO X, CHU X, et al. Fog-enabled intelligent IoT systems[M]. Berlin: Springer, 2019.
- [8] SHI W S, DUSTDAR S. The promise of edge computing[J]. *Computer*, 2016, 49(5): 78-81.
- [9] TRAN T X, POMPILI D. Adaptive bitrate video caching and processing in mobile-edge computing networks[J]. *IEEE Transactions on Mobile Computing*, 2019, 18(9): 1965-1978.
- [10] CAO N Y, NASIR S B, SEN S, et al. Self-optimizing IoT wireless video sensor node with in situ data analytics and context-driven energy-aware real-time adaptation[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2017, 64(9): 2470-2480.
- [11] WANG K, SHAO Y, XIE L, et al. Adaptive and fault-tolerant data processing in healthcare IoT based on fog computing[J]. *IEEE Transactions on Network Science and Engineering*, 2020, 7(1): 263-273.
- [12] LIANG C C, HE Y, YU F R, et al. Enhancing video rate adaptation with mobile edge computing and caching in software-defined mobile networks[J]. *IEEE Transactions on Wireless Communications*, 2018, 17(10): 7013-7026.
- [13] WANG L, JIAO L, LI J, et al. Online resource allocation for arbitrary user mobility in distributed edge clouds[C]//Proceedings of 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS). Piscataway: IEEE Press, 2017: 1281-1290.
- [14] CASADEI R, PIANINI D, VIROLI M, et al. Self-organising coordination regions: a pattern for edge computing[C]//Proceedings of International Conference on Coordination Languages and Models. [S.l.:s.n.], 2019: 182-199.
- [15] ZHANG T H, JIN J, ZHENG X, et al. Rate-adaptive fog service platform for heterogeneous IoT applications[J]. *IEEE Internet of Things Journal*, 2020, 7(1): 176-188.
- [16] WEN Z Y, YANG R Y, GARRAGHAN P, et al. Fog orchestration for Internet of things services[J]. *IEEE Internet Computing*, 2017, 21(2): 16-24.
- [17] XU J, CHEN L X, REN S L. Online learning for offloading and autoscaling in energy harvesting mobile edge computing[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2017, 3(3): 361-373.
- [18] SEIGER R, HUBER S, HEISIG P, et al. Toward a framework for self-adaptive workflows in cyber-physical systems[J]. *Software & Systems Modeling*, 2019, 18(2): 1117-1134.
- [19] LIN B, ZHU F N, ZHANG J S, et al. A time-driven data placement strategy for a scientific workflow combining edge computing and cloud computing[J]. *IEEE Transactions on Industrial Informatics*, 2019,

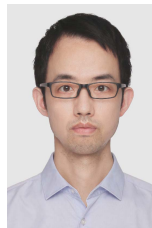
- 15(7): 4254-4265.
- [20] SODHRO A H, PIRBHULAL S, ALBUQUERQUE V H C D. Artificial intelligence-driven mechanism for edge computing-based industrial applications[J]. *IEEE Transactions on Industrial Informatics*, 2019, 15(7): 4235-4243.
- [21] CAO K, ZHOU J L, XU G, et al. Exploring renewable-adaptive computation offloading for hierarchical QoS optimization in fog computing[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2020, 39(10): 2095-2108.
- [22] ZHAO T Q, ZHANG W, ZHAO H Y, et al. A reinforcement learning-based framework for the generation and evolution of adaptation rules[C]//*Proceedings of 2017 IEEE International Conference on Autonomous Computing (ICAC)*. Piscataway: IEEE Press, 2017: 103-112.
- [23] TANG Z Q, ZHOU X J, ZHANG F M, et al. Migration modeling and learning algorithms for containers in fog computing[J]. *IEEE Transactions on Services Computing*, 2019, 12(5): 712-725.
- [24] SEIGER R, HUBER S, HEISIG P, et al. Enabling self-adaptive workflows for cyber-physical systems[C]//*Proceedings of Enterprise, Business-Process and Information Systems Modeling*. [S.l.:s.n.], 2016.
- [25] FERRÁNDEZ-PASTOR F J, MORA H, JIMENO-MORENILLA A, et al. Deployment of IoT edge and fog computing technologies to develop smart building services[J]. *Sustainability*, 2018, 10(11): 1-23.
- [26] GATOULLAT A, BADR Y, MASSOT B. QoS-driven self-adaptation for critical IoT-based systems[C]//*Proceedings of International Conference on Service-Oriented Computing*. [S.l.:s.n.], 2017: 93-105.
- [27] KIT M, GEROSTATHOPOULOS I, BURES T, et al. An architecture framework for experimentations with self-adaptive cyber-physical systems[C]//*Proceedings of 10th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*. [S.l.:s.n.], 2015: 93-96.
- [28] WEYNS D, RAMACHANDRAN G, SINGH R. Self-managing Internet of things[C]//*Proceedings of International Conference on Current Trends in Theory and Practice of Informatics*. [S.l.:s.n.], 2018: 1-18.
- [29] TAHERIZADEH S, JONES A, TAYLOR I, et al. Monitoring self-adaptive applications within edge computing frameworks: a state-of-the-art review[J]. *Journal of Systems and Software*, 2018, 136: 19-38.
- [30] ZHANG L, ALHARBE N, ATKINS A. An IoT application for inventory management with a self-adaptive decision model[C]//*Proceedings of IEEE International Conference on Internet of Things*. Piscataway: IEEE Press, 2016: 317-322.
- [31] RAVINDRA P, KHOCHARE A, REDDY S, et al. ECHO: an adaptive orchestration platform for hybrid dataflows across cloud and edge[C]//*Proceedings of international Conference on Service-Oriented Computing*. [S.l.:s.n.], 2017: 395-410.
- [32] ZOLOTUKHIN M, HÄMÄLÄINEN T, KOKKONEN T, et al. Increasing web service availability by detecting application-layer DDoS attacks in encrypted traffic[C]//*Proceedings of 2016 23rd International Conference on Telecommunications (ICT)*. Piscataway: IEEE Press, 2016: 1-6.
- [33] SONG W, YIN H, LIU C, et al. DeepMem: learning graph neural network models for fast and robust memory forensic analysis[C]//*Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. New York: ACM Press, 2018: 606-618.
- [34] CAPORUSCIO M, ANGELO M D, GRASSI V, et al. Reinforcement learning techniques for decentralized self-adaptive service assembly[C]//*Proceedings of 5th European Conference on Service-Oriented and Cloud Computing (ESOCC)*. [S.l.:s.n.], 2016: 53-68.
- [35] MU T Y, AL-FUQAHA A, SHUAIB K, et al. SDN flow entry management using reinforcement learning[J]. *ACM Transactions on Autonomous and Adaptive Systems*, 2018, 13(2): 1-23.
- [36] RAHMAN W U, HONG C S, HUH E N. Edge computing assisted joint quality adaptation for mobile video streaming[J]. *IEEE Access*, 2019(7): 129082-129094.
- [37] SILVA R A C D, FONSECAN L S D. Resource allocation mechanism for a fog-cloud infrastructure[C]//*Proceedings of 2018 IEEE International Conference on Communications (ICC)*. Piscataway: IEEE Press, 2018: 1-6.
- [38] XIAO Y H, JIA Y Z, LIU C C, et al. Edge computing security: state of the art and challenges[J]. *Proceedings of the IEEE*, 2019, 107(8): 1608-1631.
- [39] WEYNS D. *Handbook of software engineering*[M]. Berlin: Springer, 2019.
- [40] LI Q, ZHANG Y M, LI Y Y, et al. Capacity-aware edge caching in fog computing networks[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(8): 9244-9248.
- [41] XIAO Y, KRUNZ M. Dynamic network slicing for scalable fog computing systems with energy harvesting[J]. *IEEE Journal on Selected Areas in Communications*, 2018, 36(12): 2640-2654.
- [42] CHEN X, JIAO L, LI W Z, et al. Efficient multi-user computation offloading for mobile-edge cloud computing[J]. *IEEE/ACM Transactions on Networking*, 2016, 24(5): 2795-2808.
- [43] LIN R P, ZHOU Z J, LUO S, et al. Distributed optimization for computation offloading in edge computing[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(12): 8179-8194.
- [44] CHEN X, SHI Q, YANG L, et al. Thrifty edge: resource-efficient edge computing for intelligent IoT applications[J]. *IEEE Network*, 2018, 32(1): 61-65.
- [45] D'ANGELO M. Decentralized self-adaptive computing at the edge[C]//*Proceedings of the 13th International Conference on Software Engineering for Adaptive and Self-Managing Systems*. [S.l.:s.n.], 2018: 144-148.
- [46] LIU Q Y, WEI Y K, LENG S P, et al. Task scheduling in fog enabled Internet of things for smart cities[C]//*Proceedings of IEEE 17th International Conference on Communication Technology (ICCT)*. Piscataway: IEEE Press, 2017: 975-980.
- [47] YIGITOGU E, MOHAMED M, LIU L, et al. Foggy: a framework for continuous automated IoT application deployment in fog computing[C]//*Proceedings of 2017 IEEE International Conference on AI &*

- Mobile Services (AIMS). Piscataway: IEEE Press, 2017: 38-45.
- [48] MINH Q T, NGUYEN D T, VAN L A, et al. Toward service placement on fog computing landscape[C]//Proceedings of 2017 4th NAFOSTED Conference on Information and Computer Science. Piscataway: IEEE Press, 2017: 291-296.
- [49] JUTILA M. An adaptive edge router enabling Internet of things[J]. IEEE Internet of Things Journal, 2016, 3(6): 1061-1069.
- [50] CHEN L X, ZHOU P, GAO L, et al. Adaptive fog configuration for the industrial Internet of things[J]. IEEE Transactions on Industrial Informatics, 2018, 14(10): 4656-4664.
- [51] ORSINI G, BADE D, LAMERSDORF W. Cloud aware: a context-adaptive middleware for mobile edge and cloud computing applications[C]//Proceedings of 2016 IEEE 1st International Workshops on Foundations & Applications of Self Systems. Piscataway: IEEE Press, 2016: 216-221.
- [52] CHEN N X, YANG Y, ZHANG T, et al. Fog as a service technology[J]. IEEE Communications Magazine, 2018, 56(11): 95-101.
- [53] NGUYEN T D, KIM Y, KIM D H, et al. A proposal of autonomic edge cloud platform with CCN-based service routing protocol[C]//Proceedings of 2018 IEEE 11th International Conference on Cloud Computing (CLOUD). Piscataway: IEEE Press, 2018: 802-809.
- [54] DANESHFAR N, PAPPAS N, POLISHCHUK V, et al. Service allocation in a mobile fog infrastructure under availability and QoS constraints[C]//Proceedings of 2018 IEEE Global Communications Conference (GLOBECOM). Piscataway: IEEE Press, 2018: 1-6.
- [55] HSIEH Y C, HONG H J, TSAI P H, et al. Managed edge computing on Internet-of-things devices for smart city applications[C]//Proceedings of NOMS 2018 IEEE/IFIP Network Operations and Management Symposium. Piscataway: IEEE Press, 2018: 1-2.
- [56] GEROSTATHOPOULOS I, BURES T, HNETYNKA P, et al.

Self-adaptation in software-intensive cyber-physical systems: from system goals to architecture configurations[J]. Journal of Systems and Software, 2016: 1-20.

- [57] XIAO Y, SHI G M, LI Y Y, et al. Toward self-learning edge intelligence in 6G[J]. IEEE Communications Magazine, 2020, 58(12): 34-40.

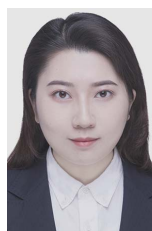
#### [作者简介]



王旭（1986- ），男，博士，中国科学院上海微系统与信息技术研究所副研究员，主要研究方向为无线通信、深度学习等。



陈南希（1989- ），女，博士，中国科学院上海微系统与信息技术研究所副研究员、云脑课题组组长，主要研究方向为泛在智能、物联网、深度学习等。



张柔佳（1995- ），女，华为技术有限公司上海研究所测试工程师，主要研究方向为通信和信息技术、服务计算等。