

边缘计算的架构、挑战与应用

李林哲^{1,2}, 周佩雷³, 程鹏¹, 史治国^{1,2}

1. 浙江大学工业控制技术国家重点实验室, 浙江 杭州 310058;
2. 阿里巴巴-浙江大学前沿技术联合研究中心, 浙江 杭州 310058;
3. 浙江省公众信息产业有限公司, 浙江 杭州 310058

摘要

边缘计算是一种在网络边缘执行计算任务的新型计算模型, 相比于云计算模型, 能够更加迅速、可靠和节能地响应用户需求。从云计算模型的不足出发, 首先介绍了边缘计算的概念和通用构架, 随后详细阐述了两个边缘计算参考构架, 总结了边缘计算面临的挑战, 并介绍了针对这些挑战的研究进展。随着边缘计算相关理论和技术的发展, 边缘计算将成为推动物联网服务升级的关键技术, 以预测性维护和安防监控为例, 对边缘计算的应用进行了介绍。

关键词

边缘计算; 云计算; 框架

中图分类号: TP393.0

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2019010

Architecture, challenges and applications of edge computing

LI Linzhe^{1,2}, ZHOU Peilei³, CHENG Peng¹, SHI Zhiguo^{1,2}

1. State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310058, China
2. Alibaba-Zhejiang University Joint Institute of Frontier Technologies, Hangzhou 310058, China
3. Zhejiang Public Information Industry Co., Ltd., Hangzhou 310058, China

Abstract

Edge computing is a new type of computing models that performs computing tasks at the edge of the network. Compared with cloud computing, it can respond to user's needs more quickly and reliably. Starting from the shortcomings of cloud computing, the concept and general architecture of edge computing were illustrated, and then two reference frameworks proposed by industry alliances were elaborated. Four challenges of edge computing and their latest research progress were summarized. With the development of theory and technology related to edge computing, it will become a key technology to promote the upgrade of Internet of things (IoT) services. For this reason, two applications of edge computing in manufacturing and security monitoring were introduced.

Key words

edge computing, cloud computing, architecture

1 引言

随着物联网、5G通信等技术的快速发展,万物互联的智能时代正在加速到来。中国经济信息社在2018年9月发布的《2017—2018年中国物联网发展年度报告》中提到,2017年物联网设备的数量增长强劲,达到84亿台,首次超过全球人口数量^[1]。伴随而来的是数据量的高速增长,数据统计公司Statista预测,到2020年,将大约有310亿个物联网设备连接,同时据思科云指数估计,2021年云计算的数据总流量将达到18.9 ZB,相比2016年的6.0 ZB增长2.15倍,这种情形对当前广泛使用的云计算模型提出了巨大的挑战。边缘计算是一种在靠近物或数据源头的网络边缘提供智能服务的新型计算模型,它能够节省网络流量、提高响应速度和保护用户隐私,在物联网应用中显示出了优于云计算的性能,受到工业界、学术界的高度关注和一致认可。

边缘计算于2015年进入快速发展期,并迅速在安防监控、智慧城市、智慧家居等行业实现了应用试点。边缘计算的快速发展得益于学术界的高度关注。近年来,计算机及通信等领域的知名国际学术会议中,关于边缘计算的文章数目和比例逐年上升,其中计算机领域顶级会议ICDCS在2015年、2016年和2017年收录的边缘计算方向的文章比例分别为6.7%、9.3%和12.7%,通信领域顶级会议INFOCOM每年约收录270篇文章,其中2016年和2017年收录边缘计算方向的文章分别为8篇和5篇,2018年收录边缘计算方向的文章高达29篇,可见学术界对边缘计算的热情。此外,2016年ACM与IEEE共同创办了首个以边缘计算为主题的学术会议——IEEE/ACM Symposium on Edge Computing (SEC),以推进

边缘计算关键技术的发展。

2 边缘计算概述

2.1 边缘计算的产生背景

云计算服务是一种集中式服务,所有数据都通过网络传输到云计算中心进行处理。资源的高度集中与整合使得云计算具有很高的通用性,然而,面对物联网设备和数据的爆发式增长,基于云计算模型的聚合性服务逐渐显露出了其在实时性、网络制约、资源开销和隐私保护上的不足。

- 云计算难以保证实时性要求。云计算模型将全部数据上传至云计算中心进行处理,其处理速度受到网络带宽、中心计算能力、总计算任务量等多因素的影响^[2],且请求至响应的链路较长,各个环节的时延累计可能造成无法接受的处理时延。

- 云计算对网络环境过度依赖^[2-3]。尽管我国4G网络覆盖率已经超过95%,但仍存在海岛、地下室等网络盲区,同时也存在山谷、隧道等无法保证网络质量的区域。由于云计算依赖网络实现数据的传输,在这些场景中其难以提供可靠的服务。

- 云计算的资源开销较大。随着数据量的攀升,数据传输带来的网络流量开销也在逐渐升高,同时云计算中心的计算、存储功能带来了极高的能耗^[4],而这些开销并不是完全必要的。以野生动物保护区中的监控图像处理为例,大量的监控图像中并未包含任何动物,然而云计算模型仍然会对每一张图片进行传输、处理和存储。

- 云计算难以保证用户隐私^[5-6]。云计算处理的数据可能是包含用户隐私的,例如家庭内的监控摄像头、工厂内的生产数据

等, 尽管存在用户隐私协议等约束, 但服务提供商对数据的实际使用情况是不透明的, 例如2017年8月, 华为技术有限公司被报道根据用户聊天内容自动加载地址、天气等信息, 侵犯了用户隐私。

为了弥补集中式云计算的不足, 边缘计算的概念应运而生, 它是指在靠近物或数据源头的网络边缘侧, 融合网络、计算、存储、应用核心能力的分布式开放平台, 就近提供边缘智能服务^[7-8]。由于传输链路的缩短, 边缘计算能够在数据产生侧快捷、高效地响应业务需求, 数据的本地处理也可以提升用户隐私保护程度。另外, 边缘计算减小了服务对网络的依赖, 在离线状态下也能够提供基础业务服务。云边协同的联合式服务能够充分利用云计算和边缘计算的联合优势, 针对不同特征的业务需求进行灵活的部署和响应, 图1是云计算模型与云边协同计算模型的对比。据思科云指数估计, 2019年人、机、物产生的数据将达到500 ZB, 网络带宽将成为云计算的瓶颈, 融入边缘计算的云边协同的联合式服务将成为更有效的服务构架。国际数据公司IDC也预测, 到2022年, 超过40%的云部

署结构将容纳边缘计算能力。

2.2 边缘计算的概念

边缘计算最早可以追溯至内容分发网络 (content delivery network, CDN) 中功能缓存的概念^[9-10], 2015年边缘计算进入快速发展期后, 以边缘计算为主题的协会与联盟相继成立, 各类定义、标准与规范逐渐形成。旨在推动云操作系统的发展、传播和使用的OpenStack基金会以及由华为技术有限公司、中国科学院沈阳自动化研究所等联合成立的边缘计算产业联盟 (Edge Computing Consortium, ECC) 等组织对边缘计算进行了定义, 尽管这些定义的描述不尽相同, 但在边缘计算的核心概念上达成了共识: 边缘计算是指在网络边缘执行计算的一种新型计算模型, 这里的边缘是指从数据源到云计算中心之间的任意资源, 其操作对象包括来自于云服务的下行数据和万物互联服务的上行数据^[7,10-12]。

与云计算模型不同的是, 边缘计算中终端设备与云计算中心的请求与响应是双向的, 如图2所示, 终端设备不仅向云计算

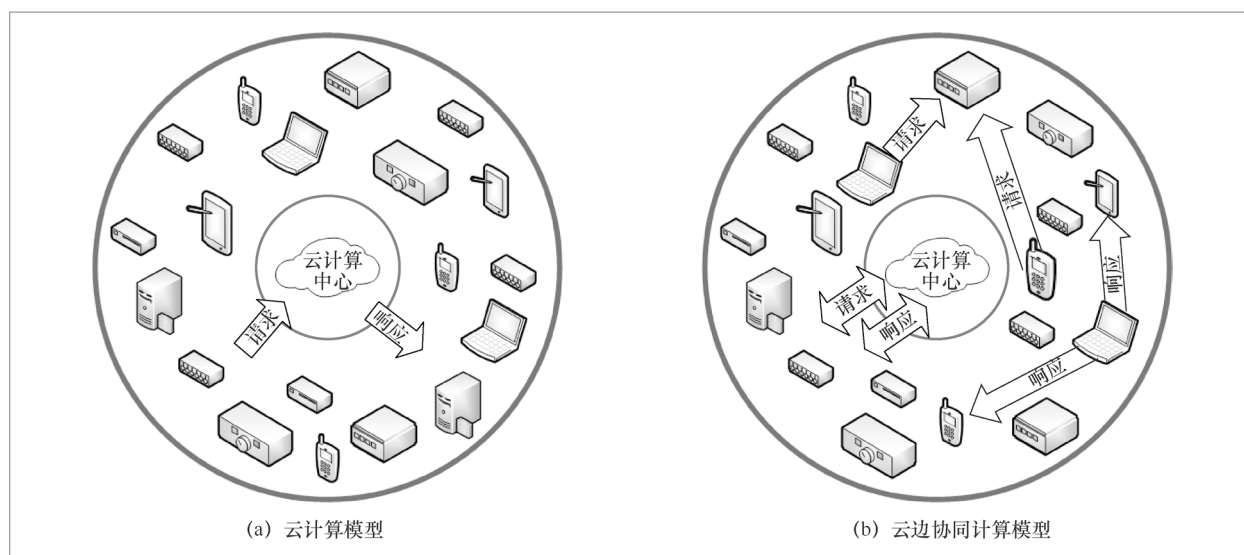


图1 云计算模型与云边协同计算模型

中心发出请求,同时也能够完成云计算中心下发的计算任务^[7]。云计算中心不再是数据生产者 and 消费者的唯一中继,由于终端设备兼顾了数据生产者和消费者的角色,部分服务可以直接在边缘完成响应,并返回终端设备,云计算中心和边缘分别形成了两个服务响应流^[13]。

边缘计算的核心是在靠近数据源或物的一侧提供计算、存储和应用服务,这似乎与雾计算将计算和分析能力扩展至网络“边缘”的定义非常相近。雾计算也是云计算模型的延伸,但雾计算的核心是将云计算中心的能力下沉至接近物的一侧,具有更平坦的架构,属于通用性较高的基础设施^[3,14-15],仍然依赖于网络,多使用本地服务器或路由器实现。从实现架构来讲,雾计算也属于边缘计算的一种,除了在雾计算中部署通用性较高的基础设施以外,边缘计算还可以将终端设备侧的能力进行升级,依赖于不构成网络的终端节点。

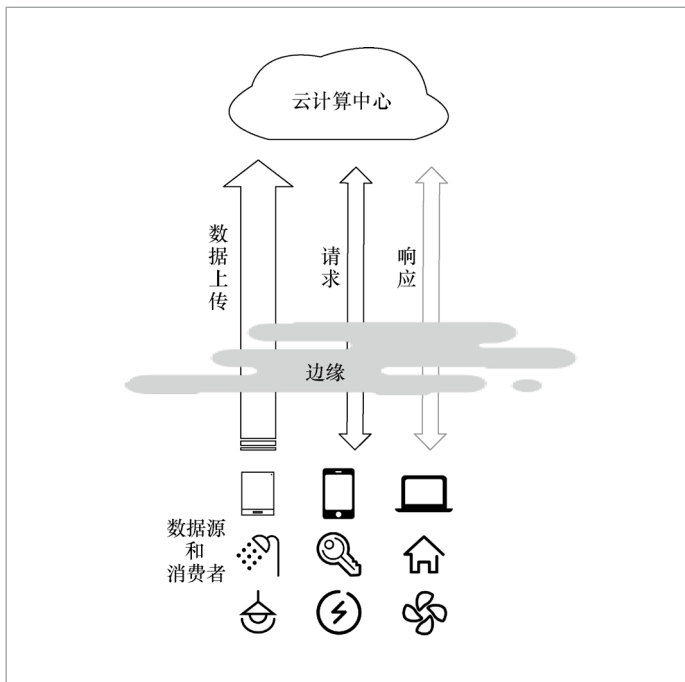


图2 边缘计算的双向计算流模型

3 边缘计算架构

3.1 边缘计算的通用架构

云边协同的联合式网络结构一般可以分为终端层、边缘计算层和云计算层^[16-18],如图3所示,各层可以进行层间及跨层通信,各层的组成决定了层级的计算和存储能力,从而决定了各个层级的功能。

(1) 终端层

终端层由各种物联网设备(如传感器、RFID标签、摄像头、智能手机等)组成,主要完成收集原始数据并上报的功能^[19]。在终端层中,只考虑各种物联网设备的感知能力,而不考虑它们的计算能力。终端层的数十亿台物联网设备源源不断地收集各类数据,以事件源的形式作为应用服务的输入。

(2) 边缘计算层

边缘计算层是由网络边缘节点构成的,广泛分布在终端设备与计算中心之间,它可以是智能终端设备本身,例如智能手环、智能摄像头等,也可以被部署在网络连接中,例如网关、路由器等。显然,边缘节点的计算和存储资源是差别很大的,并且边缘节点的资源是动态变化的,例如智能手环的可使用资源是随着人的使用情况动态变化的。因此,如何在动态的网络拓扑中对计算任务进行分配和调度是值得研究的问题。边缘计算层通过合理部署和调配网络边缘侧的计算和存储能力,实现基础服务响应。

(3) 云计算层

在云边计算的联合式服务中,云计算仍然是最强大的数据处理中心,边缘计算层的上报数据将在云计算中心进行永久性存储,边缘计算层无法处理的分析任务和综合全局信息的处理任务也仍然需要在云计算中心完成。除此之外,云计算中心还可以

根据网络资源分布动态调整边缘计算层的部署策略和算法。

边缘计算的参考框架是众多组织关注的焦点,它将抽象的边缘计算通用框架进行具象,提供了边缘计算框架的实现范式。第3.2节与第3.3节将详细介绍Linux基金会与边缘计算产业联盟提出的两种参考架构。

3.2 EdgeX Foundry

2017年4月, Linux基金会创立了EdgeX Foundry社区,旨在创建一个互操作性强、即插即用和模块化的物联网边缘计算生态系统,提出了专注于物联网边缘的标准化的微服务框架——EdgeX Foundry。该框架最早孵化于戴尔公司的物联网中间件框架中,现已将代码开源,可供开发者快速地根据自己的服务需求进行重构和部署,其结构如图4所示。架构的设计遵循了以下原则:架构应是与平台无关的,能够与多类别操作系统进行对接;架构需具有高

灵活性,其中的任意部分应该都可以进行升级、替换或扩充;架构需具有存储和转发的功能,支持离线运行,并保证计算能力能够靠近边缘。

EdgeX Foundry是微服务的集合,这些微服务分为4个层次:设备服务层、核心服务层、支持服务层、应用及导出服务层。以

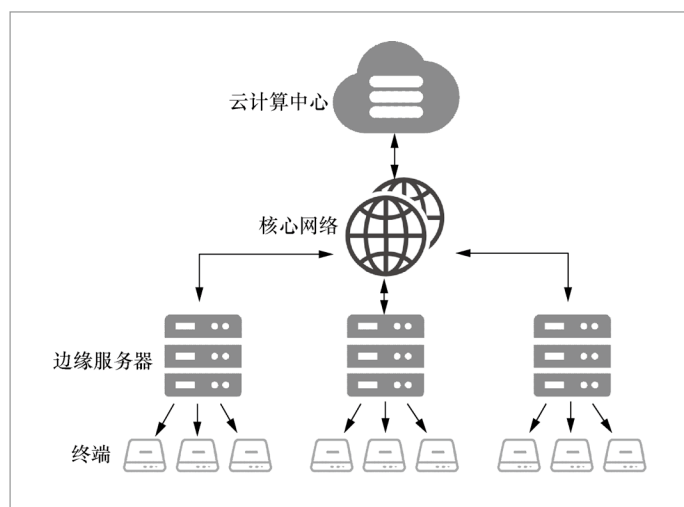


图3 云边协同的联合式网络结构

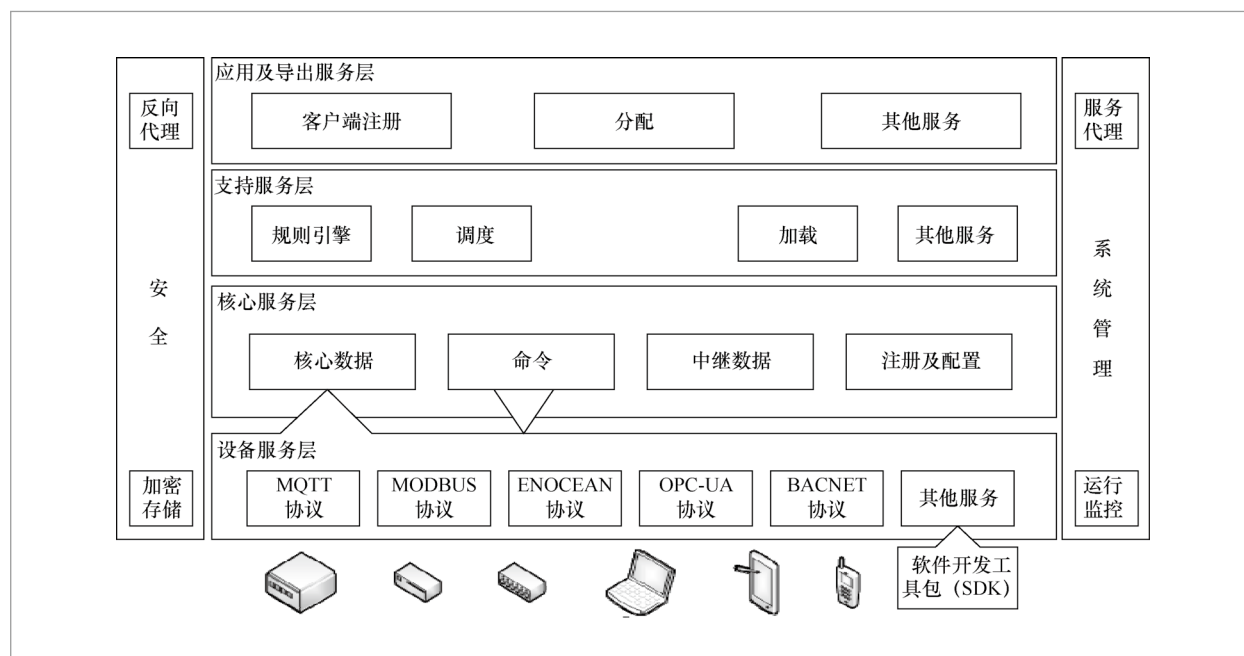


图4 EdgeX Foundry 架构

核心服务层为界, 整个服务架构可以分为“北侧”和“南侧”。“北侧”包含云计算中心和与云计算中心通信的网络, 包含支持服务层与应用及导出服务层。其中, 支持服务层包含各种微服务, 可提供边缘分析能力, 并可以为框架本身提供日志记录、调度和规则引擎等服务; 应用及导出服务层则保证了EdgeX Foundry的独立运行, 在其不与云计算中心连接时, 仍可以对边缘设备的数据进行收集, 同时, 导出服务层也负责提供网关客户端注册等功能, 并对与云计算中心传递的数据格式和规则进行实现。

“南侧”包含物理领域中的全部物联网对象以及与它们直接通信的网络边缘。其中, 设备服务层提供软件开发工具包 (software development kit, SDK), 以实现与设备的连接和通信, 设备可以是网关或其他具有数据汇集能力的设备, 同时设备服务层也可以接收来自其他微服务的命令, 进而传递到设备。作为中心的核心服务层是实现边缘能力的关键, 其中“核心数据服务”提供了持久性存储服务和对设备数据的管理服务。“命令服务”负责将云计算中心的需求驱动至设备端, 并提供命令的缓存和管理服务。“中继数据服务”为中继数据 (又称元数据, 是对数据的属性描述) 提供管理和存储服务, 信息用于为设备和服务提供配对。“注册及配置服务”为其他微服务提供配置信息。

EdgeX Foundry还包含了两个贯穿整个框架且为各层提供服务的基础服务层——安全和系统管理。安全服务中的元件为EdgeX Foundry中的各类设备提供保护, 支持认证授权计费 (authentication、authorization、accounting, AAA) 访问控制、高级加密标准 (advanced encryption standard, AES) 数据加密、证书认证、超文本传输安全协议 (HTTPS) 等保护方法。系统管理工具提供了监控EdgeX Foundry运行情况的能

力, 在未来可能会提供服务配置、为管理平台提供信息等功能。

EdgeX Foundry的主要任务是简化和标准化工业物联网边缘计算, 它提供了一个可操作的开源平台, 有效降低了边缘计算的准入门槛, 小型应用提供商也能够快速地构建和部署边缘计算服务。工业物联网推广组织工业互联网联盟 (Industrial Internet Consortium, IIC) 已经与Linux基金会达成合作协议, 共同推进工业物联网边缘服务的普及。

3.3 边缘计算参考框架3.0

除了Linux基金会外, 边缘计算产业联盟也于2018年12月发布了《边缘计算白皮书3.0》, 并提出了边缘计算参考架构3.0 (以下简称边缘框架3.0)。边缘计算产业联盟认为, 边缘计算服务框架需要达成的目标有: 对物理世界具有系统和实时的认知能力, 在数字世界进行仿真和推理, 实现物理世界与数字世界的协作; 基于模型化的方法在各产业中建立可复用的知识模型体系, 实现跨行业的生态协作; 系统与系统之间、服务与服务之间等基于模型化接口进行交互, 实现软件接口与开发语言、工具的解耦; 框架应该可以支撑部署、数据处理和安全等服务的全生命周期。

边缘框架3.0也具有贯通整个框架的基础服务层, 其架构如图5所示, 其中安全服务与管理服务的功能与EdgeX Foundry类似, 数据全生命周期服务提供了对数据从产生、处理到消费的综合管理。从纵向结构来看, 最上侧的是模型驱动的统一服务框架, 它能够实现服务的快速开发和部署。下侧按照边缘计算通用架构分为现场设备、边缘和云3层, 边缘层又划分为边缘节点和边缘管理器两个层次。边缘节点的形式、种类是多种多样的, 为了解决异

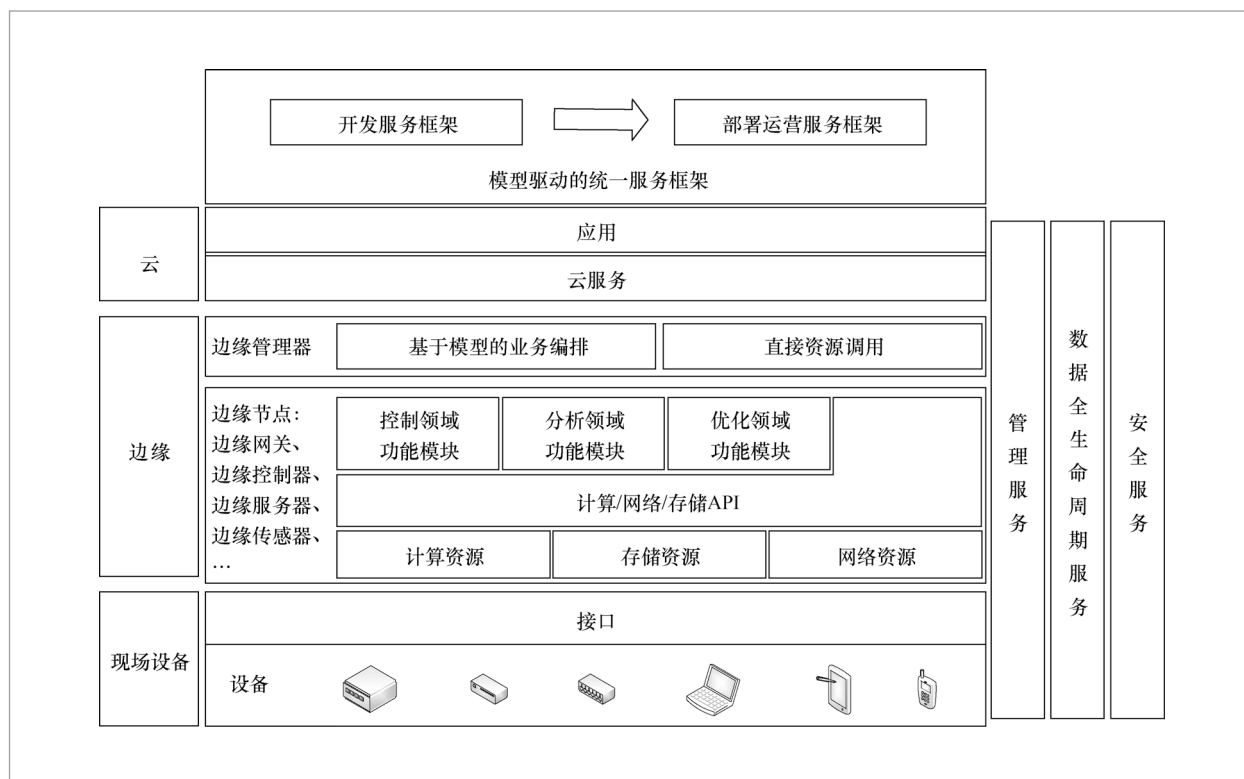


图5 边缘框架 3.0 架构

构计算与边缘节点的强耦合关系，降低物理世界带来的结构复杂性，边缘节点层中的设备资源被抽象为计算、网络和存储3种资源，使用应用程序编程接口(application programming interface, API)实现通用的能力调用，控制、分析与优化领域功能模块实现了上下层信息的传输和本地资源的规划。边缘管理器则使用模型化的描述语言帮助不同角色使用统一的语言定义业务，实现智能服务与下层结构交互的标准化。根据功能，边缘框架3.0提供了4种开发框架：实时计算系统、轻量计算系统、智能网关系系统和智能分布式系统，覆盖了从终端节点到云计算中心链路的服务开发。

4 边缘计算面临的挑战

从边缘计算的定义及架构上可以看出，“边缘”是一个相对云计算中心的概念，这意味着边缘计算的覆盖范围很广，需要多种资源的协同工作，并且需要与云计算架构实现良好的对接，因而面临着众多挑战。2016年，美国韦恩州立大学的施巍松教授团队^[7]提出，边缘计算面临着可编程性、命名、数据抽象、服务管理、隐私及安全性和性能指标优化6种挑战，其中，在可编程性、命名、服务管理和隐私及安全问题上，学术界及工业界已经取得了阶段性的成果，本节将对这4种挑战和研究进展进行详细介绍。

4.1 可编程性

边缘节点组成的计算平台类似于异构平台，边缘节点的计算与存储能力、运行时间、操作系统和支持语言等资源都可能是

不同的,这意味着开发者需要根据不同种类边缘设备的资源进行程序开发。边缘计算应该是一个动态、灵活的计算平台,能够根据当前的资源分布动态配置计算任务。显然,与硬件资源高度耦合的传统的开发模式并不适用于边缘计算的场景。为了解决边缘计算的可编程性,需要开发具有高层综合能力的编译工具,使开发者能够使用统一的语言编写程序,由编译平台根据计算任务分配情况自动编译适用于硬件的程序。TVM是一种针对机器学习的跨硬件平台编译器^[20-21],边缘计算中的机器学习算法主要运行在移动图形处理器(graphics processing unit, GPU)和现场可编程门阵列(field-programmable gate array, FPGA)两类嵌入式处理器中,然而它们使用的编程语言和操作系统架构通常是不同的,且程序部署时需要大量的手动工作。TVM能够实现面向GPU与FPGA的机器学习算法动态移植,已经在几家互联网主流企业内部开源和使用。

4.2 命名

域名系统(domain name system, DNS)^[22]等命名机制已经在云计算模型中得到了很好的应用,能够满足当前的大多数网络。但是现有命名机制并不适用于边缘计算,以智能家居中玄关灯随门打开而自动开启为例,边缘计算程序根据玄关灯的唯一ID控制它的开关,如果这个设备被更换,玄关灯的ID将会改变,此时只有更改程序才能实现原有的功能。可见,原有命名机制灵活性较低,因而不能适应边缘计算中动态变化的网络拓扑,同时一些边缘节点的资源不足以支撑原有命名机制的开销。命名数据网络(named data networking, NDN)^[23]使用内容名字代替地址,例如在玄关灯的场景中,NDN不再需要知

道玄关灯的地址,只需要将内容名字统一为“控制玄关灯”,网络就可以自动找到控制玄关灯的节点,进行数据传输。尽管NDN能够适应动态的边缘网络,但它与上层使用地址进行内容分发的网络并不匹配,而且存在安全隐患。清华大学与亚利桑那大学的学者在参考文献[24]中提出了一种使用双栈交换机搭建NDN、局域网混合网络的方法,并对双栈交换机的布局进行了优化,在保持基于IP地址的内容分发的同时,提高了网络的弹性,能够适应边缘计算动态的网络拓扑。

4.3 服务管理

服务管理是边缘计算中的关键技术,2018年IEEE/ACM SEC收录的文章中有20%与这一话题有关。边缘计算中的服务管理应该满足4种特性:差异化,即各类服务应根据其属性分为不同的优先级;可扩展性,即边缘计算中的节点是动态变化的,服务管理应该能够具有灵活的扩展性;隔离性,即应避免服务之间的耦合,当某个应用程序崩溃时,系统应仍能够保持运行;可靠性,即数据传输、设备自身的可靠性对服务非常重要。除此之外,边缘计算场景中的服务管理还面临着云计算与边缘计算目标不一致的独特问题。参考文献[25]提出了一种基于游戏理论的任务分配框架,利用动态反馈激励机制适应边缘计算的动态网络和解决边缘计算与云计算目标冲突的问题。边缘服务器的布局也对边缘计算的服务管理有非常重要的影响,参考文献[26]提出了一种基于资源需求预测的跨区域资源优化模型,首先对计算任务进行拆分和预测,然后根据预测结果使用启发式优化算法求解服务器的布局策略。

4.4 隐私及安全

相比于云计算模型,边缘计算模型可以在网络边缘完成一部分数据处理工作,这避免了用户隐私信息在云计算中心或过长的传输链路上被滥用和被窃取的风险,但是边缘计算中多类别、多数量设备的接入也带来了新的隐私及安全问题^[19,27-28]。首先物联网汇集的数据中很有可能包含用户的隐私,例如在智能家居场景中,宠物监控摄像头包含房屋结构和室内陈设信息。其次,边缘网络的安全性往往是没有保证的,仍然以智能家居为例,有数据显示,有49%的家庭无线网络是不安全的,攻击者可以轻易地破解密码,并窃取信息。即便是部署了安全策略的网络,由于一部分终端设备资源有限而无法部署安全保护方案,仍然会造成网络的不安全。最后,网络边缘的高度动态性也会增加网络的脆弱性。随着用户对隐私与安全的要求越来越严苛,学术界对隐私与安全问题的关注度也越来越高,2017年IEEE/ACM SEC收录安全与隐私主题文章6篇,2018年则多达13篇,可见边缘计算中隐私与安全方向的研究正处于蓬勃发展期。弗吉尼亚大学学者在参考文献[29]中使用了一种基于二分拓扑威胁模型和交互式对抗深度网络的分类算法实现隐私保护,提出了“隐私分区”的概念,将资源分为可信分区和不可信分区,并进行隔离。更有针对性地,加利福尼亚大学Brian Demsky教授的程序设计语言研究小组提出了一种应用于智能家居场景的隐私保护方法——Vigilia^[30],通过限制设备的网络访问增强系统的防御力,Vigilia能在保持很小的资源开销的同时,对网络实现有效的通信限制。

5 应用案例

边缘计算的特点是能够实时、高效、节能地响应用户需求,相比于云计算模型,能够更好地保护用户隐私。边缘计算的应用场景与它的优势息息相关,本节将以预测性维护、智能摄像头为例,对边缘计算在制造业和安防监控领域的应用进行介绍。

5.1 预测性维护

可靠、高效的维护方案对于制造商来说至关重要,因为设备故障带来的被动停机可能对生产效率和安全造成极大的影响。预测性维护(predictive maintenance)^[31]是工业4.0提出的关键创新点之一,它基于对系统连续的测量和分析,对设备运行状态进行预测和优化维护。

当前绝大多数制造商使用预防性维护的方案提高生产线的稳定性,预防性维护通常是以时间为基准进行维护的,通过定期检修,降低一段时间内的设备故障或停机的概率。预防性维修能够降低生产线的总停机时间和设备衰退引起的故障数量,并且具有实施方便、可操作性强的优势。但由于检修时间是根据经验确定的,可能会出现维修不足或过剩的情况。随着物联网、大数据等技术的成熟,预测性维护通过分析设备的实时监测数据,预测设备可能出现的故障,并对故障的原因和解决方案提出建议,使维护更加智能。预测性维护的关键是提前排查故障隐患并解决,这能够降低维护总成本、故障率和总停机时间,提高了设备的可靠性。然而,由于终端数量多、种类繁多,预测性维护的实际应用面临着处理海量终端的连接和管理、保证分析的实时性和保护工业数据隐私的问题。

美国Efficient Plant杂志认为,边缘计算的出现对预测性维护方案的实现具有

重要意义,边缘节点强大的感知能力和与设备的近距离能够满足预测性维护的实时性和隐私保护需求。施耐德电气有限公司的分析应用构架工程师Matt Boujonnier指出,机器学习算法通常只能在云计算中心运行,但在物联网应用中,人们希望算法可以运行在任何有需要的地方。施耐德电气有限公司与微软Azure公司共同开发的Realift Rod Pump控制器已经在网络边缘实现了设备的运行状态实时分析和预测,目前已经在油田行业中进行试点应用。华为技术有限公司也指出,基于边缘计算的物联网(如EC-IoT、Edge Computing-IoT)可以有效地构建预测性维护方案,并已经推出了设计和部署预测性维护解决方案的服务。华为技术有限公司使用智能网关提供智能服务,对维护对象的关键指标进行实时监测和分析,预测维护对象可能出现的故障,并进行信息上报。云计算中心则根据多个对象的综合信息进行全面的状态评估,并可以不断迭代和优化边缘节点运行的预测算法,实现动态部署。这种维护方案综合利用了边缘节点与云计算中心的优势,能够满足预测性维护对实时性和隐私保护的要求,同时保证故障预测的准确率。

5.2 安防监控

视觉是人类认识世界、获取信息的重要途径之一,我国公安机关部署的“天网”监控系统通过在公共区域部署的大量摄像头维持稳定、安全的社会秩序,许多家庭也主动使用家用摄像头、宠物监视器等保障房屋、家庭成员的安全。最初人们只能用人工的方式处理图像信息,然而这种方式往往具有很长的时延和波动的准确率。随着人工智能技术的不断发展,图像数据计算机已经具有非常强大的学习和处理能力,然而

传统的云计算模型难以支持图像处理在部分场景的应用,以铁路轨道异物检测为例进行更详细的阐述。首先,由于铁轨铺设在室外且包含隧道、山地等网络质量无法保证的区域,使用云计算模型很可能出现图像丢失或质量损失严重的情况。其次,铁轨异物检测对实时性要求很高,必须及时发现危险并做出报警,云计算处理链路较长,加之没有网络带宽的保证,很可能无法保证检测和报警的实时性。同时,铁轨监控图像具有一定的保密性,一旦上传至云端就存在着被窃取和被篡改的风险,为不法分子所用。最后,我国铁路长度达12.1万千米,如果全部的监控图像数据均上传至云服务器进行分析,势必要求云服务器具有极其强大的处理和计算能力。显然,在实时性要求高、网络质量无法保证且涉及隐私的场景中,基于边缘计算的图像处理能够更好地提供服务。

杭州海康威视数字技术股份有限公司(以下简称海康威视)是以视频为核心的物联网解决方案提供商,在安防监控领域深耕多年。2017年10月,海康威视首次公开发布了AI Cloud的云边协同体系架构,该架构由云计算中心、边缘域和边缘节点3个部分组成,将AI推理能力赋能边缘,从而实现快速、高效的感知,云计算中心则聚焦于全局性的认知和分析。“深眸”双目行为分析摄像机是海康威视的代表性边缘智能产品,内置高性能GPU和人工智能算法,在离线状态下也能够对穿越警戒线、徘徊、奔跑、离岗等9项行为进行分析和检测,可以应用于银行金库、警卫室、医院等场景。学术界也对边缘计算在安防监控领域的应用持乐观态度,2016年密苏里大学学者在参考文献[32]中提出了一种基于雾计算的城市交通超速监控系统,系统架构由监控应用层、雾计算层和云计算层构成,其中雾计算层包括摄像头、智能平板电脑和智能手机

等设备,雾节点可以对可疑目标进行跟踪,并实时计算行驶速度,处理结果将被发送至云计算中心,这种模式可以大量减少网络流量,并提升系统的实时性。由于边缘智能设备的资源通常是有限的,而基于深度学习的人工智能算法往往需要大量的计算和存储资源才能够运行,因此算法的轻量化也是边缘计算应用于安防监控领域的研究热点之一。2016年DeepScale公司提出了一种轻量级的目标检测神经网络——SqueezeNet^[33],经过压缩后,网络体积低至0.5 MB,然而这种算法无法满足检测的实时性。2018年宾汉姆顿大学学者在参考文献^[34]中使用了一种轻量级实时检测和跟踪算法,实现了公共区域的行人识别、跟踪和异常行为检测。其中使用的Kerman算法^[35]于2018年6月被提出,它使用一种基于决策树的混合滤波器构造了用于人体目标跟踪的轻量级卷积神经网络。

6 结束语

随着物联网成为新一轮科技发展制高点,边缘计算因给物联网领域中的海量数据传输、实时服务响应等复杂挑战提供了新的解决方案,而受到了国内外政府、工业界和学术界的高度关注和认可。以边缘计算为主题的国内外组织和会议的不断涌现显示了边缘计算领域的蓬勃发展,它们推进了边缘计算在工业、无人驾驶、安防监控等各个领域的应用。

本文首先介绍了边缘计算的背景、概念和通用架构,并对国内外边缘计算产业联盟提出的工业物联网云边协同参考框架——EdgeX Foundry和边缘框架3.0进行了详细阐述。为了使边缘计算能够在参考框架的指导下成功应用于各行业场景中,本文对边缘计算当前面临的挑战进行了总结,包括

可编程性、命名、服务管理和隐私及安全,学术界与工业界积极应对挑战,并分别从架构、网络、算法、硬件等角度提出了解决方案。最后,本文介绍了边缘计算在预测性维护和安防监控领域的应用与进展。

边缘计算将云计算的计算、存储等能力扩展到了网络边缘,提供低时延、高可用和隐私保护的本地计算服务,解决了云计算时延高、受网络环境制约等问题。笔者相信,边缘计算将大力深化各行业中的物联网应用,成为推动智能制造、智慧城市等产业升级的关键技术。

参考文献:

- [1] 于文平. 《2017—2018中国物联网发展年度报告》发布[J]. 物联网技术, 2018, 8(10): 5-6.
YU W P. Annual report on China's Internet of things development from 2017 to 2018' released[J]. Internet of Things Technologies, 2018, 8(10): 5-6.
- [2] KNORR E, GRUMAN G. What cloud computing really means[J]. InfoWorld, 2008(7): 20.
- [3] BONOMI F, MILITO R, ZHU J, et al. Fog computing and its role in the Internet of things[C]// The 1st Edition of the MCC Workshop on Mobile Cloud Computing, August 17, 2012, Helsinki, Finland. New York: ACM Press, 2012: 13-16.
- [4] BELOGLAZOV A, ABAWAJY J, BUYYYA R. Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing[J]. Future Generation Computer Systems, 2012, 28(5): 755-768.
- [5] TAKABI H, JOSHI J B D, AHN G J. Security and privacy challenges in cloud computing environments[J]. IEEE Security and Privacy, 2010(6): 24-31.
- [6] PEARSON S, BENAMEUR A. Privacy, security and trust issues arising from cloud comput-

- ing[C]// 2010 IEEE 2nd International Conference on Cloud Computing Technology and Science (CloudCom), November 30-December 3, 2010, Indianapolis, USA. Piscataway: IEEE Press, 2010: 693-702.
- [7] SHI W S, CAO J, ZHANG Q, et al. Edge computing: vision and challenges[J]. *IEEE Internet of Things Journal*, 2016, 3(5): 637-646.
- [8] HU Y C, PATEL M, SABELLA D, et al. Mobile edge computing-a key technology towards 5G[M]//ETSI White Paper. [S.l.:s.n.], 2015.
- [9] WEIN J M, KLONINGER J J, NOTTINGHAM M C, et al. Content delivery network (CDN) content server request handling mechanism with metadata framework support: U.S. Patent 7,240,100[P]. 2007-7-3.
- [10] 施巍松, 张星洲, 王一帆, 等. 边缘计算: 现状与展望[J]. *计算机研究与发展*, 2019(1): 1-21.
- SHI W S, ZHANG X Z, WANG Y F, et al. Edge computing: state-of-the-art and future directions[J]. *Journal of Computer Research and Development*, 2019(1): 1-21.
- [11] PORTILLA J, MUJICA G, LEE J, et al. The extreme edge at the bottom of the Internet of things: a review[J]. *IEEE Sensors Journal*, 2019(1): 1-20.
- [12] YU W, LIANG F, HE X, et al. A survey on the edge computing for the Internet of things[J]. *IEEE Access*, 2018(6): 6900-6919.
- [13] 施巍松, 孙辉, 曹杰, 等. 边缘计算: 万物互联时代新型计算模型[J]. *计算机研究与发展*, 2017, 54(5): 907-924.
- SHI W S, SUN H, CAO J, et al. Edge computing-an emerging computing model for the Internet of everything era[J]. *Journal of Computer Research and Development*, 2017, 54(5): 907-924.
- [14] BONOMI F, MILITO R, NATARAJAN P, et al. Fog computing: a platform for internet of things and analytics[M]//Big data and Internet of things: a roadmap for smart environments. Heidelberg: Springer, 2014: 169-186.
- [15] STOJMENOVIC I, WEN S. The fog computing paradigm: scenarios and security issues[C]// 2014 Federated Conference on Computer Science and Information Systems (FedCSIS), September 7-10, 2014, Warsaw, Poland. Piscataway: IEEE Press, 2014: 1-8.
- [16] REN J, GUO H, XU C, et al. Serving at the edge: a scalable IoT architecture based on transparent computing[J]. *IEEE Network*, 2017, 31(5): 96-105.
- [17] TALEB T, SAMDANIS K, MADA B, et al. On multi-access edge computing: a survey of the emerging 5G network edge cloud architecture and orchestration[J]. *IEEE Communications Surveys and Tutorials*, 2017, 19(3): 1657-1681.
- [18] GARCIA LOPEZ P, MONTRESOR A, EPEDA D, et al. Edge-centric computing: vision and challenges[J]. *ACM SIGCOMM Computer Communication Review*, 2015, 45(5): 37-42.
- [19] 安星硕, 曹桂兴, 苗莉, 等. 智慧边缘计算安全综述[J]. *电信科学*, 2018, 34(7): 135-147.
- AN X S, CAO G X, MIAO L, et al. Security of intelligent edge computing: a survey[J]. *Telecommunications Science*, 2018, 34(7): 135-147.
- [20] CHEN T, MOREAU T, JIANG Z, et al. TVM: end-to-end compilation stack for deep learning[J]. *Computer Science*, 2018, arXiv:1802.04799v1.
- [21] MOREAU T, CHEN T, JIANG Z, et al. VTA: an open hardware-software stack for deep learning[J]. *Computer Science*, 2018, arXiv:1807.04188.
- [22] MOCKAPETRIS P, DUNLAP K J. Development of the domain name system[M]. New York: ACM Press, 1988.
- [23] ZHANG L, ESTRIN D, BURKE J, et al. Named data networking (ndn) project[J]. *ACM SIGCOMM Computer Communication Review*, 2010, 44(3): 66-73.
- [24] WU H, SHI J, WANG Y, et al. On incremental deployment of named data networking in local area networks[C]// 2017 ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS), May 18-19, 2017, Beijing, China. Piscataway: IEEE Press, 2017: 82-94.

- [25] ZHANG D, MA Y, ZHENG C, et al. Cooperative-competitive task allocation in edge computing for delay-sensitive social sensing[C]//2018 IEEE/ACM Symposium on Edge Computing (SEC), February 24-25, 2018, Dubai, United Arab Emirates. Piscataway: IEEE Press, 2018: 243-259.
- [26] XIAO K, GAO Z, WANG Q, et al. A heuristic algorithm based on resource requirements forecasting for server placement in edge computing[C]//2018 IEEE/ACM Symposium on Edge Computing (SEC), February 24-25, 2018, Dubai, United Arab Emirates. Piscataway: IEEE Press, 2018: 354-355.
- [27] 庄小君, 杨波, 王旭, 等. 移动边缘计算安全研究[J]. 电信工程技术与标准化, 2018, 31(12): 38-43.
- ZHUANG X J, YANG B, WANG X, et al. Approach on mobile edge computing security[J]. Telecom Engineering Technics and Standardization, 2018, 31(12): 38-43.
- [28] SHI W, DUSTDAR S. The promise of edge computing[J]. Computer, 2016, 49(5): 78-81.
- [29] CHI J, OWUSU E, YIN X, et al. Privacy partition: a privacy-preserving framework for deep neural networks in edge networks[C]//2018 IEEE/ACM Symposium on Edge Computing (SEC), February 24-25, 2018, Dubai, United Arab Emirates. Piscataway: IEEE Press, 2018: 378-380.
- [30] TRIMANANDA R, YOUNIS A, WANG B, et al. Vigilia: securing smart home edge computing[C]//2018 IEEE/ACM Symposium on Edge Computing (SEC), February 24-25, 2018, Dubai, United Arab Emirates. Piscataway: IEEE Press, 2018: 74-89.
- [31] MOBLEY R K. An introduction to predictive maintenance[M]. Amsterdam: Elsevier, 2002.
- [32] CHEN N, CHEN Y, SONG S, et al. Smart urban surveillance using fog computing[C]//2016 IEEE/ACM Symposium on Edge Computing (SEC), October 27-28, 2016, Washington, DC, USA. Piscataway: IEEE Press, 2016: 95-96.
- [33] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. Squeezenet: alexnet-level accuracy with 50x fewer parameters and < 0.5 MB model size[J]. Computer Science, 2016, arXiv:1602.07360.
- [34] NIKOUEI S Y, CHEN Y, FAUGHNAN T R. Smart surveillance as an edge service for real-time human detection and tracking[C]//2018 IEEE/ACM Symposium on Edge Computing (SEC), February 24-25, 2018, Dubai, United Arab Emirates. Piscataway: IEEE Press, 2018: 336-337.
- [35] NIKOUEI S Y, CHEN Y, SONG S, et al. Kerman: a hybrid lightweight tracking algorithm to enable smart surveillance as an edge service[J]. Computer Science, 2018, arXiv: 1808.02134.

作者简介



李林哲 (1994-), 女, 浙江大学工业控制技术国家重点实验室硕士生, 主要研究方向为预测性维护、边缘计算场景中的机器视觉应用。



周佩雷 (1977-), 男, 浙江省公众信息产业有限公司智慧旅游事业部总经理, 浙江省公众信息旅游云重点企业研究院副院长, 中国电信旅游行业信息化应用基地副总经理, 主要研究方向为旅游大数据的架构及应用、物联网应用、无线网络的基础应用研究。



程鹏 (1982-), 男, 博士, 浙江大学工业控制技术国家重点实验室教授、博士生导师, 工业控制研究所副所长、网络空间安全研究中心副主任, 主要研究方向为控制系统安全、物联网及信息物理融合系统、数据安全与隐私保护。



史治国 (1978-), 男, 博士, 浙江大学工业控制技术国家重点实验室教授、博士生导师, 信息与电子工程学院电子工程系副系主任, 主要研究方向为群智感知、目标跟踪与定位、物联网系统设计、反无人机技术与系统。

收稿日期: 2019-01-18