

2024 IEEE International Conference on Edge Computing and Communications (EDGE 2024)

**Shenzhen, China
7-13 July 2024**



**IEEE Catalog Number: CFP24L50-POD
ISBN: 979-8-3503-6850-5**

**Copyright © 2024 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP24L50-POD
ISBN (Print-On-Demand):	979-8-3503-6850-5
ISBN (Online):	979-8-3503-6849-9
ISSN:	2767-990X

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2024 IEEE International Conference on Edge Computing and Communications (EDGE) **EDGE 2024**

Table of Contents

Congress Steering Committee Chair Message	ix
Congress General Chair Message	x
Congress Program Chairs Message	xii
TCSVC Chair Message	xiii
EDGE 2024 General Chairs Message	xiv
EDGE 2024 Program Committee	xvi
EDGE 2024 Reviewers	xviii

EDGE 1: System

An Empirical Study on Edge-to-Cloud Continuum for Smart Applications: Performance, Design Patterns, and Key Factors	1
<i>Norman Chen (Monash University, Australia), Adel N. Toosi (Monash University, Australia), Bahman Javadi (Western Sydney University, Australia), Daghash Alqahtani (Monash University, Australia), Mohammad Sadegh Aslanpour (Monash University, Australia), and Minxian Xu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China)</i>	
Cross Network Layer Cognitive Service Orchestration in Edge Computing Systems	12
<i>Milad Rafiee (University of Oslo, Norway), Amir Taherkordi (University of Oslo, Norway), and Özgü Alay (University of Oslo, Norway & Karlstad University, Sweden)</i>	
CargoSync: Efficient Containerd Image Updates in Low-Bandwidth Edge Environments with Rsync-based Delta Compression	22
<i>Ioannis Protogeris (TWI Hellas & National Technical University of Athens), Michalis Karamousadakis (Plaixus Ltd.), Michail Rizakis (University of Essex), and Antonios Porichis (University of Essex)</i>	
Do 5G Networks Achieve The Proclaimed Promises? An Empirical Study Using YouTube Edge Service	29
<i>Peixuan Song (University of Cambridge), Junkyu Lee (University of Essex), Ahmed M. Abdelmoniem (Queen Mary University of London), and Lev Mukhanov (Queen Mary University of London)</i>	

EDGE 2: Scheduling & Routing

Navigator: A Decentralized Scheduler for Latency-Sensitive AI Workflows	35
<i>Yuting Yang (Cornell University, USA), Andrea Merlina (University of Oslo, Norway), Weijia Song (Cornell University, USA), Tiancheng Yuan (Cornell University, USA), Ken Birman (Cornell University, USA), and Roman Vitenberg (University of Oslo, Norway)</i>	
Reliable Network Performance for Edge Networks with QoS-Aware Adaptive Routing	48
<i>Osama Abu Hamdan (University of Texas at Arlington) and Engin Arslan (University of Texas at Arlington)</i>	
Concierge: Towards Accuracy-Driven Bandwidth Allocation for Video Analytics Applications in Edge Network	59
<i>Yuyang Huang (University of Chicago, USA), Faishal Zharfan (Institut Teknologi Bandung, Indonesia), Hendrawan Hendrawan (Institut Teknologi Bandung, Indonesia), Haryadi S. Gunawi (University of Chicago, USA), and Junchen Jiang (University of Chicago, USA)</i>	
QEdgeProxy: QoS-Aware Load Balancing for IoT Services in the Computing Continuum	67
<i>Ivan Čilić (University of Zagreb, Croatia), Valentin Jukanović (University of Zagreb, Croatia), Ivana Podnar Žarko (University of Zagreb, Croatia), Pantelis Frangoudis (Distributed Systems Group, TU Wien, Austria), and Schahram Dustdar (Distributed Systems Group, TU Wien, Austria)</i>	

EDGE 3: Inference Optimization

The Impact of GPU on Containerized Computer Vision Applications on Edge Nodes	74
<i>Osamah Alqaisi (University of Texas at San Antonio), Ali Tosun (The University of North Carolina at Pembroke), and Turgay Korkmaz (University of Texas at San Antonio)</i>	
Split DNN Inference for Exploiting Near-Edge Accelerators	84
<i>Hao Liu (KAUST), Mohammed Fouda (Rain Neuromorphics Inc.), Ahmed Eltawil (KAUST), and Suhaib A. Fahmy (KAUST)</i>	
A Case for Deploying Dynamic Neural Network on Edge-Cloud Continuum Environment	92
<i>Muhammad Syafiq Mohd Pozi (Toyohashi University of Technology, Japan) and Yukinori Sato (Toyohashi University of Technology, Japan)</i>	

EDGE 4: Model Training & Compression

FedCode: Communication-Efficient Federated Learning via Transferring Codebooks	99
<i>Saeed Khalilian Gourtani (Eindhoven University of Technology), Vasileios Tsouvalas (Eindhoven University of Technology), Tanir Ozcelebi (Eindhoven University of Technology), and Nirvana Meratnia (Eindhoven University of Technology)</i>	

Characterizing Deep Learning Model Compression with Post-Training Quantization on Accelerated Edge Devices	110
<i>Rakandhiya D. Rachmanto (Universitas Amikom Yogyakarta), Zaki Sukma (University of Georgia), Ahmad N. L. Nabhaan (Universitas Amikom Yogyakarta), Arief Setyanto (Universitas Amikom Yogyakarta), Ting Jiang (University of Georgia), and In Kee Kim (University of Georgia)</i>	
Classifier Reuse-Based Contrastive Knowledge Distillation	121
<i>Zhipeng Gao (Beijing University of Posts and Telecommunications), Shu Yang (Beijing University of Posts and Telecommunications), Xinlei Yu (Beijing University of Posts and Telecommunications), Zijia Mo (Industrial and Commercial Bank of China), Lanlan Rui (Beijing University of Posts and Telecommunications), and Yang Yang (Beijing University of Posts and Telecommunications)</i>	

EDGE 5: Task Partitioning & Allocation

Markov Blanket Composition of SLOs	128
<i>Boris Sedlak (TU Wien, Austria), Victor Casamayor Pujol (TU Wien, Austria), Praveen Kumar Donta (TU Wien, Austria), and Schahram Dustdar (TU Wien, Austria)</i>	
Data Sharing-Aware Online Algorithms for Task Allocation in Edge Computing Systems	139
<i>Sanaz Rabinia (Wayne State University) and Daniel Grosu (Wayne State University)</i>	
DONNA: Distributed Optimized Neural Network Allocation on CIM-Based Heterogeneous Accelerators	149
<i>Mojtaba AlShams (King Abdullah University of Science and Technology, Saudi Arabia), Kamilya Smagulova (King Abdullah University of Science and Technology, Saudi Arabia), Suhaib Fahmy (King Abdullah University of Science and Technology, Saudi Arabia), Mohammed Fouda (Rain Neuromorphics, USA), and Ahmed Elta'wil (King Abdullah University of Science and Technology, Saudi Arabia)</i>	
SlimNet: A Lightweight Attentive Network for Speech-Music-Noise Classification and Voice Activity Detection	157
<i>Nicolas Shu (Georgia Institute of Technology, USA), You Wang (Georgia Institute of Technology, USA), Desmond Caulley (Georgia Institute of Technology, USA), and David V. Anderson (Georgia Institute of Technology, USA)</i>	
Are Large Language Models Good Neural Architecture Generators for Edge?	162
<i>Hadjer Benmeziane (IBM Research Europe) and Kaoutar El Maghraoui (IBM T.J. Watson Research Center)</i>	
EfficientMedSAM: Accelerating Medical Image Segmentation via Neural Architecture Search and Knowledge Distillation	166
<i>Arnold Caleb Asiimwe (Columbia University), William Das (Columbia University), Hadjer Benmeziane (IBM Research Europe), and Kaoutar El Maghraoui (IBM T. J. Watson Research Center)</i>	

Author Index 175