2024 21st Annual IEEE International Conference on Sensing, Communication, and Networking (SECON 2024)

Phoenix, Arizona, USA
2-4 December 2024

IEEE Catalog Number: CFP24SCN-POD
ISBN: 979-8-3315-1919-3

****** 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:	CFP24SCN-POD
ISBN (Print-On-Demand):	979-8-3315-1919-3
ISBN (Online):	979-8-3315-1918-6
ISSN:	2155-5486

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

Tuesday, December 3

Tuesday, December 3 10:30 - 11:30

Session 1A: Session 1A: Network Traffic Optimization and Scheduling

Online Scheduling of Federated Learning with In-Network Aggregation and Flow Routing

Mingtao Ji (Nanjing University, China); Lei Jiao (University of Oregon, USA); Yitao Fan (Nanjing University, China); Yang Chen (Fudan University, China); Zhuzhong Qian (Nanjing University, China); Ji Qi (China Mobile Suzhou Software Technology Company Limited, China); Gangyi Luo (China Mobile (Suzhou) Software Technology Co Ltd, China); Baoliu Ye (Nanjing University, China)

Continuously orchestrating in-network model aggregations for federated learning faces fundamental challenges such as the combinatorial nature of traffic reduction, the dynamic trade-offs between system overhead and model convergence, and the unpredictable inputs from uncertain system environments. In this work, we model a nonlinear mixed-integer program to optimize the long-term total cost of federated learning computation overhead, traffic reduction, network delay, and programmable switch reconfigurations over time. To attack the lexicographic minimax, submodular, and online nature of this problem, we propose a polynomial-time algorithmic framework to judiciously designate the timing of reconfigurations, while designing and invoking a linearized transformation for selecting routing paths, a greedy sub-algorithm for selecting aggregation locations, and an online learning sub-algorithm for controlling federated learning convergence. We demonstrate our rigorous mathematical insights behind our algorithms, and prove the competitive ratio as the performance guarantee. Using trace-driven evaluations, we have validated our approach's superiority over existing methods.

pp. 1-9

TINIEE: Traffic-Aware Adaptive In-Network Intelligence via Early-Exit Strategy

Heewon Kim, Seongyeon Yoon, Chanbin Bae, Sanghoon Lee and Sangheon Pack (Korea University, Korea (South))

In-network (or on-path) inference over programmable data planes (PDPs) allows fast and low-overhead inference using deep neural networks (DNN). To alleviate massive processing and deployment cost of in-network inference, a distributed deployment on multiple programmable network devices is mainly adopted. However, it is likely to produce considerable amount of network traffic due to the exclusive forwarding chain and intermediate data between submodels. In this work, we propose a traffic-aware adaptive in-network inference scheme, TINIEE, to maximally reduce the network traffic of in-network inference without causing a significant classification performance drop. To this end, we first devise an adaptive inference method on the data plane striking the balance between the classification performance and the network traffic cost. Furthermore, we formulate a traffic minimization problem to decide the proper location of each submodel considering each flow's exit tendency with a predefined confidence threshold. Since the problem is excessively complicated, we devise a low-complexity practical submodel placement algorithm. We implement the proposed scheme on software-programmable switches, and the evaluation results demonstrate that TINIEE reduces network traffic by up to 34.48% compared to the state-of-the-art, while maintaining sufficiently high classification performance.

pp. 10-18

RTCC: Enable End-to-end Sub-RTT Congestion Control for Next-generation Network

Yihang Zhang, Zhidong Jia, Qingyang Li, Xinggong Zhang and Zongming Guo (Peking University, China)

The advancement of next-generation networks such as 5G/6G and satellite systems has significantly increased available network bandwidth, while also exacerbating network burstiness. This surge presents a formidable challenge for congestion control (CC), a pivotal mechanism for achieving high bandwidth utilization and low latency by adjusting congestion windows or modifying sending rates. Traditional end-to-end CC algorithms fall short of optimality due to

their reliance on congestion signals in acknowledgment packets, which introduce a delay of one round-trip time (RTT).

In this paper, to mitigate end-to-end delayed feedback, we introduce a novel Real-Time Congestion Control (RTCC) algorithm that integrates machine learning with conventional model-based CC. RTCC employs a Multi-Layer Perceptron~(MLP) to model network conditions and predict current congestion signals accurately. A tailored network model then utilizes these predictions to manage packet accumulation for individual flows in the network. Moreover, an online model-updating mechanism is proposed to adapt to diverse network environments.

We integrated RTCC into QUIC and conducted comprehensive experiments in both emulated test-beds and real-world settings, including WiFi/4G/5G and cross-continent networks. The results demonstrate RTCC's efficacy, with up to a 32% increase in average throughput, a reduction in RTT by up to 21% compared to BBR V2, and the maintenance of fair bandwidth allocation.

pp. 19-27

OFC: An Original congestion-based Fine-grained Priority Flow Control

Wenli Xiao (Shanghai Jiaotong University & John Hopcroft Center, China); Yuqing Yang (Shanghai Jiao Tong University, China); Peirui Cao (Nanjing University, China); Zhuoran Liu and Shizhen Zhao (Shanghai Jiao Tong University, China); Xinbing Wang (Shanghai Jiaotong University, China)

With the proliferation of online data intensive applications and virtualized services, the growing complexity of traffic patterns in data centers increases the likelihood of congestion, especially in incast scenarios and with a combination of short and large flows. To ensure lossless transmission, RDMA over Converged Ethernet networks rely on Priority-based Flow Control (PFC) to prevent packet loss due to buffer overflow. However, it is widely acknowledged that PFC gives rise to several issues, such as Congestion Spreading, Head-of-Line Blocking, and Deadlock, which are increasingly prominent in modern highly congested data centers. In this paper, we analyze the primary causes of congestion spreading and head-of-line blocking issues associated with PFC and propose Original congestion- based fine-grained priority Flow Control (OFC) as a solution. The performance of OFC is assessed using the programmable switch Tofino and simulations carried out with a packet-level simulator across various scenarios, encompassing incast, realistic, deadlock, and in-depth scenarios. The validation of the simulation results through testbed evaluation confirmed that OFC effectively reduces flow completion time, buffer occupancy, and deadlock occurrence by up to 60.28%, 51.47%, and 48.7%, respectively

pp. 28-36

Tuesday, December 3 11:30 - 12:30

Session 1B: Session 1B: Federated Learning and Active Learning Techniques

GDFAL: Global Diversity-based Federated Active Learning

Enzhi Zhang and Liu Yang (Tianjin University, China)

The objective of federated active learning is to strategically label a subset of the dataset within clients to optimize the performance of the global model within a predetermined labeling budget. Due to disparate data distribution characteristics of the samples across different clients, the independent selection of samples from disparate clients often results in the inclusion of samples with redundant information or low diversity. Furthermore, existing methods result in additional communication overheads while enhancing annotation efficiency. In this study, we deliberately select samples that are both globally diverse and informative. We introduce the busybody score, a simple yet effective measure, to assess the global rarity of intra-client samples. We subsequently introduce global diversity-based federated

active learning (G DFAL), a framework with a sub-query structure to efficiently compute the busybody score and an information hunter score designed to balance data diversity and uncertainty. Our method outperforms existing federated active learning methods on widely recognized benchmarks. Furthermore, we evaluate the effectiveness of our proposed framework in extended settings, including different federated learning data distributions and different query ratio. We also compare different communication overheads and the results show that our approach still outperforms the best existing methods even at low communication overheads, which is sustained by the sub-query structure.

N/A

One Cycle With N Messages: A Secure and Lightweight Key Agreement Scheme for WSNs

Shenglin Gui, Meng Zheng and Yangyang Liu (University of Electronic Science and Technology of China, China)

In recent years, wireless sensor networks (WSNs) have been widely used in various critical applications. Secure and efficient mutual authentication and key agreement (AKA) schemes under resource-constrained WSNs are crucial to guarantee security. A wealth of AKA schemes have been proposed, yet most of them are often found to be either insufficient in security or poor in performance. To design a scheme both with robust security and better performance, we propose a novel secure and lightweight three-factor AKA scheme under one cycle with N messages mode, where its communication process between N participants requires only N messages in one cycle. Furthermore, we adopt enhanced Chebyshev polynomials to reduce computation overheads while ensuring computational difficulties against attacks and to provide secure access between remote users and sensor nodes with forward secrecy and multi-factor security. The security of this scheme is proved under the random-oracle model and the informal security analysis further demonstrates that our scheme succeeds in meeting all evaluation criteria in the face of the state-of-the-art threat model. At last, we conduct adequate comparisons respectively on both single-gateway and multi-gateway scenarios. Compared to other state-of-the-art schemes in terms of security and performance, to our best knowledge, our scheme achieves all the security requirements with minimum communication overheads and low computation overheads. We believe that our work takes a substantial step towards a better trade-off between security and efficiency for AKA schemes in WSNs.

pp. 46-54

ZChirp: Speeding Up LPWANs by Combining the Chirp with Binary Sequences

Zhenghao Zhang (Florida State University, USA)

Low Power Wide Area Networks (LPWAN), such as LoRa, support long range wireless connections for a vast number of low power IoT devices. One of the main challenges in LPWANs is to provide high network capacity. In this paper, a novel modulation technique, referred to as ZChirp, is proposed. ZChirp augments the chirp modulation in LoRa by multiplying the chirp with a binary sequence called the Z sequence, which serves dual purposes. First, chirps multiplied with different Z sequences have very low cross-correlations, which enables simultaneous transmissions of multiple nodes and increases the network capacity. Second, additional data can be modulated by using different Z sequences for different data values, which proves to be an effective option to increase the data rate. Experiments and simulations show that ZChirp does not increase bandwidth usage nor reduce the communication range comparing to the original LoRa. Experiments in the POWDER platform demonstrate that the ZChirp gateway successfully decoded packets from multiple nodes transmitted simultaneously at random times. Simulations further show that the capacity gain of ZChirp over the state-of-the-art solution that also modifies the LoRa modulation is 2.0X

pp. 55-63

BETA: A Novel Learning-based Adaptive Streaming Approach with Spatial and Temporal Optimization

Ziming Wang, Guanghui Zhang, Mengbai Xiao, Dongxiao Yu and Xiuzhen Cheng (Shandong University, China)

Adaptive video streaming (DASH) has become a key technology in video transmission. Given the advantages of deep reinforcement learning (DRL), streaming vendors are increasingly focusing on DRL-based adaptive bitrate (ABR)

algorithms. In practice, almost all the ABR algorithms are designed with the intention to work well in any size/shape of networks. However, far from expectations, our measurements revealed that, even with extensive training on a vast scale of real network trace data, if the network condition varies over a wide range, the achieved Quality-of-Experience (QoE) is only 43.1% ~ 48.9%, far below the optimal 100%. We termed this problem as "ABR Under-Generalization." To address this challenge, we developed BETA, a novel ABR-specialized DRL-based approach that integrates spatial and temporal modules. 1) Spatial: BETA offers a detector that identifies the potential network condition that may cause poor performance and then trains complementary ABR algorithms specifically functioning in it; 2) Temporal: At each training epoch, BETA learns from the decision experience of multiple-step segments to cover long-term environmental feedback. Extensive evaluations demonstrate that BETA achieves an average QoE improvement of 11.4% to 17.3% compared to the state-of-the-art methods, with gains of up to 244.1% in the challenging networks with dramatically variable throughput.

pp. 64-72

Tuesday, December 3 14:00 - 15:00

Session 2A: Session 2A: RFID, Sensors, and IoT Protocols

Parallel Missing Tag Identification for Anonymous Multiple Users RFID Systems

Jiangjin Yin (Huazhong Agricultural University, China); Hangyu Mao (Kuaishou Technology, China); Rongbo Zhu and Shiwei Xu (Huazhong Agricultural University, China)

Radio frequency identification (RFID) system has been widely employed in warehouse management and supply logistics. A fundamental systematic functionality is to determine the presence or absence of tagged items, referred to as missing tag identification. Although this research has attracted extensive attention, prior methods predominantly address scenarios involving a single user. In this paper, we extend the research to more common multi-user scenarios, which raise two concerns: time sensitivity and privacy protection. We propose a novel \underline{Pa}rallel \underline{M}issing tag \underline{I}dentification protocol (PaMI) that leverages lightweight and anonymous bit-vector techniques to specify the data and order of tag responses. This effectively prevents both intra- and inter-user tag collisions, enabling the identification of missing tags across multiple users in one shot while safeguarding user privacy. We carry out a comprehensive theoretical analysis to optimize the proposed method's performance. Extensive experiments show that our protocol strikingly outperforms state-of-the-art baselines.

pp. 73-81

UWB-ND: Low-power neighbor discovery protocol for Ultra-wideband radio networks

Alireza Ansaripour, Aryo Yarahmadi, Milad Heydariaan and Omprakash Gnawali (University of Houston, USA)

Due to the frequent topology changes in most wire- less networks, low-power Neighbor Discovery (ND) is essential for many Wireless Sensor Networks (WSNs) and Internet of Things (IoT) applications. In this work, we present UWB-ND, a low- power ND protocol for ultra-wideband (UWB) radio networks that are becoming increasingly popular in IoT applications. To conserve energy, these IoT applications typically rely on other low-power radio technology such as Bluetooth Low Energy (BLE) for ND, requiring the integration of auxiliary radios in all nodes. Utilizing specific characteristics of UWB radios such as efficient Channel Activity Detection (CAD) and varying preamble modulation, UWB-ND introduces a low-power ND approach specific to UWB radios. Our evaluation shows that UWB-ND can reduce ND power consumption by 60%, compared to the state-of-the-art PI-based approach.

pp. 82-90

mmMC: A Contactless Wood Moisture Content Measurement System based on COTS FMCW mmWave Radar

Xueyang Hu (Peng Cheng Laboratory, China); Tian Liu (Zhejiang Academy of Agricultural Sciences, China); Minarul Islam and Tao Shu (Auburn University, USA)

Wood moisture content (MC) estimation is a fundamental aspect of woodworking, construction, and various other industries that rely on the versatile properties of wood. In this paper, we present mmMC, a novel system that estimates the MC of wood by a single commercial off-the-shelf (COTS) mmWave radar, which provides accurate end-to-end real-time wood moisture measurement results while being non-invasive, portable, and flexible in deployment. The proposed system uses a novel target reflection feature (TRF) to determine the reflectivity of the wood, thereby correlating it with specific MC levels. To robustly and accurately estimate the TRF of the wood from mmWave signal, a signal processing pipeline is proposed. Specifically, an object detection and signal extraction module is designed to resolve the range of the object and eliminate multipath effects from the original signal. Then, a multiple chirps/antennas signals processing module is proposed to obtain a stable TRF from the extracted signals by using signals from different antennas and chirps. The TRF is associate with the MC of wood by a regression to enable real-time MC estimation. Through extensive real-world experiments, we have demonstrated that the proposed system has high accuracy and reliability.

pp. 91-99

LAORA: Location-Aware Orientation Adjustment for MIMO Magnetic Wireless Charging System

Lingchang Kong, Xinyu Wang, Hao Zhou, Fengyu Zhou, Shenyao Jiang, Peide Zhu and Qi Song (University of Science and Technology of China, China); Zhi Liu (The University of Electro-Communications, Japan)

Wireless power transfer (WPT) systems using magnetic resonant coupling (MRC) have made significant progress recently, leading to various optimization methods in scenarios involving multiple-input multiple-output (MIMO) to improve charging performance. Adjusting the coil orientation of the power transmitter (TX) is a simple but effective method due to the directional nature of magnetic field distribution, but existing approaches often require unnecessary coil rotations. In this study, we introduce a Location-Aware Orientation Adjustment algorithm, known as LAORA, to address these inefficiencies. LAORA focuses on solving the problems of charging devices (RX) localization and location-based optimization. We begin by introducing the concept of Equivalent Impedance Distribution Image (EIDI) and transform the RX localization problem into a combined matching process involving EIDI. In addition, we establish a dynamic simulation framework to predict charging performance using RX-related knowledge, enabling us to obtain optimal TX orientations through reinforcement learning without needing to rotate on mechanical devices physically. We then implement the prototype and conduct extensive experiments. The results show that, compared to other existing orientation adjustment methods, LAORA achieves an average improvement of 186% while reducing the mechanical rotations by 83.3%.

pp. 100-108

Tuesday, December 3 15:00 - 15:30

Posters and Demos 1: Posters and Demos 1

3D Detection Under Foggy Weather: The Effective Multi-Sensor Fusion of LiDAR and Radar

Jiake Tian (South China University of Technology, China); Dacheng Li (Gosuncn Technology Group Co. Ltd, China); Yi Zou, Jiale Lai and Zufeng Liang (South China University of Technology, China)

Adverse weather conditions are a significant challenge for fully automated driving systems in the domain of autonomous vehicles. While radar sensor integration can alleviate some of the adverse effects of weather, the inherent sparsity and altitude discrepancies in radar data remain significant concerns. In response to these challenges, we introduce an innovative sensor fusion method for LiDAR and radar data under foggy weather conditions. This method effectively augments radar data in voxel space and facilitates the synchronization and complementation of LiDAR and radar features in the bird's eye view space, thereby achieving high-precision 3D object detection. Through extensive experimentation on the Oxford Radar RobotCar dataset under foggy weather conditions, the effectiveness of this method has been extensively validated.

pp. 109-110

Driving Important Scene Detection

Yuta Tsubaki (Toyota Motor North America R&D & InfoTech Labs, USA); Seyhan Ucar (Toyota Motor North America R&D, InfoTech Labs, USA); Emrah Akin Sisbot (InfoTech Labs, Toyota Motor North America, USA); Xiaofei Cao and Kentaro Oguchi (Toyota Motor North America R&D, InfoTech Labs, USA)

A driving important scene refers to scenarios that pose challenges for Advanced Driver Assistance Systems (ADAS). The important scene data should be collected and used for data-driven development to address potential bugs in ADAS. However, collecting this data is costly, and user preferences for important scenes can vary, making detection challenging. In this paper, we propose a method to identify important scenes based on differing user preferences. Our approach filters driving data to extract the most relevant scenes. We annotate key scenes within the NuScenes dataset and demonstrate that our method achieves an accuracy of 62% in detecting important scenes.

pp. 111-112

Spatial Clustering of LTE Base Stations for Spectrum Usage Diagnosis in Congested Areas

Yunbae Kim (Electronics and Telecommunications Research Institute, Korea (South)); Jonghun Yoon (Electronics and Telecommunications Research Institute (ETRI), Korea (South)); Hyeyeon Kwon (Electronics Telecommunications Research Imstitute, Korea (South)); Seung Keun Park (Electronics and Telecommunications Research Institute, Korea (South))

This poster presents a method to cluster LTE base stations based on their locations to diagnose the usage of fragmented spectrum resources. The growing demand for low-frequency spectrum in next-generation networks highlights the need to evaluate the adequacy of LTE spectrum utilization. We analyze resource block usage data from a congested area in Korea, provided monthly by mobile network operators. Density-based clustering and Voronoi diagrams are applied to divide regions and assess total spectrum utilization. Probabilistic modeling using the areas of the regions provides an indicator of spectrum utilization.

pp. 113-114

Deep Learning-Based 5G SINR Prediction using Urban DEM data

Gyeong-June Hahm (Electronics and Telecommunications Research Institute, Korea (South)); Kyung-yul Cheon (ETRI(Electronics and Telecommunications Research Institute), Korea (South)); Hyeyeon Kwon (Electronics Telecommunications Research Imstitute, Korea (South)); Seung Keun Park (Electronics and Telecommunications Research Institute, Korea (South))

Accurate prediction of 5G SINR (Signal-to-Interference-and-Noise Ratio) in urban environments is essential for the effective operation of 5G networks. To date, SINR prediction has predominantly been performed through channel modeling using statistical approaches. However, these methods have limitations in adequately accounting for terrain information, such as buildings, in urban settings. In this study, we propose a deep learning model that predicts the SINR of user equipment (UE) at specific points by fully incorporating terrain information from digital elevation model (DEM) data and 5G base station (BS) location information. The prediction results demonstrate that our model provides more accurate predictions compared to existing simulation-based methods and closely aligns with actual measurements.

pp. 115-116

Demo: Prevention of Fall-on-Car Incidents

Seyhan Ucar (Toyota Motor North America R&D, InfoTech Labs, USA); Xiaoyu Zhang (InfoTech Labs, Toyota Motor North America R&D, USA); Emrah Akin Sisbot (InfoTech Labs, Toyota Motor North America, USA); Kentaro Oguchi (Toyota Motor North America R&D, InfoTech Labs, USA)

Fall-on-car incidents (e.g., trees or branches falling on a car) are an underestimated hazard during weather events, mainly affecting vehicles. Unfortunately, drivers are usually unaware of this danger until it occurs. On the other hand, connected vehicles can sense their surroundings, analyze this data with weather forecasts, and alert the driver if there

is a risk of a fall-on-car incident. This paper focuses on this use case and demonstrates the Fall-on-Car Prevention (FoP) system. FoP system detects trees and tree branches and alerts drivers when windy conditions are forecasted, allowing early preventative action to be taken while parking. Our evaluation compared to 12 human experts demonstrates that the FoP system can enhance driver awareness of the risk of falling objects.

pp. 117-119

Tuesday, December 3 15:30 - 16:30

Session 2B: Session 2B: Optimization in Hybrid and Distributed Systems

Scalable and Distributed Optimization of Shared 3D Object Quality for Large-Scale Hybrid-Metaverses

Yui Maruyama, Tatsuya Amano and Hirozumi Yamaguchi (Osaka University, Japan)

Hybrid metaverses, integrating physical and virtual spaces, face a critical challenge in managing shared 3D object quality across multiple users with diverse preferences and limited network resources. This paper addresses the problem of allocating limited bandwidth for transmitting point cloud representations while maximizing overall user satisfaction. We propose a distributed optimization method that dynamically adjusts 3D object quality based on contextual importance, available resources, and user preferences. Our approach uses Input Convex Neural Networks (ICNN) to model user utility functions and employs the Alternating Direction Method of Multipliers (ADMM) for distributed optimization. Key advantages include scalability, adaptability, and improved quality of experience. Evaluation using both real-world data we captured and open dataset demonstrates significant improvements in user satisfaction and resource utilization compared to baseline approaches. Our method achieves 93-94.6% accuracy in modeling user utility and shows up to 60% faster convergence for scenarios with 30 users, contributing to the balance between high-fidelity representation and efficient data management in hybrid metaverses.

pp. 120-128

Budget-Constrained Traveling Salesman Problem: a Cooperative Multi-Agent Reinforcement Learning Approach

King To Vincent Mak (CSUDH, USA); Christopher Gonzalez (California State University Dominguez Hills, USA); Zari Justine Magnaye and Jessica Gonzalez (CSUDH, USA); Yutian Chen (California State University Long Beach, USA); Bin Tang (California State University Dominguez Hills, USA)

This paper studies a new variation of the Traveling Salesman Problem (TSP) called the Budget-Constrained Traveling Salesman Problem (BC-TSP). BC-TSP is inspired by a few emerging network applications, such as robotic sensor networks. We design a prize-driven multi-agent reinforcement learning algorithm called P-MARL to solve the BC-TSP. The main novelty of P-MARL is making a connection between the prize maximization in BC-TSP and the cumulative reward maximization in reinforcement learning (RL) to design a more efficient RL algorithm. In particular, P-MARL integrates the prizes available at nodes into the reward model of the RL via the cooperative effort of multiple learning agents. Via extensive simulations using synthetic data of state capital cities of the U.S., we show that a) the P-MARL outperforms the existing {\em prize-oblivious} MARL work by collecting 28.8% of more prizes under the same budget constraints, b) it takes two orders of magnitudes of shorter training time than the state-of-the-art deep reinforcement learning-based approach while collecting 45.3% more prizes under the same budgets, and c) P-MARL collects prizes at least 91.9% of optimal obtained by the Integer Linear Programming (ILP) under different network parameters.

pp. 129-137

Budget-Constrained and Deadline-Driven Multi-Armed Bandits with Delays

Abdalaziz Sawwan and Jie Wu (Temple University, USA)

Many extensions of the Multi-Armed Bandit (MAB) problem were studied recently offering a strong theoretical basis for applications that require active learning. The Budget-Constrained and Deadline-Driven MAB with Delay (BD-MAB)

variation introduces a novel scenario in which a player must pull from K arms, each associated with reward, delay, and cost distributions. To the best of our knowledge, this is the first work that combines budget constraints with time delays in an MAB problem. The model involves three phases. The first phase comprises pulling arms, while incurring random cost, and observing rewards. This phase ends when the budget is depleted. The second phase includes observing some of the delayed rewards. The third phase starts at a fixed termination deadline, marking the end of reward observation, where any rewards returned after this phase are not observed and are considered dead. We present a novel solution to this problem by developing a new Upper Confidence Bound (UCB)-based algorithm. The name of the algorithm is Budget-Constrained and Deadline-Driven UCB with Delay (BD-UCB) algorithm. We provide extensive regret analysis that confirms the efficiency of our approach in managing these complexities. Lastly, we provide numerical simulations which further demonstrate the effectiveness of our proposed solution.

pp. 138-146

DeepReal: Short-form Video Streaming with Fine-grained Bitrate Adaptation

Guanghui Zhang (Shandong University, China); Jing Guo (The Hong Kong Polytechnic University); Mengbai Xiao, Dongxiao Yu and Xiuzhen Cheng (Shandong University, China)

Short-form video streaming has seen explosive growth in recent years. Typically, the short video apps (e.g., TikTok) are installed on mobile devices, with the video content delivered over wireless networks. Our measurement study shows that almost all the state-of-the-art adaptive bitrate (ABR) algorithms for short video streaming fall short of providing satisfactory Quality-of-Experience (QoE). This is due to their reliance on bandwidth-sensitive adaptation logic with over-discretized encoding bitrate levels, which hampers their ability to manage significant wireless bandwidth fluctuations. To tackle this problem, we propose DeepReal, a fine-grained ABR approach for short videos, the core of which is a novel clustering-augmented continuous control learning algorithm. On the one hand, the clustering offers more expertise for the ABR logic training. On the other hand, the continuous-valued action domain enables the trained ABR logic to support any number of discrete bitrate levels. As a result, DeepReal's bitrate decision is able to effectively adapt to the substantial network dynamics. Large-scale evaluations show that DeepReal improves QoE by 23.5% to 33.3% while maintaining strong spatial and temporal robustness. It offers a tailored and innovative solution for current short-form video services.

pp. 147-155

Tuesday, December 3 16:30 - 17:30

Session 3A: Session 3A: Quantum Networks and LoRa Enhancement

Topology Design with Resource Allocation and Entanglement Distribution for Quantum Networks

Jiyao Liu and Xuanzhang Liu (Temple University, USA); Xinliang Wei (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China); Yu Wang (Temple University, USA)

Topology is one of the most critical properties of networks. Quantum networks, as a new type of network, have fundamentally different principles for establishing connections compared to classical networks, leading to distinct challenges in topology design. Finding an optimal topology for quantum networks to meet traffic demands is a crucial yet not fully understood problem. In this paper, we explore the topology design problem for quantum networks, considering both resource allocation and entanglement distribution. We propose and investigate both flow-based and path-based formulations, along with their associated solutions, aimed at minimizing the topology cost. For the path-based formulation, we also provide the first theoretical analysis of the cost associated with swapping strategies over a quantum path. Extensive simulations demonstrate that our enhanced path-based formulation is both efficient and effective.

pp. 156-164

EGAL: Enhancing LoRa Network Lifetime with Load Balancing

Malaika Afra Taj, Kanav Sabharwal and Mun Choon Chan (National University of Singapore, Singapore)

LoRa networks have emerged as a promising solution for long-range, low-power communication in Internet of Things (IoT) applications. However, a significant challenge in LoRa networks is the uneven battery depletion among nodes, due to the varying transmission configurations required to support different distances between the nodes and the gateway. This disparity in battery consumption poses a substantial challenge to the overall network lifetime. This paper introduces \name, a relay-based approach designed to address the battery imbalance issue by ensuring balanced energy consumption across all nodes. Unlike purely multi-hop networks, \name follows a hybrid approach that selectively uses relays when necessary while maintaining direct connections to the gateway. \name employs a reward-based algorithm that dynamically adjusts relay duties and integrates predictive analytics, thereby enhancing network lifespan with minimal overhead. The proposed solution is validated through NS3 simulation and real-world testbed, demonstrating significant improvements in network lifetime. Specifically, in simulations, \name exhibits up to a 457% increase in network lifetime over standard LoRaWAN, and in hardware tests, it shows up to a 70% decrease in total energy consumption. The key contributions of this work include the development of the \name algorithm, its integration with standard LoRaWAN, and extensive validation of its effectiveness.

pp. 165-173

TransTuple: Toward Fast Packet Classification via Adaptive Tuple Replacement

Jiashuo Yu and Longlong Zhu (Fuzhou University & Quan Cheng Laboratory, China); Rongbang Wu (Fuzhou University, China); Linying Zheng and Hongyan Liu (Zhejiang University, China); Dong Zhang (Fuzhou University & Quan Cheng Laboratory, China); Chunming Wu (Zhejiang University, China)

Open vSwitch (OVS) is a widely used software switch in virtualized environments and software-defined networks. OVS uses tuple space search (TSS) for packet classification in the datapath, allowing fast network rule updates, but the increasing number of rules poses a classification performance challenge. To address this, existing methods incorporate decision trees with TSS to form a hybrid structure, enhancing classification speed. However, decision trees tend to overfit the initial ruleset, becoming unbalanced after rule updates and leading to a sharp decline in classification performance. In this paper, we propose TransTuple, a framework to optimize hybrid structures for fast packet classification under rule updates. The core idea of TransTuple is to identify bottleneck branches in decision trees that degrade performance and to replace them with lightweight tuples, providing better throughput under rule updates. These tuples maintain rules using hash tables, enabling fast updating and packet matching on bottleneck branches. We use TransTuple to optimize three state-of-the-art hybrid structured methods, i.e., CutTSS, TabTree, and MBitTree, achieving up to a 3.1x improvement in classification speed during rule updates.

pp. 174-182

JUNCTION: A Scalable Multi-access Solution using Programmable Switches

Xin Zhe Khooi, Cha Hwan Song and Satis Kumar Permal (National University of Singapore, Singapore); Nishant Budhdev (Nokia Bell Labs, Belgium); Levente Csikor (Institute for Infocomm Research & Agency for Science Technology and Research, Singapore); Raj Joshi (Harvard University, USA); Mun Choon Chan (National University of Singapore, Singapore)

Multi-access networks are increasingly important for reliable end-to-end connectivity and enhanced throughput performance. A scalable multi-access solution is required to roll out multi-access networks at scale. However, existing CPU-based solutions can no longer scale sustainably, as network traffic has outgrown the CPU performance growth. Consequently, hardware accelerators offer a compelling alternative. This paper introduces JUNCTION, a scalable multi-access solution designed using programmable switches. JUNCTION features a multipath protocol tailored to the hardware constraints and optimized for efficient memory utilization, enabling it to handle a large number of multipath sessions. We validate JUNCTION on a 5G-WiFi multi-access testbed. Our analysis demonstrates that it can scale an

order of magnitude better than existing solutions.

pp. 183-191

Wednesday, December 4

Wednesday, December 4 10:30 - 11:30

Session 3B: Session 3B: Privacy, Security, and Blockchain

Charge Me Securely: Decentralized Privacy-Aware and Publicly Verifiable Energy Trading with Electric Vehicles

Jucai Yang and Haiqin Wu (East China Normal University, China); Xiao Chen (University of Leicester, United Kingdom (Great Britain)); Zhenfu Cao and Xiaolei Dong (East China Normal University, China)

With the widespread adoption of electric vehicles (EVs), Vehicle-to-Vehicle (V2V) charging technology offers a more flexible solution for EV charging, significantly alleviating charging inconvenience, particularly in remote areas. However, the current V2V energy trading landscape lacks reliable management platforms with transparent transaction protocols. Furthermore, notable deficiencies exist in data privacy and security, which hinder the broader implementation of V2V charging services. Addressing these challenges, particularly the provision of secure and publicly auditable V2V charging without relying on centralized platforms or disclosing user privacy, has become a critical concern. For this reason, this paper presents PET, a decentralized Privacy-preserving Energy Trading system with public verifiability for EVs. PET is built upon the emerging blockchain technology to decentralize energy trading while ensuring accountability. We model energy demand-supply matching as range matching of locations and charge amounts between a buyer and sellers, and employ reverse auction to select the winner. For efficient and privacy-aware range proofs, we propose a novel batched hash chain-based range proof (BHW) primitive. In addition, PET integrates zk-SNARKs to verify payment correctness while maintaining user privacy. Our system supports public verifiability, entitling any third party to independently verify the transaction integrity. We analyze the privacy guarantees and public verifiability of PET. Extensive experiments implemented on Hyperledger Fabric further validate that PET delivers robust performance with a 4× reduction in verification cost compared to that without batch proofs.

pp. 192-200

Publicly Verify While Hiding Data: Privacy-Aware Streaming Truth Discovery with Public Verifiability in Blockchain-Enhanced Crowdsensing

Ruikai Zheng and Haiqin Wu (East China Normal University, China); Boris Düdder (University of Copenhagen, Denmark)

In crowdsensing, truth discovery (TD) has been extensively applied to resolve the data conflicts among publicly recruited workers and provide more reliable task data truths for task requesters. Data privacy and computation integrity are two major security concerns in TD due to the inherent untrustworthiness of the crowdsensing platform. Previous resolutions mostly focus on the design of privacy-preserving TD schemes while those leveraging blockchain to ensure TD integrity unfortunately sacrifice privacy or endure expensive on-chain overhead. In the presence of any misbehaviors, making TD integrity publicly and efficiently verifiable without compromising data privacy is imperative. This paper proposes P 2STD, a Privacy-preserving and Publicly verifiable scheme for generic Streaming TD (STD) in blockchain-enhanced crowdsensing. Unlike traditional TD in centralized crowdsensing, P 2STD works in a decentralized architecture empowered by blockchain and fog, in which STD is jointly performed by fog nodes while the proof information is anchored to the blockchain for public verification. We take streaming confident-aware TD (S-CATD), the latest STD, as an instance and design secure aggregation protocols based on verifiable additive homomorphic secret sharing, concealing both task data and truth and generating publicly verifiable information. P 2STD is not specific to S-CATD and can be adapted to other iterative STD algorithms. Security analysis proves the desired privacy and public verifiability achieved in P 2STD. Experimental evaluations demonstrate that P 2STD has high accuracy, lower

computational and communication overhead, offering public verifiability with additional subtle gas cost than others.

pp. 201-209

Payment Routing across IoT Blockchain Shards using Deep Reinforcement Learning

Ting Cai (Hubei University of Technology, China & Sun Yat-Sen University, China); Chuqi Li, Qiang Li and Yifei Wang (Hubei University of Technology, China); Yuxin Wu (Hubei University of Arts and Science, China); Zhiwei Ye (Hubei University of Technology, Japan); Patrick C. K. Hung (Ontario Tech Technology, Canada)

Payment channel network (PCN) routing is crucial for ensuring a high system throughput and transaction (TX) success ratio. However, designing PCN routing across IoT blockchain shards is not straightforward. First, the sharded architecture isolates information among multiple shards, increasing difficulties and costs associated with channel probing. Second, the dynamic nature of IoT, e.g., frequent changes in PCN topology and channel states, makes traditional mathematical routing approaches inefficient and even invalid. These challenges inevitably result in a high TX failure ratio and low throughput. This paper presents a novel cross-shard PCN routing IoT blockchain framework that optimizes TX scheduling to maximize long-term throughput and success ratio. Specifically, we propose an efficient cross-shard PCN routing protocol that distinctly categorizes channels into intra-shard and inter-shard types, simplifying cross-shard channel probing and reducing costs. Then, to optimize PCN routing policies in dynamic sharding environments, we propose a deep reinforcement learning algorithm, which includes: 1) multi-agent collaborative learning for the view of incomplete information across shards; 2) a two-layer network architecture to reduce computational complexity on resource-constrained IoT devices. Experiment results show that the proposed cross-shard routing improves 48.4% of the TX success ratio averaged over baselines, which is 1.56 times of average throughput compared with other routing algorithms in cross-shard PCNs.

pp. 210-218

UniQR: A Secure QR Code Payment Scheme Using Device Pose and Environmental Matching

Jingwen Wei, Lupeng Zhang, Jingchi Zhang, Bin Han and Lei Wang (Dalian University of Technology, China)

The convenience of QR codes has made them increasingly popular in the field of mobile payments, with many payment service providers (e.g., PayPal, Alipay, WeChat) offering QR code payment services. However, due to the openness of QR code scanning process, attackers can capture the code and perform fraudulent transactions before the original QR code is used, posing a significant security threat. In this paper, we design UniQR (Unique QR), which embeds device pose information and environmental information as "physically unclonable" fingerprints into the QR code to prevent attackers from unauthorized use of it. We propose a pose matching method based on Perspective-n-Point (PnP) algorithm and the integrated sensors of the phone to bind the QR code to the payment device. We introduce ambient light information to address the instability of geomagnetism and bind the QR code to the payment environment. This dual information binding effectively enhances the security of QR code payments. Additionally, we modify the QR code encoding using a segmented hybrid encoding method, allowing secure authentication information to be embedded by utilizing only a portion of the space originally designated for dummy data. We implemented UniQR on six different commercial phones and conducted experiments with 23 participants simulating both legitimate and illegitimate payment scenarios. The 97.46% success rate in legitimate user scans demonstrates the feasibility and robustness of UniQR.

pp. 219-227

Wednesday, December 4 11:30 - 12:30

Session 4B: Session 4B: Autonomous Vehicles and Image Processing

ECSeg: Edge-Cloud Switched Image Segmentation for Autonomous Vehicles

Siyuan Zhou, Duc Van Le and Rui Tan (Nanyang Technological University, Singapore)

Existing autonomous vehicles have not utilized the cloud computing for execution of their deep learning-based driving tasks due to the long vehicle-to-cloud communication latency. The vehicles are in general equipped with the resource-constrained edge computing devices which may be unable to execute the compute-intensive deep learning models in real time. The increasing data transmission speed of the commercial mobile networks sheds light upon the feasibility of using the cloud computing for autonomous driving. Our city-scale real-world measurements show that the vehicles can partially use the cloud computing via the fifth generation (5G) mobile network with the low data transmission latency. In this paper, we present the design and implementation of ECSeg, an edge-cloud switched image segmentation system that dynamically switches between the edge and cloud in which the deep learning-based semantic segmentation models are executed for understanding the vehicle's visual scenes in real time. The switching decision-making is challenging due to the intricate interdependencies among various factors including the dynamic wireless channel condition, vehicle's movement and visual scene change. To this end, we employ deep reinforcement learning to learn an optimal switching policy. Extensive evaluation based on both real-world experiments and trace-driven simulations demonstrates that ECSeg achieves superior image segmentation accuracy for autonomous vehicles, compared with four baseline approaches.

pp. 228-236

KeyCoop: Communication-Efficient Raw-Level Cooperative Perception for Connected Autonomous Vehicles via Keypoints Extraction

Qi Xie, Xiaobo Zhou, Chuanan Wang, Tie Qiu and Wenyu Qu (Tianjin University, China)

Cooperative perception is an emerging paradigm that expects to conquer the sensory limitations of individual vehicles by sharing sensor information with each other and significantly improve driving safety. However, achieving highly precise data sharing and low communication overhead remains a challenge for cooperative perception, especially when real-time communication is necessary in autonomous driving. As a result, it is essential to decrease the transmitted sensor data while maintaining the perception performance. For this purpose, we propose a communication-efficient raw-level cooperative perception system for connected autonomous vehicles (CAVs), which is able to significantly compress the raw sensor data each CAV shares with each other by only transmitting the most informative keypoints. Specifically, at the local level, a voxel-based instance-ware keypoints selection strategy is proposed to select the points that belong to regions of interest. To further supervise the local keypoints selection, we present a collaborative global-local learning strategy, enabling each vehicle to consider both the local scenario and the global context when selecting the transmitted data. Comprehensive evaluations indicate the superiority of the proposed system, which achieves more than 300 times lower communication volume compared to the raw data, with a performance degradation of less than 1 %.

pp. 237-245

Cascade Reinforcement Learning with State Space Factorization for O-RAN-based Traffic Steering

Chuanneng Sun (Rutgers University, USA); Yu Zhou, Gueyoung Jung and Xuan Tuyen Tran (ATT Labs, USA); Dario Pompili (Rutgers University, USA)

We study the traffic steering (TS) problem in open radio access network (O-RAN), leveraging its RAN intelligent controller (RIC), in which RAN configuration parameters of cells can be jointly and dynamically optimized in near-real-time. To address the TS problem, we propose a novel Cascade Reinforcement Learning (CaRL) framework, where we propose state space factorization and policy decomposition to mitigate the need for large complex models and well-labeled datasets. For each sub-state space, an RL sub-policy is trained to optimize the Quality of Service (QoS). To apply CaRL to new network areas, we propose a knowledge transfer approach to initialize a new sub-policy based on knowledge learned by the trained policies. To evaluate CaRL, we build a data-driven and scalable RIC Digital Twin (DT) that is modeled using real-world data, including network setup, user geo-distribution, and traffic demand, among others, from a tier-1 network operator in the US. We evaluated CaRL in two DT scenarios representing two different US cities and compared its performance with business-as-usual policy as a baseline and other competing optimization

approaches (i.e., heuristic and Q-table algorithms). Furthermore, we have conducted a field trial with the tier-1 operator to evaluate the performance of CaRL in two areas in the Northeast US regions.

pp. 246-254

Resource Allocation and Task Offloading for Slicing-Based Communication and Computing in Space-Air-Ground Integrated Networks

Yueqian Song and Guiyuan Yuan (Shandong University of Science and Technology, China); Qingtian Zeng (China); Geng Chen (Shandong University of Science and Technology, China)

In the upcoming 6G era, the demand for network communication and computing is expected to surge and diversify. A Space-Air-Ground Integrated Network (SAGIN) is introduced as a solution to provide seamless global connectivity. Meanwhile, network slicing technology can further enhance network capabilities by supporting customized services. However, the joint optimization of multi-dimensional resource allocation and task offloading decisions presents a significant challenge in dynamic and complex environments with diverse task types. In this work, we establish an SAGIN paradigm that integrates network slicing. A multi-level optimization scheme for resource allocation and task offloading is proposed to improve Key Performance Indicators (KPIs) and Quality of Service (QoS). Specifically, we propose a State Encoding-based Multi-Agent Soft Actor-Critic algorithm (SE-MASAC) within a Centralized Training with Decentralized Execution architecture (CTDE). This algorithm processes current sensing states and historical information, leveraging reinforcement learning to make inter-slice resource and offloading scheduling decisions. Intra-slice decisions for intelligent User Equipments (iUEs) are determined based on their sensing data-driven priority. Simulation results demonstrate that our approach outperforms baselines in terms of system utility, QoS satisfaction, and various KPIs.

pp. 255-263

An Intelligent Prefetch Strategy with Multi-round Cell Enhancement in Volumetric Video Streaming

Shuquan Liu, Guanghui Zhang, Mengbai Xiao, Dongxiao Yu and Xiuzhen Cheng (Shandong University, China)

Over the past few years, volumetric video streaming, a cutting-edge application in virtual reality (VR) and augmented reality (AR), has gained significant attention. However, its enormous bandwidth demands exceed the capabilities of current networks to support full-size transmission. As a result, streaming vendors in the industry typically use field-of-view (FoV) prediction to reduce the streaming video size. While this approach makes transmission feasible, our measurements have shown that the adopted sequential video prefetch mechanism significantly hinders streaming performance. To address this challenge, we propose PACE, an intelligent video prefetch strategy that leverages multi-round cell-enhanced downloading. PACE divides the prefetching process for each group-of-frame (GoF) into multiple rounds, downloading different cells in each round based on periodically updated FoV prediction results. This method allows the FoV prediction to operate independently of the prefetch length limitations. Additionally, PACE incorporates a greedy-policy-based video quality decision algorithm to fully utilize network bandwidth and maximize the quality of experience (QoE). Extensive evaluations demonstrate that PACE can enhance FoV prediction accuracy by up to 54.8% and improve QoE performance by up to 23.1% while showing strong robustness across a wide range of streaming environments.

pp. 264-272

Wednesday, December 4 14:00 - 15:00

Session 4A: Session 4A: Network Metadata, Sensor Fusion, and Flow Management

DeepMEND: Reliable and Scalable Network Metadata Geolocation from Base Station Positions

Orlando E. Martínez-Durive (IMDEA Networks Institute, Spain & Universidad Carlos III de Madrid, Spain); Stefanos

Sotirios Bakirtzis (University of Cambridge, United Kingdom (Great Britain)); Cezary Ziemlicki (Orange labs, France); Jie Zhang (University of Sheffield, Dept. of Electronic and Electrical Engineering, United Kingdom (Great Britain)); Ian James Wassell (University of Cambridge, United Kingdom (Great Britain)); Marco Fiore (IMDEA Networks Institute, Spain)

Metadata geolocation, i.e., the mapping of information collected at a cellular Base Station (BS) to the geographical area it covers, is a central operation in pipelines for the production of statistics from mobile network measurements, which are in turn relevant to a wide range studies across many scientific disciplines. This task requires modeling the probability that a mobile device associated with the BS is at a specific location and is presently addressed with simplistic approximations based on Voronoi tessellations. As we show in our work, Voronoi cells exhibit poor accuracy when compared to real-world geolocation data, which can, in turn, curb the reliability of research results. We propose DeepMEND, a new approach for data-driven metadata geolocation that is based on a teacher-student paradigm and combines probabilistic inference and deep learning. Our trained model is: (🛜) practical, as it solely relies on BS positions exactly like Voronoi tessellations; (🛜🛜) dependable, as it produces geolocation information that is 56% and 33% more accurate than that of legacy Voronoi and of their state-of-the-art VoronoiBoost calibration, respectively; and, (🛜🛜🛜) scalable, as it generates geolocation information for thousands of BSs in minutes. DeepMEND lets any researcher immediately and substantially repair the spatial mapping of mobile network metadata, hence its name. We assess its accuracy against real-world multi-city geolocation data of 5,947 BSs provided by a network operator and demonstrate the impact of its enhanced metadata geolocation on two practical applications.

pp. 273-281

PCI Classification in 5G NR: Deep Learning Unravels Synchronization Signal Blocks

Md Rabiul Hossain (The University of Arizona, USA); Marwan Krunz (University of Arizona, USA)

Accurate detection of the Physical Cell Identity (PCI) is critical for rapid synchronization and connection establishment in 5G New Radio (5G-NR) systems. This paper introduces a deep learning-based approach for PCI classification, aiming to mitigate the computational complexity associated with traditional methods that rely on decoding the Synchronization Signal Block (SSB). Our approach processes only time-domain baseband samples of the downlink signal, arranged in fixed-length windows. These windows are inputted into a pre-trained Convolutional Neural Network (CNN), which classifies the samples into one of several known PCI values (representing nearby cells) or into an ‘other' category (representing all non-nearby cells as well as windows that do not contain SSB samples). Because PCI-related information is contained only in the SSB symbols of a frame, it is possible for an input window to include no or a few SSB samples. Accordingly, in labeling the training set, we use a threshold Ttrain on the fraction of the samples within a window: if Ttrain or more of the samples belong to an SSB of cell with a target PCI value, the true label for that window is set to that PCI value; otherwise, it is set to ‘other.' A separate threshold Ttest is used for labeling the test windows. We also study another labeling mechanism whereby only samples of the third OFDM symbol in an SSB (which contains the Secondary Synchronization Signal) are used to determine the label. Our analysis considers two commonly used SSB formats that correspond to 15 and 30 kHz subcarrier spacings, respectively. Extensive simulations are conducted, which reveal that the proposed classifier can reliably (above 98%) identify the PCI value of a captured signal even under Signal-to-Noise Ratio (SNR) values as low as −10 dB. This performance comes with a significant reduction in computational complexity as it bypasses the need for traditional SSB decoding procedures used for PCI estimation in 5G networks.

pp. 282-290

Resource-Efficient Sensor Fusion via System-Wide Dynamic Gated Neural Networks

Chetna Singhal (INRIA France, France); Yashuo Wu (University of California Irvine, USA); Francesco Malandrino (CNR-IEIIT, Italy); Sharon Ladron de Guevara Contreras and Marco Levorato (University of California, Irvine, USA); Carla Fabiana Chiasserini (Politecnico di Torino & CNIT, IEIIT-CNR, Italy)

Next-generation mobile systems will support multiple AI-based applications, each leveraging heterogeneous sensors and data sources through DNN architectures collaboratively executed within the network. In this context, to minimize

the cost of the AI inference task subject to requirements on latency, quality, and -- crucially -- reliability of the inference process, it is vital to optimize (i) the set of sensors/data sources and (ii) the DNN architecture, (iii) the network nodes executing sections of the DNN, and (iv) the resources to use. To achieve these goals, we leverage dynamic gated neural networks with branches, and propose a novel algorithmic strategy called Quantile-constrained Inference (QIC), based upon quantile-Constrained policy optimization. QIC makes joint, high-quality, swift decisions on all the above aspects of the system, with the aim to minimize inference energy cost. We remark that this is the first contribution connecting gated dynamic DNNs with infrastructure-level decision making. We evaluate QIC using a dynamic gated DNN with stems and branches for optimal sensor fusion and inference, trained on the RADIATE dataset offering Radar, LiDAR, and Camera data, and real-world wireless measurements. Our results confirm that QIC outperforms existing approaches by over 80%, and closely matches the optimum.

pp. 291-299

A Large-Scale P2P Botnet Detection Framework via Topology and Traffic Co-Verification

Ziming Zhao (Zhejiang University, China); Zhaoxuan Li (Institute of Information Engineering Chinese Academy of Sciences, China); Tingting Li and Fan Zhang (Zhejiang University, China)

Botnets are still serious threats to infrastructure security nowadays. Recently, adversaries tend to leverage peer-to-peer (P2P) manner propagation to construct large-scale botnets since P2P-based schemes have no single points of failure. Over the past few decades, the research and industry communities have proposed a variety of solutions to detect botnets, which mainly involve communication topology identification and network traffic analysis. Yet, coping with the large-scale P2P botnets, the former suffer topology indistinguishability, and the latter struggles under massive background traffic. In this paper, we present TNT, a large-scale P2P botnet detection framework via communication topology and network traffic. As its core, TNT is powered by three tightly-coupled components: (i) tScouter is responsible for profiling the communication topology; (ii) tCommander plans the strategy for node inspection; and (iii) tPatroller investigates the traffic of the corresponding node. Taken together, TNT advances the trade-off between detection accuracy (enhance topology-based results via traffic analysis) and overhead (only check part of node traffic according to the planning). Based on 42 groups of combinations involving 6 types of botnets and 7 legitimate P2P traffic, we perform extensive evaluation and demonstrate that TNT realizes outstanding detection performance, e.g., after checking ~20K nodes, achieve ~99.9% accuracy for a communication graph (including >140K nodes).

pp. 300-308

Wednesday, December 4 15:00 - 15:30

Posters and Demos 2: Posters and Demos 2

Wednesday, December 4 15:30 - 17:00

Session 5A: Session 5A: Indoor Tracking, Localization, and Sensing

SaTrack: LoS/NLoS State-Aware WiFi Indoor Tracking System

Yinnan Zhou, Hao Zhou, Luowei Li and Siming Chen (University of Science and Technology of China, China); Zhi Liu (The University of Electro-Communications, Japan); Xiang-Yang Li (University of Science and Technology of China, China)

WiFi-based technology is appealing for indoor localization due to the widely deployed infrastructures. Recently, path separation solutions have been proposed to address the multipath effects and achieve decimeter-level localization accuracy in line-of-sight (LoS) scenarios. However, these solutions experience serious performance degradation in non-line-of-sight (NLoS) scenarios, and couldn't be used for mobile device tracking where continuous LoS/NLoS switching happens. In this paper, we propose SaTrack, a LoS/NLoS state-aware mobile device tracking system. SaTrack identifies

LoS/NLoS states based on the diversity of the strongest estimated paths when using different reference antennas. With the observation of spatial aggregation and temporal continuity for the Tx-Rx direct path, SaTrack chooses the direct path through two-step clustering, i.e., clustering in the spatial domain to form candidates and clustering again in the temporal domain to select the winner. Extensive experiments are conducted to evaluate the effectiveness of SaTrack. In a typical indoor environment with abundant multipath, SaTrack achieves 0.64m and 1.27m for the median and 90th percentile tracking errors, outperforming the state-of-the-art (SOTA) solutions.

pp. 309-317

USee: Ultrasound-based Device-free Eye Movement Sensing

Wen Cheng and Mingzhi Pang (Nanjing University, China); Haoran Wan (Princeton University, USA); Shichen Dong, Dongxu Liu and Wei Wang (Nanjing University, China)

Detecting eye movements has become a hot topic in human-computer interaction and serves as an overall health indicator, making it both appealing and challenging. In this paper, we present a system named USee that achieves high-precision capture of weak and aperiodic eye movements by utilizing fine-grained and ubiquitous ultrasound signals, capturing both blinking and more subtle saccades. We initially identify signal changes associated with eye movements by leveraging the distinctive impact of blinking on signals. Pioneeringly, we reveal the relationship between the residuals of signal decomposition and subtle eye movements. Utilizing innovative signal processing architectures, we mitigate interference and extract eye movement features. Subsequently, we employ one-dimensional convolutional operations in place of signal cross-correlation, designing filters for motion category identification and a lightweight convolutional neural network for saccade direction classification. This capability allows our system to serve as a foundational sensing layer for eye movements, making it applicable to various applications. We implement USee on both a research-purpose platform and a commodity Raspberry Pi. Extensive experimental results validate the effectiveness of our system, demonstrating a saccade recognition accuracy of 91% and a blinking recognition accuracy of 94%. The system exhibits robustness, persisting even in challenging scenarios with strong interference, such as the presence of a moving pedestrian.

pp. 318-326

GAN-Loc: Empowering Indoor Localization for Unknown Areas via Generative Fingerprint Map

JinYi Yoon (Virginia Tech, USA); Yeawon You and Dayeon Kang (Ewha Womans University, Korea (South)); Jeewoon Kim (University of California, Irvine, USA); HyungJune Lee (Ewha Womans University, Korea (South))

As indoor localization becomes a necessity to provide intelligent location-based services for in-building users, fingerprint-based positioning has been widely adopted in numerous Wi-Fi-equipped devices. However, its reliance on extensive prior site surveys limits its application in unexplored environments without prior fingerprint sampling. To address the challenge, we propose GAN-Loc, a novel generative fingerprint-based framework. We extract the underlying correlation between location and radio signal features, empowering indoor localization in unexplored or under-explored areas, including unknown data points, newly deployed APs, or unexplored sites. It involves: 1) decomposing into a signal feature map for each AP; 2) learning with a set of points and their associated signal strength data; 3) generating and integrating synthetic radio fingerprints; and 4) employing them into some existing localization algorithms. We evaluated GAN-Loc with extensive real-world RSSI experiments in seven different indoor places. GAN-Loc achieves the localization accuracy of up to 1.21m, 1.29m, and 1.28m for Wi-Fi, Zigbee, and BLE, respectively, compared to the accuracy of 1.47m, 1.58m, and 1.22m using an ideal ground-truth map, which is unachievable without site survey in unknown sites. \emph{GAN-Loc} also reduces the distance error with 34.2% at the same place by augmenting the unrevealed fingerprints.

pp. 327-335

LoDiHAR: A Low-Cost Distributed Human Activity Recognition System Based on RFID

Xiaoqi Sun, Yanwen Wang, Chenwei Zhang, Zheng Wang and Xiaokang Shi (Hunan University, China); Yuanqing

Zheng (The Hong Kong Polytechnic University, Hong Kong)
Human Activity Recognition has been extensively applied to fulfill tasks such as fall detection, human-computer interaction, virtual reality, etc. Existing radio frequency-based HAR methods, although overcoming limitations of wearable-, visual-, and acoustic-based sensing technology, still suffer from high costs and low efficiency, which limits their pervasive use. In this paper, we propose LoDiHAR, a low-cost, distributed HAR system leveraging Radio Frequency Identification technology. LoDiHAR employs low-cost and fully programmable commercial wireless components, providing full access to the PHY samples of the backscattered signals, in which signal phases can be extracted to infer different activities. Different from COTS RFID systems that adopt a polling interrogation scheme, LoDiHAR supports a distributed sensing scheme, which profiles human activities more efficiently. LoDiHAR addresses a series of technical challenges such as accurate phase extraction from backscattered signals, asynchronous distributed RF data fusion and insufficient training data. A Conditional Generative Adversarial Network framework combined with a Transformer model is designed for accurate time-series activity classification. LoDiHAR demonstrates proficiency in recognizing eight types of human activities across diverse environments, achieving an accuracy of up to 94.9% while only costing 10% of the mainstream COTS RFID systems.

pp. 336-344

OISMic: Acoustic Eavesdropping Exploiting Sound Induced OIS Vibrations in Smartphones

Ziyu Shao, Yuchen Su, Yicong Du, Shiyue Huang and Tingyuan Yang (University of Electronic Science and Technology of China, China); Hongbo Liu (Electronic Science and Technology of China, China); Yanzhi Ren (University of Electronic Science and Technology of China, China); Bo Liu (Amazon, USA); Huan Dai (Suzhou University of Science and Technology, China); Shuai Li (University of Oulu, Finland)
Optical image stabilization (OIS), powered by a special micro-electromechanical structure in the camera lenses to compensate for the optical distortion caused by camera shakes, has become an indispensable feature in many smartphones. However, we discover that this seemingly benign component can be exploited to eavesdrop on nearby audio signals, posing a significant threat to people's privacy during conversations or phone calls. Specifically, the OIS component can be influenced by external acoustic stimuli leading to slight vibrations, and at the same time, the coil and magnetized components inside the OIS induce electromagnetic leakage as they vibrate, according to Faraday's Law of Electromagnetic Induction. This electromagnetic leakage contains voice information that can be used to recover the audio signals if intercepted by individuals with malicious intent. Inspired by the above discovery, we propose OISMic, a new acoustic eavesdropping attack that takes advantage of sound-induced OIS vibrations on smartphones. Unlike other existing acoustic eavesdropping attacks, eavesdropping exploiting OIS vibrations not only overcomes the constraints imposed by system permissions for many sensor-based approaches but is also immune to ultrasonic jammer that hinders the methods relying on microwave or light reflections to sense sound-induced vibrations. To execute this non-trivial attack in practical scenarios, we developed a prototype circuit that has a compact design capable of capturing the electromagnetic leakage caused by OIS vibrations. After converting the collected leaked electromagnetic signals into audio signals, a software-based phase-locked loop (PLL) method is developed to enhance the representation of voice components. Meanwhile, to reconstruct the weak audio signals, we also designed a diffusion-based neural network to learn the distribution of electromagnetic noise within the audio spectrum. Extensive experiments indicate that OISMic can accurately reconstruct voice under various scenarios, achieving an average word correct rate of 90.57% across different devices.

pp. 345-353