Wed 2024.04.17

Egret: Reinforcement Mechanism for Sequential Computation Offloading in Edge Computing

Haosong Peng, Yufeng Zhan, DiHua Zhai, Xiaopu Zhang, Yuanqing Xia

As an emerging computing paradigm, edge computing offers computing resources closer to the data sources, helping to improve the service quality of many real-time applications. A crucial problem is designing a rational pricing mechanism to maximize the revenue of the edge computing service provider (ECSP). However, prior works have considerable limitations: clients are static and are required to disclose their preferences, which is impractical in reality. However, previous works assume user privacy information to be known or consider the number of users in edge scenarios to be static. To address this issue, we propose a novel sequential computation offloading mechanism, where the ECSP posts prices of computing resources with different configurations to clients in turn. Clients independently choose which computing resources to purchase and how to offload based on their prices. Then Egret, a deep reinforcement learning-based approach that achieves maximum revenue, is proposed. Egret determines the optimal price and visiting orders online without considering clients' preferences. Experimental results show that the revenue of ECSP in Egret is only 1.29\% lower than Oracle and 23.43\% better than the state-of-the-art when the client arrives dynamically.

link: http://arxiv.org/abs/2404.09285v1

Artificial Intelligence enhanced Security Problems in Real-Time Scenario using Blowfish Algorithm

Yuvaraju Chinnam, Bosubabu Sambana

In a nutshell, "the cloud" refers to a collection of interconnected computing resources made possible by an extensive, real-time communication network like the internet. Because of its potential to reduce processing costs, the emerging paradigm of cloud computing has recently attracted a large number of academics. The exponential expansion of cloud computing has made the rapid expansion of cloud services very remarkable. Ensuring the security of personal information in today's interconnected world is no easy task. These days, security is really crucial. Models of security that are relevant to cloud computing include confidentiality, authenticity, accessibility, data integrity, and recovery. Using the Hybrid Encryption this study, we cover all the security issues and leaks in cloud infrastructure.

link: http://arxiv.org/abs/2404.09286v1

RoofDiffusion: Constructing Roofs from Severely Corrupted Point Data via Diffusion

Kyle Shih-Huang Lo, Jörg Peters, Eric Spellman

Accurate completion and denoising of roof height maps are crucial to reconstructing high-quality 3D buildings. Repairing sparse points can enhance low-cost sensor use and reduce UAV flight overlap. RoofDiffusion is a new end-to-end self-supervised diffusion technique for robustly completing, in particular difficult, roof height maps. RoofDiffusion leverages widely-available curated footprints and can so handle up to 99\% point sparsity and 80\% roof area occlusion (regional incompleteness). A variant, No-FP RoofDiffusion, simultaneously predicts building footprints and heights. Both quantitatively outperform state-of-the-art unguided depth completion and representative inpainting methods for Digital Elevation Models (DEM), on both a roof-specific benchmark and the BuildingNet dataset. Qualitative assessments show the effectiveness of RoofDiffusion for datasets with real-world scans including AHN3, Dales3D, and USGS 3DEP LiDAR. Tested with the leading City3D algorithm, preprocessing height maps with RoofDiffusion noticeably improves 3D building reconstruction. RoofDiffusion is complemented by a new dataset of 13k complex roof geometries, focusing on long-tail issues in remote sensing; a novel simulation of tree occlusion; and a wide variety of large-area roof cut-outs for data augmentation and benchmarking.

link: http://arxiv.org/abs/2404.09290v1

Bridging Data Islands: Geographic Heterogeneity-Aware Federated Learning for Collaborative Remote Sensing Semantic Segmentation

Jieyi Tan, Yansheng Li, Sergey A. Bartalev, Bo Dang, Wei Chen, Yongjun Zhang, Liangqi Yuan

Remote sensing semantic segmentation (RSS) is an essential task in Earth Observation missions. Due to data privacy concerns, high-quality remote sensing images with annotations cannot be well shared among institutions, making it difficult to fully utilize RSS data to train a generalized model. Federated Learning (FL), a privacy-preserving collaborative learning technology, is a potential solution. However, the current research on how to effectively apply FL in RSS is still scarce and requires further investigation. Remote sensing images in various institutions often exhibit strong geographical heterogeneity. More specifically, it is reflected in terms of class-distribution heterogeneity and object-appearance heterogeneity. Unfortunately, most existing FL studies show inadequate focus on geographical heterogeneity, thus leading to performance degradation in the global model. Considering the aforementioned issues, we propose a novel Geographic Heterogeneity-Aware Federated Learning (GeoFed) framework to address privacy-preserving RSS. Through Global Feature Extension and Tail Regeneration modules, class-distribution heterogeneity is alleviated. Additionally, we design an Essential Feature Mining strategy to alleviate object-appearance heterogeneity by constructing essential features. Extensive experiments on three datasets (i.e., FBP, CASID, Inria) show that our GeoFed consistently outperforms the current state-of-the-art methods. The code will be available publicly.

link: http://arxiv.org/abs/2404.09292v1

A Novel State Space Model with Local Enhancement and State Sharing for Image Fusion

Zihan Cao, Xiao Wu, Liang-Jian Deng, Yu Zhong

In image fusion tasks, images from different sources possess distinct characteristics. This has driven the development of numerous methods to explore better ways of fusing them while preserving their respective characteristics. Mamba, as a state space model, has emerged in the field of natural language processing. Recently, many studies have attempted to extend Mamba to vision tasks. However, due to the nature of images different from casual language sequences, the limited state capacity of Mamba weakens its ability to model image information. Additionally, the sequence modeling ability of Mamba is only capable of spatial information and cannot effectively capture the rich spectral information in images. Motivated by these challenges, we customize and improve the vision Mamba network designed for the image fusion task. Specifically, we propose the local-enhanced vision Mamba block, dubbed as LEVM. The LEVM block can improve local information perception of the network and simultaneously learn local and global spatial information. Furthermore, we propose the state sharing technique to enhance spatial details and integrate spatial and spectral information. Finally, the overall network is a multi-scale structure based on vision Mamba, called LE-Mamba. Extensive experiments show the proposed methods achieve state-of-the-art results on multispectral pansharpening and multispectral and hyperspectral image fusion datasets, and demonstrate the effectiveness of the proposed approach. Code will be made available.

link: http://arxiv.org/abs/2404.09293v1

Cross-Data Knowledge Graph Construction for LLM-enabled Educational Question-Answering System: A~Case~Study~at~HCMUT

Tuan Bui, Oanh Tran, Phuong Nguyen, Bao Ho, Long Nguyen, Thang Bui, Tho Quan

In today's rapidly evolving landscape of Artificial Intelligence, large language models (LLMs) have emerged as a vibrant research topic. LLMs find applications in various fields and contribute significantly. Despite their powerful language capabilities, similar to pre-trained language models (PLMs), LLMs still face challenges in remembering events, incorporating new information, and addressing domain-specific issues or hallucinations. To overcome these limitations, researchers

have proposed Retrieval-Augmented Generation (RAG) techniques, some others have proposed the integration of LLMs with Knowledge Graphs (KGs) to provide factual context, thereby improving performance and delivering more accurate feedback to user queries. Education plays a crucial role in human development and progress. With the technology transformation, traditional education is being replaced by digital or blended education. Therefore, educational data in the digital environment is increasing day by day. Data in higher education institutions are diverse, comprising various sources such as unstructured/structured text, relational databases, web/app-based API access, etc. Constructing a Knowledge Graph from these cross-data sources is not a simple task. This article proposes a method for automatically constructing a Knowledge Graph from multiple data sources and discusses some initial applications (experimental trials) of KG in conjunction with LLMs for question-answering tasks.

link: http://arxiv.org/abs/2404.09296v1

Reap the Wild Wind: Detecting Media Storms in Large-Scale News Corpora

Dror K. Markus, Effi Levi, Tamir Sheafer, Shaul R. Shenhav

Media Storms, dramatic outbursts of attention to a story, are central components of media dynamics and the attention landscape. Despite their significance, there has been little systematic and empirical research on this concept due to issues of measurement and operationalization. We introduce an iterative human-in-the-loop method to identify media storms in a large-scale corpus of news articles. The text is first transformed into signals of dispersion based on several textual characteristics. In each iteration, we apply unsupervised anomaly detection to these signals; each anomaly is then validated by an expert to confirm the presence of a storm, and those results are then used to tune the anomaly detection in the next iteration. We demonstrate the applicability of this method in two scenarios: first, supplementing an initial list of media storms within a specific time frame; and second, detecting media storms in new time periods. We make available a media storm dataset compiled using both scenarios. Both the method and dataset offer the basis for comprehensive empirical research into the concept of media storms, including characterizing them and predicting their outbursts and durations, in mainstream media or social media platforms.

link: http://arxiv.org/abs/2404.09299v1

A Simple Strategy for Body Estimation from Partial-View Images

Yafei Mao, Xuelu Li, Brandon Smith, Jinjin Li, Raja Bala

Virtual try-on and product personalization have become increasingly important in modern online shopping, highlighting the need for accurate body measurement estimation. Although previous research has advanced in estimating 3D body shapes from RGB images, the task is inherently ambiguous as the observed scale of human subjects in the images depends on two unknown factors: capture distance and body dimensions. This ambiguity is particularly pronounced in partial-view scenarios. To address this challenge, we propose a modular and simple height normalization solution. This solution relocates the subject skeleton to the desired position, thereby normalizing the scale and disentangling the relationship between the two variables. Our experimental results demonstrate that integrating this technique into state-of-the-art human mesh reconstruction models significantly enhances partial body measurement estimation. Additionally, we illustrate the applicability of this approach to multi-view settings, showcasing its versatility.

link: http://arxiv.org/abs/2404.09301v2

High Significant Fault Detection in Azure Core Workload Insights

Pranay Lohia, Laurent Boue, Sharath Rangappa, Vijay Agneeswaran

Azure Core workload insights have time-series data with different metric units. Faults or Anomalies are observed in these time-series data owing to faults observed with respect to metric name, resources region, dimensions, and its dimension value associated with the data. For Azure Core, an important task is to highlight faults or anomalies to the user on a dashboard that they can perceive easily. The number of anomalies reported should be highly significant and in a limited number, e.g., 5-20 anomalies reported per hour. The reported anomalies will have significant user

perception and high reconstruction error in any time-series forecasting model. Hence, our task is to automatically identify 'high significant anomalies' and their associated information for user perception.

link: http://arxiv.org/abs/2404.09302v1

Monte Carlo Search Algorithms Discovering Monte Carlo Tree Search Exploration Terms

Tristan Cazenave

Monte Carlo Tree Search and Monte Carlo Search have good results for many combinatorial problems. In this paper we propose to use Monte Carlo Search to design mathematical expressions that are used as exploration terms for Monte Carlo Tree Search algorithms. The optimized Monte Carlo Tree Search algorithms are PUCT and SHUSS. We automatically design the PUCT and the SHUSS root exploration terms. For small search budgets of 32 evaluations the discovered root exploration terms make both algorithms competitive with usual PUCT.

link: http://arxiv.org/abs/2404.09304v1

In My Perspective, In My Hands: Accurate Egocentric 2D Hand Pose and Action Recognition

Wiktor Mucha, Martin Kampel

Action recognition is essential for egocentric video understanding, allowing automatic and continuous monitoring of Activities of Daily Living (ADLs) without user effort. Existing literature focuses on 3D hand pose input, which requires computationally intensive depth estimation networks or wearing an uncomfortable depth sensor. In contrast, there has been insufficient research in understanding 2D hand pose for egocentric action recognition, despite the availability of user-friendly smart glasses in the market capable of capturing a single RGB image. Our study aims to fill this research gap by exploring the field of 2D hand pose estimation for egocentric action recognition, making two contributions. Firstly, we introduce two novel approaches for 2D hand pose estimation, namely EffHandNet for single-hand estimation and EffHandEgoNet, tailored for an egocentric perspective, capturing interactions between hands and objects. Both methods outperform state-of-the-art models on H2O and FPHA public benchmarks. Secondly, we present a robust action recognition architecture from 2D hand and object poses. This method incorporates EffHandEgoNet, and a transformer-based action recognition method. Evaluated on H2O and FPHA datasets, our architecture has a faster inference time and achieves an accuracy of 91.32% and 94.43%, respectively, surpassing state of the art, including 3D-based methods. Our work demonstrates that using 2D skeletal data is a robust approach for egocentric action understanding. Extensive evaluation and ablation studies show the impact of the hand pose estimation approach, and how each input affects the overall performance.

link: http://arxiv.org/abs/2404.09308v1

Text-to-Song: Towards Controllable Music Generation Incorporating Vocals and Accompaniment

Zhiqing Hong, Rongjie Huang, Xize Cheng, Yongqi Wang, Ruiqi Li, Fuming You, Zhou Zhao, Zhimeng Zhang

A song is a combination of singing voice and accompaniment. However, existing works focus on singing voice synthesis and music generation independently. Little attention was paid to explore song synthesis. In this work, we propose a novel task called text-to-song synthesis which incorporating both vocals and accompaniments generation. We develop Melodist, a two-stage text-to-song method that consists of singing voice synthesis (SVS) and vocal-to-accompaniment (V2A) synthesis. Melodist leverages tri-tower contrastive pretraining to learn more effective text representation for controllable V2A synthesis. A Chinese song dataset mined from a music website is built up to alleviate data scarcity for our research. The evaluation results on our dataset demonstrate that Melodist can synthesize songs with comparable quality and style consistency. Audio samples can be found in https://text2songMelodist.github.io/Sample/.

link: http://arxiv.org/abs/2404.09313v2

Characterizing Soft-Error Resiliency in Arm's Ethos-U55 Embedded Machine Learning Accelerator

Abhishek Tyagi, Reiley Jeyapaul, Chuteng Zhu, Paul Whatmough, Yuhao Zhu

As Neural Processing Units (NPU) or accelerators are increasingly deployed in a variety of applications including safety critical applications such as autonomous vehicle, and medical imaging, it is critical to understand the fault-tolerance nature of the NPUs. We present a reliability study of Arm's Ethos-U55, an important industrial-scale NPU being utilised in embedded and IoT applications. We perform large scale RTL-level fault injections to characterize Ethos-U55 against the Automotive Safety Integrity Level D (ASIL-D) resiliency standard commonly used for safety-critical applications such as autonomous vehicles. We show that, under soft errors, all four configurations of the NPU fall short of the required level of resiliency for a variety of neural networks running on the NPU. We show that it is possible to meet the ASIL-D level resiliency without resorting to conventional strategies like Dual Core Lock Step (DCLS) that has an area overhead of 100%. We achieve so through selective protection, where hardware structures are selectively protected (e.g., duplicated, hardened) based on their sensitivity to soft errors and their silicon areas. To identify the optimal configuration that minimizes the area overhead while meeting the ASIL-D standard, the main challenge is the large search space associated with the time-consuming RTL simulation. To address this challenge, we present a statistical analysis tool that is validated against Arm silicon and that allows us to quickly navigate hundreds of billions of fault sites without exhaustive RTL fault injections. We show that by carefully duplicating a small fraction of the functional blocks and hardening the Flops in other blocks meets the ASIL-D safety standard while introducing an area overhead of only 38%.

link: http://arxiv.org/abs/2404.09317v1

The intelligent prediction and assessment of financial information risk in the cloud computing model

Yufu Wang, Mingwei Zhu, Jiaqiang Yuan, Guanghui Wang, Hong Zhou

Cloud computing (cloud computing) is a kind of distributed computing, referring to the network "cloud" will be a huge data calculation and processing program into countless small programs, and then, through the system composed of multiple servers to process and analyze these small programs to get the results and return to the user. This report explores the intersection of cloud computing and financial information processing, identifying risks and challenges faced by financial institutions in adopting cloud technology. It discusses the need for intelligent solutions to enhance data processing efficiency and accuracy while addressing security and privacy concerns. Drawing on regulatory frameworks, the report proposes policy recommendations to mitigate concentration risks associated with cloud computing in the financial industry. By combining intelligent forecasting and evaluation technologies with cloud computing models, the study aims to provide effective solutions for financial data processing and management, facilitating the industry's transition towards digital transformation.

link: http://arxiv.org/abs/2404.09322v1

Correlated Mean Field Imitation Learning

Zhiyu Zhao, Ning Yang, Xue Yan, Haifeng Zhang, Jun Wang, Yaodong Yang

We investigate multi-agent imitation learning (IL) within the framework of mean field games (MFGs), considering the presence of time-varying correlated signals. Existing MFG IL algorithms assume demonstrations are sampled from Mean Field Nash Equilibria (MFNE), limiting their adaptability to real-world scenarios. For example, in the traffic network equilibrium influenced by public routing recommendations, recommendations introduce time-varying correlated signals into the game, not captured by MFNE and other existing correlated equilibrium concepts. To address this gap, we propose Adaptive Mean Field Correlated Equilibrium (AMFCE), a general equilibrium incorporating time-varying correlated signals. We establish the existence of AMFCE under mild conditions and

prove that MFNE is a subclass of AMFCE. We further propose Correlated Mean Field Imitation Learning (CMFIL), a novel IL framework designed to recover the AMFCE, accompanied by a theoretical guarantee on the quality of the recovered policy. Experimental results, including a real-world traffic flow prediction problem, demonstrate the superiority of CMFIL over state-of-the-art IL baselines, highlighting the potential of CMFIL in understanding large population behavior under correlated signals.

link: http://arxiv.org/abs/2404.09324v1