SfS²

SECURE for Student Success

Undergraduate Research Program

Computer Science

Research Duration: Summer 2025 (June – August 2025)

Faculty: Xunfei Jiang

Email address: xunfei.jiang@csun.edu

Contact No: JD3519, 818-677-6769
Monday/Wednesday 9:30-11:00AM on Zoom

Title of Project: **Project 1**: Energy-efficient Workload Scheduling in Datacenters
Project 2: Point-based Approach for 3D Vehicle Detection and Classification for Roadside for Traffic Flow

Project 1: Energy-efficient Workload Scheduling in Datacenters (2 students)

Various strategies have been proposed to reduce the computing cost (electrical energy consumption of servers) of computer servers in data centers, while the cooling cost (energy consumption of cooling systems) has not been fully investigated. Predicting the cooling cost is a challenge because it is affected by not only the utilization of computer servers, but also by the temperature distribution of cluster systems [1]. For a homogenous rack mount cluster system, cooling costs would be reduced by dispatching tasks to nodes near the bottom of the rack because it results in lighter air recirculation and thus lower cooling costs [2]. With the ability to predict the temperatures of cluster nodes, the cooling cost of the cluster could be estimated. A recent study applied a gradient boosting machine learning model to predict the server temperature [3]; however, it considered only the impact of the CPU and inlet temperatures.

Goal and Objectives: This project will develop energy-efficient workload scheduling strategies based on energy consumption models that simulate the energy cost of components in a cluster system. In this program, students will apply thermal and energy machine learning models to evaluate various workload scheduling strategies on a datacenter simulator, and propose new strategies to reduce the energy consumption for GPU intensive workload. Real-world workload traces will be used as input for running simulations, and experiments will be conducted for evaluating workload scheduling strategies.

Prior work in the area: This project is part of an ongoing project to develop energy-efficient workload management for cluster systems in datacenters. I have mentored several undergraduate and graduate students working on this project in the past 3 years, which all involve in-depth study and literature review on corresponding topics. 3 papers co-authored with undergraduate students have been published from this project [4-6]**.** In summer 2024, we developed an energy prediction model to estimate the energy consumption of GPU under different workload, and modified the GPU CloudSim Plus simulator to incorporate real-world datacenter workload traces as input and produce summary of energy consumption for applying different scheduling algorithms.

Expectation and Outcomes:

Students will develop a literature review to gain knowledge and experience of DS and PDC systems. They will conduct experiments to characterize energy consumption patterns of major components in the test bed cluster through data collection and analysis.
(1) Students will apply a set of temperature and energy consumption models to validate various workload scheduling strategies by using real-world datacenter workload traces on a datacenter simulator.
(2) Students will develop a literature review to gain knowledge and experience of DS and PDC systems.
(3) Students will explore methods to feed real-world datacenter workload traces to a datacenter simulator, design and conduct experiments to evaluate the energy-efficient of workload scheduling strategies through data collection and analysis.

Student learning:

In this project, students will gain immerse research experience on problem-solving through analyzing existing research on workload scheduling for datacenters, designing and conducting experiments, and applying DS in data collection, data processing, and data analysis to evaluate different workload scheduling algorithms. Energy-efficient workload management and data placement strategies will be introduced to students. They will learn machine learning skills and apply them on energy-efficient workload management in datacenter, and work with group members through collaborative discussions with participants possessing diverse technical backgrounds and research preparation.

[1] Moore, Justin D., Jeffrey S. Chase, Parthasarathy Ranganathan, and Ratnesh K. Sharma. "Making Scheduling" Cool": Temperature-Aware Workload Placement in Data Centers." In *USENIX annual technical conference, General Track*, pp. 61-75. 2005.
[2] Tang, Qinghui, Sandeep Kumar S. Gupta, and Georgios Varsamopoulos. "Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach." *IEEE Transactions on Parallel and Distributed Systems* 19, no. 11 (2008): 1458-1472.
[3] Ilager, Shashikant, Kotagiri Ramamohanarao, and Rajkumar Buyya. "Thermal prediction for efficient energy management of clouds using machine learning." *IEEE Transactions on Parallel and Distributed Systems* 32, no. 5 (2020): 1044-1056.
[4] I. Nisce, X. Jiang and S. P. Vishnu, "Machine Learning based Thermal Prediction for Energy-efficient Cloud Computing," 2023 IEEE 20th Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2023, pp. 624-627, doi:10.1109/CCNC51644.2023.10060079.
[5] M. Smith, L. Zhao, J. Cordova, X.-F. Jiang, and M. Ebrahimi. "Machine Learning-based Energy-efficient Workload Management for Data Centers". 2024 IEEE Consumer Communications & Networking Conference (IEEE CCNC 2024), January 6–9, 2024, Las Vegas, NV, USA. (In press)
[6] M. Smith, L. Zhao, J. Cordova, X.-F. Jiang, and M. Ebrahimi. "Energy-efficient GPU-intensive Workload Scheduling for Data Centers". 2023 Symposium for Undergraduate Research in Data Science, Systems, and Security (REU Symposium 2023), Special Session at the 22nd International Conference on Machine Learning and Applications (ICMLA 2023), Dec 15-17, 2023, Jacksonville, Florida. (In Press)

Project 2: Point-based Approach for 3D Vehicle Detection and Classification for Roadside for Traffic Flow (3 students)

Traffic has been a long problem in Southern California, and the situation in Los Angeles is even worse. According to a report from INRIX, the Los Angeles urban area was the sixth most congested area in the United States in 2021 [1]. Intelligent Transport System (ITS) improves traffic efficiency because it enables users to know information about real-time traffic flow and also provides support for transportation departments to make planning for traffic management. Object detection technologies provide fundamental support for ITS systems in traffic monitoring to assistant the understanding of traffic conditions (such as speed, density, and prediction) [2]. Machine Learning technology has been increasingly used for vehicle detection, and adverse weather conditions prove to be challenging for 2D vehicle detection. Using a 3D LiDAR camera, vehicles in traffic are captured in the form of point clouds, which are more resistant to adverse weather conditions. 3D LiDAR vehicle detection has been widely used in autonomous driving, but there is a lack of research on roadside vehicle detection using 3D LiDAR point clouds. With a trained 3D vehicle detection and classification model, information such as the number of vehicles, vehicle types, traffic flow speed, or density of vehicles in each lane can be calculated and visualized to provide insights on traffic flow.
Goals: In this research, we will investigate roadside vehicle detection and classification for 3D LiDAR point cloud data collected from traffic flow in Southern California. The 3D machine learning model is a key component in the real time traffic monitoring system, and it will provide fundamental support for the real-time traffic detection. New LiDAR point cloud datasets will be collected and labeled, and processed for vehicle detection and classification. Point-based approaches for 3D vehicle detection will be investigated [3-4], and the models will be trained and tested on the new labeled datasets. Detection results generated from the trained machine learning model will be processed, and analysis will be conducted for the collected traffic flow datasets. Students will participate in research activities such as reading research papers, install and software systems for data labeling, write script programs to automate the data labeling work, process the data for machine learning models, and test the performance of selected machine learning algorithms.
Prior work in the area: This project is part of an ongoing project to build a real-time traffic monitoring system since February 2021. I have mentored 3 groups of senior design project students and 3 master students working on this project, which all involve in-depth study and literature review on corresponding topics. The senior design group in 2021-2022 won the first place of oral presentations, and the senior design group in 2022-2023 won the first place of project display in the CSUN CECS Senior Design Project Showcase. I supervised undergraduate students to revise and submit their project final papers, and two papers from the project had published in **IEEE GESSC 2022 and IEEE GESSC 2023** [5-6]**.** In addition, another two research papers from this project have been accepted for publication by IEEE GESS 2024 and IEEE ICMLA 2024. In spring and summer 2024, we investigated the projection based approach for 3D vehicle detection and classification using Complex-YOLO algorithm, implemented the Complex-YOLO model through transfer learning and normalization in handling and aligning the dataset for effective learning, and extended the object classification from 3 types to 9 types.
Expected outcomes of this research include: (1) a literature review of existing point-based 3D vehicle detection and classification machine learning models; (2) an automatic approach of vehicle labeling for 3D LiDAR point cloud datasets; (3) train and test Point-based 3D vehicle detection and classification models and analysis of the results; and (4) a poster which concludes the work in the summer research.
Student learning: Training sessions will be provided for the student to conduct literature review, such as: how to read a paper and find related resources for technologies used in the paper, how to analyze the performance of a machine learning model, and design programs to process the 3D Lidar point cloud datasets and for data visualization. We will discuss several research papers and the student will be required to read selected research papers in detail to learn the layout of a research paper and technical writing. Based on the survey paper, we will investigate applying one selected 3D machine learning model to detect and classify the vehicles for the labeled 3D traffic datasets.

Reference

[1] B. Pishue, “2021 inrix global traffic scorecard,” INRIX (December 2021), 2021.
[2] Z. Yang and L. S. Pun-Cheng, “Vehicle detection in intelligent transportation systems and its applications under varying environments: A review,” Image and Vision Computing, vol. 69, pp. 143–154, 2018. [Online]. Available:https://www.sciencedirect.com/science/article/pii/S0262885617301592

[3] Y. Wu, Y. Wang, S. Zhang, and H. Ogai, “Deep 3d object detection networks using lidar data: A review,” IEEE Sensors Journal, vol. 21, no. 2, pp. 1152–1171, 2021.
[4] Ester, Martin, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. "A density-based algorithm for discovering clusters in large spatial databases with noise." In kdd, vol. 96, no. 34, pp. 226-231. 1996.
[5] A. Rose, Ic. Nisce, A. Gonzalez, M. Davis, B. Uribe, J.Carranza, J. Flores, X.-D. Jia, B.-B. Li, X.-F. Jiang, "A Cloud-based Real-Time Traffic Monitoring System with Lidar-based Vehicle Detection". 2023 IEEE Green Energy and Smart Systems Conference, November 13-14, 2023, Long Beach, CA, USA.
[6] R. Wu, Z. Chowdhury, G. V. Sanchez, X. Gao, C. Villa and X. Jiang, "Real-time Vehicle Detection System for Intelligent Transportation using Machine Learning," 2022 IEEE Green Energy and Smart System Systems(IGESSC), 2022, pp. 1-6, doi: 10.1109/IGESSC55810.2022.9955329.