CAREER: Robustifying Machine Learning for Cyber-Physical Systems

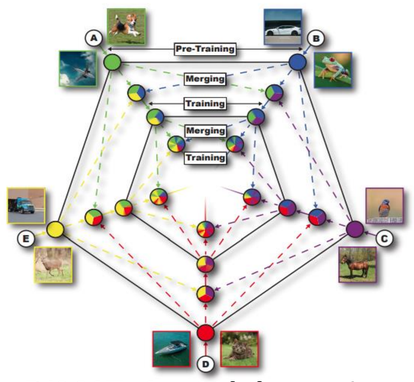
Soumik Sarkar, PhD. Department of Mechanical Engineering, Iowa State University, Ames, IA

Introduction

This project focuses on detecting and reducing the vulnerabilities of ML models that have become pervasive and are being deployed for decision-making in real-life CPS applications including self-driving cars, robots, and other autonomous systems.

Challenges

- The major technical challenges we focused on in Year 5 of this project include: (i) developing decentralized deep learning (DDL) algorithms that enable systemic robustness; (ii) developing fast certification algorithms for large vision-language models such as CLIP
- Explore 'model merging' techniques for decentralized deep learning (DDL) to reduce communication and computation overheads
- Open Vocabulary Certification (OVC): how to quickly certify large Vision-Language Models, e.g., CLIP, for novel prompts?



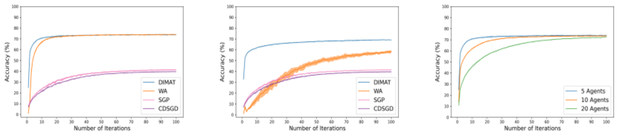
DIMAT: A model merging framework for DDL.

Scientific Impact

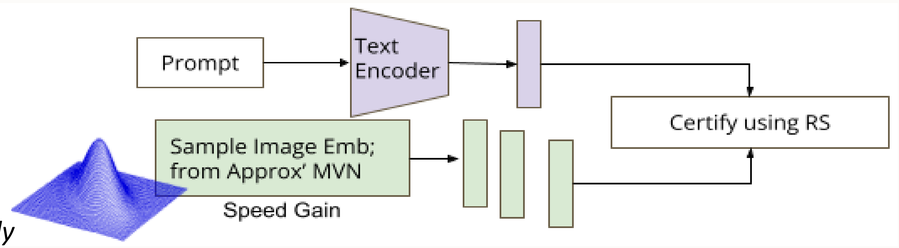
- DIMAT shows superior performance with faster initial gains due to its sublinear convergence for nonconvex functions, tighter error bounds, and sustained linear speedup compared to existing approaches.
- We present both exact and heuristic methods for fast OVC achieving up to ~130X speedup in certifying vision language models, e.g., CLIP

Technical Approach/Solution

- We proposed DIMAT for DDL with extensive theoretical and empirical analysis that utilizes advanced model merging techniques (e.g., activation matching)
- We built on the Incremental Randomized Smoothing concept for OVC that cleverly exploits certification for existing prompts to compute certificates for new prompts quickly



DIMAT achieves faster convergence for both (a) fully connected and (b) ring topologies with significantly reduced communication rounds among the agents



Fast Open Vocabulary Certification

Broader Impact

- Both DDL and OVC frameworks have real-world applicability in various CPS sectors including transportation, manufacturing and agriculture
- DIMAT with lightweight communication and computation, can significantly improve MLOps in IoT
- Fast certification of vision-language foundation models can lead to safety assurance of learning-enabled CPS

Education & Outreach

- PI Sarkar led the development of, and continues to lead an Undergraduate CPS minor at Iowa State that began in Fall 2021


IOWA STATE UNIVERSITY
College of Engineering News

New cyber-physical systems minor leverages industry ties to enhance student futures

A new undergraduate minor in cyber-physical systems (CPS) will debut in the fall 2021 semester. It will be open to all Iowa State engineering majors, and will combine teaching efforts from three different College of Engineering departments: mechanical engineering, electrical and computer engineering, and aerospace engineering – with mechanical engineering serving as the home and administrative department for the program.

- PI Sarkar continues to offer his successful graduate course: "Data Analytics and ML for CPS" using results from this research

Impact Quantification

- Project developed and open-sourced multiple robust ML software and data sets
- Partially supported PhD study of 10 students in 5 years including 3 URM students
- PI Sarkar leads the Translational AI Center (TrAC) at Iowa State; through the center, he organizes multiple ML workshops and tutorials impacting more than 200 participants each year

Award ID#: CNS-1845969