PURDUE UNIVERSITY | Department of Computer Science COLLEGE OF SCIENCE

2018 GOBOILER INTERNSHIP PROJECT LIST

MACHINE LEARNING

Project number: 1
Project title: Structured Prediction
Faculty member: Jean Honorio
More info: Learning and inference drives much of the research in diverse domains. For instance, in computer vision and image understanding, complex models are required in order to better represent real-world images. As such, one aims to obtain more representational power by expressing images as the interaction of a large number of constituent elements. While producing more realistic models, this also increases the computational cost of inference and learning. In this project, you will be able to use a family of newly created computationally-efficient randomized algorithms on several real-world image datasets.
E-mail: jhonorio@purdue.edu

INFORMATION SECURITY & PRIVACY

Project number: 2
Project title: Interconnecting Blockchains Together
Faculty member: Aniket Kate
More info: Cryptocurrencies such as Bitcoin and Ethereum have emerged as a paradigm shift for the way payment systems work today. They rely on the blockchain, a technology that has been proven useful in a vast number of applications other than monetary transactions. Many companies today are tailoring the blockchain technology to their business logic and successfully developing applications for credit settlement networks, supply chain, IoT and beyond. However, these separate efforts are leading to incompatible individual systems. This contrasts with our highly interconnected world and it is inevitable to see that soon these blockchains will need to operate with each other, effectively forming a network of blockchains where transactions can flow through a sequence of blockchains, similar how the network of networks (i.e., the Internet) works today. Towards enabling the Internet of Value, in this project, we will design and evaluate the cryptographic and distributed systems tools required to move money the same way as the information/data moves today. Project Webpage:
https://freedom.cs.purdue.edu/blockchains/
E-mail: aniket@purdue.edu

Project number: 3
Project title: Blockchain Hub
Faculty member: Bharat Bhargava

More info: Big data in various forms (including streaming) from multiple sources is received by government agencies for dissemination to many stakeholders. To ensure integrity, trust, immutability, and authenticity of information (cyber data, user data, and attack event data), research is needed to advance blockchains and use them for provenance, trust, and analytics of data. We propose an approach for tracking the provenance of data with enhanced blockchain-based mechanisms to build trustworthiness of the data and ensure identities of the parties/nodes.
E-mail: bbshail@purdue.edu

Project number: 4
Project title: V2V Safety and Security
Faculty member: Bharat Bhargava
More info: Evaluate V2V and its safety under various scenarios and attacks.
E-mail: bbshail@purdue.edu

FORMAL METHODS/PROGRAMMING LANGUAGES

Project number: 5
Project title: Robustness of Probabilistic Programs
Faculty member: Roopsha Samanta
More info: The goal of this project is to develop foundational results and practical techniques for the design of robust systems in the presence of stochasticity in the system or its environment. An example problem is ensuring that the expected behavior of a probabilistic program is "acceptable" when the input distribution is perturbed. The project will involve development and/or implementation of algorithms for verification/repair/synthesis of robust programs. The candidate must have a strong background in program analysis and probability theory.
E-mail: roopsha@purdue.edu

Project number: 6
Project title: Data-driven Repair of Web Applications
Faculty member: Roopsha Samanta
More info: PHP web applications are ubiquitous and error-prone. Bugs such as HTML validation errors, access-control bugs, SQL injection and cross-site scripting (XSS) are commonplace in PHP applications. Inspired by the success of data-driven repair of programs written in C, Java or C#, the goal of this project is to develop a framework for data-driven repair of real-world PHP-based server-side programs. The framework will learn probabilistic models of repaired/correct programs from training data, which will then be used to produce likely candidate fixes for buggy programs. The project will involve identification of open-source repositories of PHP web applications and investigation of different choices of program representations, probabilistic models and machine learning algorithms. The candidate must have a strong background in program analysis and machine learning.
E-mail: roopsha@purdue.edu

MOBILE NETWORKING/MOBILE COMPUTING

Project number: 7
Project title: Supporting Immersive Virtual Reality Over 4G LTE
Faculty member: Chunyi Peng
More info: In recent years, we have witnessed a boom in mobile virtual reality (VR) to offer panorama, immerse 3D experience for users. The most popular mobile VR form is empowered by mobile phones with VR headsets (e.g., Google Cardboard and DayDream, Samsung Gear VR), which contributes 98% of the sales. However, our current 4G LTE mobile network can't support ubiquitous and high-fidelity VR experience "anytime, anywhere" regardless of whether they roam (e.g., in cars or on trains) or remain static in indoor/outdoor settings. The biggest obstacle lies in network latency.
In this project, we study how to lower network latency in better support of mobile VR. We will quantify the sources of latency, explore the potential improvement room in a standard-compatible way and devise a solution which is ready to launch and improve VR experience right away. In the next phase, we will move to clean-slate design to further support real-time VR.
E-mail: chunyi@purdue.edu

Project number: 8
Project title: Build IoT Applications for Autonomous Driving at Scale
Faculty member: Chunyi Peng
More info: We envision a fully automated driving experience in the years to come. V2X (vehicle-to-anything) will play a critical role to make it happen. In this project, we conduct the preliminary study to use IoT technique over cellular links (here, 4G LTE) to prototype it. We will incorporate mobile edge computing and 4G networking tailored for critical communication, massive communication, low power and low latency communication to study its feasibility and disclose its challenges when it is extended to a large scale.
E-mail: chunyi@purdue.edu

COMPUTER NETWORKING

Project number: 9
Project title: Investigation of a Scalable Data Center Network Architecture
Faculty member: Douglas Comer
More info: The move to cloud computing has changed the network traffic patterns in multi-tenant data centers. Instead of primarily handling web traffic between a server in the data center and a user at an arbitrary location on the Internet, data centers now experience significant intra-center traffic that travels between pairs of servers within the data center (so called "east-west" traffic). We are investigating a novel data center network architecture that supports significant east-west traffic and allows a VM to retain its IP address as the VM migrates from one physical server to another. An intern will help conduct experiments that measure and assess the design, analyze the resulting measurements, and propose modifications or extensions.
E-mail: comer@cs.purdue.edu

Project number: 10
Project title: Assessing Mesh Routing Protocols for the Internet of Things
Faculty member: Douglas Comer
More info: Protocols for low-power wireless Internet of Things (IoT) devices use a mesh networking approach. Devices that cannot reach a central base station connect to neighbors and depend on neighbors to forward packets on their behalf. The standards body in charge of Internet protocols, the IETF, has developed a routing protocol for use in such networks. Named RPL (Routing over Low Power and Lossy Networks), the protocol uses neighbor relationships to impose a tree topology on a set of nodes. We are constructing a testbed to assess wireless protocols. An intern will help finish an implementation of RPL, and then use the testbed to assess how RPL reacts to continuous changes in topology (nodes joining and leaving the mesh, and nodes moving from one location to another).
E-mail: comer@cs.purdue.edu

COMPILERS, ARTIFICIAL INTELLIGENCE & DATA MGMT.

Project number: 11
Project title: Flare: A High-Performance Big Data and AI System
Faculty member: Tiark Rompf
More info: Flare is a big data and machine learning platform developed here at Purdue (flaredata.github.io). Flare can transparently accelerate pipelines implemented in Apache Spark and TensorFlow, and provides speedups of 10x-100x, thanks to cutting edge compiler technology. The goal of this summer research project will be to extend Flare in one of several possible dimensions: implement code generation for distributed execution (using MPI or similar), implement streaming abstractions for incremental data processing, implement code generation for GPUs, implement a cost-based query optimizer, implement new internal compiler optimizations, implement case studies based on state of the art AI kernels.
E-mail: tiark@purdue.edu

COGNITIVE INTELLIGENCE & LEARNING

Project number: 12
Project title: Intelligence Autonomous Systems
Faculty member: Bharat Bhargava
More info: Intelligent Autonomous Systems (IAS) are highly cognitive, reflective, multitask-able, and effective in knowledge discovery. Examples of IAS include software that is capable of automatic reconfiguration, autonomous vehicles, network of sensors with reconfigurable sensory platforms, and an unmanned aerial vehicle respecting privacy by deciding to turn off its camera when pointing inside a private residence. Research is needed to build systems that can monitor their environment and interactions, learn their capability, and adapt to meet the mission objectives with limited or no human intervention. The systems should be fail-safe and should allow for graceful degradations while continuing to meet the mission objectives.
The main objective is to realize a vision based on the following approaches.
(1) Employ machine learning techniques on sensor and provenance data to learn and

understand the underlying patterns of interaction, conduct forensics to detect anomalies, and provide assistance in decision making by on-the-fly semantic and probabilistic reasoning.
(2) Apply advanced data analytics techniques to incomplete and hidden raw system data (provenance data, error logs, etc.,) to discover new knowledge that contributes to the success of the IAS mission.
(3) Enhance the autonomous system's self-awareness, self-protection, self-healing, and self-optimization by learning from the knowledge discovered through data analytics.
(4) Utilize blockchain technology for storing provenance data for providing monitoring, trust, and verification, using the WaxedPrune system developed for Northrup Grumman.
E-mail: bbshail@purdue.edu