Abstract of the Keynotes

Stable and Efficient Recurrent Neural Networks
Prof. Inderjit S. Dhillon
University of Texas at Austin, USA

Despite having remarkable performance on many sequence learning tasks, recurrent neural networks (RNNs) are hard to train with long sequences due to limited expressive power and the vanishing and exploding gradient issues. Previous work has focused on stabilizing the gradients by encouraging orthogonality of weight matrices via re-parameterization techniques. However, two major issues remain in these methods. First, the re-parameterization often relies on a chain of operations on small matrices or vectors that are not friendly to hardware accelerators. As a result, it becomes a source of performance bottleneck for training. Second, these methods fix the singular values of the transition matrix throughout the temporal dimension, which further restricts the expressive power of the model and wastes the potential of encoding useful information into the singular values. In this talk, I will present the Singular Value Gated RNN that can efficiently encode temporal information into singular values, as well as mitigate the vanishing and exploding gradient problems. In addition, we can design novel forward and backward propagation algorithms that are friendly to hardware accelerators. This leads to 3-4 times speedup on GPUs and greatly reduces memory cost. On contemporary applications like voice recognition and text summarization, where long term dependencies are hard to capture, the proposed method outperforms other recurrent models with similar or smaller model sizes. Joint work with Jiong Zhang of UT Austin

Understanding Visual Appearance from Micro Scale to Global Scale
Prof. Kavita Bala
Cornell University, New York, USA

Mixed reality environments require understanding scenes, and then seamlessly blending the rich visual appearance of real and virtual materials to create a compelling user experience. In this talk I will describe our work on modeling and recognizing complex materials, and visual recognition. Using these algorithms as core building blocks we can understand appearance at a global scale by mining social media to discover visual patterns across geography and time. This talk will describe our work on understanding global fashion styles and trends.

Determinism and Energy Efficiency in Industrial IoT
Prof. Christos Douligeris
University of Piraeus, Greece

Industrial applications, such as smart grid, smart manufacturing or automotive industry applications, require ultra low-latency, ultra low-power consumption and high network reliability. The Time Slotted Channel Hopping (TSCH) mode of IEEE802.15.4 can provide these characteristics; however, as all the wireless technologies, it is prone to internal and external interference. A number of techniques to mitigate or avoid potential collisions since the formation of the network is described. Moreover, promising wireless power transfer methods are introduced to provide energy efficiency and even power autonomy to Industrial Internet of Things (IIoT) applications.
AI and ML in Manufacturing – Past, Present and Future

Prof. Soundar Kumara
Professor of Information Sciences & Technology (Affiliate), Pennsylvania State University, USA

Artificial Intelligence and Machine Learning have played a significant role since the early 1980s in Manufacturing. From the first NSF ERC on Intelligent Manufacturing to the current DoE’s CESMII, AI has played a significant role. This talk will focus on the application of AI and ML to solve several relevant manufacturing problems and discuss the future impact of this new paradigm. We will discuss the history of AI starting with the Intelligent Manufacturing Systems of the 1980s to look into the future from a personal viewpoint. We will emphasize the need for a new science of ML using IoT data, real-time nonlinear data analysis and community detection in networks. We will discuss the smart manufacturing of the future which will be combining the matching features of Uber, Google search mechanisms, and Amazon’s distribution formalisms leading to manufacture anytime, anywhere a reality.

A Multi-sensor System for Traffic Analysis at Smart Intersections

Prof. Sanjay Ranka
University of Florida, USA

We present a multi-sensor system for vehicle and pedestrian traffic analysis and visualization at intersections to discover trajectory patterns and anomalous traffic behavior. Augmenting these data with signal and phasing information, we show how clustering in the context of signal information may help us to detect anomalies with respect to vehicles violating signals. We demonstrate our workflow on two intersections with very different traffic composition. The system may be leveraged by a number of other applications, including conflict detection in object movements, turn movement counts, incident detection and management, and demand profiling, for better traffic management through the adjustment of signal timing.

Two Tales of Cache Optimization: Technology-oblivious and Technology-aware

Prof. Mainak Chauduri
IIT Kanpur, India

An important class of processor cache optimization techniques attempts to improve the cache hit rate. These techniques are based on reuse-based principles that abstract away the technology used to build the cache. In the first part of the talk, I will follow this traditional cache design principle and present a technique that dynamically learns reuse probability at run-time and employs the learned probabilities to design high-performance cache management policies. We have applied this generic technique to a number of different scenarios employing SRAM caches such as shared last-level cache of multi-core CPUs, last-level cache of discrete GPUs, and shared last-level cache of CPU-GPU heterogeneous multiprocessor SoCs.

Model-Based Safety Analysis in Cyber-Physical Systems

Prof. Ratnesh Kumar
Iowa State University, USA

Cyber-Physical Systems abound us. They combine Cyber---Software for networking and decision-making, and Physical components. Accordingly, their behaviors comprise of hybrid of discrete and continuous states, and are modeled as hybrid automata. Simulink/Stateflow is a commonly used modeling and simulation platform for cyber-physical systems consisting of reactive embedded code that interacts with its environment in real-time fashion. This talk will present an approach for safety analysis of Simulink/Stateflow models, based on its automated translation to input-output extended finite automaton (I/O-EFA), followed by automated test-generation, guaranteeing user-defined code as well as requirements coverage, and also support for automated test-execution and error-localization. While testing is useful for design-time error analysis, the talk will further discuss our model-based approach for run-time error monitoring, detection and localization. Monitoring at CPS level (as opposed to software level) is necessarily stochastic, and a more general I/O-Stochastic Hybrid Automaton (I/O-SHA) model is used, and condition for bounded-delay detectability, and achieving desired levels of false-positives/-negatives will be discussed. The talk also presents our simulation-based approach for safely analysis of hybrid systems, where a finite number of simulation runs are used to confirm bounded-horizon safety, useful for run-time assurance.