## Constantine Dovrolis

**Short Biography:** Dr. Constantine Dovrolis is a Professor at the School of Computer Science at the Georgia Institute of Technology (Georgia Tech). He is a graduate of the Technical University of Crete (Engr.Dipl. 1995), University of Rochester (M.S. 1996), and University of Wisconsin-Madison (Ph.D. 2000). His research combines Network Science, Data Mining and Machine Learning with applications in climate science, biology, neuroscience, sociology and machine learning. More recently, his group has been focusing on neuro-inspired architectures for machine learning based on what is currently known about the structure of brain networks.

**Presentation Title:**

If the brain is a very sparse network, why does deep learning use dense neural networks?

**Abstract:**

What we know from neuroscience ("connectomics") is that the brain is, overall, a very sparse network with relatively small locally dense clusters of neurons. These topological properties are crucial for the brain's ability to perform efficiently, robustly, and to process information in a hierarchically modular manner. On the other hand, the artificial neural networks we use today are very dense, or even fully connected, at least between successive layers. Additionally, it is well known that deep neural networks are highly over-parameterized: pruning studies have shown that it is often possible to eliminate 90% of the connections (weights) without significant loss in performance. Pruning, however, is typically performed after the dense network has been trained, which only improves the run-time efficiency of the inference process. The previous points suggest that we need methods to design sparse neural networks, without any training, that can perform almost as well as the corresponding dense networks after training. This talk will first provide some background in the pruning literature, either after training or before training. Then, we will present a recently proposed (ICML 2021) method called PHEW (Paths with Higher Edge Weights) which creates sparse neural networks, before training, and that can learn fast and generalize well. Additionally, PHEW does not require access to any data as it only depends on the initial weights and the topology of the given network architecture.