Accurate forecast of data center energy consumption can help eliminate waste caused by overprovisioning. However, due to the nonlinear nature of its dynamics, energy prediction remains a challenge for data centers. Within a data center, multiple electrical and mechanical systems interact in a complex manner in supplying power and cooling to the physical servers consolidated within virtual machine cloud environment. An added layer of complexity further comes from dynamically changing workloads.

These practical factors coupled with constraints on physical ability to measure and capture meaningful data make it difficult to model the energy characteristics of the data center. Yet, two approaches in the data center modelling have stood out. One approach is the physical modelling of data centers using clear-box model. This approach has the advantage of representing the nonlinear data center system whilst retaining useful information related to its energy consumption. However, data center systems are complicated and usually involved unknown nonlinearities and dynamics that often make it impossible to build an accurate clear-box model. The other approach is using the black-box models such as neural networks models employed to approximate the non-linear data center system. An advantage of this approach lies in its simplicity, where little or no a prior knowledge of the system is required. Nonetheless, an apparent drawback is that the physical meaning of input parameters and their impacts on the system outputs could not be captured, thereby placing a limit to further investigate the underlying system.

In this paper, we build on our previous successes in neural network based modelling and predictions where neural networks, evolved using Evolutionary Algorithms, were applied to accurately forecast financial data trends and energy yield of photovoltaic cells. By combining the merits of the black-box model with conventional clear-box approaches, we develop a ‘grey box’ that can preserve the physical significance of the system parameters while at the same time model the nonlinear complexities of the data center to accurately predict its energy consumption. An advantage of this approach is that key factors affecting the energy performance can be accurately extracted; allowing for a better understanding of the system. It is anticipated that such an approach would produce superior results than when either models were used on their own.
Primary authors: Mr. FOO, Yong Wee (University of Glasgow)
Co-authors: Dr. GOH, Cindy (University of Glasgow); Prof. LI, Yun (University of Glasgow)
Presenter: Mr. FOO, Yong Wee (University of Glasgow)

Session classification: Highly Distributed Computing System II
Track classification: Highly Distributed Computing Systems
Type: Oral