March 01, 2016 | By Tiffany Fox

ECE Student, Professor Win Best Paper Award at NIPS Conference

Together with his advisor, Professor Alon Orlitsky, Ananda Theertha Suresh, a Ph.D. student in Electrical and Computer Engineering at the University of California, San Diego, won the Best Paper award at the 29th Annual Conference on Neural Information Processing Systems (NIPS). NIPS, which took place this year in Montreal, is an international gathering of researchers in the fields of machine learning and computational neuroscience.

Alon Orlitsky

Suresh and Orlitsky beat out more than 1,800 submissions from researchers around the world to receive the honor for their paper titled "Competitive Distribution Estimation: Why is Good-Turing Good?" Good-Turing refers to a statistical technique developed by British computer scientist Alan Turing and his assistant I.J. Good as part of their efforts to crack German encryption codes during World War II. The technique addresses an uncertainty problem called discrete distribution estimation, and makes it possible to estimate the probability of encountering an object not previously encountered, given a set of past observations of objects with different qualities.

Ananda Theertha Suresh

By way of explanation, Suresh provides a theoretical example: "Imagine you're trying to calculate the best possible vehicle to buy, one that will move quickly over all types of terrain— snow, desert, roads, etc. You will only be buying one vehicle, so you need to optimize for a wide distribution of probabilities.

"With the prior computational approach, known as the MinMax algorithm, you would look at a range of 100 miles and minimize the time it would take to travel on all types of terrain by different types of vehicles. In other words," he continues, "you're trying to minimize how long it will take in a worst-case scenario. Eventually, you'd calculate that the best possible vehicle

would be a tank, since a tank can travel on these worst-case scenario terrains, but a tank is not practical because it can't travel well on the road, which is where most people are driving. Thus, with the MinMax approach, you lose functionality."

Another example is a simple coin toss. "You have a coin and you want to find out the probability of heads" explains Suresh. "In a coin toss, the worst-case distribution for heads would be close to zero, but in reality, you're not likely to have a worst-case distribution. This is often reflected in practice, where the best MinMax optimal estimator performs not as well as some heuristics."

In their paper (which Suresh presents at the conference here), Suresh and Orlitsky showed that instead of favoring the MinMax approach, where one optimizes for the worst-case distribution, the Good-Turing frequency estimation can be used to optimize for every distribution, not the just worst-case.

"So if the MinMax estimator calculates you should have a tank, the Good-Turing estimator is like having a Transformer, where whatever situation you're in, it adapts to it."

Applications for the Good-Turing estimator include ecology (estimating the probability of animal populations with a certain physical trait, for example) and a field known as natural language processing, which is a branch of computer science concerned with interactions between computers and human (natural) languages. The "autocorrect" function on a smartphone is one example of a probability estimator at work, where the computer tries to estimate which word the user is trying to type, given a set of past observations.

"Good and Turing came up with this estimator in World War II and people have been using it for more than 60 years," says Orlitsky, who holds dual appointments in the departments of Electrical and Computer Engineering and Computer Science and Engineering. "This paper provides perhaps the clearest explanation to date as to why it works."