Speaker: Jan Vondrak, *Stanford Mathematics*

**Title:** A concentration inequality for generalization error of uniformly stable algorithms

**Abstract:**

Generalization error is a fundamental issue in machine learning. How much can the performance of a learning algorithm on previously unseen data differ from its performance on the data it was trained on? It has been known that “stability” is a beneficial property in this context: a learning algorithm is uniformly stable, if it is not too sensitive to any particular data point. The study of generalization error for uniformly stable algorithms boils down to the following question. Suppose that $f(x_1, \ldots, x_n, z)$ is a function with range $[0, R]$, where $x_1, \ldots, x_n, z$ are independent random variables, and each argument $x_i$ can affect the value of $f$ by at most $\gamma$. Let $Z = g(x_1, \ldots, x_n) = E_z[f(x_1, \ldots, x_n, z)] - (1/n) \sum_{i=1}^n f(x_1, \ldots, x_n, x_i)$. How well is $Z$ concentrated?

The best previous result [Bousquet–Elisseeff 2002] was that $Z$ is sub-gaussian with standard deviation $O(\gamma \sqrt{n} + R/\sqrt{n})$. We prove that $Z$ is in fact concentrated with standard deviation $O(\gamma \log^2 n + R/\sqrt{n})$, which is optimal except for the $\log^2 n$ factor. This implies a near-optimal bound on the generalization error of uniformly stable algorithms, in particular providing new results for stochastic convex optimization algorithms.

This is joint work with Vitaly Feldman (Google Brain).