Information Theory/Coding Theory/High-dimensional Statistics

Sidharth Jaggi

My research has focused on an array of problems in Information and Data Sciences viewed through the lens of Information Theory, with an emphasis on both deriving fundamental performance limits and also on designing algorithms approaching these fundamental limits. A particular focus is on information storage/communication/processing systems that may be under attack by eavesdropping and/or jamming malicious parties – the tools I have helped develop over the last two decades provide unconditional information-theoretic security guarantees (independent of cryptographic security guarantees often considered in security scenarios, which usually rely on computational hardness assumptions that are sometimes fragile).

My research group calls itself the CAN-DO-IT team: Codes, Algorithms, Networks – Design and Optimization for Information Theory. The work that I and collaborators focus on lies in the intersection of and impacts, among other fields, Information Theory, Coding theory, algorithm design, high-dimensional geometry, estimation theory, and optimization. Despite the tools of my trade being mathematically abstract and theoretical, they have tangible real-world implications and a broad range of data-driven applications, such as for large-scale data processing, secure distributed computing, and robust distributed data storage.

Specific projects I am involved in, and which may potentially lead to research projects (feel free to email me), include:

Covert & stealthy communication: The task of communication, of talking and being heard and understood, is a fundamental human need. But sometimes one is expected to be silent, and the consequences of being detected communicating can be severe. Covert and stealthy communication consider the problem of communicating to one's intended recipients in a manner such that even the fact of communication (not just the contents) is undetectable. Using the tools of information theory, coding theory, and cryptography, one can pose this problem mathematically (for instance, steganography falls into this framework, as do cognitive radio systems). The goal of this project will be to characterize whether or not such communication can happen in various scenarios (and if so, derive fundamental limits on how much), and design computationally efficient coding schemes.

Adversarial communication: Communication in the presence of a malicious adversary who wishes to jam communications can be viewed as a high-dimensional packing problem – the transmission can be viewed as a point in a suitable high-dimensional space, perturbed by the jamming noise into a "noise ball"; hence the problem is equivalent to packing these noise balls. Recent work in information theory, coding theory, and theoretical computer science has led to insights into the fundamental limits of such packing problems, with applications in both the fundamentals of high-dimensional geometry, and also in applications such as the design of robust data storage, processing, and communication systems. In this project we will investigate the impact of adversarial limitations (such as the adversary's lack of perfect knowledge of the transmission, or computational bounds on the adversary's actions), or of additional resources available to the communication system (such as feedback).

Estimating sparse patterns from sparse data: Consider, first, the problem of group-testing, which aims to identify the (hopefully) small number of individuals carrying a disease in a large population via as few "group" tests as possible (potentially due to lack of population-scale testing resources); these group tests involve pooling together samples from different subsets of individuals into a small number of pooled tests (each pooled test has a positive test outcome if and only if at least one individual whose sample was included was a disease carrier) and inferring the set of diseased individuals from the set of test outcomes. This problem of group-testing is one example of nonlinear sparse signal estimation, wherein a sparse (mostly zero) input is to be inferred from a small number of outputs, where the input-output relationship is non-linear. Fundamental limits (on the minimum number of tests required for reliable estimation) for the classical group-testing problem have only recently been obtained; in this project we will investigate how to translate insights from these recent works to broad classes of estimation problems.