EMINA SOLJANIN, Rutgers University

Multiple Concurrent (Local) Data Access with Codes

Distributed storage systems strive to maximize the number of concurrent data access requests they can support with fixed resources. Replicating data objects according to their relative popularity and access volume helps achieve this goal. However, these quantities are often unpredictable. In emerging applications such as edge computing, even the expected numbers of users and their data interests extensively fluctuate, and data storage schemes should support such dynamics. Erasure-coding has emerged as an efficient and robust form of redundant storage. In erasure-coded models, data objects are elements of a finite field. Each node in the system stores one or more linear combinations of data objects. This talk asks 1) which data access rates an erasure-coded system can support and 2) which codes can support a specified region of access rates. We will address these questions by formulating them as some known and some new combinatorial optimization problems on graphs. We will explain connections with batch codes. This talk will also describe how, instead of a combinatorial, one can adopt a geometric approach to the problem.