Electrical & Computer ENGINEERING

ENERGY AND INFORMATION SEMINAR

eis.ece.cmu.edu

Thurs, Nov. 10

Hamerschlag Hall Room 1107 at 12:30 p.m.

Viveck Cadambe

Pennsylvania State University

Erasure Coding for Consistent Distributed Storage

Viveck Cadambe is an Assistant Professor in the Department of Electrical Engineering at Pennsylvania State University. His research involves understanding of data storage and communication using the tools of information theory and coding theory. His interests include applications to wireless communication networks, distributed storage systems such as data centers, and to distributed computing.

Dr. Cadambe received his B.Tech and M.Tech from IIT Madras in 2006. He received a Ph.D from the University of California, Irvine in 2011. He was a postdoctoral researcher in the Research Laboratory of Electronics (RLE) at the Massachusetts Institute of Technology (MIT) between 2011 and 2014. He also held a position as a visiting researcher in the Department of Electrical and Computer Engineering at Boston University between 2011 and 2014. He was an intern at the Communications, Collaboration and Systems Group at Microsoft Research, Redmond WA during June-September of 2010.

Dr. Cadambe is a recipient of the 2014 IEEE International Symposium on Network Computing and Applications (NCA) Best Paper Award, the 2009 IEEE Information Theory Society Paper Award and the UCI Electrical Engineering and Computer Science Department Best Paper Award for 2008-09. His dissertation received the 2011 CPCC Best Dissertation Award in the UCI Electrical Engineering and Computer Science Department.

Erasure Coding for Consistent Distributed Storage

Modern data services store data in distributed storage systems. In such systems, it is important to ensure that the data is available to the users even though the system components can be unreliable, for instance, the servers can crash. In addition to reliability, a common requirement, especially in applications to databases and distributed computing systems, is the following property known as consistency: when the data is being constantly updated, a client that reads from the system should obtain the latest version of the data. The goal of this talk is to describe algorithms that provide a consistent (specifically, atomic) data service minimizing overheads in terms of communication and storage costs incurred by the distributed storage system.

Erasure coding is a well known tool for designing fault tolerance in distributed storage systems at a low storage cost. The classical coding-theoretic model, however, does not consider the case where there are multiple versions of the data and a consistent version of the data is desired to be decoded. In this talk, we will present an erasure coding based algorithm that provides a consistent (atomic) data service and characterize its communication and storage costs. Time-permitting, we will outline directions of future research by describing a relatively new coding-theoretic formulation with potential applications to design of consistent data services over distributed storage systems.

ECE Energy and Information Seminar Hosts

Pulkit Grover <pulkit@cmu.edu>
Marija Ilic <milic@ece.cmu.edu>
Soummya Kar <soummyak@ece.cmu.edu>
José Moura <moura@ece.cmu.edu>
Rohti Negi <negi@ece.cmu.edu>
Aswin Sankaranarayanan <saswin@ece.cmu.edu>

Student Coordinator

June Zhang <junez@andrew.cmu.edu>

Electrical & Computer ENGINEERING

Carnegie Mellon University