Finding Needle in a Haystack - Data Provenance at Internet Scale

Speaker: Wenchao Zhou
Georgetown University, Department of Computer Science
Friday, Oct 12, 2018
1:00PM - 2:00PM, T3

Abstract
Provenance is a way to answer “why” questions about computations. It has found a number of uses in the database community, such as debugging query answers or tracing unexpected results to database tuples. In fact, the ability to ask “why” can be useful for a much broader range of applications. This talk summarizes our experiences over the past few years in adapting provenance for diagnostic and forensic uses in networks and distributed systems. Our work draws inspirations from database provenance, yet the deployment scale, use cases, and the distributed nature of networks require a significant re-design of traditional data provenance models. We review the challenges and lessons we encountered in a number of use cases, ranging from investigating intrusions to diagnosing (and even automatically fixing) distributed systems.

Biography
Dr. Wenchao Zhou is an Associate Professor in the Computer Science Department at Georgetown University. He received the BSE degree in computer science from Tsinghua University in 2006, and the MSE and PhD degrees in computer science both from the University of Pennsylvania in 2009 and 2012 respectively. He is a recipient of the NSF CAREER Award. His research interests center on the use of data-centric and formal techniques towards ensuring safe and secure distributed systems.