



**The George Washington University  
Department of Computer Science  
Colloquium**

**February 28, 2008, 11:00am  
Room 736 Academic Center**

**Nan Zhang, Ph.D.  
University of Texas at Arlington**

**Protecting Privacy in Data Mining Systems**

**ABSTRACT**

Data mining has been successfully applied to support a variety of applications, including marketing, medical diagnosis, and homeland security. Mining data without violating the privacy of data being mined, however, is still a critical challenge. Emerging privacy legislation, such as the Health Insurance Portability and Accountability Act (HIPAA), as well as the heightened public concerns about privacy protection, require immediate and resolute attention from the computing community on the protection of private information in data mining. This talk explains some of the author's ongoing research projects in privacy-preserving data mining. He will first provide a brief overview of the baseline architecture and design principles of privacy-preserving data mining systems. After that, he will discuss his recent results for the key components of privacy-preserving data mining systems, including data collection, inference control, data publishing, and information sharing. No prior knowledge of data mining is required for the talk.

**BIOGRAPHY**

**Dr. Nan Zhang** is an Assistant Professor at the Computer Science and Engineering Department of the University of Texas at Arlington. He received the B.S. degree from Peking University in 2001 and the Ph.D. degree from Texas A&M University in 2006, both in computer science. His current research interests span security and privacy issues in databases, data mining, and computer networks, including privacy and anonymity in data collection, publishing, and sharing, privacy-preserving data mining, and wireless network security and privacy. He received the NSF CAREER award in 2008.