Cybercasing the Joint: Language Technologies, Multimedia Retrieval, and Online Privacy

In this talk, I present recent case studies that highlight the potential for (multimedia) retrieval of online (social network) data to support real-world attacks. Both language-based and multimedia-based retrieval has rapidly emerged as a field with highly useful applications in many different domains. Researchers from different areas in signal processing and computer science have invested significant effort into the development of convenient and efficient retrieval mechanisms. While retrieval speed, flexibility, and accuracy are still research problems, this talk will demonstrate that they are not the only ones. This talk aims to raise awareness for a rapidly emerging privacy threat that we termed "cybercasing": leveraging information available online to mount real-world attacks. Based on the initial example of geo-tagging, I will show that while users typically realize that sharing information, e.g., on social networks, has some implications for their privacy, many users 1) are unaware of the full scope of the threat they face when doing so, and 2) often do not even realize when they publish such information. The threat is elevated by recent developments that make systematic search for information (either posted by humans or by sensors) and inference from multiple sources easier than ever before. However, even with relatively high error rates, retrieval techniques can be used effectively for different real-world attacks by using "lop-sided" tuning; for example by favoring low false alarm rates over high hit rates when scanning for potential victims to attack. This talk presents a set of scenarios demonstrating how easy it is to correlate data, especially those based on location information, with corresponding publicly available information for compromising a victim's privacy.

Bio: Dr. Gerald Friedland is a senior research scientist at the International Computer Science Institute, a private lab affiliated with the University of California, Berkeley, where he leads multimedia content analysis research, mostly focusing on ("non-speech, non-music") acoustic techniques as an aid for video analysis. He is currently leading a group of 6 multimedia researchers supported by NSF, DARPA, IARPA, and industry grants. Gerald has published more than 100 peer-reviewed articles in conferences, journals, and books and is currently authoring a new textbook on multimedia computing together with Dr. Ramesh Jain. Gerald co-founded the IEEE International Conference on Semantic Computing and is a proud founder and program director of the IEEE International Summer School on Semantic Computing at UC Berkeley. He is associate editor for ACM Transactions on Multimedia Computing, Communications, and Applications, is in the organization committee of ACM Multimedia 2011, 2012, and 2014. He is also serves as TPC Co-Chair of IEEE ICME 2012. He is the recipient of several research and industry recognitions, among them the European Academic Software Award and the Multimedia Entrepreneur Award by the German Federal Department of Economics. Most recently, he lead the team that won the ACM Multimedia Grand Challenge in 2009. Gerald received his doctorate (summa cum laude) and master's degree in computer science from Freie Universitaet Berlin, Germany, in 2002 and 2006, respectively.

*Find more information about the LTI colloquium at: http://www.cs.cmu.edu/~roni/11700/*