Panel 1: Data Privacy and Security

Thursday, October 27, 2016 at 9:30 a.m.

University of Michigan Law School, Hutchins Hall 100

Moderator:
Alfred O. Hero, John H. Holland Distinguished University Professor and R. Jamison and Betty Williams Professor of Engineering, University of Michigan; Co-Director, Michigan Institute for Data Science (MIDAS)

Panelists:
John Carlson, Vice-Chair, Financial Services Sector Coordinating Council (FSSCC); Chief of Staff, The Financial Services Information Sharing & Analysis Center (FS-ISAC)

Jonathan Katz, Professor of Computer Science, University of Maryland; Director, Maryland Cybersecurity Center

Michael Reitblat, Co-Founder & CEO, Forter, Inc.

Peter Swire, Nancy J. and Lawrence P. Huang Professor of Law and Ethics, Georgia Institute of Technology Scheller College of Business

As the number and scope of available data sets proliferate, can we help ensure that proprietary data remains private and secure on one hand, and usable and shareable within and among organizations on the other hand? This panel will explore challenges of financial data privacy and security in an increasingly interconnected world.

Narrative and questions

The international financial system uses massive amounts of data from a wide variety of public and private sources, including: corporate reports; public transactions; finance newsfeeds; securities and exchange data; summarized economic data; personal credit records; personal social network data; and other databases and data streams. By correlating public data with private data, companies can use Big Data analytics to learn about their markets, improve customer service, and better compete. As examples, Big Data enables improvements in areas including:

- Cybersecurity (early detection of attacks and intrusions on financial datacenters);
- Suspicious activity detection (fraudulent transactions or insider trading);

- Investment analytics (automated portfolio design, risk management, electronic trading);
- Credit analytics (customer profiling and actuarial statistical analysis);
- Customer development (targeted advertising and clickstream analytics).

Sharing information derived from Big Data can make the industry healthier, improving collective awareness of trends, opportunities and dangers. However, the lack of information-security safeguards will put individuals and institutions at risk to unauthorized disclosure of sensitive information. Risks range from theft of a company's proprietary technology to breach of privacy of a client's personal data. As data becomes more voluminous, diverse, and inter-connectable, securing the data becomes more difficult - creating unique challenges for the finance industry. These include: sharing information across the industry while protecting proprietary data; providing enhanced financial services while meeting the privacy expectations of clients and others from whom the industry collects information; and enabling regulators to monitor and assess companies from reported data. Some specific challenges are highlighted below.

Ensuring security in the era of Big Data. It has been common practice to protect information by maintaining all data behind a thick firewall, guarding against disclosure by using information security methods such as hashing and strong encryption. However, this approach has become unworkable since all the data that is potentially relevant to a "Big Data" analysis is too big to be assembled and protected in a single place. For example, a financial analyst may wish to use customized data mining methods to correlate parts of his company's proprietary data with data on the cloud that might be relevant to an investment or business opportunity. This exposure to the network creates new security vulnerabilities such as interception of data packets or vulnerability to attacks, e.g., man-in-the-middle (MIM). Indeed, a clever observer of the stream of queries to the cloud may be able to infer sensitive information. Financial networks are especially vulnerable to interception and disruption due to their high value to the attacker, their high volume of traffic, and their global scale.

This raises several questions. 1) What are the emerging security vulnerabilities in financial networks as they become increasingly diverse and widely distributed? 2) What existing cybersecurity tools are available to address these vulnerabilities? 3) What new tools need to be developed? 4) As the tools become more complex what will their impact be on real-time networks supporting rapid trading, credit card verification, and other financial services? 4) What are the differences, if any, in cybersecurity requirements for consumer banking, investment banking, and their supporting data infrastructures? 5) How can data sharing between companies be encouraged while protecting proprietary information?

Protecting privacy in the era of Big Data. Individuals and institutions who entrust their data to an organization expect their private information to be protected. However, recent surveys of the American public suggest that there is low confidence in the ability of organizations to guarantee privacy of their data. Notably, only 38% of the respondents stated that have confidence that their credit card companies will ensure privacy and security of their personal credit card activity records [Pew 2015]. This proportion of confident respondents

drops to 16% and 11%, respectively, when asked about their trust of search engine providers and social media sites. As personal information becomes increasingly shared, bought and sold as a commodity, incidents of unauthorized disclosure are likely to grow. Such incidents will further erode public confidence in privacy of their information.

While a major concern, identity theft by cyber-intrusion is only one of the privacy challenges faced by the public in the era of Big Data. This era raises new privacy risks that have been collectively denoted the "4 R challenges of Big Data" in [Steinmann 2016]: Reuse, Repurposing, Recombination, and Reanalysis. Reuse, repurposing and reanalysis risks arise when data that is collected with prior consent for one purpose are analyzed for another purpose that may cause harm to the individual. Recombination denotes the risk of personal re-identification achieved by combining collected data with information from other data sources. By launching an "inference attack" on the data, recombination can succeed, at least partially, even when the data is summarized, anonymized and encrypted." Recent research on new methods of privacy protection, e.g., differential privacy [Dwork 2008], shows that an individual's information can be protected from recombination with some loss in data fidelity. However, associated losses in data quality may decrease data utility for the organization and its data-sharing partners.

This raises several questions. 1) What are the "privacy-at-risk" data sharing practices in the investment, commercial banking and financial services sectors? 2) How can these risks be reduced and consumer confidence be improved? 3) Should the banking industry consider adopting IRB-style broad consent regulations governing the use and sharing of personal data? 4) Should the consumer have more control on how their financial data will be anonymized, used and shared? 5) What new categories of personal data are emerging, e.g., personal accessory data (IoT) and health data (wearable sensors), that could be properly used or shared?

[Pew 2015] M. Madden and L. Rainie, "American's attitudes about privacy, security and surveillance," Pew Research 6, May 20 2015, http://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/

[Steinmann 2016] M. Steinmann, S.A. Matei and J. Collmann, "A theoretical framework for ethical reflection in Big Data research," in Ethical Reasoning in Big Data, J. Collmann and S.A. Matei (eds), Springer series in Computational Social Sciences, pp. 11-27.

[Dwork 2008] C. Dwork, "Differential privacy: a survey of results," Theory and Applications of Models of Computation, vol. 4978, pp. 1-19, April 2008. https://www.microsoft.com/en-us/research/publication/differential-privacy-a-survey-of-results