ENEE 759D
Fall 2013

CyberSecurity Data Science

Course Goals:
Targeted cyber attacks are increasingly sophisticated, while traditional security technologies (e.g. firewalls, password-protection systems, or other passive measures) have limited utility against skilled and persistent targeted hackers. Today, many organizations are seeking data scientists, who are able to use Big Data techniques for identifying threats and attacks. This specialized field demands skills and knowledge in multiple areas, including (a) systems, to develop the technologies needed to store and process massive data sets; (b) data analytics, to extract information from these data sets; and (c) security, to ask the right questions about cyber attacks.

This course will provide an introduction to cybersecurity data science from the perspective of the security expert. We will focus on the current tends in cyber attacks, we will survey some of the recent literature and we will get hands-on experience with using data analysis systems and techniques for solving security problems.

Course prerequisites
One of ENEE459C or CMSC 414, one of ENEE459M or CMSC 422, or permission from the instructor (see below for topic prerequisites).

Topic prerequisites

- Good programming skills and exposure to some of the languages and systems used for data analysis, such as Python, R, Weka, Matlab, Hadoop.
- Familiarity with basic statistics and machine learning concepts, such as regression, sampling, clustering, classification.
- Familiarity with basic security concepts, such as authorization, authentication, buffer overflow exploits.

If you are concerned about these prerequisites please contact the instructor for guidance. We want everyone who is serious about taking this course to get in; we just don't want students to be lost due to lack of background.

Textbook:
No required textbook. *Reading materials will be provided on the course website and/or distributed in class.*

Course Topics

- Vulnerabilities and exploits
- Failures of cryptosystems
- Worms
- Denial of service
- Botnets
- Spam infrastructures
- Pay per install
- Attacks against physical infrastructure
- Targeted attacks
- Economic implications of cybercrime

Course structure
This is a graduate-level course, designed to emphasize skills that are critical for succeeding in a Ph.D. program:

- Ability to understand and interpret scholarly publications, to explain their key ideas, and to provide constructive feedback.
- Ability to apply some of these ideas in practice.

The students will read papers on selected topics in security, summarize their strengths and weaknesses using a defined written template, and present this critique in front of the class. The students will also form teams and work on a semester-long project to investigate a security problem of their choosing using data analysis techniques. Basic knowledge of data analysis systems and techniques will be delivered through lectures.

Grading method
- Written paper critique and class discussion: 50%
- Project: 50%