COSC 4397/6397, Security Analytics
Department of Computer Science, University of Houston
Project Guidelines, Spring 2016, Multiple deadlines below

All work **must** be your own. You may discuss the topic with only the other students in class and the instructor, but you cannot copy materials from anyone or anywhere (journals, magazines, conferences, web, etc., this list is not exhaustive).

The task is to:

1. Select a cyber security problem that has been tackled using zero or only one or at most two of the data analytics techniques out of the four types of data analytics approaches: data mining, natural language processing, machine learning and statistics. There must be a clear plan for obtaining a dataset for the problem: either one is already available or it is feasible for you to collect it itself (I will call this "the data plan"). Deadline for selection and results of literature search (this is how you demonstrate that your problem has been studied using only 0-2 analytical methods so far): Mar 31, 12pm. You must turn in: the security problem, the dataset plan, which two techniques will be employed and a brief plan on how they would be employed ("the action plan"), and the results of the literature search (your search queries, search results and databases used in the searches).

2. If the problem is not approved for any reason (e.g. if it is not relevant to the course, or it is too well-studied, or if it is not an important problem), you must find an alternative by Monday April 4, 12pm or accept a problem from me. Hence you are strongly encouraged to choose the problem carefully and to discuss it with me well before the deadline.

3. Apply, to the approved problem, two different type of analytical approaches that have *not* been employed so far. Deadlines for these so that you can plan ahead are: April 18 and April 25, 11.59pm CDT.

4. Document all your work and share it with the rest of the class in a short presentation of 10 minutes, more on this aspect later.

Formatting. The problem, the data plan, the action plan, and the results of the literature survey should be turned in as an organized and readable document of no more than 5 pages. Note that if your problem already has 4 pages or more of references on it, it is probably too well studied. Use only one side of each sheet, single spacing and 1 inch margin on all sides with no less than 10 point font. **Staple** all sheets together. No unstapled term papers will be accepted. Follow MLA or computer science literature format for the bibliography section of the paper carefully. Remember to include your name and email address (one that you check at least once daily) on all submissions.

Some starting points. Besides the Google Scholar and the DBLP databases, the following journals and conferences are known to publish Security research and may be used as starting points for a literature survey (this is not an exhaustive list): ACM Transactions on Information and System Security, Journal of the ACM, Journal of Computer Security, IEEE Transactions on Dependable and Secure Computing, IEEE Transactions on Information Security and Forensics, International Journal of Information Security, IEEE Symposium on Security and Privacy, ACM Conference on Computer and Communications Security (CCS), USENIX Security Symposium, IEEE Computer Security Foundations Symposium (used to be a workshop in the past), Network and Distributed System Security Symposium (NDSS), Annual Computer Security Applications Conference (ACSAC), and European Symposium on Research in Computer Security (ESORICS).

N.B. Do **NOT** tear papers from books or journals in the library! The library has excellent CDROM databases and free printing facilities and low cost copying services.