CS 5435:
Security and Privacy Concepts in the Wild
Homework #4

Due: 5 Nov. 2015
100 points

Instructor: Ari Juels TA: Fan Zhang

In Homework #3, you invented algorithms for generating honeywords. In this assignment, you'll explore *attacks* against them.

Your task is to try to build attack algorithms that can distinguish between the true passwords and the honeywords in sweetword sets produced by the algorithms crafted in Homeword #3. You may use any external data / training set you like for this exercise.

As you're not being given access to the generation algorithms, you might test your attack algorithms against your own generation algorithms and against the simple algorithm at http://people.csail.mit.edu/rivest/honeywords/gen.py.

You should collaborate with the same team as in Homework #3. Your team should work together to devise a master attack algorithm for the honeywords generated by the algorithms crafted by classmates in Homework #3. You should submit your code as a team. Additionally, each team member should submit an *individually* written (at most) one-page description of your attack strategy / techniques.

As in Homework #3, you will be graded on your conceptual understanding and the quality of your ideas, primarily as reflected in your writeup. The performance of your attack algorithm will be considered, but will not be the major determinant of your grade. Again, we're interested in your conceptual understanding and ideas. That said, we will award a prize (to be determined) to the team that achieves the best attack results.

```
> python your_program.py n m filename
```

Here `filename` is a CSV file containing m sweetword sets (rows / records), each with n sweetwords (fields). The program `your_program` should output a string of m integers $x_1, x_2, \ldots, x_m$, delimited by commas. Here $x_i \in [1, n]$ represents your algorithm's guess at the true password in sweetword set i.