Name: Adwait Nadkarni

**Hard Problem Area:** Resilient Architectures

**Overall Research Goal(s):**
What do you ultimately hope to show with your research? This section can involve some jargon, but try to relate these goals to your broad impact section. Start with the larger goal(s) and narrow the scope towards your specific study (and specific goals, below).

Our prior work has been in building access control frameworks that provide isolation and protection for application-specific user data (e.g., email attachments, notes, document scans). The problem of getting policies from the user that fuel the enforcement is generally unsolved for such systems. A user-specific policy is necessary, as 1) users use diverse kinds of applications, each with their own types of data and 2) users may have different security preferences, even for the same application’s data type.

The applications that generate data often directly or indirectly gather contextual information about the data from the user. The ultimate goal of this work is to assist the user in specifying security policies for their application-specific data, by learning the policy from a small set of examples provided by the user. The goal of this study is to demonstrate the feasibility of such an approach.

**Broad Impact:**
Why is your research important? This section should avoid any technical jargon and should be meaningful to the general public. Try to keep this down to five sentences. This should be hierarchical: the broad impact decomposed into more specific impacts that connect your overall research goals to your more specific goals.

Security systems often require the user to specify the policy. Prior work has shown that users are unable to specify the policy unassisted. Our work makes the problem of policy specification tractable, by predicting the security policy from a few examples provided by the user. As a result, it makes complex security systems practical.

**Specific Research Goals:**
Lay out the steps you are going to take to achieve your overall research goal. You can get technical here.

Our specific research goals are as follows:
1) To devise an approach to train on context-policy examples given by the user, and to predict
2) To measure the accuracy of the algorithm mentioned above, in order to demonstrate feasibility, via a study with expert users.
3) To discover other challenges that may affect the feasibility of the approach, apart from the accuracy.
**Proposed Data Collection (if applicable):**
What data will you collect to answer your research goals? How will you collect it? Will it be an observational study, randomized comparative experiment, or simulation study? Include potential biases and be prepared to explain how the data will achieve your specific research goals.

Our first study involves expert users from the security lab. We will collect an initial training sample consisting of policy examples from these users.

Additionally, we also collect information on which policies are more important (e.g., sensitive), to make our predictions more precise.

Our approach involves an interactive review of the user’s examples, to find and correct errors. Our algorithm asks questions to the user, to identify user error in the initial data collection phase.

Finally, we test our algorithm with random, unlabeled samples. Users will be required to label these samples, and we will compute accuracy based on the correct predictions by our algorithm.

**Success Criteria:**
How will you determine whether you satisfied your specific and overall research goals?

The primary success criterion is the accuracy of our algorithm; i.e., our predictions must be fairly accurate.

In addition, we plan to discover future challenges in extending this approach to non-expert users, for which the criteria is as follows:

1) We discover cases of user-error in the initial training samples.
2) Where the prediction is incorrect, we should be able to identify if the fault is with the algorithm, the initial training dataset, or if it is a separate research challenge.

**Anticipated Difficulties, Limitations, and Criticisms:**
What will make the above specific research goals difficult to achieve? How do you plan on dealing with these difficulties if they arise?

1) User Error: Users may make errors in the training phase. We plan to resolve some of these errors through a manual review phase involving the user, as well as a semi-automatic phase where the algorithm will make suggestions to the user.
2) Size of the expert user base: Our initial experiments have been performed with 8 expert users. As a result, although our results are promising, we cannot make general claims. We plan to scale this up to a fairly large set of users in the future.