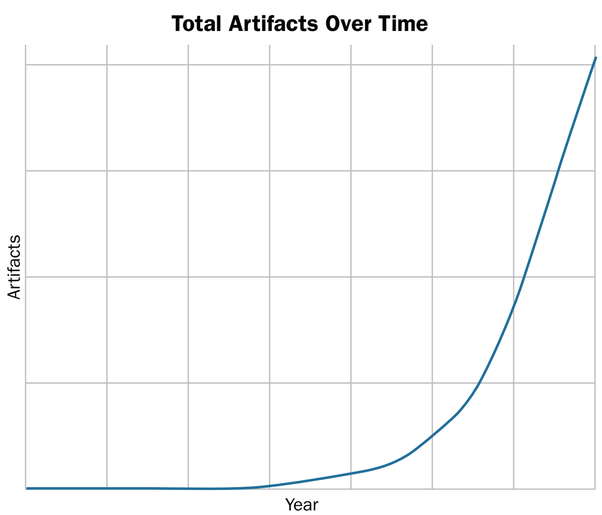
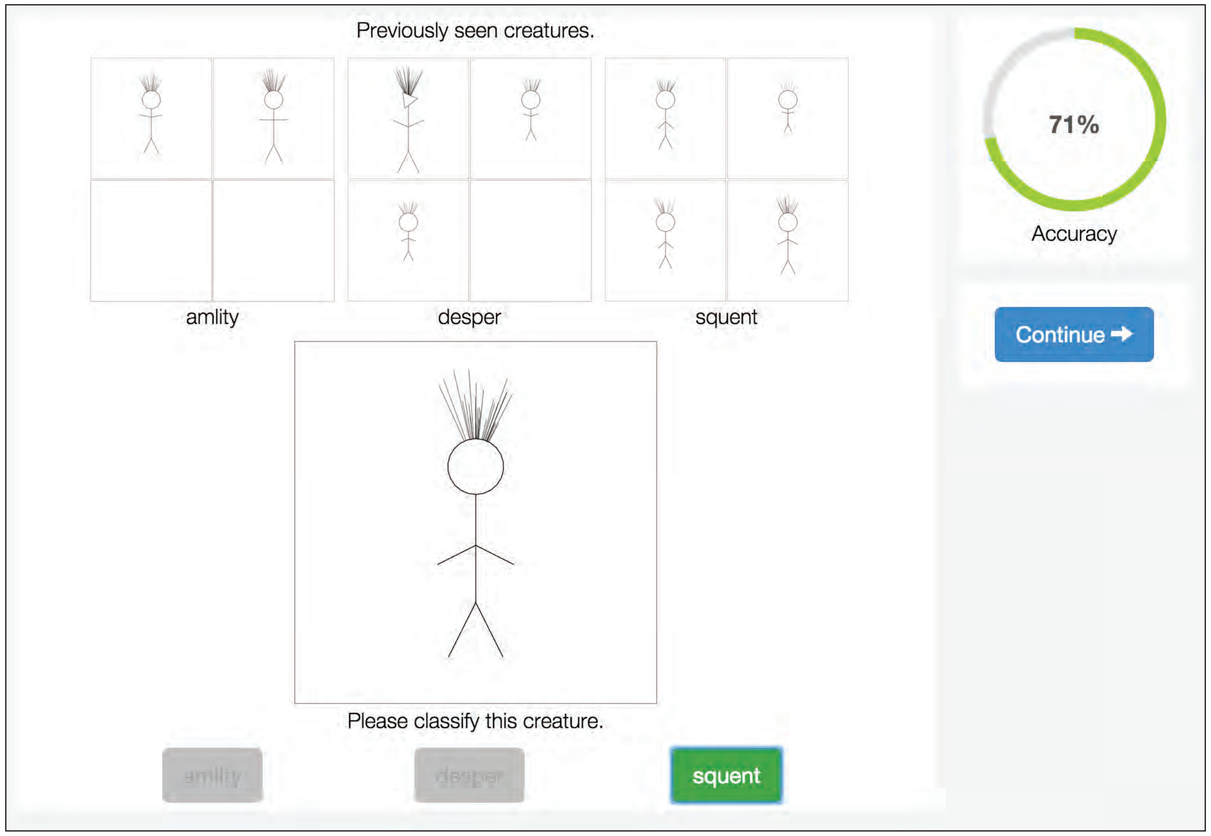
Human-Computer Decision Systems

Security decision systems aim to distinguish malicious activity from benign and often use a combination of human expert and automated analysis, including machine learning (ML). Systems using only human experts scale badly; pure ML systems are susceptible to structured attack by adversaries and, in most cases, have unsatisfactory performance on their own.

- Many operational security problems depend on a small number of skilled analysts to process a large and growing firehose of potentially malicious data.
- Traditional active learning tries to address this situation by suggesting allocation of limited analysis resources that optimize the convergence of a machine learning classifier.

Growth of CERT Artifact Catalog

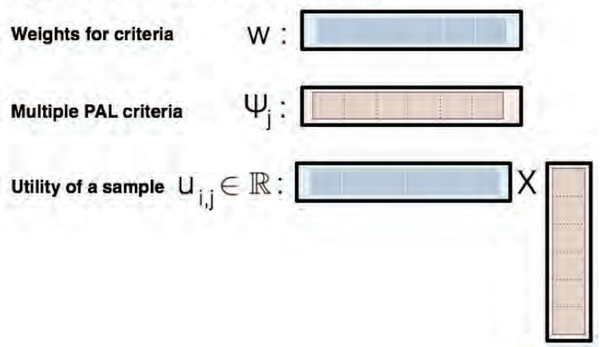




A screenshot of the experimentation system built using Mechanical Turk and Psiturk.

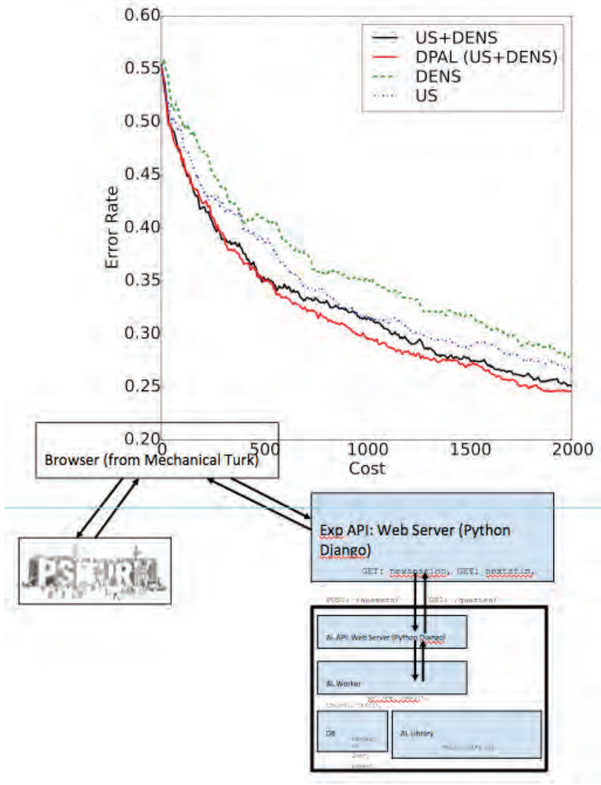
- The human-computer collaboration model will improve upon traditional active learning by optimizing not simply for convergence of the ML component, but also for future performance of the overall system, including mutable human analysts.
- We test the performance of new models not only through simulation, but also through human-subject experiments.
- Because conducting these experiments using real security analysts performing their normal tasks would be prohibitively expensive, we instead developed a proxy problem of identifying fictional creatures and leveraged non-experts on Amazon's Mechanical Turk platform. The process of generating the fictional creatures adheres to the statistical distributions of real malware classes.

Dynamic Proactive Learning



DPAL provides a framework to combine multiple factors in choosing points, including factors related to analyst performance. It shows promise in simulation and will be put to the test in a human-subject experiment.

Preliminary DPAL Results



Future work includes joint optimization of classifier and analyst objectives, extension of the experimentation software to support multi-session and team experimental trials, and a test of transferability of the model problem results to the target domain.

To keep pace with adaptive adversaries, our cybersecurity defenses must take advantage of both machine learning and human analyst strengths. Future solutions should optimize for success of the overall system.

Contact: Brian Lindauer lindauer@sei.cmu.edu

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM-0002856