FARM: Finding the Appropriate level of Realism for Modeling

Jim Blythe
Information Sciences Institute
University of Southern California
blythe@isi.edu

Sean Smith
Department of Computer Science
Dartmouth College
sws@cs.dartmouth.edu

Ross Koppel
Department of Sociology
Univ. of Pennsylvania
rkoppel@sas.upenn.edu

Christopher Novak
Department of Computer Science
Dartmouth College
novak.17@dartmouth.edu

Vijay Kothari
Department of Computer Science
Dartmouth College
vijayk@cs.dartmouth.edu

http://shucs.org

DASH MODELS OF HUMAN BEHAVIOR

Why Do We Need to Model Human Behavior?

Understanding human behavior to critique and suggest better security policies.

Security practitioners must make decisions that reflect the interplay between human behavior and systems. As our other posters show, this can be significant and complex. How can we predict it well enough to test or discover better, human-aware policies and security tools?

The role of human subject experiments.

Our understanding of human behavior in security must be rooted in observation and experimental work. However:

- Human subject experiments are too expensive to use in every case, to test every potential solution.
- This is particularly true for teams, or more than a few individuals, or organization-wide effects or the interplay between a number of organizations.

Models based on experimental data.

When it is infeasible to test security decisions with human participants before going live, a valuable approach is to build models of human behavior, allowing better informed decisions.

To develop models that are useful in practice, we focus on modeling real-world scenarios and ensuring that the models capture the relevant facets of human behavior.

We aim to base each model on relevant experimental data, by our group or others, and to validate any extensions required to fit .

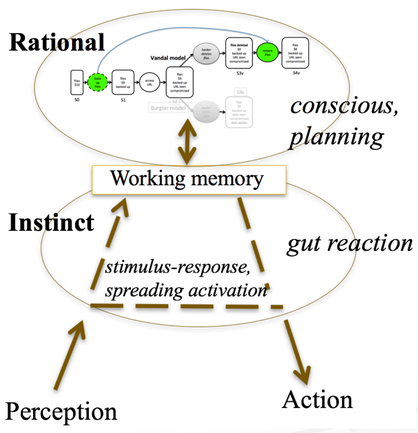
Platform for reuse of experimental data.

This approach makes experimental data more useful to practitioners. Our platform for building agent models, below, also provides a repository where this data can be available in an easily extensible, plug-and-play for for practitioners.

DASH: Dual-Process Modeling Toolkit

Our platform for agent-based modeling of security decisions is DASH: Deter Agents Simulating Humans.
DASH is inspired by work in cognitive science that models some distinguishing features of humans:

- **Dual process**: We sometimes think carefully about our next action, but more often simply follow routine.
- **Mental models**: When we think about security issues, most people do not have a good understanding of the situation but reason from analogy with physical security, or use models like healthcare (e.g. 'virus').
- **Bounded rationality**: we are affected by cognitive load, deadlines, fatigue and emotion.
- **Replanning:** When we follow a plan, we continually check that it is working and can re-plan on the fly.



A dual process model combines thoughtful and instinctive behavior

Mental models capture approximate reasoning

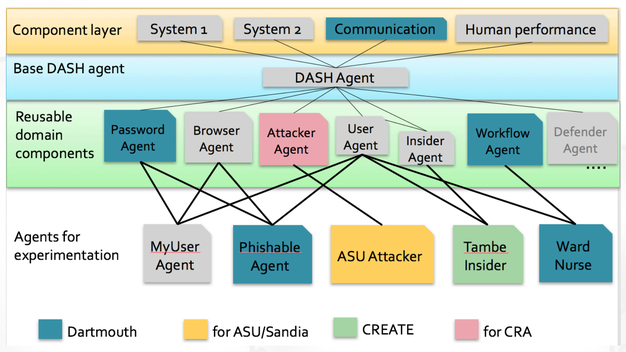
Reactive planning models human ability to adapt and re-plan.

Effects of cognitive load, fatigue etc. typically based on experimental data.

Plug-and-play module reuse

- of experimentally validated data
- of fragments of agent behavior, *e.g.* attacker, password user etc.
- of world models

Further simplifies agent modeling while supporting best practices



Recent improvements:

- Easier modeling language based on python
- Support for experiment management
- Discrete-event model supporting telescoped time
- Use in human subject experiments mixing DASH agents and subjects:

Current Experimental Work

Pilot testing a behavioral experiment to uncover circumvention behavior when humans manage passwords on multiple accounts under cognitive load and time pressure. We will explore the consequences of this behavior in our simulation.

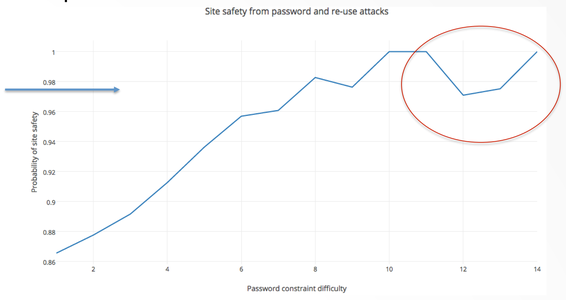
We will use the data to learn Markov models of password behavior that can be incorporated into DASH agents, improving our simulation (collaboration with U Penn)

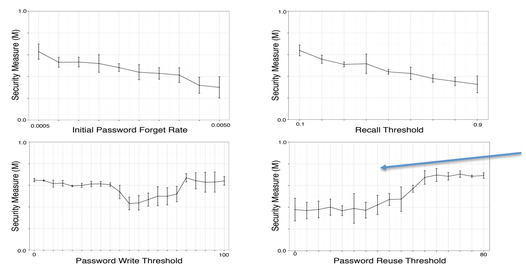
Human subject experiments with both humans and DASH agents networked on DETER, with humans alternately playing the roles of defenders or attackers, to understand human behavior in cyber attacks. (one completed in collaboration with Arizona State, one proposed with USC)

Modeling Network Effects in Password Use

What is the best password policy for an organization? We have developed and are continuing to improve a password management simulation to uncover the likely consequences and value of a policy. The simulation is also used to estimate how a policy impacts the aggregate security of a group of services.
DASH agents create and use accounts on various services and manage their passwords. The cognitive burden associated with remembering many passwords, along with a tendency to forget over time, drives users to circumvent recommended password advice. Agents use coping strategies observed in human subjects (e.g., Florencio et al., 2014); they write passwords down and reuse them between sites.

Security vs. Target Service's Password Constraints





Sensitivity analysis and comparison to the literature helps validate the model. [Korbar et al., 2016]

FARM (Finding Appropriate Realism for Modelin

What confidence can we place in simulation results?

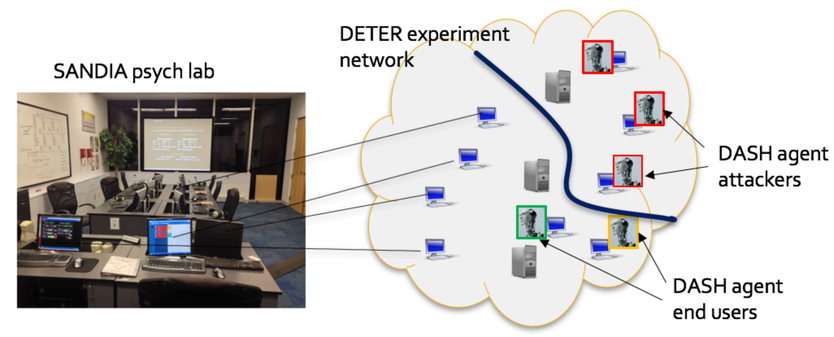
Our simulations must match reality to be informative (external validity).
FARM keeps an explicit record of experimental work that backs up parameters in the simulation, and probability distributions of possible values.

Phishable Agent
P(id phish)
P(open)
Caputo et al. 14

Aim: quantify tradeoffs in modeling and experimentation

Given data on human behavior FARM will tune agent and agent community behaviors through optimization of parameters to best match the data.
FARM can help explore how results from a simulation depend on confidence in model parameters.
will also optimize the agent specification to find the *simplest* model that fits given data well.
Given uncertain results, FARM can help pinpoint underlying parameters whose uncertainty reduction would best improve simulation results.
This approach is a step towards *evidence-based cybersecurity policy*.

Example from a human-agent experiment with defenders



Experimenter implements a phishing attack. FARM chooses the most plausible parameters given the experiment scenario by picking MLE values. E.g. 15 phish emails sent.

FARM also samples the parameter space to show ranges of derived values,
e.g. number of phish emails N: 10 < N < 30

User can interrogate the simulation model by conditional sampling, e.g. P(N <= 25),
or most plausible scenario *given* N > 20

Allows user to explore confidence levels in recommendations from simulations, and plausible alternative scenarios given experimental data.

Simulations using plug-and-play agents can be `informed' when new experimental data is available. FARM can check whether this is significant for the simulation.

http://shucs.org

HoTSoS Symposium and Bootcamp
HOT TOPICS *in the* **SCIENCE OF SECURITY**
APRIL 4-5, 2017 | HANOVER, MARYLAND