Carnegie Mellon University

Adaptive Phishing Training for Simulation Campaigns: Combining ML with Cognitive Models

JULY 18, 2023

Edward Cranford, Ph.D.
Special Faculty Researcher
Department of Psychology
Carnegie Mellon University

Carnegie Mellon University

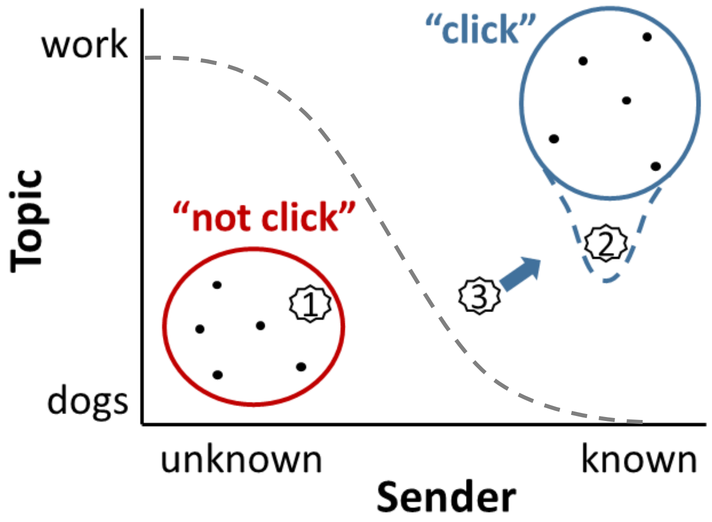
Harvard John A. Paulson School of Engineering and Applied Sciences

Drexel University

This material is based upon work supported by the Army Research Office under MURI Grant Number W911NF-17-1-0370.

Personalizing Anti-phishing Training

- Organizations typically use simulation campaigns to train employees to detect phishing emails
  - Employees selected randomly to be sent a simulated phishing email
    - If they click on the malicious link, they are given immediate feedback and training
  - Non-personalized – humans are adaptive and learn from experience
- Personalized training requires a representation of the cognitive states of each individual in the organization
  - We propose that phishing classification decisions are similar to other kinds of decisions from experience
  - Instance-based learning theory (IBLT)[1]



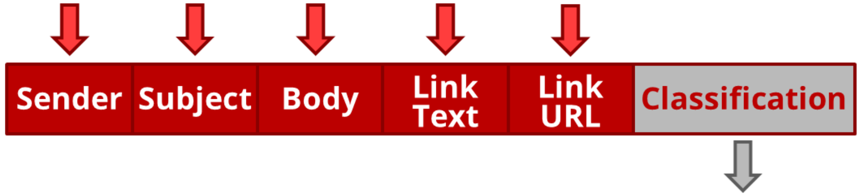
[1]Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance based learning in dynamic decision making. *Cognitive Science*, *27*(4), 591-635

Cognitive Model of Phishing Susceptibility

- Generalizable IBL model built in ACT-R cognitive architecture[1]
  - Classifications made by generalizing across past experiences in memory
    - Influenced by matching and retrieval mechanisms (i.e., *blending*[2])
      - Similarity of current instance to past instances
      - Recency of past instances
      - Frequency of past instances
  - Similarities based on the semantic similarity between email features
    - UMBC Semantic Similarity tool[3]
      - Combination of LSA and WordNet
  - Feedback
    - Classification slot changed to "Phishing" after incorrect classifications of phishing emails, prior to saving instance to memory

$$A_i = \ln \sum_{j=1}^{n} t_j^{-d} + MP * \sum_k Sim(v_k, c_k) + \varepsilon_i$$

$$P_i = \frac{e^{A_i/t}}{\sum_j e^{A_j/t}} \qquad \underset{V}{\operatorname{argmin}} \sum_i P_i \times \left(1 - Sim(V, V_i)\right)^2$$



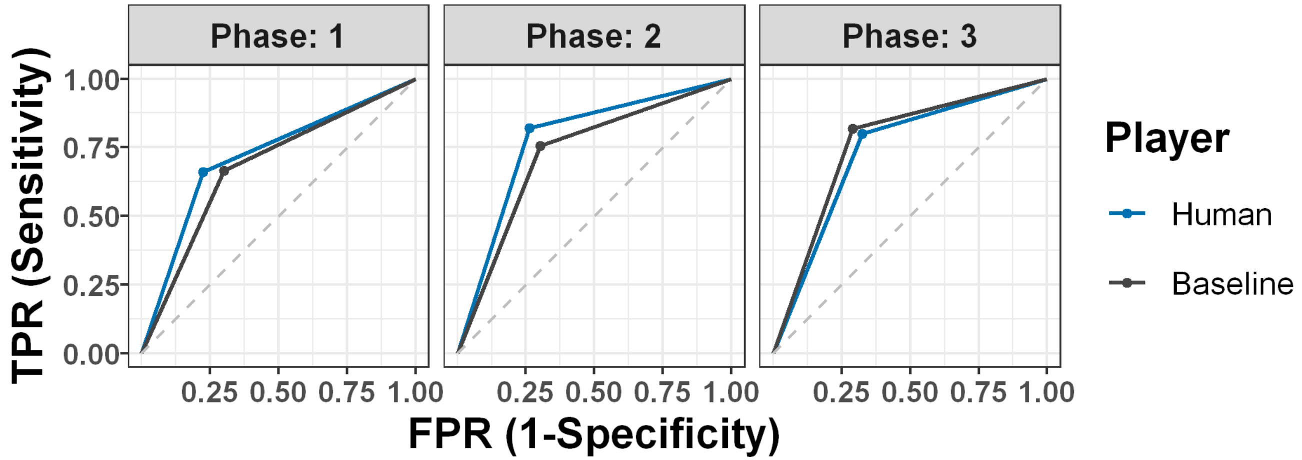
[1]Cranford, E. A., Lebiere, C., Rajivan, P., Aggarwal, P., & Gonzalez, C. (2019). Modeling cognitive dynamics in end-user response to phishing emails. In *Proceedings of the 17th Annual Meeting of the International Conference on Cognitive Modeling* (pp. 35–40). Montreal, CA.

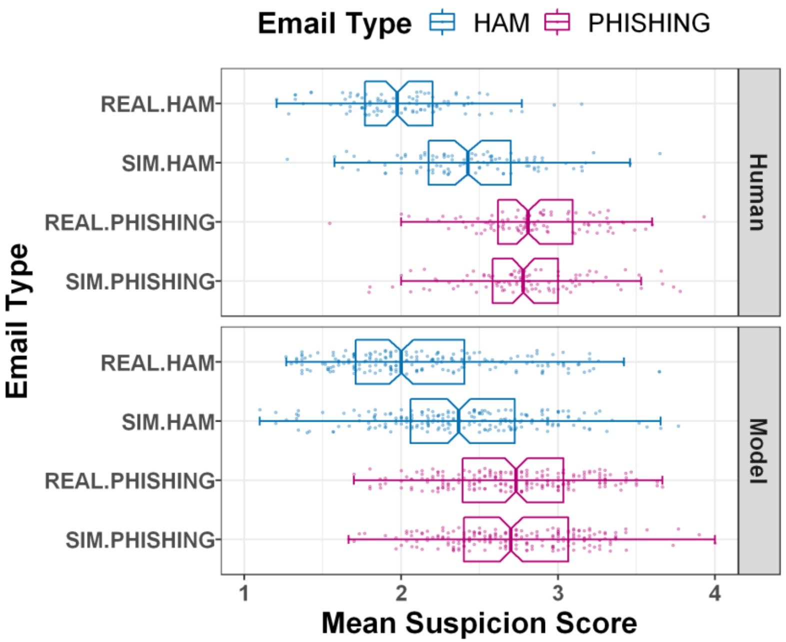
[2]Lebiere, C. (1999). A blending process for aggregate retrievals. In *Proceedings of the 6th ACT-R Workshop*. George Mason University, Fairfax, Va.

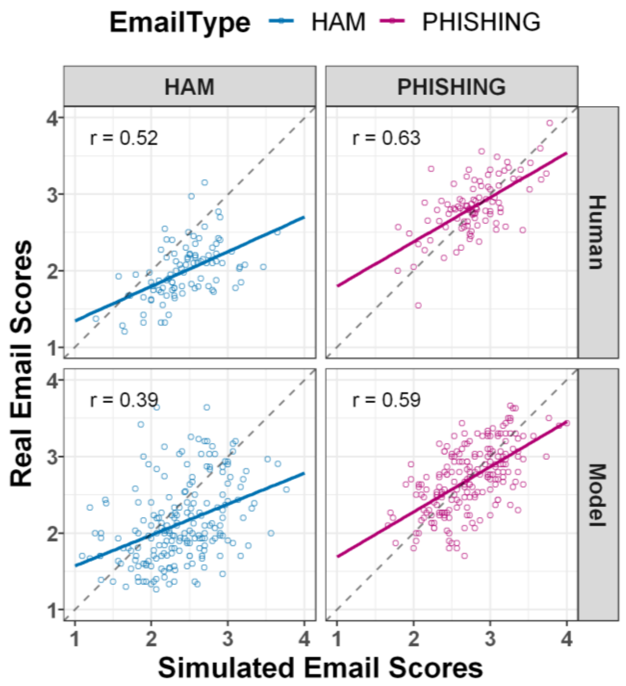
[3]Han, L., Kashyap, A. L., Finin, T., Mayfield, J., & Weese, J. (2013). UMBC_EBIQUITY-CORE: Semantic Textual Similarity Systems. In *Proceedings of the 2nd JCLCS* (pp. 44-52). Atlanta, GA.

Generalizable Model

- Same IBL model predicts human decision making across different tasks, with different databases of emails, and with different pools of participants[1]
  - Phishing Training Task - PTT (Singh et al., 2019)
    - 3 phases: Pre-Test, Training, Post-Test
  - Phishing Email Suspicion Test - PEST (Hakim et al., 2020)
    - Testing Phase only
    - Continuous suspiciousness ratings instead of binary classification decision



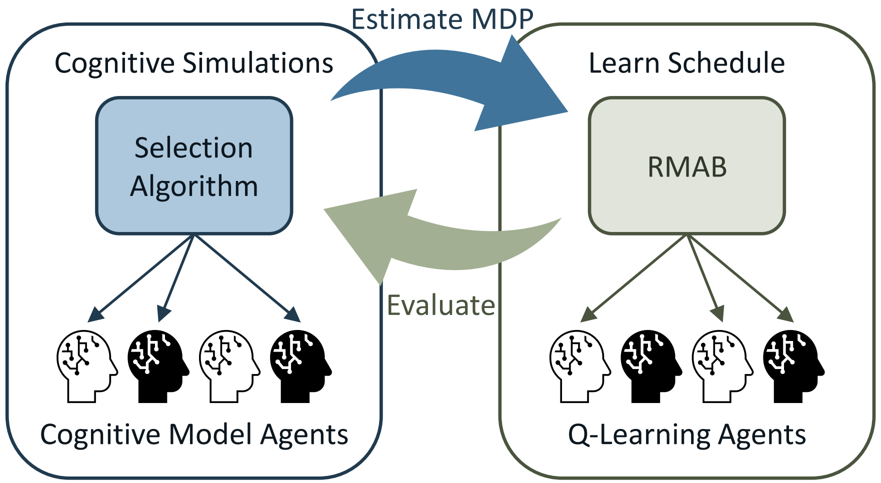




[1]Cranford, E. A., Singh, K., Aggarwal, P., Lebiere, C., & Gonzalez, C. (2021). Modeling phishing susceptibility as decisions from experience. *Proceedings of the 19th Annual Meeting of the ICCM* (pp. 44–49). Virtual.

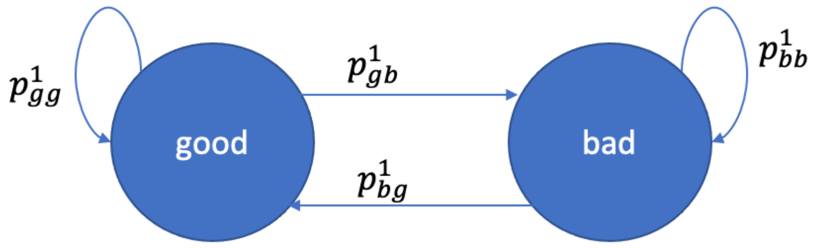
How can we strategically schedule interventions?

- Effectively a scheduling problem
  - Individuals require different amounts of training; timing is important
  - Who to target at each time step?
- We combine cognitive modeling with machine learning methods to improve training
  - Framed as a Restless Multi-Armed Bandit (RMAB)
    - Employees (i.e., arms) modeled as Markov Decision Process (MDP)
  - Cognitive model used to estimate transition probabilities
- Simulation study to compare effectiveness of solutions
  - Cognitive Model of phishing susceptibility as simulated participants
  - Presented either a phishing email (intervention) or ham email (no intervention) on each trial
    - 100 trials – 20 pre-test, 60 training, 20 post-test
    - Selection algorithm determines which users to send phishing interventions
    - Feedback provided after only after incorrect classification of phishing email



Restless Multi-Armed Bandit formulation

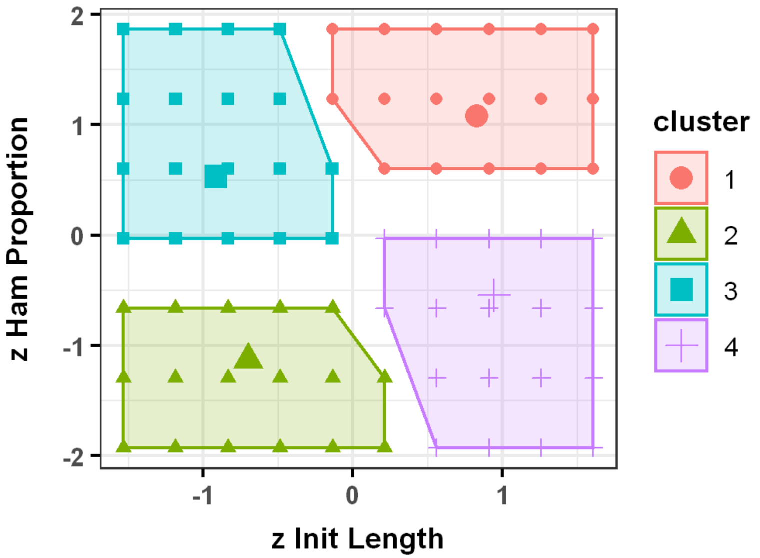
- Each round, learner selects an arm for intervention and receives feedback/reward
  - Goal to maximize total reward observed by learner
  - Budget of 20%
- Each arm (i.e., employee/user) is modeled as a Markov Decision Process
  - 2 possible **States**:
    - Good or Bad
      - Roughly, in the good state, the user always labels emails correctly and the reverse for the bad state
  - 2 possible **Actions**:
    - 1) intervention (send phishing email)
    - 2) no intervention (send ham email)
  - **Rewards**:
    - Value of being in each of the states
      - 1 in a good state and 0 in a bad state
  - **Transition Probabilities**:
    - Distribution over the possible next states given the current state and action
    - p_{gb}^1, p_{gb}^2, p_{bg}^1, and p_{bg}^2, where p_{xy}^i denote the probability of transfer from state *x* to state *y* when action *i* is taken
    - We use the Cognitive Model to estimate these transition probabilities



RMAB – User specification

- SuperArm-WIQL used to solve the RMAB
  - Reduces complexity and computational costs of learning the parameters for each user
  - SuperArm – users clustered into groups (as described earlier) and learning experiences combined
  - WIQL – Whittle Index Q-Learning[1]
    - *Q*(s, a)* values capture the quality of taking action *a* from state *s*
    - Selects users from SuperArms based on Whittle Index
      - *Q*(s, 1) – Q*(s, 2)*

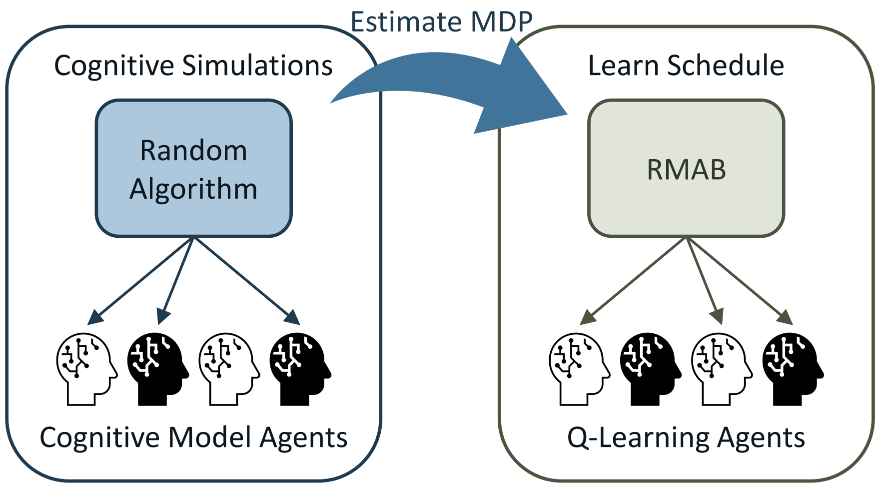
- Simulated users initialized with different amounts/types of emails
  - Represents individual differences in experience
    - ***Email usage*** (*Init Length*)
      - 10-100 emails in increments of 10
      - More emails = greater overall email usage, but new emails have less impact on learning
    - ***Phishing & network security experience*** (*Ham Prop*)
      - 70%-100% ham normally distributed (*M* = 0.85, *sd* = 0.05)
      - Fewer ham emails = greater phishing experience and greater likelihood to classify phishing emails correctly



[1]Biswas, A., Aggarwal, G., Varakantham, P., & Tambe M. (2021). Learn to intervene: An adaptive learning policy for restless bandits in application to preventive healthcare. *Proceedings of the 30th IJCAI*, (pp. 4039–4046).

RMAB – Using Cog Model to derive transition probabilities

- In the absence of data to train an MDP, we use a cognitive model to simulate data
  - A priori predictions based on constrained mechanisms resulting from a theory of cognition
- Simulated 1000 cognitive agents performing the task
  - Paired against a random selection algorithm
  - On each trial an agent is deemed in a good state if their classification of a test phishing email is correct or a bad state if their classification is incorrect
  - Transition probabilities based on the model's sequence of decisions
    - Proportion of transitions from a good or bad state at time *t* to a good state at *t+1*, depending on the action



Cluster ID	Cluster Label	p^1_{GG}	p^1_{BG}	p^2_{GG}	p^2_{BG}
1	high-high	0.783	0.610	0.659	0.458
2	low-low	0.877	0.824	0.849	0.715
3	low-high	0.824	0.645	0.738	0.461
4	high-low	0.871	0.818	0.830	0.761

Comparing alternative transition probabilities

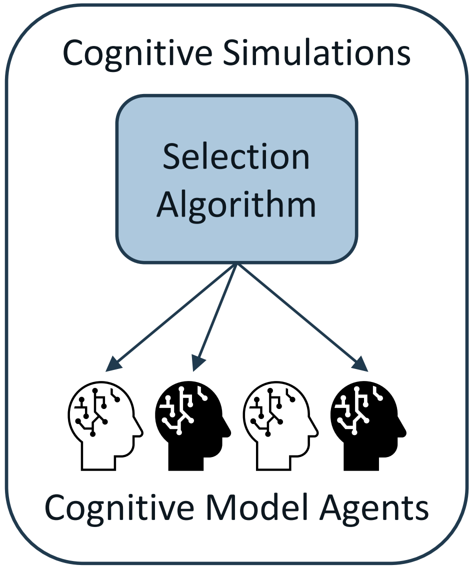
- Defining only 2 states limits effectiveness of RMAB
  - In reality, users are typically somewhere between “good” and “bad” states
  - As a first step, increase to 3 states: Good, Intermediate, Bad
    - Two test phishing emails given on each trial to determine state

Cluster ID	Cluster Label	p^1_{GG}	p^1_{GI}	p^1_{GB}	p^1_{II}	p^1_{IG}	p^1_{IB}	p^1_{BB}	p^1_{BI}	p^1_{BG}	p^2_{GG}	p^2_{GI}	p^2_{GB}	p^2_{II}	p^2_{IG}	p^2_{IB}	p^2_{BB}	p^2_{BI}	p^2_{BG}
1	**high-high**	0.663	0.287	0.050	0.377	0.520	0.104	0.248	0.423	0.329	0.516	0.367	0.117	0.414	0.359	0.228	0.472	0.348	0.180
2	**low-low**	0.793	0.191	0.016	0.278	0.682	0.040	0.167	0.258	0.575	0.754	0.218	0.028	0.304	0.623	0.073	0.238	0.393	0.369
3	**low-high**	0.737	0.236	0.027	0.360	0.536	0.105	0.236	0.429	0.334	0.635	0.296	0.070	0.395	0.407	0.199	0.495	0.340	0.164
4	**high-low**	0.793	0.188	0.019	0.259	0.711	0.031	0.049	0.343	0.608	0.711	0.253	0.035	0.326	0.612	0.063	0.119	0.379	0.502

- Static transition probabilities represent average across time horizon
  - Users learn and adapt over time, and thus transition probabilities should reflect this learning rate
  - As a first step, derive transition probabilities for each block of 20 trials
  - Using 3 states
  - Reflects improvement in phishing classification ability (and some decrease in ham classification ability) from start, to middle, to end of training phase

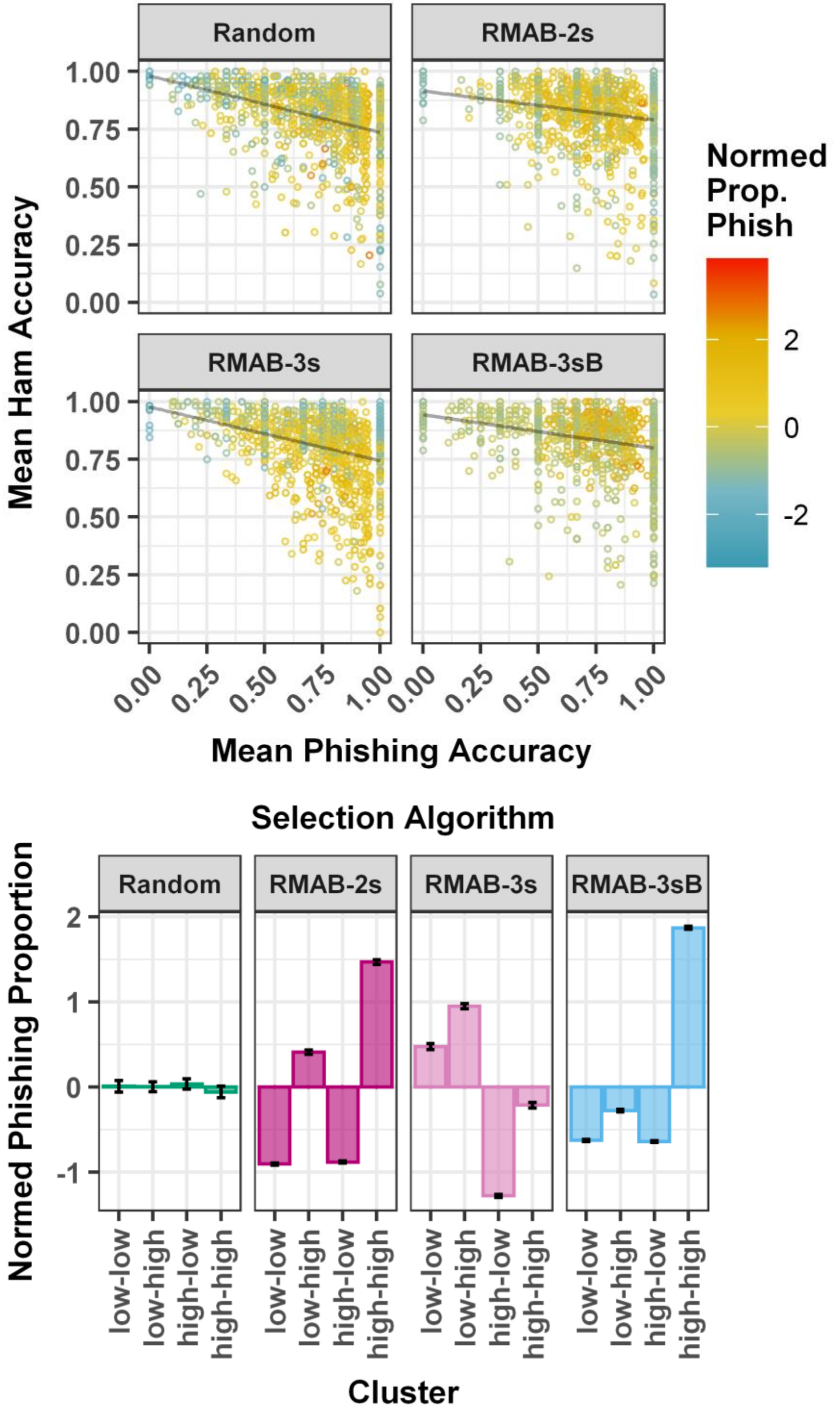
Cog Model Simulation Design

- Cognitive Model used to simulate 1000 users paired with each selection algorithm
  - Same set of initialized users for each simulation
- Multiple model agents run in parallel via ACT-R's built-in mechanism
  - On each trial, selection algorithm determines which users to send phishing intervention
  - Agents are sent appropriate type of email
  - Each agent makes a decision before moving to the next trial
- 5 selection algorithms compared
  - NoAction
  - Random
  - RMAB-2s (2 states)
  - RMAB-3s (3 states)
  - RMAB-3sB (3 states – Blocks)



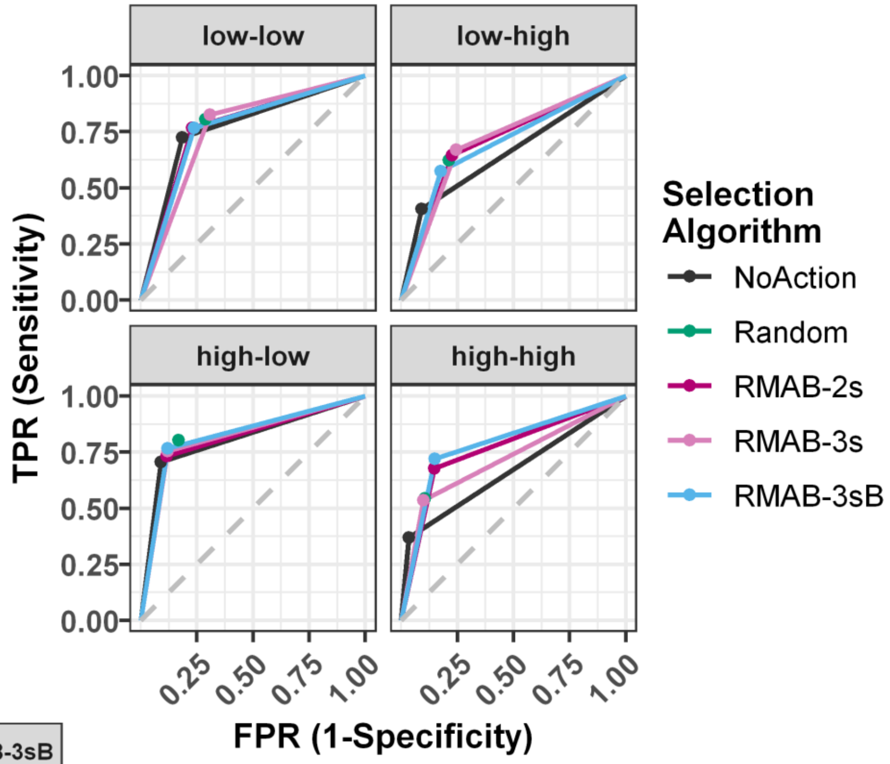
Selection Preferences

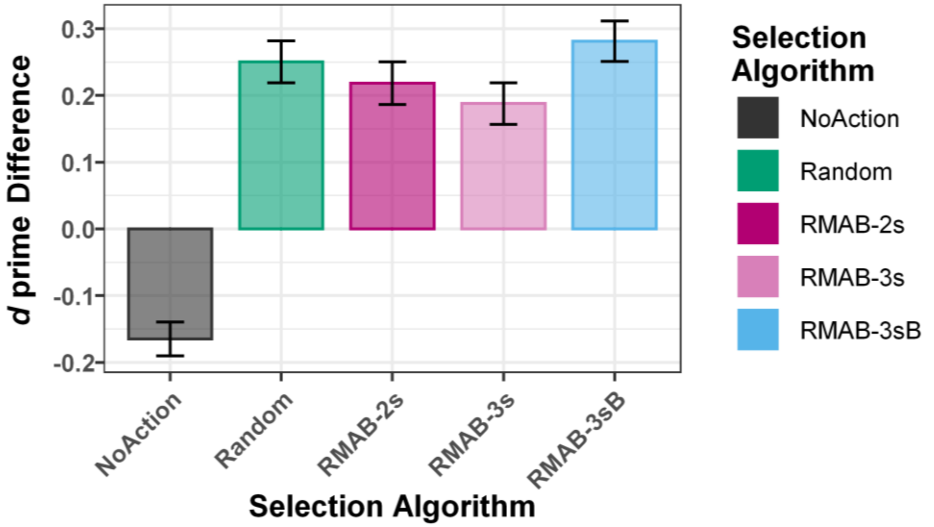
- Accuracy metrics are dependent on which users are selected
- Analysis of selection preferences reveal different patterns between selection algorithms
  - RMAB-2s tends to select users with high HamProp
    - Results in improving users who have least phishing experience
  - RMAB-3s tends to select users with low InitLength
    - Results in improving users who have least experience with emails in general (training has easier impact)
  - RMAB-3sB tends to select users with high InitLength and high HamProp
    - Results in improving those users that need the most phishing training to overcome large history of experience (recency) with ham emails and little experience with phishing emails (frequency)

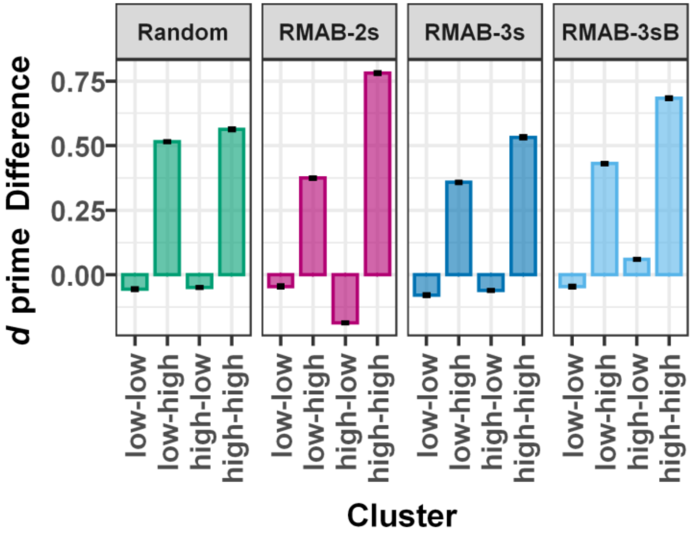


Signal Detection Measures

- Intervention improves signal detection compared to no intervention (NoAction)
  - Most impact on High HamProp groups (Low-High & High-High)
- Difference in D-prime scores from first 20 trials to last 20
  - RMAB-3sB only condition predicted to improve overall classification ability compared to Random intervention
    - Also, only condition to improve High-Low group







Conclusions

- Overall, the RMAB-3sB solution proved most successful at increasing phishing detection accuracy while minimizing false alarms across the groups
  - Likely that RMAB would do better
    - as the number of clusters approaches the number of users
    - as number of states increases
  - Future research will explore optimal tradeoff in number of clusters, number of states, and computational costs
- Future research will
  - Validate simulation results in human laboratory experiments
  - Refine the Cognitive Model based on experimental results of Random condition
  - Refine the RMAB formulation based on updated Cognitive Model
  - Explore methods of further combining the pros of cognitive models with pros of RMAB
    - e.g., Using cog model to provide additional learning rate estimations
  - Compare the RMAB formulation to purely cognitive solutions
    - Cognitive solution have lower computational overhead and the advantage of selecting users at the individual level
  - Explore methods to further personalize training by selecting specific emails based on type and content

Carnegie Mellon University

Collaborators:
Christian Lebiere
Coty Gonzalez
Shahin Jabbari
Han-Ching Ou
Milind Tambe

Questions?

cranford@cmu.edu