2.2 Application to Digital Security

We can apply this mechanism to end-user digital security. When users use a website they are at some risk of being hacked (W_r). The website itself has some level of security (S_s) that reduces W_r and some level of overall quality S_q (e.g., how long it takes for a user to login). There is a set of protective (i.e., security) behavior(s) that users can enable within the system B in the system (e.g., turning on two factor authentication (2FA)). The quality B_q of these behaviors is (e.g., how long it takes for a 2FA text to get sent) as well as the overall system quality S_q influences the cost of the behavior to the user. The protection value (B_s) is e.g., how much using 2FA reduces the risk of hacking, influences the value of the behavior to both the user and the firm.

System private parameters can be adjusted by greater engineering investment (to increase S_s, S_q, B_q, and potentially, B_s) or, for R, monetary investment (e.g., in support staff to help people learn the security behavior, provision of easier to use but more expensive tools like Ubikeys).

The User's Parameters. The user's TYPE includes their capability to do a security behavior, their typical security behavior on similar websites, their risk tolerance (e.g., how much risk of hacking they are comfortable with), and their investment in the system (e.g., if this is a bank account, how much money they have stored in the account).

The User's Utility and Goal. Where the user's goal is to maximize their utility, $max(f^u)$ over the space of behaviors that they can possibly adopt.

The Firm's Utility, Goal, and Levers. The firm sustains losses from users getting hacked (as they have to pay back users and they lose good PR), and may have costs from the level of risks of the system users (cyber-insurance plans are often priced based on estimated risk of the firm, which is based in part on user risk of hacking) and thus the firm gains utility from users adopting security behaviors. However, the firm also sustains costs from adoption of security behaviors as these behaviors may decrease user engagement or have direct costs (e.g., cost to send 2FA text message). The firm's goal is to maximize their gains over the set of users and behaviors by adjusting the m shown to the user and the true values of the systems parameters.

2.3 Constrained Mechanism: Introducing Policies

In the mechanism we have defined thus far, there are no restrictions on what the firm can communicate in m (that is, how much the firm can lie to users in order to encourage them to behave in a particular way), no requirements on fairness for equity in risk between different users of the same system, and no requirements around equity of effort or cost of behaviors to different users of the same system.

In many cases, we may wish to impose such restrictions, however. False advertising laws restrict the "lies" firms can tell consumers, and maximum tolerable risk is often defined for a multitude of publicly used systems (e.g., power plants [2]). Similarly, then, we could imagine imposing constraints such that online websites cannot put their users at more than a H% risk of being hacked.

To achieve H% may require user behavior. In this case, we may wish to impose additional policies. We may wish to impose additional policies, a fairness constraints. For example, a risk fairness constraint: all people in the system should have as equal as possible risk of a negative outcome (e.g., a user of one race should not have a greater risk of being hacked than one of another race), given the same protective behaviors. Or, effort fairness: resources should be assigned and behaviors should be created such that user variance in cost (effort) of behavior required to reach H is minimized. In this context, a policymaker wishes to set optimal policies and maximize compliance.

3 FUTURE WORK

Using data collected from prior empirical experiments (see Redmiles et al.), we can model how varying the message shown – e.g., showing different values of B_s or S_s – can alter participant behavior. In these empirical experiments, u_i creates an account in the system (bank.cs). During this process some information is collected about the user's TYPE, t_i, through observation. Users are provided a set investment in the system I. The system's private parameter S_s is set as H, the percent chance that a users' account will get hacked and they will lose their investment, I; S_q is a function of the engineering of the system: users have some average login speed. The system has a binary set of security behaviors B: enable or do not enable two-factor authentication (2FA) that users can turn on or off. B_{2FAon} reduces the risk of hacking by some protection percent ($B_s = P$), costs the user some amount of time, dependent on their type and the quality of the behavior and the system, $c(B_{2FAon}, t_i) = B_q \times t_i \times S_q$, and costs the firm some amount of money (e.g., \$0.09 cents per 2FA text message sent). For both the user and the firm, $u(B_{2FAon}, t_i) = t_i \times S_q \times S_s \times B_q \times B_s \times R$. The system attempts to adjust user behavior by revealing S_s and B_s to the user via a message m.

In our future work, we will incorporate these simple empirical-based estimations into a linear optimization framework to solve the optimization problems presented in the prior section. This will allow us to optimize over both firm and user variables to establish a set of potential optimal parameters for security nudges and policies within a particular context. Once validated, we are in the discussion stages of using these values in a live test with an end-user facing software platform.

REFERENCES

[1] [n. d.]. Back to basics: Multi-factor authentication (MFA). https://www.nist.gov/itl/tig/back-basics-multi-factor-authentication

[2] Baruch Fischhoff. 1984. *Acceptable risk.* Cambridge University Press.

[3] Nicole Kobie. 2017. Why Two-Factor Authentication Is So Important. https://www.teenvogue.com/story/why-two-factor-authentication-is-important

[4] Elissa M Redmiles, Michelle L Mazurek, and John P Dickerson. 2018. Dancing Pigs or Externalities?: Measuring the Rationality of Security Decisions. In *Proceedings of the 2018 ACM Conference on Economics and Computation.* ACM, 215–232.