Some ideas from others

Lorenzo Alvisi

- Understanding an appropriate response model for handling cybercrime from nation states
- Revisit 2003 grand challenge (esp 3)

Cormac Herley

Meaningful cyber-crime statistics

The one is close to my heart, but I keep coming back to the observation that we've been saying this for some time. See the article by Steve Bellovin and Adam Shostack, where they point out that this was one of the asks in a 1991 NRC report. https://www.cs.columbia.edu/~smb/papers/Current_and_Future_States_of_Cybersecurity-Bellovin-Shostack.pdf I think they're onto something when they suggest that figuring out why we keep failing at a goal that so many agree on is worth serious consideration. This point alone seems worthy of vigorous investigation. "We need better cybercrime stats" reminds me a bit of "we need to get rid of passwords": everyone agrees how important it is, and yet nothing happens over and over.

Cormac Herley

Having said that I do think that we can make progress on the cybercrime stats, and the current doc is surfacing some really good points: identifying durable categories that can be collected over decades time-scales, improving measurement of harm and defender effort. Given the long history of sub-optimal progress here I think it's worth considering shaking the question up some. I.e., instead of saying "if we had better cyber-crime stats we could make better decisions" tell me precisely what decisions we're currently blocked on and what data precisely we need to make them. It also might be worthwhile to prioritize the data wishlist. E.g. I think it might be profitable to think harder on what we mean by overcoming barriers to reporting and consider the possibility that some of the barriers might be too big to overcome.

Cormac Herley

I would put more emphasis on trying to discover without pre-conceptions what the barriers are, rather than supposing we know what they are as current doc does (embarrassment, loss of reputation, etc.). E.g. the current doc seems to suggest that individuals and organizations just need some encouragement and assistance with reporting. But it could be the case that for individuals law enforcement actually can't do much, so it's just a burden for no real return. For organizations it could be that you do get punished for openness and transparency (e.g., does anyone voluntarily sign up for the kind of shellacking that Yahoo took?). I have the strong feeling that the barriers and incentive stuff isn't just a question of a "raising awareness" campaign.

Cormac Herley

Preserving Individual Agency

Reading the Exec Summary is hard. I think there's a good and important grand challenge here, but it's getting totally lost. The density of abstract nouns is way too high.

I think a core of the challenge is in #2 Privacy preservation etc. The problem of how to return/retain some control over how they interact with services that log and mine everything is going to be of growing importance.

Cormac Herley

I think it's worth drilling down on one or two target problems to figure out what do users understand is going on, what are the harms they may be subject to, and what's the minimum they need to understand to be well-informed and whether it is feasible to get them to that level of understanding. This area seems like one where we will have to pick our battles. Some phenomena are too technical and some change too quickly to have bringing large user populations up to speed. Consider the questions involving encryption for example; too appreciate the possibilities you have to understand not just some idea of scrambling with secrete keys but public-key encryption, key management, certificate authorities, chains of trust. That's probably too much to realistically convey. Similarly, for understanding ads and tracking you might need to understand a lot more about how routing and browsers work than you want.

Stefan Savage

Doing the hard but necessary work to put security on a quantified footing. The cybercrime statistics proposal is a specific case of this (i.e., that we don't actually know what our costs/losses due to one particular kind of security action is) but so is the incentivization one. But similarly, we don't have good understanding of how much we spend (in time/money) or cybersecurity. We don't know how various practices (defenses, training, IT expenditures, etc) impact these costs on either side. We don't know if problems are distributed uniformly or concentrated. We don't know if there are long-term impacts of cyber events, etc...

Bottom line: for a critical issue we have much less data about cybersecurity than we do about traffic, health care, trade, etc. There are technical, social, political and economic reasons for this but I think there is a grand challenge to be had around gathering enough data that an enterprise's policy decisions about cybersecurity investments can be based on data with at least as much fidelity as a hospitals policy decisions about how to invest its resources to maximize patient outcomes.

Stefan Savage

Another challenge we have is bad information. There are multiple kinds of bad here. Studies with poor or undocumented methodologies, generalizing from point measurements when there are large variances in the overall distribution, and then plain old misinformation (aka fake news). Short of gathering information ourselves how do we convince ourselves that the information we get is good.

There are some low-level technical issues here (e.g., validating sensors, etc) but the larger problems are socio-technical. What mechanisms are most effective and helping people sift between correct data and misinformation, particularly when the difference may not be black and white and when people's own confirmation biases work against the correct outcome.

Stefan Savage

Finally, the third side of the coin is how information is understood and used. This issues is implicit in the incentives proposal, the risk proposal and the agency proposals. We are flawed beings and neither are able to understand/contextualize all the information we receive, nor are we rational actors about the data we do understand (nor is this any different in organizations). Thus, it isn't sufficient to simply put together all the best information and expect that an ideal outcome will result. This is both true in the small scale (i.e., people reusing passwords), in the medium scale (all the work showing limited efficacy in notifying system administrators of vulnerabilities in their systems) and in the large (routine underinvestment in IT security or overinvestment in solutions with no evidence-basis).

Stefan Savage: cybercrime statistics

This is a straightforward call for measurement. Its hard for me to not like it because of my own biases. However, it doesn't have that sexy "grand challenge" feel (as few measurement proposals do). However, I think that could be done by operationalizing it. For example, lets ask the larger question of what we expect out of law enforcement wrt cybersec?

Inherently it seems like there are two factors: to stop ongoing criminal actions where they are significant and to deter future criminal actions of the same flavor. Moreover, "significant" here can be either "the biggest losses" or "what citizens are most worried about".

We need better data on both crime statistics and enforcement actions to make a determination of how much LE action is doing what we expect of it – both in enforcement and in deterrence.

So a grand challenge might be something like "Can we develop the data and analysis techniques to understand and improve the effectiveness of law enforcement in stopping and deterring cybercrime?" or something like that... Such a grand challenge also lets you open things up to issues like "what does law enforcement need, what does the public expect, etc" . Alternatively, you could propose a predictive grad challenge here (e.g., where you use data to predict the growth and focus on different cybercrime waves).

Stefan Savage: incentives

This is another really important topic that needs a couple of more concrete challenges to make them more enticing. There are a few one might do:

- Build a model for effectively predicting cyber-loss risk and thus allow insurance pricing to follow risk
- Build a tool to effectively estimate the cost a breaches (losses, remediation, legal and brand damage) based on a parsimonious set of features (train on past data and test on new... the idea being to provide a tool that an organization can understand its exposure with

I think all of the points in the proposal are great for study however even if many of them don't have the sex appeal that one might otherwise like.

Stefan Savage: agency/usability

Here's one idea:

- Devise an approach or approaches (system, training, hybrid, etc) that can meaningfully protect against even targeted forms of deception (e.g., phishing, BEM scams) by modeling, anticipating or adapting to limited human mental models (i.e., negligence).

Bottom line: do something meaningful about the phishing problem by incorporating an understanding of human behavior. Phishing is a true grand challenge problem. Almost all data breaches can be traced to phishing.

Franky, if a program come out that improved the phishing problem by some significant amount it would be a huge boon.I think the "deceptive agent" thrust is captured by a bunch of the insider threat work that's been funded by DoD for years, but it certainly might be interesting to see if there is some social science take on this that makes it work better.

Stefan Savage: mapping cyber risk

The rresearch that would allow one to build a tool that could ingest an organization's e-mail and chat logs, the configuration of access rights (i.e., who was allowed to access what) and from this make an assessment of the security culture.

I can imagine a bunch of interesting ways to do this: e.g., how does It respond when new vulnerabilities disclosed, how do people talk about security trainings to each other, are security-related messages deleted, etc...