The Economics of Cybersecurity — Lecture 11 Notes

Adam Hastings

April 2, 2024

1 Game Theory Recap

1.1 Backwards Induction

2 Control Systems

What does this look like? (A system diagram! Remember HW1?)

3 Mechanism Design

3.1 The Divide-and-Choose Mechanism

- The paper begins with a description of the divide-and-choose mechanism. Can someone tell me what this is? (A mechanism for creating a fair split)
- We learned game theory last week. Let's write this out as a game:
  - Players: $\mathcal{I} = \{1, 2\}$
  - Actions:
- Does this work for more than two players? What if we wanted to split a cake three ways?
  - I could cut a cake into three pieces. What does this look like as a game?
    - Players: $\mathcal{I} = \{1, 2, 3\}$
    - Actions: $\mathcal{A} = \{a_1 = \text{Choose piece } 1, a_2 = Choosepiece2, a_3 = Choosepiece3$
    - Utility: $u_i = \text{size}(a)$

3.2 Auctions

Auctions are a classic area of study in mechanism design. Auctions are one of those things that may seem so simple you don't need to study them at all, but it turns out that there are many different types of auctions and many reasons why you might choose form over another.

3.3 Simple Auctions

Let's start with an open-air auction at, say, and art gallery:

- Bidders will place bids that are monotonically increasing. There is usually minimum "bid increment" of about 10%, so that if the previous bid is \$100, you have to bid at least \$110, not \$100.01.
- The bidding repeats until a highest bid is placed. The person who places the highest bid

3.4 Mathematical Definition

The definition of a mechanism is similar to that of a game:

- A set of players $\mathcal{I}, |\mathcal{I}| = n \in \mathbb{N}$
- A set of outcomes $\mathcal{O} = \{o_1, o_2, o_3 ... o_k\}$
- Each agent $i \in \mathcal{I}$ has private information $\theta_i \in \Theta_i$
- Each agent i has a utility function $u_i : \mathcal{O} \times \Theta_i \to \mathbb{R}$
- Utility is assumed to be of the quasi-linear form

  $$u_i(o, \theta_i) = v_i(o, \theta_i) - p_i(\theta_i)$$

  where $v_i : \mathcal{O} \times \Theta_i \to \mathbb{R}_{\geq 0}$ represents an arbitrary valuation function and $p_i \mapsto \mathbb{R}$
  - If $o \in \mathcal{O}$ is an outcome, then p_i can be thought of as the cost (or price) imposed upon player i for that particular outcome o.
- The social welfare W can be defined as the collective summation of all agents valuations, that is

  $$W(o, \theta) = \sum_{i \in \mathcal{I}} v_i(o, \theta) \tag{1}$$

What we've talked about previously in the class is that the state of security is often an economic failure, where we are in a socially suboptimal state because of things like misaligned incentives,

prisoners dilemmas, and tragedies of the commons. So if you're a social planner and you want to maximize for everyone, your goals is to find the outcome o that maximizes

$$\max_{o \in \mathcal{O}} SW(o, \theta)$$

.

Question: If you're the social planner, why are you trying to find the o that maximizes social welfare? Why not try to find the θ as well? Answer: Because you can't control people's private information and can't control their internal preferences!

This leads to a somewhat obvious problem though: If we are the social planner and are trying to maximize welfare, you are relying on the players' self-reported θ_i to find the o that maximizes welfare. A self-interested non-cooperative player may lie about their θ_i to try to increase their own personal welfare.

For example, consider the naive cake cutting protocol. Player X may

4 Workshop Report

4.1 Formula 1 and "dirty air"

- The workshop report mentioned something about Formula 1 racing. *Anyone remember the name of the "problem"*
- Last class, Gabe mentioned how game theory is used in international relations, but not strictly in a mathematical or quantitative way.
  - Which is still a perfectly valid application, in my opinion.
- This workshop report is maybe an analogue of that idea but applied to mechanism design and security.
- As we've just seen, mechanism design can be a purely mathematical pursuit.
- Or we can ditch the math and just apply the themes.

5 Braess's Paradox

Ask: What do we think—do we need a social planner (i.e. government) to come in and solve all our problems?

I want to point out that if you don't fully understand the system you're modifying, you can very easily make things worse.

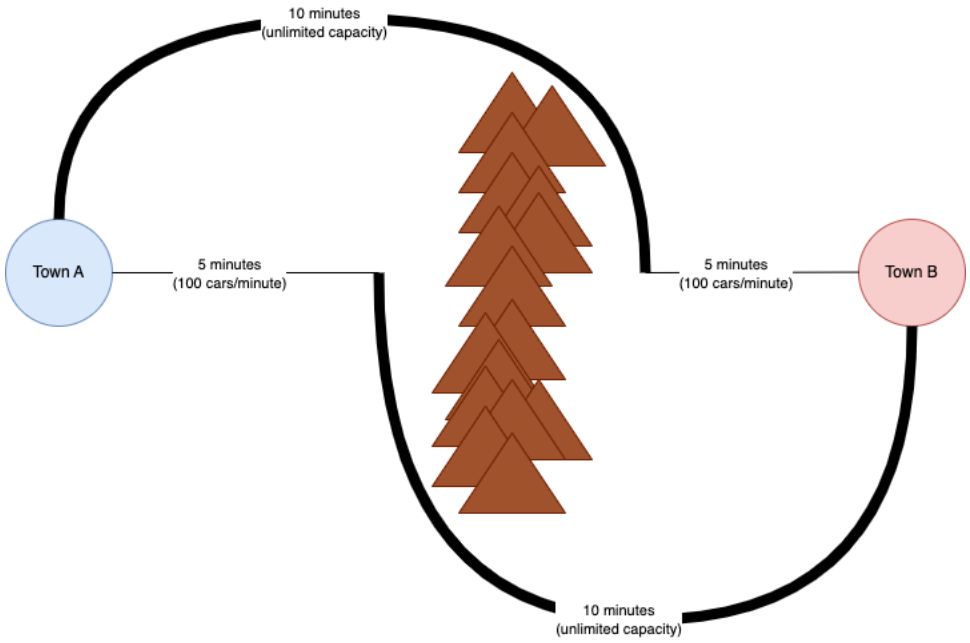

Figure 1: Before the tunnel, it takes 1200 cars 14 minutes to travel from Town A to Town B

Imagine there are two town, separated by a mountain range. There is one highway to the north, and one highway to the south. Connecting the highways to the towns are smaller country roads. The highways take 10 minutes to drive, and the country roads take 4 minutes, so no matter on which route you take, it takes 14 minutes to drive from Town A to Town B.

Suppose 1200 cars drive from Town A to Town B every hour.

Let's look at this as a game from the persepctive of the players.

- *What is the number of players?* $\mathcal{I} = \{1, 2, \ldots 1200\}$.
- *What is the set of actions available to the players?* Take the north route or take the south route: $a = \{\text{North}, \text{South}\}$.
- *How much utility do players get from either strategy?* In both cases the players are trying to minimize negative utility (lost travel time). We can write this as $u(R_N) = u(R_S) = -14{:}00$.
  - *Why is the utility negative?* Negative utility represents a bad thing. In this case, the bad thing is travel time. So the longer the drive time is, the greater the negative utility is, and the worse it is.

So what is the strategy? The strategy is to maximize utility, i.e. choose $\text{argmax}_a(u(a))$.

Since the utility of both is equal, we can assume that 50% of players will choose the north route and that 50% of players take the south route. In terms of game theory, we call this a *mixed strategy* since strategies are probabilistic in nature. Just in case you end up reading or working more on game theory later on and you see the term *mixed strategy*, that's all it means.

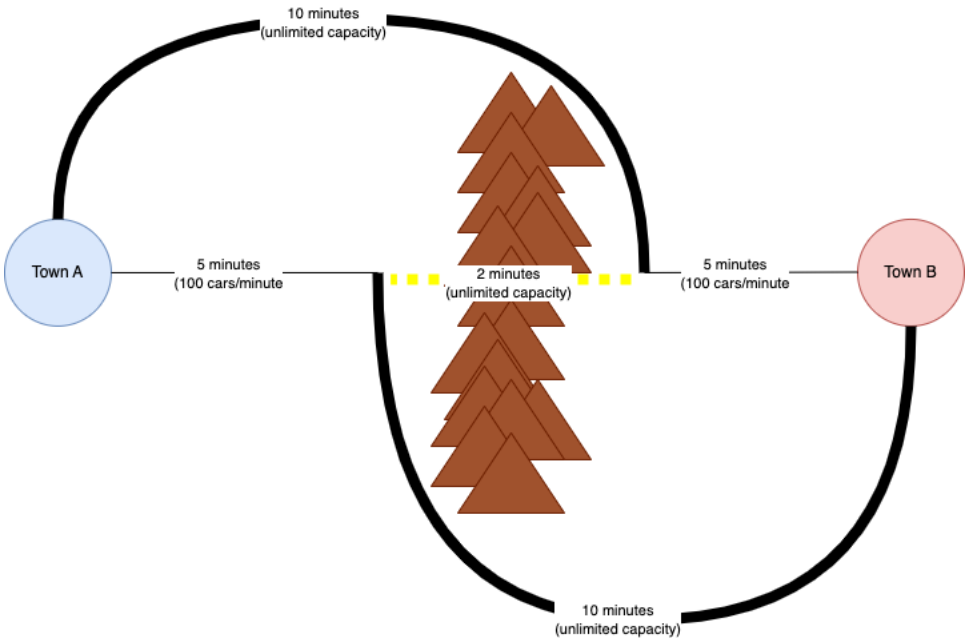

Figure 2: A naive urban planner might expect the tunnel to reduce the drive time to 10 minutes

Let's say you're an urban planner, and you just been hired by the Inter-Town Transit Authority to reduce the travel time between the two towns. What might you do to reduce travel times? You could dig a tunnel through the mountains.

Does anyone see a problem here? (If a student notices: We're going to incentivize all the traffic to go on the small country roads)

Let's suppose that you, the urban planner, did a little bit more research on the roads before digging the tunnel, and you found that the country roads have a limited capacity. Suppose they can only allow 12 cars/minute.

Recall that there are 1200 cars per hour. Since there are two routes, and they're identical, we can assume that every hour 600 cars take the north route and 600 cars take the south route per hour, or 10 cars per minute.

$$\frac{600 \text{ cars}}{\text{hour}} = \frac{10 \text{ cars}}{\text{minute}}$$

Can our country roads handle this capacity? Yes. In other words, the travel time is unconstrained by the number of cars. We can write the travel time of the country roads as

$$t = \max(4, 4n/12)$$

where n is the number of cars per minute.

TODO draw graph and explain!!

Any questions?

What happens if *every* car takes the tunnel? I.e. instead of 10 cars per minute, the country roads now have 20 cars per minute? The travel time for each stretch of the country road becomes

$$t = \max(4, 4(20)/12) = 6{:}40 \text{ minutes}$$

What happens if we dig the tunnel? In the best case the travel time will be 10 minutes, if there were no other cars on the road. But if everybody takes the tunnel, the travel time will be 6:40 + 6:40 + 2:00 = 15:20 minutes. This isn't any faster than the old route, which was 14 minutes.

But what is the time of the old route now? 6:40 + 10:00 = 16:40. Taking the highway is still the slowest option, but it too is slower than before!

Let's look at the game.

- $\mathcal{I} = \{1, 2, \ldots 1200\}$
- $a = \{\text{North}, \text{South}, \text{Tunnel}\}$
- $u(R_N) = -16{:}40, \quad u(R_S) = -16{:}40, \quad u(R_T) = -15{:}20$

Recall that the strategy is to maximize utility. The highest utility is achieved by taking the tunnel. But note that this is no longer a mixed strategy. This is a dominant strategy! Every player is going to take the tunnel.

Can someone tell me the plain-English interpretation of what I just said? We built a tunnel. The tunnel incentivized everybody to take the tunnel. But because of the road capacity issues, this route ended up being slower than the old routes around the mountains. But the tragedy is that the old routes now also became slower as well! You can't get from Town A to Town B in 14 minutes anymore.

Note that the old roads still exist! Everyone could cooperate and have 50% take the north road and 50% take the south road. But that's really hard to enforce! Because then what happens? That tunnel is still there, and is just too tempting—If you're the only one taking the tunnel, your travel time is 10 minutes!

5.0.1 Price of Anarchy

In economics, game theory, and mechanism design, there is a concept called the *POA*, or *Price of Anarchy*. This is a quantitative measure of the overhead imposed by agents choosing selfish behavior, and is defined as the ratio between the Nash equilibrium outcomem and the socially optimal outcome. In this game, the socially optimal outcome was -14 minutes while the equilibrium

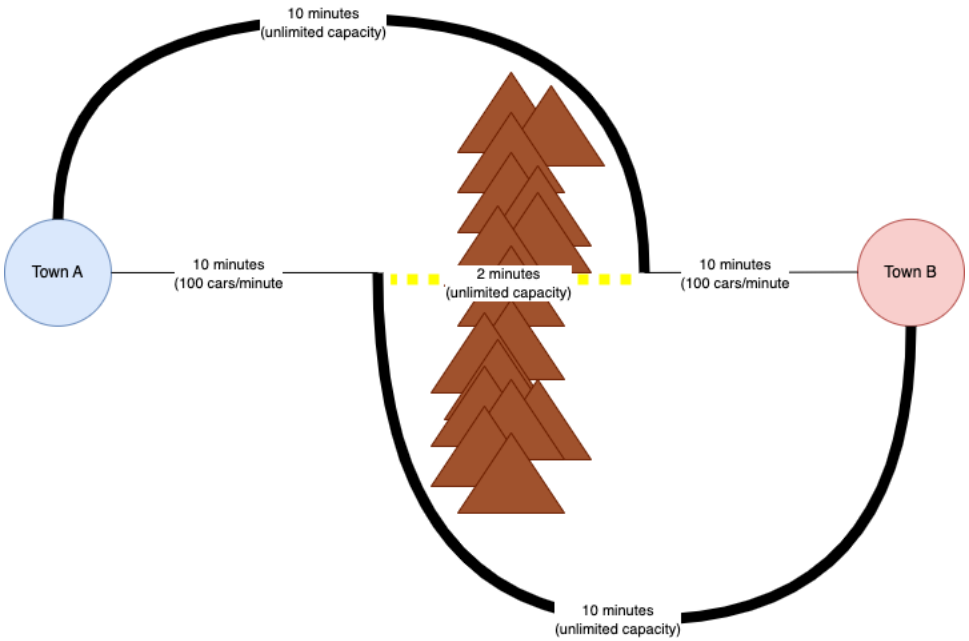

Figure 3: A naive urban planner might expect the tunnel to reduce the drive time to 12 minutess

outcome was -15 minutes 20 seconds. *(Note that the socially optimal outcome was -14 minutes even after the tunnel is built).* So we have

$$PoA = \frac{15{:}20}{14{:}00} \approx 1.095 = 9.5\%$$

In other words, the system is 9.5% worse off because the rules of the game allow for selfish behavior which comes at the expense of the other drivers.

In the prisoner's dilemma, we had a (-1,-1) payoff for (cooperate, cooperate) and a (-2,-2) payoff for (defect, defect). The social cost of (cooperate, cooperate) is 2 and the social cost of (defect, defect) is 4. This is a $4/2 = 2\text{X} = 100\%$ price of anarchy. *Any questions?*

Ask: In the prisoner's dilemma, who is the social planner? The party who has a degree of influence or control over the players, so in this case, the mob :)

As a fun fact, the term *Price of Anarchy* comes from a 1999 paper "Worst-case equilibria" by Elias Koutsoupias and our very own Christos Papadimitriou (although the concept of measuring the inefficiency of selfish behaviour predates this paper by quite a bit. But within the algorithmic game theory and algorithmic mechanism design, this is a very commonly-used term.)

5.0.2 Braess' Paradox

This is called Braess' Paradox. You may have also seen this with two springs and a weight. *(Has anyone seen this before?)* It's a physical analog to this situtation with roads and tunnels.

What are the takeaways from Braess's Paradox?

- If we're going to meddle with systems, we have to be very confident that we fully understand the system we are meddling with. Otherwise we can make things much worse. I suppose this is a mathematical example of a perverse incentive.
- *Any others*?
- The takeaway that I've led you to thus far is that a misguided planner can make things worse. Equivalently, a smart planner can make things better by doing counterintuitive roads. Imagine that you're the city planner hired to improve travel times between Town A and Town B, *and the tunnel already existed.* How do you think the public would react if your solution is to *close the tunnel*?
  - Sometimes you need the political will and determination to do things that are beneficial but unpopular.
  - *In general though, how do you know if you're the smart planner who is going to make things better, or the naive planner who is going to make things worse??* I don't know. Would love some answers to this question!

6 Mechanism Design for Hardware Security

This is not a peer-reviewed paper. This is a report from a workshop event a couple of summers ago, where me, Prof. Sethumadhavan, and about 50 other people from academia, industry, and government got together and discussed what mechanism design means for hardware security.

Aside: For anyone who doesn't know, *Hardware Security* is exactly what it sounds like: It's a subfield of security that focuses on securing systems from the ground up, starting from the circuit level to the physical device level and typically including up to computer architecture. If this is a topic that interests you, and you are not graduating this semester, there is a hardware security class in this department, taught by Prof. Sethumadhavan. (Today's class is really just an advertisement for other graduate-level classes at Columbia, isn't it?)

The workshop was from people who work in the hardware security area. I.e. topics like side channel detection and exploitation, hardware trojans, speculation safety, Rowhammer, hardware-based defenses like CFI, secure embedded systems, tamper-proof hardware, hardware enclaves and trusted execution environments (*to Simha: any others I'm missing?*) To prevent an echo chamber, there were also some people from economics and a few people who worked in government policy (next week's topic).

Ask: Why did I have you read a workshop report rather than a peer-reviewed paper on mechanism design for security? Mostly because there is no work on this topic! This is an emerging research area. The purpose of this workshop is to spur research in this direction. But mechanism design is such an interesting tool and one that security practitioners ought to have in their toolbox. It's economics applied directly to improving security. Extremely relevant to this class, (hopefully) extremely important to the field of security moving forward.

6.1 Recommendations

The format of the event was a mix of presentations and discussion, with the end goal of coming up with a set of recommendations for what the hardware and computing communities can do to improve hardware security through mechanism design.

These were the recommendations we came up with *(write on board)*:

6.1.1 Foster Diverse Educational, Professional, and Industrial Communities in Hardware Security

6.1.2 Lay the Scientific Foundations for Work that Combines Incentives and Technology

6.1.3 Make Security Accountable and Explainable

6.1.4 Co-Develop Emerging Technologies with the Understanding of their Hardware Security Ramifications

6.1.5 Prioritize the Human Impact of Hardware Security

6.2 Discussion

What do we think? Did we get it right? What did we miss? What did we get wrong? *Write recommendations on board.*

One thing I'd like to point out is that you should not assume that just because the audience of people who contributed to this list was a bunch of established professors that their level of wisdom or insight in the topic is much better than yours. Don't think of this list as infallible, and don't think of your intuitions as invalid. Like I said, this is an emerging area, so someone who has spent all their life working in security may barely have a leg up on you (if at all!) when it comes to taking an economic perspective at security problems.

Ask Simha: In retrospect, what changes would you make to the set of recommendations?

7 Homework