Privacy and Poverty

Professor Rachel Cummings
Georgia Institute of Technology
School of Industrial and Systems Engineering

January 23, 2018

These notes are based off a presentation by Professor Rachel Cummings (Georgia Tech, ISyE) in the Mechanism Design for Social Good Reading Group. The notes are taken by members of the reading group with some figures and texts taken from the accompanying papers or presentations. Questions and comments from reading group members during the presentation are labeled as such. Please contact the reading group organizers for any questions or comments.

1 Introduction

The world of privacy is much broader than what we typically think about in the TCS community. Today, we will talk specifically about how privacy and poverty intersect, and discuss papers from law and the social sciences, to brainstorm what techniques we can use from the Algorithmic Game Theory toolkit to understand and mitigate the negative impact of privacy loss on low-SES people. The first two papers we discuss talk about how privacy violations in practice disproportionally affect the poor, and the third paper puts forth a game-theoretic model of privacy loss that our research community is well-posed to analyze.

2 Class Differential in Privacy Violations

We notice invasions of privacy everyday from overly targeted ads, to TSA body-scanners, to our shopping behavior being tracked, and social media data being sold to third parties. In a sense, these are "first world problems" of privacy.

There is a lot of work showing that the privacy loss of low-SES individuals are much more invasive. For instance, [1] talks about class differentials in privacy law and the privacy violations that people have to incur to receive government benefits such as welfare. Individuals on such assistance programs are frequently visited by inspectors to make sure that the information they provide is true; they are questioned about their relationships, subjected to invasive testing (parental testing, drug testing), and their spending is tracked. Data collected in this way is shared across federal and state lines, and individuals lose their autonomy.

There is some argument that this is necessary to mitigate fraud. However, there is minimal evidence of fraud in the cases that have been investigated, so this concern might not be founded in reality. Furthermore, other forms of government assistance that are offered to the upper and middle classes do not have the same level of invasive monitoring: e.g., tax deductions, farm subsidies, government-backed mortgages. That is, there is no home inspection to verify that you in fact gave to charity. Government intervention and surveillance is disproportionally applied to impoverished communities.

Low wage workplaces similarly experience a disproportionally higher rate of surveillance and monitoring (e.g., close circuit cameras, computer keystrokes, GPS tracking, etc., which are not experienced in most middle- and high-wage workplaces.) In order to get a low wage job, people are often required to undergo invasive personality assessment tests where they are asked invasive and intimate questions, or required to take drug tests with urine samples produced in front of technicians. Gilman, who engaged in several low wage jobs for her research, says:

> It is unsettling, at the very least, to give a stranger access to things, like your self-doubts and your urine, that are otherwise shared only in medical or therapeutic situations. [1]

The privacy violations experienced by the poor are much more invasive and degrading than those experienced by middle- and upper-class. This is unfortunate given that the poor are among the most vulnerable members of society.

2.1 Algorithmic Game Theory for Privacy?

Given this body of work from the social sciences and law, we might ask what techniques AGT has to offer to address some of these problems. We find that our tools cannot be directly applied to address some of the above concerns. Revealed preference theory says that individuals chose to receive welfare/low-wage employment, so they must be happy to exchange privacy for money at these prices. However, no one chooses to go onto welfare, so it would be erroneous to assume that this is a free choice.

> Just because people give up information in exchange for jobs and other valued outcomes should not be construed as meaning that doing so is voluntary. [1]

Gilman also points out:

> Accepting welfare can subject one to humiliation, but refusing it can result in hunger. This 'choice' hardly promotes autonomy or dignity. [1]

If we are viewing this problem from an AGT lens, then it is important to take into account that this is not a free choice.

There is already a precedent for not letting individuals choose things simply because they are willing. In practice, markets are regulated; we impose minimum wage requirements to make sure that people don't work for too low a wage, and we have a ban on pricing and selling organs since we believe that it is morally repugnant to do so.

Question: What is the reason behind minimum wage requirements? And, can those reasons be translated to the privacy case?

Answer: Both of the above examples are concerned with people being desperate and having very few outside options, so they are willing to accept something that is bad for them or bad for society. The two examples are different in that you can only give away so many organs before you die, while the issue with minimum wage is a question of market unraveling, which we will talk about later.

Question: What are the assumptions about how people value privacy? Are we just assuming that invasion of privacy is a bad outcome for everyone or that everyone has the same value for losing privacy?

Answer: We'll want to specify a model for privacy first before we delve into this. For this talk, we can say that we want privacy for moral reasons. Later, we will hopefully discuss the right model for the value of privacy. In general, there is a very active research area about how we can model and quantify the value of privacy. Here, since we're talking about a particular market for privacy, perhaps we can find some model that makes sense in our particular case.

The overarching question that we ask here is:

Question 1. Should we regulate the market for privacy to restrict people from sharing information, even when doing so is a best response?

2.2 Other Shortcomings

In addition to the AGT shortcomings above, there are other failures. One is the failure of legal principles, especially for the poor.

A common notion in law is the Brandeis "right to be left alone." This does not match the needs of poor people who need government assistance and, in that sense, do not want to be left alone. Indeed, Brandeis was a sort of celebrity of his time, and the ways in which discussions evolved are phrased more like things that a celebrity might want to preserve from the paparazzi. This doesn't map to the needs of the poor. There still needs to be some privacy notion that accounts for both their economic/financial need and preserves their privacy.

Similarly, the fourth amendment, which discusses unreasonable searches and seizures is interpreted by US courts as being a property-based privacy protection. This is, again, ill-suited in our context.

There are also some laws and privacy statues focusing on misuse of existing data (e.g., what it can be used for, who it can be shared with), rather than invasive data collection methods.

2.3 A Brief History of Privacy for the Poor

In the colonial era, the poor were servants to the rich. They had no autonomy, no freedom, and definitely no privacy. Then, in the 18th century, we had poorhouses. The poor had uniforms; there were behavioral rules and forced labor. Again, we had no autonomy and no privacy for the poor.

Around the 19th century came a period of "scientific charity" where middle-class visitors entered the homes of the poor to provide moral and religious counseling. Here, we see a switch from people being completely stripped away of any autonomy to a situation where there was some autonomy, but in order to live, you had to accept invasions by the rich who are going to come in and judge your life and determine whether you are worthy of receiving aid.

In modern day privacy, we have both online and offline surveillance. In [2], they interviewed low-SES youth (17-27) in NYC and drew a comparison between online and offline police surveillance. Low-SES neighborhoods have high police presence, and low-SES individuals are more accustomed to offline surveillance in daily life. They showed that perceptions of online vs. offline surveillance are vastly different. In the online world, a lot of the comments of the youth had to do with personal responsibility in their interactions. They expressed ownership of the responsibility for their actions and the resulting surveillance. Below are some quotes from the interviewees:

> The thing I like about the Internet is that you have kind of this power to avoid situations that make you uncomfortable.
>
> You abstain, you control. The only thing you actually have control over is [...] what you put out there in the world.
>
> Don't say stuff that you don't want other people to hear.
>
> It boggles my mind how people can just put [...] provocative pictures of themselves on the Internet. [2]

The interviewees exhibited a lot of awareness of being observed online, but they also saw it as a personal responsibility to take charge of how they were viewed.

Their perception of offline surveillance was very different. They expressed sentiment that it was unavoidable and outside of their control.

> We're not doing anything wrong, but we're going to get busted for nothing.
>
> I've seen the bad things where cops will just stop others for no reason. [2]

Marwick et al. describe,

> Many participants had extensive experience with policing and physical surveillance and were aware they could not avoid such encounters through their own efforts.
>
> They are aware that avoiding police harassment cannot be achieved through individual action. [2]

Another observation is that the "nothing to hide" privacy fallacy—which says that if you're not doing anything wrong, then you have nothing to hide and you don't need privacy—fails in the offline world. For example, seeing police lights in your rear-view mirror can still make your heart race, even if you aren't speeding. Nonetheless, this fallacy is often used to incentivize information-sharing and reduce privacy protections.

> The whole 'Oh, I'm not doing anything wrong, so I have nothing to fear,' thing. I believe that, but in another way, every time I'm on a train car, I'll feel intensely uncomfortable. [2]

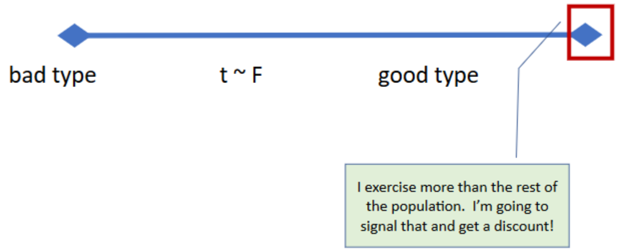
3 Game Theoretic Model for Privacy Loss

In this section, we talk about [3], which is a law paper that puts forth an economic model of privacy that we can formalize using tools from AGT.

This model has to do with readily available personal signals. Increasingly, and in a variety of contexts, we have a way of sending low-cost, verifiable information about ourselves to companies who might want to use that information for personalization, such as price discrimination or decision-making. Companies can incentivize individuals to share their information. For example, Fitbit tracks heart-rate and exercise rates. Health insuance providers can use this information to give discounts for healthy patients. Another example is GPS trackers in your vehicle, which can be used by car insurances to give discounted premiums for safe drivers.

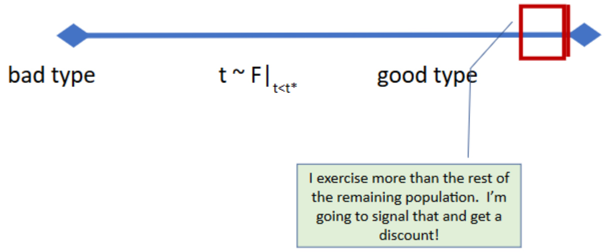
Due to the ease of information collection and sharing, we can imagine it extending to other cases like credit scoring, employment contexts, etc. Instances like the Equifax breach coupled with the ease of information-sharing described above move us into a world where we need to rethink information-sharing models. It might not be shocking if we moved in the future to a decentralized setting of information sharing as described above, since we are already there in select markets. In this setting, people have control over their own information and could authorize sharing it.

The paper suggests the following model: imagine a signaling model where a company wants to screen customers. There is a continuum of consumer types and the company has some prior over types. The company wants to assign an outcome to each individual based on the company's belief of that individual's type. As a running example, think of pricing health insurance based on how much you exercise. If the company has no personal information, then it only has its prior belief and it has to set the same price for everyone. In this setting, it will be individually-rational for the people who exercise the most (on the right of the scale below) to reveal information in order to get a lower price.



I exercise more than the rest of the population. I'm going to signal that and get a discount!

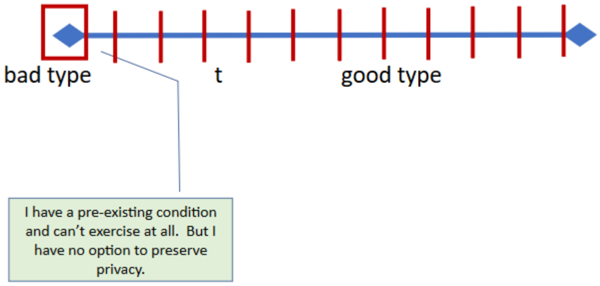
The company then updates their belief based on this signaling scheme, so they assume that everyone who has not revealed information must be below this type. They reset the price based on this, but we are back to the above case where there will be a new set of agents who want to signal their information in order to get a discount.



I exercise more than the rest of the remaining population. I'm going to signal that and get a discount!

This leads to unraveling, and we ultimately end up with a case where the only non-revealed agents that remain are of a "bad type." The company now has perfect information and the "bad type" has no option to preserve privacy. Their decision to not signal is itself a signal.

We might be in a setting where we believe that it is morally unacceptable to discriminate against the "bad type." (E.g., individuals have a pre-existing condition that should not be used


bad type
t
good type
I have a pre-existing condition and can't exercise at all. But I have no option to preserve privacy.

to discriminate against them.)

This is a very stylized model. We know that in general that there are several possible barriers to unraveling. Unfortunately, they don't really exist in this model. Some examples for these barriers are:

1. *Cost of signaling*: If you were in good health in the 1980s and wanted to prove that to your insurance company, then you would have to go to a cardiologist to get an expensive and time-consuming test. Healthy people may choose to not signal because the process is too costly. In modern times, people only need to hit 'agree' on their FitBit. The cost of signaling is diminishing and entirely disappearing in many domains.

2. *Cheap talk*: If the information is unverifiable, then it may not be as valuable. We don't have the option to do cheap talk in this setting since signals can be verified more easily. This problem will continue to diminish as devices and sensors improve.

3. *Moral resistance*: perhaps the example that has the most bite here is that of moral resistance. An envelope is only suspicious if everyone else uses a postcard. People who don't have anything to hide can choose to mail things in an envelope. However, companies may provide a significant economic incentive to share data, which can outweigh morals. This incentive is getting bigger and bigger as we are moving into a more data-driven world.

4 Discussion

With this in the background, we now ask what our research community can contribute. The running theme has been that low-SES people choose to give up their privacy, but we have seen that, in many settings, it's not really a choice. The overarching question is: Given that this is the case, *should the market for privacy be regulated? And, if so, how?*

Below are some guiding questions:

1. How do we quantify or model the disparate impact of privacy violations on the poor? This is currently not captured in the proposed economic model of unraveling. We want to account for an element of disparity and unfairness that is happening here.

2. In signaling model of Peppet [3], lower types may be correlated with poverty (or having fewer outside options). How should we quantify social welfare or disutility from privacy loss? We need to understand why people want privacy and why the poor are more affected by privacy loss.

3. Are there alternative models that more closely match the welfare story (or other practical applications)? Some disenfranchised groups are more affected more by privacy losses in practice.

4. What about data-driven analysis? What data would be useful and where might it exist? Take a look at: "Privacy, Poverty, and Big Data: A Matrix of Vulnerabilities for Poor Americans."

Question: Are there any existing regulations on privacy that we can use as a sort of example?

Answer: Not as much as we might hope. Privacy law is very application-driven. For instance, for historical reasons, there are strong protections for video rental data, but not, say, FitBit data. HIPAA does not cover wearable technologies. A part of the problem is that privacy laws have not caught up with technological advances. We have better technology for collecting data than we have laws to protect data. Many privacy laws are concerned with consent, and we have seen that this does not apply in contexts where individuals might be economically coerced into consenting when they don't want their data to be shared.

Question: We just covered a section on bias and discrimination, where we also discussed regulations. A lot of the issues that are coming up here related to privacy seem very closely tied to the discrimination literature as well. How are the two connected and is it possible to piggyback off of some of the work in the discrimination literature and the infrastructure put in place to initiate conversations across academic boundaries to deepen our understanding of how privacy and poverty intersect? The people being negatively impacted by algorithmic bias and privacy loss are likely a lot of the same communities of people.

Answer: Yes, this is indeed a discrimination question. I hope that the tools from the discrimination work will be helpful here. Research in our community on the intersection of privacy and poverty is very sparse, and there is potential to connect it to other fields.

Question: Why is so much information asked of low-wage workers? Is wealth seen as a sign of less risky behavior or trust-worthiness?

Answer: There are a variety of reasons. In low-wage work, there is a huge amount of substitutability among workers. If the job is low-skill, then one might believe that ability to do it is correlated with very simple personality traits. Therefore, it might be seen as beneficial to do such a test, whereas it might otherwise be difficult to obtain this level of relevant information for high-skill jobs.

If you are in a medium- or high-skilled job, your boss might still want to run these tests, but workers in these settings have more outside options, and there are fewer qualified applicants. Therefore, the company runs a larger risk in asking employees more information than is the industry standard.

Question: The argument that there is a moral direction of revealing information discussed in the envelope-postcard example can go in the opposite direction at times. Take, for instance, the case of insurance companies who want to know how safe a driver is. But, in the case where everyone uses an envelope, i.e., refused to provide information about how safe of a driver they are, then the insurance companies might decide to instead use information about your car (e.g., age, quality, etc.) to determine how safe of a driver you are. This would discriminate against low-SES individuals. But, if you are someone who is low-SES and a safe driver, you might want to offer evidence that you are a safe driver. In this situation, why would we consider this person to be morally dubious when, in fact, they are just providing information to mitigate some of the bias against them due to their low-SES status?

Comment: This does bring up the question of whether information that firms are collecting is causally linked (rather than correlated) to things that they are actually trying to predict. (e.g., are the personality tests that are currently conducted necessary for low-skill jobs?)

Answer: It's not quite saying that they are engaging in a morally dubious activity. This example might indicate that we should allow people to share information, or perhaps share information in a more nuanced way by using multi-dimensional ranking rather than this simple model.

Another point is that, we might also want to challenge that the company is using the age and quality of the car as a substitution if it is not shown to be causally liked to safety of the driver. If it is not, then this would be an instance of discrimination that we should address independent of the concerns about privacy unraveling.

There is also a distinction between providing information about driver safety and, say, the sort of day-to-day invasions that low-SES people who are receiving government assistance experience.

This also suggests an avenue for research, which is to understand how unraveling unfolds depending on how biased the mechanism is.

Question: What is the effect of a minimum wage (say on the bottom x percent of the population)? Is there a model for why minimum wage does not stop people from hiring and how it improves social welfare?

Answer: The minimum wage increases the wages of the bottom x percent that is employed. The welfare effects in this case are negative due to disemployment. The idea is that, as you increase the minimum wage, it might be higher than the marginal productivity of some workers, and employers will choose not to hire them. This leads to more unemployment in the market. These two effects are balancing each other out. If the disemployment is too large, then you end up having too negative of a social welfare effect (and vice-versa). There are some models of the labor market that are becoming popular which say that when you increase the minimum wage, then you increase overall welfare. In this model, with the increase in minimum wage, you attract more people to work, and there is a net increase in welfare in this model.

Comment: I can imagine using our tools to study different mechanisms and their welfares and how much monitoring and invasions is happening here. What is hard to imagine is how we can trade off people wanting privacy and invasions versus wages.

Comment: It seems like the hard thing to capture is how people trade-off privacy and wages. Another hard balance is fairness and privacy in these settings that we have discussed.

Comment: There is some literature on privacy as a public good and that there are individual and group incentives that are diverging, as discussed above. One way to think about mitigating the negative outcomes is to think of this in a repeated interaction way. Agents could cooperate to reap benefits of not revealing information in this repeated interaction setting, even though the individually rational action in each isolated setting might be different. We can consider impact of informing individuals about the long-term benefits from not revealing their information in this repeated-interaction context.

Comment: Another issue along the same lines is, if you consent to sharing your information for a specific purpose, it might be used for other purposes down the line that you cannot foresee at the moment, so the cost of sharing information might be underestimated. As above, this again assumes that the individual is not so desperate that taking into account this repeated-interaction or potential future cost framework would change their mind. It might not be the case, and the question of whether and how to regulate markets for privacy should address this in order to take those with few outside options into account.

This point indicates that there should also be regulations on what kind of information be passed around between different entities, rather than what kind of information is being collected and how we should regulate it. Information passing has also been studied.

References

[1] Michele Estrin Gilman. The class differential in privacy law. *Brooklyn Law Review*, 77, 2012.

[2] Alice Marwick, Claire Fontaine, and danah boyd. nobody sees it, nobody gets mad: Social media, privacy, and personal responsibility among low-ses youth. *Social Media + Society*, 3(2):2056305117710455, 2017.

[3] Scott R. Peppet. Unraveling privacy: The personal prospectus and the threat of a full-disclosure future. *Northwestern University Law Review*, 105(3), 2011.