Benefits of Privacy and Anonymity in Information Technology
Project Proposal

Today we live in a world where we use electronics in all facets of our life. From communication to shopping, many aspects of life have been simplified greatly by the Internet. However, by using many of these systems, we are exposing many details of our personal lives. Many industries, such as credit card companies and online shopping sites, record every transaction and have compiled them into massive databases. Other social networking sites have even more details that people consider private, but are actually part of large databases that are combed over to find patterns. As computer memory storage has become cheaper and processors faster, analysis of this glut of data has become easier and patterns have become easier to recognize. These patterns can reveal behavioral traits that are not easily apparent.

This information is very valuable to advertisers and marketers, as even small amounts of past purchases can reveal a detailed profile of an individual. From this profile, advertisers can tailor ads to specific customers by showing them ads that they are most likely interested in. One such store to do this is Target. Target stores every purchase a customer makes and sends targeted coupons to them. One example of this is that one Target statistician found a bundle of products that allowed him to assign every shopper a likelihood of pregnancy\(^1\). The algorithm eventually was able to predict a due date in a very small window and this allowed Target to send coupons to expecting mothers for essential baby items. Tracking past purchases can also allow stores to offer new customer specials or reward loyal customers for their business. It can also make purchasing more efficient, as previous purchase options are already known.

While this seems like it would benefit customers, as targeted ads are more likely to be relevant than random ads, it can actually work to the detriment of the customer. The first problem is that if a company knows what an individual is likely to buy, they can increase the price of the item for the one customer. This is known as price discrimination, where a company will charge one customer more because they are more likely to pay more for the item. One example of this was an attempt by Amazon in 2000 to charge customers more for DVD's that past purchase history suggested they would like. This incident caused an uproar and forced Amazon to stop the practice.

Beyond price discrimination, the fundamental issue is one of privacy. Very few customers explicitly allow companies to collect their data. Most are not even aware they are being tracked, analyzed and many times taken advantage of.

\(^1\) [http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/](http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/)
However, consumers are able to fight back. Through various techniques, such as anonymizing browsers or using new credit cards, customers can hide their identity and appear like new customers. Nevertheless, while it is still possible to stay anonymous, it is costly to do so in terms of time, effort and sometimes even money.

Thus, this interaction can been seen as a conflict between a firm, which attempts to identify a customer in order to price discriminate, and a customer, which does not want to be identified but must pay a cost in order to stay anonymous. This problem is tackled in three different papers: “An Economic Analysis of User-Privacy Options in Ad-Supported Services”, “Condition Prices on Purchase History”, and “Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases”. In these papers, a model is proposed where a firm can set the cost of anonymity. In the first paper, the authors look at where the firm should set this cost in order to maximize their revenue from targeted and untargeted ads. In the other two, the authors look at whether a firm should price discriminate when the customers have the option to be anonymous and not be tracked.

The models in the three papers have both similarities and differences. In all three models, a firm has monopoly control of a good. This firm tracks its customers unless they are anonymous and can set the price of anonymity. This firm wants to maximize its profits. The customers all derive utility from the firm’s good. However, beyond this the models tend to differ. The model of the first paper from above only occurs in one period and looks at how a firm should set the cost of privacy in order to maximize its revenue from targeted and untargeted ads. By controlling this cost, the firm can increase revenue and give users the option of privacy. The paper found that if the users have normally distributed utilities and a privacy-sensitive user is worth at least $\sqrt{2} - 1$ as much as a privacy-insensitive user, the site should try to make the cost of privacy as small as possible in order to maximize revenue.

In the other models, the interaction between firms and customers occurs over two periods. In the first period, the customers decide whether to buy the product and whether to stay anonymous. Based on this data, firms can decide to price discriminate in the second period. These papers found that the firms realize higher profits when everyone stays anonymous than when they fully price discriminate. This is because higher value customers will not purchase in the first period when they know they will be discriminated against in the future. So, the firm will need to offer a lower introductory price to incentivize the higher value users to buy in the first period. The firms, thus do not recoup the losses realized in the first period from the lower prices when they price discriminate in the second period.

However, both of these models can be further extended to consider other economic factors. One of the biggest element that is not considered is the myopic consumer. A myopic consumer does not realize or care that they are being tracked. Thus, they base their purchasing decisions on the current price, not the possibility of a future price that is greater. Intuitively, many consumers seem to act like this. Thus, one way these models can be extended is to include the possibility of a myopic consumer and see if in this case, price discrimination might increase revenue.

Beyond the myopic customer, these papers only touch upon competitor firms and do not consider firms with differentiated products. Another direction is to
consider an extension where the consumer gains a benefit from being identified, such as what happens in loyalty programs. This can be taken even further by considering instances where a firm might pay a customer to be identified. One other direction to take would be to consider a game that is longer than two periods, either with multiple interactions between a fixed set of customers or over-lapping generations. This can also be extended by considering decreasing marginal utility for the product.

In my project, I will look deeper into the above models and try to extend them to better capture economic realities. I will examine the natural extensions to the model as outlined above and see if the papers’ conclusions still hold in those cases. This is important because advertising and data mining is becoming an increasingly large and important industry. It is significant for firms to know whether the money they spend on it is well spent. Additionally, it is important for customers in order to be aware of both price discrimination and the privacy concern.

My deliverable at the end of the project will be a paper that examines in detail the work described above, generalizes the models therein and addresses one or more of the open questions.

References
