2018 ICM
Problem F: Cost of Privacy

Pervasiveness of, and reliance on, electronic communication and social media have become widespread. One result is that some people seem willing to share private information (PI) about their personal interactions, relationships, purchases, beliefs, health, and movements, while others hold their privacy in these areas as very important and valuable. There are also significant differences in privacy choices across various domains. For example, some people are quick to give away the protection of their purchasing information for a quick price reduction, but at the same time are unlikely to share information about their disease conditions or health risks. Similarly, some populations or subgroups may be less willing to give up particular types of personal information if they perceive it posing a personal or community risk. The risk may involve loss of safety, money, valuable items, intellectual property (IP), or the person's electronic identity. Other risks include professional embarrassment, loss of a position or job, social loss (friendships), social stigmatization, or marginalization. While a government employee who has voiced political dissent against the government might be willing to pay to keep their social media data private, a young college student may feel no pressure to restrict their posting of political opinion or social information. It seems that individual choices on PI protection and internet and system security in cyber space can create risks and rewards in elements of freedom, privacy, convenience, social standing, financial benefits, and medical treatment.

Is private information (PI) similar to private personal property (PP) and intellectual property (IP)? Once lawfully obtained, can PI be sold or given to others who then have the right or ownership of the information? As detailed information and meta-data of human activity becomes more and more valuable to society, specifically in the areas of medical research, disease spread, disaster relief, businesses (e.g. marketing, insurance, and income), records of personal behaviors, statements of beliefs, and physical movement, these data and detailed information may become a valuable and quantifiable commodity. Trading in one's own private data comes with a set of risks and benefits that may differ by the domain of information (e.g. purchasing, social media, medical) and by subgroup (e.g. citizenship, professional profile, age).

Can we quantify the cost of privacy of electronic communications and transactions across society? That is, what is the monetary value of keeping PI protected, or how much would it cost for others to have or use PI? Should the government regulate this information or is it better left to privacy industry or the individual? Are these information and privacy issues merely personal decisions that individuals must evaluate to make their own choices and provide their own protection?

There are several things to consider when evaluating the cost of privacy. First, is data sharing a public good? For example, Center for Disease Control may use the data to trace the spread of disease in order to prevent further outbreak. Other examples include managing at risk populations, such as children under 16, people at risk of suicide, and the elderly. Moreover, consider groups of extremists who seek to hide their activities. Should their data be trackable by the government for national security concerns? Consider a person's browser, phone system, and internet feed with their personalized advertisements; how much is this customization worth?

Overall, when evaluating cost of privacy we need to consider all of these tradeoffs. What is the potential gain from keeping data private and what is lost by doing so?

As a policy analysis team for a national decision maker, your team's tasks are:

Task 1: Develop a price point for protecting one's privacy and PI in various applications. To evaluate this, you may want to categorize individuals into subgroups with reasonably similar levels of risk or into related domains of the data. What are the set of parameters and measures that would need to be considered to accurately model risk to account for both 1) characteristics of the individuals, and 2) characteristics of the specific domain of information?

Task 2: Given the set of parameters and measures from Task 1, model for cost of privacy across at least three domains (social media, financial transactions, and health/medical records). In your base model consider how the tradeoffs and risks of keeping data protected affect your model. You may consider giving some of the tradeoffs and risks more weight than others as well as stratifying weights by subgroup or category. Consider how different basic elements of the data (e.g. name, date of birth, gender, social security or citizenship number) contribute to your model. Are some of these elements worth more than others? For example, what is the value of a name alone compared with value of a name with the person's picture attached? Your model should design a pricing structure for PI.

Task 3: Not long ago, people had no knowledge about which agencies had purchased their PI, how much their PI was worth, or how PI was being used. New proposals are being put forth which would turn PI into a commodity. With the pricing structure you generated in Task 2, establish a pricing system for individuals, groups, and entire nations. With data becoming a commodity subject to market fluctuations, is it appropriate to consider forces of supply and demand for PI? Assuming people have control to sell to their own data, how does this change the model?

Task 4: What are the assumptions and constraints of your model? Assumptions and constraints should address issues such as government regulations (e.g. price regulations, specific data protections such as certain records that may not be subject to the economic system) and cultural and political issues. Based on your model and the political and cultural issues, consider if information privacy should be made a basic human right when thinking about policy recommendations. Consider introducing a dynamic element to your model by introducing the variations over time in human decision-making given changing personal beliefs about the worth of their own data (e.g. personal data such as name, address, picture), transaction data (e.g. on-line purchases, search history), and social media data (e.g. posts, pictures).

Task 5: Are there generational differences in perceptions of the risk-to-benefit ratio of PI and data privacy? As generations age, how does this change the model? How is PI different or similar to PP and IP?

Task 6: What are the ways to account for the fact that human data is highly linked and often each individual's behaviors are highly correlated with others? Data on one person can provide information about others whom they are socially, professionally, economically, or

demographically connected. Therefore, personal decisions to share one's own data can affect countless others. Are there good ways to capture the network effects of data sharing? Does that effect the price system for individuals, subgroups, and entire communities and nations? If communities have shared privacy risks, is it the responsibility of the communities to protect citizens' PI?

Task 7: Consider the effects of a massive data breach where millions of people's PI are stolen and sold on the dark web, sold as part of an identity theft ring, or used as ransom. How does such a PI loss or cascade event impact your model? Now that you have a pricing system that quantifies the value of data per individual or loss type, are agencies that are to blame for the data breach responsible to pay individuals directly for misuse or loss of PI?

Task 8: Write a two-page policy memo to the decision maker on the utility, results, and recommendations based your policy modeling on this issue. Be sure to specify what types of PI are included in your recommendations.

Your submission should consist of:

- One-page Summary Sheet,
- Two-page memo,
- Your solution of no more than 20 pages, for a maximum of 23 pages with your summary and memo.
- Note: Reference list and any appendices do not count toward the 23-page limit and should appear after your completed solution.