

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.

2018 年 ICM

问题 F：隐私成本

电子传播和社交媒体的普及和依赖已经变得普遍。其中一个结果是，有些人似乎愿意分享关于他们的个人交往，关系，购买，信仰，健康和运动的私人信息（PI，private information），而另一些人则认为这些方面的隐私非常重要和有价值。不同领域的隐私选择也有显著差异。例如，一些人会迅速放弃购买信息的保护，以便迅速降价，但同时他们不太可能共享有关其疾病状况或健康风险的信息。同样，如果某些人群或子群体认为个人信息构成了个人或社区风险，他们可能不愿意放弃特定类型的个人信息。风险可能涉及安全，金钱，贵重物品，知识产权（IP）或个人电子身份的损失。其他风险包括职业尴尬，失去职位或工作，社交损失（友谊），社会耻辱或边缘化。虽然对政府表示异议的政府工作人员可能愿意为保留其社交媒体数据而付费，但一位年轻的大学生可能不会对限制他们发表政治观点或社会信息感到任何压力。网络空间中个人信息保护，互联网和系统安全的个人选择似乎可以在自由，隐私，便利，社会地位，经济利益和医疗等方面创造风险和回报。

私人信息（PI）是否与私人财产（PP）和知识产权（IP）类似？一旦合法获得，PI 可以被出售或给予给他人，然后他人便拥有该信息的权利或所有权？随着人类活动的详细信息和元数据对于社会越来越重要，尤其是在医学研究，疾病传播，救灾，企业（如营销，保险和收入），个人行为记录，信仰和身体活动这些方面，这些数据和详细信息可能成为一种有价值 and 可量化的商品。交易自己的私人数据带有一系列风险和收益，这些风险和收益可能因信息领域（例如购买，社交媒体，

医疗)和小组(例如公民身份,专业背景,年龄)而有所不同。

我们能否量化整个社会电子通讯和交易的隐私成本?也就是说,保持PI的货币价值是多少,或者其他人拥有或使用PI的成本是多少?政府应该调整这些信息还是留给隐私行业或个人呢?这些信息和隐私问题只是个人决定,个人必须评估自己的选择和提供自己的保护?

评估隐私成本时需要考虑几件事情。首先,数据分享是否是公共物品?例如,疾病预防控制中心可以使用这些数据来追踪疾病的传播,以防止进一步的暴发。其他例子包括管理危险人群,如16岁以下的儿童,有自杀风险的人和老人。此外,还要考虑一些试图隐藏自己活动的极端主义分子,为了国家安全问题他们的数据是否可以由政府追踪?考虑一个公民的浏览器、电话系统和有个性化的广告的互联网,这个定制值多少钱?

总的来说,在评估隐私成本时,我们需要考虑所有这些权衡。保持数据隐私的潜在收益是什么?这样做会导致什么损失?

作为国家决策者的政策分析团队,您的团队的任务是:

任务1:在各种应用程序中制定保护个人隐私和PI的价格点。为了对此进行评估,您可能希望将个人分类到具有相当相似风险水平的亚组或数据的相关领域。为了准确地模拟风险,需要考虑哪些参数和措施?1)个人的特征,2)特定信息领域的特征?

任务2:从任务1给出的参数和度量,至少在三个领域(社交媒体,金融交易和健康/医疗记录)建立隐私成本模型。在您的基础模型中,考虑保持数据保护的

权衡和风险如何影响您的模型。您可能会考虑给予一些权衡和风险更多的权重，并按分组或类别对权重进行分层。考虑数据的不同基本要素（例如姓名，出生日期，性别，社会保障或公民身份号码）如何对您的模型作出贡献。这些元素中是否有些比其他元素更值钱？例如，与附有该人的照片的姓名的价值相比，单独一个姓名的价值是多少？您的模型应该为 PI 设计一个定价结构。

任务 3：不久之前，人们不知道哪个机构购买了 PI，他们的 PI 值多少，或者 PI 如何使用。现在提出的新的提案，会将 PI 变成商品。利用您在任务 2 中产生的定价结构，为个人，团体和整个国家建立定价体系。随着数据成为受到市场波动影响的商品，考虑供应和需求的力量是否合适？假设人们有控制权来销售他们自己的数据，这是如何改变模型的？

任务 4：你的模型有哪些假设和限制？假设和约束条件应该解决诸如政府法规（例如价格法规，具体的数据保护，例如可能不受经济制度约束的某些记录）以及文化和政治问题等问题。根据你的模式和政治和文化问题，考虑一下在考虑政策建议时是否应把信息隐私作为基本人权。考虑改变个人对自身数据价值（如姓名，地址，图片等个人数据）的个人信念，交易数据（如在线），引入人类决策随时间变化的动态因素（购买，搜索历史）和社交媒体数据（例如帖子，图片）。

任务 5：在 PI 和数据隐私的风险收益比的观念上是否存在代际差异？随着年龄的增长，这是如何改变模型的？PI 和 PP 和 IP 有什么不同？

任务 6：如何解释人类数据高度相关的事实，并且每个人的行为往往与他人高度相关？一个人的数据可以提供与其有社交，专业，经济或人口统计学方面联系

的其他信息。因此，个人决定分享自己的数据会影响到无数其他人。有没有很好的方法来捕捉数据共享的网络效应？这是否会影响个人，小组和整个社区和国家的价格体系？如果社区共享隐私风险，社区是否有责任保护公民的 PI？

任务 7：考虑一个大规模的数据泄露的影响，数以百万计的人的 PI 在黑暗的网络上被盗窃并出售，作为身份盗窃环的一部分出售，或作为赎金使用。这种 PI 损失或级联事件如何影响您的模型？现在您已经有一个定价系统来量化每个人或每个损失类型的数据价值，那么机构是否有责任因 PI 的丢失或者滥用直接向个人付钱？

任务 8：根据你在这个问题上的政策建模，向决策者写一个两页的政策备忘录，介绍效用，结果和建议。请务必指定您的建议中包含哪些类型的 PI。

您的提交应包括：

- 1) 单页汇总表
- 2) 两页备忘录
- 3) 您的解决方案不超过 20 页，最多 23 页的摘要和备忘录。
- 4) 注意：参考文献列表和任何附录不计入 23 页的限制，应在完成的解决方案后出现。