Chapter 1

PRIVACY AND SECURITY OF PERSONAL INFORMATION

Economic Incentives and Technological Solutions

Alessandro Acquisti
H. John Heinz III School of Public Policy and Management
Carnegie Mellon University
acquisti@andrew.cmu.edu

Abstract I discuss the evolution of the economic analysis of privacy, and then distinguish between the on-line and the off-line identities of an individual. Each type of identity raises different privacy concerns and economic implications. While market forces might ensure fair use of data connected to the on-line identity, they do not guarantee optimal use and appropriate protection of the off-line identity.

Keywords: Privacy, Cryptography, Economics, Externalities

Preliminary draft. Final version forthcoming in J. Camp and R. Lewis (eds), The Economics of Information Security, Kluwer, 2004

Introduction

Several technological approaches have been proposed to solve the problem of personal privacy. In almost any conceivable scenario - when making purchases, browsing the Internet, responding to surveys, or completing medical tests - the identity of an individual can be dissociated from the rest of the information revealed during the transaction. The companies based on those technologies, however, have struggled to balance the differing needs of the various parties in the privacy equation, eventually failing to gain widespread adoption. While privacy and security of personal information remain a concern for many, the economic

incentives have not generated widespread adoption, and government intervention has increased the responsibilities for companies to collect personal information, without determining their liabilities for misuses of those data. Privacy, so it seems, is more difficult to "sell" than to protect.

One of the causes of these difficulties lies in the ambiguity of the very concept of privacy. Privacy means different things to different people, including the scholars who study it, and raises different concerns at different levels. Hence "protecting privacy" is a vague concept. Not only different parties might have opposite interests and views about the amount of information to disclose during a certain transaction, but also the same individual might face trade-offs between her need to reveal and her need to conceal different types of personal information.

But trade-offs are the domain of economics - even when not all dimensions of a problem are economically measurable. Posner, 1978, Posner, 1981, and Stigler, 1980 (as well other contributors to the Spring 1978 issue of the *Georgia Law Review* and the December 1980 issue of the *Journal of Legal Studies*) were among the first to discuss privacy from an explicitly economic perspective. The orthodox economic view suggested that market forces and economic laws, if left alone, would eventually result in the most efficient amount of personal information being exchanged. Individuals and entities interested in information about individuals would converge to that equilibrium regardless of the initial allocation of privacy rights.

After a long silence, economic analysis focused again on privacy at a moment (roughly, the second half of the 1990s) when both privacy intrusions and technologies for privacy protection were dramatically expanding. Concepts such as encryption, National Information Markets, and secondary use of personal information appeared in the analysis. While some (like Noam, 1996) maintained that technology such as encryption would "not create privacy," but simply cause consumers to be paid more to give it up, others started noticing the emergence of externalities (Varian, 1996) and even the possibility of market failures (Laudon, 1996).

The panorama today, with both anecdotal evidence of growing privacy costs and intrusions (Gellman, 2002) and reports of scarce adoption and success of privacy technologies and initiatives, offer arguments to all sides: those who believe that individuals act rationally when they choose not to adopt privacy technologies; and those who consider individual customers stuck in an impasse they are unable to cope with alone. At the same time, however, a *new* economics of privacy has emerged, its novelty being the application of formal micro-economic modelling

to various privacy considerations (Acquisti and Varian, 2002, Calzolari and Pavan, 2001, Taylor, 2002, and a growing literature thereafter). In what follows I will consider the insights offered by these recent economic approaches to discuss the market for the technological protection of individual information.

1. On-line and Off-line Identities

While my analysis is not restricted to privacy and personal information security issues that arise in e-commerce or Internet transactions, I find it useful to draw from the cryptographic literature on pseudonyms and (un)linkability and distinguish between the "on-line" and "off-line" identities of an individual. The on-line identity might carry information about an individual's tastes, her evaluation of a certain good, her browsing behavior, her purchase history, etc.: the on-line identity is what in an economic model would be called the customer "type." In e-commerce transactions the on-line identity is often associated to cookies or IP addresses used to track customer behavior during and across sessions. On the other side, the off-line identity represents the actual identity of an individual, as revealed by identifiers such as credit card numbers and social security numbers. When I login to Amazon.com with a Hotmail.com email address, for example, I am revealing my on-line identity. When I complete a purchase at Amazon.com with my personal credit card, I am revealing my off-line identity.

Of course, this distinction has several gray areas. In the majority of real life instances the off-line and on-line identities of a same individual are linkable (or, in fact, linked) together because of legacy applications and existing infrastructures. Re-identification or "trail" attacks can expose an otherwise anonymized identity by matching data from different sources. In the Amazon case, I might login with a certain unidentifiable email address and then receive a certain cookie on my computer (two items potentially representing on-line identities). The cookie and the email address could then be linked to my credit card information (the off-line identity) released when I check-out. Now not only Amazon, but possibly also other third parties may be able to link my on-line behavior to my real identity.

Information technology, however, can be used not only to track, analyze and link vast amounts of data, but also to split and un-link pieces of data and keep on-line and off-line information separate in ways that are both effective (in the sense that matching, linking back, or re-identifying information becomes either technically impossible or just costly enough to be no longer profitable) and efficient (in the sense that the transac-

tion can be regularly completed with no additional costs for the parties involved). A purchase history at a merchant site, for example, can be associated to an on-line account whose balance is paid through one of many anonymous payment technologies. Or, information sharing between merchants can be realized through coupons and referrals that do not reveal the identity of the customer. Or, individuals can share files and recommendations in ways that hide their personal identities and yet track their contributions to the system. And so on.

While I will not discuss here the many privacy enhancing technologies that can be used to ensure anonymity and protect individual privacy in several scenarios, I will analyze the economic incentives of the various parties to adopt such technologies.

2. The Economics of On-line Identities

Some recent economic studies (Acquisti and Varian, 2002, Calzolari and Pavan, 2001, Taylor, 2002) have shown something interesting about the economics of privacy in relation to purchase transactions: when information about customers' tastes and purchase history is available and can be shared among sellers, market laws alone might produce Pareto-optimal outcomes. For example, in Acquisti and Varian, 2002, under general conditions allowing firms to use cookies make society better off, because the buyer can benefit from the seller knowing him better and thereby providing him targeted services. In Calzolari and Pavan, 2001, sharing information between sellers reduces the distortions associated to asymmetric information between buyer and seller. In Taylor, 2002, when the seller is facing strategic customers, she will autonomously tend to adopt a policy that protects the privacy of her customers. In a more abstract framework, Friedman and Resnick, 2001 have found that "the distrust of newcomers is an inherent social cost of easy identity changes," but persistent pseudonyms can help both the society and the individual. Do these results then support the 1980s economic view of an eventually self-regulating market for privacy? Something must be noted: what these papers have in common is that they all deal with individuals as (economic) agents whose profiles might include information on taste, purchase histories, price sensitivity or risk aversion, etc., but not necessarily information about those individuals' off-line identities. This literature shows that, while distortionary forces might also be in action, for several types of transactions market laws tend towards fair use of on-line information. To put it another way, this literature tells us that there might be economic benefits from sharing and increasing the use

of on-line information, and that these benefits would not be harmed by the protection of the off-line information.

Existing information systems, however, are built in ways that link on-line and off-line identities of their users. With the growth of e-commerce and the diffusion of the Internet these linkages have caused increasing concerns about the practices and protection that other parties (such as merchants) will adopt for an individual's off-line, personal information. At the peak of the privacy scare in the late 1990s, several surveys found that identity thefts and credit card frauds were the main concerns of individuals using new information technology, and that billions of dollars were lost in missed sales because of these concerns. These surveys supported the view that there are in fact economic reasons to protect the off-line identity of individuals.

On the other side, a number of more recent surveys, anecdotic evidence, and experiments (see Spiekermann et al., 2002), have also shown that individuals are actually less concerned about privacy than what they claim to be: many are willing to provide very personal information, in exchange for small rewards. From an economic perspective, one could make the argument that those individuals who demand privacy but take no action to protect theirs, are actually acting rationally. They discount the potential losses from losing control of their personal information (uncertain, but possibly large) with the probability that such an outcome will take place (uncertain, but perceived as low). Then, they compare the resulting value with the implicit or explicit costs of using an anonymizing technology, which are certain and immediate. All things considered, most individuals will therefore decide not to go through the hassle of hiding their off-line information. Some might simply decide not to purchase on-line (or not to use credit cards). Only a few will choose the anonymizing technology.

So: personal preferences respected and market equilibrium re-established even in absence of wide protection of the off-line information? Well, not necessarily. As progresses in information technology make the dissemination and use of information so inexpensive, new complexities arise.

3. The Economics of Off-line Identities

First, given that the individual loses control of her personal information and that information multiplies, propagates, and persists for an unpredictable span of time, the individual is in a position of information asymmetry with respect to the party she is completing transaction with. Hence, the negative utility coming from future potential misuses of off-line personal information is a random shock practically impossible

to calculate. Because of identity theft, for example, an individual might be denied a small loan, a lucrative job, or a crucial mortgage.

In addition, even if the expected negative utility could be estimated, I put forward the following hypothesis: when it comes to security of personal information, individuals tend to look for immediate gratification, discounting hyperbolically the future risks (for example of being subject to identity theft), and choosing to ignore the danger. Hence, they act myopically when it comes to their off-line identity even when they might be acting strategically for what relates to their on-line identity.

If individuals are myopic about the future potential risks related to their off-line identities, and do not act optimally, the other parties they interact with have little incentive to take the burden of protecting the personal data of those individuals. The database of a merchant, for example, might be hacked and the credit card numbers stored there might be stolen and then illegally re-used, without the individuals being able to know where the "leak" took place and without the merchant (in almost all occasions) having to pay for it. This implies that without liability for misuse, abuse, or negligence in handling personal information, moral hazard ensues on the side of the other parties.

Finally, since the market of privacy conscious individuals willing to pay for their protection is small, it ends up not being satisfied. The economic rationale can be described in the following way. Since the only economic interest in protecting personal information seems to belong to the owner of that information, who is also subject to "immediate gratification," the profit margins in this area of business are low. Since few people are so conscious about their information security needs to be willing to pay for it, the size of the market is in addition very small. Low margins and small demand make it very hard for any technology to succeed - except in niche (and possibly disagreeable) markets. Now: while actual usage costs of privacy enhancing technologies are low once adopted, their adoption fees are high because they involve significant switching costs. Hence, as merchants decide against offering anonymizing technologies to their customers, the privacy concerned customers choose not to purchase on-line, or to purchase less. A latent, potentially large market demand remains therefore unsatisfied.

4. Economics and Technology in Privacy Protection

While market forces might ensure fair use of data connected to the on-line identity of individuals, they do not guarantee optimal use and appropriate protection of the off-line identity. In fact, the evaluation of

current dominant practices in the handling of privacy and personal information (on-line and off-line) shows that self-regulation has not provided the results expected by the Federal Trade Commission (2000). Information technology, on the other side, can be used to split on-line and off-line identities or make the linkages between the identities of an individual too costly for any practical application. But without economic incentives no technology reaches widespread adoption.

So, what can economics do?

Firstly, in specific instances, economics can be used to define mechanisms which are privacy enhancing. For example, in anonymous protocols based on the interaction of many agents (see, e.g. Acquisti et al., 2003), economics can assist in the design process of mechanisms to solve the impasse when no party alone would have the incentive to perform certain actions (for example, sending dummy traffic to other parties in order to increase the level of anonymity in the system). Under an appropriate incentive compatible contract, different parties might be induced to support each other and therefore the anonymity of the system. Secondly, and more generally, in the framework of socially-informed design of privacy technologies economics can be used to define what information should be shared, and what protected.

Thereafter economics will need to be assisted by law and technology to actually achieve the balances it proposes. Market forces might ensure fair use of data connected to some pseudo identities of individuals. However, because of the adoption costs and trade-offs analyzed in the previous section, they do not guarantee optimal use and appropriate protection of her legal identity. In these cases, legal intervention, on the model of the EU directive on data protection, or as proposed in Samuelson, 2000, should place constraints and liabilities on the side of the parties receiving private information, calibrating them in order to compensate the moral hazard and asymmetric information in the market of personal data, and combining them with information technology as a "commitment" device in the system.

By generating incentives to handle personal information in a new way, appropriate legal intervention can allow the growth of the market for third parties providing solutions that anonymize off-line information but make it possible to share on-line profiles. By designing the appropriate liabilities, that intervention can also fight the tendency of "trust-me" or self-regulatory solutions to fail under pressure. If privacy is a holistic concept (Scoglio, 1998), only a holistic approach can provide its adequate protection: economic tools to identify the areas of information to share and those to protect; law to signal the directions the market should thereby take; and technology to make those directions viable.

References

Acquisti, Alessandro, Dingledine, Roger, and Syverson, Paul (2003). On the economics of anonymity. In *Financial Cryptography - FC '03*, pages 84–102. Springer Verlag, LNCS 2742.

Acquisti, Alessandro and Varian, Hal R. (2002). Conditioning prices on purchase history. Technical report, University of California, Berkeley. First draft: 2001. Presented at the European Economic Association Conference, Venice, IT, August 2002.

Calzolari, Giacomo and Pavan, Alessandro (2001). Optimal design of privacy policies. Technical report, Gremaq, University of Toulouse.

Friedman, Eric J. and Resnick, Paul (2001). The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy*, 10(2):173–199.

Gellman, Robert (2002). Privacy, consumers, and costs - how the lack of privacy costs consumers and why business studies of privacy costs are biased and incomplete. `http://www.epic.org/reports/dmfprivacy.html`.

Laudon, Kenneth C. (1996). Markets and privacy. *Communications of the ACM*, 39(9):92–104.

Noam, Eli M. (1996). Privacy and self-regulation: Markets for electronic privacy. In *Privacy and Self-Regulation in the Information Age*. National Telecommunications and Information Administration.

Posner, Richard A. (1978). An economic theory of privacy. *Regulation*, pages 19–26.

Posner, Richard A. (1981). The economics of privacy. *American Economic Review*, 71(2):405–409.

Samuelson, Pam (2000). Privacy as intellectual property. *Stanford Law Review*, 52(1125).

Scoglio, Stefano (1998). *Transforming Privacy: A Transpersonal Philosophy of Rights*. Praeger, Westport.

Spiekermann, Sarah, Grossklags, Jens, and Berendt, Bettina (2002). E-privacy in 2nd generation e-commerce: Privacy preferences versus actual behavior. In *3rd ACM Conference on Electronic Commerce - EC '01*, pages 38–47.

Stigler, George J. (1980). An introduction to privacy in economics and politics. *Journal of Legal Studies*, 9:623–644.

Taylor, Curtis R. (2002). Private demands and demands for privacy: Dynamic pricing and the market for customer information. Technical report, Department of Economics, Duke University.

Varian, Hal R. (1996). Economic aspects of personal privacy. In *Privacy and Self-Regulation in the Information Age*. National Telecommunications and Information Administration.