The Paradoxical Value of Privacy

Paul Syverson
Naval Research Laboratory
syverson@itd.nrl.navy.mil

March 14, 2003

Abstract

We consider some common assumptions about the value placed on privacy in society. We observe that:

1. Contrary to popular accounts, individuals are not obviously irrational in how they value privacy.
2. Current governmental and economic structures do not properly place the cost of privacy, thus skewing incentives and behavior.
3. Security of institutions may decrease and infrastructure costs may be increased by a reduction in privacy.

1 Individual Valuation of Privacy

It is commonplace to note that in surveys people claim to place a high value on privacy while they paradoxically throw away their privacy in exchange for a free hamburger or a two dollar discount on groceries. The usual conclusion is that people do not really value their privacy as they claim to or that they are irrational about the risks they are taking.

I claim that there need be no inconsistency inherent in such behavior. Suppose a hamburger is worth two dollars, a full blown identity theft costs an average of 100K dollars, and the probability of such identity theft from giving name, address, and phone number to the hamburger vendor is 10^{-10}. In this case, the rational action is to trade the information for the hamburger. Expected value of such a transaction is still effectively two dollars.

But even assuming these numbers are reasonable, this example reflects a short-sighted consumer. Suppose the incremental probability given a previous history of such transactions is on average slightly higher, say 10^{-9}. A thousand such transactions reduce the long term average expected value to a dollar. Thus even in the relatively long run, the consumer made no mistakes.

This is a very simplistic example. It overlooks the cost of discomfort the individual feels from her information being held by the vendor, the inconvenience from receiving resulting unwanted junk mail or the positive value if the consumer actually desires, e.g., the resulting coupons she receives, etc.

The cost of the discomfort felt at the collection of information is especially difficult to quantify. But, it may be reasonable to completely remove it from any analysis. For it is the expectation of how that information will be used that is significant. If such data were collected such that the individual felt genuinely sure that it would simply be filed away and never accessed, never correlated with any other actions of hers, never used in any way, it is unclear that she would care. Of course there is always some expectation that if an effort is made to collect the data, then someone intends to use it in some way. In any case, even adding such costs as the

increase in junkmail, the expectation of unpleasant inferences about her by marketers, financial institutions, etc. it is at best unclear that the expected cost exceeds the value of the hamburger.

So, what is going on? Are privacy advocates just fanatics, themselves irrational about such things? Some have concluded as much with less justification. But there are other aspects to this issue.

First, the above numbers, however plausible, are made up. A shift of a few orders of magnitude could change things drastically. Second, real numbers are virtually impossible to come by. It might be possible to collect data on occurrence of identity theft correlated with consumer behavior so that probabilities of at least such clear privacy problems could be assigned to some actions. However, this is at best unclear and has not been done yet. And even this would ignore the other types of privacy cost, a few of which we have mentioned. Also, limiting ourselves to identity theft for the moment, any data collected would be of limited predictive value. According to the US FTC, the rate of identity theft is doubling every year. Obviously if true, that cannot continue for long. The situation is just too dynamic right now. And, the market typically needs to learn from experience, so consumer behavior is likely to lag behind any current reality. So one answer is that the expected cost of privacy compromise, both large and small, is increasing. Privacy advocates are just ahead of their time.

Third, the example we have been considering is one involving the assessment of low probability but high value events. This is difficult enough for those who have good numbers and good understanding. Individuals may be somewhat polar in response to these circumstances. Horror stories of lost livelihood are met with sympathy but an expectation that it won't happen to me. And historically that has been statistically accurate. But, there may come a tipping threshold that will make this a major issue not just in polls but in individual behavior and in individual demands of government and business. Alternatively, the right sort of individual soundbite may resonate through society. A recent story in MSNBC [5] recounts the plight of Malcolm Byrd who besides economic suffering, job loss, etc. has been arrested many times and spent time in jail more than once as the result of an identity theft.

2 Allocating the Cost of Affecting Reputation

Why have we focused so much on identity theft? In addition to the above points, it illustrates how the allocation of the costs in protecting privacy do not currently reflect the value and incentives of those with control over its protection.

Malcolm Byrd ended up in jail because the primary cost of misidentifying him was not born by the criminal who used his name, nor by the police who misidentified the criminal as Byrd, nor by any of the police, prosecutors, employers, credit issuers or others who continue to misattribute crimes to Byrd and act accordingly. The cost has been primarily born by Byrd. In general, while individuals are primarily legally responsible for their reputation, the actions of others (government entities, businesses, etc.) are increasingly causally responsible for how that reputation is constituted. This absurdity has absurd implications.

Current advice to protect oneself against identity theft includes checking one's credit record twice a year (up from once a year only a few years ago). Though prudent in the current US socio-economic environment, making individuals responsible for protecting their identity and reputation by such means is akin to requiring them to leave their homes unlocked while suggesting they check with the local pawn shop to see if any of their things are fenced as stolen. It is not a tremendous comfort that the 'pawn shop' in identity theft is larger, more centralized, and has in recent years made some efforts to return goods to their owners, i.e., correct credit records. Worse, as the far from unique case of Malcolm Byrd illustrates, it may only be a short time before one is well advised to check one's criminal record twice a year as well.

One aspect of a solution is more accurate authentication. This could be taken to mean that every action we take should be scrutinized and properly bound to us. However, the costs of such an approach, both literal and intangible are astronomical. Alternatively, our responsibility

for any action could (at minimum) be proportional to the degree of authentication associating us with that action. Criminal and other personal records are currently reputation management systems with no probabilities (in compiling the entries). However, building such probabilities in is a daunting, perhaps hopeless, task especially given the dynamics of how reliable identifications of various types are.

Another part of the solution would be to structure the incentives in collecting, attributing, and dissemination information to accurately reflect costs. We have been looking at criminal records, but the same applies to other areas. If the sending of preapproved credit offers required that the senders bear the expected cost not just of duly reported fraudulent charges but of the resultant reputation damage, such offers might not be worth sending. Similarly if the expected damage caused by sharing of personal financial data were figured into the value of such sharing, there would be no need to push for legislation to allow people to opt in rather than opt out of such sharing. It would not be worthwhile for institutions to share; indeed the amount of data that is even worth collecting would probably greatly diminish as the responsibility not just the benefit for the correct value of that data were accounted.

How might this more accurate accounting be instituted? This is hard to say. Litigation is an easy answer. Another possibility is government reform of standards of evidence, not just for criminal trial but also for arrest, for attributions in best practice business accounting, etc. Many activities such as misdemeanor crimes and small value economic transactions might better be handled without affecting reputation at all. But any suggestion here would be very speculative.

3 Infrastructure Cost

We have already noted how accurate reflection of the costs of assigning, storing, and disseminating reputation would affect the incentives and behavior of infrastructure elements such as businesses and the components of the justice system. However, even without such reallocation, a more accurate assessment of infrastructure costs might lead to an increased emphasis on privacy.

Spam is a large privacy issue. (This is more from the right-to-be-let-alone aspect of privacy than the personal reputation aspect we have been discussing so far.) But, it's not just an issue of personal inconvenience. Recent estimates of spam put it at approaching half of all email traffic in the US [3]. This is a tremendous overhead born by business, government, and individuals. And, part of it comes from the distribution of email addresses without the consent of those who hold the addresses. (Another part is due to the easy compromise of machines to make them zombie mailers, but that's a subject for others at this workshop to discuss.)

Adam Shostack has noted that criminals already know how to communicate anonymously and privately. For example, they can just steal cell phones for brief use, then toss them and steal more. Another technique noted in the general press is to compromise a web host and leave files there for others to retrieve. Thus, monitoring communication primarily eavesdrops only on the law abiding.

One answer to this is that such activity by criminals involves transactional risk [4]. Thus, providing general private and anonymous Internet communication removes a disincentive to crime. True enough, but the analysis in [4] does not account for the cost of privacy loss. If incorporated, an anonymous communications infrastructure may be more cost effective for the infrastructure providers.

Reduction in privacy also has a cost to security. A commonplace in recent polls is to ask how much privacy people would exchange for increased security. However, it is assumed rather than argued that decreasing privacy increases security. Just the opposite may be true. Law enforcement has made use of anonymous tips for years with the recognition that much of the information so gathered would not have been given without a plausible expectation of anonymity. Very shortly after September 11^{th}, the Anonymizer set up an Web interface "providing anonymous access to the FBI's Terrorism Activity tip page to over 26,000 individuals around the world" [1, 2]. They have since added anonymous interface to the Utility Consumer's Action

Network. Similarly, the Witness Protection Program relies on the ability to assign people a new identity. In an environment in which all commercial and public actions by individuals is monitored, this possibility becomes far less plausible. To effectively monitor to the degree necessary for effective authentication as discussed in section 2, the creation of a new identity would likely be noticed in a commercial database (whose entries would be shared without disincentives to do so). The person who recently turned in Khalid Sheikh Mohammed and received a new identity might not have risked doing so without a plausible new identity possible.

We have argued in this abstract that assumptions about privacy are not justified without further analysis: that individual behavior is inherently irrational with respect to claimed valuation of privacy, and that trading away privacy will enhance security. We further provided initial arguments that the opposite may be true in each of these cases. Finally, we observed that the cost of protecting privacy is not allocated in an accurate way and that a correct reallocation would provide government and business with incentives to increase rather than decrease protection of individual privacy. These are just preliminary observations that require further analysis if they are to be confirmed.

References

[1] Secure tips online program (STOP). `http://www.anonymizer.com/tips/`.

[2] Web site passes anonymous tips to FBI, September 14 2001. `http://www.cnn.com/2001/TECH/internet/09/14/anonymous.tips/`.

[3] Jonathan Krim. Spam's cost to business escalates. *The Washington Post*, page A01, March 13 2003.

[4] Stuart E. Schechter and Michael D. Smith. How much security is enough to stop a thief? In *Financial Cryptography, Pre-Proceedings*, January 2003. Final proceedings forthcoming in Springer LNCS.

[5] Bob Sullivan. The darkest side of ID theft. `http://www.msnbc.com/news/877978.asp?0si=-&cp1=1`.