Internet Privacy: Big Brother and Little Brother

Steven M. Bellovin
http://www.cs.columbia.edu/~smb
Columbia University

March 14, 2008
Introduction

What is Privacy?
A Security Perspective
Big Brother is Watching You
Little Brother is Watching You
One Vision of the Internet
Another Vision of the Internet
A Third Vision of the Internet
Three Visions of the Internet

Types of Privacy Violation

The Privacy Threat from Governments

The Privacy Threat from Corporations

Defenses

Conclusions
“The right of individuals to control or influence what information related to them may be collected and stored and by whom and to whom that information may be disclosed.” (OSI Reference Model)

“Privacy is the interest that individuals have in sustaining a ’personal space’, free from interference by other people and organisations.” (Clarke)

“The right to be let alone.” (Future U.S. Supreme Court Justice Louis Brandeis, 1890)
A Security Perspective

- What are you trying to protect?
- Who is the enemy?
- How much is your privacy worth? How much can your enemy spend?
From George Orwell’s classic book *1984*

Refers to the threat from dictators

Is that a risk on the Internet?
Little Brother is Watching You

- Refers to corporations
- Often harder to stop
- But is it a threat?
“On the Internet, no one knows you’re a dog.”
It wags its tail, it barks, and it acts the same as last time.
A Third Vision of the Internet

Introduction
What is Privacy?
A Security Perspective
Big Brother is Watching You
Little Brother is Watching You
One Vision of the Internet
Another Vision of the Internet

A Third Vision of the Internet

Three Visions of the Internet

Types of Privacy Violation

The Privacy Threat from Governments

The Privacy Threat from Corporations

Defenses

Conclusions

It snarls and bites; we have to deal with it.
Three Visions of the Internet

“On the Internet, no one knows you’re a dog.”

It wags its tail, it barks, and it acts the same as last time.

It snarls and bites; we have to deal with it.

All three perspectives are valid; all three co-exist on the Internet.
Types of Privacy Violation

Introduction

Types of Privacy Violation
Reading Full Content
Behavior Tracking
Example
Web Cookies
Cookie Linkages
Traffic Analysis
Database Linking

The Privacy Threat from Governments
The Privacy Threat from Corporations

Defenses

Conclusions
Types of Privacy Violation

- Full content monitoring
- Behavior tracking
- Traffic analysis
- Linking multiple sources
Typically done via a wiretap
In theory, easily prevented by encryption
But — encryption is hard to use; very few people actually use it
It’s even harder to use correctly
It can also be inefficient
Result: rarely used except to protect passwords and credit card numbers
Behavior Tracking

- Monitor what sorts of sites users visit
- Can be done by login or “cookies”
- Often hard to block
- Very common on sites that require subscriptions
Example

- The New York Times and Wall Street Journal track what articles you read, and select appropriate ads
- Amazon.com tracks your purchases and recommends related items
- Google looks at queries and mail keywords; then serves up appropriate ads
Technically: cookies are small text files sent by a site to your web browser

Intended uses: login name; shopping cart

Actually: used to track you from session to session

Third party cookies: a lot of Internet advertising is done by sites like Doubleclick.com, which set their own cookies

Result: Doubleclick sees your behavior on many sites
Cookie Linkages

- If the web site ever learns who you are, they can associate that with your cookie.
- Web bugs from third parties: one-pixel “clear GIFs”
- If you visit another Doubleclick-using site, Doubleclick will know you’ve been to the N.Y. Times, and even something about what pages you viewed.
Traffic Analysis

- Use wiretaps to see what site a user visits
- Track email the same way
- Very hard to block with encryption
- Frequently used by law enforcement or intelligence agencies
Database Linking

- Combine multiple sources of information
- Example: link address from online purchase to cookies
- Privacy principle: the secondary uses of the data are much more troublesome
The Privacy Threat from Governments

Introduction
Types of Privacy Violation

The Privacy Threat from Governments
Goals
Government Purposes
Considerations
Threat Level

The Privacy Threat from Corporations

Defenses

Conclusions
Goals

- Many governments spy on some Internet traffic
- Some, of course, do it more than others
- Generally aimed at government purposes
Government Purposes

- Foreign intelligence
- Law enforcement
- Tracking domestic dissidents
- These categories can overlap
- Scope depends on type of government
Considerations

- Generally must be completely invisible to person being monitored
- Result: done by wiretaps and by obtaining ISP and web site logs
- (Cookies aren’t used)
Threat Level

- Not a major threat in most democracies
- Some democratic governments abuse their authority, but this can be dealt with politically
- Volume of traffic limits the amount of monitoring possible
- Even democracies monitor foreign traffic
- Dictatorships monitor much more internal traffic
The Privacy Threat from Corporations
Corporate Goals

- Primary goal: making money
- Privacy issue: can corporations make money by tracking users?
- Many of them think so!
Attitudes

From a Wall Street Journal article on how Nielsen will use set-top boxes to monitor viewership:

Supporters of set-top-box data say it is more useful to marketers and less burdensome to participants than traditional Nielsen ratings. The set-top boxes cover many more households, and, unlike the panels, researchers don’t have to secure agreement from those households to participate.
Techniques

- Cookies
- Third-party cookies
- Cooperation with third-party advertising sites: pass information both ways
- Web site registration
Wiretapping?

- Some ISPs monitor user traffic
- Often for traffic engineering
- Example: Comcast (a US ISP) interfering with BitTorrent
- Some ISPs (i.e., Rogers in Canada) insert ads in some traffic
- Some British ISPs, including BT, are watching users’ web traffic for sale to advertisers
- In some countries, monitoring is done to deal with copyrighted material
Corporations can sell data about users

Often, they won’t do that — but they’ll sell aggregate data

User behavior data used for targeted ads, online and offline

Sometimes, they sell data to the government that the government can’t collect by itself
Defenses
Defenses

- Encryption
- Onion routing
- Discarding cookies
- Lying
- Laws
Encryption

- As noted, hard for most users
- Only protects against wiretappers
- Does not protect against attacks from end-site, such as cookies or third-party cookies
Onion Routing

- Forward through random path
- Defeat wiretappers and traffic analysts
- But — slow
- Doesn’t protect against attacks from end-site
To G, both traffic streams are coming from D. However, the two streams would have different cookies.
Discarding Cookies

- Some people delete all of their cookies periodically
- Most people don’t
- Other tracking technologies exist, i.e., Macromedia Flash cookies
Some people lie when registering for web sites

Give fake ages, location, gender

Some sites notice, some don’t; most don’t care
Some countries have strong privacy laws
But — the net is international
Does EU privacy law apply to American companies?
Laws won’t stop foreign intelligence
Conclusions
The Three Visions

- Without outside information, there is some privacy against ordinary attackers — they don’t know if you’re a dog
- Corporations mostly care about behavior, and track that
- Governments have to deal with criminals
What is the balance?
How can governments track real criminals without invading the rights of others?
Advertising pays for the net; it’s more effective with tracking
What is the right answer?