Chapter 5, Tracking and Surveillance

Lorrie Faith Cranor, Manya Sleeper, Blase Ur

This is the authors' draft of a chapter that will appear in the IAPP Privacy and Information Technology text book, expected to be published in 2013.

Chapter 5, Tracking and Surveillance

Lorrie Faith Cranor, Manya Sleeper, Blase Ur

5.1. Introduction

In today's digital world, both electronic and non-electronic actions and communications have the potential to be tracked and surveilled. Reasons for tracking are manifold. On one level, advertising companies wish to profile users so that they can better target relevant ads. Other groups or individuals might wish to spy on a person for the purpose of blackmail, extortion, or to cause embarrassment. On a grander scale, a government or other organization may hope to gather intelligence to thwart an act of terrorism, or perhaps to spy on a group or individual for more insidious purposes. Our goal in this chapter is to provide IT professionals with a richer understanding of the techniques that enable tracking and surveillance on a number of different levels, alongside an explanation of countermeasures and their limitations. As someone responsible for developing IT within a company, government agency or other organization, an IT professional must be aware of the impact that surveillance technology has on individual privacy.

We start by looking at how usage of the Internet can be monitored. Following a brief overview of the main technologies underlying communication on the Internet, we discuss how network traffic can be surveilled on a large scale using techniques such as Deep Packet Inspection. We also discuss more localized approaches for tracking all of a user's communications, such as eavesdropping on a Wi-Fi connection or monitoring a school or workplace network, before explaining how anonymizing systems can help defend against this sort of tracking.

Next, we turn our attention to web tracking, in which companies work to profile the websites that a particular user visits, most prominently for advertising purposes. We discuss how a combination of HTTP cookies and an ecosystem in which a handful of advertising companies serve advertisements on many popular websites enables this sort of tracking. We also explain a number of additional mechanisms that enable tracking even when HTTP cookies are disabled or deleted. We briefly survey how a user's web

searches can be tracked, and we show how the sender of an email can employ tricks to determine exactly when an email recipient opens the message.

A number of web blocking tools have been designed to combat the myriad techniques through which a user's web browsing can be tracked. We introduce a range of these tools, some of which work automatically or are as simple to use as clicking an obvious button in software that a user has already installed. However, many privacy tools must be installed separately and require time-consuming and potentially confusing configuration. Along with pointing out the features of the assortment of tools available for privacy-conscious users, we highlight the shortcomings and usability issues of each approach.

Tracking and surveillance are not limited to communications being intercepted or web-browsing behaviors being logged. We show how the mere presence of a cellular phone or RFID chip in an individual's pocket can reveal his or her location, in addition to explaining the relationship between GPS technologies and tracking. Next, we discuss how the use and misuse of location-sharing features on social media sites can leak potentially private information. We also explain how users might control the disclosure of location information at a finer-grained level. Finally, we provide a brief overview of audio and video surveillance techniques, as well as emerging surveillance issues in medical devices and smart energy meters.

It is essential that IT professionals understand the technologies that enable tracking and surveillance in order to prevent privacy violations. From a corporate standpoint, it is critical to avoid privacy violations that could lead to negative press reports, a tarnished reputation, or regulatory or legal consequences. It is also important to avoid making users feel that their privacy has been violated since they may stop using products from companies that they do not trust. In addition, knowledge of tools that limit tracking and surveillance is valuable both in illuminating possible opportunities for a company to compete on privacy and in understanding steps users take to protect their own privacy.

5.2. Internet Monitoring

The Internet provides a variety of opportunities for tracking and surveillance. This includes activities that range from automated monitoring by network administrators for detecting malicious software to illegal monitoring by criminals who are trying to steal passwords and account information. Employers may monitor their employees' Internet connections to enforce company policies. Law enforcement may obtain warrants that allow them to monitor Internet traffic to investigate a crime. Internet Service Providers (ISPs) or email providers may monitor web browsing behavior or email to deliver targeted advertising. This section describes the basics of how data travels across the Internet and how such data can be monitored in transit. It also outlines several defenses against Internet monitoring and surveillance.

5.2.1. An Overview of Computer Networking

To travel from source to destination on the Internet, data must be directed across intermediate networking links. Monitoring tools may be placed anywhere along the path that data travels.

5.2.1.1. Packets and IP Addresses

The Internet is designed so that many senders and receivers can share networks efficiently. To achieve this goal, messages sent across the Internet are broken into small segments and encapsulated in packets. In addition to the data itself, a packet contains information needed to deliver that segment. The Internet Protocol (IP) that governs data transmission on the Internet specifies both a hierarchical system for assigning addresses to Internet users and a particular format for these packets. The address assigned to users is their IP address, and the packets of data are known as IP packets.

An IP address is a numerical identifier for a particular device. IPv4, which is currently in wide use, specifies addresses consisting of four numbers separated by periods. As an example, 128.2.42.10 is the IP address for Carnegie Mellon University's main website. Each of the four numbers ranges from 0 to 255, which means that over 4 billion unique addresses can be created. IP addresses are assigned hierarchically, which enables them to be routed to their destination. For instance, all IP addresses beginning with 128.2 belong to Carnegie Mellon University. The Internet is currently transitioning to IPv6, which is somewhat similar to IPv4, although it can support many more addresses.

Each IP packet consists of a header and the data payload. The header includes the IP addresses of the sender and recipient, as well as a checksum for verifying that the packet does not contain errors. Using the information included in the header, each router on the Internet passes a packet on to the next router closer to

its final destination. Once packets reach their final destination, they are reassembled into the original message.

5.2.1.2. Email

One of the myriad types of communications split into packets for transmission across the Internet is electronic mail, or email. Emails are structured in two parts: a header containing information about the message, and a body that contains the message itself.

A user creates an email message using a Mail User Agent (MUA). Examples of mail user agents include desktop email clients like Microsoft Outlook and webmail clients like Gmail. The email message is made up of a message header and a body, all of which is contained within the body of one or more IP packets. The header includes a variety of addressing fields, such as the sender and recipients' email addresses, the subject, and cc'd recipients. The body includes the email message.

The email message is transmitted to the user's mail server and then sent across the Internet to its destination using the Simple Mail Transfer Protocol (SMTP). This protocol uses packet routing to transfer the message to its destination. Once the email reaches its destination mail server, it is available for access either directly or using a mail server protocol, such as the Post Office Protocol (POP) or Internet Message Access Protocol (IMAP). In POP, the MUA removes the emails from the server after storing them locally. When using IMAP, the emails remain on the server for access later, or for access by multiple mail clients across different devices.

5.2.1.3. The HTTP and HTTPS Protocols

Like emails, webpages are split into packets as they are sent across the Internet. However, whereas SMTP specifies how emails are sent between servers, the Hypertext Transfer Protocol (HTTP) specifies how webpages are transmitted to browsers.

Users typically visit webpages using web browsers, such as Microsoft Internet Explorer, Mozilla Firefox, Google Chrome, or Apple Safari. The address that a user types into the browser is known as the URL, or Uniform Resource Locator, and it contains four main parts. As an example, a user might send a request to the IAPP's website using the following URL:

https://www.privacyassociation.org/about_iapp
[___] [_____] [_____]
service host resource

The *service* component of a URL specifies the protocol that will be used for the request. Most commonly, webpages use HTTP for communication between a web browser and the web server that hosts the page. Messages sent over HTTP are sent in plain text, and thus are susceptible to monitoring and tampering by any of the intermediary nodes through which the packets are sent. To prevent data traveling over the Internet from being monitored or tampered with, the HTTPS protocol can be used. The main difference between HTTPS and HTTP is that data in HTTPS is encrypted using the Secure Sockets Layer (SSL) protocol or its replacement, Transport Layer Security (TLS). Thus, the presence of "HTTPS" in place of "HTTP" in the browser's URL indicates an encrypted connection.

The *host* portion of the URL specifies who will receive the request, most often a computer server owned or contracted by the group represented by the website. The host can also be referred to as the site's domain. Along with the host, a *port* can optionally be specified. Similar to the way apartment numbers enable mail to be routed to a particular resident of a large apartment building with a single street address, ports allow numerous programs and processes on one computer to communicate simultaneously without jumbling the conversations. Although a single computer usually has tens of thousands of different ports, there are default ports to which requests following a particular protocol should be made. For instance, HTTP requests are sent to port 80 by default, while HTTPS requests are sent to port 443. Since no port was specified in the URL above, the default port for the HTTPS protocol will be used. Finally, the *resource* portion of the URL specifies exactly which page, image, or other object should be returned.

An HTTP request for web content is first initiated by the user's web browser, which sends a request message to the host. First, the name of the *host* is converted to an IP Address. For instance, the IAPP's hostname of *privacyassociation.org* is converted to the address *50.116.48.191*. Once the host's IP Address is known, the request can be broken into packets and accurately routed to its recipient.

A browser can make one of several types of HTTP requests, of which GET and POST requests are most relevant for our discussion of surveillance techniques. In the case of a GET request, the browser simply requests that a particular resource be returned. A POST request is similar to a GET request except that the browser also sends information to the server to be processed, such as the data that a user typed into an online form. For either type of request, the server replies with a short status code indicating the success of the request (e.g., the code 200 indicates a successful request, whereas 404 indicates that the resource was not found). For successful requests, the content is also included in the body of the response. In the HTTP protocol, a single request results in a single reply. Furthermore, the connection is stateless, which means that the server is not required to recall anything about past requests to fulfill future requests.

Both HTTP requests and HTTP responses can include headers, which are short messages containing additional information. For instance, HTTP requests can include the date and time the request was sent in the *Date* field. When a user clicks on a hyperlink on a page, an HTTP request for the new page is initiated. For these sorts of requests, a *Referer* header is included to indicate to the new page the identity of the page on which the link was clicked. An HTTP response header can include fields like *Content-Language*, which specifies the natural language in which the page is written.

5.2.2. Network-Scale Monitoring and Deep Packet Inspection

Network hardware only needs to examine the IP header, the first part of a packet, to route a packet to its destination. However, it is possible for network hardware to look further into the packet and examine header information for other protocols, or the full body of the packet itself. This practice is known as deep packet inspection**.**

Deep packet inspection can serve a variety of purposes. For example, deep packet inspection can be performed at the edge of an organizational network. Examining the contents of packets before they pass into a local organizational network can help determine whether or not the packets contain malicious content, such as known viruses. Alternatively, examining packets before they leave an organizational network can help prevent data leaks if the organization searches the packets for potentially sensitive content.

Deep packet inspection can also be used for non-organizational purposes. It is used by advertisers to track users' online behavior to better target ads [W+08] and by government entities to censor or track citizens' online behaviors. Both of these activities raise privacy concerns. In China, deep packet inspection is used as part of the "Great Firewall," which the government uses to perform large-scale censorship on potentially sensitive topics [S+11].

Opponents of deep packet inspection view it as a violation of net neutrality, the principle that packets should be treated the same regardless of their contents. For instance, in the U.S., the Internet Service Provider Comcast used deep packet inspection to determine the type of content being sent and subsequently rate-limiting the connections of customers using peer-to-peer networking applications or sharing files. The U.S. Federal Communications Commission ordered Comcast to stop using deep packet inspection to throttle peer-to-peer connections [FC+08].

5.2.3. Wi-Fi Eavesdropping

Internet monitoring can also occur on wireless networks. It is possible to eavesdrop on, or capture data being sent over, a wireless (Wi-Fi) network at the packet level. Several systems for Wi-Fi eavesdropping, including packet sniffing and analysis tools, are available off the shelf.

Unsecured communications sent over a shared wireless network can be intercepted easily. This risk is often present at Wi-Fi hotspots in public spaces, such as hotels or coffee shops, where many users share a common Wi-Fi network that is either unprotected by a password or protected using a password known to a large group of users.

Packet-sniffing systems capture packets sent over Wi-Fi networks. If the data is unencrypted, these packets can then be examined and reassembled. These reassembled packets may provide information about all the network user's activities, including websites he or she visits, emails and files sent, and the data included in session cookies, including website authentication information. Wireshark[1] is one example of a packet-sniffing and network-analysis tool. It captures packet-level data on wired or wireless networks to

[1] http://www.wireshark.org/

which a user has access, allowing a user to examine and reassemble packet content. Other examples of packet sniffers include Kismet[2] for Unix and Eavesdrop[3] for Mac.

There are also more specialized Wi-Fi eavesdropping systems. One such tool enables HTTP session hijacking, or "sidejacking," attacks. When a user logs into an Internet site, the initial login process is usually encrypted. Sites often then store a token on the user's computer, and this token is sent along with future HTTP requests as proof that the user has logged in. However, some popular sites use HTTP, rather than HTTPS, to send this token, which means the token is sent unencrypted. Firesheep[4] is a Firefox extension that enables an adversary listening to this Wi-Fi connection to hijack these tokens and impersonate a user who is logged in. It captures tokens it sees transmitted across the wireless network to which the user is connected. It then displays information about the captured accounts (e.g., the site, the username, and the user's photo) and allows the adversary to send requests to the applicable website as if they had originally logged in as that user. At the time of its deployment, it was commonly used to allow users to log into Facebook using other users' accounts.

Tools like Wireshark, Eavesdrop and Firesheep allow users to purposefully eavesdrop on network information included in the body of packets for both benevolent and malicious purposes. However, when companies are not careful about how much data they record, they may violate consumer privacy unintentionally or run afoul of regulations. For instance, in 2010, Google sought to capture public Wi-Fi network data using packet sniffers installed on their Street View cars. When auditing the data at the request of the German government, Google discovered that they also had recorded information contained within the packets.[5] This discovery led to a public outcry and an investigation by the U.S. Federal Communications Commission, as well as actions by other governments across the world.[6]

There are several potential defenses against Wi-Fi eavesdropping. First, Wi-Fi eavesdropping requires that the eavesdropper have access to the Wi-Fi network and be able to read the packets that are sent. Ensuring that Wi-Fi networks are protected by strong passwords can therefore limit the danger of Wi-Fi eavesdropping by preventing some adversaries from reading network traffic. However, strong Wi-Fi passwords are often not sufficient to protect this communication channel. For instance, the WEP Wi-Fi encryption scheme that is still in use has significant vulnerabilities [BGW+01] and can often be broken within seconds. More modern security schemes for Wi-Fi routers can also sometimes be defeated [CERT+11]. One countermeasure is the use of Virtual Private Networks (VPNs), which allow users to create secure, encrypted tunnels to send data, offering a defense against interception on unsecured networks. Additionally, encrypting web requests using HTTPS can prevent eavesdroppers from intercepting sensitive or personally identifiable data regardless of the security of the network itself.

5.2.4. Internet Monitoring for Employers, Schools, and Parents

There are also systems available for people in positions of authority to monitor local networks. Some of these systems are specifically designed for employers, schools, and parents. Such monitoring often occurs both for security purposes and to ensure appropriate behavior on the network, often by blacklisting, or blocking access to, websites considered inappropriate.

Employers in the United States are legally allowed to monitor their employees' Internet usage on their organization's network or company-owned machines [NMN+06]. Companies monitor e-mail and Internet usage for a variety of reasons, including tracking employee productivity, maintaining security within the corporate network, and ensuring appropriate behavior among employees. As of 2007, the American Management Association found that 66% of surveyed companies monitored Internet browsing and 65% blocked access to blacklisted websites. Forty-three percent of companies also monitored e-mail, both to limit security breaches and to prevent potential lawsuits [AMA+08]. For example, some companies decrypt HTTPS traffic coming from their internal networks, enabling them to read communications that would otherwise be secure. These companies add their own certificates to workplace computers' web browsers, enabling the

[2] http://www.kismetwireless.net
[3] http://www.baurhome.net/software/eavesdrop/
[4] http://codebutler.com/firesheep
[5] http://googleblog.blogspot.com/2010/05/wifi-data-collection-update.html
[6] http://online.wsj.com/article/SB10001424052702304356604577344171454221422.html

company to impersonate the external web server and intercept traffic before forwarding it to the external server. This technique is a type of man-in-the-middle attack.

Similar monitoring occurs in schools. The Children's Internet Protection Act (CIPA) requires U.S. schools and public libraries to install filters to prevent children from viewing inappropriate content online [FCC]. Many schools also track students' Internet usage to prevent inappropriate behavior, such as illegal downloading [S+01].

Parents can also monitor their children's Internet usage. There are a variety of tools available that allow parents to limit the types of sites their children are allowed to visit, typically using blacklists. Parents can also track the sites their children visit and view the emails and chat messages the children send. These tools are available for protecting children from online predators as well as for allowing parents to supervise their children's online activities.

5.2.5. Spyware

Beyond monitoring the network connection, malicious software can also surveil data before it even leaves the user's computer. Spyware is malicious software that is covertly installed on a user's computer, often by tricking users through social engineering attacks. Spyware can then monitor the user's activities through a variety of methods. It can track online activity in a number of ways, including capturing cookie data to determine browsing history or directly monitoring and reporting on browsing behavior. Spyware can also monitor what a user is doing on his or her computer by transmitting an image of the user's screen back to the attacker or by keylogging. In keylogging, malicious software records all of a user's keystrokes. This keystroke data, which is sent to the attacker, often includes passwords and other sensitive information.

Anti-virus providers offer anti-spyware programs that can be used to protect against known spyware. These systems use anti-malware signatures that are created from components of the malware code to detect and remove spyware from installed software. However, these programs are often reactive; they successfully detect well-known malware, but fail to detect new types. Systems also defend against spyware by tracking known spyware websites and blocking installations coming from those sites at the firewall level [M+03].

5.2.6. Preventing Network-Level Surveillance

Even if packets are encrypted when a user accesses the Internet, it is often possible for an observer, such as an ISP or government, to learn what sites the user accesses by examining the headers of the packets. Such tracking is problematic in a variety of circumstances, such as when citizens who live under a hostile regime wish to access political websites secretly.

To prevent such tracking, anonymizers can be used to mask the link between the user and the destination of his or her network traffic. Two popular types of anonymizers are anonymous proxies, which forward a users' traffic through an intermediary, and onion routers like Tor, which forward traffic over a series of relays in a way that is very difficult to trace.

Anonymous proxies allow users to anonymize their network traffic by forwarding the traffic through an intermediary. Thus, the user's traffic appears to come from the proxy server's IP address, rather than the original user's IP address. JonDonym[7] is a service that anonymizes traffic by routing packets through a mix of multiple, user-chosen anonymous proxies. However, the use of an anonymous proxy requires that the user trust the anonymous proxy, and this approach runs the risk of presenting a single point of failure.

An alternative to anonymous proxies is the use of onion routing systems. The most popular onion routing system is Tor,[8] which uses a series of volunteer-run relays to enable anonymity. Similar to the layers of an onion, packets that are sent through Tor are encrypted in layers that each have a different encryption key. These encrypted packets are then sent through a series of relays. When a relay receives a Tor packet, it strips off the outermost layer of encryption, revealing the next relay to which the packet should be forwarded. Assuming that the relays do not collude, each relay along the way will only know from whom it received a packet, and to whom it forwarded the packet. This configuration allows a layer of anonymity to be inserted between the source and destination of traffic. However, encryption is still required to keep the data itself anonymous once it leaves Tor [GRS+96].

[7] https://anonymous-proxy-servers.net/
[8] https://www.torproject.org/index.html.en

5.2.7. Best Practices

There are a variety of tools available for performing Internet monitoring. Thus, there are several best practices to keep in mind when performing or trying to prevent such monitoring.

Internet monitoring presents a tradeoff with privacy. It should have a stated and narrow goal. If it is deemed necessary to monitor online activities, such as within a corporate environment, best practice dictates that it be done in a manner that protects individual privacy while achieving the desired goals. Neglecting privacy in pursuit of a goal can result in negative consequences.

It is essential to determine what data is necessary to achieve the desired goal, and then capturing only this data. This advice was not followed in the 2010 Google Street View example, and this oversight led to substantial bad publicity. Whenever possible, data should be captured in an aggregate or anonymized fashion (see Chapter 4 for more details on data aggregation, sanitization and anonymization). For example, if deep packet inspection is being performed for network statistics, data should be examined in aggregate; the contents of individual packets should not be examined. To further protect privacy and limit liability, data should not be retained or shared beyond a necessary level.

There are several best practices to prevent monitoring. First, HTTPS should always be used for transmitting sensitive information between clients and servers. Additionally, sensitive information should not be sent over unsecured Wi-Fi networks, nor should a user rely on a Wi-Fi network's security to provide confidentiality. Virtual private networks provide some protection in such an environment, although they require that one trusts the operator of the VPN. Finally, sensitive information, such as confidential emails, should also be encrypted to provide an additional layer of protection.

5.3. Web Tracking

Companies and websites have many reasons for wanting to track users on the Internet, ranging from analysis and personalization on an individual website to targeting advertisements to users based on their activities across multiple websites.

Many websites personalize their content or user experience for each visitor. For instance, if a user has changed his or her language settings for a website from English to Romanian, the website would want to identify that user when he or she returns and automatically show the page in the correct language. Furthermore, many websites calculate analytics about their pages, for example to understand how users navigate the page layout. Websites also might want to show a particular user the stories or products that they believe are most relevant to him or her based on past actions or purchases on that website.

Tracking is also used to profile a user's browsing across many different websites. Online advertising companies create detailed profiles about the websites a particular user visits for the purpose of targeting advertisements and offers to that user. If a company tracks exactly what pages a user visits across the Internet over a long period of time, the company can infer information ranging from general categories of user interests to potentially detailed personal information.

On a technical level, the amount of information a determined company could collect about a particular user is nearly limitless. For instance, based on which articles a user views on health websites and the terms he or she enters into search engines, a company could likely infer what diseases the user might have. Furthermore, even seemingly anonymous data can sometimes be connected to a particular individual. For example, in 2006, AOL publicly released supposedly anonymized search queries for research purposes, yet some of the users were identified based on searches they made about themselves or their neighborhoods [BZ+06].

In practice, many companies on the Internet say that they collect far less information than is theoretically possible, usually outlining in a privacy policy the information they collect and how they use it. However, it is not possible for a user to verify the information actually collected, and a user would need to expend hundreds of hours a year to read the privacy policies for every site visited [MC+08].

To better understand the types of information that can be collected and by whom, we spend the rest of this section delving into common mechanisms for tracking Internet users. We begin by discussing how webpages remember settings for a particular user using HTTP cookies, and then show how these cookies can be used for tracking that user across the Internet. These techniques have applications that even privacy-sensitive users might consider benign, yet they can also be used for privacy-invasive purposes.

5.3.1. HTTP Cookies and Web Tracking

The most common protocol for accessing websites is HTTP. Although HTTP is stateless, which means the protocol is not expected to remember past transactions, it is useful for websites to be able to save state about a particular user. For instance, a website should be able to remember when a particular user is logged in. A user would likely find it quite cumbersome to log in anew each time he or she clicked on a link or navigated to a new page. Similarly, if a user places an item into his or her online shopping cart or updates his or her preferences for how the site is displayed, it is useful for the website to remember these changes not just for that particular visit, but for all future visits from the same user.

To remember state, websites can request that the web browser save small text files, known as HTTP cookies, on the user's computer. In an HTTP response from a server, one possible header is *Set-Cookie*, which is followed by the value(s) that the site would like to store in the web browser. Along with the value(s) to be stored, a cookie also contains an expiration date, as well as the domain and path for which that cookie is valid. Taken together, the domain and path define the scope of the cookie, which specifies for which parts of a domain the cookie applies.

Every major web browser has a particular location on the hard drive of a user's computer where it stores these cookies. Users can examine their cookies, which contain plain text. However, while some cookies may contain words indicating what has been saved, cookies more commonly contain codes that are intended to be parsed only by the web site that originally set that cookie.

When a particular website has set one or more cookies on the user's computer, the contents of those cookies may be included as a header in HTTP requests sent to that domain if the resource being requested falls within the cookie's scope and if the same browser is being used. Cookies are included regardless of the type of content being requested. Therefore, requests for content ranging from the main page of a website to a single image from the web server may include cookies.

Cookies can be set just for a particular visit to a website or for extended periods of time. *Session cookies* are cookies that are stored only until the web browser is closed and thus contain information only for a particular visit to a page. For instance, a token that can be used to prove to a website that a particular user has successfully logged into his or her email account would often be saved as a session cookie and sent along with every HTTP request to that site. In contrast, *persistent cookies* can be saved indefinitely and are often used to save website preferences or a unique identifier for correlating multiple visits over time.

Web domains can only read and write cookies that they themselves have set, a practice known generally as the *single-origin policy*. As such, Facebook cannot directly read cookies set by Google, nor can Google directly read cookies placed by Facebook. However, it is often the case that visiting a single website will result in cookies from multiple companies being placed on a user's computer because websites that appear as a single entity to the user may actually be comprised of HTML from many different sources. For instance, a news website might load articles from its own Internet domain (for example, *www.news-website.com*). These sorts of cookies from the primary page that the user is visiting are known as *first-party cookies*. However, images on this page might be downloaded from another company (such as *www.photojournalism-aggregator.com*), while each advertisement on the page might be served by a different advertising network (such as *www.xyz-advertising.com*). Each of these domains can also set their own cookies. *Third-party cookies* are those that are set by domains other than the domain whose URL is displayed in a browser.

The content contributed by third-party companies need not even be visible to the user. Invisible elements of a webpage that are used for tracking are known as *beacons* or *web bugs*. Beacons are loaded onto a page using elements of the HTML markup language, which is the most widely used language for specifying the layout of a webpage. "Tags" are used to specify page layout in HTML, and these tags allow text and multimedia from a variety of different sources to be brought together to form a webpage. For instance, the *<img>* tag instructs a web browser to download an image from a remote web server and include it in the webpage, which is how images are most commonly displayed on the Internet. Similarly, the *<iframe>* tag specifies that an entire webpage should be included inside another webpage, while the *<script>* tag runs computer code written in certain other languages, such as JavaScript.

Since HTTP requests for all types of content include the cookie as a header, and since the cookie can contain an identifier unique to a user, beacons can enable tracking. The most canonical example of a beacon is a one-pixel image whose sole purpose is to generate an HTTP request containing the cookie with the user's unique identifier. However, beacons can also come in the form of fragments of computer code or other files embedded in an HTML page.

Although a company can only track a user's visits to websites on which it serves content, widespread tracking is still possible since a small number of companies include their content and beacons on many popular websites. For instance, a 2009 study identified ten domains that served content on 70% of 1,000 popular webpages examined [KW+09].

Much of this tracking across popular sites supports *online behavioral advertising*, the practice of targeting advertisements using a profile of a user generated from the websites he or she visits. The canonical method of profiling involves having a list of interest categories, such as "home and garden." These interest categories are either selected or unselected for a particular user. However, some companies may be collecting more detailed information. Online behavioral advertising has attracted substantial attention in the media and from regulators. In the United States, both the Federal Trade Commission [FTC+12] and the White House [WH+12] released privacy reports in 2012 discussing concerns about online behavioral advertising.

In recent years, social media sites have also begun to serve content on many different websites, allowing these social media companies to track a user's activities across a broad swath of the Internet. The content served by social media companies often takes the form of social widgets. For example, the social network Facebook places a 'Like' button on many pages, allowing users to click a button and share that article on his or her Facebook profile. Whether or not a user clicks on the button, and whether or not the user is logged into Facebook at the time [V+11], Facebook notices the user's visit to that page. On the one hand, these widgets provide a frictionless way for a user to share articles and information with his or her social circle. On the other hand, this functionality raises privacy concerns since it allows the social networking companies to track a user's movement across many websites.

Websites can also track the links a user clicks on a webpage. Using a technique known as *URL Rewriting*, a website can be crafted to determine whether or not a user has clicked on an individual link. Understanding how a user has navigated a page can be useful for analytics, such as helping a search engine determine which of the links it presented were actually clicked. As an example of this practice, if a user goes to Google and searches for "privacy organizations," the results would likely include a link to the IAPP. However, rather than presenting a direct link to the IAPP's website at *https://www.privacyassociation.org/*, Google might instead present a link of the form:

http://www.google.com/url?query=privacy_organization&user=2fE65Da&url=privacyassociation.org

Such a link would automatically redirect the user to the IAPP website. However, by first directing the user's browser to Google's own server, Google would be able to learn which link a particular user chose from a particular set of results. The unique identifier for the user can be the same unique identifier contained in a cookie that Google stored on that user's computer, allowing that action to be associated with that particular user. Furthermore, by analyzing the time and IP address associated with this redirection request, Google could reconstruct the order in which a user clicked on links in search results, as well as the geographic location and Internet Service Provider from which the user was connecting to the Internet.

JavaScript, a programming language used to create dynamic and interactive websites, can be used to track how a user navigates a webpage in even greater detail than simple URL rewriting. To enable webpages that dynamically change as a user interacts with the page, there are JavaScript functions for determining where a user's mouse cursor is placed on the page, when a user has placed the mouse cursor over a particular element of the page, what the user has typed, and in what sequence. These functions can be used to capture navigation information in great detail and then to send it to a remote web server for analysis.

5.3.2. Web Tracking Beyond HTTP Cookies

While third-party cookies are a primary means of tracking a user's web browsing, many additional technological mechanisms enable tracking. Some of these technologies have been used to respawn cookies that the user deleted, while others present entirely different methods for tracking and surveillance. Many of these techniques are related to caching, which is the idea of saving web data locally. When a user views a website, browsers generally cache the page, or save a copy of the page on the user's hard drive so that identical content does not need to be re-downloaded on subsequent visits. Similarly, it can be beneficial to cache content or settings for webpages that use plugins like Adobe Flash. Of course, these techniques can also store tracking information.

While a user's IP address might initially seem to be a promising mechanism for tracking, the potentially large number of users who share an IP address and the frequency with which users acquire new

IP addresses as they move between locations makes the use of IP addresses for tracking less attractive. Rather, users are tracked using some of their computers' local storage mechanisms that are accessible to web browsers, as well as through techniques that subtly misuse features of web browsers to glean information about a user.

Web browsers and their plugins can write data to the hard drive of the user's computer in a number of ways beyond simple HTTP cookies. Many of these mechanisms have been used for tracking purposes. For example, the Adobe Flash plugin that is used to display videos and other interactive content on a number of sites has its own means of storing information, commonly called either Local Shared Objects (LSOs) or "Flash Cookies." LSOs can store approximately twenty-five times as much information as a standard HTTP cookie. Unlike HTTP cookies, LSOs are shared between all web browsers on a user's computer, and they are also stored in a location on the hard drive separate from HTTP cookies. While LSOs can be used for purposes like remembering the volume setting for watching videos on a particular website, they can also be used for storing unique identifiers that may not be deleted when a user deletes his or her cookies.

LSOs have been used widely. A 2009 study examined the 100 most popular websites and found that more than 50% of these sites were using LSOs [SCMTH+09]. While some of these sites seemed to be using LSOs for caching data, many others, including U.S. government websites, were storing unique identifiers about the user in LSOs. Until recently, LSOs were not deleted when a user cleared the cookies in his or her web browser. Therefore, some of these sites used LSOs to respawn HTTP cookies that a user had deleted. These respawned cookies often contained the same unique identifier as before deletion, seemingly contradicting the user's wish to clear this information. Following controversy over the use of LSOs for tracking, the Flash plugin and some browsers, including Firefox and Chrome, were updated so that choosing to clear the history in those web browsers would also delete LSOs.

A number of similar technologies also store web tracking information on a user's computer. For instance, Silverlight Isolated Storage provides an analogous function to LSOs for Microsoft's Silverlight framework, which competes with Adobe Flash. In addition to these plugins, core features of the web programming languages JavaScript and HTML5 can enable tracking. The JavaScript programming language can store data using a property called window.name. This property was originally created for web designers to assign each browser window a name that they could reference in computer code. This JavaScript property allows up to two megabytes of data to be stored during a particular web browsing session [F+08]. Unlike HTTP cookies, the *window.name* property is initially empty each time a new browser window or browser tab is opened, limiting the amount of time a unique identifier will persist. However, this method can be used in conjunction with other methods to retain data over a long period of time.

The Internet Explorer browser itself is able to store information on the local hard drive with userData storage, which enables each domain to store up to one megabyte of data. Like many other alternatives to HTTP cookies, information in userData storage is not deleted when a user deletes his or her cookies. A more subtle tracking method is a *pixel hack*, in which a unique identifier is written into a miniscule image that is generated on the fly. The tracking information is stored in these images as the color values for one or more pixels. Since images are often cached locally, these tracking values can often be retrieved later.

An updated specification of the HTML markup language known as HTML5 enables video, audio, and graphics to be embedded into webpages more easily. It also specifies methods for storing information locally, and these methods can be used for tracking. These storage methods, which have been separated from HTML5 for specification purposes and are currently known as either DOM Storage or Web Storage, are supported by all major web browsers as of 2012. Modern web browsers enable users to view multiple webpages at the same time in different tabs or different windows, and the two main types of DOM Storage work differently across different windows. The Session Storage method saves information for a particular window only; other windows cannot access this information even if the same website is being viewed. Furthermore, this data is removed when the window is closed. In contrast, the Local Storage method stores data semi-permanently. This stored information is available across windows that contain pages from the same domain.

Additional methods for DOM storage have been proposed and debated, yet are in flux at the time of press. For example, Mozilla Firefox previously supported a storage mechanism called Global Storage, but support for this feature was removed in version 13 of the browser [M+12]. Similarly, Database Storage using SQLite was considered as a possible W3C standard and was implemented in some browsers, but official efforts towards its standardization ended in 2010 [W+10].

Yet another mechanism for a browser to store information locally leverages the way web browsers cache data. ETags are HTTP headers that allow a browser to tag a previously-viewed webpage or object with

a permanent identifier. They were originally designed to enhance performance when visiting websites that have been cached. A site can tag content with an HTTP ETag identifier that changes each time the content is updated on the server. As a result, a browser can request a resource that should only be returned if its ETag has changed. If the resource has not changed, the site can use the local copy. To track a user across multiple visits, a server can insert a unique identifier as an ETag. Although ETags are not deleted when a user clears his or her cookies, they may be deleted when a user clears the browser's cache of previously viewed pages. As such, ETags enable tracking even if cookies are deleted or disabled.

In recent years, there have been rapid changes in the tracking techniques being used. For instance, a 2011 study found a substantial decline in the proportion of popular websites using LSOs [MC+11] compared to the 2009 study that initially brought attention to the issue [SCMTH+09]. However, a separate 2011 study noted that HTML5 (DOM Storage) and ETags were being used for the same tracking purposes [AWSGH+11]. In 2010, a security researcher prototyped a tracking mechanism, the "Evercookie," designed to be extremely difficult to delete [K+10]. The Evercookie combined many of the above techniques, storing unique identifiers in over ten different locations on a user's computer. If data from one storage location were deleted, this data would be respawned from other locations.

Features of web browsers designed to enhance users' experiences on the web can also be misused for tracking purposes. By default, web browsers show links on a page that have already been visited in one color, while links that have not yet been visited are displayed in a different color. Although it cannot directly access a user's browsing history, JavaScript can access the color of any element on a webpage, including links. Therefore, in a technique known as *browser history stealing* or *sniffing*, an unscrupulous page can include thousands of invisible links to popular sites and then use JavaScript to query the color of those links to learn whether a particular page has been visited by the client browser [G+06].

Another technique that misuses features of JavaScript and HTML for tracking purposes is *browser fingerprinting*. So that websites can adjust their pages to match the configuration of a particular user's computer, there are JavaScript functions that reveal the timezone, screen resolution, as well as fonts and plugins that have been installed on a particular computer. A 2010 study found that, even among a sample of potentially privacy-conscious users, 94.2% of browsers with Flash or Java installed could be uniquely fingerprinted using their configuration information [E+10].

5.3.3. Tracking Email Recipients

Some of the mechanisms that can be used to track the websites visited from a particular computer can also be used to determine whether or not an email has been opened. Knowing when an email has been opened or when a particular link in an email has been clicked can be useful for advertising companies to evaluate the effectiveness of a campaign, but it can also have more pernicious uses. Two common techniques for tracking email recipients are variants of the beacon and URL rewriting techniques used for web tracking.

Popular email programs, such as Microsoft Outlook and Gmail, can display emails containing HTML code. Similar to its use on websites, HTML code enables emails to contain different colors, advanced formatting, and images. Images can be attached to the email, or they can be downloaded automatically from a remote server. When an email is opened, its HTML code can request an image uniquely tied to a particular user from a remote server, enabling the remote server to know when the recipient has opened the email.

For instance, when Bob opens an advertising email from ABC Advertising, the email's HTML code can instruct the email program to download a 1-pixel image with a filename unique to Bob from *ABCAdvertising.com*. Since the advertising company controls the server, it can tell whether it has ever received a request for this image, which is unique to Bob. If it has, then it knows that Bob opened the email, along with exactly when this occurred. If it has not received a request for the image, Bob may not have opened the email, or he may have opened the email using a mail client configured not to download images.

As on webpages, links to websites in an email can also be customized to track whether or not a user has clicked on them. An email might contain a link that will eventually bring the email recipient to a specific website, such as *www.big-sale.com*, if he or she clicks on the link. However, rather than containing a direct link to that page, the email might contain a link to *www.big-sale.com/174cx3a*, where *174cx3a* is an identifier sent only to Bob. Therefore, if *big-sale.com* receives a request for the page *174cx3a*, it knows this request originated from the email that it sent to Bob.

5.3.4. Best Practices

In crafting a system that tracks the activities of users on the Internet for any purpose, it is essential to adhere to the core privacy principles of "notice" and "consent." Key goals include providing clear notice about any tracking that occurs and asking users to consent to the collection of data. Furthermore, this privacy notice should explain what information is collected, how it is used and stored, and with whom it may be shared.

While best practice dictates following these guidelines for all users, regulatory and legal requirements differ based on jurisdiction. For instance, an amendment to the European Data Protection Directive that was written in 2009 and that went into effect in May 2012 specifies how companies should handle storing or reading information from a user's computer. Since HTTP Cookies are the most common type of local storage, this requirement has been informally called the "Cookie Directive." In particular, storing or reading information is "only allowed on condition that the subscriber or user concerned has given his or her consent, having been provided with clear and comprehensive information" about the storage or access of information [D+09]. As a result of these changes, some websites have begun to provide opt-in boxes when visitors from the EU first visit their websites, only setting certain types of cookies if a user has clicked a button or performed a similar opt-in action.

Violating users' privacy expectations or trust can lead to serious legal and public relations repercussions. For example, a series of class-action lawsuits were brought over the use of LSOs for surreptitious tracking [V+10], while privacy concerns about some of the tracking technologies discussed in this chapter led to front page coverage in major media outlets [e.g., [V+10 (2)]]. On the other hand, communicating clearly with users can directly benefit companies and organizations. For instance, a 2012 study revealed that a confluence of privacy concerns and misconceptions that advertisers collect more personal information than is typically collected leads users to oppose behaviorally-targeted advertisements [ULCSW+12]. These findings suggest that openly communicating about practices the average user would deem reasonable might lead to increased user acceptance of tracking.

5.4. Blocking and Controlling Web Tracking

A user who wishes either to block or to limit web tracking has a number of options at his or her disposal. Unfortunately, there is not one silver bullet that protects a user against most tracking. Software that automatically protects users' privacy, such as an email client that blocks images from being downloaded by default, is simplest from the perspective of a user. More commonly, a user is able to control web tracking to some degree by adjusting default settings in software they're likely to already have, configuring settings using a web-based interface, or by downloading additional software designed primarily for privacy. Of course, for all but the automatic tools, the user must know to take certain steps, often having learned to do so from the news media, from a software vendor, or from a friend.

Tools for controlling web tracking are either provided by groups who would not benefit directly from the tracking, or by companies and organizations that conduct tracking. In the former case, some of the tools are provided by groups that protect consumer privacy as a public service or as a business, while other tools are provided by companies with a broad mission, such as the makers of web browsers. In the latter case, tools are provided by the same group that would be tracking the user. An example of such a tool is the Digital Advertising Alliance (DAA) Consumer Choice page.[10] Although groups sometimes offer privacy mechanisms to distinguish their organization in a competitive market or to enhance the end-user experience, it is often the case that suggestion or writ from legal and regulatory authorities leads companies to craft these tools. For instance, pressure from the Federal Trade Commission in 2009 led the advertising industry to develop principles for self-regulation, and these principles were realized in part by the DAA Consumer Choice page [JM+11].

In this section, we discuss ways that a user can block or limit web tracking by third parties – domains other than the primary one being visited. Beyond not visiting a particular company's website, tracking by first parties is much more difficult to prevent. Tools that block all requests to trackers, including first parties, would break the functionality of most websites. Furthermore, in the United States, first parties are generally exempt from legal and regulatory frameworks concerning tracking. As a result, one of the few resources left to a user is reading those sites' privacy policies to understand their practices. Unfortunately, past research has found reading long privacy policies leads to unrealistic opportunity costs for the user

[10] http://www.aboutads.info/choices/

[MC+08], that these privacy policies are written at an advanced reading level [GDJ+02], and that privacy policies sometimes are not translated into languages to which the rest of the site has been translated [USC+12]. After focusing primarily on how a user can block third-party web tracking designed to profile behavior across many websites, we discuss how a user can prevent his or her web searches from being tracked, as well as how a user can block email tracking.

5.4.1. Blocking web tracking

Web browsers' cookie settings are one way for users to disable tracking based on HTTP cookies. In some browsers, privacy-protective processes are automatic; for instance, by default Apple Safari blocks third-party cookies from being set. Users of other browsers who wish to block third-party cookies must change the browser's default setting, which many users never do [L+11]. Of course, these methods only prevent tracking methods that exclusively use cookies.

5.4.1.1. Privacy Settings in Browsers and Do Not Track

Browsers differ in the prominence and type of cookie settings that they provide. At a high level, most browsers enable users to allow all cookies, allow only first-party cookies, or block all cookies. Complicating matters, some browsers that allow users to control third-party cookie settings block third-party cookies from being sent back to websites, while others block third-party cookies from being set. If third-party cookies are blocked only at the time they are set, a cookie set by a first party may still be sent back to that same website even when it is in a third-party context. Browser cookie settings generally do not make it clear whether they apply only when cookies are set, or also when cookies are sent back.

Mozilla Firefox's cookie settings are featured prominently in its "Privacy" menu, with a box that can be unchecked to block the use of cookies. An additional checkbox specifies whether or not third-party cookies should be accepted. On the other hand, Google Chrome's cookie settings are buried in the "Under the Hood" menu's "Content Settings" sub-menu. In this sub-menu, Chrome's recommended setting is to allow cookies to be set. However, cookies can instead be restricted to the current browsing session, restricted to first-party sites, or blocked entirely. Chrome independently allows a user to choose to delete cookies and plugin data, such as LSOs, when the browser is closed.

In contrast to other browsers, Microsoft Internet Explorer's cookie settings provide users with a slider on which they can choose their privacy level. Settings for directly disabling third-party cookies, or disabling all cookies, are buried in an "Advanced" menu. A user's preferences based on this slider are evaluated using websites' P3P tokens. P3P, the Platform for Privacy Preferences Project, is a machine-readable language with which websites can express their privacy practices, such as the information that they collect and how this information is used. Microsoft Internet Explorer uses these P3P policies to decide whether to accept or block a cookie from a particular site. Although P3P is a World Wide Web Consortium (W3C) standard, Internet Explorer is the only major web browser that uses P3P, and P3P has not been widely adopted by websites. Furthermore, some sites misrepresent their privacy policies by using nonsensical or sample P3P tokens [LCMM+10], rendering P3P unreliable.

In contrast to privacy-protection mechanisms that attempt to block or circumvent tracking technologies, an idea that has gained momentum in recent years is a "Do Not Track" header that requests that a user's browsing not be tracked. Do Not Track, abbreviated as DNT, is an HTTP header that is sent alongside requests for web content. This header can indicate that the user has requested not to be tracked, that the user has chosen not to send a DNT request, or that no preference has been set. However, the exact details of both the meaning and mechanism of Do Not Track are currently in flux as the W3C works to define these concepts [W3C+11]. For instance, one debate about the Do Not Track proposal has centered on whether enabling DNT would mean that trackers could not collect information about a user, or whether they could still collect information but not use this information to target advertising. Representatives of the advertising industry have argued in favor of the latter definition, which has lead to an outcry from some privacy advocates [NS+12]. Furthermore, since Do Not Track does not actually block communication with a particular tracker, a user who enables this feature relies on companies to honor this request. At the time of press, only a small number of companies had committed to do so.

Despite the uncertainty about the meaning of Do Not Track, Firefox, Safari, and Internet Explorer have all implemented the feature, and Google has announced that Chrome will implement DNT by the end of 2012. However, browsers differ in the user experience of enabling DNT. For instance, Microsoft has announced that Do Not Track will be automatically enabled as the default setting in Internet Explorer 10

[B+12b], although there has been a protracted and currently unresolved dispute about whether or not trackers may ignore such default settings [EB+12]. In Firefox, DNT may be enabled by clicking the "Tell web sites I do not want to be tracked" checkbox, which is the first option available in the browser's privacy menu. Similarly, Safari's privacy menu contains an "Ask websites not to track me" option.

In addition to cookie controls and Do Not Track, web browsers often include other privacy features. For instance, Internet Explorer 9 includes Tracking Protection Lists (TPLs). With TPLs, third parties can curate lists of domains to which web requests should be blocked. Users can then download one or more of these lists. In contrast to DNT, TPLs do not require users to rely on websites to honor their requests since communication to those sites is actually blocked [C+11]. However, it is also worth noting that enabling the "Tracking Protection List" feature in Internet Explorer 9 also enables DNT.

Furthermore, all major browsers offer private browsing modes that disable the storage of browsing data on a user's computer while the mode is enabled. However, a 2010 study demonstrated that, at best, these modes offer minimal protection from tracking and surveillance. The privacy protections provided by these modes often differ from users' expectations, and the implementation of these modes is sometimes flawed [ABJB+10].

5.4.1.2. Web-based Privacy Tools

Beyond software tools already available on their computers, users can take advantage of web-based tools to control tracking. Among the most visible web-based tools is a system of opt-out cookies offered by some companies engaged in tracking. Opt-out cookies are HTTP cookies that signal that a consumer has opted out of receiving behavioral advertising. Although users who have opted out do not receive targeted ads from a particular company, some companies still track those users' online activities. A primary way for consumers to learn about opt-out programs and set opt-out cookies is through industry-standardized icons and taglines. However, a 2012 study found that these icons and taglines poorly communicated that advertisements were being targeted or that consumers could click on these icons as a step in choosing not to receive targeted ads [LCCGHUX+12]. Opt-out cookies are also problematic from a usability perspective since users who delete their cookies, as many privacy-conscious users might, also delete their opt-out cookies. Furthermore, setting opt-out cookies for each of the hundreds of tracking companies a user might encounter would take a long time. Centralized websites organized by industry groups offer a single place at which a user can opt out from many companies at once [NAI, DAA, EV]. However, research has identified major usability problems with these centralized websites [LUBCSW+12].

Some companies that track users also provide web-based dashboards through which users can view and sometimes edit the profiles of their interests that these companies have constructed. For instance, Google [GOOG], Microsoft [MS], and Yahoo! [YAH] all provide dashboards for advertising preferences, and Evidon's Open Data Partnership allows a number of trackers to show users the behavioral profiles they have created [EVb].

An amendment to the EU Data Protection Directive has also provided consumers with more prominent web-based tools. These 2012 updates to the law state that implicit consent is no longer considered sufficient notice to consumers in certain cases when websites wish to set cookies [ICO+12]. As a result, in 2012, some sites began providing consumers with conspicuous notices about the use of cookies, often with options to disallow certain types of cookies on that site. These notices are provided the first time a user visits that website.

5.4.1.3. Third-Party Browser Add-ons

Browser add-ons designed for privacy purposes are an additional mechanism for preventing web tracking. A number of companies offer tools specifically designed to stop web tracking conducted by advertising networks, social networks, and other companies interested in tracking which websites a user visits. For example, the company Evidon offers Ghostery, while the company Abine makes DNT+. These tools work by blocking the mechanisms used for tracking. Some tools, such as DNT+, maintain a blacklist of domains or resources tied to tracking and completely prevent the user's browser from communicating with those domains [DNTP]. In contrast, tools like Ghostery allow the request to go through, yet prevent the request from including cookies. Other subtle modifications to requests, such as removing the HTTP referrer field, can also protect the user's privacy in limited ways.

Some general-purpose browser add-ons can also limit web tracking to an extent. For instance, the popular Firefox and Chrome add-on Adblock Plus, designed to block nearly all advertising on the web, blocks requests to the domains of a number of advertisers and consequently limits data collection by those particular advertisers. Similarly, NoScript, a Firefox add-on that prevents websites from executing JavaScript code and plugins like Flash, can prevent tracking that occurs using these techniques. Notably, HTTP cookies are sometimes created using JavaScript, and blocking the Flash plugin can prevent LSOs from being set.

A major dilemma with all of these tools is the burden they impose on users, who must take a number of steps to protect their privacy. Generally, these tools first require a user to install the tool, either from the developer's website or from a centralized repository. Following installation, tools often ask a user to choose configuration options. Over time, the tool may automatically update itself, or user intervention may be required to install the newest version.

Unfortunately, neither specific privacy tools nor general-purpose add-ons are necessarily easy for an average consumer to use. For example, a 2012 study of web-tracking privacy tools revealed serious usability flaws in all nine of the popular tools it tested [PUBCSW+12]. Pervasive usability flaws identified by the authors included non-protective default configurations incommensurate with the goals of a privacy tool, confusing interfaces, and the use of jargon. Study participants also found it difficult to decide which advertising companies to block, which blacklists or whitelists to choose, and which tracking technologies to disallow.

5.4.1.4. Deciding What to Block

The majority of tools ask users to configure settings and decide what to block. However, a user's decision about what to block can be fraught with complexity. Based on interviews with 48 non-technical users, a 2012 study concluded that users have great difficulty reasoning about tracking technologies and evaluating tracking companies. Study participants also had major misconceptions about how online behavioral advertising worked [ULCSW+12]. However, users are not alone in facing this difficulty. Software designed for blocking also must decide what to block. If a tool blocks too little, it will be ineffective. However, if it blocks too much, it may break web functionality, annoying users and potentially leading them to abandon the tool. Furthermore, widespread adoption of a tool that thoroughly blocks existing methods might lead trackers to adopt increasingly surreptitious and subtle methods of tracking.

Some tools use partially automated metrics for determining what to block. For instance, Microsoft Internet Explorer determines what to block using P3P tokens specified by websites. Some versions of Internet Explorer determine when to block cookies from a domain using frequency counts, which are running totals of the number of different first-party domains on which a particular tracker has been observed by that user's browser [D+10]. Other tools are designed simply to help users visualize this sort of frequency information. For instance, Mozilla's Collusion tool [COL] presents a visualization of which third parties track users across particular websites.

Rather than asking the user to make a decision about each potential tracker, organizations or individual experts can compile a list of domains or patterns to block. The approach of using pre-compiled lists has been used by tools ranging from Internet Explorer Tracking Protection Lists to the Adblock Plus browser plugin. While lists can ease the decision-making process for users, they do have drawbacks. First of all, users need to be aware that they must select a list, and they also must be able to evaluate potential choices. Both of these tasks are problematic for many users [PUBCSW+12]. Furthermore, by their nature, lists will not conform exactly to an individual user's privacy preferences, which tend to be complex. It is also possible for lists created by one organization to cause competitors' pages to be blocked or to serve business interests, further complicating the compilation of lists. Of course, the advantage of lists is that they provide users a simple one-size solution that users may later tailor to their needs.

Tools differ in the extent to which they block tracking, adding additional complexity to users' efforts. A tool that blocks too little leaves the user vulnerable to tracking, whereas a tool that blocks too much can break the functionality of websites a user hopes to visit. Some tools, including the general-purpose script blocking tool NoScript and web-tracking-privacy tool Ghostery, break functionality on websites in ways that may not be obvious to users. To proceed with their browsing, users might try to unblock specific elements until a page loads, disable the tool entirely, or adopt ad hoc solutions, such as using a different browser without the privacy tool.

Researchers have coined the term "functional privacy" to capture users' willingness to aim for as much privacy as they can get without breaking the functionality of the web [WCL+12]. However, it remains to

be seen whether such a conception would incentivize trackers to break the functionality of pages intentionally. Tools that aim to provide privacy without breaking website functionality have also been proposed. For example, ShareMeNot blocks social widgets on websites until a user actually wishes to use the widget [RKS+12], which is a promising direction for the design of privacy tools.

In place of web-based tools or software, some users adopt overriding or ad hoc strategies. For instance, some users might use different browsers with different privacy settings. These individuals primarily use the more privacy-protective setup, yet switch to the less privacy-protective browser when a particular site's functionality breaks. In contrast, other users employ privacy-preserving proxies that scrub from their web requests potential identifying information, such as cookies, HTTP headers, and their IP address. These proxies also tunnel requests from many users through the same channel to provide a limited degree of anonymity. Like all other solutions, proxies are not a silver bullet and can certainly still leak information about a user. They can also break web functionality.

Implicit among the challenges of protecting user privacy is the complex nature of privacy decision making. Past research has demonstrated that privacy is often a personal process in which different users will have very different preferences [ULCSW+12]. Furthermore, privacy depends on context [N+04], which is a notion currently unsupported by major privacy tools. Overall, a user has many tough choices and few easy solutions when attempting to stop web tracking.

5.4.2. Blocking tracking of web searches

Users who wish to protect their privacy when using search engines also have a handful of mechanisms available. However, this task is more complex since most privacy tools for limiting web tracking focus on third-party tracking. Disabling certain tracking mechanisms on first-party sites, such as search engines, will often break those sites' functionality. As a result, the simplest and most secure way for a user to minimize having private information leaked through web searches is to use a search engine that promises not to track the user. Alternatively, users can download tools that help obscure their searches by inserting a large number of decoy requests, or they may use proxies or other general-purpose tools to make their searches more private to an extent.

In recent years, a handful of search engines have begun to use privacy-protective practices as a competitive advantage. Most popular search engines generally save a user's search history, including his or her search queries, when these queries occurred, the user's IP address, and unique identifiers from cookies. In contrast, privacy-protective search engines, such as DuckDuckGo, promise neither to collect nor share a user's personal information. By default, DuckDuckGo does not use HTTP cookies except to save preferences about the page layout a user has chosen, nor does it allow the HTTP referrer field to contain information about the search query. However, users must trust DuckDuckGo and similar sites to fulfill their privacy promises.

Users who wish to hide their search history can also download a tool to assist them, although few such tools exist. TrackMeNot, an add-on for Firefox and Chrome, protects a user's privacy by issuing decoy queries to major search engines [HN+09]. It operates by achieving security through obscurity, creating ambiguity about whether a particular query was issued by a user, or whether it was issued automatically by the program. The plugin's behavior is meant to mimic that of a real user. For example, it sometimes performs a large number of queries in a short amount of time, and it selectively chooses whether or not to click through to a link.

Users can also use general or ad hoc techniques to prevent their searches from being tracked. For instance, a proxy or an anonymizing network, such as Tor, can strip some or all of the identifying information from web traffic, making a user's searches more difficult to track. However, it is still possible for private information to leak even using techniques described in this chapter. For instance, users who enter their own name or other personally identifiable information in searches can still be vulnerable to having their searches tracked even if they would otherwise be anonymous [S+07].

5.4.3. Blocking email tracking

In contrast to web tracking, the issue of email tracking has become a less pervasive problem in recent years without direct user intervention. A number of modern email clients block beacons, images, and other content loaded from external sites since this external content could be used for tracking. This blocking disables one of the most ubiquitous techniques for determining whether or not an email has been read.

However, since tracking can still be accomplished through URL rewriting, it is important that a privacy-conscious user not follow links contained in emails. Furthermore, due to the threat of phishing attacks, it is generally considered good practice not to follow links in emails if the user expects to enter sensitive information at the destination. Of course, even if a link in an email does not seem to contain any type of unique identifier, users who follow the link or otherwise access that site are still subject to web-tracking techniques.

Finally, a common ad hoc technique to prevent email tracking is for a user to maintain multiple personas online. As it is often free to create an email account, many users have multiple email accounts. If a user has created an account exclusively for receiving email solicitations, he or she may not mind if a company tracks its emails since that account may not be tied to any real-life identity. Of course, subtle information leaks are possible. For instance, the date and time that the email has been read, as well as a user's IP address, browser configuration, and HTTP cookies, can tell a company substantial amounts of information about that user. In addition, a person who uses a separate email account for solicitations but then makes a purchase tied to that address is at risk for having his or her identity tied to that account. This privacy leak can even spread to other companies with whom data is shared or to whom this data is sold.

5.4.4. Best practices

Currently, a user who hopes to prevent his or her activities on the web from being tracked is put in a difficult position. To protect against the myriad different tracking threats, a privacy-conscious user will often need to use a patchwork of tools. Even with substantial effort on the part of a user, it is unlikely that he or she can feel fully in control of his or her privacy. For instance, there are always potential tracking mechanisms that a researcher or newspaper may soon reveal have been put into use, and the data that has been collected about a user's activities is rarely made available to the user. While some tracking companies provide dashboards that allow users to see what information the company is storing in their profile and make changes to that profile, it is difficult for users to get a comprehensive view of what data has been collected by all trackers. This presents an opportunity for developers to design a means to present the results of tracking data to a user. Such a system would provide additional clarity about the tracking process, as well as an opportunity for the user to correct mistakes. Furthermore, while some tools provide a technological solution by blocking tracking mechanisms, other tools require a user to trust that a company or website respect a preference or adhere to this advertised practices.

While there are no perfect solutions for protecting against tracking, there are many best practices that should be followed by groups providing users with privacy tools. The inclusion of a graphical user interface and colorful icons does not automatically make a privacy tool easy to use. Rather, to craft a tool that successfully supports consumers in protecting their privacy, careful consideration and substantive feedback from user studies is essential, albeit often overlooked.

It is important for users to be made aware of tracking in the first place. User education can take many forms, ranging from prominent privacy disclosures to media stories warning of privacy risks. Regardless of the type of communication, it is essential that communication be both clear and prominent. This issue has taken on particular importance following debate over the European Union's disclosure requirements for cookies, in which regulators deemed insufficient the notion that users implicitly consent to tracking after being notified only by a lengthy privacy policy. In contrast, it is considered best practice to provide users with conspicuous, clear, and understandable notice about tracking, and to give them the opportunity to choose not to be tracked.

Of course, the mechanisms for allowing users to choose not to be tracked have their own best practices. First of all, users should not be expected to make decisions about particular technologies. When a user decides he or she does not wish to be tracked, the developer should understand that this preference covers all tracking technologies, including HTTP cookies and LSOs. It is essential that privacy tools match this expectation to the greatest extent possible. Furthermore, privacy-protective default behaviors should be chosen when they will not interfere with functionality desired by the user. For instance, many email clients automatically block tracking beacons without requiring the user to do anything. Any actions for which user intervention is required should be able to be completed quickly, and any interfaces presented to the user should be tested extensively in usability studies in order to ensure that they can be used with few or no errors.

5.5. Location Tracking

In recent years, the types of tracking that occur have broadened to include a person's location. As people carry mobile phones with location-tracking capabilities, it has become possible to collect a person's whereabouts at nearly all times. Social networking applications, employee-tracking systems, and location-enabled media, such as photos, are among the technologies that use this location data to enhance their systems.

Location tracking also extends beyond the mobile phone. Radio-Frequency Identification (RFID) chips that are embedded in smartcards and in consumer products can also be used for location tracking. Furthermore, Global Position System (GPS) technologies are often included in consumer hardware, including cars and cameras. These GPS-enabled devices are able to know their own location at all times. In such an environment, groups deploying such technologies should be aware of the capabilities as well as the potential privacy implications of their uses. Therefore, this section outlines location tracking technologies and services, techniques for blocking location tracking, and best practices for those employing location-tracking technologies.

5.5.1. Location-Tracking Technologies

Devices can contain a wide variety of types of location-tracking technologies, each of which relies on slightly different underlying systems. We describe several of the most common location-tracking technologies: Wi-Fi and cell tower triangulation, GPS, RFID chips, phone tracking, and the use of location data stored in content metadata.

5.5.1.1. Wi-Fi and cell tower triangulation

Cellular and Wi-Fi signals can be used to track the location of cellular phones. Cellular phones communicate wirelessly with towers owned by phone companies. These cell towers receive a phone's signal and connect the phone to a global network. However, a phone's communication with a cellular tower gives the phone many clues about its location. The time it takes messages from a particular tower to arrive, the strength of the signal from that tower, and, most simply, which towers a phone can communicate with all reveal information about the phone's location relative to particular cell towers. Therefore, after determining a phone's position relative to a handful of towers with known locations, the position of the phone can be determined geometrically through triangulation.

In addition to signals from cellular towers, the identity of Wi-Fi networks with which a phone can communicate may pinpoint its location. Wi-Fi signals have a shorter range, allowing for more fine-grained location information, whereas cell towers provide a more permanent location marker, but they provide less granular location data. Wi-Fi and cell tower triangulation requires the creation of a pre-existing database of Wi-Fi access points and cell tower locations that cover the region over which the location tracking will occur. Thus, this type of location tracking is primarily beneficial in urban areas where there is a high density of Wi-Fi access points and cell towers.

5.5.1.2. GPS

Global Positioning System (GPS) satellites can also be used to determine location, specifically a device's longitude, latitude, and altitude. Many consumer devices, including mobile phones, are equipped with GPS capabilities for location tracking. Cameras and similar devices can also include GPS capabilities for tagging the location of photographs taken, and automobile infotainment systems can use GPS to identify which region's content, such as weather and news, is relevant to the driver at a particular time.

Devices using GPS calculate location using signals received from at least four geosynchronous satellites positioned in space out of a set of dozens [FAA+10]. Based on the differences in time it takes messages from different satellites to arrive to a GPS-enabled device, the device can determine its position relative to the satellites. Since these satellites' positions are known and constant relative to the earth, the GPS receiver can determine its own position geometrically. Since devices only receive, but do not transmit, signals in the GPS process, devices do not automatically reveal their location by using GPS. However, after learning their own location, devices with GPS can reveal this location to other parties and services.

5.5.1.3. RFID

In contrast to GPS, Radio Frequency Identification (RFID) chips can be used to track devices locally. An RFID chip is a tiny microchip that can reach a size of 0.4mm square. Each microchip is identified by a

unique serial number and contains an antenna with which it transmits information to RFID readers. An RFID chip can be placed on products or cards, built into passports, or implanted in animals for tracking purposes. They are often used in supply chain management to allow companies to track inventory.

Most commonly, an RFID chip transmits a signal containing its unique serial number to RFID readers. Particular RFID chips differ in a number of characteristics that affect their applicability for tracking. For instance, passive RFID chips do not contain batteries and have a smaller range of transmission than active RFID chips, which do contain a power source. The frequency of transmission is also important since RFID chips transmitting at low frequencies have a range of about half a meter, while those that transmit at ultra-high frequencies can be read from dozens of meters away [J+05].

The unique serial number associated with each RFID tag allows for location tracking. Unlike GPS, an RFID chip does not know its own location. Rather, the RFID chip repeatedly broadcasts its own serial number. If an RFID reader is within its transmission range, that RFID reader knows the tag is nearby, and that RFID reader's location is known. If additional information is stored on a tag, the reader also receives that information and associates it with the tag's location.

5.5.1.4. Phone tracking

The location of a mobile phone and the individual to whom the cell phone belongs can be tracked using receivers installed within a building. The United States FCC also requires that phone companies be able to track phones when a 911 emergency call is placed [J+09]

An application of phone tracking technology came to light during the 2011 holiday shopping season. Two U.S. shopping malls tested a system to track shoppers within the mall based on the location of their cell phones as part of a "mobile phone survey." Signs notified the shoppers that the tracking was taking place, and shoppers had the option to opt out by turning off their cell phones. The malls ended the survey early after concerns about shopper privacy and the legality of the survey were raised [G+11, G+11 (2)].

5.5.1.5. Metadata

Location information can also be automatically stored in the metadata of content, such as photos. Metadata is information about content, such as the date and time it was created. This information can be automatically or manually added to content and accessed by applications.

Devices with GPS, such as cell phones and GPS-capable cameras, often automatically store the location a picture was taken in its metadata. When the photos are loaded into photo-browsing or editing applications, this location information is accessible.

5.5.2. Location-Based Services

Location-based services use location data to augment a variety of systems, including social media and applications that revolve around location data. Emerging uses of location data include tracking individuals, such as employees or children, and presenting advertisements specific to a particular location.

5.5.2.1. Social Media

A variety of social media applications use location tracking as part of their services. Some applications focus on enabling users to notify others of their location and to track others' locations. For example, Foursquare is a mobile application that lets users "check in" at locations and view their friends' check-ins. Users are further motivated to check in by the ability to earn badges and receive coupons. Similarly, Find My Friends[11] is an Apple application for iPhones, IPads and iPods that shows the location of a user's friends who use the service.

Other applications augment their services with location-based features. For instance, Facebook Places[12] enables users to check in at locations, as well as to tag shared items with location data. Yelp, a review site for restaurants and other services, also includes a feature with which users can check into a restaurant or other location.

[11] http://itunes.apple.com/us/app/find-my-friends/id466122094?mt=8
[12] http://www.facebook.com/about/location/

5.5.2.2. Location-Based Applications

A variety of other applications rely on location-based services for functionality. For instance, some applications provide maps and directions, give local weather forecasts, or provide information on nearby services, items, and individuals based on a user's location.

Applications that provide maps or directions typically rely on location-based services to pinpoint a user's location. The application then provides a map of the user's current location based on a database. Location data is often used in automobiles to help users navigate street maps. The user's location is provided using GPS positioning, and maps are provided from a database. Mapping algorithms are then used to compute routes between destinations.

Many smartphones also provide map and direction functionality using a combination of GPS, Wi-Fi, and cell tower triangulation to calculate a user's location. Google Maps for mobile devices is one example of such a mapping application. Many mapping applications allow users to search for the nearest restaurants, gas stations, or other services.

Other applications use location information to provide location-specific content. For example, weather applications track users to provide location-specific weather alerts or updates. For instance, iMapWeather Radio[13] uses location tracking to localize weather forecasts and provide critical weather alerts for a user.

5.5.2.3. Tracking Kids, Employees

Location-based services can also allow users to track other people. Parents can use location-based services to track their children, and employers can use them to track employees. Using the GPS in their children's cell phones, parents can track their children's location. Online services allow them to see where their children are throughout the day using either a specialized or standard cell phone [P+06]. Just as employers can use online surveillance to monitor employee computer usage, they can also use location tracking to monitor the location of employees. This can be done to reduce company liability, address potential threats to security, and track operational efficiency. Employee tracking can be performed using RFID chips or GPS trackers carried with the employees, among other techniques [K+05].

5.5.2.4. Location-Based Ads

Location information also provides a rich data source to advertisers, allowing them to create advertising that takes an individual's location into account. Location information can be sourced from mobile devices with location identification, or from a user's self-identified location on social networks. Advertisers can then offer advertisements or marketing offers tailored to a user's specific location.

Advertisers can take advantage of user location in a variety of ways. The first approach is to identify the consumer's physical location exactly, often using the consumer's smart phone. As opposed to tracking location at a fine granularity, an advertiser can create a "geofence," which is the process of defining a broad area in the vicinity of a particular location. If configured in a privacy-protective manner, the geofence would inform the advertiser whether or not the consumer is inside the geofence, but not the consumer's exact location. Geofences can also been used to trigger alerts when an individual crosses the perimeter of the geofence, presenting an advertisement for a nearby store. An additional approach is for advertisers to use consumers' check-ins on services like Facebook or Foursquare to determine the consumer's location and offer targeted advertising.

Mobile devices equipped with Near-Field Communication (NFC) can also support location-based advertising. Devices with NFC that are in close proximity can transmit information via radio waves, enabling consumers to access content when at a specific location [L+12].

Location-based advertising presents privacy concerns. In 2011, a study of theoretical sharing preferences by Kelley et al. found that users had significant privacy concerns about sharing their location with advertisers. However, they found that these concerns were somewhat mitigated when users had the ability to restrict sharing based on their location or the time of day [KBCS+11].

[13] http://itunes.apple.com/us/app/imapweather-radio/id413511993?mt=8

5.5.2.5. Combining with Data from Other Sources

Location information drawn from mobile devices can be combined with data from other sources to make inferences that were not previously possible. For instance, the "Please Rob Me" website gained media attention in 2010 for its possible privacy implications. This site aggregated information from users' Foursquare and Twitter accounts to create a list of people who were likely not at home since they had checked in elsewhere. Controversy over the privacy implications of combining location data with data from other sources erupted again in early 2012. This time, the "Girls Around Me" phone application combined gender information from Foursquare accounts with location information from check-ins to enable users to search for women in their vicinity. Foursquare found that the app violated their policies and shut off the developer's access to Foursquare data [B+12a]. However, this scenario provides an example of how inferences drawn from location-based data can lead to privacy concerns.

5.5.3. Geographic Information Systems

A geographic information system (GIS), such as a computer database or imaging tool, is a technology used to view and manipulate stored geographic information. Such geographic content could relate to any quantities associated with a particular location, including maps, population or other census statistics, or data about a specific resource at a location.

Uses for GIS are wide-ranging. They can include logistics systems used for businesses that need to track passengers. For instance, airlines need to track passengers, and utility companies need to direct crews. They also have applications in making agricultural decisions about planting crops [G+97].

5.5.4. Preventing and Controlling Location Tracking

The range and use of location tracking technologies can present privacy concerns. It is possible to block some types of location tracking on mobile devices. However, even when a mobile phone is turned off, it is often still possible to use the triangulation techniques discussed in this chapter to track its location as long as the phone's battery is connected. Furthermore, it is not always preferable to block location tracking since location data can augment a service. Thus, systems are being developed to allow more granular control over location sharing for location-based applications. Additionally, current research examines ways to preserve privacy in location sharing technologies.

5.5.4.1. Blocking Location-Based Tracking

Depending on the type of location-tracking technology, users can block or limit tracking to varied degrees. For location-based services that automatically include location data, users can often opt out of location tracking. The mechanism for opting out of location tracking varies based on the technology used. For location-based services that rely on check-ins, users are opted out by default and are required to check in to be tracked. In contrast, for actions like adding location data to tweets on Twitter, the user can choose to turn location tracking on or off. Other services, like some smart phone applications, require that users decide whether or not to enable location tracking when they download the application. In some cases, a user can only opt out by choosing not to use the service.

Users can also remove location data from some content after the fact. For example, location data is automatically added as metadata to photos on GPS-enabled cameras and mobile phones. It is possible to use photo-editing applications, such as iPhoto or Picasa, to delete this metadata after the photo has been taken. Other location-based services, such as Google Latitude, also allow users to view the information that was tracked and remove location data after the tracking has occurred.

To prevent tracking using RFID chips, the chips' communication can be physically blocked or the RFID chip can be physically removed. Because RFID chips use wireless communication, a protective sleeve placed over an item containing RFID can prevent the chip from being read. This approach is useful for items like passports whose chips contain information a user might not want to be accessible. In particular, shielding devices that contain RFID can prevent attacks similar to a 2009 demonstration in which a researcher carried his own RFID reader through San Francisco to clone passports [R+09]. RFID chips can also be physically removed from items like clothing to prevent tracking. However, removing the chip also disables beneficial uses of the technology.

5.5.4.2. User Controls

Although it is possible to block tracking to various degrees, it is sometimes more desirable to control who has access to different types of location data at different times. Interfaces on location-based applications can enable users to set privacy settings with various degrees of granularity. These settings specify who may access location information. Technologies like geofences can also be configured to allow users to control the boundaries of location tracking.

Applications that use location information also enable users to configure privacy settings to control who has access to their location information, and at what times. At a basic level, Foursquare[14] allows users to prevent anyone they have not accepted as a "friend" from viewing their location data. This technique of only allowing previously identified users from viewing information is known as setting a "whitelist."

Researchers have found that users' location sharing preferences vary based on the type of information shared, the group with whom they share, the time of day, the day of week, and the location [BKSC+11]. Loccacino,[15] a prototype based on these findings, allows for constant location tracking. However, it also enables users to control their location-sharing preferences based on who can view them, where they are located, and the time.

5.5.4.3. Research into Privacy Protections for Location

Beyond identifying methods for controlling location sharing, researchers have investigated how to provide the benefits of location sharing while preserving a level of user privacy. However, privacy preservation is difficult to achieve in location-based services. For instance, patterns in an individual's location, such as apparent home and work locations, can uniquely identify many people.

In 2011, a group of researchers prototyped a privacy-protective system for detecting whether two users of a location-based service were near each other. They used "location tags," which are signatures associated with a location, to allow users to detect whether they were near another user without giving away their own exact location or detecting other users' exact locations [NTLHB+11].

5.5.5. Best Practices

The wide range of location-tracking technologies presents many promising applications. However, when creating and using systems that track a user's location, it is necessary to consider the privacy implications of location tracking.

As was apparent when the shopping malls attempted to implement a customer tracking system, privacy is an important issue for location-tracking technologies. Location tracking should be included only if it provides a direct benefit. Wherever possible, applications should ask users to opt in to location tracking. Once location data has been collected, users should be able to see what has been stored about them and delete or update this data.

Collected location data should be treated as privacy-sensitive. Users should be informed through a privacy policy or other means of how this information will be used. If this data will be used in an unexpected manner, it is best practice to ensure that users have been provided clear and conspicuous notice. Additionally, before making location data more publicly available, it is best practice to consider carefully how it might be reused or combined with other datasets. Combining location data with other sources can provide a rich dataset, as many GIS systems show. However, as the Foursquare "Girls Around Me" application showed, it can also violate privacy.

When using location-based applications to track others, such as in a workplace setting, it is best practice to limit tracking to instances where there is a clear need. Furthermore, it is essential to inform employees about this tracking. Tracking should only take place while the employee is working. If tracking occurs on a mobile phone that an employee also carries during non-work hours, tracking should be limited to the workday. Once tracking data is collected, it should only be used for the intended purpose, and access should be minimized.

[14] https://foursquare.com/privacy/
[15] http://locaccino.org/

5.6. Audio and Video Surveillance

While a users' online activities and location are prime targets for tracking and surveillance, their conversations and real-world activities can also be surveilled using technology. For instance, audio and video recording devices might be placed in public or private areas. More insidiously, the webcams and microphones on modern laptops and phones can also be co-opted for this purpose.

5.6.1 Hidden Cameras and Microphones

Hidden cameras and microphones can provide video and audio surveillance without a person's awareness. Such devices can be very small and disguised as another object, as is the case with "nanny cams" intended to surveil the home. They can record information or wirelessly transmit data to a person performing surveillance in real-time.

5.6.1.1 Smartphones as Hidden Cameras and Microphones

Since smartphones often include microphones and webcams, they have also been used for audio and video surveillance. These devices' internal microphones and cameras enable an individual to perform remote surveillance around the device.

An individual's smartphone can be transformed into a hidden camera. Smartphones contain microphones and cameras that can be remotely activated, as well as a connection to the Internet that allows for the remote activation and transmission of surveillance data. This use of a remotely-activated smartphone microphone is called a "roving bug."

Within the United States, it was found legal under the federal Wiretap Act for the FBI to use remotely-activated audio surveillance on a cellphone in the case of United States v. Jon Tomero [K+06]. In the case that led to US v. Tomero, the FBI used roving bugs on two mob bosses' cell phones to surveil conversations that occurred near the phones [MB+06].

Through phone malware and viruses, remote attackers can also install roving bugs. Smartphones have little antimalware protection, and smartphone malware is a growing attack vector [L+05]. Through smartphone malware, an attacker can gain control of the microphone and camera to surveil an individual [FW+09].

5.6.1.2 Monitoring Through Laptop and Desktop Computers

Laptop and desktop computers also enable audio and video surveillance. Like smartphones, computers typically have microphones and cameras, as well as a network connection. This combination makes them useful for eavesdropping.

Surveillance can occur after a user inadvertently installs malware on his or her computer, allowing an attacker to take control of the camera and/or microphone. Farley and Wang describe an example of how malware can be uploaded onto a computer, take control of a user's microphone for audio surveillance, and simultaneously hide its own existence [FW+09].

Surveillance using computer cameras and microphones can also occur when the computer is owned by one entity and used by another, such as when computers are distributed by an employer or school. As illustrated by a recent case, the use of hidden surveillance has legal implications for privacy that are still becoming apparent. In 2010, it was discovered that a Pennsylvania school district was using the webcam on district-owned laptops to remotely take pictures while the laptops were in student homes. This remote monitoring became apparent after pictures taken from a student's home were used to confront the student about inappropriate behavior at home. While the district claimed that the purpose of this feature was to track lost laptops, they had not informed the parents or students about their ability to use the laptops for monitoring, despite using the feature 42 times over a 14-month period [C+10]. The parents of the student whose surveillance brought the issue into the open sued the school district for violation of privacy. They settled for over $600,000 [M+10]. There were additional federal and state investigations of whether the district had violated federal wiretap, or other privacy, laws [C+10].

Simply being in the same room as another individual's computer can be sufficient for surveillance. For instance, in a 2010 incident, a student at Rutgers University tweaked the settings of the iChat program on his computer to automatically answer calls and darkened the computer's screen. When he called his computer from a friend's room and observed his roommate with a romantic partner, he publicized this event on Twitter. This incident was widely covered in the news media after the roommate, who had been surveilled, committed suicide a few days later [P+12].

5.6.2 CCTV

Closed Circuit Television (CCTV) cameras are a type of video monitoring system commonly used by governments for preventing crime or terrorism. However, such systems can also be used for preventing shoplifting, providing security for a private facility, or tracking employees in a workplace [L+01, W].

In CCTV systems, cameras transfer images to a remote destination, where footage is available to law enforcement or security personnel for either real-time or post-hoc monitoring. A CCTV system can encompass a variety of different types of cameras of different levels of sophistication, including both mobile and permanently-mounted cameras. The quality of the images captured varies [GS+05].

Some examples of systems that use CCTV include traffic enforcement, such as red light cameras, and efforts to prevent terrorism, such as license plate recognition programs. The analysis of CCTV systems can be augmented with additional technology. For example, biometrics and other facial recognition capabilities are used to try to recognize terrorists in a crowd or in an airport. In the United States, general use of CCTV in public spaces by the government is legal as long there is not a "reasonable expectation of privacy" [L+01].

Workplace surveillance can also employ CCTV to observe employees. These systems capture both criminal acts and daily activities that occur in a store or workplace. In this environment, use of video surveillance is legal as long as it is limited to public areas [W].

5.6.3 Protecting Against Audio and Video Surveillance

Protection against audio and video surveillance on computers and smartphones can occur on several levels. First, it is possible to prevent some surveillance techniques by using antivirus software to keep a computer free from malware that would allow attackers to take over a camera or microphone. Additionally, malware is often delivered through untrustworthy downloaded software, especially on smartphones. Avoiding such software can help lower the risk of an attacker conducting audio or video surveillance.

It is also possible to protect against audio and video surveillance by blocking the hardware and software necessary for it to occur. For example, a user can disconnect or physically block a computer's webcam when it is not in use. Using firewall software to track incoming and outgoing network connections and block any suspicious activity can also help prevent surveillance [FW+09].

5.6.4 Best practices

When performing audio or video surveillance, especially within a work environment, it is best practice to ensure that the amount of surveillance performed is minimized and that it is being performed in a legal manner. Video and audio surveillance can be very privacy-invasive and should not be performed unless a particular objective (e.g., workplace security) outweighs the privacy drawbacks. Wherever possible, individuals under surveillance should be informed about the system. Additionally, local privacy laws should be checked before putting surveillance in place. As an employer in the United States, a first step is making sure that the surveillance is not taking place in an environment in which employees have an expectation of privacy, such as inside a bathroom stall.

Once audio and video surveillance data has been gathered, it is best practice to take proper measures to maintain data security and limit exposure of the content. Whenever possible, use automated systems to analyze the data or employ a separation of duties in which an analyst examines the audio or video and only reports findings to others. Such a separation of duties minimizes the chance the raw audio and video will be repurposed for unauthorized purposes. Of course, it is important to retain the raw audio or video in a secure location in the event the finding is challenged. To prevent misuse, one should limit access to the data and audit attempts to view this surveillance data. A clear policy should be established identifying who has access to the data and under what circumstances, and for how long the data will be retained. Data should be purged when no longer needed for the intended purpose.

5.7. Surveillance in Emerging Technologies

As microcontrollers and computers are increasingly embedded in devices inside the human body and around the house, privacy issues have begun to emerge. On the frontiers of medicine, implantable medical devices with wireless communication functionality can assist in monitoring and diagnosis, yet they can also reveal private information about an individual's health status. Similarly, smart electrical meters promise to optimize energy consumption, yet the real-time measurements that enable power savings can also reveal what's happening inside a home.

5.7.1. Medical Device Surveillance

Implantable medical devices, such as pacemakers and drug delivery systems, can benefit from remote monitoring by medical professionals. However, enabling wireless communication on these devices can lead to privacy issues. The measurements logged by a medical device, any unique identifiers tying a device to a particular patient, and even the existence of the device can all be considered private information [HBFKM+08]. Although these types of information are private, they might provide essential information to doctors in an emergency situation, creating a tension between patient privacy and medical utility [CF+11]. Of course, patients' wishes should be given great consideration. A 2010 study asked participants about hypothetical medical systems attempting to balance privacy and utility. The study found that many participants did not want to broadcast their medical conditions publicly, yet some systems met many of their requirements [DBFGKM+10].

While no large-scale attacks on implantable medical devices have been reported to date, malware and other viruses have been observed on computer-controlled medical equipment used in hospitals [T+12]. These computer infections, often caused by unpatched or outdated software, leave patients vulnerable to privacy violations and medical equipment malfunctions.

5.7.2. Surveillance of Smart Meters

As with medical devices, the proliferation of smart energy meters holds both promise and peril. Smart meters can optimize energy use by providing real-time monitoring of both the home and the electrical grid, yet these devices may also negatively impact privacy in the home. In particular, detailed measurement of power usage can reveal information about household activities, such as pinpointing when a particular appliance is switched on or off [PRKRA+07]. Researchers have considered the potential of power analysis techniques to reveal private information at a deep level. For instance, a 2011 study of eight LCD and plasma televisions found that the program being watched on the television could generally be identified just by analyzing the powerline [EGKP+11]. A 2012 study similarly measured power consumption to determine with high accuracy what websites were being visited on a personal computer [CRSXLF+12]. While some of these techniques require finer-grained data than provided by many smart meters, even measurements taken every one second on a smart meter can be used to profile the activities in a house, such as routines for eating and sleeping [MSFCI+10].

5.7.3. Best Practices

As with technologies that have been widely deployed, emerging technologies require careful consideration of the tradeoff between utility and privacy. Thankfully, systems can be designed to respect privacy while achieving the desired benefit with both medical devices and smart meters. For instance, researchers have proposed storing the wireless access keys for medical devices on a patient's skin using ultraviolet pigment [S+10] in an attempt to preserve patient privacy while still enabling quick access to the medical device during emergencies. Similarly, energy monitoring that measures consumption at longer intervals or at the level of a neighborhood, rather than a single home, can prevent certain monitoring attacks while still providing many of the benefits of smart meters. Overall, solutions that carefully consider ways to achieve the desired goals while protecting consumer privacy can sometimes alleviate tension between privacy and utility.

5.8. Chapter Summary

In this chapter, we presented an overview of tracking and surveillance techniques and countermeasures. We began by explaining how the packets that encapsulate communication over the Internet can be surveilled at a large scale, such as when these packets leave a workplace network or travel over a wireless connection. We then demonstrated how surveillance of low-level communications over the Internet can be made more difficult by encrypting web requests by using the HTTPS protocol, sending traffic through an encrypted VPN tunnel, or by using an anonymizing network like Tor to separate the sender of a message and the content of the message.

We next discussed numerous technologies that can be used to collect information about the websites that a particular user has visited, often for the purpose of better targeting advertisements to that person.

While we explained in depth how HTTP cookies enable this sort of tracking, we also introduced alternative technologies that can either be used as alternatives to HTTP cookies or can be used in conjunction with HTTP cookies to respawn deleted cookies. These technologies included LSOs, DOM Storage, and more subtle techniques, such as browser history stealing and browser fingerprinting. We also explained a number of ways in which users can block certain types of web tracking. For instance, we introduced the cookie controls built into web browsers, illuminated the debate over a proposed Do Not Track mechanism, discussed opt-out cookies offered by some advertising companies, and delved into third-party browser add-ons. Alongside these methods, we explained their shortcomings, as well as best practices in the design of privacy tools. We also briefly described ways to prevent web searches and email messages from being tracked.

We then provided an overview of technologies for tracking location, describing how a cell phone's location can be triangulated based on the signal from cell towers. We also described how devices can use GPS to determine their location. We also explained the complex interaction between utility and privacy as users share information about their location on social networking sites, in addition to current research on limiting the privacy perils of these disclosures. We then touched briefly on methods for audio and video surveillance. In addition to more canonical examples of surveillance, such as placing hidden cameras and microphones in public or private locations, we demonstrated how both smartphones and personal computers can be co-opted to surveil users. Finally, we discussed emerging surveillance techniques for medical devices and smart energy meters. Throughout the text, we also noted best practices for any group considering deploying tracking and surveillance technologies.

5.9. About the Contributors

Dr. Lorrie Faith Cranor is an Associate Professor of Computer Science and of Engineering and Public Policy at Carnegie Mellon University where she is director of the CyLab Usable Privacy and Security Laboratory (CUPS) and co-director of the master's degree program in privacy engineering. She is also a co-founder of Wombat Security Technologies, Inc. and previously was a researcher at AT&T Labs-Research. She has authored over 100 research papers on online privacy, usable security, and other topics. She has played a key role in building the usable privacy and security research community, having co-edited the seminal book Security and Usability (O'Reilly 2005) and founded the Symposium On Usable Privacy and Security (SOUPS). She has served on a number of boards, including the Electronic Frontier Foundation Board of Directors, and on the editorial boards of several journals.

Ms. Manya Sleeper is a third-year Ph.D. student at Carnegie Mellon University in the School of Computer Science. She is advised by Lorrie Faith Cranor. Her research focuses on usable security and privacy, specifically improving access control and privacy decision-making. She is the recipient of an NSF Graduate Research Fellowship and an ARCS Fellowship. She received her undergraduate degree from Dartmouth College in government and computer science, and, prior to coming to Carnegie Mellon, worked as a consultant at McKinsey & Company.

Mr. Blase Ur is a second-year Ph.D. student in the Carnegie Mellon University School of Computer Science, advised by Lorrie Faith Cranor. He researches usable security and privacy, including passwords, online behavioral advertising, and privacy decision-making. He is the recipient of an NDSEG Fellowship, a Fulbright Scholarship, and the Yahoo! Key Scientific Challenges prize. Some of his recent privacy research has been selected for the Future of Privacy Forum's 2012 "Privacy Papers for Policy Makers" compendium and has received a "Best Paper Honorable Mention" at CHI 2012. Before coming to CMU, he received his undergraduate degree in computer science from Harvard University and taught at Rutgers University.

References

[GRS+96] David M. Godschlag, Michael G. Reed, Paul F. Syverson, "Hiding Routing Information," in *1st* International *Workshop on Information Hiding*, Cambridge, U.K., 1996, pp. 137--150.

[M+03] Michael McCardle, "How Spyware fits into defense in depth," SANS Institute: InfoSec Reading Room, 2003.

[W+08] Peter Whoriskey, "Every click you make," *The Washington Post*, Apr. 3, 2008, http://www.washingtonpost.com/wp-dyn/content/article/2008/04/03/AR2008040304052.html.

[S+11] Robert Sheldon, "The situation is under control: Cyberspace situation awareness and the implications of China's internet censorship," *Strategic Insights*, vol. 10, no. 11, 2011.

[FCC+08] Federal Communications Commission, "Commission orders Comcast to end discriminatory network management practices," Federal Communications Commission: News Media Information, 2008.

[AMA+08] American Management Association, "Over half of all employers combined fire workers for e-mail & Internet abuse," Press Release, 2008. http://press.amanet.org/press-releases/177/2007-electronic-monitoring-surveillance-survey/ Accessed June 22, 2012.

[NMN+06] G. Daryl Nord, Tipton F. McCubbins, Jeretta Horn Nord, "E-monitoring in the workplace: Privacy, legislation, and surveillance software," *Communications of the ACM,* vol. 49, no. 8, 2006.

[FCC] Federal Communications Commission, "Children's Internet Protection Act," http://www.fcc.gov/guides/childrens-internet-protection-act Accessed June 23, 2012.

[S+01] John Schwartz, "Schools get tool to track students' use of Internet," *The New York Times*, May 21, 2001. http://www.nytimes.com/2001/05/21/business/schools-get-tool-to-track-students-use-of-internet.html

[FAA+10] Federal Aviation Administration. "Navigation Services – Global Position System," 2010. http://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/gnss/gps/ Accessed July 22, 2012.

[J+05] Ari Juels, "RFID Security and Privacy: A Research Survey," *IEEE Journal on Selected Areas in Communications,* vol. 24, no. 2, 2005, pp. 381--395.

[G+11] Sean Gallagher, "Mall owners pull plug on cellular tracking (for now)," *Wired*, Nov. 2011. http://www.wired.com/business/2011/11/mall-pull-plug-cell-tracking/

[G+11 (2)] Sean Gallagher, "We're watching: malls track shoppers' cellphone signals to gather marketing data," *Wired*, Nov. 2011. http://www.wired.com/business/2011/11/malls-track-phone-signals/

[P+06] David Pogue, "Cellphones that track the kids," *The New York Times*, Dec. 21, 2006. http://www.nytimes.com/2006/12/21/technology/21pogue.html

[K+05] Gundars Kaupins and Robert Minch, "Legal and ethical implications of employee location monitoring," In *Proc. of the 38th Hawaii International Conference on System Sciences*, 2005.

[KBCS+11] Patrick Gage Kelley, Michael Benisch, Lorrie Faith Cranor, Norman Sadeh, "When are users comfortable sharing location with advertisers," In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, Vancouver, BC, Canada, 2011, pp. 2449--2452.

[L+12] "Location-based advertising – relevant trends and technologies," Frost & Sullivan, 2012.

[B+12a] Nick Bilton, "Girls around me: an app takes creepy to a new level," *The New York Times*, Mar. 30, 2012. http://bits.blogs.nytimes.com/2012/03/30/girls-around-me-ios-app-takes-creepy-to-a-new-level/

[G+97] Michael F. Goodchild, "Unit 002- What is geographic information science?" *The NCGIA Core Curriculum in GIScience*, 1997, http://www.ncgia.ucsb.edu/giscc/units/u002/ Accessed July 29, 2012.

[BKSC+11] Michael Benisch, Patrick Gage Kelley, Norman Sadeh, Lorrie Faith Cranor, "Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs," *Personal Ubiquitous Computing*, vol. 15, no. 7, 2011, pp. 679--694.

[FW+09] Ryan Farley, and Xinyuan Wang, "Roving bugnet: distributed surveillance threat and mitigation," *Emerging Challenges for Security, Privacy and Trust*, IFIP Advances in Information and Communications Technologies, Springer, 2009, pp. 39--50.

[K+06] Lewis A. Kaplan, United States District Court, S. D. New York, United States of America, v. John Tomero et al., Defendeants, No. S2 06 Crim. 0008(LAK), Nov. 27, 2006.

[MB+06] Declan McCullagh and Anne Broache, "FBI taps cell phone mic as eavesdropping tool," *CNET*, Dec. 1, 2006. http://news.cnet.com/2100-1029-6140191.html

[L+05] Neal Leavitt, "Mobile phones: the next frontier for hackers?" *Computer*, vol. 38, no. 4, Apr. 2005.

[C+10] Suzan Clark, "Pa. school faces FBI probe, lawsuit for using webcams on laptops to watch students at home," *ABC Good Morning America*, 2010. http://abcnews.go.com/GMA/Parenting/pennsylvania-school-fbi-probe-webcam-students-spying/story?id=9905488

[M+10] John P. Martin, "Lower Merion district's laptop saga ends with $610,000 settlement," *The Inquirer*, Oct. 12, 2010. http://articles.philly.com/2010-10-12/news/24981536_1_laptop-students-district-several-million-dollars/2

[GS+05] Martin Gill and Angela Spriggs, "Home Office Research Study 292: Assessing the impact of CCTV," Home Office Research, Development and Statistics Directorate, 2005.

[P+12] Ian Parker, "The Story of a Suicide," *The New Yorker*, Feb. 6, 2012. http://www.newyorker.com/reporting/2012/02/06/120206fa_fact_parker

[L+01] Joyce W. Luk, "Identifying terrorists: privacy rights in the United States and the United Kingdom," *Hastings International and Comparative Law Review*, vol. 25, 2002, pp. 223--259.

[W] "Workplace privacy," Epic.org: Electronic Privacy Information Center. http://epic.org/privacy/workplace/#technologies Accessed August 27, 2012.

[J+09] Joel Johnson, "Is your carrier tracing you via GPS and 911 calls?" *Popular Mechanic*, Oct. , 2009. http://www.popularmechanics.com/technology/how-to/4258805

[NTLHB+11] Arvind Narayanan, Narendran Thiagarajan, Mugdha Lakhani, Michael Hamburg, and Dan Boheh, "Location proximity via private proximity testing," In *Proc. 18th Annual Network and Distributed Systems Symposium (NDSS '11)*, San Diego, CA, 2011.

[BZ+06] Michael Barbaro and Tom Zeller, Jr., "A Face Is Exposed for AOL Searcher No. 4417749," *The New York Times*, Aug. 9, 2006.

[MC+08] Aleecia McDonald and Lorrie Faith Cranor, "The Cost of Reading Privacy Policies," *I/S: A Journal of Law and Policy for the Information Society*, vol. 4, no. 3, 2008, pp. 540--565.

[KW+09] Balachander Krishnamurthy and Craig E. Wills, "Privacy Diffusion on the Web: A Longitudinal Perspective ," In *Proc. of the 18th International Conference on World Wide Web (WWW '09)*, Madrid, Spain, 2009, pp. 541--550.

[V+11] Jennifer Valentino-DeVries, "Facebook Defends Getting Data From Logged-Out Users," *The Wall Street Journal Blog*, Sept. 26, 2011. http://blogs.wsj.com/digits/2011/09/26/facebook-defends-getting-data-from-logged-out-users/

[SCMTH+09] Ashkan Soltani, Shannon Canty, Quentin Mayo, Lauren Thomas, and Chris Jay Hoofnagle, "Flash Cookies and Privacy," SSRN, Aug. 10, 2009.

[F+08] Thomas Frank, "Session variables without cookies," Jan. 20, 2008. http://www.thomasfrank.se/sessionvars.html

[M+12] Mozilla Developer Network, "DOM Storage," https://developer.mozilla.org/en-US/docs/DOM/Storage Accessed August 15, 2012.

[W+10] W3C, Web SQL Database Working Group Note 18, Nov. 2010. http://www.w3.org/TR/webdatabase/.

[MC+11] Aleecia McDonald and Lorrie Faith Cranor, "A Survey of the Use of Adobe Flash Local Shared Objects to Respawn HTTP Cookies," Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-CyLab-11-001, 2011.

[AWSGH+11] Mika Ayenson, Dietrich James Wambach, Ashkan Soltani, Nathan Good, and Chris Jay Hoofnagle, "Flash Cookies and Privacy II: Now with HTML5 and ETag Respawning," SSRN, 2011.

[K+10] Samy Kamkar, "Evercookie," 2010. http://samy.pl/evercookie/.

[G+06] Jeremiah Grossman, "I know where you've been," Blog post, Aug. 11, 2006. http://jeremiahgrossman.blogspot.ro/2006/08/i-know-where-youve-been.html.

[E+10] Peter Eckersley, "How unique is your web browser?" In *Proc. of the 10th International Conference on Privacy Enhancing Technologies (PETS 2010)*, Berlin, Germany, 2010, pp. 1--18.

[D+09] Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009.

[V+10] Tanzina Vega, "Code That Tracks Users' Browsing Prompts Lawsuits," *The New York Times*, Sep. 20, 2010. http://www.nytimes.com/2010/09/21/technology/21cookie.html

[V+10 (2)] Tanzina Vega, "New Web Code Draws Concern Over Privacy Risks," *The New York Times,* Oct. 10, 2010. http://www.nytimes.com/2010/10/11/business/media/11privacy.html

[ULCSW+12] Blase Ur, Pedro G. Leon, Lorrie Faith Cranor, Richard Shay, Yang Wang. Smart, Useful, Scary, Creepy: Perceptions of Online Behavioral Advertising. In *Proc. of the Eighth Symposium on Usable Privacy and Security (SOUPS '12)*, Washington, DC, 2012.

[GDJ+02] Mark A. Graber, Donna M. D'Alessandro, and Jill Johnson-West, "Reading level of privacy policies on Internet health web sites," *Journal of Family Practice*, vol. 51, no. 7, 2002, pp. 642--645.

[USC+12] Blase Ur, Manya Sleeper, Lorrie Faith Cranor, "{Privacy, Privacidad, Приватност} Policies in Social Media: Providing Translated Privacy Notice," In *Proc. of the WWW Workshop on Privacy and Security in Online Social Media (PSOSM '12)*, Lyon, France, 2012.

[L+11] Steve Lohr, "The Default Choice, So Hard to Resist," *The New York Times*, Oct. 15, 2011. http://www.nytimes.com/2011/10/16/technology/default-choices-are-hard-to-resist-online-or-not.html

[LCMM+10] Pedro G. Leon, Lorrie Faith Cranor, Aleecia M. McDonald, and Robert McGuire, "Token Attempt: The Misrepresentation of Website Privacy Policies through the Misuse of P3P Compact Policy Tokens," In *Proc. of the 9th Annual ACM Workshop on Privacy in the Electronic Society (WPES '10)*, Chicago, IL, 2010, pp. 93--104.

[W3C+11] W3C Tracking Protection Working Group. http://www.w3.org/2011/tracking-protection/

[NS+12] Natasha Singer, "Do Not Track? Advertisers Say 'Don't Tread on Us,'" *The New York Times*, Oct. 13, 2012. http://www.nytimes.com/2012/10/14/technology/do-not-track-movement-is-drawing-advertisers-fire.html

[B+12b] Peter Bright, "Microsoft sticks to its guns, keeps Do Not Track on by default in IE10," *Arstechnica*, Aug. 8, 2012. http://arstechnica.com/information-technology/2012/08/microsoft-sticks-to-its-guns-keeps-do-not-track-on-by-default-in-ie10/

[EB+12] Ed Bott, "The Do Not Track standard has crossed into crazy territory," *ZDNet*, Oct. 9, 2012. http://www.zdnet.com/the-do-not-track-standard-has-crossed-into-crazy-territory-7000005502/

[LCCGHUX+12] Pedro Giovanni Leon, Justin Cranshaw, Lorrie Faith Cranor, Jim Graves, Manoj Hastak, Blase Ur, and Guzi Xu, "What Do Online Behavioral Advertising Privacy Disclosures Communicate to Users?" In *Proc. of the 2012 ACM Workshop on Privacy in the Electronic Society (WPES '12)*, Raleigh, NC, 2012, pp. 19--30.

[JM+11] John Montgomery, Testimony before the Senate Commerce, Science, & Transportation Committee Hearing on "The State of Online Consumer Privacy," Mar. 16, 2011. http://www.iab.net/media/file/DC1DOCS1-432016-v1-John_Montgomery_-_Written_Testimony.pdf

[NAI] Network Advertising Initiative. http://www.networkadvertising.org/choices/

[DAA] Digital Advertising Alliance. http://www.aboutads.info/choices/

[EV] Evidon. http://www.evidon.com/consumers/profile_manager#tab3

[EVb] Evidon. http://www.evidon.com/consumers/profile_manager#tab1

[LUBCSW+12] Pedro G. Leon, Blase Ur, Rebecca Balebako, Lorrie Faith Cranor, Richard Shay, and Yang Wang, "Why Johnny Can't Opt Out: A Usability Evaluation of Tools to Limit Online Behavioral Advertising," In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, Austin, TX, 2012, pp. 589--598.

[GOOG] Google. https://www.google.com/settings/ads/onweb/

[MS] Microsoft. https://choice.live.com/data/Dashboard

[YAH] Yahoo! http://info.yahoo.com/privacy/us/yahoo/opt_out/targeting/details.html

[ICO+12] Information Commissioner's Office, "Guidance on the rules on use of cookies and similar technologies," May 2012. http://www.ico.gov.uk/for_organisations/privacy_and_electronic_communications/the_guide/~/media/documents/library/Privacy_and_electronic/Practical_application/cookies_guidance_v3.ashx

[DNTP] Frequently Asked Questions about DNT+. http://www.donottrackplus.com/faqs.php. Accessed August 27, 2012

[MOZ] Mozilla Firefox Add-ons. https://addons.mozilla.org

[COL] Mozilla Collusion. http://www.mozilla.org/en-US/collusion/

[WCL+12] Robert J. Walls, Shane S. Clark, and Brian Neil Levine, "Functional Privacy or Why Cookies Are Better with Milk,"In *Proc. of the 7th USENIX conference on Hot Topics in Security (HotSec '12)*, Bellevue, WA, 2012.

[RKW+12] Franziska Roesner, Tadayoshi Kohno, and David Wetherall, "Detecting and Defending Against Third-Party Tracking on the Web," In *Proc. of the 9th USENIX Conference on Networked Systems Design and Implementation (NSDI '12)*, San Jose, CA, 2012.

[N+04] Helen F. Nissenbaum, "Privacy as Contextual Integrity," *Washington Law Review*, vol. 79, no. 1, 2004.

[HN+09] Daniel C. Howe and Helen Nissenbaum, "TrackMeNot: Resisting Surveillance in Web Search," In Ian Kerr, Carole Lucock, Valerie Steeves (Eds.), *Lessons from the Identity Trail: Anonymity, Privacy, and Identity in a Networked Society*, Oxford University Press, USA, 2009, pp. 417--436.

[S+07] Christopher Soghoian, "The Problem of Anonymous Vanity Searches," *I/S: A Journal of Law and Policy for the Information Society*, vol. 3, no. 2, 2007, pp. 297--316.

[D+10] Jennifer Valentino-DeVries, "How to use Microsoft's InPrivate filtering," *The Wall Street Journal Digits Blog*, Aug. 1, 2010. http://blogs.wsj.com/digits/2010/08/01/how-to-use-microsofts-inprivate-filtering/

[FTC+12] United States Federal Trade Commission, "Protecting Consumer Privacy in an Era of Rapid Change," Mar. 2012.

[WH+12] United States White House, "Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy," Feb. 2012.

[BGW+01] Nikita Borisov, Ian Goldberg, David Wagner, "Intercepting Mobile Communications: The Insecurity of 802.11," In *Proc. of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom '01)*, Rome, Italy, 2001, pp. 180--189.

[CERT+11] US-CERT, Wi-Fi Protected Setup (WPS) PIN brute force vulnerability (Vulnerability Note VU#723755), Dec. 27, 2011. http://www.kb.cert.org/vuls/id/723755

[C+11] Lorrie Faith Cranor, "A First Look at Internet Explorer 9 Privacy Features," *Technology | Academics | Policy Blog*, Mar. 16, 2011. http://www.techpolicy.com/Blog/March-2011/A-first-look-at-Internet-Explorer-9-privacy-featur.aspx

[ABJB+10] Gaurav Aggarwal, Elie Bursztein, Collin Jackson, and Dan Boneh, "An analysis of private browsing modes in modern browsers," In *Proc. of the 19th USENIX Conference on Security (USENIX Security '10)*, Washinton, DC, 2010.

[R+09] Thomas Ricker, "Video: Hacker war drives San Francisco cloning RFID passports," *Engadget*, Feb. 2, 2009. http://www.engadget.com/2009/02/02/video-hacker-war-drives-san-francisco-cloning-rfid-passports/

[HBFKM+08] Daniel Halperin, Thomas S. Benjamin-Heydt, Kevin Fu, Tadayoshi Kohno, and William H. Maisel, "Security and Privacy for Implantable Medical Devices," *IEEE Pervasive Computing*, vol. 7, no. 1, pp. 30--39, Jan.--Mar. 2008.

[CF+11] Shane S. Clark and Kevin Fu, "Recent Results in Computer Security for Medical Devices," In *International ICST Conference on Wireless Mobile Communication and Healthcare*, Kos Island, Greece, Oct. 2011.

[DBFGKM+10] Tamara Denning, Alan Borning, Batya Friedman, Brian T. Gill, Tadayoshi Kohno, and William H. Maisel, "Patients, Pacemakers, and Implantable Defibrillators: Human Values and Security for Wireless Implantable Medical Devices," In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, Atlanta, Georgia, 2010, pp. 917--926.

[T+12] David Talbot, "Computer Viruses Are 'Rampant' on Medical Devices in Hospitals," *MIT Technology Review*, October 17, 2012. http://www.technologyreview.com/news/429616/computer-viruses-are-rampant-on-medical-devices/

[EGKP+11] Miro Enev, Sidhant Gupta, Tadayoshi Kono, and Shwetak N. Patel, "Televisions, Video Privacy, and Powerline Electromagnetic Interference," in *Proc. of the 18th ACM Conference on Computer and Communications Security (CCS '11)*, Chicago, IL, 2011, pp. 537--550.

[CRSXLF+12] Shane S. Clark, Benjamin Ransford, Jacob Sorber, Wenyuan Xu, Erik Learned-Miller, and Kevin Fu, "Current Events: Identifying Webpages by Tapping the Electrical Outlet," University of Massachusetts, Amherst, M., Tech. Report UM-CS-2011-030, Jul. 2012.

[MSFCI+10] Andrés Molina-Markham, Prashant Shenoy, Kevin Fu, Emmanuel Cecchet, and David Irwin, "Private Memoirs of a Smart Meter," in *2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys '10)*, Zurich, Switzerland, Nov. 2010.

[PRKRA+07] Shwetak N. Patel, Thomas Robertson, Julie A. Kientz, Matthew S. Reynolds, Gregory D. Abowd, "At the flick of a switch: detecting and classifying unique electrical events on the residential power line," in *Proc. of the 9th International Conference on Ubiquitous Computing (Ubicomp '07),* Innsbruck, Austria, 2007.

[S+10] Stuart Schechter, "Security that is Meant to be Skin Deep: Using Ultraviolet Micropigmentation to Store Emergency-Access Keys for Implantable Medical Devices," in *1st USENIX Workshop on Health Security and Privacy (USENIX HealthSec '10)*, Washington, DC, 2010.