The Decline of Internet Privacy

Amey A Rairikar

Department of Electrical and Computer Engineering, Rutgers University
New Brunswick, New Jersey, USA
ameyr@eden.rutgers.edu

Abstract—The growth of Internet has enabled more people to access information freely over the web. Amongst the plethora of data available, there is also a significant amount of Personally Identifiable Information which can be misused. An average user is not aware of his private data being collected when he or she surfs the internet. Lot of services and applications over the web use some kind of Personal Information to provide service to the end user. The information that is collected can be viewed by many other unauthorized entities and used for their own benefit. As more people gain access to the Internet there will be more personal information collected by various methods and hence ‘Privacy’ as we know it is on a rapid decline.

In this paper we enumerate the causes for the decline of privacy. We inspect the tools and regulations to protect the user from privacy attacks and analyse why these methods are not sufficient to stop the decline and the death of Privacy over the Internet.

I. INTRODUCTION

Privacy is a notion which has been associated with human beings long before the advent of the information and the internet age. Over the years the issue of Privacy has been addressed in context of lot of economic and social backgrounds. According to Roger Clarke, Privacy can be defined as the interest of individuals in sustaining their personal space free from interference by other people or organizations.

If we look deeper into the topic of Privacy it can be applicable to several dimensions. Some of the issues include

- Personal Privacy
- Privacy of Personal Behaviour
- Privacy of Data
- Privacy of Communication.

The issue of privacy is an important to the people for social, economical and political reasons. Privacy means that people are free to have private space, they are free to communicate and think and associate with others. Lack of privacy leads to strict regimes, hardships for the people and loss of free human spirit.

The Internet has evolved at an exponential rate since the years. It has indeed encompassed almost all aspects of our lives from buying groceries to watching live satellite feed of the traffic movements across busy streets. The naissance of Internet was a collection of computers connected by a network to share information to be used for research purposes. Now the amount of data and information available over the internet is vast and is easily accessible to any user from any country at any geographical location in the world. According to internet usage statistics the Internet usage growth has been tremendous over the last few years with a 305% growth worldwide and the percentage of growth reaching 4 figures in some countries in Asia, Africa and the Middle East.

The tremendous growth in the internet usage and its application to day to day lives is indeed a feat of technology. But there is a growing concern about the free availability of personal and critical data over the internet. These concerns are exalted by the existence of large databases which can store personal information in many forms. Moreover it is very easy to demonstrate that just with a click it is possible to obtain information from the user and store it in the database. Post 9/11 there is a growing concern about security in general. This has prompted the watchdogs to monitor and collect information about people using freely available data and create a profile based on that person. This has led to more monitoring and control over a persons activity over the internet and hence is a significant intrusion into the private space as we know it. Moreover many companies sell their databases to advertising firms, corporate firms which means that your information as a user is now available to anyone for a price. Thus the concern about privacy over the internet is justified.

An interesting statistic that was found was the percentage penetration of population by the internet.The figure shows the percentage of populations of the regions which have access to internet and use it in their daily lives. From the figure it is obvious that as the number of users increases so does the percentage of total population using the internet. Hence there is more information available over the internet and also correspondingly more amount of private information can be available easily. Consider North America for example.

We see that 74 % of the population use the internet, the increased use of applications which require some kind of personal information to be entered and the effectiveness of the data collection and mining methods means that the personal information of a significant portion of population is available on the internet. With the increase in penetration every year it is clear that the notion of Internet Privacy is on the decline and is Privacy is dying. Even Scott Mc Nealy, CEO Sun Microsystems famously quoted “Privacy is dead, Get over it." There have been efforts in the recent past to protect the users privacy over the Internet. The governments over the globe have put forth a series of Internet Privacy Legislations.
There has also been a significant increase in the amount of Privacy Protection tools available on the internet.

In this paper we address the issue of decline in Internet Privacy with respect to the main causes in the present time. We base our analysis on various studies and reports available in the community. We also take a look at the methods available to protect the user and why they might not be adequate enough to stop or hinder the decline in the users privacy over the internet.

II. THE DECLINE OF PRIVACY

There has been a significant increase in the amount of Internet use over the past few years. Most of the users provide their personal information over the web without pondering over who can access this information. Personal Information including hobbies, employment, preferences etc. can be collected when a person is online.

Most of the information over the internet is collected by organizations to help them in their day to day business activities. But businesses, government and also any person with a malicious intent can use this information for their own advantage whether it will be to advertise a new product, or simply to monitor your activities for security reasons. There are several reasons why there is so much private and potentially critical data available over the Internet. In this section, the causes behind the decline of privacy are investigated. The reasons range from a users attitude while he or she surfs the net to a harmless looking record of what you surf: cookies. We try to outline why these reasons are a serious threat to online privacy.

1. User knowledge and attitude

It is important to analyse a users attitude when he/she surfs the net because ultimately a user is responsible for what information he/she puts on the web. Most of the users are unaware of the technology that keeps the internet running and hence are ignorant to most of the privacy issues associated.

In this section we present a study conducted in [3], in which a survey was conducted asking users about the reasons for which they surf the Internet, how privacy concerns them and about privacy policies of the websites that they visit. From the survey it is found that over 60% of the sample use the web to surf or to use their email. Even when surfing or registering for an email service a user might be prompted to give information to the application provider in order to use the application properly. In the figure below we see how comfortable users are in providing certain information over the internet.

![Figure 1: World Internet Penetration Rates.](source: Internet World Stats - www.internetworldstats.com/stats.htm)

Penetration Rates are based on a world population of 6,876,120,289 for mid-year 2008 and 1,463,632,361 estimated Internet Users.

**Sources:**

- Internet World Stats - www.internetworldstats.com/stats.htm
- Penetration Rates are based on a world population of 6,876,120,289 for mid-year 2008 and 1,463,632,361 estimated Internet Users.

Copyright © 2008, Internet Marketing Group

**FIGURE 2: Internet Users Comfort Level while Providing Certain Data**

From the above figure we make the following observations:

- Close to 90% of the easily provide information about their hobbies, interests and general information.
- 60 – 80% of the people are comfortable with providing their name and email addresses.

So considering a large population of Internet users we can assume that up to 70% of the users provide their name, email, favourite programs, hobbies etc to websites or applications over the internet. Now one might think that this information is not enough to steal your privacy but on the contrary this information is enough to generate a pattern on the user which profiles him according to his likes, dislikes etc. This information gets stored in databases which can be retrieved by firms to market or advertise their products based on the information that they know prior, generating annoying ads and spam mail over the emails which is indeed Invasion of the user privacy.

Another important factor in the user attitude are the opinions about the websites he or she visits while surfing. We analyse a statistic provided by the survey in [3]. It essentially captures the users awareness about the constancy of the website with respect to their data privacy. The table below shows the attitude of the users towards website policies.
Users generally access internet through ISPs or the Internet Service Providers. So all the data that the user accesses flows through the ISP connections to the network. It is possible for an ISP to monitor and control all the data that a particular user might access or use. However it is a policy of the ISPs not to monitor the data. To what extent this policy is followed is not yet clear.

The ISPs do collect a lot of personal information before a connection is given to a user. Information such as the IP address, user ID, and sometimes website access flows through the ISP connections to the network. It is possible for an ISP to monitor and control all the data that a particular user might access or use. However it is a policy of the ISPs not to monitor the data. To what extent this policy is followed is not yet clear.

From the above table we can see that the users tend to think that data availability and sharing over the website is more important that data privacy policy.

It is also from the survey conducted that 80% of the users are very concerned about the privacy and tend to mistrust the internet over privacy matters. In spite of that most of the users neglect the privacy policy and give basic information about them over the net. Thus it can be derived that even concerned users unwarily divulge information which can be used to invade privacy. If access to data is the primary goal of the user then soon privacy will be neglected from the users side itself leading to an even more availability of data and decline of user privacy.

2. **Cookies: Not so sweet!**

Cookies are perhaps the most dangerous threats to your privacy on the internet. The main reason which makes them dangerous is that a normal user is not aware of the presence of cookies and the importance of them. In fact according to the survey [3] 48 % of the users did not know what cookies were.

Although not intended for malicious use, ironically cookies are perceived as significant threat to Internet Privacy. HTTP cookies are used by Web Servers to maintain data related to a visit by a user to the particular server. Cookies are pieces of a data sent by the Web Server to the Browser. They are sent by the browser to the Web Server denoting the state of the communication. Next time when the same Web Server is accessed the cookie help to load the information from the previously stored state which helps to increase throughput and decrease loading times. They are also used to customize web pages, log in authentication etc.

Although cookies provide a good browsing experience they are prone to lot of attacks and faults which mainly result in user data being misused. If more than one browser is used on a computer, each has a separate storage area for cookies. Hence cookies do not identify a person, but a combination of a user account, a computer, and a Web browser. Thus, a single person has multiple sets of cookies if they use multiple accounts, computers, or browsers. On the other hand, cookies do not differentiate between multiple users who share a computer and browser, if they do not use different user accounts.

One more disadvantage of a cookie is cookie theft. The intent of a cookie is to traverse back and forth between the Browser and the Server updating the user state information. Generally a cookie is a sensitive data packet which contains user information and should not be available to any unauthorized person. But in a normal HTTP session we often observe that cookies are visible to all the users and they can “listen” using packet sniffers. Cross site scripting allows the value of cookies to be sent to servers that are normally not sent these values. Modern browsers allow execution of segments of code retrieved from the server. If cookies are accessible during execution, their value may be communicated in some form to servers that should not access them. Encryption methods are generally not useful with respect to cookies.

Attacks can be done by people using websites which allow the users to post their HTML content. By using a suitable piece of code the attacker can effectively divert the cookie traffic towards himself. The information in these cookies can be extracted using the stolen cookies and connecting to the websites. Thus effectively harnessing potentially important information of the user.

Cookies are supposed to be stored and sent back to the server unchanged, an attacker may modify the value of cookies before sending them back to the server. If, for example, a cookie contains the information about a transaction online, an attacker might access this information to alter the values of the transaction. This might lead to losses financially as well as loss of privacy. The process of tampering with the value of cookies is called cookie poisoning, and is sometimes used after cookie theft to make an attack persistent.

Given the dormant nature of cookies on our system one way to prevent harm is to disable the use of cookies in the browser. However a recent article on Microsoft’s website states that when cookies are deleted some cookies are left untouched thus exposing a flaw in the solution of disabling or deleting the cookies. The information on the cookies is again useful for determining browsing patterns. Once known some websites might enforce this profile on you by using pop up ads related to your browsing patterns. Also it changes the way you see browsing as a whole. Instead of looking for information yourself you end up doing what the Server thinks you will do based on the profile created using data captured from the cookies. Less number of people are aware of this problem. As we have seen not a significant majority know the use and possible misuse of cookies. This results in more users being drawn in a type of ‘Suggested’ Browsing without knowing that their privacy and thinking process was altered by the use of a visibly harmless file.

## Table I
**User Attitude Towards Website Policies.**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Very Important</th>
<th>Quite Important</th>
<th>Not Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy policy</td>
<td>20</td>
<td>36</td>
<td>44</td>
</tr>
<tr>
<td>Access to data</td>
<td>68</td>
<td>24</td>
<td>8</td>
</tr>
<tr>
<td>Sharing data</td>
<td>76</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Parental consent</td>
<td>48</td>
<td>28</td>
<td>24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attack</th>
<th>Description</th>
</tr>
</thead>
</table>
| Cross site scripting | Allows execution of segments of code retrieved from the server.
| Cookie poisoning | Modifies the value of cookies before sending them back to the server.
| Cookie theft | Diverts cookie data to another location.

<table>
<thead>
<tr>
<th>Encryption Methods</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not useful with respect to cookies.</td>
<td></td>
</tr>
</tbody>
</table>
address, Billing details are available with most of the ISPs. An IP address is a unique identifier assigned to each computer that is connected to the network. Every transfer of information over the Internet must include the capture of the IP address. Some examples of automatic logging are visiting a web site, sending or receiving e-mail, using a chat room, or reading and posting to newsgroups.

Transferring IP addresses to a third party can also be accomplished by sending a web page via e-mail. When the user opens the attachment (if they are connected to the Internet) the e-mailed web page could make a request to a web site anywhere on the Internet (such as requesting an image file). This transfers the user's IP address to that web site along with the date and time that the user opened the message. A cookie can also be placed on the user's system at that time. This enables advertising firms to directly advertise to the user whose IP address is known.

IP address collection by ISPs and by a third party are violation of personal privacy. After attaining the IP addresses anyone can look up for the specific IP address/ Computer name on the Internet and generate a profile on that person. Monitoring by ISPs for security might lead to your IP address being taken by security enforcers without your consent, thus keeping a constant surveillance over your Internet activity.

4. Social Networking

With the increasing popularity of staying connected over the internet and sharing multimedia, lot of social networking websites have become very popular over the last decade. To give an example of how popular the social networking sites really are, in a survey by Nielsen/Net ratings it was said that there are over 90 million users of the social networking website www.Myspace.com. This is approximately one third of the current US population.

Social Networking is a collection of websites which allow users to put up information about themselves such as their academic interests, age, phone numbers etc. This information then can be viewed by their connections or by anyone who visits their customised page. MySpace is the 3rd most visited website while Facebook stands at 7th spot. They both beat Google with respect to generating hits.

In an survey to understand the attitudes toward social networking sites and privacy, a classroom attitudinal survey was conducted to collect data about student attitudes about Facebook [.]. In an exploratory survey, conducted with 64 undergraduate students, it was learned that 65 percent of the students used Facebook or MySpace. Another scale was utilized to collect data about student attitudes about privacy when they use social networking sites. The students were asked questions which included the topics of putting their private information on the website and trusting the people who view their private information.

The subjects of the survey did not seem to care much about sharing their private information on social networking websites. But they did show a concern on privacy without realizing that social networking helps to spread their information over the web and makes them a target of advertisements, spams, suggestions and other such suggested surfing links based on the profile of the person.

If a given user creates a profile on a typical website like Facebook he enters his age, sexual orientation, race, education background and all such information which can be used in identifying the person. Such information will be used by companies to target the user for specific advertisement. If the profile that you created mentions that your relationship status is ‘single’, there will be dating services advertisements on your Facebook Page. Similarly in Google mail chat, the subject content of the emails is often mined for keywords which help in putting advertisements on the Webmail page related to the keywords found in the mail.

It is critical to notice that the content in your mail or chats or messages exchanged through social networking tools can easily be watched by the application providers themselves to perpetually exploit the advertising and marketing opportunities. Teenagers and other people using social networking sites tend to think that unwanted people cannot read their personal entries, but it is quite easy to access their profile through indirect means. It is interesting to know that social networking sites are generating more hits annually and they are competing with giants like Yahoo and Google.

5. Security and Watchdogs.

Post 9/11 it is obvious that Governments all over the world are making efforts to obtain more information about people so
that any possible threat to security is dealt with accordingly. An effective way to do this is eavesdropping or tapping the telecommunication resources. The aim of the government is to obtain enough information through data mining techniques so that individuals can be monitored with respect to the profile that they create by surfing the internet.

Governments have tied up with major corporate players in the telecom industry to keep a watch over the internet users. For example, according to the New York Times, the NSA has worked with "the leading companies" in the telecommunications industry to collect communications patterns, and has gained access "to switches that act as gateways" at "some of the main arteries for moving voice and some Internet traffic into and out of the United States.

A representation of how NSA can achieve surveillance over the internet is shown in the following figure.

**Figure 4: NSA Surveillance Representation.**

By tapping the nodes and exchanges it is possible for the NSA to effectively monitor all of your internet traffic in case it decides to do so.

Data mining algorithms are used by security organizations to create profiles and alerts based on some keywords that might appear in your private voice calls, emails, messages and other communication over the internet. So it is possible for an individual to be arrested on basis of his or her profile as created from the keywords. Use of these statistical fishing methods has been made possible by the access to communications streams granted by key corporations. The NSA may also be engaging in worldwide targeting in which they listen in on communications between the United States and a particular foreign country or region. More broadly, data mining has been greatly facilitated by underlying changes in technology that have taken place in the past few years.

There is also a new concept of Biometrics in which information regarding to your physical attributes such as fingerprints, iris, speech metrics can be stored on a machine readable chip. The biometric passport has already been launched for USA. If this information stored on a Government Database can be accessed by any unauthorized person then it can lead to identity theft and such other serious infractions.

In an age where privacy concerns are growing, the laws proposed by the Government are hardly enough to protect the users privacy in addendum the Government by using high end surveillance and data mining tools over the internet is effectively violating privacy of an individual on all levels.

6. **Phishing, Spyware and other threats to Privacy.**

Phishing is a criminal process in which personal information such as a user Social Number, Credit card detail etc can be obtained by posing as a message or a mail originating from a trustworthy entity. Phishing is more common in emails where unsuspecting users are often the prey to maligners. From the survey presented in [ ] we have seen that the attitude of the average internet user is casual with respect to his privacy and most users want data to be available as quickly as possible.

![Phishing Reported between October 2004 to June 2005](image)

**Figure 5: Trends in Phishing over period of 9 months.**

Phishing causes people to disclose private data to websites or servers who appear as a popular trustworthy source but are actually just mere lookalikes. Its hard to distinguish between the authentic and the false website, even SSL methods employing cryptography fail to identify the fake. The graph in Phishing reported over the internet is shown in figure 5. It clearly shows increasing attacks and attempts to gain private information.

Spyware is software or a bug which can be employed over the users internet connection to get information about their machines and data stored on the host computer. Most of the spywares are embedded with lot of free softwares and sharewares available on the internet. An example of a spyware was the Netbus bug which when put into a host allows the controller to monitor the host actions and even control the normal functioning of the host.

Apart from the above there are also other factors and methods to invade privacy of a user. They include Browser Chatter, ad networks, web bugs etc. Thus there are plenty of threats to a users privacy over the internet, from our study of these causes it is found that the number of attacks or attempts
to invade privacy has increased giving rise to less secure internet experience for the average user.

III. PROTECTION METHODS

The threat to data security and user privacy is serious and there are increasing concerns in the user community about these issues. These concerns are alleviated by the availability of various privacy protection tools available over the internet. In this section we mention a few popular tools used to protect a users privacy.

A protocol developed for protection of privacy is called the P3P or Platform for Privacy Preferences. P3P helps the user to choose what information he or she puts on the Internet for general surfing and activities in which data is collected from the user. P3P aims to create a user agent which will control the access to different websites based on the Privacy Policy of that particular website.

Another useful way to protect ones privacy is to surf the web anonymously. The basic principle behind anonymous surfing is that the users IP and other details are hidden from the destination server. Instead a proxy server called as an anonymous proxy is responsible for acting as an intermediary in which the information stays hidden from end users.

Anti spyware tools have been of quite importance lately due to the increase in the number of spyware attacks. The anti spyware software is responsible for searching the user memory for any bugs or spyware and also removing it from the system. Users can also join a crowd of users which prevent the destination server from getting information about any single user. It also protects individual data from the other users in the crowd.

Firewall is another popular method to ensure data privacy. It is essentially a collection of protocols which prevent unauthorized access to a computer on the network. Firewalls implement data filters on several communication layers like the network layer, application layer. They also provide NAT or the Network Address Translation to hide the identity of protected hosts. In addition some mentioned tools above there are many resources available to protect privacy over all aspects of the internet. Some of these include Cookie busters, Voice Encryptors, Email privacy, SSL, Snoop Proof Email etc.

In addition to the tools available to protect the privacy recent concerns about the invasion of privacy and misuse of personal information has forced the Governments to address the issue. The US government has issued sector specific laws which deal with the issue of privacy like the Privacy Laws of 1974 and the Electronics Communication Privacy act of 1994. The EU has also issued more stringent laws to protect privacy over the internet which include the Data Protection Directive of 1995, The first data protection law in fact was issued in Germany in 1970.

IV. SHORTCOMINGS AND ANALYSIS

In the preceding section we mentioned some methods and laws which were developed to protect a users privacy. However there are several shortcomings and problems related to the methods discussed and their application.

P3P was considered to be an effective solution to protect user privacy but the report [ ] found that P3P is not based on any standards. The usage and application of P3P is confusing and complex for the average internet user. Thus it is desirable to use any other technique over P3P. Firewalls do protect internet against malicious content but they also cause a decrease in throughput and bandwidth utilization. Anti spyware tools are pretty common but there exist lot of rogue anti spyware tools that will do more harm than good.

The biggest shortcomings in defining practical regulations have been in the legislatures related to Internet Privacy. Although laws have been passed, it is found that US anti privacy laws are lenient and address specific areas of the problem. Also Government has left some sectors untouched fearing too many regulations in the private sector.

There has been a significant increase in the number of anti spyware software and data protection utilities. The increased number itself is warning to the steady decline of privacy. Users are now aware that their information can be misused. Self regulation and control are also supposed to hinder private information availability. But the attitude of the users to the subject of Privacy is still casual. In a study [ ] it has been found that Google ranks last in terms of Privacy. Firewalls, Anti Spyware and other such tools cannot effectively prevent websites like Google from generating an individual profile based on illicit data collection. Considering that Google is one of the most popular websites in the world it is indeed apparent that more people will be profiled for advertisement and marketing purposes which profits Google.

While there exist number of steps to try and protect privacy entirely, more number of people register to social networking sites divulging their private information to web servers. The ultimate goal of social networking site is to connect all the users to each other. It is a fact that people trust users of these sites and assume that their information is safe, but this is far from the truth. When all people are connected to each other ultimately via these websites, your information is available to anyone and everyone. Web based applications will target you for advertising and that will mean destruction of a significant portion of privacy.

Post 9/11 paranoia about security has engulfed the world. In these times when freedom is endangered and privacy is violated, Government agencies are using extensive data mining and collection tools to create an identity for each person that leaves some kind of trail on the internet. Instead of protecting citizen’s privacy they are trying to watch and record information about citizens to exercise tight control over them, in a way compromising freedom and privacy.

We noticed growth in the causes hampering privacy over the internet. Due to tremendous business opportunities there are lucrative incentives for websites or web services to sell
private information to any organization which can use it for their own benefit.

V. CONCLUSION

We have seen the reasons for concern amongst users and networking communities. These reasons are just and are ever present. Privacy as we defined it in Section I is not just related to one particular aspect but encompasses a lot of variations in its definition. From the data presented in this paper it is clear that basic private information like a users Name, Email, Age, Education are easily available over the net to almost everyone. This aspect of Privacy has perhaps died long ago with the advent and popular rise of social networks, Messaging, and such websites. It is obvious that the users by disclosing all such information have exposed their privacy and will continue to do so.

The tools available are effective only to a certain limit. The increase in use of privacy protection tools is constant but it might not be enough to cope up with new methods to collect data and eavesdrop on the users through the Internet. Legislations and Government measures are just mere formalities. Security agencies on the other hand are constantly engaged in invasion of Privacy under the pretext of security.

So long as users want information they will continue to be prone to data collection methods and profiling by a lot of third parties. The need for more potential buyers will force many online businesses to create databases based on profiling of what a user surfs over the Internet. There are no significant breakthroughs in protection of privacy and the self regulation methods have failed. So it is safe to assume that if this trend of decline in personal space continues, then Privacy will soon be a thing of the past.

References