CMI working paper no. 7, 2015

Where Does My Private Data Go?

- Visualization of Users' Privacy

Samant Khajuria and Lene Sørensen

Center for Communication, Media and Information technologies (CMI), Electronic Systems, AAU Copenhagen, Denmark

CMI Working Paper no. 7:

Samant Khajuria and Lene Sørensen (2015) Where Does My Private Data Go? - Visualization of Users' Privacy. AAU, Copenhagen

Paper presented at the 48th Annual Hawaii International Conference on System Sciences (HICSS), workshop of Wireless Applications and Services Beyond 2020, 5 January, 2015

ISBN: 978-87-7152-058-3

Published by:
center for Communication, Media and Information technologies (CMI)
Department of Electronic Systems,
Aalborg University Copenhagen,
A.C. Meyers Vænge 15,
DK-2450 Copenhagen SV
Tel +45 99403661
E-mail cmi@cmi.aau.dk
URL http://www.cmi.aau.dk

Presented at the 48th Annual Hawaii International Conference on System Sciences (HICSS), workshop of Wireless Applications and Services Beyond 2020, 5 January, 2015

Where Does My Private Data Go?

- Visualization of Users' Privacy

Samant Khajuria and Lene Sørensen

Aalborg University, CMI

skh@cmi.aau.dk; ls@cmi.aau.dk

Abstract

Privacy has become a leading concern for many users using online services of any kind: social media activities, health care or shopping. As a consequence, targeted services offer information visualizations of private data for users on a browser or app level. This paper focuses on comparing existing information visualization services to see how they comply with supporting users' control of privacy. The visualization services, however, only address specific areas of user privacy. Therefore, there is a need for a more holistic approach to privacy in which elements such as a Trusting Authority becomes an integrated part of an ecosystem for future services that support users' control of their private data.

1. Introduction

The daily life of online browsing drags a long list of personal information behind on location, email address and IP address, registration of shopping habits and what you clicked on, to mention a few. More often users find specialized ads popping up on different websites, the ads corresponding to the keywords that were used for searches in other websites. For a long time, social media users have been aware of issues in privacy of for example Facebook. As a result many users have become aware of privacy as an element in being online [17]. However, there are many more situations in which private data is collected by third parties: when using special apps on for example Facebook, the data is handed/sold to third parties; how stores, via WIFI, track customers when they are in the store or pass by [9], or when Samsung "Smart TV" records

personal conversations in the living room and transmits these to third parties [20]. Privacy is part of the everyday.

Information Privacy also known as data privacy relates to one's ability to have control and freedom of choice about the collection, user and disclosure of information our selves – what we might call our personal data flows [5]. According to the data protection law, "Data are personal data if they relate to an identified or at least identifiable person, the data subject"[11]. One of the most used privacy definitions from Westin [22] defines privacy, as "the right of the individual to decide what information about himself should be communicated to others and under what condition". The definition was further developed by Smith et al. [18] focusing specifically on perspectives in online privacy concerns – unauthorized secondary use of personal information, improper access of digitally stored personal information, collection of personal information and errors in collected personal information.

Looking into Westin [22] and Smith et al.'s [18] definitions, today an online user is reluctant to disclose their personal online information because the user's privacy concerns are neglected constantly by the organizations. Many users have experienced or heard of friends on daily basis who have experienced violations of privacy one-way or the other. Ideally the owner i.e., the user should be in control of their personal data that has become an economic commodity alias today it is in the hands of the service providers whose business case often revolves around the use of user personal data they collect via social networks, search engines, online retailers etc.

In an online world, privacy is not about hiding or keeping information secret from the service providers but is about understanding the processes and being able to see through unwanted disclosure of private data. One of the tools that can be implemented in that respect is the so-called visualization tools.

Visualization is a well-known and used technique for showing data/information in another form. It is used widely within computer network and computer management [8] and has been for years. Visualizations can be in forms of color maps, coordinate plots, histograms, scatterplots, colored bars and the like. One example of a visualization most users know is the changing of passwords indicator showing the strength of the password; the indicator changes color according to the complexity of the password. However, within computer networks and management many visualizations require expert knowledge in understanding, reading and acting on the visualization tools [8]. These tools are not meant for ordinary users.

The clear strength of using visualization tools is that it is easier to find new and unexpected patterns and a color representation of a series of data, for example, is much easier to grasp than the data. The raised attention on users' privacy and security calls for ways that the users can handle and manage their online activities. Therefore, the visualization tools have entered the field of usable security and privacy where there is a focus on providing user-friendly interfaces representing security and privacy information [1]. In particular when it comes to security and policy policies these can be difficult to read and understand for users, which is why the visualization tools can be more eligible.

Common for privacy visualization tools is that they look at the control users have over information about themselves and the threat can be related to either social or organizational threats [13]. Social threats are those related to other individuals on for example a social networking site. Organizational threats come from the service, social networking sites as example, themselves or by its partners who for example can disclose or sell personal profile data [13]. Most users target the social threats since they do not believe they can do anything with the organizational privacy threats [17]. However a number of currently available privacy visualization tools target particularly the organizational privacy threats (see following section).

In order to increase user trust in present and future services it is necessary to deploy a new privacy ecosystem based on digital technologies that can innovatively – enhance the end users' understanding and control of their personal data in online transactions by using visualization tools, gives users possibility to pay for privacy at desired levels and enable service providers and app developers to engage in explaining their requests, treatments and protections for end users' personal data, ultimately increasing the general acceptability of the user-provider liaison. These privacy ecosystems will not only protect and promote privacy but also at the same time encourage socio-economic opportunities and benefits by rejecting the dated "all or nothing" acceptance policies to a positive-sum mindset focused on win-win solutions. This ecosystem will be able to accommodate mutual interests of user and service providers' in future user-centric services.

For present and future Internet services the most prominent design paradigms are: Security and Privacy by Design [3]. In general nothing is 100 % secure in terms of security and privacy. Instead the requirements from the security and privacy concerns of the participating entities needs to be considered, analyzed and optimally balanced from service to service. In order to

minimize the risk of data breaches in future services, Cavoukian and Chanliau [3] stated that addressing security objectives – i.e. Confidentiality, integrity, availability – as well as privacy goals – i.e. transparency, unlinkability, intervenability – in every system design one may save reputation, resources and revenue. It is very critical to understand that the security does not equal privacy; both the paradigms are converging and complementary.

In this paper we take the stance that the personal data have become an economic asset belonging to the service providers. To increase privacy awareness and reduce ramifications associated with these services there is a clear call for user controlled privacy services. There exists several user controlled privacy services – one of these are visualization tools, one category that has drawn attention with the presentation of a number of new services on the market (such as LightBeam [14] etc. – see more later).

This paper has the purpose to suggest perspectives for user controlled privacy in future services. To do this, the paper analyzes selected visualization tools for user controlled privacy. The analysis is based on a combination between literature study as well as hands-on experience of the authors in respect to the services. The services are compared and discussed and the analysis used as a basis for suggesting a more integrated solution for user controlled privacy for future services including the application of Privacy by Design Principles [3].

The remainder of this paper is structured as follows. In section 2, we present and compare a selection of existing privacy visualization tools that users can apply to either browser or mobile phone in order to follow their data and find out if they disclose information to unwanted third entities. Section 3, looks into how future services can be designed in order to focus on privacy as a key element, visualization tools being one element in a more integrated ecosystem.

2. Privacy Aware Tools

Privacy engineering has become a central part of IT systems and has grown out of the necessity to guarantee people control of their private data [10]. Building IT systems that store and handles personal data, is one thing but building tools/systems, which provides insight into where private data go when users are online, is another thing

In order to make users aware about their online data privacy, a number of privacy aware browser add-on’s and mobile apps have been developed. These tools allow users to see the first

and third party sites the user interacts with on the web. Additionally, some of these tools can also detect and stop third-party trackers from secretly tracking users–and they do this with visualizations so it is more easy for the user to understand. The following tools are representative for services that exist now and they provide an insight to the variety of existing privacy aware tools. General for the approach for these tools is that they receive data from other websites and/or applications and express these in a relatively simple interface.

Privacy Badger [6] is a plug-in build on a tracker protection approach for browsers developed by the Electronic Frontier Foundation (EFF). The fundamental purpose is to analyze and block trackers or ads that violate the user consent. The extension is designed to automatically protect user privacy from third party trackers by letting the user block trackers that may be surreptitiously keeping track of the user's web activity. It can function without any setting, knowledge or configuration by the user and is therefore relatively easy to use for anyone. A so-called third party tracker (trackers which track browsing habits in order to display customized ads) is key in this tool. The user is in this case, presented with a slider in the Privacy Badger menu that shows either a green, yellow or red dependent on level of tracking from different third parties.

The LightBeam [14] tool is a browser plugin developed by Mozilla that enables graphical representation of the first and third party sites interacting with the browser, revealing the full depth of the Web today, including parts that are not transparent to the average user. Using three distinct interactive graphic representations (Graph, Clock and List) it provides insights of individual third parties over time and space, and allows users to identify where they connect to their online activity. The graph gives a real time visualization of all third party requests in the moment a user visit a specific website. The clock allows for examining connections over a 24 hours period. And the list view enables the user to block sites from connecting with the Firefox browser. The user has the possibility to set up filters to see more types of data. The LightBeam tool uses lists and plots over the data information.

Various steps are taken in the app, F-secure Freedome [16], to respect the user's privacy and security. The F-secure Freedome is a VPN service that keeps the user invisible for anonymous browsing by masking the IP address under the protective cloud. Additionally, the app also gives the user a possibility of safe browsing and being untrackable by scanning for malwares and

blocking 3rd party / data brokers. The app has an interface with a large button in the middle that shows whether the user is protected or not. The features are provided against a monthly fee.

F-Secure App permissions [16] is another application that displays the permission of the apps installed on an Android device. It categorizes and ranks the apps based on the permissions it requests. It also informs about the ramifications of the given permissions. App Permissions analyses only apps that have already been installed.

Recently AVG has also developed an online privacy dashboard. AVG PrivacyFix [2] is a browser add-on and mobile app for privacy issues based on a user's Facebook, Google and Linkedin settings. The dashboard gives a user visual representation of what personal data one has exposed and gets advice on how to fix it. The PrivacyFix also lets user know what their data is worth (economically) for example to Facebook and Google.

Another privacy-aware tool for the protection of users' personal online information is MyPermissions [15] – it creates an online privacy shield by Online Permissions Technologies for browsers and applications on Android and Apple devices. The application offers an interface, so the user is able to manage all services permissions in one screen. The app provides information about the permissions requested by other apps like – if an app is acting on the user's behalf, knows the user's location, can access inbox or contact information, and basic permissions, e.g., posting on social media websites on the user's behalf. The app gives users the possibility to Revoke, Trust or Report the permissions requested by other apps. Additionally, the MyPermissions app keep track of other apps' updates, where they might ask for more information about the user's personal data than originally granted, when the app was first installed.

Well known to most online users, the Terms of Service is the first step to privacy awareness. In most cases, Terms of service is often too long to read and the users simply accept the terms without reading it – giving the service providers complete power. However it is important to understand what's in the Terms of Service and that the rights of the user depends on them. Terms of Service; Didn't read [19] is a browser extension that rate and analyze Terms of service and Privacy policies in order to create a rating from class A to class E. Terms of service are reviewed by legal experts and divided into small points that can be discussed, compared and ultimately assigned a score with a badge: + Good, - Mediocre, ¤ Alert , or -> Informative.

In Table 1, the six privacy aware tools referred to here are compared and displayed.

Table 1 Overview of visualization tools and their characteristics

Category Tool	Standalone Privacy Tools	Visualization Type
Privacy Aware - Browser extension	LightBeam	Graphical Representation of 1st and 3rd Parties interacting with the browser
	Terms of Service; Didn't read (ToS; DR)	A rating system that rate and analyze terms of service and privacy policies
Privacy Protection - Browser extension and Mobile apps	F-secure Freedome	A paid VPN app for anonymous browsing, keeps the user untrackable by blocking 3rd parties and trackers.
	AVG PrivacyFix	Specially designed for Social networks; gives user visual representation of what personal data more have exposed and how to fix it.
	My Permissions	Privacy management tool for all the service permissions; allows user to revoke, trust or report the permissions requested by other applications.
	Privacy Badger	Designed to protect user's privacy by letting user block 3rd party trackers.

What is common for all these tools is that they address components of a privacy-aware ecosystem, but miss out on innovation addressing the ecosystem as a whole. This has implications for the users' own control of privacy using perhaps one of the above tools and then missing out on other privacy threats.

It shall be mentioned that the above-mentioned privacy aware tools builds on a number of different security requirements and polities. A review of the existing privacy oriented approaches can be found in Kalloniatis et al. [12].

3. User Privacy in the Future

The future of internet technologies might strongly depend on their ability to handle issues of power and freedom concerning authority over personal information, transparency, identification and interpretation based on others' data collection [7]. As shown in the previous section, all the privacy-aware tools are addressing just a segment of a privacy-aware ecosystem, but not the ecosystem as a whole. Following segments are discussed in to the privacy aware tools in the section above:

- Visual representation of invisible trackers: As mentioned, there are multiple invisible trackers on many websites and applications user visits. Hidden requests from the 3rd parties make users' personal information vulnerable to privacy and security threats.
- Safe and Secure internet using Encryption: With the help of encryption, an encrypted tunnel can be established which can:
    - Protect users from wireless eavesdropping
    - Encrypts sensitive personal information, identity and passwords there are vulnerable to hidden trackers
    - Block malicious trackers, sources of malware and identity theft.
    - Anonymize user search queries and let the user browse privately through search engines of choice without being tracked.
- Automatic Permission Alerts: The user get alerts whenever any application tries to access personal and private information.
- Evaluation of privacy policies: Evaluating the privacy friendliness of the services by reviewing the privacy policies/terms of use.
- Economic value of user's data: Estimates user's data worth to the social networking sites.

In this paper a privacy aware ecosystem is suggested for future privacy services. An approach is needed that can balance the interests between service providers or Internet companies on the one hand and end user on the other end i.e., a Positive-sum model which creates a "win-win" scenario for the involved parties than an "either/or" involving unnecessary trade-offs. Some strategic objectives are led down to Empower Users' privacy for future services by deploying a new privacy ecosystem based on digital technologies that can innovatively:

- Educate the end users' to understand and control their personal information in online transactions by using visualization tools.

- Introduce the possibility of assigning value to personal data so that users may pay and service providers may charge for privacy at desired levels.
- Enable service providers and mobile application developers to engage in explaining their requests, treatments and protections for end users' personal data, ultimately increasing the general acceptability of the user-provider liaison.

To fulfill the above objectives, a Trusting authority is needed that can be located between the end user and the service providers as a distributed system. The concept is similar to User managed access (UMA) developed under the Kantara Initiative with the purpose to "develop a set of draft specifications that enable an individual to control the authorization of data sharing and service access made between online services on the individual's behalf [20]. Similar to Authorization server in UMA the Trusting authority will act as a front man for the user. The Trusting Authority (TA) approach is depicted in Figure 1.

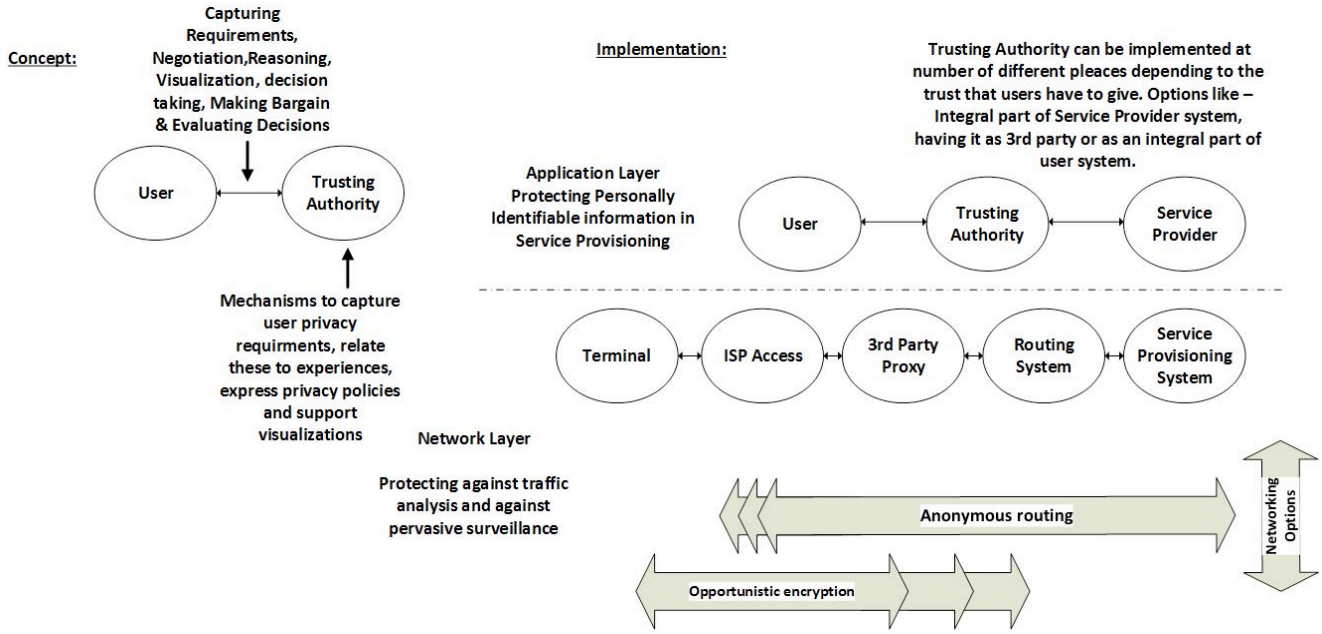

Figure 1 Overview of the Trusting Authority approach.

The left-hand side outlines the concept, while the right-hand side details its implementation. TA will perform following tasks:

- Collect Data : Service Provider policies, personal attributes requested, traffic / network data
- Record usage history data: Name of service provider, transaction performed, attributes requested, rating of the experience (after completion)
- Analyze and process data: Compliance with data protection laws and privacy-by design guidelines, app scanning.
- Calculation of trust level: Based on previous interactions with this service provider (if any)
- Classification of the service based on usefulness and trust: Assigning value to the personal data involved
- Policy handling: Proposing new policies for personal data disclosure (if needed) or identifying the proper policy from the existing repository of policies
- Visualizing the bargain for the user
- Setting the privacy level
- Bargaining: Optionally choose between different versions of the service to match the privacy level: Anonymous, pseudonymous, personalized or personalized and context-aware.

The Trusting Authority Approach will in that way deal with the privacy issues that online users of today try to handle themselves.

4. Conclusions

Privacy awareness is a key area for current and future service development. Numerous examples on privacy disclosure and the knowledge that third parties /data brokers everyday, at all times, requests, sell and use private user data puts the user privacy on a top priority for future services.

This paper has looked into one element of current privacy aware tools – visualization tools that can inform the users about trackers, third parties and other privacy violating elements. However, the existing privacy aware tools do address single elements of privacy violations. Using only one tool, the user may have a false sense of being private.

Therefore, the paper proposes a more coherent approach to privacy awareness in future services, namely an ecosystem of elements that all together shall provide necessary privacy for users. The elements of the ecosystem consist of a Trusting Authority that is located between the user and the service provider and can secure the user controlled privacy.

5. References

[1] Angulo, J., Fischer-Hübner, S., Wästlund, E., & Pulls, T. (2012). Towards usable privacy policy display and management. Information Management & Computer Security, 20(1), 4-17

[2] AVG PrivacyFix: http://www.avg.com/ww-en/privacyfix

[3] Cavoukina, A. and Chanliau, M. Privacy and Security by design: A convergence of paradigms. 2013.

[4] Cavoukian, A. and Chibba, M. A regulartor's perspective: Leading the way with Privacy by Design. Cyber security in future Internet, security and privacy by design. OUTLOOK, Visions and research for the wireless world, no 11. 2014.

[5] Cavoukian, A. Privacy by Design and the Emerging Personal Data Ecosystem. 2012

[6] Electronic Frontier Foundation/Privacy Badger: http//www.eff.org(privacybadger

[7] Fritsch, L. The clean Privacy Ecosystem of the future internet. Future Internet 2013, 5(1), 34-45

[8] Garfinkel, S. Security tools: Visualization is power. http:www.csoonline.com/article/2121183/data-protection/security….., 2007.

[9] Gosk, S. Stores may be tracking you through your cellphone. http://www.today.com/money/stores-may-be-tracking-you……, 2013.

[10] Guarda, P. and Zannone, N. Towards the development of privacy-aware systems. Information and Software Technology, vol. 51 (2009), pp. 337-350.

[11] Handbook on European Data Protection Law. 2014

[12] Kalloniatis,C. Kavakli, E., and Gritzalis, S. Methods for designing privacy aware Information Systems: A review. IEEE Computer Society, 13th Panhellenic Conference on Informatics, 2009.

[13] King, J., Lampinen, A. akellnd Simolen, A. Privacy: Is there an app for that? Symposium On Usable Privacy and Security (SOUPS) 2011, July 20-22, 2011, Pittsburgh, PA, USA.

[14] Mozilla Lightbeam: http://mozilla.org/en-US/lightbeam

[15] MyPermissions: http://mypermissions.org/

[16] PC: http://www.pcworld.com/article/2147306/f-secure-freedome-review-vpn-and-security-for-mobile-devices.html, 2014.

[17] PewResearchCenter. Privacy management on social media sites. http://pewinternet.org/Reports/2012/Privacy-management-on-social-media.aspx. 2012.

[18] Smith H.J., Milberg, S.J., and Burke, S.J. Information Privacy. Measuring individuals' concerns about organizational practices. MIS Quarterly, 1996, Vol. 20, no.2, ppp. 167-196.

[19] Terms of Service Didn't Read: https://tosdr.org/

[20] The Kantara Initiative: User Managed Access - Kantara Initiative. http://kantarainitiative.org/confluence/display/uma/Home

[21] Watson, P.J. Samsung "SmartTV" records "personal" conversations to third parties. http://www.declothesline.com/2014/11/03/samsung-smart-tv-record……

[22] Westin, A. Privacy and freedom, New Your: Atheneum. 1967