Privacy Threat Model in Lifelogging

Md Sadek Ferdous
University of Southampton
Southampton SO17 1BJ, UK
S.Ferdous@soton.ac.uk

Soumyadeb Chowdhury
10 Dover Drive, Singapore, 138683
Soum.Chowdhury@singaporetech.edu.sg

Joemon M Jose
University of Glasgow
Glasgow G12 8RZ, UK
Joemon.Jose@glasgow.ac.uk

Abstract
The lifelogging activity enables a user, the lifelogger, to passively capture multimodal records from a first-person perspective and ultimately create a visual diary encompassing every possible aspect of her life with unprecedented details. In recent years it has gained popularity among different groups of users. However, the possibility of ubiquitous presence of lifelogging devices especially in private spheres has raised serious concerns with respect to personal privacy. Different practitioners and active researchers in the field of lifelogging have analysed the issue of privacy in lifelogging and proposed different mitigation strategies. However, none of the existing works has considered a well-defined privacy threat model in the domain of lifelogging. Without a proper threat model, any analysis and discussion of privacy threats in lifelogging remains incomplete. In this paper we aim to fill in this gap by introducing a first-ever privacy threat model identifying several threats with respect to lifelogging. We believe that the introduced threat model will be an essential tool and will act as the basis for any further research within this domain.

Ubicomp/ISWC'16 Adjunct , September 12-16, 2016, Heidelberg, Germany
ACM 978-1-4503-4462-3/16/09.
http://dx.doi.org/10.1145/2968219.2968324

Author Keywords
Privacy; lifelogging; threats; threat modelling.

ACM Classification Keywords
K.4.1. Public Policy Issues: Privacy; K.4.2. Social Issues

Introduction
The right to privacy is one of the fundamental human rights in any modern society. It advocates and facilitates mechanisms to uphold the privacy of all individuals within the society. However, what is private is highly debated. This is because privacy has social, legal, psychological, political and technical connotations [1]. Even more, privacy is of dynamic nature. What is considered private in a society can change considerably with time. Many of these changes are driven by technological advancements.

Lifelog definitions:

By Dodge *et al.*: A lifelog is defined "*as a form of pervasive computing consisting of a unified digital record of the totality of an individual's experiences, captured multi-modally through digital sensors and stored permanently as a personal multimedia archive*" [3].

By Gurrin *et al.*: A lifelog has been defined as "*a form of pervasive computing which utilises software and sensors to generate a permanent, private and unified multimedia record of the totality of an individual's life experience and makes it available in a secure and pervasive manner*".

One such technology is lifelogging that has further-reaching privacy implications than any existing technologies. The lifelogging enables a user, the lifelogger, to passively capture multimodal records (e.g. images, videos, etc.) from a first-person perspective and ultimately create a visual diary encompassing every possible aspect of her life with unprecedented details. The possibility of ubiquitous presence as well as discreet size of many lifelogging devices especially in private sphere has raised serious concerns with respect to personal privacy.

Understandably there is an urge to understand the privacy implications, in the form of privacy threats, of such a ubiquitous technology. Different practitioners and researchers have explored different ways to propose, design and develop different frameworks to mitigate the identified threats. However, none of the existing works has considered a threat model of privacy in lifelogging. Without such a model, it is difficult to understand, identify, assess and address the risk of privacy threats comprehensively.

In this paper, we aim to fill in this gap by introducing a first-ever privacy threat model in lifelogging using a systematic approach. At first, we present a brief analysis on different aspects of lifelogging. Next, we examine different aspects of privacy and its associated dimensions in order to formulate a definition of privacy with respect to lifelogging. Then, we present our threat model along with the identified assets and their associated threats. Finally, the inter-relation between the identified threats and other aspects of lifelogging is presented.

Lifelogging

In general, lifelogging is a solipsistic activity that utilises pervasive computing technologies to capture the first-person view of the daily activities of a user in an automatic and continuous fashion. The main motivation for any user to engage in lifelogging is to create a digital representation of her daily experience that can be stored in a preferred storage medium for future recollecting, reminiscing, retrieving, reflecting, and remembering intentions [2] and/or for other purposes. To better understand and study the privacy implications in lifelogging, at first, we need to define the notion of lifelogging and study its different aspects.

Lifelogging is the process of creating a lifelog. There are several definitions in the existing literature. A couple of such definitions are provided in the sidebar. Between these definitions, we prefer the definition by Gurrin *et al*. as it is more expressive. Even so, this definition has shortcomings. For example, a lifelog has been defined as a *permanent* and private *multimedia* record. We argue that providing a user with the capability to store massive amount of digital data at an ever-decreasing cost not necessarily guarantees the permanent storage of such data, since this completely depends on the reliability of the storage medium as well as the willingness of a lifelogger. Similarly, we argue that many users would be willing to share their lifelogs in social networks if there is a technical capability to do so. Hence, a lifelog is not completely private. To rectify the stated shortcomings, we propose a revised definition, based on the definition of Gurrin *et al*., where our revised connotations are highlighted in bold.

Definition 1. *A lifelog is a form of pervasive computing which utilises software and sensors to generate a*

(potentially) *permanent, private* ***yet (potentially) shareable*** *and unified multimedia record of the totality of an individual's life experience and makes it available in a secure,* ***privacy-friendly*** *and pervasive manner.*

Actors in lifelogging are the involved entities during lifelogging. Gurrin *et al.* identified four different actors [1] presented in the sidebar.

Privacy in Lifelogging

What is the most appropriate definition of *Privacy* is highly debated. This is because privacy has social, legal, psychological, political and technical connotations [1]. A complex entanglement of these connotations dictates what can be considered private in a society. Interestingly, what is considered private in a period of time may not be considered as a private in another period. The involvement of such different perspectives and their highly volatile dynamic nature make it harder to define a one-size-fits-all definition of privacy. Hence, there exist a number of definitions of privacy from different perspectives and from different time periods. Next, we explore a few influential definitions of privacy and analyse their relevance and suitability in terms of lifelogging.

Motivated with the availability and popularity of modern photography and printing press and their implications on the privacy of people, Samuel Warren and Louis Brandeis wrote the seminal, influential paper *The Right to Privacy* in which they defined privacy as: "the right to be alone" [4]. It is thought to be the first definition of privacy [1] and devised with the motivation to protect people from nosy reporters who would take their photographs without their consent [5]. Unfortunately, this definition has lost its effectiveness in the modern day society where taking photographs of other people in public places is no longer considered a breach of privacy of those people, legally as well as socially. This notion of privacy is all about capturing the one's right to be in solitude and to protect the person from intrusion in a physical domain. Hence, it is viewed as the privacy of personal sphere [6].

With the popularities of computers and computing systems and the possibilities of storing large amount of personal data into these systems and the capability of advanced data processing mechanisms, a new notion of privacy, called *Information Privacy*, in the domain of technology started to gain attention from 1960s onward. In this regard, one of the most influential definitions of privacy was given by Alan Westin in [7] where privacy was defined as: "*the right to select what personal information about me is known to what people*".

Next, we explore how the concept of privacy is applied in the physical world using the concept of privacy dimensions. Privacy dimensions denote the different modes of privacy. Based on the four modes (*Solitude, Intimacy, Anonymity and Reserve*) of privacy introduced by Westin in [7], Pedersen conducted an empirical study and identified six dimensions of privacy in the social setting [8]. The dimensions are presented in the sidebar of the next page.

These six modes altogether define different aspects of privacy of a person in the social setting. Ensuring the gratification of these aspects can enable the right for a person to be private according to her needs. This is facilitated by social norms and legal practices. These social norms and practices draw a line, often imaginary,

Lifelogging actors:

The Lifelogger. A lifelogger is the entity which utilises a lifelogging device to capture and store lifelogs. Here, we assume that a lifelogger is a person.

The Bystander. A bystander is any person who is captured (intentionally, incidentally or accidentally) in a lifelog of another person (lifelogger) without engaging in interactions with the lifelogger. Examples of bystanders are strangers in an environment, family members, friends, colleagues, etc.

The Subject. A subject is any person who is captured (intentionally or incidentally) in a lifelog of the lifelogger during their interaction.

The Host. A host is the entity who bears the responsibility of storing a lifelog of a lifelogger.

Privacy dimensions:

Reserve. This represents the unwillingness of a person to be with and to interact with others, especially strangers.

Isolation. This represents the desire of a person "*to be alone and away from others*".

Solitude. This represents the state of a person when she is "*alone by oneself and free from observations by others*".

Intimacy with Family. This represents the state of a person being alone with members of the person's family.

Intimacy with Friends. This represents the state of a person being alone with her friends.

Anonymity. This represents the expectation of a person not being recognised or to remain unnoticed in a crowd and hence "*not wishing to be the centre of group attention*".

between what is private and what is public. However, advocates of personal privacy have witnessed a tension or even a threat to this imaginary line with the advent of modern technologies allowing devices, especially cameras, camcorders, mobile phones and tablets, to blur the distinction between what is private and public.

Defining Privacy in Lifelogging

To define privacy with respect to lifelogging, we, at first, analyse the only one definition in the existing literature presented by Gurrin *et al.* [1] in which privacy in lifelogging has been defined as the "*the right to choose the composition and the usage of your lifelog and the right to choose what happens to your representation in the lifelogs of others*".

This is a simple literal definition that captures the notion of user empowerment (especially data control) by enabling the lifelogger to capture a lifelog and the other actors (the subject and bystander) with the right to dictate what to do with the lifelogs in which they appear. However, this definition fails to capture other privacy dimensions. Based on this definition, we have formulated an elaborate definition of privacy in lifelogging that captures all dimensions of privacy in a physical world. The definition is presented below.

Definition 2. *Privacy of captured lifelogs in an information system is the right to exercise anonymity when desired by any involved actors (lifeloggers, bystanders and subjects) as well as to empower each respective actor with the required capability to exert privacy considering all (appropriate) dimensions while the lifelogs are stored in a storage medium, processed in a system, visualised in an interface and (optionally) shared among different users.*

Threat Modelling in Lifelogging

Threat modelling is an integrated process of designing and developing a secure and privacy-friendly system. A well-defined threat model helps to identify security and privacy threats on different assets of a system. In essence, a threat modelling consists of the following steps [9, 10, 11]:

- listing assets of the system and
- identifying possible security and privacy threats on those assets.

Each single step of the threat modelling process is described in the following sub-sections.

Listing assets

An asset is the abstract or physical resource in an information system that needs to be protected from an adversary (attacker) [9]. It is the resource for which a threat exists and represents the target of the adversary in the system. The motivation behind this step is to highlight such assets in the system. The corresponding assets of a lifelogging system is presented in the sidebar of the next page.

Identifying Threats

A threat represents the activity or capability of an adversary onto an asset of a system with an intention to invade the security of the system or invade the privacy of a user in the system [9]. The main motivation behind this step is to identity possible threats on different assets of the system. Based on the threat modelling process presented in [11], we identify the following threats:

Identified assets:

Lifelogs. In a lifelogging system, the lifelogs are the core assets since the main purpose of such a system is to deal with captured lifelogs.

Identity of a user. The identity of a user is defined as a representation of the user in a specific application domain [12]. Since lifelogs can be used to identify users, the identity of a user is also a crucial asset.

Information embedded within a lifelog. Meta-information (e.g. gps coordinates) within each lifelog represents a valuable asset as it can abused to infer unforeseen knowledge about a user.

Access control mechanisms. The deployed access control mechanisms determine which lifelog is exposed to which user(s) and thus, can be regarded as a crucial asset.

T1: Unnoticed Capture. A lifelogging device can be quite discreet in nature. This will allow a lifelogger to carry on the unnoticed capture of lifelogs in which other actors may appear even without their knowledge and/or consent.

T2: Unaware Identification and unforeseen inference. An attacker can identify a person using a lifelog and ever-powerful image search online services. Combining the identity of the person with other meta-information (especially gps coordinates) embedded inside a lifelog the attacker can create a profile of the person without theirr knowledge. Such profile can be used to create inference for future occurrence of events which otherwise were not possible.

T3: Lack of control. Many lifelogging devices (e.g. Narrative Clip) require the lifelogger to upload data in a cloud server maintained by the manufacturer even before the lifelogs are accessible to the lifelogger. Once lifelogs are uploaded to the server, the lifelogger has limited control over them and may not be aware how such lifelogs are being abused by the manufacturer. The very nature of the lifelogging makes it very difficult for other actors such as subjects or bystanders to express their consent explicitly while lifelogs are captured, stored and analysed n a system.

T4: Inaccessibility of lifelogs. Lifelogs are inaccessible to subjects and bystanders until they are shared by the lifelogger.

T5: Determining sensitivity. Sensitivity in a lifelog will determine if the lifelog can be considered private. For example, a lifelog captured in a private and/or intimate setting can be considered as of highly sensitive. Technically, lifelogs are created in large volume. For example, modern lifelogging devices such Narrative Clip allow capturing nearly 3000 lifelogs each day. This sheer volume of lifelogs makes it extremely difficult even for the lifelogger to pinpoint each sensitive lifelog.

T6: Security. There is a strong inter-relationship between security and privacy. In many ways, different security measures safeguard the privacy of users in an information system. The threats related to security are presented below.

- **T6.1: Secure storage.** The information system should take great care to securely store the captured lifelogs so that attackers cannot access them inappropriately.
- **T6.2: Confidentiality and integrity.** An attacker can intercept shared lifelogs while being transmitted, allowing the attacker to get hold of lifelogs in an unauthorised fashion and may alter a lifelog before they are transmitted to the destination system.
- **T6.3: Unauthorised disclosure.** Lifelogs are disclosed to another unauthenticated and/or unauthorised user allowing the second user to get hold of such lifelogs inappropriately.

Some threats are applicable to all actors whereas others apply to a specific actor. For example, the dimension of *Solitude* is not applicable to a subject and a bystander since, otherwise, they would not appear in a lifelog. Similarly, the dimensions of *Reserve* and *Isolation* are not applicable to a subject. By combining these two arguments, we summarise which threats are

	Lifelogger					
	R	I	S	FM	FR	A
T-1	-	-	-	-	-	-
T-2	-	-	-	-	-	-
T-3	√	√	√	√	√	√
T-4	-	-	-	-	-	-
T-5	√	√	√	√	√	√
T6.1	√	√	√	√	√	√
T6.2	√	√	√	√	√	√
T6.3	√	√	√	√	√	√

Table 1: Threats for lifeloggers

	Bystander					
	R	I	S	FM	FR	A
T-1	√	√	-	√	√	√
T-2	√	√	-	√	√	√
T-3	√	√	-	√	√	√
T-4	√	√	-	√	√	√
T-5	√	√	-	√	√	√
T6.1	-	-	-	-	-	-
T6.2	√	√	-	√	√	√
T6.3	√	√	-	√	√	√

Table 2: Threats for bystanders

	Subject					
	R	I	S	FM	FR	A
T-1	-	-	-	√	√	√
T-2	-	-	-	√	√	√
T-3	-	-	-	√	√	√
T-4	-	-	-	√	√	√
T-5	-	-	-	√	√	√
T6.1	-	-	-	-	-	-
T6.2	-	-	-	√	√	√
T6.3	-	-	-	√	√	√

Table 3: Threats for subjects

applicable within which dimensions for which actor in Table 1, Table 2 and Table 3. In these tables, *R* represents the *Reserve* dimension whereas *I*, *S*, *FM*, *FR* and *A* represent *Intimacy*, *Solitude*, *Intimacy with family*, *Intimacy with friends* and *Anonymity* respectively. To indicate a threat is applicable to a particular dimension with respect to an actor, "√" symbol has been used whereas "-" indicates the particular threat does not apply for the corresponding dimension with respect to an actor.

Conclusions

The motivation of this paper is to identify potential privacy threats and then analyse their implications on different aspects of lifelogging. Being a nascent technology, it is still not clear how the lifelogging technology will be shaped and what privacy implications it will expose in future. One thing is certain that there will be many more interesting use-cases of lifelogging apart from being a tool of personal recollection and ramification. As such, it has the potential to gain mainstream traction just like photography. To realise this potentiality, privacy threats need to be identified and addressed. This paper aims to meet these goals and lay out the foundation for subsequent research to design and develop a privacy-preserving lifelogging system.

References

1. Gurrin, C., Albatal, R., Joho, H., & Ishii, K. (2014). A privacy by design approach to lifelogging. Digital Enlightenment Yearbook 2014: Social Networks and Social Machines, Surveillance and Empowerment, 49.
2. Sellen, A. J., & Whittaker, S. (2010). Beyond total capture: a constructive critique of lifelogging. Communications of the ACM, 53(5), 70-77.
3. Dodge, M., & Kitchin, R. (2007). 'Outlines of a world coming into existence': pervasive computing and the ethics of forgetting. Environment and planning B: planning and design, 34(3), 431-445.
4. Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. Harvard law review, 193-220.
5. Langheinrich, M. (2001, September). Privacy by design—principles of privacy-aware ubiquitous systems. In Ubicomp 2001: Ubiquitous Computing (pp. 273-291). Springer Berlin Heidelberg.
6. Pötzsch, S. (2008). Privacy awareness: A means to solve the privacy paradox?. In The future of identity in the information society (pp. 226-236). Springer Berlin Heidelberg.
7. Westin, A. F. (1968). Privacy and freedom. Washington and Lee Law Review, 25(1), 166.
8. Pedersen, D. M. (1979). Dimensions of privacy. Perceptual and Motor Skills, 48(3c), 1291-1297.
9. Myagmar, S., Lee, A. J., & Yurcik, W. (2005, August). Threat modeling as a basis for security requirements. In Symposium on requirements engineering for information security (SREIS) (Vol. 2005, pp. 1-8).
10. De Cock, D., Wouters, K., Schellekens, D., Singelee, D., & Preneel, B. (2005). Threat modelling for security tokens in web applications. In Communications and Multimedia Security (pp. 183-193). Springer US.
11. Desmet, L., Jacobs, B., Piessens, F., & Joosen, W. (2005). Threat modelling for web services based web applications. In Communications and multimedia security (pp. 131-144). Springer US.
12. Ferdous, M. S., Norman, G., & Poet, R. (2014, September). Mathematical Modelling of Identity, Identity Management and Other Related Topics. In Proceedings of the 7th International Conference on Security of Information and Networks (p. 9-16). ACM.