Design strategies for Data Sensitive Applications safeguarding User's Privacy through Privacy Regulations

P. Ram Mohan Rao[1,4], S Murali Krishna[2], A P Siva Kumar[3]

[1] Research Scholar , Department of Computer Science and Engineering, JNTUA Anantapuramu,
[2] Professor, Department of Computer Science and Engineering, Sri Venkateswara College of Engineering Tirupathi, Andhra Pradesh, India,
[3] Assistant Professor, Department of Computer Science and Engineering, JNTUA Anantapuramu, Andhra Pradesh India,
[4] Associate Professor, Department of Computer Science and Engineering, Sphoorthy Engineering College, Hyderabad.

Abstract: Rapid technological advancements and digitization have brought new challenges for the protection of personal data. Private and Public sector firms like banks, insurance, hospitals, e-commerce, mobile app vendors, social media etc. collect a lot of personal and transactional data from the users and data analytics is applied to gain more real time insights about the data, to offer value added services and facilitate efficient decision making. However privacy concerns limit the usage of data analytics and privacy hazards continue to increase, which even led to creation of data privacy regulations by many countries including, The Personal Data Protection bill, 2018 of India and General Data Protection Regulation, 2016 of European Union Law. Some of the key privacy threats include Digital Profiling, Cyber Stalking, Search Engines Surveillance, Recommendation Systems leading to disclosure of sensitive data and sharing of data without consent. In order to overcome these privacy challenges by incorporating privacy regulations, we have designed guidelines for application development, incorporating key features of privacy regulations along with the implementation strategies which will help in developing data sensitive applications which can offer strong and coherent privacy protection of personal data.

Keywords: Privacy, Privacy legislations, Digital applications, Design Guidelines, Privacy Protection

1 Introduction

Smart Phones, Internet and Communication Technologies (ICT), e-commerce, social media has transformed our day to day life. Even though digital applications ease our work they are prone to privacy vulnerabilities. The key privacy threats include surveillance, disclosure and discrimination. The consequences of the privacy threats are activity monitoring and targeted advertisements [1], identity theft, information disclosure without consent, personal abuse through cyber stalking [2], studying

emotions and mood of the people by accessing profile pictures, tweets, likes and comments to find emotionally weak, personally deserted people and trap them using various cyber attacks like ransom ware, sexual abuse etc. [3]. Extensive usage of smart phones, social media and other digital applications have definitely increased the privacy concerns. In this section we will illustrate the modern privacy threats in detail. Section 2 describes the key features of various privacy regulations which are in force across different countries. In section 3 we presented our research work, which facilitate implementation of data sensitive applications to incorporate key features of privacy regulations. In section 4 we have highlighted the benefits of applying our implementation strategies in comparison with the existing implementations and in Section 5 we presented the future scope in the design of data sensitive applications with at most emphasis on privacy protection.

1.1 Modern Privacy Threats

i. Digital Auto Profiling: Digital Auto Profiling is the process of analyzing the transactional data of a person and classify him/her into some class. For example, non banking firms will calculate our Cibil score based on our credit card transactions and decide whether to offer or deny a loan. This is a serious privacy violation where third party firms are able to access our transactional data. Digital profiling also influences group privacy where in a person may be a member of one or more groups [4]. Digital online advertising agencies use cookies to track user behavior across multiple web sites to build a profile of the user and target more appropriate advertisements. In this context, Google has announced to end its support for third party cookies in chrome browser which will make it further more difficult for the digital marketing companies to build user profile. This move of Google was amid privacy legislations and to enhance user privacy [5]. Microsoft Edge has also announced tracking prevention to be enabled by default in its latest version. Article 22 of GDPR facilitates right to the individual that, no automated data processing including profiling is allowed without consent from the user and it is one of the important reason behind prominent browsers like Google Chrome, Microsoft Edge to withdraw support to third party cookies and prevent Digital Profiling.

ii. Social media privacy issues: Social media platforms are the most vulnerable places where tons of personal data is uploaded by millions of users every hour. Social media platforms analyze data to recommend groups, suggest friends etc. However these social media platforms are highly prone to stalking attacks. Stalking attack is an act of exploiting, social media vulnerabilities to harass an individual or organization and the common form of stalking involve online mob of anonymous people self-organized to target individuals with defamation, sexual abuse, threats of violence and technology-based attacks. These attacks include publishing lies and fake photographs, threats of violence and rape, posting personal, abusive and sensitive information, e-mailing damaging statements about victims to their employers. Social media is used to build trust between the perpetrator and the victim. When victim transmits confidential data including pictures and videos, the perpetrator abuses them for blackmail purposes [6]. Most of the times, the victims are the youngsters who are

emotionally disturbed, socially isolated or depressed. The social well being of a person can be found by analyzing the kind of posts and tweets made by the person in social media [7].

2 Privacy Regulations

With growing number of privacy threats and grave consequences, awareness among users have increased and in turn increased the demand for privacy protection which eventually led to creation of privacy laws and legislations in many countries. The list of privacy legislations issued across ten countries is described in Table 1. Most of the privacy legislations are based on the European Union's General Data Protection Regulation (GDPR).

Table 1: Privacy Regulations across ten countries.

S no.	Country	Privacy Legislation
1	European Union	General Data Protection Regulation (GDPR)
2	India	The Personal Data Protection Bill
3	Canada	Personal Information Protection and Electronic Data Act
4	France	The Data Protection Act
5	Germany	The Federal Data Protection Act
6	Australia	Privacy Principles
7	Japan	The Personal Information Protection Act
8	Malaysia	Personal Data Protection Act
9	Mexico	Federal Law for the Protection of Personal Data
10	Philippines	Republic Act no. 10173

More than hundred countries have passed legislations to protect individual privacy. However Table 1 contains the list of only ten countries where privacy legislations are in force. The key features of all privacy legislations and privacy rights are described below.

1. **Right to forget or erase data:** Personal data gets uploaded in many digital applications. For example, people upload certain private photos and videos, buy certain products online and if people want, the data or transactional records can be removed from their databases.
2. **Users consent before sharing the data:** Data holders share and exchange data for real time insights, but in many applications the data owner is not aware of it. It is required to take the consent of the data owner before sharing.
3. **No surveillance without consent:** Many applications will monitor their user's behaviour including location, device type etc. Data profiling companies and digital advertisement companies do surveillance without consent from the user

and most of the users are not even aware of surveillance. It is now mandatory for all firms to take user consent for surveillance, in the countries where privacy legislations are in force.

4. **Right to restrict the data processing:** Many data intensive applications, process data without prior consent of the data owner. It is mandatory to take prior permission from the data owner to use data for further processing.

3 Related Work

The important aspect of application design is to ensure the application is inherently private by design, which is a common feature found in all privacy legislations. Some of the applications which are prone to privacy threats are listed in Table 2. For each such application we have suggested design strategies to ensure privacy of an individual.

Table 2 Privacy vulnerabilities in various application types.

S no.	Application Type	Privacy Risk involved
1	Smart phone apps.	Information theft, Intrusion
2	e-Commerce sites	Inference attacks, Disclosure
3	Social media	Cyber stalking, Ransom ware
4	Data capturing systems like banking, hospitals, insurance, government portals etc.	Disclosure, Discrimination

3.1 Design guidelines

As mentioned in Table 2, all the applications where personal data is captured are vulnerable to privacy risks and also they are not in line with the privacy legislations. Design guidelines for each application are provided in this section.

3.1.1 Design guidelines for Smart Phone Apps

It is a common practice that most of the users do not read the privacy policy and the network permissions which an app demands before installation. People ignore and will agree for all permissions the app demands which may lead to serious privacy concerns. In a recent study it is observed that there is growing concern among smart phone users with respect to privacy hazards and also the apps access the data which is not needed for them to function [8]. To ensure inherent privacy protection, smart phone apps must be designed with following features.

a. Seek only the minimum permissions for the app to be functional.
b. Do not collect any metadata including location, type of device, time etc.

c. No auto profiling of the user by any app is allowed.
d. Accept and abide the federal laws of the region or state pertaining to data access and sharing.
e. Design to ensure no access to any free Wi-Fi which is not registered by the user.
f. Do not transfer any data from the phone without consent from the user.
g. Privacy policy should not be a text document. Privacy policy should be an audio file played in the language opted by the user, ensure the user listens it completely and finally accepts or rejects the privacy policy. polling can be used to ensure user's attention. i18n (internationalization) applications are required and easy to develop with present open source technology frameworks to offer privacy policy as an audio file in the language opted by the user.
h. Handling of EXIF data: With every picture taken in a smart phone, a special meta data called EXIF (Exchangeable Image File Format) data is associated, which contains information like type of device, date and time of picture taken, location etc. The EXIF data will also be uploaded along with the picture whenever user uploads or shares them. This is also a serious privacy concern because it may disclose confidential and private data like location, date and time, device used etc. Typical EXIF data for an image file is shown in Figure 1.

Figure 1: Typical EXIF data for an image file

Most of the smart phone users are not even aware of the EXIF data and tend to upload tons of images in to the public domain. EXIF data discloses important information which can be private and sensitive for a user. It is very important to note that, except BITMAP format all other image formats carry EXIF data. Hence it is recommended to convert images into BITMAP format before uploading or sharing, which is tedious and most of the users are unaware of the existence of EXIF data itself. Hence it is recommended that the mobile operating systems must provide an inherent feature of converting every image into BITMAP format before it is being uploaded or shared. By this we can avoid disclosure of meta data associated with every image or picture stored in smart phones.

i. Finally the companies hosting the apps must ensure registration of a data protection officer before allowing the any app into the store.

3.1.2 Design guidelines for e-Commerce sites

All e-commerce sites use recommendations to offer value added services to the customers. Recommendations are used as part of improved service, however there is always a possibility of information disclosure. For example, A person wanted to buy some product for personal use. He/she wanted this to be confidential and by virtue of recommendations, he/she may see a pop up or alert showing a better offer on that product which is visible to the people sitting nearby and this will lead to discrimination and personal embarrassment. Based on the type of products bought the gender of the person can also be inferred which is an unwanted disclosure. In order to ensure privacy protection, following features need to be incorporated in the design of the e-commerce sites in line with the privacy legislations.

1. Privacy Quotient (P_μ) : Recommendations are used by ecommerce firms to provide value added services and best possible offers to the customers based on their buying habits and transaction history. Recommendation systems lead to serious privacy concern which are not addressed by any ecommerce firm and the same is illustrated here. For example, A person regularly bought some product online, related to personal care and does not want to disclose this to anyone. However since it is a regular transaction the ecommerce firm would like to recommend the same product to him by offering decent discount on the product and the same will displayed on his screen when he/she logs into their account and it is a privacy breach if someone else see the same. It can lead to discrimination of the person in the family or profession. To address this problem, we introduce the concept of privacy quotient. For every product the ecommerce firm should provide an option where in user can opt, whether this product and purchase is to be made private or not, thereby excluding it from any form of analytics or recommendations. If 40% buyers of a product opt for transaction privacy i.e. the product purchase is not to be used for recommendations, then the product must be considered as private and for all buyers of this product, the transaction must be made private. This percentage of transactions which decide the transaction privacy is called as privacy quotient (P_μ). Sample user interface designs for the ecommerce site is described in figure 2 and figure 3.

Figure 2: User has opted for personal care products

Figure 3: User consent for transaction privacy is taken

2. Most of the users of e-commerce sites are not even well educated and may not be aware of the privacy threats involved in e-commerce applications. In such cases the e-commerce sites must define a sensitivity threshold for every product. Privacy Quotient (P_μ) can be used as sensitivity threshold to transaction privacy.
3. No sharing of data without users consent: No e-commerce site, must share customers data without consent. However data can be shared with federal authorities for any investigation purpose [9].
4. Meta data: E-Commerce sites tend to collect meta data including location, type of device used, IP address etc. without the permission and knowledge of the user. It has to be avoided.

3.1.3 Social media platform design issues

Social media has emerged as the most vulnerable platform of privacy abuse especially cyber stalking, ransom ware, sexual abuse etc. Important issues to be addressed in social media applications are

1. Identification of fake accounts and stringent mechanism of anomaly detection.
2. Deep neural networks can be used in indentifying the private and sensitive information in the images uploaded by the user, remove them and store the modified image. User consent is mandatory. Users must be advised of privacy threats every time when they upload photos or videos.

3.1.4 Data Capturing Systems

Disclosure and discrimination are the common threats related to data capturing systems. Organizations like hospitals, banks, retail supply chain etc collect a lot of person specific data while offering respective services. This data will be analyzed to gain deep insights and come up with better decisions and offer value added services. However, privacy threats limit the applications of data analytics especially discrimination through credit scoring, marketing, employment etc. In order ensure privacy protection and be in line with privacy regulations, following important features to be embedded in the data capturing systems as inherent components.

1. As per the privacy regulations across many countries, it is recommended to use non anonymized and model based solutions for privacy preservation.
2. Sensitive attributes must be tokenized before sharing with any other third party for analytics.
3. Quasi identifiers must be synthesized before sharing. [10]

4 Results and Discussions

As part of our research we identified potential data sensitive applications which may cause privacy threats. The key vulnerabilities and consequences were also identified and categorized into four types. For each category of applications it is important to embed few critical changes in the design to ensure privacy protection and also abide privacy laws. In this regard, we have enumerated the standard guidelines for each category to ensure privacy as a inherent feature in each application. However, it is the responsibility of the firms to reinvent their application designs to ensure privacy protection. The awareness among the users regarding privacy threats and consequences is growing especially among women and youngsters using social media applications. In this digital rich data intensive world the usage of digital applications is experiencing an unprecedented growth which in itself is a grave challenge for individual privacy. The guidelines given in this paper will definitely improve the privacy preservation if employed by all the application developers whose deal with person specific private and sensitive data.

5 Conclusion

Privacy threats are continuously growing and privacy awareness among users have also increased which eventually led to creation of privacy legislations across many countries. As part of our work we proposed few guidelines for application design which will support individual privacy in many data intensive applications in line with many privacy legislations. These days more number of privacy violations and abuse is being reported in social media where people upload lot of personal photos and videos. Huge number of fake profiles were also reported who may indulge in activities like cyber stalking, ransom ware etc. There is a need for strong and coherent privacy preservation mechanism for social media applications and has enough scope for research especially employing deep learning models.

References

1. Ducange, Pietro, Riccardo Pecori, Paolo Mezzina, “A glimpse on big data analytics in the framework of marketing strategies”, *Soft Computing* 22.1 (2018): 325-342.
2. Dhillon, Gurpreet, Kane J. Smith, "Defining objectives for preventing Cyber Stalking", *Journal of Business Ethics* 157.1 (2019): 137-158.
3. Yan, Song Y, "Offensive cryptography", *Cybercryptography: Applicable Cryptography for Cyberspace Security*, Springer, Cham, 2019, 413-429.
4. Mavriki, Paola, Maria Karyda, "Automated data-driven profiling: threats for group privacy", *Information & Computer Security* (2019).
5. https://www.cnbc.com/2020/01/14/google-chrome-to-end-support-for-third-party-cookies-within-two-years.html
6. Huber, Edith, and Roman H. Brandtweiner, "Cyberstalking: The New Threat on the Internet", *Encyclopedia of Criminal Activities and the Deep Web*, IGI Global, 2020, 628-639.
7. Chen, Lushi et al, "Building a profile of subjective well-being for social media users", *PloS one* 12.11 (2017).
8. Furini, Marco et al, "Privacy Perception when Using Smartphone Applications", *Mobile Networks and Applications* (2020): 1-7.
9. Abdul-Ghani, Eathar, "Consumers’ Online Institutional Privacy Literacy", *Advances in Digital Marketing and eCommerce*, Springer, Cham, 2020. 40-46.
10. Walters, Austin et al, "Systems and methods for synthetic data generation for time-series data using data segments", U.S. Patent Application No. 16/405, 989.