UNIVERSITY OF PIRAEUS

SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

DEPARTMENT OF INFORMATICS

PHD THESIS

Thesis Title:	**(English)** **Personalized Health Services in IoT and Privacy** **(Greek)** **Προσωποποιημένες Υπηρεσίες Υγείας στο Διαδίκτυο των Πραγμάτων και Ιδιωτικότητα**
Student's name-surname:	**ACHILLEAS PAPAGEORGIOU**
Father's name:	**KONSTANTINOS**
Student's ID No:	**ΠΛΔ/1407**
Supervisor:	**Constantinos Patsakis, Associate Professor**

July 2021

PhD Thesis
was prepared during the Programme of Doctoral Studies
of the Department of Informatics
of the School of Information and Communication Technologies
of the University of Piraeus
for the degree of Doctor of Philosophy

3-Member Counseling Committee

Constantinos Patsakis
Associate Professor
University of Piraeus
School of Information and Communication Technologies
Department of Informatics
(Supervisor)

Georgios Tsihrintzis
Professor
University of Piraeus
School of Information and Communication Technologies
Department of Informatics
(Member)

Agusti Solanas
Associate Professor
Rovira i Virgili University
Department of Computer Engineering and Mathematics
(Member)

PhD Thesis
was presented before the 7-Member Examination Committee and approved on July 30, 2021

7-Member Examination Committee

Constantinos Patsakis
Associate Professor
University of Piraeus
School of Information and Communication Technologies
Department of Informatics

Georgios Tsihrintzis
Professor
University of Piraeus
School of Information and Communication Technologies
Department of Informatics

Agusti Solanas
Associate Professor
Rovira i Virgili University
Department of Computer Engineering and Mathematics

Efthimios Alepis
Associate Professor
University of Piraeus
School of Information and Communication Technologies
Department of Informatics

Nikolaos Kolokotronis
Associate Professor
University of Peloponnese
School of Economics and Technology
Department of Informatics and Telecommunications

Evangelos Sakkopoulos
Assistant Professor
University of Piraeus
School of Information and Communication Technologies
Department of Informatics

Edgar Galvan Lopez
Assistant Professor
Maynooth University
Faculty of Science and Engineering
Department of Computer Science

This thesis is dedicated to
my family and my beloved Angeliki
for their continuous support and motivation.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Patsakis Constantinos for the continuous support of my Ph.D. study. I am deeply grateful for his endless patience and motivation and direction throughout this academic path.

Besides my advisor, I would like to thank the rest of my supervising thesis committee, Prof. Georgios Tsihrintzis, and Prof. Agusti Solanas for their insightful comments and encouragement.

I thank my friends at the University of Piraeus, Dr. Athanassios Zigomitros, Dr. Kleanthis Dellios and Alkaios Anagnostopoulos, M.Sc. for their motivation and advice.

I would also like to thank Dorothea Kalogianni, Ph.D. Candidate in Architecture, at the University of Edinburgh for her willingness to contribute with reviews and support whenever requested.

Last but not the least, I would like to thank my family for supporting me spiritually throughout writing this thesis and my life in general.

Abstract

The secure and privacy-oriented introduction of personalized health services in the Internet of Things (IoT) is one of the most emerging topics globally. Currently, a field that is offered for great investigation is the establishment of trust, confidentiality, accuracy, integrity, and interoperability within the multidisciplinary domains and technologies involved in the delivery of personalized health services. The thesis examines and evaluates several heterogeneous applications, towards the successful provision of a higher level of privacy and security for the applications' end-users within ubiquitous and smart environments.

Keywords: Personalized Health, Ubiquitous Computing, Internet of Things, Mobile Application Privacy

Abstract

Ένα από τα κυρίαρχα θέματα παγκοσμίως αποτελεί η εισαγωγή ασφαλών και με έμφαση στην ιδιωτικότητα, υπηρεσιών υγείας στο αναδυόμενο περιβάλλον του Διαδικτύου των Πραγμάτων. Σήμερα ένα εξαιρετικού ενδιαφέροντος πεδίο έρευνας, που διατίθεται για περαιτέρω διερεύνηση, είναι η ενίσχυση της εμπιστοσύνης, της εμπιστευτικότητας, της ακρίβειας, της ακεραιότητας, αλλά και της διαλειτουργικότητας μεταξύ των διαφορετικών περιοχών εφαρμογής και των τεχνολογιών οι οποίες εμπλέκονται στην παροχή προσωποποιημένων υπηρεσιών υγείας. Η εργασία εξετάζει και αξιολογεί διάφορες ετερογενείς εφαρμογές στην κατεύθυνση της επιτυχούς προσφοράς ασφαλών και ιδιωτικών υπηρεσιών σε τελικούς χρήστες εφαρμογών που εκμεταλλεύονται τη διάχυτη υπολογιστική και τα χαρακτηριστικά των έξυπνων περιβαλλόντων.

Λεξεις κλειδια: Προσωποποιημένη Υγεία, Διάχυτη Υπολογιστική, Διαδίκτυο των Πραγμάτων, Ιδιωτικότητα Κινητών Εφαρμογών

Contents

List of Figures

Part I

Prelude

Chapter 1

Introduction

Personalized Health Services are a major part of the new emerging smart environments, for example, smart cities and smart homes. These smart environments that are included in the main focus of this work are considered part of the wider context of the Internet of Things (IoT), a large variety of devices connected to the Internet using mainly sensors. These devices to provide ubiquitous and context-aware services to humans, may interact and exchange information using several protocols, domains, and applications [14, 70].

In most of the cases, the use of the well-known Machine to Machine (M2M) protocols and Wireless Sensors Networks (WSNs) is needed in order to leverage low-power radios and multihop communication to cover large areas with small, inexpensive, autonomous sensor nodes and interconnection between machines with automation characteristics [70]. In the IoT, the interconnected things can be vehicles, smart devices (smartphones, wearables), traffic lights or even buildings within a smart city.

This thesis examines how synchronous Personalized Health Services within the IoT can affect privacy and could also be affected by privacy-related frameworks and laws. The sensitivity in combination with the amount of the data they are collected, processed and stored within services and systems, that someone can make use in the Internet of Things, is demanding the introduction of special and robust security and privacy mechanisms, that can ensure confidentiality, integrity, and availability.

The recent advances in ICT and the market trends have developed the base for numerous proposed solutions that can be integrated into smart environments like

Smart Cities, where the health services provision is challenging and demanding. While the health domain is still under investigation in Smart Cities, several crucial services are missing. Thus, we, also, explore the proposal of innovative smart cities' services, as part of the Smart Health services initiative. By doing this exercise it is expected that we could better understand this emerging area of research. Afterwards, an in-depth examination, from a privacy perspective, is performed on popular technologies that today citizens use to share, store and exchange personal and sensitive data. All the above have been stressed in the previous months during the rise of the pandemic with the unprecedented need to use privacy-preserving methods for handling very sensitive health and location data, e.g. COVID-19 contact tracing applications, and now with the necessary processing of vaccination history for the Digital Green Certificate[1].

1.1 Thesis structure

This thesis starts with an introduction focused on the emerging domain of the Internet of Things (IoT) and the development of several smart environments, as subdomains, that are framing the IoT landscape over the latest years. Smart health, smart cities and smart homes are just some of them, still important case studies when we are investigating the transformation of the traditional e-health services to the today's ubiquitous and context-aware health applications based environment.

The next chapter examines smart health schemes and approaches for the provision of health services in Smart Cities that could lead to better health services to citizens.

In Chapter 4, we investigate the social mechanisms that users can potentially use to exchange data. Social Networks (SNs) are definitely a powerful channel that needs a careful examination regarding the way that should be implemented in synchronous smart applications, like health applications. Multimedia is, as expected, one of the main information that is stored and processed by SNs. This chapter pro-

[1] https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/safe-covid-19-vaccines-europeans/covid-19-digital-green-certificates_en, (last accessed on 18/04/2021)

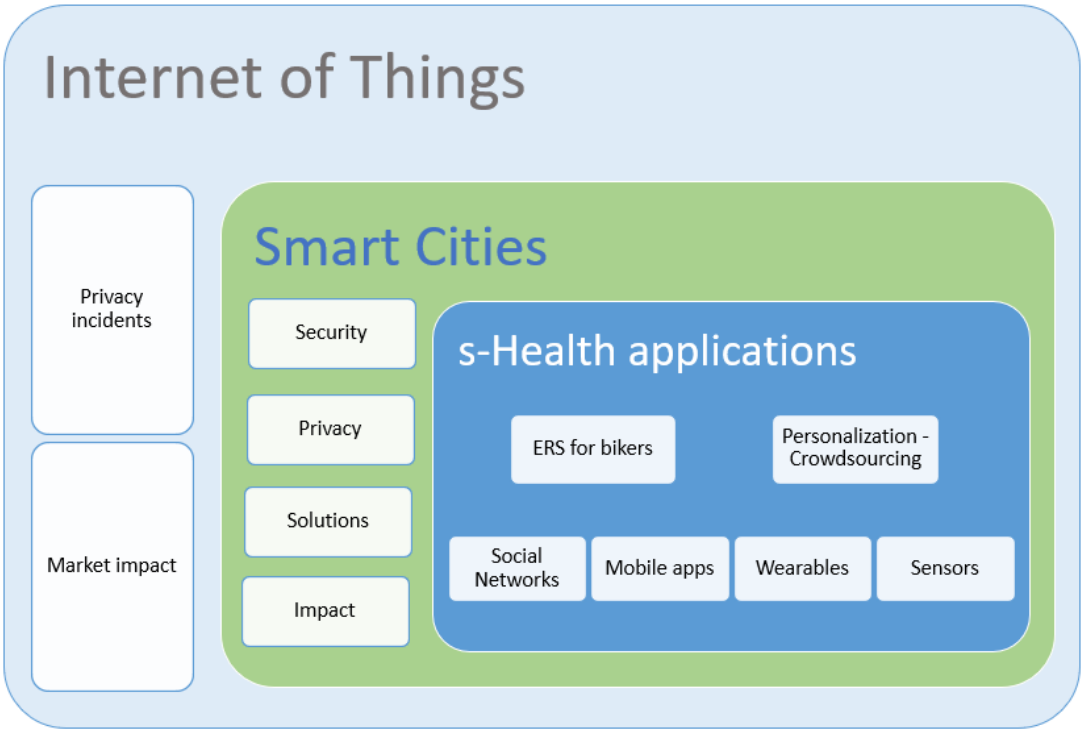

Figure 1.1: Thesis research topics.

vides an analysis of the security and privacy issues combined with their potential countermeasures.

The Chapter 5 analyses the need for a new mechanism that will minimize the risk of sharing multimedia files by users on SNs even when they are shared between different SNs. A cross-platform SN privacy mechanism is proposed. The mechanism is based on watermarks and a framework of policies that could be used to provide trust across multiple socia media.

In Chapter 6 an in-depth analysis is provided regarding the security and privacy mechanisms that popular mhealth applications are base their functions responsible for the health-related services delivery. Also, a structural privacy-related evaluation is performed heavily based on GDPR requirements and the readiness of the mhealth applications in this emerging domain. Furthermore, an investigation of the way that the app developers respond to the security reports is included.

Chapter 7 presents the impact of privacy incidents to the market. More specifically, it presents the results of the revelations of Snowden on the market from a

financial point of view, which are examined as the largest privacy incidents to for-profit data-controlling entities in modern history.

Finally, Chapter 8 concludes this thesis, highlighting the open question and future directions in this research line.

The dissertation ends with the bibliography of the publications and books used as a background for this thesis.

1.2 Research projects

During my PhD research I participated in the following research and development projects:

- OPERANDO Online Privacy Enforcement, Rights Assurance & Optimization operando.eu. Supported by the European Commission under the Horizon 2020 Programme (H2020), as part of the OPERANDO project (GA no. 653704).
- IDEA-C http://idea-c.weebly.com/ Co-funded by the Europe for Citizens Program of the European Union.

1.3 List of publications

Part of the findings during my research for this thesis have been published in peer reviewed journals and conferences. More precisely, the following publications were made in JCR indexed journals:

- Achilleas Papageorgiou, Michael Strigkos, Eugenia Politou, Efthimios Alepis, Agusti Solanas, and Constantinos Patsakis. Security and privacy analysis of mobile health applications: The alarming state of practice. IEEE Access, 2018.
- Constantinos Patsakis, Athanasios Charemis, Achilleas Papageorgiou, Dimitrios Mermigas, and Sotirios Pirounias. The market's response toward privacy and mass surveillance: The snowden aftermath. Computers & Security, 73:194–206, 2018.

- Constantinos Patsakis, Athanasios Zigomitros, Achilleas Papageorgiou, and Edgar Galván-López. Distributing privacy policies over multimedia content across multiple online social networks. Computer Networks, 75:531–543, 2014.

- Constantinos Patsakis, Athanasios Zigomitros, Achilleas Papageorgiou, and Agusti Solanas. Privacy and security for multimedia content shared on osns: issues and countermeasures. The Computer Journal, 58(4):518–535, 2015.

Moreover, the following publications have been made in peer reviewed international conferences:

- Athanasios Zigomitros, Achilleas Papageorgiou, and Constantinos Patsakis. A practical k-anonymous recommender system. In 2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA), pages 1–4. IEEE, 2016.

- Jose Javier Martinez, Peio Lopez-Iturri, Erik Aguirre, Leire Azpilicueta, Constantinos Patsakis, Achilleas Papageorgiou, Agusti Solanas, and Francisco Falcone. Analysis of wireless sensor network performance embedded in motorcycle communication system. In 2015 USNC-URSI Radio Science Meeting (Joint with AP-S Symposium), pages 210–210. IEEE, 2015 (poster presentation).

- Achilleas Papageorgiou, Athanasios Zigomitros, and Constantinos Patsakis. Personalising and crowdsourcing stress management in urban environments via s-health. In Information, Intelligence, Systems and Applications (IISA), 2015 6th International Conference on, pages 1–4. IEEE, 2015.

- Constantinos Patsakis, Achilleas Papageorgiou, Francisco Falcone, and Agusti Solanas. s-health as a driver towards better emergency response systems in urban environments. In 2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings, pages 214–218. IEEE, 2015.

Furthermore, I participated as co-author to the following open access book chapter that was a result of the successful EU Horizon 2020 COST Action called CRYPTACUS:

- Agusti Solanas, Edgar Batista, Fran Casino, Achilleas Papageorgiou, and Constantinos Patsakis. Privacy-Oriented Analysis of Ubiquitous Computing Systems: A 5-D Approach, pages 201–213. Springer International Publishing, Cham, 2021.

Moreover, I participated in the following workshop as a presenter:

- Achilleas Papageorgiou, Michael Strigkos, Eugenia Politou, Efthimios Alepis, Agusti Solanas and Constantinos Patsakis, How private is your mobile health advisor? Free popular m-Health apps under review, CRYPTACUS: Workshop & MC meeting, 16 - 18 November 2017, Nijmegen - Netherlands https://cryptacus.cs.ru.nl/program.shtml .

Chapter 2

Privacy in personalized health services

2.1 Introduction

From the early years before the Internet became part of humans' life, there was a debate between the need for the respect of users' privacy and surveillance. One of the most famous phrases is the "*nothing to hide argument*" that describes the huge phenomenon that we are witnessing on the web. The phrase "*I have nothing to hide*" often mislead the majority of people [140] who tend to expose themselves to several privacy risks.

Privacy is a fundamental human right[1].

E. Houghes in [73] defines privacy as follows:

> *"Privacy is not secrecy. A private matter is something one doesn't want the whole world to know, but a secret matter is something one doesn't want anybody to know. Privacy is the power to selectively reveal oneself to the world."*

Moreover, privacy is one of the most serious –usually missing– characteristics within a service. When we are evaluating a health-related service, then privacy should be in the core that any application should guarantee to its users. Privacy

[1] Universal Declaration of Human Rights - Article 12 *"No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. Everyone has the right to the protection of the law against such interference or attacks."*

through the years, the huge growth of the Internet, and personal data sharing have been upgraded and transformed into a demanding fundamental feature that users and also, the law requires in their online daily life..

But before the deep dive into privacy issues on personalized health services, a brief historical review of some important milestones of this domain through the years is deemed necessary and will be presented below.

The first milestone is the establishment of the Right to Privacy [163]. Samuel D. Warren and Louis Brandeis argued in favour of the "*right to be left alone*", using the phrase as a definition of privacy in 1890. In 1948 The Universal Declaration of Human Rights is adopted, including the 12th fundamental right; the Right to Privacy [2]. In 1967 a very important decision was made, and the Freedom of Information Act (FOIA) came into effect in the US[3]. It was then that the public was granted the right to request access to records from any federal agency. Several countries followed this practice to provide advanced privacy rights to their citizens.

In 1980 the Organization for Economic Cooperation and Development (OECD) introduced several guidelines to harmonize national privacy legislation among international laws and ensure data protection[4].

In 1981, the Council of Europe adopted the Data Protection Convention (Treaty 108)[5] and introduced the right to privacy as a legal imperative. The Treaty was not limited but highly focused on the automatic processing of personal data. Another milestone on privacy and data protection was achieved in 1983 by the Federal Constitutional Court of Germany. The court made a fundamental decision regarding the *census judgment*, known as the 1983 Census Act. In 1995 the Directive 95/46/EC[6], a European Union directive which regulates the processing of personal data within the European Union (EU), was adopted. The Directive was characterised as a major component of European citizens' privacy and data protection.

[2] http://www.claiminghumanrights.org/udhr_article_12.html, last accessed on 18/04/2021.

[3] https://www.foia.gov/about.html, last accessed on 18/04/2021

[4] https://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflo
htm, last accessed on 18/04/2021.

[5] https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108, last accessed on 18/04/2021.

[6] https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A31995L0046, last accessed on 18/04/2021.

In 2002, the EU implemented the Privacy and Electronic Communications Directive 2002/58/EC on Privacy and Electronic Communications[7]. The directive is known as ePrivacy Directive (ePD) and it is highly focused on data protection and privacy in the digital age, including several important issues that were under debate until then, like information confidentiality, spam communications, and cookies installed on devices by applications. In 2014, a ruling by the Court of Justice of the EU concluded that users have the right to request by search engines like Google, to remove results for queries that their personal data are included. This right that is connected to the EU law is also known as "the right to be forgotten". In 2016 the European Commission adopted a new legal framework (Regulation EU 2016/679) for protecting individuals' personal data, the General Data Protection Regulation (GDPR) which will replace the existing 1995's Data Protection Directive[8]. GDPR became applicable on 25 May 2018, harmonizing this way the various national regulations across the European Economic Area (EEA).

2.2 Internet of Things (IoT) in healthcare

A massive hype around the emerging ubiquitous, context-aware communication technologies that include interactions and interoperability between things has led to the generation of *Internet of Things* (IoT) concept. IoT is defined[9] as the proposed development of the Internet in which everyday objects have network connectivity, allowing them to send and receive data. The majority of users today, after the second era of the web, widely known as Web 2.0, manage several social media accounts and generate rich content online. Thus, there is a demanding need for fast, efficient, mobile and continuous network interaction. Moreover, the introduction of smart devices and applications that users can easily make use for their daily needs forced the market and the global vendors to embed sensors and actuators, Machine-to-Machine (M2M) communication interfaces and wireless protocols towards the

[7] https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058, last accessed on 18/04/2021.

[8] https://eur-lex.europa.eu/eli/reg/2016/679/oj, last accessed on 18/04/2021.

[9] https://www.lexico.com/definition/internet_of_things, last accessed on 18/04/2021.

scope of a highly interoperable Web of Things[10] that operate autonomously over the Internet as a major part of the IoT.

IoT, as a massive collection of devices connected to the Internet using mainly sensors, develop a whole new ecosystem of interactions and services. The devices to provide ubiquitous and context-aware services to humans may interact and exchange information using several protocols, domains, and applications [14, 70]. In most of the cases, the use of the well-known (Machine to Machine) M2M protocols and (Wireless Sensors Networks) WSNs is needed to leverage low-power radios and multihop communication to cover large areas with small, inexpensive, autonomous sensor nodes and interconnection between machines with automation characteristics [70]. The results of these interactions between humans, devices, several things that through sensors are characterized as 'smart', and more sophisticated devices that implement extensive architectures, when interacting altogether, they collectively form a smart ecosystem that based on societal trends, and their scope can be distinguished and implement different features and services, e.g. energy, industry, domotic, health, transport, mobility, entertainment, sports, and lifestyle. Within IoT, the interconnected things can be vehicles, smart devices (smartphones, wearables), traffic lights or even buildings within a smart city.

Today, these individual smart environments have developed several completely alternative areas of research. More specifically, some of the major IoT domains that have been developed based on researchers and market forces are listed below:

- Smart Cities; Cities that utilize the IoT technology to automate and improve city services and citizens' quality of life. Within smart cities, someone can find many 'smaller' smart environments of smart 'things' (e.g. home appliances, industrial machines/smart-grids, connected vehicles, wearables) that interact to provide to citizens better-personalized services.

- Smart Health; The introduction of smart services within urban, suburban, rural, and exurban environments developed the need for a new generation of health services. Smart Health (s-Health) is the provision of health services by using the context-aware network and sensing infrastructure of smart cities [137]. Additionally, the use of IoT and ubiquitous services by users using

[10] https://www.w3.org/WoT/, last accessed on 18/04/2021.

their smartphones or wearable devices to track health-related issues could lead to the provision of more efficient health services.

- Smart Homes; Homes can automate several functions and control services like lighting, security, air temperature, energy management. Homeowners can also use services based on mobile or web-based applications through their smart devices (smartphones, wearables) to control smart home. Several alerts that assist them in managing their smart homes better, are also available to their devices.

- Smart Transport; From intelligent infrastructure to vehicles and roads (for example traffic lights, road sensors, radars) to connected vehicles, the area smart transportation aims to provide safer, intelligent, and faster transportations. Smart transportation is taking advantage of the adaption of ICT technologies (sensor networks, context-aware, and location-based services) to provide the improvement of all the passengers, drivers, and pedestrians' quality of mobility.

- Smart Industry; Industrial manufacturing and the introduction of upgraded production systems led to advanced connectivity, interoperability, and decentralized systems and services applied in the production cycle. Additionally, data analytics and artificial intelligence capabilities advanced several areas of production and finally improved critical industrial procedures.

Personalized healthcare, precision healthcare, patient-centric healthcare, health 2.0, medicine 2.0, mhealth, or smart health, are terms that while they have differences mainly refer to or include patients in their center of interest and the advance of information technologies for information sharing and health delivery. Meanwhile, in the age of big cities' development, large populations gather there, and a whole new domain of research has risen based on the need for more efficient, proactive, and cost-effective delivery of health services. Notably smaller societies like villages and sub-urban areas could also be advanced by the use of personalized services to predict a health-related situation or to deliver emergency response services to health incidents. Technology can be used to better understand and respond to situations where the transitional services could not.

One key element for the success of personalized health services is the great familiarization of the citizens with synchronous technologies. It is fair to say that SNs, web services, mobile applications, wearable technology and other newly introduced sensors to smart devices or environments have already been integrated to the majority of citizens lives. In fact, some of them play a crucial role in everyday life. Some examples of this familiarization can be considered the increase of Facebook, Twitter and Linkedin (social media) usage and the growth of the e-governmental and e-health services as a result of the growth of the web services. Moreover, Android and iOS software and generally the use of mobile devices as one major communication channel by citizens, in combination with the emerging market of smartwatches or other wearable devices, that are mostly known for wellness and athletic/sport activities or clinical use by professionals to their patients to better understand their condition, are included to the familiarization examples. We are witnessing the continuous integration of smart sensors and services within smart cities, smart hospitals, smart homes and more smart environments where users can exploit a plethora of useful services based on their needs.

The concept of Smart Health (s-Health) that Solanas et al. [138] proposed, extends both Smart Cities and electronic health (e-Health) in exploiting urban sensors and Smart City infrastructures to provide health care services, can be considered as the next generation of Personalized Health Services in the era of Smart Cities. The concept is based on the fact that systems with context-awareness can have a great impact on citizens [117] and city management. The wide deployment of sensors can provide citizens with vital information and novel services, while simultaneously facilitating the organization of the city to achieve its goals for smart participation, sustainability, and more.

2.3 Smart Health services: Smart Cities should provide services for healthier citizens

The urbanization of modern cities is rapidly progressing. This significant change, along with the population growth, is expected to increase current urban life problems and introduce new challenges. Smart Cities will possibly make our cities sustainable and improve citizens' lives. The major difference between regular cities

and Smart Cities is the extensive use of Information and Communication Technologies (ICT) in their management and the deployment of sensor and communication infrastructures to allow real-time data collection, communication with actuators, and dynamic resource allocation.

We performed an investigation regarding any missing, still demanding, services that could deliver even more efficient and personalized health services to citizens, and we conclude to two major proposals. The first one is based on how a smart city or community could centrally manage a transportation accident to provide better emergency response services to its citizens. The case study is focused on motorcycles and the technologies that could be used to support such an accurate and robust emergency response system. The second proposed solution is based on how the technology could be used to better feed with valuable data a smart city service that could recommend routes to citizens that suffer from specific symptoms or disease. Below, there is a brief introduction to each smart health solution.

2.3.1 Smart Health and Emergency Response Systems

In this section, we study emergency response systems that could save riders' lives after an accident, and we propose a solution based on a real-world scenario. By simulating several scenarios based on real routes that the ambulance could follow to reach the motorcycle and then deliver the injured rider safely to the closest hospital, we illustrate the improvements that s-Health can bring to emergency response systems.

Our research begins with the assumption that in big cities today, due to high traffic congestion, the use of motorcycles is very frequent in urban environments, especially in areas where there are few rainfalls. While motorcycle drivers significantly reduce their commuting times, they are exposed to many risks as small human errors or even bumps on the road can lead to severe injuries or even casualties. Definitely, the response time to such events can be turning points from casualties to saving citizens' life.

Systems like Event Data Recorders (EDR) also commonly known as "black boxes", are being used in transportations for years. Their role is to keep a record of all the

important information so that in the event of an accident, experts will be able to reconstruct and/or simulate the events that led to it. For example, in cars, an EDR can store information such as vehicle and engine speed, the force of the impact, steering input, airbag deployment lapse, accelerator position, brake status, seatbelt status, passenger's airbag, etc. Additional advanced EDRs have been introduced in the market by insurance companies that provide more features such as alert systems on a crash.

On the contrary technologies like EDRs have not been installed in motorcycles while there are more prone to accidents, and the driver is far more exposed. In the bibliography, someone can find proposals that are mostly based on wearables devices [125, 101]. Unfortunately, wearables are more liable to faults and physical damage, while their accuracy depends on the way the user will adopt them which is also based on how well trained will be to fit it to the right position.

Even if computational systems have been by the years integrated to motorcycles [15], the interaction that they have with the infrastructure is mainly related to push notifications to the user, the arrangement of mechanical revisions, requests for information, etc [96]. Based on our findings, there is a missing link to systems that are focused on accident reporting or their management.

Through the years, several cities are implementing many features towards their transformation to the so-called Smart Cities. Handling emergencies is not only a crucial and very thorny issue but a demanding service for urban management to decrease the number of accidents that result in deaths. Even small delays can lead to casualties. Of course, delays can be caused by several issues, with the most important being traffic jams.

In this regard, we propose a novel scheme, that takes advantage of the s-Health concept to report and respond to emergencies of motorcycle riders. Our proposed scheme detects an emergency, reports it to the corresponding hospital and caters for the safe delivery of the injured with the least possible latencies. Moreover, we performed a number of simulations based on real data retrieved by Open Street Map (OSM) data with a scope to collect data regarding the percentage of reduction of the response time to increase the accuracy of our scheme. We have studied the actual impact of prioritizing emergency vehicles on traffic lights. Contrary to many current solutions, which allow emergency vehicles to bypass parts of the

traffic through the use of tokens, we have gone one step further and rearrange the traffic of the whole city to cater to the needs of the emergency vehicles.

Our solution is unique in that it introduces a novel emergency response system for motorcycles, a vehicle that is more prone to accidents, and that highly automates the procedure. The simulations indicate that our proposal reduces the response time efficiently. Further research will consider the experimentation with real-time traffic data so that the simulations will be more realistic. Also, we will evaluate the cost of an average driver in terms of time and consumption.

2.3.2 Smart Cities and citizen-centric personalized health services

Over the years, most mobile applications' providers focus on the mass use of these applications, which, consequently, have developed, their discrete market with their marketplaces. Furthermore, this trend has triggered the development of a variety of new applications that have improved citizens' everyday lives. Most of these applications are offered freely, and they monetize users' preferences, their demographic characteristics, as well as the extent of their user database by mainly providing them to advertising companies. The extraordinary targeting abilities that these applications offer to their providers have transformed them to become or be part of successful business models, due to their targeted audience.

As the main challenge of the mobility that modern devices provide to their users is their satisfaction, dozens of applications are offered to them, aiming to serve as informational agents to make their lives better. Within smart cities, some citizens already use mobile applications to find the exact time that the bus will pass and buy their tickets[11], the closest pharmacy to their location[12] or even the shortest route to their workplace[13].

[11] https://play.google.com/store/apps/details?id=com.lothianbuses.lothianbuses, last accessed on 18/04/2021.

[12] https://play.google.com/store/apps/details?id=com.guardianpharmacy.android&hl=el&gl=US, last accessed on 18/04/2021.

[13] https://itunes.apple.com/us/app/inrix-xd-traffic-maps-routes/id324384027?mt=8, last accessed on 18/04/2021, https://www.google.com/maps/about/#!/, last accessed on 18/04/2021

Following the extreme mobile usage of the latest years, wearable devices, are becoming more and more accepted by modern users for their everyday life activities. In this regard, several applications are exploiting this new information to provide new features, more accurate measurements, and even more personalized applications. Nonetheless, the most famous applications is focused on sports and healthcare as most wearables can measure heartbeat rate, sweat and motion [131].

In this section, we propose a concept that could provide a more targeted informational channel to the existing information flow that a smart health service provides. The main scope is to exploit urban sensing with wearables to provide novel functionalities and more personalized user experience by introducing a novel scheme for context-aware mobile applications based on services with ubiquitous characteristics, as an s-Health solution, to citizens with stress or anxiety disorders. The parts involved in our scheme are a mobile app, one or more wearable devices, a platform that could collect, manage and aggregate information and the appropriate city infrastructure that will collect environmental and urban data.

The unprecedented urbanization is pushing towards the realization of Smart Cities. The vast deployment of sensors can provide a real-time overview of the city and a wealth of information that researchers are trying to mine to extract new knowledge. Nevertheless, urban life is quite complicated and more stressful than life in the suburbs. As a result, many people feel trapped in a rat race from which they cannot escape. In some cases, this feeling can lead to nervous breakdowns and stress crises. In this work, we introduced a novel framework that offers a more holistic approach to stress management compared to the current state of the art, as it incorporates many technologies and could provide more advanced intervention.

The realization of the framework faces many challenges, such as data accuracy; sensors might not be properly calibrated or user-contributed data might not be accurate, device interoperability, seamless user experience, user acceptance, and privacy just to name a few. These challenges consisted of the inspiration for the work in Chapter 6. For device and protocol interoperability, platforms like Anypoint[14] could facilitate the development while protocols such as Ardagna et al. [11] could provide users with the needed security and privacy. However, it has the potential

[14] https://www.mulesoft.com/, last accessed on 18/04/2021.

to make a significant impact in the lives of individuals and provide an insight into how we should make our cities more sustainable and what are the key aspects in designing a human-friendly urban environment.

2.4 Health information sharing and privacy applications

Without any doubt, the most significant part of the shared information within the above services belongs to the user-generated content. Also, most of their functionalities are based on a user profile creation. The transformation or the user from reader to writer, evaluator or manager of these services lead to an empowerment of his profile and this is one of the main reasons that the latest years gave security and privacy a great role that developers and businesses should re-evaluate and establish as a fundamental characteristic of their applications and services.

European Union has tried to investigate and support for several years the introduction of privacy-related mechanisms, guides and tools directions to developers and integrators that are a key part of this emerging ecosystem to better secure their users' personal data. Without any doubt, the introduction of the General Data Protection Regulation (GDPR) [61] was one of the greatest milestones in the modern history of privacy. GDPR was adopted on 14 April 2016 and became applicable on 25 May 2018. GDPR introduced several new requirements to controllers and processors of personal data while introduces the Privacy by Default and by Design as part of the Regulation. It also introduced specific requirements for sensitive data such as health and medical-related data in the context of data that fall in the special categories (Article 9) of personal data and should be only processed under specific conditions and purposes.

More specifically, the Article 9 of GDPR states that: "*Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data to uniquely identify a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.*"

As expected, based on the period of the research and writing of this thesis and by 2016 that GDPR was announced to be adopted, a great focus was also given to

research towards the directions and the readiness of the commonly used technologies by citizens for health services provisions like SNs and mobile applications.

2.5 Social networks in healthcare

During the last years, we witnessed the phenomenon where the digital persona of users has moved from their personal websites or blogs to their SNs profiles. Definitely, a key factor in explaining this shift resides in the simplicity that SNs provide to their users to manage their social lives. Users that created and manage social profiles can easily share worldwide their personal information with others. At the same time, they have several options to the uploaded content like the ability to edit or delete it, at least in the majority of the cases.

The design and the development of new applications and services brought radical changes to the way that the users interact with each other, but at the same time generated a whole new market of services that created the so-called second generation of the Web (Web 2.0). In the healthcare domain, this movement was represented by the terms Medicine 2.0 and Health 2.0 [156]. User-generated content, the use of social networking platforms, the networking between health professionals, the collaboration between patients and professionals are just a few of the massively introduced characteristics of this notion. From the early years of SNs, someone could find several health-specific SNs that provided their users the ability to connect with each other and exchange information that can be used for their health benefit. Examples of health SNs can be considered patientslikeme.com, inspire.com, medhelp.org, dailystrength.org.

Over the last few years, social networking, not only became an essential part of the digital strategy of health organizations but also brought an evolution in the way of communication between all parties involved (patients, professionals, companies). SNs like Facebook and Twitter are widely used, also, in health communication [65, 7, 52] and give a gateway for patients or professionals to seek support if needed. As a result, users found a whole new way to communicate about health issues without the need to change the platform they use. Undoubtedly SNs have established a big market share of the Internet. Millions of users are using them daily and their traffic, as well as their influence, is continuously growing. While this is

a benefit in terms of flexibility and easiness for the user, we cannot say the same thing for his privacy, and generally security [69, 112, 124, 147, 74].

2.5.1 Security and privacy in social networks

Since SNs are being used every day by millions of users, but the exchanged information is not limited to just messages between friends or typical small talk socialising but as we said above as platforms to share their health data. Therefore, this emerging phenomenon leads to the need for the creation of new standards that should be applied by vendors to protect their users' personal and sensitive data.

Today we can find malicious users that try to exploit software vulnerabilities of SNs' infrastructure or other non-technical advanced users still trying to bypass privacy measures and gain access to sensitive information. Although hacking is not an easy practice on SNs, we often witnessed several attacks like identity theft and impersonation.

The secure management of users' health data and other profile data that can be used to impersonate them to health or reputation related attacks to their personal or professional environment is today a major privacy issue. Multimedia is also a crucial part of the user information, but also medical information that can be used (not limited) for the following reasons:

- to be identified by his/her friends
- to share visual medical information (for example x-Rays or CT scan's images)
- to share medical prescriptions or medical advice

Thus, the protection of such multimedia content can be considered a demanding feature. This section is focused on the security and privacy risks and issues of SNs while particular research has been conducted based on the multimedia protection within SNs. Numerous studies have already highlighted many of these issues; however, few of them are focused on the core of the shared information, the multimedia content. An up-to-date categorised mapping of security and privacy of these risks is provided.

2.5.2 The need for a more interoperable approach to social networks' security mechanisms

Security risks are increased by the years and users' privacy is exposed in several ways. Thus, a whole new approach for the secure management of multimedia content is demanding. This section includes a research proposal with a very specific authentication scheme that in case of the health-related SNs can provide data protection in several ways. Today, health professionals and patients usually maintain multiple accounts on several SNs based on their professional and personal needs. An attack scenario on SNs used for health-related multimedia sharing could be the following.

Bob is a malicious user that wants to harm the professional fame of Alice based on some previous events between them. Bob is included in the friend list of Alice on Facebook. Alice uses Facebook without considering that a shared with their friends multimedia, could be considered as a potential privacy issue for her. Bob, on the other hand, has started to plan his attack scenario to manipulate Alice's professional fame. Bob already knows that someone can download and re-upload users' photos to another profile or page within a SN without any limitation regarding the ownership or authentication within the same SN. But, Bob wants to perform an impersonation attack by creating a profile of Alice to a professional network and by uploading some photos that Alices has shared to a close group of job 'friends' regarding some serious health symptoms she had. After some days that Bob worked on his attack scenario, Alice has a fresh new profile on one of the most well-known professional SNs that HR recruiters usually check before their decisions and as we can easily understand she does not know about that and it would be very hard to learn for that as there is any related mechanism, like the one presented on [176], that could notify her for such an attack against her privacy. Alice could learn about this attack only by chance.

The core privacy issue stems from the fact that the SN does not check which multimedia are being uploaded and as a result, there is no privacy-related management on how any of the uploaded multimedia are being shared and/or re-uploaded by the SN's users. Also, in this section we approach an additional issue; modern SNs do not care about their users' privacy out of their context and "borders". From

a marketing point of view, this could be something accepted, as each SN tries to enlarge its own users' database. Nevertheless, this situation has developed a whole new hacking scene where social engineering, impersonation and ID theft are more than frequent.

A major issue is the authentication mechanism that could be implemented to protect the published or privately shared health and medical images. Additionally, the proposed authentication mechanism can also be used to protect the identity-related multimedia for a user to be protected by ID theft. Multimedia content authentication today in SNs is a significant privacy and security open issue, as apart of the privacy settings that someone can apply based on the SN's policy and privacy-related features, there are several ways to bypass the security of those settings and attack to the owner of the multimedia.

User's multimedia content is not only prone to attacks from skilled malicious users, but even reckless users, that can arbitrarily re-upload any image they have access to. In [176], we proved how easy it is for not only a hacker but also for a simple user to leak multimedia content outside the trusted zone of a user profile. Furthermore, we experimentally tested several setups of multimedia and potential protection mechanisms and resulted that fine-grained privacy policies can be implemented using digital watermarks. This solution can be applied without the need to build a SN from scratch.

Definitely watermarks as a protection mechanism for medical-related multimedia content are not something new [10]. This means that watermarks can be a mature way to protect data while their implementation of a secure scheme for the protection and authentication of health multimedia can be considered without doubt innovative.

In this section, we mainly focused on achieving more secure and private sharing models and their feasibility within current structures, in terms of implementation effort, processing needs, and economic constraints. Thus, a new scheme that enables collaboration between SNs to enhance users' privacy is proposed. The proposed scheme uses no Trusted Third Parties (TTP) and introduces the new feature of the shared ownership of multimedia. The goal of this work is to deliver a whole new approach that will let cooperative authentication between different SNs that can effectively and securely manage the privacy of their users both in terms of their

personal data protection, but also regarding their sensitive health/medical multimedia information. Another important advantage of this solution is its implementation, as it does not require the redesign of a current SNs' architecture on scratch.

2.6 Mobile health applications

Based on the special focus of this thesis to the health-related applications that could be used by users to interact in several schemes, in this section, we study, and in-depth analyze, the security and privacy protection that mobile health applications can bring to their users. The emerging environment of sensors embedded on smart devices, like smartphones, that can sense changes in environmental and human body measurements, with the scope to provide useful and personalized alerts and reports constitute a great severity to the risk that this information may have in case of miss-protection. Smartphones, today, are also used for storing and processing sensitive information as well, such as user's location, lists of contacts and personal multimedia files that associated with the health-related users' information. Thus, m-health apps depending on their features and settings have to deal with a great amount of data, which are considered very sensitive and are highly protected by national and international regulations such as the GDPR.

The new mobile interactive environment of devices creates several difficulties to the security and privacy [143]. On the other hand, someone would expect that when we are examining m-health apps, then security and privacy should be at their core principles. Nevertheless, there are many examples where applications (not limited to health-related apps) failed to protect their users' privacy due to either inappropriate implementations or poor design choices [58, 124, 120, 39, 76].

An in-depth analysis of how m-health apps managing user's personal and sensitive data was performed in [107]. Android's Operating System (OS) was selected as the most popular OS the period this research was performed. More specifically, we retrieved data and apps from Google Play, the official apps store of Google. The applications were selected based on quality, popularity and content-related criteria. Finally, 20 apps were selected and evaluated under a number of manual and automated analysis to study and understand the way each app manages and protects users' data security and privacy. Surprisingly, the majority of the analyzed apps

do not meet the expected standards for security and privacy, leaving unprotected personal and sensitive data to hackers or by sharing sensitive data to third parties without any security protection and without the upfront user consent or even their knowledge.

Our study is innovative and has unique features with respect to previous articles in this area. We provide an analysis of security and privacy concerns in m-health apps through long term evaluation, monitoring and recording of the full life cycle of the apps (from January 2016 to August 2017), assessing the quality of all communication channels. Moreover, we investigate the way that app developers responded to the security reports we submitted them. Finally, we perform a GDPR compliance auditing procedure to determine whether the reviewed apps conform to the new EU legal requirements.

The results of our research were that most of the mobile applications found not to protect their users' personal and sensitive data. Not only they did not apply security mechanisms and best practices that could protect data by attackers, but they also were found to share their users' data with third parties without their upfront consent or any other notice.

In light of our findings, security experts and privacy advocates raise the alarm about the potential privacy harms that derive from m-health apps processing personal and sensitive data and the urge for suitable countermeasures. To build solid foundations and easily implementable privacy standards for the development of m-health applications, and especially for fostering trust among their users, European Commission issued in 2016 a draft "*Code of Conduct on privacy for mobile health applications*" [38]. Although its final version is yet to be adopted, as it is subjected to the implementation of Article 29 Data Protection Working Party comments [64] and to its conformance to the GDPR's provisions, it is still a good reference point for providing practical guidelines to app developers to build reliable applications compliant with data protection standards and principles.

Part II

Personalized Health Services

Chapter 3

Smart Health services and Smart Cities

3.1 Smart Health and Emergency Response Systems for bikers within cities

The recent years we are coming across the phenomenon where cities are shifting towards more smart and flexible ICT-based solutions that can advance domains like energy, health, environment/pollution, traffic, emergency response, and more. Health care today is shifting towards more personalised services that better fit the needs of each patient. In this section, we will explore some hybrid smart solutions that could potentially be implemented within smart cities.

In this section, the main focus is the improvement of the emergency response systems offered to motorcycle riders.

Big cities are suffering from high traffic congestion, a situation that not only leads to several negative situations like delays daily, air pollution due to high traffic, increased risks for the citizens' health even their psychology.

As expected and due to high traffic congestion, the use of motorcycles is widespread in urban environments, especially in areas where there are few rainfalls. Even if motorcycle drivers significantly reduce their commuting times, they are exposed to many risks as small human errors or even bumps on the road can lead to severe injuries or even casualties. For this reason, a city's authorities must collaborate on improving the response time to such events immediately as seconds can be, in

many cases, turning points from casualties to saving citizens' life.

For many years, cars have been using Event Data Recorders (EDR), commonly known as "black boxes". The role of such devices is crucial, as they are responsible for recording of all the important information included in an accident. This gives experts the ability to reconstruct and/or simulate the events and what led to it. Example of the information that is stored in the case of a car accident:

- vehicle and engine speed
- the force of the impact
- steering input
- airbag deployment lapse
- accelerator position
- brake status
- seatbelt status
- passenger's airbag

Unexpectedly and in contrast with the damage that bikes can lead to, as the driver is far more exposed, technologies like EDRs have not been installed in motorcycles. The time of the research solutions that found to exist in the market are mostly based on wearables devices [125, 101], which are more susceptible to faults and physical damage. Even if some systems have been integrated into motorcycles to made them more "computerised" [15], at the end are mostly limited to interactions with the infrastructure, providing push notifications to the user, arrangement of mechanical revisions, requests for information, etc. [96].

EDRs are usually described in standards as responsible to store data in a well-bounded interval, e.g. five seconds before the crash and five seconds after it. This limitation is put in place to avert arbitrary driver monitoring and privacy invasion, nevertheless, other measures have also been proposed [116].

Nevertheless, the computerised interaction systems do not give a complete solution to an accident reporting or its management after that. That is the main motivation of the current work, where a novel scheme is introduced and which takes

advantage of the s-Health concept. The scheme is aiming to support the detection of an emergency, manage its reporting to the closest hospital, and take care of the safest delivery of the injured biker without latencies. The simulations included in our work prove that such a scheme can reduce the response time, increases the accuracy of the reporting values, and further automate the accident management.

Below there is a summary of reasons that constitute the problem of the current state:

- Technologies/mechanisms that are implemented in cars have not yet been used in motorcycles even though the latter are more prone to accidents.
- Motorcycles do not come pre-equipped with "black boxes" (Event Data Recorders-EDR).
- The total response time for the caters to deliver an injured rider has to be significantly reduced.
- It is crucial to respond to such events immediately as minor delays can lead to casualties.

The main contribution of our proposal is the introduction of a novel scheme which takes advantage of the s-Health concept. Our scheme detects an emergency, reports it to the corresponding hospital, dynamically regulates traffic lights, for the ambulance to meet only green lights, so that to caters for the safe delivery of the injured with the least possible latencies.

3.1.1 Related work

Emergency response systems have an important role in handling accidents and more precisely when coming to road accident events. While the majority of such systems are operated by phone call services, someone can find more automated solutions specially designed for cars[1]. These systems exploit cars' sensors to report incidents that can include car damages, accidents, or theft attempts to the service

[1]Examples of emergency response systems on cars are OnStar (https://www.onstar.com/, last accessed on 18/04/2021.) and AcuraLink (https://acuralink.acura.com/, last accessed on 18/04/2021).

provider. However, such systems are not widely used and are not designed to be used by motorcycle riders.

Below, there is a summary of basic actions that are taken place after a bike accident and the estimated time that each action costs for an injured citizen until the time that could receive treatment to the closest hospital.

- Time to contact emergency services: From the time that an accident happens to the time that someone will call the emergency response call centre, there is a wide range, and it depends on several conditions. First of all, a witness should be in the place of the accident. In such cases, the time that a witness needs to call the emergency response call centre can range from 30 seconds to several minutes, as in most cases, people will try to identify what happened and if possible try to help the injured before reporting the incident.

- Time to answer the call: This is related to the time that the call centre of emergency service needs to receive and successfully register the call. There are some country-based examples of the time needed. For instance, in 2012-2013 in Australia 90.90% of emergency calls were answered within 10 seconds, while 98.02% of emergency calls were answered within 45 seconds[2].

- Activation time: This is the time needed by the operator to record the incident and request the closest unit to go to the location of the accident.

- Response time: This is the time needed from the response unit to arrive at the accident's location.

Within the above actions, there is also a critical element that can affect the required time for an ambulance unit to reach the accident area; the traffic lights. Currently, there is a lot of research on how to manage traffic lights with urban environments. Also, several algorithms have been proposed to optimise the use of traffic lights by using machine learning [166], road-to-vehicle communication [160] or vehicle-to-vehicle-to-traffic-light communication [63].

Today, also in many countries, authorities try to install traffic signal preemption systems as part of their current infrastructures aiming to succeed in the decrease of

[2] https://web.archive.org/web/20150419090916/http://www.ambulance.nsw.gov.au:80/Our-performance/Response-Times.html, last accessed on 18/04/2021.

emergency response time. This decrease is mostly based on the function that allows emergency vehicles, usually based on RFID tokens, to bypass the normal operation of traffic lights by changing them to green the time the vehicle passes [165]. As expected, the above methodology can reduce response times, but there are some drawbacks. For example, one of the drawbacks is that for an emergency response vehicle to pass a previously red light, the vehicle should approach the traffic light to transmit the signal to the traffic light infrastructure successfully. This is not always easy, for example, in case of a congestion situation. Additionally, this is not a global solution that could support the synchronisation of the other traffic lights that the vehicle should pass to approach the accident location. Moreover, security-wise it is almost infeasible to restrict attacks, e.g. from other drivers to reduce waiting times or other nefarious[3] and sinister activities [62].

The above indicates the need for the introduction of advanced automation that could manage such an incident. A notable improvement would be if the emergency response systems would not limit their operation functions to phone calls. Automation would minimise the time to contact emergency services and that to answer the call under the barrier of a second. EU has already planned to fund projects aiming to improve and extend the functionality of traditional "112" products[4]. A good idea would be to couple such a solution with a recommender system, as it will be also discussed in this section, towards the reduction of the response time.

3.1.2 Use case scenario

The scenario that can be used to evaluate the need for such a solution is the following; Bob is riding his motorcycle in a suburban road of his city $\mathcal{A}$. For an unspecified reason, Bob lost control of his motorcycle and fell in a given location (x, y). Bob is now injured and not able to call for an ambulance. Only witnesses could help in his situation. Similar to Bob's accident, in many accidents in suburban areas, where the traffic is very low, there might be no witnesses. Also, even in cities' centres when an accident occurs very late, there will likely be no witnesses. It is not rare that bikers,

[3] http://www.wired.com/2014/04/traffic-lights-hacking/, last accessed on 18/04/2021.

[4] http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/1077-drs-19-2014.html, last accessed on 09/05/2015.

to be left on the road heavily injured, either because they involved in the accident driver's panic or because they leave their victims for other reasons. As a result, it is almost impossible to estimate the final time needed for an ambulance to receive a signal to start moving to the accident location and pick the injured biker.

The most critical issue that this solution aims to address is that even if many cities are equipped with traffic cameras it is almost impossible to 24/7 monitor them in real-time to successfully report an accident in the first minutes this happens to save injured lives. Also, in case there are no witnesses in the accident location when it occurred, the possibility to be immediately reported in case of heavy injures is almost zero. But even there are witnesses, nothing can guarantee that the accident will be automatically reported, letting injured drivers at great risk for their lives.

3.1.3 The proposed scheme and architecture

In our proposed scheme, some basic assumptions are made. First, we assume that a smart infrastructure dynamically can manage traffic lights and that infrastructure could be re-arranged depending on traffic [166] or other events. As a result, traffic events and traffic lines could potentially be blocked or released, depending on each specific management needs.

Apart from the road infrastructure, the motorcycle should be equipped with some special hardware to make our scenario work successfully. In our scheme's architecture, we propose the use of impact sensors, as presented in Figure 3.1. Impact sensors are activated on an event that will cause a huge impact or when one of the sides of the motorcycle hits the ground. Additionally, a weight sensor determines whether someone is on the motorcycle or not.

The use of a gyroscope sensor and its provided data can aggregate useful information regarding the following events:

- The motorcycle has collided with another vehicle (sensors back and forth).
- The rider or the co-rider of the motorcycle has fallen (sudden loss of weight while the motorcycle is on the move).

- The motorcycle has fallen on one of its sides while moving (side sensors and speedometer).

The above-mentioned sensors are connected to a mobile processing unit $\mathcal{M}$ which process their measurements. The proposed sensors can give a highly accurate report by a biker's accident with a very low false-positive rate. The combination of the information gathered by additional measurements like the high speed and the long-distance crawl of the motorcycle can lead to estimate the severity of the accident.

3.1.4 Solution architecture

Everything starts after a motorcycle accident detection, something that can be detected based on the sensors embedded on the motorcycle. After the accident detection, our service allows $\mathcal{M}$ to contact a predefined emergency service $\mathcal{E}$ to report the incident. Immediately a number of reported data, like the ID of the rider, the GPS coordinates of the event, and its exact time are reported to $\mathcal{M}$. Additionally, the embedded weight sensor can help with the estimation of the number of riders.

To make our system even more efficient and based on the location data, a recommender system $\mathcal{R}$ is used by $\mathcal{E}$ to better select and inform the closer to the accident ambulances. At the same time, it can share metadata that could help the medical staff. Moreover, the architecture takes advantage of $\mathcal{A}$'s smart infrastructure and synchronises the traffic lights so that the ambulance will only meet lights which are turned to green while at the same time blocks other streets to reduce the traffic that the ambulance has to face towards the accident location. The proposed system will not only take care of the traffic lights towards the accident location but also while the ambulance is turning back to the hospital, by sending a signal to $\mathcal{E}$ to rearrange the traffic lights to deliver the injured biker in a much faster and safer way.

A recommender system will be responsible for approaching the accident first. Additionally, $\mathcal{R}$ could also provide several smart services, as follows:

- Aggregate live observational data regarding the traffic and the road quality,
- Collect information regarding the accident, such as the speed of the vehicle and/or the impact force,

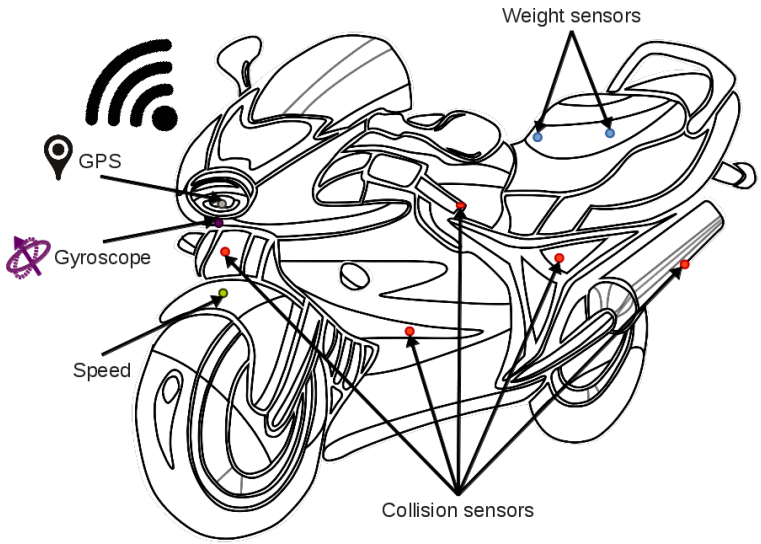

Figure 3.1: Proposed sensor positions [114].

- Collect sensors' information regarding the number of riders and other similar critical information that can help the corresponding ambulance to take the best possible decisions to handle and prioritize the accident management,
- Collect possible available information regarding the event like gas leakage or other metrics that can influence decision making to better manage the response to such an event. An example could be the need, not only of an ambulance but, also of a fire brigade.
- An additional and very useful feature is the ability to send a backup of its EDR measurements to guarantee that the data will not be lost.

Finally, to minimise the report of false positives, riders will have the ability to cancel a report by using a "break the glass" approach. Therefore, using their credentials to authenticate themselves securely, they could inform that they are not hurt and as a result flag the automated report by a manual action the sooner possible.

3.1.5 Simulation results

Towards the simulation of our approaches, we used the well-known SUMO road traffic simulator [88]. For our simulation experiments, a highly dense urban area was selected and more specifically, a place in the centre of Athens. Another useful tool we used to achieve our scope was the OpenStreetMap[5] (OSM). We used the OSM to get the map of Athens, and we specifically used the location close to hospital "Evaggelismos"[6], that is a well-known public hospital in the city centre.

Our main scope was to estimate the time that an ambulance would need to reach the location of the event and bring the patient to the hospital, with and without any management of the traffic lights. Our experiments consisted of routes in the area close to the hospital and included two journeys for each case study one with standard traffic lights and one with all the traffic lights turning to green for the experiment's route. It is expected that the latter time to be significantly lower, but the immediate question is; can we estimate the result and this result should be studied in combination with the the response time from the calling centre, then the final value and outcome of our scheme become apparent.

Some basic assumptions were made to facilitate our experiments. While we picked a central hospital, we simulated traffic like the one that exists in such places the night late hours. In our case that fitted well with cases where the accident happened far away of the selected for our simulations hospital. Nevertheless, in late-night hours it is common practice that only some of the hospitals are open 24 hours to deliver patients. On the other hand, our simulation situations and specifically the amount of traffic used to perform our experiments, are closely same to the situation existed in suburban environments.

Below the parameters, we used to our experiments that were used for the vehicle in our simulations that are typical for the simulation of a van in an urban environment and are illustrated in Table 3.1.

First we set the *accel* parameter of the vehicle to 0.8 m/s^2. Furthermore, the *decel* parameter, which indicates the deceleration ability of the vehicle. The deceleration ability, *decel*, was set to 4.5 m/s^2. The *sigma* parameter, which indicates the

[5] www.openstreetmap.org/, last accessed on 18/04/2021.
[6] https://www.evaggelismos-hosp.gr/, last accessed on 18/04/2021.

driver's imperfection is set to 0.5, so it simulates the behaviour of less deterministic behaviour for the driver.Additional inputs in our simulations were the inclusion of a vehicle of five meters long. We set the *minGap* parameter of this vehicles to 2.5. The *minGap* parameter is the space from the leaders back-bumper to the followers front-bumper. We also set the *maxSpeed*, that is the vehicle's maximum velocity, to 60 km/h. The maxSpeed parameter stands for the maximum speed of the vehicle. The actual speed that it would have is regulated by the maximum speed allowed for each vehicle by the edges of the map, which in our simulations are the actual speed limits in this city.

To better simulate our results, we made several scenarios where we randomised the timing and the states of those traffic lights. Nevertheless, in our random selection algorithm for the state of the traffic lights, we block impossible combinations, e.g. in a junction of four roads all red, all green, or three of them green etc. The number of vehicles used in our scenarios was 711 vehicles, which start and finish from random locations close to our hospital.

To achieve the most possible realistic results in our measurements, we created ten completely different trips where our vehicle followed different routes from and back to the hospital. In each scenario, the ambulance starts from the hospital and finishes at a randomly selected point, which would be the same for every repetition of the simulation. In each experiment, 710 random vehicles were generated and followed random routes better to simulate the traffic scenarios around our case's location.

The benefit on an average time derived by our proposed scheme was 42.54% on average of the original time and varied in the range from 20% to 77%. Table 3.2 represents our experimental results. The results have shown that the adoption of our approach could lead, to the reduction of the overall end-to-end time, even below the barrier of 50%. Additionally, compared to other recent attempts to improve the final performance of an emergency response system which not only do not improve the current situation but lead to the exact opposite result[7].

Motorcycles are vehicles that are more prone to accidents. As expected, our unique solution introduces a novel emergency response system for motorcycles

[7] http://www.theguardian.com/society/2014/jan/31/ambulances-taking-longer-reach-stroke-heart-attack-victims, last accessed on 18/04/2021.

Parameter	Value
accel	0.8
decel	4.5
sigma	0.5
length	5
minGap	2.5
maxSpeed	60

Table 3.1: SUMO parameters [114].

Scenario	Current state	Proposed
1	443.2	199.7
2	333.4	146.5
3	496.7	145.2
4	550.7	116.0
5	129.0	77.7
6	641.0	232.2
7	660.9	218.9
8	205.5	83.2
9	370.3	146.5
10	155.6	119.1

Table 3.2: Average time of the simulations [114].

that highly automates the procedure. As described above, even small delays can lead to casualties. One crucial issue is the existence of traffic jams that lead to delays in better accident management. In our scheme, we did not restrict our scenarios to bypass parts of the traffic through the use of tokens, but we have gone one step further and simulated the rearrangement of the traffic of the whole city and result to more secure and friendly accident management for all the participant vehicles to these scenarios.

3.2 Citizen-centric Smart Health services

The recent years there is a noticeable growth of services and applications that offer users the ability to track and manage their health conditions. This growth is not limited to smartphones [78, 47, 18] only, but it is also captured on the wearable devices market [34, 167, 19]. As a result, a lot of interest has been arisen by companies that led them to develop novel products related to health and wellbeing with the use of user-friendly and user-oriented components such as smartwatches or fitness tracking devices. Personalized and evidence-based medicine allows for the provision of personalized treatments and, when properly applied, it could help to reduce costs by enabling early release from hospitals, by fostering medication adherence or, by reducing relapse rates. Information and communication technologies (ICT) and, more specifically, mobile ICT are fundamental enablers of this shift towards personalized and evidence-based medicine.

The market today is mainly focused on the applications that a user can install to exploit several useful features. In this section, our interest is based on the different ways that a provider can deliver services to its users to help them minimize or avoid stressful situations and events within crowded places. Before our approach description, we performed research on the current situation on the mobile apps that are focused on stress management or reduction.

According to Muaremi et al. [104], similar smartphone applications applications can be categorized into four main groups

1. *Diaries* refer to applications that can be used to collect and aggregate data related to stressful situations.

2. *Guides* are applications that offer several tricks and tips on how a user can handle stressful situations. Diaries and guides can be found to be combined in a unique application for an application to be more informative and useful for its users.

3. *Relaxation* applications provide sets of relaxation exercises that can help users to manage their stress levels.

4. *Sensor measures* are applications that offer a sensor-based solution to users to track stress-related behaviors.

Several studies regarding the use of smart devices that examine stress levels [22, 104, 129], or regarding the relationship between stress and daily habits or activities of users that can affect their mood [102, 103, 35] have been published. In the latest years, an additional link of information coming from the data collected by wearable devices are under consideration to be included in the mass or personalized health services offered to users. Examples can be considered the EEG headsets or the torso wearables. Their interoperability with smartphones seems to be their strongest asset in today's market.

More precisely, there is a lot of research on the chemical analysis of body fluids, either invasive [67, 20] or non-invasive through e.g. patches or biosensing textiles [128, 44, 32, 21]. Also, many companies embed sensors to their devices that can measure galvanic skin response to monitor sweat changes and determine how sweaty one is.

Consumer wearable devices constitute an emerging market and they are becoming ubiquitous and widely accepted by users [141, 170][8]. Devices today integrate software able to measure, store and share users' biophysical and activity data including heartbeat rate, sleep efficiency, brain activity, physical activity, (*e.g.*, steps walked, stairs climbed, calories burned, calories intake, etc.). In many cases, health-related data are augmented with complementary *metadata* such as location information, measurements' precision, or measurements' interpolation policy; and it can be shared in general-purpose and specialized social networks.

The EEG approach has also been used as stress indicators, for instance for children with Asperger syndrome [154]. The experiments have shown that children with Asperger syndrome seem to have a greater reaction to a stressful situation. Even if EEG devices could be used in public as they might not be intrusive; still they are not discreet enough. The results for the users could be to make them feel more anxious, augmenting the negative impact that mobile devices can have such as stress, sleep disturbances or symptoms of depression [153, 77].

[8] https://ec.europa.eu/digital-single-market/en/blog/what-about-future-wearables, last accessed on 18/04/2021.

3.2.1 Use-case scenarios

According to our research goals, we will make some scenarios to introduce stressful situations in which our service is likely to improve the everyday life of people who suffer from stress and anxiety disorders in big cities.

Based on [121], the urban way of life has a significant impact on mood and anxiety disorders compared with rural areas. Two use case scenarios were performed to highlight the necessity of our proposed framework.

Bob as a very busy professional has to conduct several meetings across the city he lives in. Unexpectedly, some changes in his life led him to become very anxious and have a panic crisis. One of his symptoms' results, when walking in very crowded and noisy places, was that he had to stop walking, to calm down and focus to retrieve his body control. After he tried to identify the source of his symptoms, it turned out that the cause of his bad health is the noisy urban environment [146].

On the flip side, Alice is a teenager that lives in a less crowded place, but she has to cross very crowded and noisy streets on her way to school or to take part in sports activities, which usually make her feel anxious. This continuous stress that crowded places bring to her can potentially damage her health.

It is considered that both cases can not move away from the city based on personal, economic, and social factors. Moreover, as these symptoms can be temporary, such solutions are not the best.

3.2.2 Conceptual approach

One of the basic problems in the current state of the art is that the applications do not unlock the full potentials of the devices, nor do they provide advanced intervention. Our conceptual approach, depicted in Figure 3.2 provides a more holistic approach than the current state of the art. More precisely, the data is supplied by three sources:

1. the *user*, via data input and wearables
2. *urban sensors*, for instance, pollution measurements, traffic and crowd sensors, and

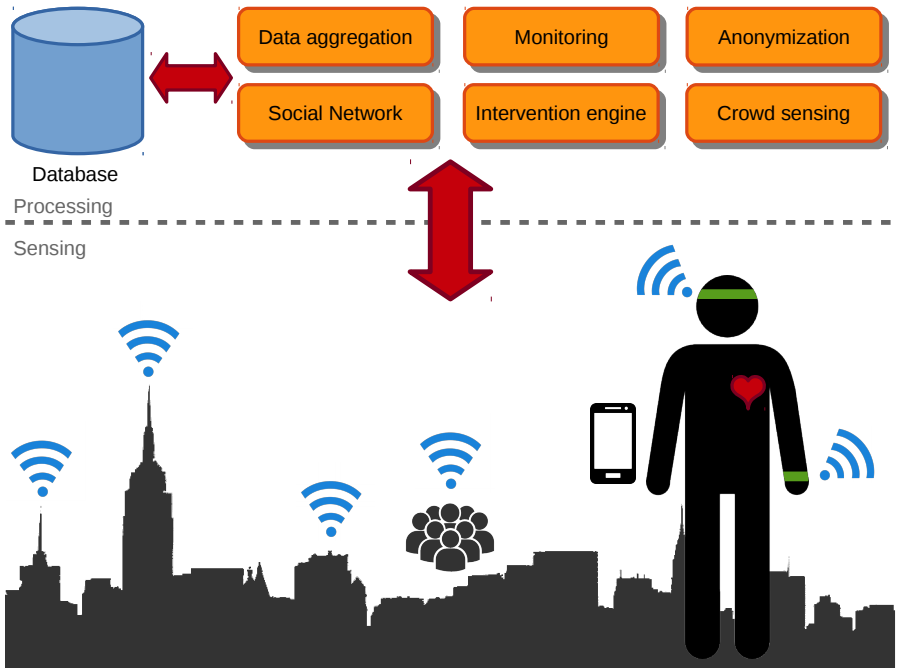

Figure 3.2: A conceptual approach of the proposed scheme [108].

3. *crowd sourced* data.

To specify the design needs of such a service, we firstly had to clarify all the involved data resources. The idea is to use personal data to determine the user's sentimental condition and use additional data sources for prevention. For example, the application can capture and determine via the user's wearables if the user is stressed according to the change in galvanic skin response, even if there is no actual change in his temperature or movement. The majority of applications have an informational role. This is where our conceptual framework is aiming to contribute the most.

In our approach, the application would suggest the reroute of the user, similarly to the concept of [117]. The idea is that the application will propose different routes to avoid stressful locations. Thus, the application would monitor a user's activity and routine to create knowledge and better learn the user's routes. Knowing the most commonly visited routes of a user, the application would try to get measurements like noise levels and crowdedness from the urban sensors. Then the application will aggregate them to users. According to these measurements,

the application will propose alternate routes to users. While the proposed routes might be more lengthy, it is promising that will offer a less stressful experience for the application's users.

Social networking features could bring value to the application. For example, for users that share their location, the application could notify their friends who are in proximity or inform them of events or even stories on media to relax the user. Also, the application could have the ability to change the music and its loudness for the users that are wearing ear-pads and listening to music.

Another feature of the application will be the provision of additional information between the users. Users, for example, will be able to share areas that made them feel distressed to notify others. Additionally, special functions will allow users to provide audio input to facilitate noise mapping or even images to indicate the most or less crowded places. The crowd sensing approach has already been used in psychological trials in projects like the "Track your Tinnitus"[9] with serious input for future research. Finally, the crowd sensing [60] aspect of the framework complements the sensing and prevention features.

Six main components are responsible for the proposed service provision, as follows:

- The *Monitoring* engine is running on the users' device and keeps track of their location, movement, routine, vital measurements, and preferences. Based on input data, the platform will determine the stress level of a user and decide whether an intervention is needed.

- The *Data Aggregation* engine is responsible for gathering the urban data collected by the smart city's sensors, aggregating them, and proving an insight to the platform regarding probable issues and alternate routes.

- The *Crowd Sensing* engine provides input to the platform from other users. Users, acting as 'sensors' would be able to provide useful insights into events and locations that common sensors cannot. The result would be a collective intelligence to the framework with valuable content for all participants.

[9] https://www.trackyourtinnitus.org, last accessed on 18/04/2021.

- The *Intervention* engine is providing the platform with proactive mechanisms in situations when a user might face distress. Several proactive mechanisms, such as music, social interaction, news, funny videos to games, or relaxing exercises, would be offered to users.

- The *Social Network* engine collects information from SNs and enriches the users' feed with events, news, and interaction with other users as the result of the intervention engine's functions.

- Finally, the *Anonymization* engine is the responsible component that ensures users' privacy. It is responsible to hide real users' identities in crowdsourced data as well as other sensitive information, such as their medical condition, their faces from images, or obfuscates users' location.

For interoperability reasons, the platform should be able to support several devices and protocols. While smartphones can consume and generate information in many formats based on their processing power, on the other hand, sensors have pretty limited computational resources. For these reasons, to make the platform compliant with the IoT approach and support as many protocols, especially lightweight ones, middlewares such as the Mule ESB[10], would be involved to facilitate these tasks.

In this line, research and development of health-related applications in mobile devices are becoming a hot topic. Being a very young research area, its primary goal is to improve many aspects of mobile devices that are not yet up to clinical standards, ranging from accuracy and reliability [25, 157] to security and privacy [53, 83, 107].

[10] https://www.mulesoft.org/, last accessed on 18/04/2021.

Part III

Social Networks Applications and Health Information

Chapter 4

Social Networks and Health Information Privacy

4.1 Introduction

The huge acceptance of SNs, after the Web 2.0 era, has transformed users from simple readers to writers and evaluators of online information. Users today can manage their social lives while they are able to share content with others and the public. Additionally, they have the ability to edit or even delete the shared content online. Nevertheless, the use of social media potentially can lead to security and privacy threats for users' data. It is not rare for malicious users to exploit software vulnerabilities of the infrastructure to gain access to personal users' information.

In [57] Boyd and Ellison define online SNs as:

> *web-based services that allow individuals to:*
>
> 1. *construct a public or semi-public profile within a bounded system,*
> 2. *articulate a list of other users with whom they share a connection, and*
> 3. *view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site.*

To highlight the dynamic nature of SNs, in [119] we provide an updated definition for SNs as follows:

Online Social Networks are web services that provide their users with mechanisms, subject to specific context constraints to:

1. *construct and manage the content and visibility of their profiles within their systems,*
2. *define and organize the type of connection with other users,*
3. *interact with other users, sharing content and information or even by altering their profiles.*

This new definition highlights some of the key ingredients of the online SNs, their dynamic nature, the interaction, the shared content, and the context, which were not highlighted in the previous one.

Today, SNs were not limited to desktop applications, but are becoming part of ubiquitous computing. Users can easily exploit SNs features via their mobile or even wearable devices. With the recent years' advances in mobile health (mhealth) and Mobile Health Networks (MHNs), users fulfil their need for context-aware social networking to share, exchange, and discuss health information with other patients or health professionals. Consisting of ubiquitous wearable devices like smart wristwatches, bracelets, rings, and hair caps, heterogeneous mobile networks (e.g., cellular network, WiFi, and device-to-device [D2D] communications), and powerful computational servers (e.g. cloud servers), MHNs collect the health information sensed by wearable devices, analyze/process for health monitoring and diagnosis, and enable users' social interactions [174].

The information that a user shares in an SN can be targeted by malicious entities unrelated to the user. Additionally, the attacker might be in the user's "neighbourhood", making the need for more complex security measures and for customizable privacy policies imminent, more essential than ever.

4.2 Health Social Networks

The latest years, Health Social Networks (HSNs) have been developed. HSNs was the beginning of a journey where patients could be the focal point in more patient-centred healthcare. Patients over the years were empowered through the usage

of such SNs with several advantages for them, arisen by their collaboration with other patients within online communities or directly with health professionals [55]. Nonetheless, these advantages didn't come without any concern. Some of the most common concerns that users of such networks are facing are the miss-trust of the shared information, the invasion of their privacy, potential security risks, and more.

On the other hand, users seem to trust for their daily interactions the most commonly used SNs that are not focused on a specific scope. General-use SNs like Facebook or Twitter are, also, highly used by users for health-related content distribution and sharing. While the use of general-use SNs is something that has an added value for users in terms of easiness and social engagement, at the same time these networks usually have not taken all the appropriate measures to control and protect sensitive data. Also, special measures that will protect health-related information and should be applied based on national or global data protection frameworks and laws are usually missing.

When dealing with health data, the data transparency should be managed in a special way to protect the identities of their data owners, as the main value that could be arisen by the data analysis should be greatly focused on the anonymous profiles and not to the real identities of their owners. Undoubtedly, images, including the user's profile photo, belong to the content that can easily expose users' identities, especially when these images include personal information.

In HSNs, patients can share their health condition with others, provide feedback about diseases and treatments to their doctors. Moreover, the feedback that users (patients or professionals) have by these networks can affect the referral of patients to other levels of health care or other vital processes in decision-making [97].

Based in [150] a HSN is a website where consumers may be able to find health resources at a number of different levels (meaning i. clinical trials access, ii. quantified self-tracking, iii. physician Q&A and iv. emotional support and information sharing). Services may range from a basic tier of emotional support and information sharing to Q&A with physicians to quantified self-tracking to clinical trial access.

There are many examples of HSNs that users can find online. While they may present some differences between them, like some special functionalities or different disease coverage among that their communities are focused on, in general,

they all have the basic characteristics of an SN. Some popular examples of HSNs are PatientsLikeMe[1], MedHelp[2], Inspire.com[3], HealthUnlocked[4].

Beyond the web-based HSNs, many ubiquitous devices and services have already provided new ways for users to interact and share their health conditions. Already HSN or new mobile-based applications keep many of the already features of SNs or integrate their features with well-known general-use SNs like Facebook, Twitter. The new ingredient in the HSN market is mainly focused on devices like smartphones and wearables. Moreover, Mobile Social Networks (MSNs) are a reality for the health market, while patients and professionals are already exploiting their context-aware features. Also, MSNs can enhance their services with functions by the well-known SNs or integrations [122, 71].

In this section, a summary of the possible attacks, the privacy issues, and the major security issues that exist to the recent HSNs, is presented. All of the section's content has a main focus on multimedia protection, as one of the most valuable content existed in SNS, that at the same time can expose users' identity. Additionally, there is an examination of the impact that such security and privacy issues can lead to and their possible countermeasures.

4.3 Attack vectors

In what follows, we report a list of the primary attack vectors that found to exist in SN, giving a particular focus on HSN and attack vendors that can also affect the multimedia shared on it.

Multimedia content : As multimedia files can contain sensitive and personal content and while this kind of content is between the most used and shared on SNs, can be considered a threat to the users. Moreover, in health-related SNs multimedia files can obviously contain very sensitive information, as they can represent health or medical results. As expected, this kind of information should be strictly protected by unauthorized access, alterations or other

[1] PatientsLikeMe, https://www.patientslikeme.com/, last accessed on 18/04/2021.
[2] MedHelp, https://www.medhelp.org/, last accessed on 18/04/2021.
[3] Inspire.com, https://www.inspire.com/, last accessed on 18/04/2021
[4] HealthUnlocked, https://healthunlocked.com/, last accessed on 18/04/2021.

attacks. Even minor alterations to the content of a medical multimedia file could lead to several wrong assumptions or potentially change the diagnosis results of a person. Thus the confidentiality, integrity and availability of the multimedia hosted in an HSN are more than critical, as they can cause real-life health issues to their owners.

Malware : Malware, a well-known way to perform attacks to users' devices. Examples include keyloggers, rootkits, ransomware and other types of malicious software may be used to exploit vulnerabilities of the user's operating system bypass any potential security measures and leak sensitive information to the attackers. Malware is also a very popular attack to mobile apps, and operating systems were attackers are exploiting their software design or security weaknesses to steal, damage or even encrypt users' data through ransomware malware attacks.

Misplaced Trust : In interactions where humans are involving, trust will always be an essential factor in exchanges involving risk [134]. Based on the lack of verification on the real identity, users are checking information measuring as trust metrics digital elements such as profile picture, friends in common, job descriptions or other elements that can raise some trust. Additionally, within the IoT environment, it is usual for devices to get connected and exchange data automatically based on criteria that in the most cases have been configured either by the vendor company or their admin users. In this environment, users are coming across with issues like the lack of ownership on data [27], either because of the architecture and nature of the IoT systems or even after an attack execution. So, trust within such an environment is such a critical thing to succeed sustainability and trust. Focusing on SNs, when you add a new friend on your list you also, based on your privacy settings, give him access to information and multimedia content that they could potentially be used for malicious actions.

Phishing : Phishing, is one of the most well-known attacks on the web. On SNs, phishing is included under the social engineering umbrella, as the attacker is usually pretending to represent a legitimate and credible entity that the

victim trusts. Popular results of a phishing attack can be the credentials of users, for example to e-banking, emails or SNs accounts. But the latest years, phishing is possible to lead to more sophisticated attacks like ransomware attacks. A typical example includes a scenario where the attacker sends a message pretending to be a well-know trust sender or after highjacking a trusted account and the victim downloads, for example, a zip file that includes the 'useful' for him files. When the victim tries to extract the zip files the malicious ransomware software locks the operating system and encrypts all the user's data on his personal computer and only a screen can be reached with some directions on how to pay the attacker (usually in bitcoins) in order to unencrypt his files.

Hijacking : When talking about an account on SN then hijacking is when an attacker breaks into the account and impersonates the owner usually to run a scam or to harm his reputation. But in case of IoT devices and SNs in more mobile environments, hijacking can potentially lead to several unwanted results as there are much more info and data that an attacker can use to cause more significant damage to the victim. Data arisen by the ability to exploit or extract additional info by the embedded sensors on a mobile phone, soy by the use of its embedded camera, listen through the microphone or capture or/and share user's location are just some of the attacks that can be performed to a mobile device.

URL redirections : This category of attack is widespread on SNs and in summary, is a short domain name followed by a short unique string that is linked to a long URL. As the real destination is not visible to the user, therefore in case a user trust this link or accidentally click on this short link he can easily become a victim of scams, malicious sites, or other sites that the user did not intend to visit.

Lack of Policies : Even if the current situation on the most famous and popular SNs, like Facebook and Twitter, has been improved in terms of policies by the years, due to the wide range of possible scenarios of human interactions, their security can, still, be exploited by malicious users. Moreover, in most

cases, the design of their integrated services are missing the best practices of privacy by design and privacy by default, two mandatory requirements of GDPR. Additionally, SNs that provide social features for special categories of networking (like dating, hobby-wise) and the new wave of apps that offer SN features for mobile users usually lack policies that would protect when needed their users. An example is the lack of policies for content re-uploading that is not handled by any SN policy, exposing users greatly to attacks like impersonation attacks.

Software vulnerabilities : SNs are software-based platforms and apps. Due to their very dynamic nature and rapid evolution, it not rare that several bugs can be exploited by malicious users to gain access, bypassing users' privacy settings to steal personal data[5]. For this reason the majority of the big SNs, maintain bug bounty programs to award hackers that find and responsibly reports bug for their applications.

Open access : The majority of SNs is based on the "freemium" model that let users create their account easily and free. Users' authentication is mainly based on their email address verification or in some case to other also free services. This tactic is not, as expected, as strict as it should restrict any potential malicious user for taking access to an SN and start exploring attacks like the above mentioned.

Misinformation: Based on [169] the term of misinformation in social media can be used as an umbrella term to include all false or inaccurate information used on these channels of information. Attackers that are spreading misinformation are aiming to avoid being detected and frequently are targeting certain groups, ideas, or facts. The motivation of an attack can be differentiated. Some example could be the following; i. spammy economical potential benefits through manipulating information of other brands, ii. the spreading of fake news between users, that is a very usual phenomenon on SNs, iii. trolling among groups of users that are spreading inaccurate information to others, iv. media competition can also be a source of misinformation and more.

[5] http://www.neowin.net/news/facebook-photo-exploit-allows-you-to-view-any-albums-of-non-friends, last accessed on 18/04/2021.

4.4 Privacy exposure categories

In many cases, SNs treat users like their products, as their primary income source derives from selling users' preferences to advertising companies. On the other hand, SNs are usually bringing responsibility to the user by having a section where they documented the end-user license agreement. So, SNs assume that users agree to this policy, even if they have never read it or in some cases even saw it. Today GDPR has brought several functional and non-functional requirements in order for the SNs to document that users have been informed about the context of the policies that govern an SN.

In this section, a list of privacy exposure categories that can be found to health SNs applications will be presented.

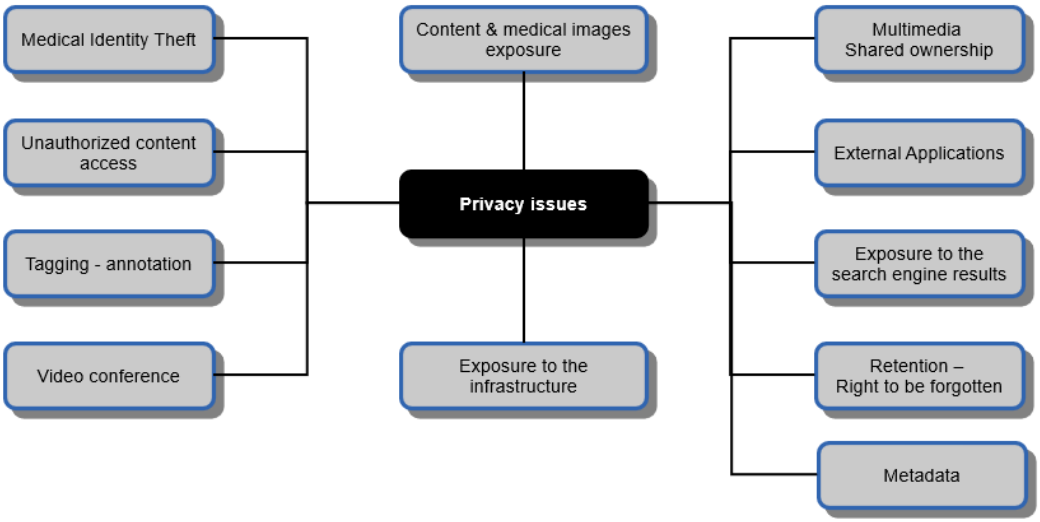

Figure 4.1: Privacy exposure categories.

4.4.1 Content and medical images exposure

It is well-known that users on SNs, even if they are trying to be careful about the content they publish, in the end several personal related information can be leaked to strangers or even publicly on the web. In case of health and medical-related

content like texts that include conditions or prescriptions, the way that the SN is developed and the fields where the user should fill in his sensitive information, is critical from a data protection point of view. Additionally, health-related multimedia is one of the most sensitive content that users share or store within health SNs and that in most cases there are not any special security measures to protect such content by unauthorized access.

SNs users can easily become a subject of profiling, something that can be achieved from the shared content and much sensitive information based on [79].

4.4.2 Medical identity theft

Medical identity theft consists of a crime that can have a significant impact on the victim either financially or even by the harm of the data owner through other harmful activities like false and erroneous entries in their medical files [54]. An attacker targets data, like health insurance information or social security number to obtain medical services or goods, or to obtain money through the malicious unauthorized use of the victim's data. This kind of attacks is a rising concern mainly in healthcare institutions that, as expected, keep a large amount of data and also are responsible for offering services to specific persons based on their identity. Medical identity theft based on the many different entities that are involved can be characterized a complex crime. As a result, a collaborative effort among individual victims, health information management technologists, institutional security officers, law enforcement, healthcare providers and payers is required in order to collaborate and exchange information to fight and succeed the limitation of this phenomenon [95].

4.4.3 Metadata

Metadata are data that are used to describe other data, usually used to help organize web resources, provide digital identification, and archive and preserve resources. Between the web professionals, "meta description tags" are very well-known for their ability present a brief summary for the search engines to include on their results when a user performs a search query about something relevant. As their main goal is to provide content, it is expected that in case of sensitive information existence on an SNs, metadata can be used by malicious users to exploit information

about their potential victims. Depending on the SN[6], multimedia metadata can also present information regarding the location, the clinic, the health condition or other sensitive information that follows a health-related multimedia file.

4.4.4 Unauthorized content access

Content that is shared through an SN is accessible, usually, to a specific group of people, based on privacy settings that its uploader has given. Nevertheless, the way each SN has applied its policies and security mechanisms, could let malicious or in certain circumstances even advanced users, access content through tricks that apply by using their browsers. From changing some URL parameters to study and reproduce some specific SN application's behaviour, malicious users may expose other users' private content or even in some cases hijack accounts as security researchers claim[7].

4.4.5 Tagging - annotation

A feature that SNs bring to users in the latest years is the tagging of other users to shared multimedia content. While this works fine for situations where friends are tagging each other, still is a privacy issue when that happens from malicious users that tag unsuspecting users, usually to share a spammy message as part of a broader phishing attack[8]. On the other hand, some users don't want to get tagged by others, even their friends, especially when this tag is taken place to a publicly accessible content, as they don't want to get crawled by search engines or their username to become searchable inside the SN the other users worldwide.

[6]Example of how different each SN can manage multimedia metadata: http://www.embeddedmetadata.org/social-media-test-results.php, last accessed on 18/04/2021.

[7]Facebook flaw could have allowed an attacker to hijack accounts, https://nakedsecurity.sophos.com/2019/02/19/facebook-flaw-could-have-allowed-an-attacker-to-hijack-accounts/, last accessed on 18/04/2021.

[8]https://www.thesun.co.uk/money/9323475/facebook-scam-clearance-sales/, last accessed on 18/04/2021.

4.4.6 Video conference

A popular way of communication among SN users, apart from messages, is by using video conference tools. While this might allow more interaction between users, patients, and doctors, the problem that arises is that more information can be leaked. One security issue is that the video stream could be intercepted over the network. Additionally, possible vulnerabilities in the protocol, or malware could allow the attacker to arbitrarily access the camera and microphone of the victim without notifications. Moreover, the conference call could be easily stored by one of the involved parties to either extort the victim or to manipulate the content and present it accordingly.

4.4.7 Multimedia shared ownership

While SNs can also be described as a multimedia sharing heaven, at the same time, some complex legal issues have arisen. One of these issues is the shared ownership of the multimedia content. For several reasons a multimedia file may belong to more than one users. Such a case is, for example, a family photo or a photo that presents a couple. In the case of health and medical-related content, an example case study could involve the multimedia shared by a doctor, including their patients or even the opposite. However, inside an SN, at least for the majority of the SNs today, the ownership is assigned to one user that is responsible for this content in terms of privacy settings, tagging, sharing settings and more.

4.4.8 External applications

SNs have developed over the years special tools for a developer to exploit functionalities through their own created apps, or through APIs that allows developers to request data from an SN that can be used to external apps (for example websites, mobile apps) based on the needs of each development project. Nevertheless, researchers have proven that malicious applications can be developed [112]. Additionally, one of the greater privacy scandals in the history of SNs, named Cambridge Analytica, was mostly based on the ability of the developer to exploit personal data by the use of external apps inside on the bigger SN, the Facebook.

The big issue is that in the case where data are transferred outside an SN, users' privacy, can't be retrieved and in most of the cases (especially when user don't have clear information on who govern his data) the user or even the SN itself has already lost the control.

4.4.9 Exposure to the search engine results

The majority of SNs allow search engines like Google or Yandex to crawl their content (mostly their users' uploaded content). While this tactic is essential for the information and knowledge that is mined by the public through their queries to search engines, on the other hand, it is usual that users that don't have advanced knowledge on how or even why to adjust the most private settings on their data, can easily be exposed to the search engines results even without creating an account to the SN.

4.4.10 Right to be forgotten

One of the practices that SNs used but still can be found to exist (mostly on smaller SNs) is the difficulty for a user to delete his account. Firstly, as expected, there is a need for SNs to say that they maintain as biggest user database as they can, so it is preferable for their profit model. More recently GDPR, introduced even more clear and strict directions for the so-called "Right to be Forgotten" for data processors like SNs. So the deletion policy on an SN should be clear, the deletion progress easy to be followed and should be part of the user's rights. Not surprisingly some any SNs either prohibit users from removing shared content, or they provide the facility with some obstacles (time frames, for example a photo will not be immediately removed, or users have to pay to remove content, as mentioned on MedHelp's Terms of Use[9]: "If you disregard this warning and post personal or confidential information (yours or others) on the Website, which you later want removed, there is a fee of up to $25 to remove each posting.").

[9] https://www.medhelp.org/legal/terms_of_use, last accessed on 18/04/2021.

4.4.11 Exposure to the infrastructure

Part of an SN in order to be functional is the infrastructure that maintains in order to provide services to its users. Some SNs choose a renting model where maintain third party infrastructure or host their applications based on well-known cloud solutions. Infrastructure as a Service (IaaS), Platform as a Service (PaaS) or Software as a Service (SaaS) is some of the different available models that SNs can use to deliver services to their users. Of course, there are cases that SNs choose to maintain their completely own infrastructure. The policies that a third-party hosting or infrastructure provider applies are influencing the final privacy and security compliance of an SN. While SNs are profiling their users in order to become profitable by serving to them personalized ads, they are a major target for hackers, cybercriminals or even for governmental spying. As expected, there are cases where the infrastructure is responsible for several leakages and privacy incidents. Recent disclosures about the role of secret agencies in the Internet[10], brought to light a series of such spying incidents.

Moreover, the more service providers are used by an SN the most complicated result may be delivered to the final policies that users should read, understand, and finally consent to, to use a service. For example, Google Plus, which was part of the Google services and was terminated in the same period that was suffered by a security-related incident[11], informed its users about their multimedia files:

> *"When you upload or otherwise submit content to our Services, you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content. The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones. This license continues even if you stop using our Services (for example, for a business listing you have*

[10] http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data, last accessed on 18/04/2021.

[11] https://www.wired.com/story/google-plus-bug-52-million-users-data-exposed/, last accessed on 18/04/2021.

added to Google Maps). Some Services may offer you ways to access and remove content that has been provided to that Service. Also, in some of our Services, there are terms or settings that narrow the scope of our use of the content submitted in those Services. Make sure you have the necessary rights to grant us this license for any content that you submit to our Services."

The result is that while the provider is trying to simplify its policies, users' content can be processed and stored to so many third parties that at the end is not easy (even for the provider) to know where are his data in a particular time or how many different providers have taken part in their processing. Having GDPR inactivation, the later is more than a demanding requirement to be feasible, as the user can request to be informed about the processors that process his data, or to edit, delete and take a copy of them without delays.

4.5 Security issues

In this section, the results of the research regarding the possible security issues on SNs is presented.

4.5.1 Unencrypted traffic

Someone would expect that years after many different campaigns and actions[12] towards the enforcement of HTTPS over the web, also the rise of many tools such as Firesheep[13] which proved how easily is for an attacker, not only to read everything that is transmitted unencrypted but even hijack sessions on social or other personal accounts, the majority of popular SNs upgraded their infrastructure to work only over HTTPS. Nevertheless, as our research [107] findings present is not unusual for mobile apps to send or request personal data over HTTP. The sensitive nature

[12] The Electronic Frontier Foundation had already warned the Council of Europe for the lack of SSL/TLS adoption from SNs and the impact to the privacy of their users, https://www.eff.org/node/58437, last accessed on 18/04/2021.

[13] http://codebutler.github.io/firesheep/, last accessed on 18/04/2021.

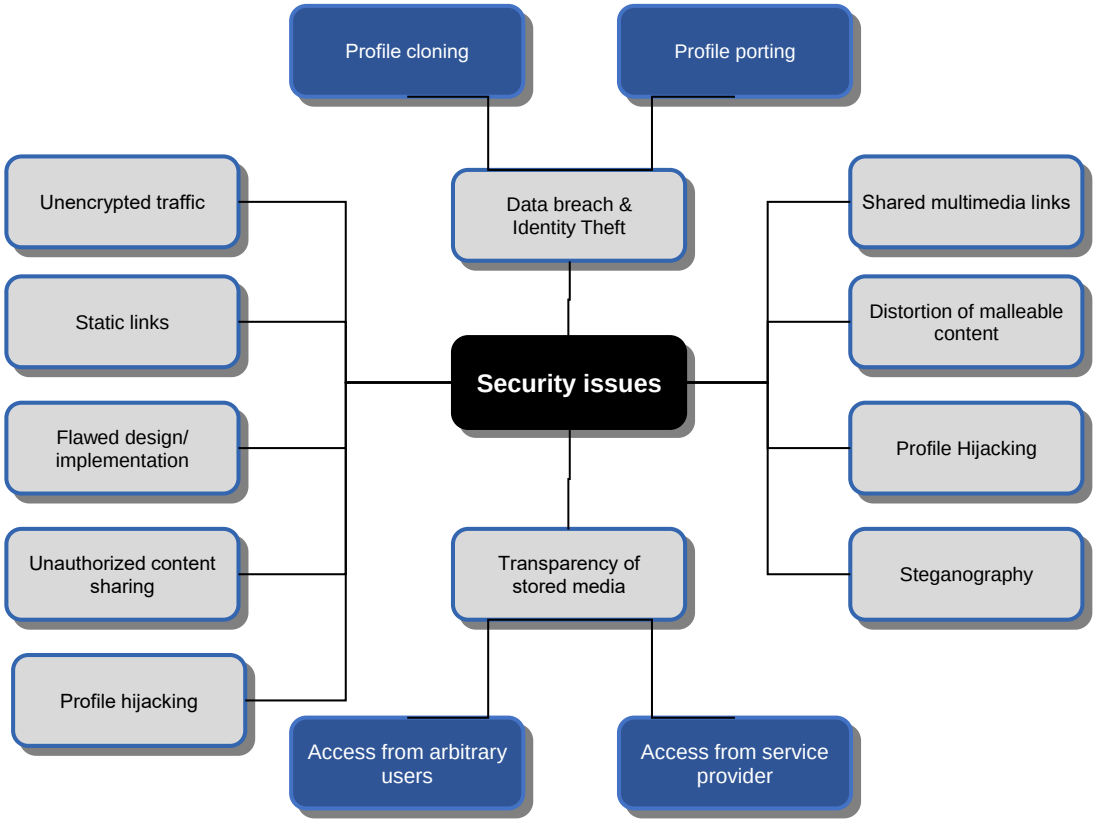

Figure 4.2: Security issues.

of the shared content that can be health conditions or medical exams in combination with the use of unencrypted traffic can lead to additional security and privacy issues.

4.5.2 Static links

A major security issue is the way that web applications, and in our case SNs, are hosting users' multimedia. It is usual for attackers to study the way that an SN works in order to expose the personal data that users host on their profile. History saw to us that even the biggest SNs had failed to ensure the privacy of their users' multimedia. For example Facebook and Google Plus use static links to host

their users' photost[14]. So, as expected, it is not difficult for an attacker to find, send outside or store the photos of their victims. In such cases, SN doesn't provide the expected security measures that users think that exists.

4.5.3 Sybil attack

In a Sybil attack, a user creates multiple accounts to manipulate and affect a result as desired by him, and his purpose [56]. Sybil attack is a node that claims multiple fake identities, and that can be highly harmful to the real users of a health-related application. Adversary attacks vary from a simple voting scenario to a de-anonymization attack to users.

4.5.4 Flawed design/implementation

Application developers and/or publishers seem to keep repeating the same mistakes over every new software environment[15]. While they guarantee on their applications' security and privacy that provide to their users (usually through their policies included on the application content), common pitfalls in the application side can lead to jeopardizing the privacy rights of millions of users.

In case of health-related SNs, where users are trusted SNs applications to share their sensitive data with other patients, their doctors or even use them as agendas to store data for their future reference to their doctors, the impact on user privacy can be much more significant than in an SN that is used for social-only related interactions. The difference is focused on the sensitivity of data and the impact that can cause in case of a data breach. Of course, we can easily admit that a data breach that includes a combination of social-related personal data combined with health-related sensitive data can occur even more damage to users' privacy.

[14] https://www.neowin.net/news/facebook-photo-exploit-allows-you-to-view-any-albums-of-non-friends, last accessed on 18/04/2021.

[15] https://cryptacus.cs.ru.nl/slides/How%20private%20is%20your%20mobile%20health%20advisor%20-%20Free%20popular%20m-Health%20apps%20under%20review.pdf, last accessed on 18/04/2021.

4.5.5 Transparency of stored media

As it was stated in the intro of this section, the majority of the social media that are used for networking, like Facebook and Twitter, are heavily promote the power of data sharing between people that is also the result of the transparency that these networks bring to our lives. Nevertheless, when we are handling health or/and medical-related multimedia content, then the sensitivity of this content impose their protection against unauthorized access. On the contrary, it is usual that SNs' stored multimedia contents are not encrypted. The result of this practice allows anyone that has a direct link to them to be able to access this content without the use of any credentials, bypassing any privacy or security policies set by the user with the use of the SN.

Another issue is the transparency towards the service provider. While a big company like Google or Facebook might have the ability to maintain their own data centres, nevertheless, smaller ones do not have this luxury, so they resort to outsourcing their data centres using virtualization or cloud-based technologies. These technologies might reduce scalability and maintenance costs. However, many concerns arise regarding their provided security [16],[17]. This usual practice leads to the phenomenon where the end-user might trust the SN, but not the cloud service provider which has access to his data[18]. The issue becomes even more thorny due to geospatial and political constraints. It is not unlikely that governments and agencies may be granted arbitrary access to foreign citizens' multimedia content without their approval or any kind of notification, as the data centres that host this information do not belong to the same country or even continent.

4.5.6 Profile hijacking

This attack can be achieved in many ways such as brute force attacks, phishing or social engineering. An attacker can use the information that the victim's multime-

[16] https://cloudsecurityalliance.org/research/working-groups/top-threats/, last accessed on 18/04/2021.

[17] http://www.enisa.europa.eu/activities/risk-management/files/deliverables/cloud-computing-risk-assessment, last accessed on 18/04/2021.

[18] https://it.slashdot.org/story/13/09/11/1657228/the-windows-flaw-that-cracks-amazon-web-services, last accessed on 18/04/2021.

dia provide to him in several harmful ways. For example, someone can find crafted tools such as CUPP (Common User Passwords Profiler)[19]. By using these tools the attacker can give some words related to the victim's profile, and the tool will provide him with a curated dictionary for possible user passwords. Weak passwords are one of the most usual to acquiring control over a victim's profile.

4.5.7 Data breach and identity theft

Identity theft is an attack where a malicious user might aim to mislead other users that he is another user by taking over the victim's account. It is usual that this kind of attack is a result of a data breach. The attacker usually targets to cause reputational damage or to exploit the trust that other users have in his authority and obtain money or credit. The nature of SNs can enable malicious users to automate such attacks [26]. The attacker can also deduce a lot of information to use in their attacks from the shared multimedia content. Moreover, since the shared multimedia contents are usually of high quality, they can be used to launch attacks in real-life, e.g. print fake ID cards or company passes.

4.5.8 Impersonation

These attacks are about the creation of fake profiles with the scope of deception of other SN's users. Attackers are expecting other users to be misled and connect with the malicious profile and start building some trust with them based on the fake information that the profile presents to them. These attacks can lead to an extensive range of attacks. Spam and phishing attacks are some of the easily performed attacks after building trust with some users based on a fake profile. Moreover, attacks like these are used even for crimes, sometimes not limited to the online world, but result in dangerous real-life crimes.

The protection of the user's identity against several attacks that can lead to impersonation attacks, as well as the profile cloning and profile porting attacks should be a priority in SNs. Profile cloning consists of the replication of the victim's profile in the same SN. Profile porting is the export of the personal data of the victim and

[19] http://www.remote-exploit.org/articles/misc_research__amp_code/index.html, last accessed on 18/04/2021.

the creation of a fake profile to another SN. Both profile cloning and profile porting are including in the category of impersonation attacks [45].

4.5.9 Distortion of malleable content

As expected, multimedia content shared on SNs, can be malleable. A simple user today can find a wide range of powerful tools available for image or audio processing. From tampering images maliciously to the production of high-resolution fake images that can harm or derogate their real owners. Such tampering is more or less expected for photos; however, new SNs such as Clubhouse[20] could be an audio source for attackers. Within the content and scope of health SNs, this content can lead to really harm patients as it can potentially make them take wrong decisions regarding their medication.

4.5.10 Shared multimedia links

Based on the need that users want to share links to internal or external to the SNs content, it is not feasible to just disapprove such functionality. Additionally, it is one of the most powerful ways for users to share sources of knowledge or news. Nevertheless, there is a wide range of formats that can be vulnerable to attacks. Examples of the formats that users share are PNG, JPEG or GIFs. Given that users are redirected, with their will, outside SNs, it is not difficult for attackers to trick them into installing malicious software (i.e. using clickjacking techniques) or visiting sites that perform cross-site scripting (XSS) attacks which attempt to steal client cookies, highjack sessions etc. Furthermore, another feasible scenario is that users could be redirected from their original link destination to another page that could contain harmful content.

4.5.11 Steganography

Hiding information in other media not only to hide the content but even to hide its existence is a pretty old practice on the web. The latest years, the advance of

[20] https://www.bloomberg.com/news/articles/2021-02-22/clubhouse-chats-are-breached-raising-concerns-over-security, last accessed on 18/04/2021.

technology made steganography a really well-known technique. That is why this technique can also be found on SNs to cover malicious activities.

An example of steganography usage can be found on [158]. While the aforementioned work is focused on providing users with more privacy, nevertheless it indicates that it can be exploited by malicious users. As a result, embedded messages can range from terrorist messages to child pornography.

4.6 Impact

In this section, we summarise the possible impact of the attacks mentioned above both for the general-use SNs like Facebook & Twitter that users still use to exchange and discover health information [151, 24] and the health-related SNs.

Information leakage A major risk of SNs regarding privacy is the information leakage inside and outside users' network. While users may control information exposure through their privacy settings that each SN provides, this is not always the case.

General-use SNs: Given the amount of multimedia information that is shared, a lot of sensitive information can be inferred with great accuracy [87] or through data fusion with other users [90].

Health SNs: In case of health SNs images are expecting to be sensitive by default as in their majority are connected with health or medical information. Thus, their protection through special mechanisms should be a priority by SNs.

Location awareness In a worldwide connected mobile world, location is a piece of information that may characterise users' profiles. It can also be used to improve user experience significantly or by advertisers to target their audiences better.

General-use SNs: Location information on SNs like Facebook, is mostly related to information the user submits to his profile, and it is mainly an issue of protection by malicious users that could exploit this information for

criminals. For instance, location information can be used by buglars[21]. Nevertheless, in the past few years, the advance of mobile technology brought capabilities for automated processing of data, without, usually, user's concern. Applications and advertisers can easily exploit location by the GPS of the user's device mainly for profitable growth.

Health SNs: As expected on health-related SNs and especially in the mobile era, users' location can be combined with several other data that users provide or SNs automatically exploit through several mobile's sensors. Especially Location-Based Social Networks (LBSN) have emerged, and researchers have already focused on how they can advance on their knowledge regarding the health behaviour of users [36]. Nonetheless, based on laws like GDPR, users should provide an explicit consent to collection and analysis of their location by SNs. Moreover, their third party processors should have taken all the possible measures to protect users' data against malicious practices or malicious advertisers that are trying to exploit location for their profitable goals.

Reputation The amount of data shared by one user on SNs can be used by other users to target him and harm his reputation.

General-use SNs: In general-use SNs the category of data shared it is to be mostly general data related to behaviours or opinions that a user exchanges. Nevertheless, there are plenty of incidents where users were targeted in their professional or other environments based on their content shared on SNs.

Health SNs: By default, health SNs is expected to require apart of personal data, the sharing of health or and medical data that are characterised as sensitive. Malicious users that will want to harm the reputation of a person it is possible to use such data to perform attacks.

Account loss Some of the possible attacks on SNs, can lead to account lock or even its loss.

General-use SNs: Account loss on a general-use SN like Facebook or Twitter can be a great issue for the everyday activities of a user. Of course, it depends

[21] https://www.pleaserobme.com/, last accessed on 18/04/2021.

on how much each user makes use of the SN to manage contacts and organise his life (personal or professional).

Health SNs: In the case of health data a user can have a great impact by an SN account loss, as this account could be the repository of data used as a personal health records application. In such cases, users may lose their historical medical history, and that could potentially have a significant impact on their health.

Loss of ownership/control of content It is almost impossible for someone to predict how popular will the content that he publishes in an SN will be, as other users (friends or friends of friends) can download, share in and out the SN. Therefore, users should be very rather cautious and responsible for what they share.

General-use SNs: On SNs like Facebook, malicious users can perform several attacks based on shared content like identity theft, social engineering and more. Today SNs can, also, be used to authenticate users for example, to other services through social login or to our network of contacts inside an SN (friends, colleagues etc.). In such case multimedia protection from SNs is more than critical and can if an SN can succeed it, its users will safely distribute more content without the risk of their privacy.

Health SNs: On health SNs and based on the level that a user or already patient count on it to store and share his health data with other patients or his doctors, the control of such a sensitive content is something more than critical. Thus, any alteration or loss of users' data is a major security incident. Additionally, the debate on the ownership of data in such cases is huge as many voices fight for the free anonymous sharing of health data towards the advance of health research, but on the other side, several concerns exist on users' privacy. Moreover, the ownership of data is still under debate [126, 82].

Blackmailing/extortion A malicious user that wants to perform blackmailing attacks against users, it is expected that the most sensitive the content that will gain access the most harmful will the attack be.

General-use SNs: This trust makes them share a lot of sensitive or even embarrassing content, which, if leaked, can be used as a threat. The threats, depending on the attacking nature, can impact his financial, sexual, professional or social status.

Health SNs: The sensitivity of health data that users exchange or store within a health SN is the main reason why health SNs are one of the main targets of malicious users. The impact on these categories of networks is definitely more significant than in general-use SNs as the value of exchanged information is directly connected to their users' health, and it is expected to impact their lives.

Cyberbullying Cyberbullying is the situation when a user is threatened, humiliated or harassed online by another user.

General-use SNs: The phenomenon of cyberbullying is most often and has a greater impact on young ages. Also, there are several cases where children have to lead to extreme acts of violence or other children to depression and suicidal ideation [42]. Researchers have shown that the older the student, the more likely he/she was to bully others or to both bully and be bullied online than to be neither a bully nor victim [100].

Health SNs: Cyberbullying is mostly connected with teenagers and health-related SNs and sources of content is mainly connected to adults than children [127]. Nevertheless, the continuous intrusion of smart devices (like smart toys, wearables) that provide connectivity to younger ages, it is expected to bring to light health-related cyberbullying incidents the next years.

Cyberstalking Social media is the place where someone can find what users share globally and learn more about them and in most cases, about their daily activities. As expected, the information shared on SNs could be used to stalk or harass the victim in real-life attacks [124].

General-use SNs: On generic-use, SNs users share news and opinions, personal files and personal data. The user-generated content shared within SNs can be categorised in a great range of information that includes, for example,

social-economical, financial, personal, sexual, political, religious, even health-related data. This large scale data sharing transform SNs to one of the most powerful tools for cyberstalking by malicious user and criminals.

Health SNs: Despite the fact that health data can also be shared through generic-use SNs, the powerful tools and ways that health SNs offer to their users to organise their health and medical data with their use is an excellent source for cyberstalkers when their subject of interest is focused on health-related issues. The findings can be related to the health condition of a user, his symptoms or historical health events.

4.7 Towards Privacy Enhanced Technologies and tools

The goal to succeed privacy-oriented design and development within complex projects and systems like SNs or mobile apps that deliver content and data through several different channels and networks led to the need of new modern tools which could ensure principles like data anonymization, pseudonymization, data minimization, user authentication and data encryption. The term Privacy Enhancing Technologies (PETs) is used to cover this wide range of technologies that are designed to deliver and support privacy and data protection.

Searching for such solution on SNs, someone can find several examples that have been proposed by and delivered to the community, some of which are presented in the following paragraphs.

Persona [17], offered users the ability to encrypt their data and only exchange a public key with authorized users. In this way, attribute-based encryption was offered to users' data, that also bring them more control over their privacy setting among their shared data. EASiER [75] was a Persona-based extended idea, that used the creation of decryption keys, associated with each user, allowing data access, only when a user uses the appropriate key to content the proxy. FlyByNight [92] was an encryption tool that proposed a trade-off between security and usability in the interests of minimally affecting users' workflow. Its main goal was the use of public-key encryption algorithms to exchange users' messages on Facebook. There also browser-based paradigms that were proposed to enhance privacy

through existed SNs. Scramble [23] is one of those solutions. It was a Firefox extension which allowed users to encrypt their previous uploaded to an SN data, storing it either at a TinyLink server or the SN.

Another proposed SNs focused PET solution was the PrivacyJudge [86]. PrivacyJudge allowed users to manage who can access their shared content. To deliver this service, it hosted users' content or used a trusted third party server.

Towards the provision of privacy tools to users, someone can also found Lockr [155] an access control system and Facecloak's [93] which act as data pseudo anonymization tool for users' profiles by providing fake information to the SNs and storing personal information on an application server in encrypted form. Patsakis and Solanas [115] also proposed a privacy solution for SN users. Their proposed service could encrypt all of the users' data, and by creating small encrypted keyword dictionaries, the service will offer more control on the data that they are willing to share. Sharing the dictionaries' decryption keys with advertising companies, users allow them to mine their data. In this way, only if advertisers find a promising profile, they can place a bid to access the full data.

Other solutions were mostly based on social trust and the connections between users. An example can be found in [50], where the use of a private set intersection (PSI) protocols to disclose only the common connections that two users have was proposed. Another data minimization solution was proposed by Li *et al.* that introduced a recommender system for SNs, which matches users with similar interests, without disclosing their preferences [91].

Based on other privacy principles like the right to be forgotten and similar to the right to restrict processing X-pire! [16] allowed users to set expiration dates for their shared multimedia content to make them unavailable after that date. On the other hand, unFriendly [152] proposed a solution to bring multi-party privacy in published photos so that they could co-managed by the people who are depicted in them.

Moreover, there are some completely decentralized SN architectures examples like Diaspora[22], Safebook [46], OneSocialWeb[23] and NYOB [66][24] that based their

[22] Diaspora, https://joindiaspora.com, last accessed on 18/04/2021.
[23] OneSocialWeb, http://onesocialweb.org/about.html, last accessed on 21/12/2014.
[24] NYOB, https://noyb.eu/, last accessed on 18/04/2021.

scope on the way they could decentralize data architectures to change the rules on privacy and empower users' role on data protection decision.

Finally, one project that delivers some interesting open-source tools for consumers and businesses is the OPERANDO project[25]. The OPERANDO project aims to implement and validate an innovative privacy enforcement framework that can deliver; Privacy as a Service (PaS). OPERNADO includes a platform that will be used by independent Privacy Service Providers (PSPs) to provide comprehensive user privacy enforcement in the form of a Privacy Authority. The OPERANDO platform will support flexible and viable business models, including the targeting of individual market segments such as public administration, SNs and Internet of Things.

PlusPrivacy[26] is a product that was delivered as part of the OPERANDO project and provides the end-users with an anonymously used unified dashboard for their protection, against a variety of privacy threats. Part of its features is the easy control of the most privacy-related setting in SN accounts, the ability of a user to hide his email identity and use alternative in the concept of pseudonymization, the adblock feature as well as the trackers and malware. Additionally, Privacy Plus offer the prevention of unwanted apps and browser extensions from tracking users and collecting private data.

4.8 Countermeasures

Recently and after 25th May 2018, GDPR enforcement introduced several upgraded requirements regarding the provided privacy and security that offers to their users. GDPR demands by service providers to clearly define their role and choose, based on their responsibilities, whether they are the data controller, the data processor or the joint controller. The main role of SNs is to act as data controllers and ensure their compliance with the GDPR requirements. Nevertheless, the business models that SNs have introduced by the years have created some complex schemes that can lead an SN to act also as a data processor in some cases or as a joint controller. Below some possible, technical and organizational measures that could potentially

[25] OPERANDO http://www.operando.eu/servizi/notizie/notizie_homepage.aspx, last accessed on 18/04/2021.

[26] PlusPrivacy https://plusprivacy.com/,last accessed on 18/04/2021.

contribute to ensuring users' privacy and security are summarized. A great focus is given to the multimedia side as one of the most valuable information shared on SNs that can authenticate a user. Additionally, table 4.3 presents the following countermeasures for each already present security and privacy issue.

4.8.1 Encryption of transmitted media

As discussed in Section 4.5, several SNs are either not encrypting their traffic or are partially using SSL/TLS. The need for using SSL/TLS encrypted traffic for all their interactions is undeniable, as well as the use of secure cookies policies to provide the minimum level of security and privacy to their users. This way, users have a guarantee that when uploading or downloading multimedia content, this content will not be intercepted.

4.8.2 Storage encryption

As discussed in Section 4.4 the multimedia content that users are sharing in many cases can be stored in data centres which are not owned by the SN and geospatial or political events may expose a lot of users to agencies without their will or any type of notification. The issue is critical, given that there are currently many health and medical-related SNs, and the shared information is very sensitive. Therefore, whether the user has to be protected from foreign agencies, malicious providers or developers working for the providers, their data should be stored encrypted. There are many cryptographic solutions, mainly based on public-key algorithms, which can provide users of SNs with the required functionality to store and efficiently recover their users' files, without leaking any information to the cloud service provider [161, 173, 110]. Additionally, proxy re-encryption based schemes [13] can guarantee that the users' information will not be leaked within the SN infrastructure.

Another approach, more focused on multimedia, would be the encryption of the multimedia content. While the previous methodology provides arbitrary encryption of data, there exist more focused solutions such as [133]. The advantage of such solutions is that even if someone manages to get a direct link to the shared

multimedia content, then the content will not be available unless the user holds the proper decryption key.

4.8.3 Data anonymization

Anonymization is a great feature that protects users' profiles from been identified after the time this countermeasure applied. While today the health digitalization has resulted to the so-called big data with several challenges onboard, there are many robust data anonymization algorithms [3, 175] that can anonymize results efficiently and finally protect personal and sensitive data that users do not longer wish to be stored, for example by a service, or should not be kept as the law demands that should be anonymized.

4.8.4 Data pseudonymisation

Separating the real data that a service wants to protect and associate them with identifiers or aliases is a tactic that can provide an adequate level of protection in several cases. Pseudonymization involves replacing part of or all the users' data with pseudonyms. In many cases like in the case of GDPR enforcement pseudonymization (or/and anonymization) is required to provide the best possible level of privacy when processing sensitive data like health and medical.

4.8.5 Steganalysis

Modern cameras and SNs enable users to upload high-resolution images, which are large files without raising any suspicions. However, as previously discussed, they can be used as cover objects to distribute malicious content. Therefore, the use of steganalysis software using multimedia content is considered essential. Experiments conducted by the authors indicate that such mechanisms do not seem to exist currently in the bulk of major SNs, or at least their output is not reported to the user. Many SNs, such as Facebook, may forbid users to use such methods in their terms of service; however, they do not seem to block such actions, something that can be exploited. A typical example of the latter is SecretBook[27], a Chrome ex-

[27] https://mashable.com/2013/04/09/secret-message-facebook/?europe=true, last accessed on 18/04/2021.

tension that allows users to exchange secret messages within Facebook, through steganographic methods.

4.8.6 Watermarking

Digital watermarking is the process of embedding information into multimedia to prove the ownership of the content. Watermarking can be visible or invisible. Visible watermarks can be used to mark a file with usually with a piece of meaningful information, for example, a logo of its creator. Invisible watermarks are used to trace or authenticate a file based on hidden information attached to it.

4.8.7 Co-ownership

To allow users to apply privacy settings, which are closer to their preferences and real-life scenarios, SNs should apply co-ownership models [144, 145]. Such models could allow more than one user to enforce their privacy policies on the co-owned photos, videos, etc. so that the permissions and restrictions on media are not dictated by the choices of one user, and the privacy of all involved users is respected.

4.8.8 Dynamic links to content

As highlighted in Section 4.5, the use of static links exposes users to many risks. Given that the aforementioned solutions, which are based on encryption, might be very demanding in terms of processing, dynamic links should be used to allow users to access multimedia content. For instance, by creating dynamic links to photographs when they are requested, that are subject to the time of the request, the IP and MAC of the user and his credentials, arbitrary access to content by users within and beyond the SN could be minimized. The cost of such solutions can be considered minimal as they involve encryption and decryption of small texts.

4.8.9 Metadata and background removal

While many SNs provide tools to embellish the shared photographs, from simple cropping to applying filters, they do not provide additional functionalities that could help in giving additional privacy to other people. Typical examples are photos from

public demonstrations that are uploaded, disclosing the location and political or even religious beliefs of many people. SNs could provide the functionality for automated detection and removal of faces through, e.g. blurring while keeping the necessary information intact. The same functionality could be extended to blurring objects in the background in case the user is interested in hiding some background context.

Additionally, given that not all SNs follow the same policy towards metadata, all uploaded multimedia files should be stripped of the embedded data unless the user indicates that some of it should be disclosed.

4.8.10 Digital oblivion

In an attempt to offer digital oblivion, several solutions have been proposed. Mayer-Schönberger argues that the use of expiration dates is enough to enforce digital forgetting [98]. Moreover, he proposes the implementation of storage devices that can store information with a pre-determined limited lifetime, so that after the lapse of that time frame, the information is automatically deleted.

Privacy issues	Encryption of transmitted media	Storage encryption	Data anonymization	Data pseudonymisation	Steganalysis	Watermarking	Co-ownership	Dynamic links to content	Metadata and background removal	Digital oblivion
Content and medical images exposure									x	x
Medical identity theft										
Metadata									x	
Unauthorized content access	x	x				x		x		
Tagging - annotation							x			
Video conference	x									
Multimedia shared ownership							x			
External applications	x									
Exposure to the search engine results	x	x	x						x	
Right to be forgotten										x
Exposure to the infrastructure		x								
Security issues										
Unencrypted traffic	x									
Static links		x						x		
Sybil attack										
Flawed design/implementation										
Transparency of stored media		x	x							
Profile hijacking	x									
Data breach and identity theft			x	x		x			x	
Impersonation						x				
Distortion of malleable content						x				
Shared multimedia links										x
Steganography					x					

Figure 4.3: Solutions per Privacy and Security issue.

Chapter 5

A cross-platform SNs privacy mechanism for multimedia protection

This thesis is examining emerging IoT services in Healthcare domain, that, as mentioned in the previous chapters, this domain implements a plethora of the well-known applications to succeed characteristics like interoperability, connectivity, socialising, mobility for the provision of context-aware services to end-users. Non-surprisingly, a plethora of users' privacy risks are related, not limited, to authentication and data management mechanisms. The extremely fast-growing rhythms that social and mobile applications are experiencing, due to the high acceptance and use by users, lead to several poor implementations that in several cases are missing basic reporting features that could protect users from malicious actions and attacks. Today, multimedia data are one of the most valuable contents that are distributed online between users and networks. When we are talking for users' multimedia content, and in case of re-uploading and re-publishing a user's images, without any form of notification, that can result in to harm the original owner both socially and economically.

In [176] we tried to detect whether such acts could be traced and to what extend. The experimental results led us to propose that more fine-grained privacy policies can be implemented using digital watermarks. Therefore SNs can become more privacy-aware, without the need to build them from scratch.

Users expect that uploading their data like their personal photos on SNs, is a

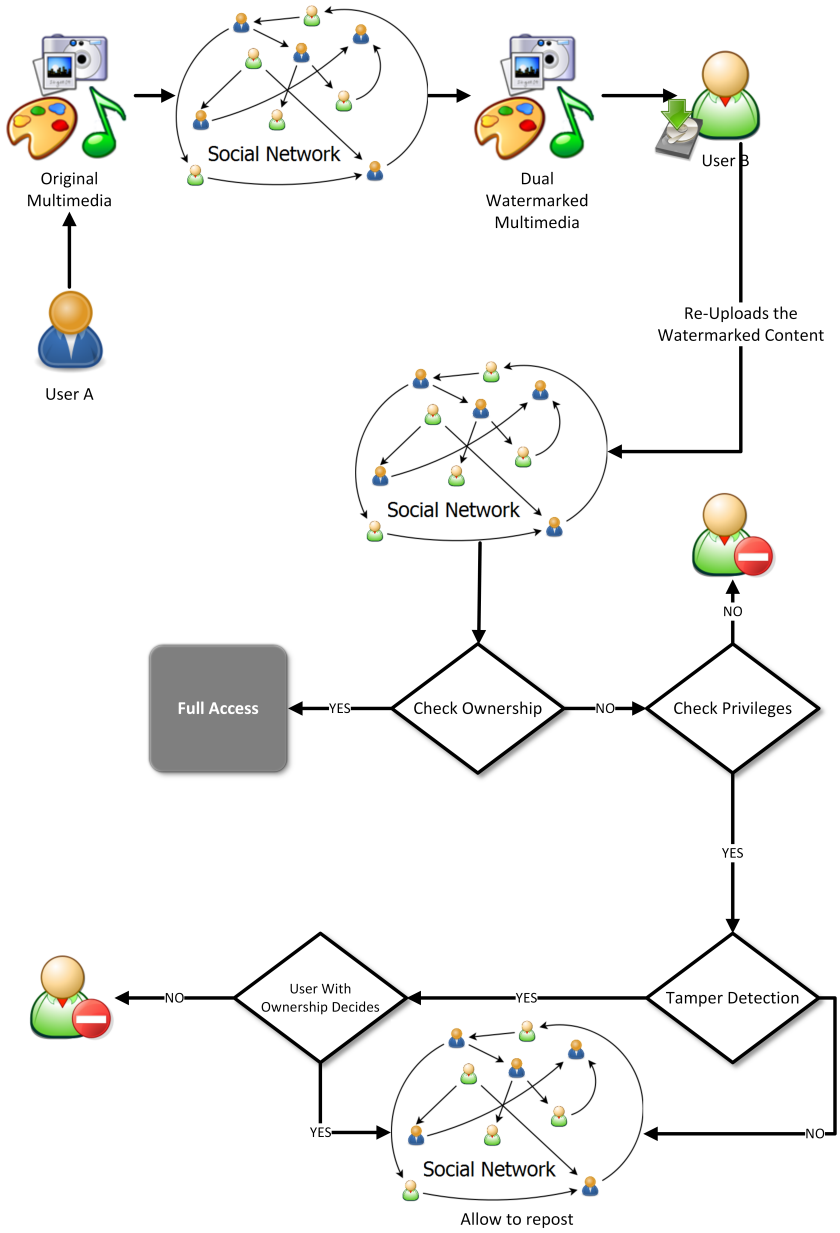

Figure 5.1: Watermarking scheme [176].

secure procedure that can be controlled by them through privacy policies, allowing access only to the users that they decide. Nevertheless, this is not entirely true as there are several risks for their privacy coming either from SN infrastructure or the malicious users inside the SN as we examined in the previous section.

5.1 Problem setting

The current situation on modern SNs normally does not offer a single account for users to sign-in, like a Single- Sign-On (SSO) mechanism to provide them with the ability to manage their content shared across multiple SNs. The only existed technology that offers users a single point of authentication, is the so-called Social login, meaning the ability that some SNs provide and let their users register or login by using their accounts on other services like their Google, Twitter or Facebook account and by considering these services as Trusted Third Parties (TTPs). Nevertheless, there are several concerns on security and mainly privacy issues that make the social login vulnerable or non-functional in certain cases like when the social networking sites are blocked in a local network like a school, hospital or even a country (for example China's Great Firewall blocks some of the most popular SNs like Facebook). The current social login services, do not offer a function that could let users take control and manage their content shared through multiple SNs from a single point, with the ability to apply their preferable privacy settings.

Users today maintain multiple accounts on different SNs. In [176], we focused on privacy issues that a user can have within a single SN. But the current research extends our investigation to a more significant privacy issue that users today can potentially and realistically experience.

A malicious user, today, can easily attack user privacy by downloading or copy part of his content and create a new fake profile to another SN. Based on the current situation, the attacker can use for example the user's name, surname, profile image or more personal images and personal details, depending on the access level that the victim authorised to the attacker (for example on Facebook someone can restrict access by using content settings like public, friends of friends, or friend-only accessible content) and then upload this content to a profile created to another SN. Today there no mechanism that can automatically inform a victim user about

this attack. In this section, we will heavily focus on the development of multimedia sharing protection mechanisms that can be used by different SNs to provide privacy services to their users.

Our tests showed that a modern SN does not check what multimedia are being uploaded, *e.g.* whether a photo has already been published, by whom, what are the privacy policies, etc. On the contrary, the largest modern SNs treat themselves as a separate entity and do not offer any privacy mechanism to protect users' data when users are interacting with more than one of them. Even if the scope of each SN is not always the same and different SNs offer different orientations like socializing, health issues, professional or academic profiles, SNs do not tend to interact, in their attemp to gather more users.

Specifically focused on Health SNs, networks can be used by patients, also, through mobile or even wearable devices. Thus, it is expected that more than the traditional functions and features are available to today's users. Sensors based functions that can track users' biological conditions are today accessible through the advance of mobile devices. It is a common practice for today's health-related mobile applications to provide social networking features, to target users that are suffering by a specific health condition or relevant conditions. Nevertheless, it is expected that users' needs are not limited to one SN and may suffer by additional health issues or even want to socialise with others that do not suffer by the same health condition or even create profiles to several other SNs with different, other than health-related, scope.

Bob suffers from diabetes and downloaded a mobile app that offers several functions to help him monitor his symptoms and his medication. Bob create his profile to this application by inserting his name, surname, email, location and profile image. Additionally, a chat section inside the application let him socialise with others, by presenting his profile photo and his personal information to others. Additionally, Bob can share image files with them, in a community-based environment, for others to better understand and discuss his symptoms. While this section provides him with the ability to exchange information with other users that suffer by a relevant health issue, Bob suffers, also, by anxiety-related symptoms. For this reason, Bob decided to download another application that provides access to a well-known SN that provides several stress relief techniques and a calendar to

manage his anxiety-related symptoms. This application requires the creation of a user profile by requiring, also, the uploading of a profile image. Additionally, it provides users with the ability to share symptoms on a calendar way and review other symptoms that are also submitted through a calendar way. Each user can see how other users manage their condition to better decisions for their future actions. As expected based on the current practice on the web and mobile world, Bob should create and maintain two completely different profiles on these two distinct applications. On the other hand, there is no interoperable mechanism between these two applications that could notify him, in case a malicious user attempts to create a fake profile of Bob to one of these two or other applications. This could result to harm his profile to others or to use his ID to exploit any additional medical-related offered services. In this section, we will highly focus on the multimedia protection issue that is shared between different SNs and propose a mechanism to mitigate the privacy risk of multimedia sharing between different SNs that could result to unwanted situations like identity theft, unauthorised content sharing, distortion of malleable content or other results that we extensively discussed on the previous section.

The proposed mechanism of this section is based on the extension of the solution proposed on [176]. The main contribution of this extended solution is the introduction of a novel distributed scheme, without TTP, which allows multiple SNs to apply the privacy policies of their users among them, even if one user is registered to only one of them. The proposed scheme aims to automatically resolve issues related to the ownership of the multimedia files in terms of privacy.

Two major business-wise concerns that should be discussed are the necessity of such a solution and the feasibility, not limited to the technical side of the solution, but also in terms of a robust business model that will be performed between the participant companies to protect users' multimedia files and their identity. First, we can easily understand how important is such a solution, as globally the data protection laws are becoming more strict and specific regarding the measures that data controllers or processors should take to protect their users' data. General Data Protection Regulation and the California Consumer Privacy Act are just some of the latest examples. On the other hand, someone may argue that the current business model does not allow for such integrations as big companies behind their SNs do

not have the proper incentive to proceed with such solutions. They are just trying to increase their market shares. Nevertheless, someone can find examples of cooperation like in Schema.org[1] in which some of the biggest companies on the web like Google, Microsoft, Pinterest, and Yandex! are cooperating. Additionally, the recent cooperation between Google and Apple regarding a COVID-19 related contact tracing technology [2] is another example that collaborations can be a reality, especially when all the participants are advanced by them. Laws like GDPR seems to make a demanding ground for collaborations to provide the required user data protection level. The recent deal between EU anti-monopoly authorities and Google[3] signifies that big players can be forced to play with more "open" rules. Thus, developing a common privacy-aware framework for SNs under the pressure of regulatory authorities[4] is not a far-fetched plan.

Furthermore, while major SNs may not interact with each other have let other services and SNs act like their authentication mechanisms. Therefore, the majority of smaller SNs are not registering their users directly, but rather obtain user authorisation through *e.g.* OAuth[5] to use some of the information from bigger SNs.

5.2 Watermarking

Digital watermarking is the process of embedding information into media. It is a well-known method to prove the content's integrity and validity of its information. Today, watermarking has been proposed for a variety of applications and mainly for copyright protection, authentication, and tamper detection, copy and device control, fingerprinting and metadata/feature tagging [43]. Watermarks are divided into visible and invisible. An example of a visible watermark can be a logo of a

[1] Schema.org, https://www.schema.org, last accessed on 18/04/2021.

[2] https://www.apple.com/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/, last accessed on 18/04/2021.

[3] http://europa.eu/rapid/press-release_IP-14-116_en.htm, last accessed on 18/04/2021.

[4] https://edps.europa.eu/sites/default/files/publication/14-03-26_competitition_law_big_data_en.pdf, last accessed on 18/04/2021.

[5] OAuth, https://www.oauth.net, last accessed on 18/04/2021.

company embedded in a published image. One example of watermark usage by SN is the dating site Badoo[6].

Invisible watermarking techniques provide minimal possible distortion, while the information is still embedded. Invisible watermarks can be robust, fragile or semi-fragile. Robust watermarks provide mechanisms to retrieve information after common signal processing or malicious attacks. Fragile watermarks are not robust against signal processing and can not be authenticated. Semi-fragile watermarks are something between the two above presented methods, and they are mostly used in tamper detection schemes [164].

The advantage of using image watermarks to authenticate user-related information is their invisibility from attackers. Several different robustness watermarks can be used to achieve the desired security level of the authentication functionality. Fragile watermarks are used to check the integrity of multimedia information, as the slightest modification can break them, triggering an alert to the watermarking system. Semi-Fragile watermark systems detect malicious modifications on the host image, *e.g.* object insertion or cropping, while common image processing as random noise and/or lossy compression does not trigger any alarm. Finally, *robust* watermarks are made to withstand a wide range of possible attacks as they are mostly used for proofs of ownership. An attack from a malicious user would be the removal of a watermark or to make it undetectable. However, this should not be possible without the great degradation of the host image.

The *capacity* of the watermark refers to the maximum number of information bits that can be embedded into a multimedia file of a given size. Depending on the application, the minimum capacity that is required can range from 1 bit, in a copy control application, to a whole photograph. Additionally, two types of algorithms exist. *Non-blind* algorithms compare the original with the watermarked image to extract the information, while *blind* algorithms do not need access to the original image.

Table 5.1, summarizes the needed properties for the aforementioned applications. For more on watermarking and possible attacks, the interested reader is referred to [164, 159].

[6]Meet People on Badoo, Make New Friends, Chat, Flirt, https://www.badoo.com, last accessed on 18/04/2021.

Application / Properties	Invisibility	Robustness	Capacity	Blind/Non-Blind
Copyright	Both	Robust	*	Both*
Auth. - Tamper Detection	Invisible	(semi-)Fragile	*	Both*
Copy control	Invisible	Robust	Low	Blind
Device Control	Invisible	Any*	Low	Blind
Fingerprinting	Invisible	Robust	*	Both*
Metadata - Feature Tagging	Invisible	Any*	High	Blind

Table 5.1: Summary of needed properties for application [118]. **Note:** *: Varies

5.3 Enforcing privacy policies within a single SN

The proposed solution by Zigomitros *et al.* [176] ti is not limited to deterring privacy leaks but also provides a notification mechanism for users to become aware of how their shared information is treated by others. While the proposed solution is mainly focused on images, can also be applied to audio and video files. Furthermore, the proposed scheme has additional value when we examine their value within health SNs where the majority of shared information is sensitive that its protection is governed for example by article 9 of GDPR (Processing of special categories of personal data) or by special data protection laws like HIPAA in U.S. The proposed scheme uses a dual watermarking scheme, a robust and a semi-fragile, for users' information storage.

Analysing the decision to propose a dual scheme, we can make a use case scenario. We assume that user A uploads to an SN, one original health-related multimedia that would also be shared with a group of users that suffers by the same symptoms. The SN starts the embedding process and embeds a robust watermark, associating the multimedia information with a unique identifier that at the same time associates it with the uploader. Additionally, a semi-fragile watermark is embedded in the multimedia file at the same time or afterwards [81], since the robust watermark can tolerate this kind of process. SN's servers are responsible for storing the dual watermark, and later, the result becomes available to other users based on its owner's privacy settings that originally uploaded. Robust watermark's use is related to the ability of the system to identify each multimedia owner even if their processing, while the semi-fragile can detect alterations.

The scheme is illustrated in Figure 5.2.

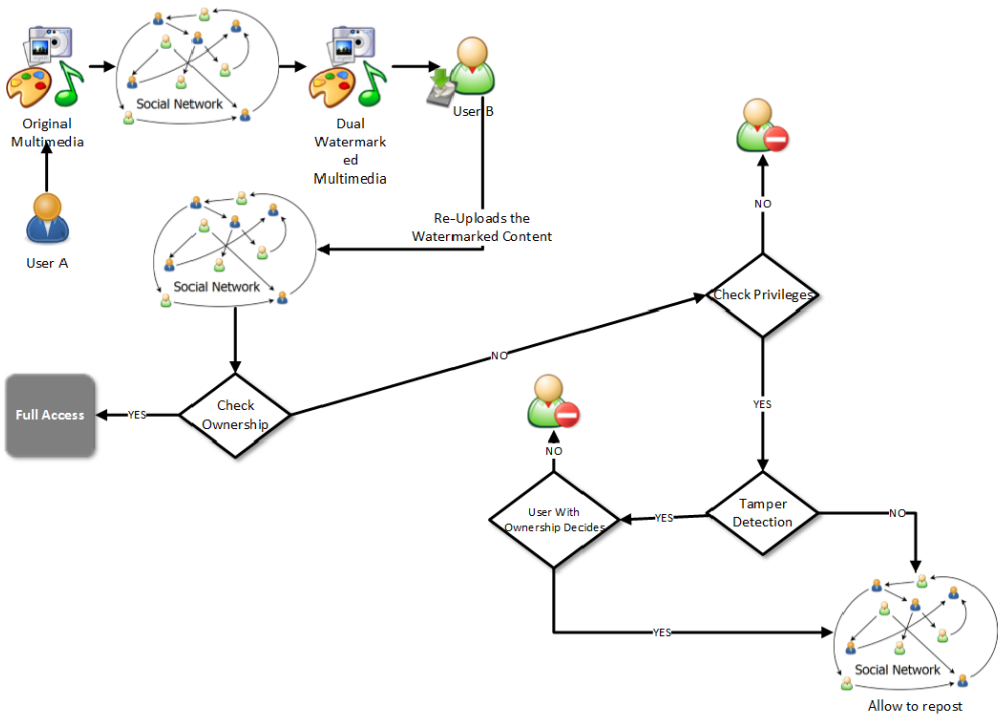

Figure 5.2: The scheme proposed in Zigomitros *et al.* [176].

5.4 Experiments

5.4.1 The process

To ensure our experiments included in [176] a series of experiments were performed on general-use SNs. The original tests were made on three major SNs of that period; Facebook[7] and Google+[8] and VK [9]. So, we repeated our tests to those

[7]Facebook, https://facebook.com, last accessed on 18/04/2021.

[8]Google+, https://plus.google.com, last accessed on 18/01/2019. In Ocotber 2018, Google announced the sunsetting of the consumer Google+ in August 2019. Nevertheless, in December 2018, Google decided to accelerate the sunsetting of consumer Google+ and its APIs because of the significant challenges involved in maintaining a successful product that meets consumers' expectations, as well as the platform's low usage, from August 2019 to April 2019. The announcement also included information regarding a software bug that was confirmed to impacted approximately 52.5 million users in connection with a Google+ API and an investigation into the issue, https://blog.google/technology/safety-security/expediting-changes-google-plus/, last accessed on 18/04/2021

[9]VK, https://vk.com, last accessed on 18/04/2021.

networks. For our tests, we repeated our unique methodology that we first introduced in [176].

Two groups of images were used named Test Set 1 and Test Set 2. **Test Set 1** includes 40 computer generated and grayscale images from TESTIMAGES [12]. The resolution of 20 of these images is 1200x1200 pixels, while the rest of them have resolution 600x600 pixels. **Test Set 2:** also has 40 images but is closer to what could be characterised as typical user images. This set consists of 20 images with resolution greater than 1200x1200 pixels, which range from 2048x1536 pixels to 3648x2736 pixels. These images were taken from 4 different devices, 7 were taken from the camera of an Apple iPhone 3GS, 6 from a Casio EX-Z1050 camera, 4 from a LG KU 990i mobile and 3 with a Canon IXUS 130 camera. The rest of the images were taken again from TESTIMAGES, 10 images of 1200x1200 pixels and 10 of 600x600 pixels.

Additionally, two user accounts were used, user A and B. The scenario includes a comparison of the images between the two user use cases. For this reason, the test images uploaded both on the two accounts and then downloaded from each users' profile. First, we downloaded user A's images and compared them against their originals. Then, we compared user B's images to the original. Afterwards, we compared the downloaded images by both two users, trying to trace possible differences. The same procedure was repeated for each SN, including different PCs and different time frames. These steps allowed us to avoid computer fingerprinting and exclude the time factor from our experiments.

The basic image characteristics that are reported in the experimental results were conducted with Matlab.

5.4.2 Results

The results are different for each SN. For the Test Set 1, the comparison between the downloaded users' images showed that there was no difference in their size or resolution for Google+.

Regarding the differences in filesizes of the downloaded images compared to the original ones in figure 5.3 a histogram is presented for Test Set 1 case.

As said above, no difference compared to the original ones in their filesize when they were uploaded on Google+.

Nevertheless, almost all of the images conclude with a reduction in their filesize, when they were uploaded on Facebook.

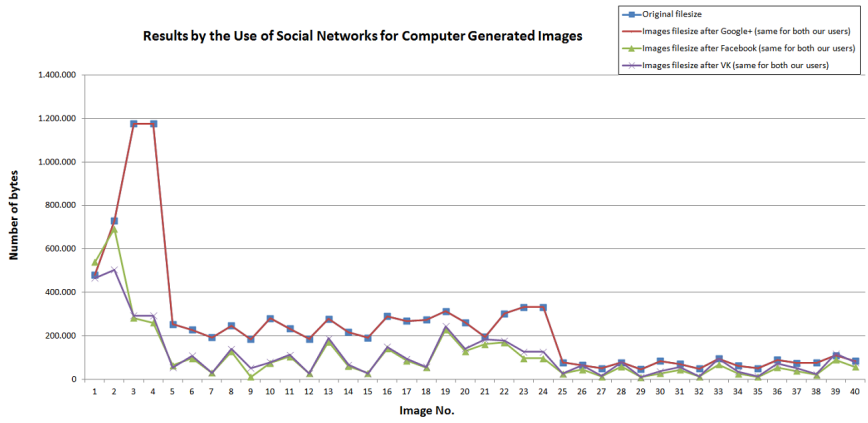

Figure 5.3: Test set 1, image file sizes [118].

In the case of Test Set 2, changes in the downloaded images were identified. First, we confirmed the fact that SNs have thresholds on the image resolution that can be shared. This is a rectangle of 2048x1536, in portrait or landscape orientation. After these dimensions images are resized by both SNs to fit the optimal rectangular. In Figure 5.4, we confirm that Google+ does not make any change in the image size if the image dimensions are smaller than the width of 2048 pixels and height of 1536 pixels. However, on Facebook, a big reduction in the filesize is observed in any case, meaning even if the image was smaller than the above rectangular's limits. Table 5.2 presents and summarizes the results.

On the contrary to the two above SN, VK included three resolution thresholds for uploaded images and that beyond these thresholds, images are resized to fit these boundaries. Therefore, only 30 cases (20 for Test Set 1 and 10 for Test Set 2) fit these boundaries and could be compared against the original ones, all of them being identical. Testing the downloaded images from the profile of user A to the respective from user B, showed again that they are identical, even in the case of size reduction.

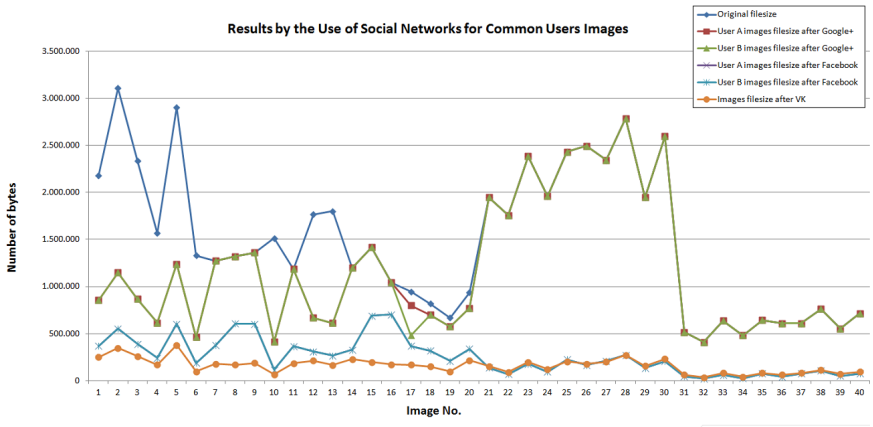

Figure 5.4: Test set 2, image file sizes [118].

	Test Set 1 Original vs FB	Test Set 2 Original vs FB	Test Set 1 Original vs G+	Test Set 2 Original vs G+	Test Set 1 Original vs Vk	Test Set 2 Original vs Vk
Mean Square Error	18,081	14,6884	0	0	4,6918	14,0569
Peak Signal to Noise Ratio	42,2241	41,4557	∞	∞	49,3408	41,8773
Normalized Cross-Correlation	1,0013	0,9993	1	1	0,9986	0,9993
Structural Content	0,9975	1,0005	1	1	1,0027	1,0004
Average Difference	-0,5513	-0,0441	0	0	0,0265	-0,0313
Maximum Difference	34,525	55,3333	0	0	18,55	55,6
Normalized Absolute Error	0,0139	0,0259	0	0	0,008	0,0225

Table 5.2: Mean values of basic image characteristics [118]
The table refers to the images that had no change in their resolution.

5.5 Proposed solution

5.5.1 Overview of the solution

The need for a proposed solution that could provide a robust framework that will be used by several SNs arises by several important facts.

- First, the privacy issue itself. A user that shares a multimedia file on a specific SN will not be notified in the case that another user download this image, re-upload it to the same SN or even worse to another SN and, as a result, has no control on his content after the time when another user decides to download his property (for example his photo, medical image files).

- Secondly, SNs themselves do not seem to worry or take some action regarding this problem. On the contrary, the large SNs like Facebook mostly provide tools to their users or spread content without any special protection all over their network, making their networks generally not suitable for health or medical data sharing. Nevertheless, many users decide to open share data and connect with other patients or even their doctors, also, through these kinds of networks. So, the issue still exists and, if we accept that patients will also make use of such SNs then the solution should also be provided to users of such networks.

- Recent data protection laws and frameworks, like GDPR, introduce more strict requirements also for such networks, and multimedia are considered to be some of the most important types of data within SNs, for example on Facebook, Instagram, Twitter, VK. Some of these requirements include the data subjects' rights to be able to rectification, to restrict access, to object, to request their data portability, to erasure, to access and restrict processing to their data, even these data were previously shared. So, as expected and especially based on the fact that in the content of this thesis we examine the case of health and medical data sharing within SNs, serious privacy-related functional measures should be developed to ensure the previously mentioned requirements.

When trying to extend the watermarking scheme to more than one SNs, we should decide which entity will act as the trusted party that will generate and apply the watermark. A TTP would sound a nice idea, but this would demand the generation of new data centers and additional communication costs. Towards a solution without TTP, each SN could alter the applied watermark to the uploaded media.

Our case study includes n SNs that are participating under a common framework of policies regarding multimedia, that share a common watermarking key K[10]. An example could be the following: A user makes use of one or more SNs that cooperate under our framework, and his userID allows each SN to determine the owner of the media. Additionally, a mediaID field notifies the SN that originally

[10] K is used to watermark each image with a dual watermark, a robust and a semi-fragile as in the original Zigomitros *et al.*[176] scheme.

hosts the multimedia file. Also, a timestamp field is indicating when the media was watermarked. A publication license ID could also be mandatory for an SN to get notified for each user's policy.

For the most possible security level protection, we assume that each SN has its own private and public key pair $(Priv_{SN_i}, Pub_{SN_i}), i \in \{1, ..., n\}$ and a symmetric key Sym_{SN_i}.

Let us assume that Alice uploads an image to SN_1, then SN_1 creates a vector v as follows:

$v = \big(E_{Sym_{SN_1}}(UserID||rnd), MediaID, Timestamp,$
$PublicationLicense, E_{Pub_{SN_i}}(SN_{m_1}Data), ..., E_{Pub_{SN_i}}(SN_{m_k}Data)\big)$
where:$\{m_1, ...m_k\} \subseteq \{1...n\}$ and rnd a random value.

Now, if we encrypt the first field with Sym_{SN_1} we can lately have SN_1 recover the UserID quickly. UserIDs can be salted with a random value to protect the original UserID's value. Salted UserIDs, is a safer way to protect users' identity against their potential tracing from other users or even the other participated or not SNs as only the original SN can find the owner's UserID value and the related to user's profile multimedia. MediaID, Timestamp and PublicationLicense are not encrypted, so that every application can retrieve this information for a specific user. Any other related information contains information that is connected and valuable to each SN that participates to this framework and can be retrieved only by them. The vector is signed by SN_1 so the information that is embedded in the watermark w is $w = v, E_{Priv_{SN_1}}(H(v))$, where H is a secure hash function. Finally, SN_1 embeds in the image the dual watermark using K and publishes it.

In the case where another user somehow gains access to the previously published image and tries to upload the same image to SN_2, then SN_2 will use K to extract the watermark. By analysing the watermark information, SN_2 will get the vector w and verify weither this is a protected image that is authenticated as another users' image.

The publication license ID and the message that SN_1 has encrypted for SN_2, will be the factor based on which SN_2 will decide whether or not it will publish the photo and with what privacy settings will apply, while it will notify SN_1 about these actions.

5.5.2 User rights and framework's contribution

The main advantage of the proposed scheme is that the user's privacy is greatly enhanced. Users can have full control over his shared content, and he can edit whenever he wants its privacy settings, meaning who can see, reshare and edit this content. User, also, has a full report of who is willing to gain access and use his multimedia content, by keeping track of where his media files are being used. A great benefit that this framework offers is the ability for users to revoke or grant access to their content on real-time, independently of the SN that he is registered.

The proposed scheme allows SNs to respond to changes in the legal system automatically. While already many changes have started to become reality in the privacy laws in national and international level, these solutions can have severe implications to SNs as they really will have to change the way they distribute content to protect users' rights. Based on the fact that major general-use SNs, like Facebook or Twitter, are still letting their users' exposed to many risks, to express the free and open sharing of multimedia information, the proposed unification, might seem on first sight scary for some of the Health SNs, that already try to take the most possible security measures to protect their users, based on their budget and resources. Nevertheless, if we admit that the same user that maintains an account to health SN, at the same time may have one or more accounts to other general-use SNs for socialising needs, even if there is a differentiation of the services that each of SN provides, this unification can only enhance their status, as they can provide an end-to-end data protection to their users even to their outter environment meaning the other particpated to the framework SNs. Users should have the ability to choose if the content that they will share should be watermarked as personal-related data that should be protected by SNs or could be distributed openly by the community or each connected SN. The decentralised nature of the scheme enables the equal treatment of all the participants, which is very crucial for its continuity, creating a web of trust not only among the SNs, but among their subscribers as well.

Furthermore, the proposed solution could enable shared ownership schemes. For example, if two users are mentioned as the owners of a multimedia file, as they declare this to the SN they are registered to, they can set their privacy preferences independently, and the SN will enforce the intersection of their policies. Having in

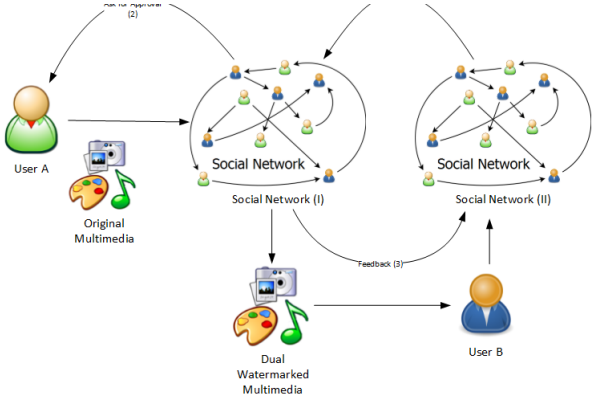

Figure 5.5: Managing media files in two Social Networks [118].

mind that health-related information is considered sensitive and should be highly protected by SNs, each SN should enforce the most strict policies and privacy settings among their users.

Even if the proposed scheme is aiming to establish trust between SNs and their users, we should think what will happen, not only in case of a malicious user tries to bypass its security, but what will happen in case of an attack coming from one of the participatory SNs, even if it might decide not to play fairly or because of a software bug or even an attack that alter how its functions work to take advantage by users' shared data. In the proposed scheme, the key is common for all SNs, therefore, malicious SNs can track where the watermark is stored and alter it, so that it appears as the medium belongs to their users. As far as this is a possible scenario, if SNs use a public watermarking scheme, the decision to have an embedding key that is different from the extraction key can provide an extra layer of protection for our scheme's protection. Additionally, this act can easily be traced, and the misbehaving SN prosecuted not only by the users but from the other SNs who have economic advantages to close down one of their competitors. It becomes apparent that enforcing the scheme by some SNs may force others to act accordingly. Finally, as already highlighted, according to the European Data Protection Supervisor P. Hustinx[11], "controllers will therefore also require them to think better about

[11]Peter Hustinx, European Data Protection Supervisor, Ensuring more effective data protection in an age of big data, Contribution to European Voice online debate on big data and consent,

the legitimacy of what they intend to do... the new framework will also provide for strong sanctions - administrative fines of millions of euros - for the most serious cases where these rules have not been respected.". Therefore, regulatory authorities are expected to enforce such policies soon. In this context, misbehaving SNs are expected to have serious law implications.

Finally, while SNs could use a public watermarking scheme, such as in [168], the adoption of the dual watermarking scheme provides another layer of security that can protect the common watermarking key k from unauthorised access.

5.6 Conclusions

The proposed framework scheme gives users the ability to set their privacy settings on their multimedia content and control the way it will be shared across multiple SNs. The major advantages of the proposed scheme can be summarised as follows:

- SNs users do not need to be registered to all SNs to allow this functionality.
- Users control the way their multimedia will be shared across the participatory SNs. Also, the publication license ID can be crucial for users to apply standard licenses such as Creative Commons[12] or define their custom, by selecting or excluding specific users that will have access to their content or SNs from distributing the content.
- Users can be notified of any attempts to violate their privacy.
- The scheme does not need any TTP; therefore, there is no further trust dependency.
- SNs that participate in the proposed framework can have different policies, without publicly disclosing them. SNs, depending on their scope, marketing goals, conflicts, and policies, may choose to cooperate under different schemes, without exposing critical information to the rest of the SN participants.

July 14, 2014, https://edps.europa.eu/sites/edp/files/publication/14-07-14_ph_for_ev_online_en.pdf, (last accessed on 18/04/2021)

[12] Creative Commons, https://creativecommons.org/, last accessed on 18/04/2021.

- Timestamp information in watermarks, let users also to control their sharing settings based on time criteria. For example, a personal photo could be shared for one month based on the scope that is shared for and then, based on its owner's settings, becomes inaccessible.

- Trust can be established across an ecosystem of SNs, no matter if these SNs are exclusively health or medical specific or are considered general-use for networking, discussion, or even dating, that can also include sexual-health related information shared between users.

- The proposed solution can convert SNs to Content Certification Authorities (CCAs), as they can authenticate and certify the original owner of a file between the users of the affiliated SNs and also detect alterations based on the dual watermarking mechanisms.

While in a scenario of users that are making use to share health information through several different channels, both on general-use SNs and health only SNs, is valid, the solution proposed in this section is more than demanding for users to have the control of their multimedia data across several SNs (those that will participate to such a framework), even if they have not registered such an account to all of them. In a health SNs related scenario, Bob tries to upload a personal health photo that indicates one of the health symptoms of Alice's profile in SN_1. Bob found this photo in one of where her characteristics are quite clear into the health SN_2, where Alice is not registered. When Alice registered to the SN she wanted to register an account, she took advantage of the cross-SN framework and applied private settings to her uploaded photos and chose to watermark it with a non distribute license. As a result, while Bob wants to perform an identity theft attack to Alice by creating a fake account on SN_2, that is the Health SNs blocks his malicious actions by reading the embedded watermark.

Part IV

Mobile Devices' Applications; A Vulnerable IoT Endpoint

Chapter 6

Mobile health applications security and privacy

6.1 Introduction

The intrusion of mobile devices in citizens everyday life, as a crucial part of their daily activities, has demonstrated a completely new era of data sharing arena. Additionally, this radical mobilization accompanied by the advance of sensors that can be embedded on a smartphone combined with their, advanced by the years, accuracy. Specifically, in the healthcare domain, health apps are one of the most popular categories in the famous app stores today [1].

Mobile health (m-health) apps today have the capabilities of sensing environmental changes, while they can collect human body measurements that gives the potential to their users to assess their health status or conclude regarding their health condition, generate alerts, store their historical health data or even connect his app with several external devices or services. Additionally, capabilities like the geolocation are coming to provide added value to the collected information, as users are able to dynamically monitor information regarding health-related data that have collected to specific places and as a result under different conditions. Geolocation tracking services are also famous to wellness-related apps because, for example, users can track their fitness activities on maps. Nevertheless, the above technological advances are opening a big debate regarding privacy and also the security level of the provided apps.

Despite the fact that the establishment of a secure and private environment for the mobile end users is not an easy task [143], someone would expect that when sensitive data are involved, m-health apps will be in place to protect their users' data. Nevertheless, someone can find popular apps not limited to m-health apps, that while they process sensitive data fail to apply the appropriate security and privacy practices for their users' data protection [58, 124, 120, 39, 76].

To investigate whether popular m-health apps are committed to their users' privacy, as they operate sensitive personal data, we focused on the apps provided for Android devices. We chose Android as it is the most popular Operating System (OS) for mobile devices. A number of criteria were set to evaluate the selected apps. These were based on quality, popularity and content-related criteria. Twenty popular m-health apps were selected. After the selection process, we proceeded with the in-depth analysis to conclude on their provided data security and privacy. Unfortunately, we found that for the majority of apps, the level of the provided privacy and security measures that could protect users' data, were poorly implemented or even not implemented at all.

The study we performed [107] has unique and innovative features with respect to previous articles in this area. An analysis of security and privacy concerns in m-health apps is provided by a long-term process that includes evaluation, monitoring and recording of the full life cycle of the apps (from January 2016 to August 2017). Additionally, a full capture and evaluation of the communications were performed to map all connections to first or third parties and assess the quality of these communications. Furthermore, we investigated how app vendors are responding to the privacy and security reports. Finally, we performed a GDPR compliance auditing to evaluate the apps' compliance with the EU legal requirements.

6.2 Background and related work

As also mentioned in previous chapters, a whole new software market of mobile apps has been arisen, where m-health apps are already a famous integral part. Moreover, there is an emerging shift towards the "connected health" model [139], where the goal is to achieve flexible, effective and affordable healthcare services by following the notion of the context-aware smart health (s-health) paradigm [137]. The lat-

est technological trends and many devices (not limited to mobiles) are using common OS platforms, and their apps are frequently considered to be part of the IoT ecosystem [85, 162]. Mobile apps are used in IoT environment to provide several extra services like API connectivity, visualization of data gathering through sensors by providing a user interface to the data collected by various sensors or wearable devices, security, privacy and authentication services and more. That is why mobile apps usage have been remained significant in IoT [105]. Undoubtedly, mobile apps can provide health professionals with several benefits like convenience, better clinical decision making, improved accuracy, increased efficiency and enhanced productivity [157]. Additionally and considering mobiles as a great wallet to collect and store information to share with others, health professional have grown their interest to maintain Personal Health Records (mPHRs) through their mobiles for their patients [30].

There is a worldwide growth of interest regarding the provided security and privacy by mobile devices (*e.g.,* smartphones, wearables) and their apps. Both the E.U. and U.S. have been enforce laws regarding the processing of personal data. EU adopted in 1995 the Data Protection Directive [49] that became applicable the next years among the EU member countries. On the other hand, US introduced in 1996 the national US standards of Health Insurance Portability and Accountability Act (HIPAA) that defines policies, procedures and guidelines for maintaining the privacy and security of individually identifiable health information. Nevertheless, there references that even after years of HIPAA enforcement, m-health apps fail to be aligned with the regulatory protection of the HIPAA [68].

Lately, the EU adopted in 2016 the General Data Protection Regulation (GDPR) [61], which replaced the existing 1995's Data Protection Directive. GDPR became applicable to the European Economic Area (EEA) on May 25 2018 with the greater goal to harmonize the various previous national applicable regulations of EU country members. Several upgraded requirements have introduced both for data controllers and processors. Nevertheless, the GDPR came within a period and technological environment of great growth, and there is a lot of scepticism regarding its applicability in the age of big data and the IoT [51], making even more challenging its applicability. Additionally, mobile devices and their apps are falling to the IoT

era, and there still several issues that experts are examining towards their GDPR compliance.

Below there is a list of references to some of the latest reports and articles that exclusively focus on the privacy and security of mobile apps.

1. A report of the European Commission about citizens' data protection within the 28 EU Members States [2] affirms that over half of the respondents in 16 of the surveyed countries stated that they were concerned about the recording of their everyday activities via mobile phone use or mobile applications.

2. In [5] 20 apps both Android and iOS were evaluated regarding their level of data protection while they listed possible risks and desirable features, based on a list of eight analysis criteria that could help users choose the most secure m-health app to use.

3. In [83] the authors provide a threat analysis regarding possible attack scenarios. Their results concluded to the security and privacy vulnerabilities of 154 selected diabetes and hypertension apps based a testing methodology of four axes: a static analysis applied to the 154 apps, a dynamic analysis performed to the 72 most frequently downloaded apps, security evaluation of web server's security and, a privacy policy inspection applied to the 20 of the 154 selected apps.

4. Additionally, in [84] the authors summarized their findings for the top 20 downloaded apps with a score based on the identified privacy and security issues.

5. in [106] 43 health and fitness apps both for iOS and Android were evaluated. The results showed a high risk to user's privacy arisen by 40% of the tested apps. Moreover, 32% of the apps implies a medium to high risk, 28% of the apps low to medium risk. Surprisingly, none of the apps found to have zero privacy risks. Unencrypted traffic, embedded advertisements and third-party analytics services are the three major security issues that expose users to high privacy risks.

6. In [68], the authors selected 160 m-health Android apps and evaluated their security and privacy level based on a list of seven attack surfaces: the Internet, third party services, Bluetooth, logging, SD card storage, exported components and side channels. From their results, 63.6% of the apps used unencrypted channels to transmit data and 81.8% were using third party storage and hosting services.

7. Another study [72] assessed the extent to which certified m-health apps were compliant with the data protection principles mandated by the UK NHS Health Apps Library. The analysis performed on a list of 79 apps, certified by the UK NHS as clinically safe and trustworthy, showed systematic gaps in compliance with data protection principles, revealing thus security and privacy issues.

8. In [53] 24,405 health-related apps, 21,953 iOS and 2,452 Android devices were assessed in terms of security and privacy based on a number of factor like their access to medical or other sensitive user information, their potential damage through information leaks, information manipulation or information loss, and their access to information valuable to third parties. At this number of apps, researchers decided that the manual testing of all the apps would be infeasible, and as a result, they focused on the presented information where the apps are hosted. The results showed that 95.63% of the apps could lead to at least some security and privacy infringements, whereas 11.67% of them could lead to high privacy risks.

The work that we present in this article is an extension of the above-referenced articles concerning the security and privacy assessment of m-health apps available in online marketplaces. In this respect, we investigate the privacy and security risks in the 20 most popular m-health apps by focusing in the area of privacy and personal data protection when sharing sensitive health information with third-party entities. In respect to the above-presented works, in this section, a special focus will be given on evaluating how the apps request, handle and disseminate the sensitive personal information, the level of implementation of required countermeasures for protecting users' data, the level of GDPR readiness and how mobile apps companies are responding to bugs reporting at the end.

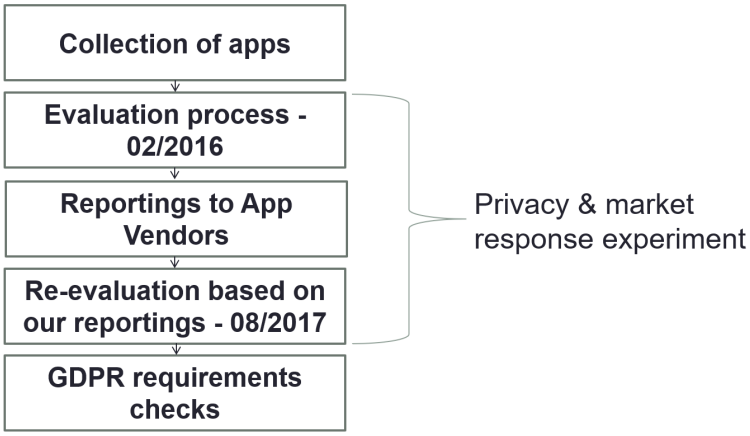

Figure 6.1: Steps of our apps assessment methodology.

6.3 Collection and assessment methodology

In this section, a presentation of mobiles apps collection criteria is included. Moreover, there is an analysis of the assessment methodology that was performed to investigate the security and privacy features for the case of each app. The period of the initial tests was from January to February 2016, and the selected apps were downloaded from the official Android marketplace (*i.e.,* Google Play). After our tests, an email notification was sent to each app's vendor to get informed regarding the findings. In response to the notification, a re-evaluation of the apps was performed from July to August 2017, based on dynamic analysis tests, to compare our previous discovered findings. That was the part of the market response experiment that we included in the steps of our assessment methodology. Finally and based on the time period of our second testing phase (that was some months before became applicable) a GDPR assessment was performed to evaluate the apps' compliance with the upcoming data protection EU law.

6.3.1 Collection methodology

As for the first step, we collected a set of 1080 of the most popular apps from the "Medical" and "Health and Fitness" sections of Google Play. We carefully analyzed the scope and features of each one app, and we kept in our list only those that

Criterion 1	The app must be free.
Criterion 2	The app's content must be in English.
Criterion 3	The app must require health and/or personal data input in order to be functional and based on its description is expected to transmit the users´data to a remote host.
Criterion 4	The app must have at least 100.000 downloads and a minimum rating of 3.5/5 stars on Google Play.

Table 6.1: Inclusion criteria

Downloads	**App. Number**
100.000 - 500.000	II - IV - V - VI - X - XX
500.000 - 1.000.000	IX - XVI - XVII
1.000.000 - 5.000.000	I - III - VII - VIII - XII - XIII - XIV - XVIII - XIX
5.000.000 - 10.000.000	XI - XV

Table 6.2: Number of Downloads of the Analyzed Applications [107]

informed that they collected users' biomedical data. We also decided to focus on apps that provide m-health managing functionalities, exclude the fitness-only related apps and include only apps that collected data regarding health conditions or specific medical diseases.

The final inclusion criteria are listed in Table 6.1. Twenty apps were finally selected, categorized into three main areas: (i) pregnancy and baby growth, (ii) personal/family members' health agenda and symptoms assistants/checkers, (iii) blood pressure and diabetes support. Due to legal issues, we cannot disclose the names (or other identifiers) of the analyzed apps. Hence, we refer to them as *App. I, App. II, . . . App. XX*. Below, in Table 6.2, there is a summary of the number of downloads (captured in 01/2016 on Google Play).

6.3.2 Assessment methodology

The assessment methodology was designed base on three main research questions:

- Which parties have access to personal data from the app?

- What exact data can each party access?
- How safe is each communication channel?

The following steps included an analysis of the technical way the assessment was performed.

Privacy policies inspection Dynamic analysis (web debugging tool) SSL/TLS assessment (ssllabs.com) Reporting and Re-evaluation Examination of critical GDPR functional and non-functional requirements

1. First we carefully read the scope and objectives of each app and emulate a typical user's behaviour. To successfully test the selected apps we created fake emails and/or Facebook accounts and fully emulated a typical user flow.

2. By installing each application, we collected each app permissions and inspected its privacy policies, if it existed on Google Play. A link to a privacy policy became required action since early 2017 by Google for the apps that request or handle sensitive user or device information.

3. As for the next step, we performed automated static code analysis. For this reason, we used MobSF [1] to detect possible vulnerabilities.

4. Afterwards, we performed dynamic analysis for each app using *Fiddler* [2]. The testing flow involved the installation of each app in a cleanroom environment to achieve the most accurate results. Having collected all the communications between each app and third parties, we analyzed and documented all the domains that the apps were communicating with, and we examined their ownership status and their regulating authority. For each captured communication, we listed the type of transmitted data and we analyzed the kind of each data exchange request in terms of its encryption (*plaintext* vs *ciphertext*) and its method (*e.g.,* GET vs POST).

[1] MobSF, https://github.com/MobSF/Mobile-Security-Framework-MobSF, last accessed on 18/04/2021.
[2] Fiddler, https://www.telerik.com/fiddler, last accessed on 18/04/2021.

5. For each communication channel, we performed a web server configuration to evaluate the security level of the HTTPS data transmission. Our tests were base on *SSL Labs* [3].

6. Next, we inspected each packet to determine the contents exchanged in each message. The scope of this step was to identify and evaluate whether the exchanged information was necessary for the intended application purpose, and what kind of data third parties may have access to.

7. Based on our findings, we communicated a report of issues to each app vendor. A list that includes each response about its content, response time, and attitude towards changes was maintained.

8. Finally, a number of checks regarding the GDPR compliance of the apps was performed against the GDPR's requirements.

Our work is heavily focused on coding style and development process. Within our research procedure, we aim to provide useful feedback to developers in order to better secure and protect their apps' features and as a result, provide safer m-health apps that will impact the lives of millions of users.

6.4 Results

The evaluation methodology that was followed generated several results that should be examined to better understand the level of data protection the selected apps offer to their users. Below we will go through the results we captured by the manual and dynamic analysis.

6.4.1 Manual analysis

6.4.1.1 Privacy policies

Based on our methodology's steps, we also examined the existence of a reference link to a Privacy Policy and its relevancy of content with the apps' functionalities. Meanwhile, Google notified by email the developers since early 2017 to provide a

[3] SSLlabs, https://www.ssllabs.com/ssltest/, last accessed on 18/04/2021.

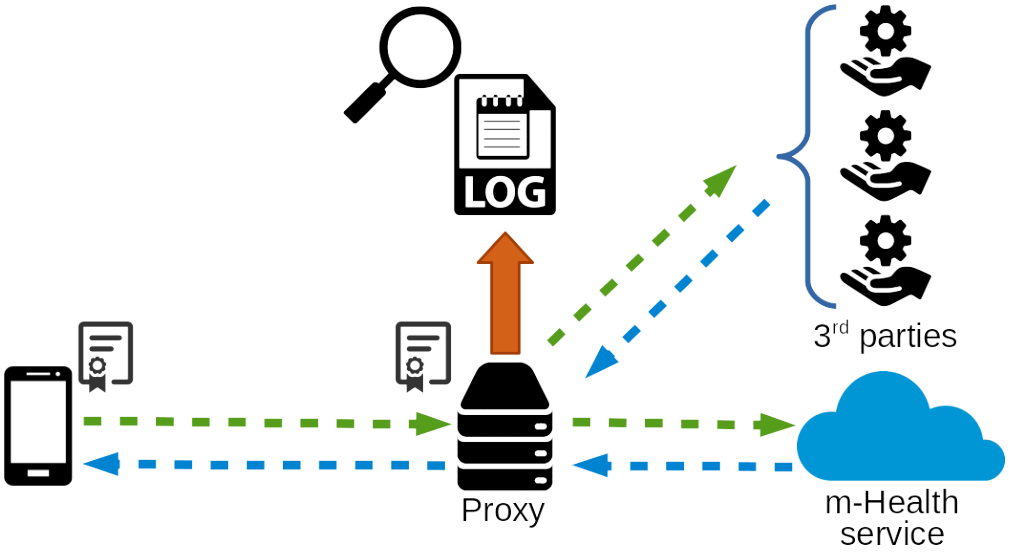

Figure 6.2: Scheme of the Interception Setup [107].

valid privacy policy when they are requesting sensitive permissions or user data, either their apps are at risk of removal from the Play Store on March 15, 2017. First, this fact tells us that it is impossible to check everything automatically. Secondly, there is a lot of interest of what happened after the deadline that Google gave to app developers to include a link to their apps' privacy policy. On February 2016 we examined each app on Google Play Store and concluded that 10% of the analyzed apps didn't have any reference to a privacy policy page. Furthermore, 5% of the apps had a link to a URL that responded with a 404 error page. Finally, 5% of the apps had a link to a privacy policy page that wasn't translated into English, while their in-app content was offered in English.

Moreover, some apps provided a non-valid content to their privacy policy, since the quality of their content structure, their coverage and the relevance of their policy were not up to the required ones for protecting users from privacy issues. Similar to [84] and [149], we also concluded that the problem of missing or invalid privacy policies mainly affects the less popular apps.

6.4.1.2 Permissions analysis

Based on [8] there are several dangerous permissions, and thus we made a list of "normal" and "dangerous" permissions requested per app in Figure 2. For this anal-

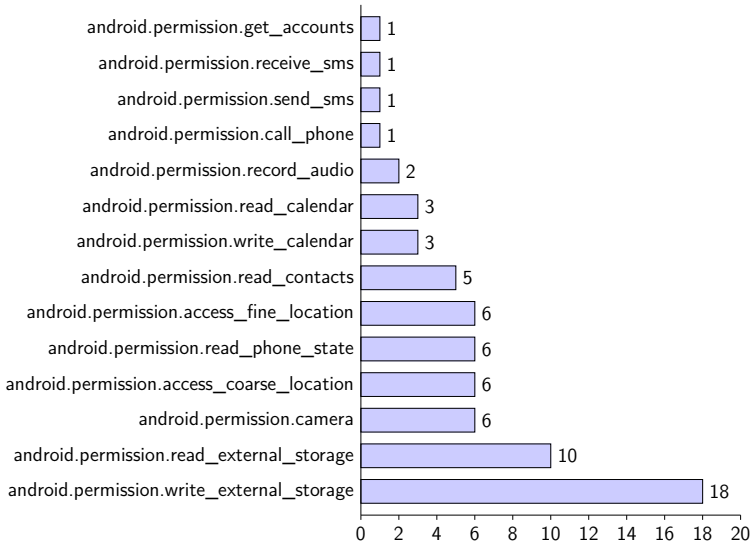

Figure 6.3: Summary of dangerous permission requests [107]

ysis, we collected the permissions listed in the *Manifest* files of the apps´APKs using python scripts. In several cases, we captured permissions requested by apps while phenomenically were beyond applications' scope. For example, two apps requested access to the microphone, but one more app was also requesting it without any obvious reason. Additionally, even if none of the apps made use of Bluetooth advantages, for example, connectivity, two applications requested this permission. This permission requests could be connected to the requirements of ad libraries which exploit Bluetooth devices to track user's location [28, 29]. Since *Marshmallow*, Google required apps that performed scanning for hardware identifiers, like via WiFi or Bluetooth, to request the location permission, leaving out though other indirect approaches for obtaining location information [9]. Six of the app samples requested permissions to access location and coarse location.

Moreover, we examined the calendar access request. While one app made use of calendar, two more asked for it and five requested access to the contacts list (*i.e.,* that is another popular ad library tactic). The most popural permission request was the access request to the devices' external storage. Below, there is a full list of our findings.

Code Analysis	Percentage
The App logs information. Sensitive information should never be logged.	100
The App uses an insecure Random Number Generator.	95
Files may contain hardcoded sensitive informations like user names, passwords, keys etc.	85
App uses SQLite Database. Sensitive Information should be encrypted.	85
App can read/write to External Storage. Any App can read data written to External Storage.	85
This App may have root detection capabilities.	45
Insecure WebView Implementation. Execution of user controlled code in WebView is a critical Security Hole.	30
Insecure Implementation of SSL. Trusting all the certificates or accepting self signed certificates is a critical Security Hole.	30
Insecure WebView Implementation. WebView ignores SSL Certificate Errors.	15
Remote WebView debugging is enabled.	10

Table 6.3: Results of the Static Code Analysis [107]

6.4.2 Static Code Analysis

Each app's APK was evaluated independently with the use of MobSF. The results of this analysis are briefly presented in Table 6.3.

As presented above the tests revealed several security issues. Despite the fact that some of these issues, for example, the insecure random number generators, it is was a found to be a popular issue, not all can be qualified as significant as, most of the times, the use of random number generators is not necessarily connected to security or privacy infractions. On the other hand, many of the apps do not use HTTPS for their connections and have several issues concerning Android WebViews components. Additionally, it worths mentioning that based on MobSF 45% of the tested apps were trying to identify whether the device was rooted that based on

our manual tests this was irrelevant to their scope.

6.4.3 Dynamic Analysis

Dynamic analysis was performed to determine whether an app transmitted personal or sensitive data over the network and to whom. For the data inspection, we used Fiddler, a well-known debugging tool, that helped us made a full analysis of data transmitted from and to the apps. Moreover, below, we discuss the findings per data category like health-related data, emails, multimedia and more.

6.4.3.1 Health-related data

To succeed the best possible study and analysis of the captured transmitted health information while the user interacts with the app, we focus on keywords and/or phrases related to the health status or the medical condition that found to be transmitted. 80% of the apps tested, transmitted their users' health-related data over the internet, while the rest 20% stored them only on the mobile device. Only 50% of those apps transmit health-related data over HTTPS connections for all of their conections. To better understand the findings, table 6.4 summarizes them. The second column *"Sent to Vendor"* presents which apps sent the health-related data they collected to the vendor's domain, while the third column *"Sent to vendor over HTTP"* presents wheither and which apps sent health-related data to the vendor's domain insecurelry (over HTTP). The fourth column *"Share data with third party"* listed the apps that transmitted health-related data to at least one third party domain, whereas the fifth column *"# 3rd party domains"* indicates the number of third party domains that each app sends data. Finally, the sixth column *"# 3rd party domains over HTTP"* presents the number of third party domains that received health-related data over HTTP. Moreover in figure 6.4 we present an example of a JSON response to a POST request over HTTP by one of the apps when we requested when we used the appś back-up function in order to send data to our email address. The numbers showed that 50% of our sample of apps send data to third parties. Based on our analysis, we could categorize the third parties to the following main categories: i. Marketing related platforms that provide mobile analytics

App No.	Sent to Vendor	Sent to Vendor over HTTP	Share data with third party	# 3rd party domains	# 3rd party domains over HTTP
App. I	✓			0	0
App. II	✓	✓	✓	1	1
App. III	✓	✓		0	0
App. IV			✓	1	0
App. V	✓	✓	✓	1	1
App. VII	✓			0	0
App. IX	✓			0	0
App. X	✓	✓		0	0
App. XII			✓	1	0
App. XIII	✓			0	0
App. XV			✓	2	1
App. XVI	✓			0	0
App. XVII			✓	1	1
App. XVIII	✓			0	0
App. XIX	✓	✓	✓	1	1
App. XX	✓	✓	✓	1	1

Table 6.4: Health data transmission [107]

or performance-related data, and ii. Cloud-based back-end solutions used to configure applications' functionalities. Strange enough, one of the tested apps found to sent health-related data to an IP that was impossible to identify any authority based on online resources. From that found to transmit data to remote hosts;

1. 7 of them transmit health-related data to their vendors using GET requests
2. 4 send data to third parties using GET requests

All the above-mentioned apps transfer their users' health data as variables included in URLs. This means that identifiers and sensitive users' data can be visible to everyone having access to the URLs. In the plain HTTP case, the threats are obvious and independent of GET/POST requests. In the case of HTTPS, while the URLs parameters are encrypted, the data is stored in the log files of the webserver. That means that they could potentially be exposed to unauthorized entities.

```
POST /apps/***/users/al***@gmail.com/backups
{
"database": {
"period": [],
"day_record": [
{
"symptoms": "20512",
"weight": "-1",
"intercourse": "1",
"_id": "9",
"headache": "1",
"mood": "1107427330",
"month": "0",
"pms": "1",
"year": "2016",
"day": "26",
"note": "Pain in stomach",
"temperature": "39.8"
}
]
},
```

Figure 6.4: Part of a JSON response to a POST request over HTTP containing health-related data [107].

6.4.3.2 Multimedia data transmission

Multimedia content is one of the crucial content of an app. First, it contains all the mandatory multimedia content the app needs to functionally presents the user interface. Second, all multimedia content that users' submit as part of their account usage, for example, the profile photo. Finally, all the medical images that an app requests by a user of a medical app based on its scope. The unencrypted transmission of multimedia content could easily lead to the exposure of the scope of the app or even the condition of the user. As an example, an eavesdropper could understand the nature and the scope of each app and better organize his attack or associate the user with the app's health category. On the next two cases, the data are falling under the category of sensitive personal data.

App No.	Sent to Vendor	Sent to Vendor over HTTP	Share data with third party	# 3rd party domains	# 3rd party domains over HTTP
App. I	✓		✓	1	0
App. II	✓	✓	✓	1	0
App. VII	✓			0	0
App. VIII			✓	2	2
App. XVII			✓	1	1
App. XVIII	✓			0	0
App. XIX	✓	✓		0	0

Table 6.5: User's location transmittion [107]

Our experiments showed that 20% of the apps ask users to submit personal photos. The half of them used connection securely to send health-related multimedia over HTTPS for all of their transmissions. Three of the examined apps transmitted multimedia content (N=3) to third parties' cloud-based solutions. Additionally, the hosted multimedia could be retrieved by static links, which is a major privacy issue [119].

6.4.3.3 Location privacy

Table 6.5 presents the results regarding the transmitted users' location data. 35% of the apps transmitted users' geolocation information or their postal address either to their vendors or to third parties. Moreover, 4 of the apps send their users' location to 5 third party domains. 3 of them are transmitted over HTTP.

Moreover, 5 out of the 7 apps used GET request to transmit user's location data. One of the apps sent user's location to 2 of its third party advertising services at a rate of almost one request per 3 seconds over HTTP connections via GET requests. Figure 6.5 shows an example of a GET request that leaked location information (latitude, longitude) over HTTP. Additional identifiable information that the same request leaked is, (*i.e.,* mobile device model, OS, device version, local IPv6 Aaddress).

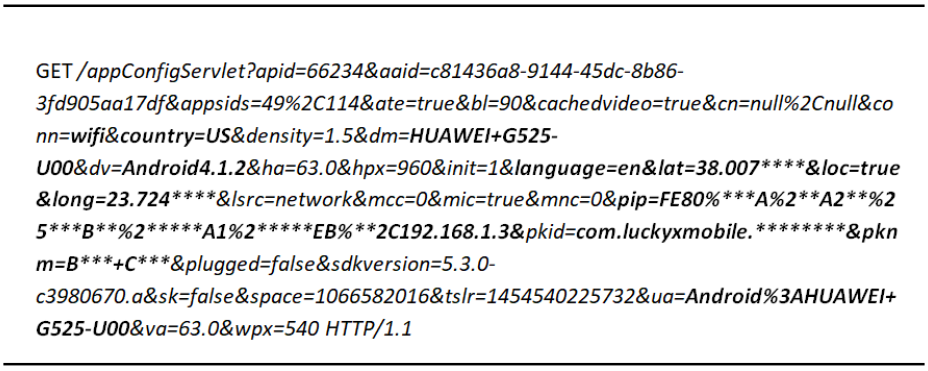

Figure 6.5: Location transmission via a GET request over HTTP to an Ad service [107]

6.4.3.4 User's registration and login security

When an app requires a user log in, then users should register their account through a form. 55% of the apps tested requested and transmitted users' passwords. 27% of the previous apps do not use HTTPS connection for the registration procedure. 45% of the apps that transmit users' passwords used GET requests, not the best possible option due to security concerns.

6.4.3.5 Email and Device Id. transmission

Based on our test results 75% of the apps were found to transmit at least to one domain the user's email address, 33% of these apps used HTTP (5/15) and 60% of the same apps (5/15) of them sent it to a third party domain. One of them sent it an unknown IP couldn't be identified based on online resources. Additionally, one of these apps transmits the user's email address to an IP that was unable to identify its owner through online sources.

Special research was conducted regarding the unique IDs that can expose a specific device's identity. 45% (9/20) of the apps transmitted at least one of the device's IDs (IMEI, GSF ID, Secure ID). 66% (6/9) of those apps used HTTP to transmit IDs while 89% (8/9) of those apps sent it to third parties.

6.4.3.6 Users' search query privacy and OS type

Search queries that users performed were found to be transmitted by 25% (5/20) of the apps. Moreover, only one app transmitted the search queries over HTTPS. 80% (4/5) of these apps sent the searches to third parties while two of the apps sent the health-related queries to 16 different 3rd party domains. Unfortunately, all of the tested apps that found to transmit their users' search queries used GET requests which in combination with the HTTP use make those search queries highly exposed to eavesdroppers.

Furthermore, the OS type was transmitted at least one time per app, and at least one of the connections used was over HTTP. So, it is not impossible for an eavesdropper that tries to identify its victims within a specific area, to be able to understand who is using a specific type of device.

6.4.3.7 Chat sessions transmission

Two apps were found to include chat functionalities to their users. While it was not so frequent for an app to offer chat functionalities to their users, based on our findings, we were pleased to have the opportunity and test their level of privacy and security compliance. Chat is the place where users discuss their health issues and occasionally ask questions or help. As a result, someone would expect that the apps will have taken all the possible affordable security measures to protect their users. Unfortunately, our tests indicated that no encryption was implemented to the chat sections leaving unprotected to eavesdroppers their users' transmitted messages and profiles information. Moreover, another phenomenon that is usual when developers are not focusing on best practices is the insecure database queries that can lead to unnecessary exposure of personal, sensitive and health data even together in only one request. An example of our findings is presented in Figure 6.6, and more specifically, the figure shows a part of a GET request's response over an insecure connection (HTTP). As presented, email addresses, health-related, images and health-related questions are leaked through this way. For privacy reasons, the users' sensitive information included in the figure have been blinded out.

```
[{"post_id":"41313", "title":"confused", "body":"hi ladies I had a miscarriage
a week ago however I did not have the symptoms for it but
I bleed like a normal period but with plenty clot...",
"lastactivity":1453831350000, "links":"none", "groupid":6,
"groupname":"***", "posttype":0, "childposts":21,
"thumbpath": "https://graph.facebook.com/102079*****/picture?type=square",
"name":"Slim ***", "datecreated":"1453480820000", "likes":"4", "posts":"0",
"poststext":[{"post_id":52752,"title":"none" ,
"body":"I spotted for four days and passed clot for two" ,
"datecreated":1453831350000,"lastactivity":1453831350000,
"expiresat":null, "groupid":6, "published":1, "languageid":1,
"postedby":12566,"adminapproved":1,"parent_post_id":41313,"pinned":0,
"closed":0, "links":"none" , "spamreport":0, "posttype":0,"child_posts":0,
"idusers":12566, "name":"Slim***" , "email":"chris****@gmail.com" ,
"password":"1020*****" ,
"thumbpath":"https://graph.facebook.com/1020796***/picture?type=square" ,
"isbanned":0,"applicationid":6,"gender":"female" , "dob":"01011970"},
{"post_id":52751,"title":"none" ,
"body":"you may have passed some old tissue you should go to..."
```


Figure 6.6: Part of a transmission of private information of users chatting over HTTP [107]

6.4.4 SSL web server configuration

As for the next step of the followed methodology, we studied the web server configuration to determine the security level regarding the connections that were sent data back and forth the apps. All the connections made by each app were captured and their domains evaluated by the use of SSL Server Test service from Qualys SSL Labs. Each domain was evaluated and took a letter grade scale (A, B, C, D, E, F, M, T) that this service use to categorize the level of secured configuration. Tests include i. the assessment of the certificate, ii. the server configuration in three categories: a. protocol support, b. key exchange support and c. cipher support.

We splitted our findings to domains owned by the apps' vendors and domains owned by third party servers. Table 6.6 presents the the tests results and the results include only those apps that made at least one request over HTTPS.

In Table 6.7 the number of HTTPS connections to third parties for each app per

App No.	Grade A	B	C	D	E	F	T
App No. I	1	0	0	0	0	0	0
App No. II	0	1	0	0	0	0	0
App No. VII	1	0	0	0	0	0	0
App No. IX	1	0	0	0	0	0	0
App No. XIII	0	0	0	0	0	0	2
App No. XVI	0	2	0	0	0	0	0
App No. XVIII	0	0	1	0	0	0	0
App No. XIX	1	0	0	0	0	0	0
App No. XX	1	0	0	0	0	0	0
N	5	3	1	0	0	0	2

Table 6.6: Number of HTTPS connections to Vendors' domains per SSL grade result [107]

SSL grade result is presented. Additionally, the number of HTTPS connections for each data category per SSL grade result is presented in Table 6.8.

6.4.5 Market response to our security and privacy reporting

One of the most interesting steps of our methodology was to provide each app vendor with a report of our findings. The reports were sent to the email that the app's vendor-provided publicly to the Google Play. Meanwhile, between the date periods, we re-evaluate the apps, one of them was unpublished from the Google Play, and we excluded it to from our future test and results.

6.4.5.1 Privacy policy

Based on the evaluation methodology we designed and followed, we reported all the issues found to exist to the apps' vendors. One category of issues was related to the privacy policies that some of the apps did not maintain or found to include but still have issues like irrelevant content. Meanwhile, in 2017 and after almost one year from our initial experiments, Google notified by email the developers that they should provide a valid privacy policy in case their apps are requesting sensitive permissions or user data. If they would not comply with this, then their apps would have the risk of removal from the Play Store on March 15 2017. The result

App No.	Grade A	B	C	D	E	F	T
App No. I	4	5	0	0	0	0	0
App No. II	1	2	0	0	0	0	0
App No. III	0	4	0	0	0	0	0
App No. IV	0	1	0	0	0	0	0
App No. V	0	2	0	0	0	0	0
App No. VI	0	2	0	0	0	0	0
App No. VII	6	4	0	0	0	0	1
App No. VIII	0	2	0	0	0	0	0
App No. IX	0	1	0	0	0	1	0
App No. X	0	2	0	0	0	0	0
App No. XI	0	5	0	0	0	0	0
App No. XII	0	1	0	0	0	0	0
App No. XIII	0	5	0	0	0	0	0
App No. XIV	0	1	0	0	0	0	1
App No. XV	1	0	0	0	0	0	0
App No. XVI	0	1	0	0	0	0	0
App No. XVII	7	9	3	0	0	0	1
App No. XVIII	2	2	0	0	0	1	0
App No. XIX	7	8	4	0	0	2	0
App No. XX	5	8	3	0	0	2	0
N	33	65	10	0	0	6	3

Table 6.7: Number of HTTPS connections to third party domains per SSL grade result [107]

that someone would expect will be that by the 5th of July 2017 (*i.e.,* the date we performed our re-evaluation process regarding the existence of a privacy policy link on Google Play) all the apps would finally add a valid link to a proper privacy policy. But the results showed that there were still apps that have not comply with that order. More specifically, one of the apps kept missing a privacy policy, another app provided a link to an error page and, another app kept having a link to a privacy policy page not translated in English. It seems that is not possible to check everything automatically even if you hold the position of one of the biggest multinational technology companies in the world.

Grade	Email	Password	Location	Health data	Search queries	Unique ID
A	3	2	1	4	0	0
B	7	5	2	2	2	2
C	1	1	0	1	0	0
F	2	0	0	0	0	2
T	0	1	1	1	0	1

Table 6.8: Total number of HTTPS connections for each data category per SSL grade [107]

6.4.5.2 Secure transmission of user data

A re-evaluation of the apps by running a dynamic analysis on their updated APKs was performed. To better organize and present our findings, we categorized them into minor and major issues, and we analyzed them in numbers per category. Examples of what we characterize as minor and what as a major issue can be found in Table 6.9.

In figure 6.7 we present the number of major issues found before and after we notified the app vendors. While we capture some improvement on major issues based on our reporting, only 5 out of the 12 apps with minor issues have partially or completely solved the reported problems.

6.4.6 GDPR-readiness assessment

During our re-evaluation process and based on the fact that our experiments took place a small period before the GDPR was about to become applicable, we performed an additional evaluation of the apps against a number of functional and non-functional requirements of that GDPR introduced for data controllers.

Below, the results of our evaluation for the 19 remaining on the Google Play apps is presented.

Example of major and minor issues	Major	Minor
Transmission of Device IDs or Personal or Health data (in any way) to 3rd parties	✓	
Transmission of Device IDs or Personal or Health data insecurely to Vendor (i.e. over HTTP via GET or POST request)	✓	
Transmission of Device IDs or Personal or Health data to Vendor via GET request over HTTPS		✓
Transmission of anonymous behavioral data to 3rd parties		✓

Table 6.9: Example cases of major or minor issues [107]

6.4.6.1 Functional requirements

- **Consent (I):** 11 (58%) out of the 19 apps provide, at least, a piece of introductory information regarding their privacy policy or/and term of use before the user's registration process or as part of it.
- **Consent (II):** Only one of the apps requested by users to consent upfront each time the app required additional information by them.
- **Consent (III):** None of the apps required users to answer specific questions, in a electronic form, about their willingness to participate.
- **Right to withdraw consent:** 7 (37%) out of the 19 apps of the apps provide a mechanism to the user to withdraw its consent and allow the erasure of any previously consented information. Nevertheless, in 1 out of the 7 apps providing this option, the deletion functionality didn't work during our tests. 3 (25%) of 12 apps informed their users that their data could be deleted only by an email request to the app vendor. 2 out of these 12 apps offer users functionality to delete their submitted data individually, one at a time, and

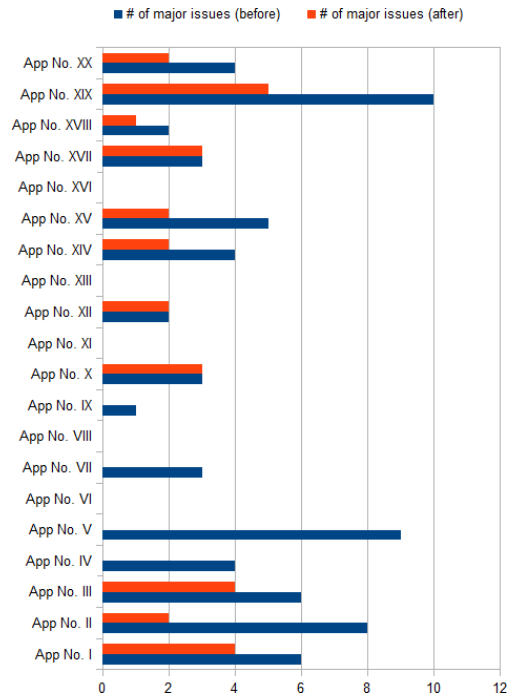

Figure 6.7: Number of major issues per app before and after our reportings [107]

not all at once. Moreover, one app stated that: *"Sometimes we were asked about deleting all records, to start a new series of measurement. But there is no reason to do this. The time range can be changed to view only the wanted part of all records. So old readings can stay"*, and in this way, it refused its users' data deletion.

- **Right to data portability.** 7 (37%) out of the 19 apps provide a mechanism to send, upon request, the personal data to another entity in a machine-readable format (*e.g.,* XML or CSV format). 2 of these apps offered this function via a web-based platform. Some apps found to advertise this functionality in its paid version.

6.4.6.2 Non-functional requirements

- **Data Protection Officer:** None of the apps found to provide any contact details to such a role. Despite that, 12 (63%) out of the 19 apps offer some point of contact for security and/or privacy-related requests.

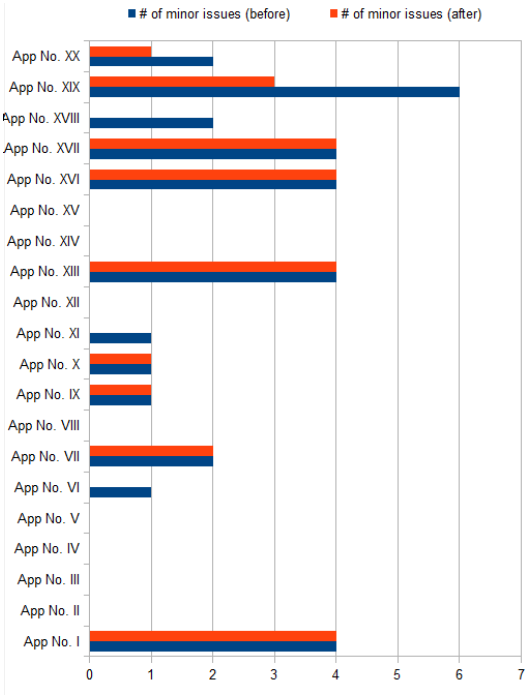

Figure 6.8: Number of minor issues per app before and after our reportings [107]

- **Profiling and marketing:** 11 (58%) out of the 19 apps provided information on the collection and processing of user data for profiling, that was found mainly as part of their privacy policy section.
- **Transfer to third countries:** 8 (42%) out of the 19 apps of apps notifies their users in advance, even before their registration, that they are sharing data with third parties. Only 4 of them (21%) of apps in a functional manner, for example, in a pop up with a checkbox.

6.5 Conclusions

Mobile health (m-Health) apps are gaining by the years even more market as there are not only limited to the mobile devices but in many cases are compatible with other devices like smartwatches, as devices like these are based on the same OS system. Additionally, many different smart applications and medical devices, including IoT installations, are communicating with mobile apps to provide a user-

friendly interface to their users to better control their functions ubiquitously.

This section provides the findings of an extended study that conclude to several security and privacy-related gaps and pitfalls mobile apps may include. For instance, the way some of the apps are violating users' privacy by revealing sensitive information like health conditions, medical symptoms, photos, location, emails and passwords to several third-party domains. Moreover, a big issue is the insecure way that the apps use to transmit their users' sensitive information, for example via HTTP and/or by using GET instead of POST requests, both to apps' vendors' servers or even to third parties.

The results indicate that very sensitive data are vulnerable to simple sniffing attacks, while most of the detected vulnerabilities have very simple solutions that in their majority do not require much effort to fix, but only a few of the apps' vendors finally fixed them promptly. One very serious finding is that users can be victims of user profiling, blackmailing, stalking, defamation, and even identity theft for economical or reputation attacks, only because they chose to download and use an insecure app that was intended for different uses.

Simple insecure software decisions by app developers/publishers are still an issue and they seem to keep repeating the same mistakes over every new software environment. Even some period before the enforcement of privacy laws like GDPR, the major question regarding the security and privacy provided to their users is open. Additionally, based on the fact that we are in the IoT era and already wearables are part of this market and in most cases are sharing apps by the same market app stores, will this market achieve to become secure and privacy ready to better protect their users?

Part V

The Market's Response on Privacy Related Incidents

Chapter 7

The Market's Response on Privacy Related Incidents

7.1 Privacy Incidents: From Data Breaches to Privacy Revelations

In this Chapter, going a step further from considering the impact of data leakages on a single individual, we are examing the case of the market impact to the company that made this leak possible. A special focus is given to privacy incidents and their impact. In Chapters 4 and 6, we mainly focused on the security and privacy issues that exist to SNs and mobile applications, that are representing a huge market share. This problem has been already researched by other scholars in the literature[4, 172, 142, 61] and researchers have concluded that in all cases the companies involved in data breaches suffered also a negative market impact. Additionally, results regarding the impact that Snowden's revelations had on the involved companies[113] are shared. Having in mind that the Ubiquitous Computing Services, similar to the Smart Health services presented in Chapter 3, could bring solutions to major issues within the context of Smart cities and that they are expecting to become even more integral for citizens, the Snowden revelations and their impact can act as an alarming state for the future communities and the emerging need for privacy by design.

The Department of Homeland Security (DHS) defines[1] the 'privacy incident' as

[1]Privacy Incident Handling Guidance, DHS Instruction Guide 047-01-008, Published by the U.S. DHS Privacy Office, December 4, 2017, https://www.dhs.gov/sites/default/

follows:

"The loss of control, compromise, unauthorized disclosure, unauthorized acquisition, or any similar occurrence where (1) a person other than the authorized user accesses or potentially accesses [PII][2] or (2) an authorized user accesses or potentially accesses [PII] for an unauthorized purpose. The term encompasses both suspected and confirmed incidents involving PII, whether intentional or inadvertent, which raises a reasonable risk of harm.

Moreover in U.S. DHS' guidance, the terms "privacy incident" and "breach" are used interchangeably, as the privacy incident occurs when someone actually or imminently jeopardizes, without lawful authority, the integrity, confidentiality, or availability of information or an information system. It constitutes a violation or imminent threat of violation of the law, security policies, security procedures, or acceptable use policies[3].

In the latest years and after GDPR enforcement in 2018, data breaches were one of the most popular news on the global media[4] as the new regulation was very clear and very strict regarding their disclosures and the related penalties in case of data privacy compliance failures. The GDPR defines [61] the 'personal data breach' as a "breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed".

Reading carefully the above definitions for data breaches, we can identify a closely common framework on how to define a privacy incident that occurred by a data breach. Unlawful and unauthorized acts and users can lead to privacy incidents and data breaches. Also, in both definitions, we read that even a breach is occurred inadvertent (U.S. DHS) or accidentally (GDPR), we still have a data breach that can lawfully harm the data owner.

---

files/publications/047-01-008%20PIHG%20FINAL%2012-4-2017_0.pdf, last accessed on 18/04/2021.

[2] PII is the used for the Personal Identifiable Information, https://www.dol.gov/general/ppii, last accessed on 18/04/2021.

[3] 44 U.S.C. § 3552(b)(2), https://www.govinfo.gov/content/pkg/USCODE-2014-title44/pdf/USCODE-2014-title44-chap35-subchapII-sec3552.pdf, last accessed on 18/04/2021.

[4] https://www.forbes.com/sites/bernardmarr/2018/06/11/gdpr-the-biggest-data-breaches-and-the-shocking-fines-that-would-have-been/, last accessed on 18/04/2021, https://www.theguardian.com/technology/2018/jun/26/european-regulators-report-sharp-rise-in-complaints-after-gdpr, last accessed on 18/04/2021.

By using the term 'privacy incidents' we are, usually, referring to events that are associated with the exposure, loss, or damage of personal or/and sensitive data. One would expect that a privacy incident would affect negatively not only the data owners by also the data controllers and processors involved in their management.

Nevertheless, the findings of the previous chapter presents that the risk of the impact is not always very well-estimated by software companies, as in many cases reported security issues, even if they are categorized as easy-to-fix issues, are not fixed promptly.

One of the maybe, first most comprehensive study [4], was focused on the market reaction to privacy breaches. The study included 79 breaches that also involved unauthorized access to sensitive data and were reported from 2000 to 2006. The analysis was performed using CAPM and a t-statistic for the statistical analysis. The outcome of the study was that there was a negative stock market reaction hypothesis for the event windows that included either the event day or both the event day and the day after, based on pieces of evidence. Adding some limitations to their study, the authors, reported that due to the small sample of breaches, the possibility of generalization of the findings is constrained.

In another study [172], their authors examined a sample of 123 security breach announcements for a large period, from 1994 to 2006. The announcements were founded from search engines' results and the Lexis-Nexis business news database for listed companies in NYSE, AMEX, and NASDAQ. The authors also based their analysis of the negative impact coming from breaches on CAPM and a t-statistic. One of their results was that the impact was bigger on e-commerce and technology firms. Another outcome was that more recent events have a smaller impact than former events.

In [142], a thorough literature review of 45 event studies was conducted following the application of the event analysis methodology. The study reported the highly critical aspect of information security incidents to stock prices. Their results showed that 75.6% of the studies included in their research, report the statistical significance of the impact of security events on the stock prices of the involved companies.

Some years ago and before the GDPR enforcement, on June 5, 2013, when the Guardian newspaper published a secret court order which directed Verizon to hand

over all its telephone data to the NSA on daily basis, one of the most popular news (not only in cybersecurity) in the world was Snowden's revelations. One of the highest interests for the context of these revelations was that many of them were connected to huge well-known companies that even the average user is making use of.

We would expect that the shock of the news would affect users and make them change their consumer behaviors. Additionally, the impact of these revelations could damage the reputation of these companies in a way that the companies would have a great loss on markets and users' loyalty.

7.2 The Impact of Privacy Related Revelations on Markets; The Snowden Case Study

Mass surveillance is one of the most greater risks that citizens may encounter in a future fully interconnected world. Snowden, a former National Security Agency (NSA) employee, revealed a series of top-secret documents to the public, revealing many NSA actions but also others like the British and the Australian intelligence services. These revelations are just some of the most recent shreds of evidence that large-scale surveillance programs can be real and can potentially harm a big number of users worldwide. Highly accepted technologies, that users toady use to effectively connect with others by sharing their news, opinions, schedules, photos, and videos using devices like mobiles, wearables, devices embedded with sensors and any other communication technologies implies the risk of mass surveillance in which either secret agencies or even big tech companies, under special circumstances, are involved in.

The Snowden case can be characterized as a very unique case, as for the first time we had so many revelations regarding mass exposure of data not by for example hackers or accidentally, but by leakages driven by governments. This series of revelations could occur on big companies in terms of economic costs for the companies involved. Snowden revelations were expected to cause big changes for the companies and general the markets as it was on the top of the news for a long period.

Public authorities and organizations are still battling over them, but as data analysis reveals, they do not have any impact on the equity returns. In [113] there was an extended research regarding the impact that Snowden's revelations had. In this study, there were included negative events based on Snowden's revelations that concern technology and ICT listed companies in NYSE and NASDAQ. A critical step of our study was to determine the event window. We reported the day of each revelation based on the official announcement and report it as "day zero". Moreover, with the term "day zero date" we referred to the date of the first closing price after the revelation occurred. The "Day zero" is, actually, the day that the event is publicly announced and not the date that the incident occurred, no matter when the event took place, as there no information of the exact time when was the event took place.

7.3 Methodology

To analyze the data collected through our study, we applied a methodology well-known methodology of event study aiming to evaluate the potential costs of security breaches. Event study has become a great tool used by researchers to evaluate the impact of technological and security-related events in conjunction with the value of a firm. Famar in [59] introduced an efficient market hypothesis; The stock price of an included firm should illustrate the true value of that firm. So, based on Fama, the weak form of efficiency should include the set of information based on historical prices, the semi-strong form of efficiency includes all the publicly available information to all market participants [94] and the strong form of efficiency includes both all private and public information.

For market participants, the event study analysis constitutes an effective tool for evaluating the information context of events. More frequently, reactions of equities under the occurrence of information breaches on the event day and the days surrounding the incident, usually have a predictable expected pattern. That means, that in a case that media news is positive for a firm, the economic participants' reaction is expected to be, also, positive. Additionally, in the case of a negative event for a firm, a downside reaction is expected.

A crucial step for the event study process was to set the event window. Our methodology was to report the day of the official revelation's announcement and report it as "day zero". The "day zero" consists of the actual date of each revelation, while the "day zero date" represents the date of the first closing price after the revelation. As expected, the two previously mentioned days can be different days and concern the active trading hours after the official announcement of a revelation when the market reacts to this new information. Additionally, for events that happened to took place during a weekend or before a bank holiday, we defined the first working day as the first day of the privacy-related event.

Moreover, based on a well-known practice for not publicly noticeable security breaches, we assumed that day zero is the day that the event is publicly announced and not the date that the incident occurred, no matter when the event took place.

Finally, the sample size of our study, compared to the period we chose to investigate, is solid in comparison to other studies, such as in [172] where the study refers to a much longer period of 12 years, with a sample of 123 events.

7.4 Collection methodology

The sample of the events included in this study was retrieved by media resources like Bloomberg and Reuters. Also, the original leaked documents as archived in the Snowden Digital Surveillance Archive[5] were used to cross-check our final results. The events were analyzed and only companies that publicly traded on any US stock exchange were finally included in our sample.

The empirical analysis, the equity daily prices and the benchmark were based on Thomson Reuters Datastream. Moreover, the Fama–French factors were based on the Kenneth R. French website [6]. Finally, the risk-free rate of return was based on the monthly 3-month Treasury Bill, adjusted on a daily basis to become suitable for the event study analysis.

To collect our sample, a keyword/phrase strategy was developed to define strategical best search terms to be used on search engines, with a great focus on Google

[5] https://snowdenarchive.cjfe.org/greenstone/cgi-bin/library.cgi, last accessed on 18/04/2021

[6] French K.R. Personal Website, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/, last accessed on 18/04/2021.

Companies					
Activision Blizzard	1	Google	6	Seagate	1
AOL	4	Facebook	16	Twitter	3
Apple	6	Level3	1	Verizon	4
AT&T	2	Mastercard	1	Visa	1
Blackberry	2	Microsoft	13	Western Digital	1
Cisco	1	Oracle	1	Yahoo	16

Table 7.1: List of the affected companies included in the sample. Note that AOL was delisted on 23/06/15 [113].

and Bing. We did not use academic search engines for this step because the sample of our research was publications on the press that include data leakages by companies of ICT and Internet-related sectors. Each publication consists of one event of our sample. The second step was to use the well-structured Snowden digital surveillance archive and validate our findings and the archive's sources to include any relevant event we may miss on the first step. To get the most relevant to the U.S. market results a VPN service was used, with an exit node in the USA.

Another step was to exclude any general news or general comments for leaks by companies that were not well-documented and keep only facts and revelations where the companies involved was a documented fact, with or without their consent. Additionally, based on the results of the same search engines we determined whether one of the events was confounded or not with other major firm-specific events. During the period of the event window, there was no event that confounded with one of the examined events.

7.5 Results

Our dataset included 80 events and 19 companies, after excluding all events of companies, which are not listed in any exchange other than the US (for example UK, Germany etc). A number of well-known firms included as shown in Table 7.1. The first event is on June 6, 2013 and the last event is on September 25, 2015.

As shown in this study the insignificant (zero) impact of day 0 seems to turn

into a positive statistically significant CAR the day after (using FF-model) and the next 3 days (using CAPM). The outcomes are aligned with other findings in privacy in the post-Snowden era.

Snowden revelations, due to their content, were expected to cause "noise" in the markets and the broadcast news. Their impact is so large that public authorities and organizations are still battling over them, but as data analysis reveals, they do not have any impact on the equity returns. On the contrary, the insignificant (zero) impact of day 0 seems to turn into a positive statistically significant CAR the day after (using FF-model) and the next 3 days (using CAPM). While some surveys indicate that there are some changes for the individuals [132], as they have started using or at least considered using more secure and private technologies, few new users have embraced privacy-enhancing technologies [123].

The response of the market against these revelations can be considered sceptical, as we could not locate any impact on the stock prices. Undoubtedly the market today is mature enough to understand and manage the results of cybersecurity events, compared to the past. The risk someone would easily identify in this case is that today, privacy is not highly regarded in our society, by both individuals and the market. As a result, the outcome of this historical privacy incidents can only be weighted in a macroeconomic fashion, and maybe is very soon now to understand what economic changes were triggered by them. Maybe the neutral findings of our study can be further justified by the fact that the study is performed in the USA market, considering the in EU there is a more strict law regarding privacy, as it is considered a fundamental right. Nevertheless, the result is raising an alarming situation for future societies.

7.6 Conclusions

The neutral response can not underestimate footage of security breaches. First the findings can be further justified by the fact that the study is performed in the USA market. In general, the European culture is considerably more fixed on privacy, something that is reflected by the fact that privacy is protected as a "fundamental right" in the EU. Therefore, the impact of the U.S. market examined in our study, it was expected to be significantly dicreased.

Another possible fair explanation is that this kind of breaches, regardless of their size, does not affect directly the value of the affected firms as they are a result of voluntary action, including in some cases also the consent of those companies and the cooperation of the government agencies. In that direction, the Snowden revelations seem not to be evaluated by the market as security breaches with their traditional form, but as the exposed cooperation of firms with government agencies into large surveillance programs.

Based on our results, we could explain the neutral response by the assumption that the markets consider this cooperation as compulsory for the involved companies, and as a result, they do not penalize them immediately for exposing sensitive data to government agencies. That means that the actual impact of these revelations should be re-evaluated in a macro economical context in the next coming years.

Moreover, ubiquitous computing is expected to become even more embedded in all aspects of citizens' lives. This could benefit peoples' lives, but on the other hand, the tolerance on behalf of the markets to privacy incidents like Snowden's revelations should generate great concerns. More specifically, the future global challenges such as the current one, the COVID-19 pandemic, may still concern governments. In such a scenario, well-known everyday users' tools like SNs and mobile-based sensors may be used, with the risk of the creation of a huge panopticon environment [40][7] for the citizens' lives trying to control the movements and behavior.

[7] Thomas McMullan, What does the panopticon mean in the age of digital surveillance?, 23 July 2015, https://www.theguardian.com/technology/2015/jul/23/panopticon-digital-surveillance-jeremy-bentham, last accessed on 18/04/2021.

Part VI

Closure

Chapter 8

Open Questions and Future Directions

8.1 Open Questions

This thesis investigated privacy technologies in current and emerging smart health environments, exploring a series of different tools that users utilise to collect, manage, and share information. First, we explored the novel capabilities that new smart services can bring to smart cities and contributed by proposing additional services that could bridge some gaps between citizens and healthcare services. Moreover, while social networking is still a great and popular tool that empowers patients' communication, we studied the privacy and security issues which are still open, their impact, and potential countermeasures in the context of metadata. Additionally, we introduced a new scheme for multimedia protection that could bring better and robust control of the multimedia shared by users across different social networks. The IoT and ubiquitous computing era bring forth the need for effective heterogeneous and interoperable solutions that current platforms could embed without the need to be developed from scratch, as users utilise even more different tools and services by creating multiple accounts to social networks. Furthermore, we evaluated the level of personal and sensitive data protection that m-health apps provide to their users and how mobile applications' vendors are responding to security and privacy bug reports. Moreover, we evaluated the GDPR readiness of popular m-health apps to understand better the level of compliance with demanding privacy regulations, like GDPR.

The distribution of m-health and medical applications, is rather alarming requiring the enforcement of more strict data protection rules. The examples are numerous [135, 130] and the rise of the pandemic with the need for various mobile health apps highlighted this gap. The contact tracing apps are a fine example of the issue as they involve sensitive personal information, usually combined with location data, while other apps may also require data from body-embedded sensors. Also, healthcare is shifting towards more personalized services, to better support each patient's special needs. This kind of healthcare service delivery has additional benefits as it can fit evidence-based medicine better, allowing the provision of personalized treatments and reduce costs by enabling a better scheme for patients management from medical clinics. Therefore, patients can greatly benefit from this personalized patient-centric model. Of course, still, there are many aspects of mobile devices that are not yet up to clinical standards, ranging from accuracy and reliability [25, 157] to security and privacy [53, 83, 107]. Nevertheless, millions of users today trust mobile apps, even the apps which are not certified by official governmental bodies for their reliability and accuracy and share their data. So, the extent of the problem is rather big and diverse.

Despite the criticality of security and privacy technologies, they are usually missing from health applications, something which has become apparent with, e.g. the devastating effects of ransomware and other cyber attacks to health institutions. Moreover, as IoT is getting more integrated into citizens' lives, consumer wearables constitute an emerging market, becoming ubiquitous and widely accepted by users [141, 170]. The high tech functionalities and ability to connect to external services allow them to act as IoT nodes but at the same time convert them to special targets for hackers. Hackers often exploit vulnerabilities that are based on the poor software implementations or even to their architecture that in most cases lacks the security and privacy design. However, in the case of s-health personalized services the risk and the impact of such attacks are notably greater and in most cases connect to the victim's health and clinical state. Therefore, the need for enhanced privacy and security safeguards is critical. The latter stems from the sensitivity of health data and the emerging threats that smart environments contain due to their high connectivity.

Another major issue of s-health and IoT-based services is that many of them, especially those that they are offered by well-known big tech companies, are the most popular targets, as they impact even more users and exploit larger databases containing personal and sensitive data. It is usual for health and medical databases to include personal data that identify certain people. Moreover, in the case of paid services, databases may include patients' financial data. These kinds of data are some of the most valuable data for cybercriminals in our days [41, 148].

The danger of an interconnected society where citizens will not change their consumer behaviors, based on the evidence that their preferred 'trusted' parties, that are used to store, share, and exchange personal and sensitive data, are compromising their privacy by sharing their data with no-governmental agencies for further un-authorized processing, is illustrating the dangerous situation of a great mass-surveillance interconnected world. It seems as EU has moved drastically, towards the delimitation of this phenomenon, by introducing laws and legislations like GDPR and NIS Directive, requiring data controllers and processors to comply with several strict security and privacy-related requirements. The latter is maybe one of the most critical and demanding research direction which comes out of the context of this thesis: the development of robust and flexible, easy to adapt security and privacy frameworks and legislations, across jurisdictions, without technological restrictions to easy plug and play in any new emerging ubiquitous and interconnected environment.

8.2 Future Directions

Inarguably, to impove the security and privacy, despite raising the user awareness, the development practicies have to be significantly improved. To this end, DevSecOps, an extension of the well-known and widely used paradigm of development lifecycle DevOps is a perfect candidate. The interweaving security tightly in the security in the design, development, and deployment of software solutions, along with automated tests guarantees that a minimum set of security controls will be integrated in the software from early stages and that bypassing these measures would be far more difficult for an adversary. Given that the concept of DevSecOps is quite new and not many automated security tools exist, a whole new field for research

is open. This includes topics including, but not limited to automated vulnerability detection, self-healing software, binary armouring, and automated fuzzing to name a few.

Despite the changes that GDPR may require [89], it is necessary to strenghten the controls and audits to software solutions. Notably, even if several big players, e.g. social networks, have been repeatedly found to violate GDPR or sensitive data have been leaked, the corresponding measures have not been applied and are either debaded or no significant changes have been introduced in the platforms to mitigate them. In fact, the dependence on social networks and media has been augmented, especially during the recent lockdowns, showcasing their critical role in our everyday lives. In this regard, user-centric approaches to notify the user of the privacy exposure towards online users, assess the aggregated information that online entities may have for individuals, and privacy-preserving methods to collect, process, and share content are in the research spotlight.

During the COVID-19 pandemic, it became apparent that the use of contact tracing apps can serve as a means to timely notify users and help in containing its spread [111, 48, 31, 80]. Despite the numerous apps that have been introduced and the sensitivity of the underlying data, not all of them have the high privacy standards and serious privacy issues emerge regarding to whom has access to what information and what can be infered from the exchanged information [6, 37]. With the introduction of vaccinations against COVID-19 and the upcoming use of the Digital Green Certificate to facilitate citizens traveling, the same privacy issues are shifted from contact tracing apps to the use of Digital Green Certificate.

Finally, with the rise of personalized medicine, the use of recommender systems and artificial intelligence algorithms introduce a set of privacy and ethical issues. For instance, and machine learning algorithms are known to have biases [99, 136, 171] and they are very relevant for healthcare [33, 109]. Many of these biases stem from imbalanced datasets used for training, however, due to the significant impact that these may have in the well-being of citizens, a more focused research in this field is necessary. Moreover, further research on whether sensitive information of individuals can be extracted from trained models is necessary to assess the exposure to such risks and possible identification.

Bibliography

[1] Developers and publishers are flocking to the mHealth app market, BI Intelligence, 2016. http://www.businessinsider.com/developers-and-publishers-are-flocking-to-the-mhealth-app-market-2016-10.

[2] Special Eurobarometer 431: Data protection, Directorate-General for Communication, 2015. https://data.europa.eu/euodp/el/data/dataset/S2075_83_1_431_ENG.

[3] Karim Abouelmehdi, Abderrahim Beni-Hssane, Hayat Khaloufi, and Mostafa Saadi. Big data security and privacy in healthcare: A review. *Procedia Computer Science*, 113:73–80, 2017.

[4] Alessandro Acquisti, Allan Friedman, and Rahul Telang. Is there a cost to privacy breaches? an event study. *ICIS 2006 Proceedings*, page 94, 2006.

[5] Rajindra Adhikari, Deborah Richards, and Karen Scott. Security and privacy issues related to the use of mobile health apps. ACIS, 2014.

[6] Nadeem Ahmed, Regio A Michelin, Wanli Xue, Sushmita Ruj, Robert Malaney, Salil S Kanhere, Aruna Seneviratne, Wen Hu, Helge Janicke, and Sanjay K Jha. A survey of covid-19 contact tracing apps. *IEEE Access*, 8:134577–134601, 2020.

[7] Mohammad Al Mamun, Hamza M Ibrahim, and Tanvir Chowdhury Turin. Peer reviewed: social media in communicating health information: an analysis of facebook groups related to hypertension. *Preventing chronic disease*, 12, 2015.

[8] Efthimios Alepis and Constantinos Patsakis. Hey doc, is this normal?: Exploring android permissions in the post marshmallow era. In *International Conference on Security, Privacy and Applied Cryptographic Engineering*. Springer, 2017.

[9] Efthimios Alepis and Constantinos Patsakis. There's wally! location tracking in android without permissions. In Paolo Mori, Steven Furnell, and Olivier Camp, editors, *Proceedings of the 3rd International Conference on Information Systems Security and Privacy, ICISSP 2017, Porto, Portugal, February 19-21, 2017.*, pages 278–284. SciTePress, 2017.

[10] Deepthi Anand and UC Niranjan. Watermarking medical images with patient information. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286)*, volume 2, pages 703–706. IEEE, 1998.

[11] Claudio A Ardagna, Mauro Conti, Mario Leone, and Julinda Stefa. An anonymous end-to-end communication protocol for mobile cloud environments. *IEEE Transactions on Services Computing*, (3):373–386, 2014.

[12] Nicola Asuni and Andrea Giachetti. Testimages: A large data archive for display and algorithm testing. *Journal of Graphics Tools*, 17(4):113–125, 2013.

[13] Giuseppe Ateniese, Kevin Fu, Matthew Green, and Susan Hohenberger. Improved proxy re-encryption schemes with applications to secure distributed storage. *ACM Transactions on Information and System Security (TISSEC)*, 9(1):1–30, 2006.

[14] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things: A survey. *Computer networks*, 54(15):2787–2805, 2010.

[15] D Audino, F Baronti, F Lenzi, R Roncella, R Saletti, and O di Tanna. A perspective on in-motorcycle electronic systems. In *Automotive Electronics, 2007 3rd Institution of Engineering and Technology Conference on*, pages 1–12. IET, 2007.

[16] Julian Backes, Michael Backes, Markus Dürmuth, Sebastian Gerling, and Stefan Lorenz. X-pire!-a digital expiration date for images in social networks. *arXiv preprint arXiv:1112.2649*, 2011.

[17] Randy Baden, Adam Bender, Neil Spring, Bobby Bhattacharjee, and Daniel Starin. Persona: an online social network with user-defined privacy. In *ACM SIGCOMM Computer Communication Review*, volume 39, pages 135–146. ACM, 2009.

[18] Mirza Mansoor Baig and Hamid Gholamhosseini. Smart health monitoring systems: an overview of design and modeling. *Journal of medical systems*, 37(2):1–14, 2013.

[19] Mirza Mansoor Baig, Hamid Gholamhosseini, and Martin J Connolly. A comprehensive survey of wearable and wireless ecg monitoring systems for older adults. *Medical & biological engineering & computing*, 51(5):485–495, 2013.

[20] Amay J Bandodkar, Denise Molinnus, Omar Mirza, Tomás Guinovart, Joshua R Windmiller, Gabriela Valdés-Ramírez, Francisco J Andrade, Michael J Schöning, and Joseph Wang. Epidermal tattoo potentiometric sodium sensors with wireless signal transduction for continuous noninvasive sweat monitoring. *Biosensors and Bioelectronics*, 54:603–609, 2014.

[21] Amay J Bandodkar and Joseph Wang. Non-invasive wearable electrochemical sensors: a review. *Trends in biotechnology*, 32(7):363–371, 2014.

[22] Gerald Bauer and Paul Lukowicz. Can smartphones detect stress-related changes in the behaviour of individuals? In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pages 423–426. IEEE, 2012.

[23] Filipe Beato, Markulf Kohlweiss, and Karel Wouters. Scramble! your social network data. In Simone Fischer-Hübner and Nicholas Hopper, editors, *Privacy Enhancing Technologies*, volume 6794 of *Lecture Notes in Computer Science*, pages 211–225. Springer Berlin Heidelberg, 2011.

[24] Natalie Berry, Fiona Lobban, Maksim Belousov, Richard Emsley, Goran Nenadic, and Sandra Bucci. # whywetweetmh: understanding why people use twitter to discuss mental health problems. *Journal of medical Internet research*, 19(4):e107, 2017.

[25] Rachel Bierbrier, Vivian Lo, and Robert C Wu. Evaluation of the accuracy of smartphone medical calculation apps. *Journal of medical Internet research*, 16(2), 2014.

[26] Leyla Bilge, Thorsten Strufe, Davide Balzarotti, and Engin Kirda. All your contacts are belong to us: Automated identity theft attacks on social networks. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, pages 551–560. ACM, 2009.

[27] Irena Bojanova and Jeffrey Voas. Trusting the internet of things. *IT Professional*, 19(5):16–19, 2017.

[28] Theodore Book, Adam Pridgen, and Dan S Wallach. Longitudinal analysis of android ad library permissions. *arXiv preprint arXiv:1303.0857*, 2013.

[29] Theodore Book and Dan S Wallach. A case of collusion: A study of the interface between ad libraries and their apps. In *Proceedings of the Third ACM workshop on Security and privacy in smartphones & mobile devices*, pages 79–86. ACM, 2013.

[30] Nidhi Bouri and Sanjana Ravi. Going mobile: how mobile personal health records can improve health care during emergencies. *JMIR mHealth and uHealth*, 2(1):e8, 2014.

[31] Isobel Braithwaite, Thomas Callender, Miriam Bullock, and Robert W Aldridge. Automated and partly automated contact tracing: a systematic review to inform the control of covid-19. *The Lancet Digital Health*, 2020.

[32] Michele Caldara, Claudio Colleoni, Emanuela Guido, Valerio Re, and Giuseppe Rosace. Development of a textile-optoelectronic ph meter based on hybrid xerogel doped with methyl red. *Sensors and Actuators B: Chemical*, 171:1013–1021, 2012.

[33] Robert Challen, Joshua Denny, Martin Pitt, Luke Gompels, Tom Edwards, and Krasimira Tsaneva-Atanasova. Artificial intelligence, bias and clinical safety. *BMJ Quality & Safety*, 28(3):231–237, 2019.

[34] Marie Chan, Daniel Estève, Jean-Yves Fourniols, Christophe Escriba, and Eric Campo. Smart wearable systems: Current status and future challenges. *Artificial intelligence in medicine*, 56(3):137–156, 2012.

[35] Gokul Chittaranjan, Jan Blom, and Daniel Gatica-Perez. Who's who with big-five: Analyzing and classifying personality traits with smartphones. In *Wearable Computers (ISWC), 2011 15th Annual International Symposium on*, pages 29–36. IEEE, 2011.

[36] Eunjoon Cho, Seth A Myers, and Jure Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090. ACM, 2011.

[37] Hyunghoon Cho, Daphne Ippolito, and Yun William Yu. Contact tracing mobile apps for covid-19: Privacy considerations and related trade-offs. *arXiv preprint arXiv:2003.11511*, 2020.

[38] European Commission. Draft Code of Conduct on privacy for mobile health applications. ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=16125.

[39] Mauro Conti, Luigi Vincenzo Mancini, Riccardo Spolaor, and Nino Vincenzo Verde. Analyzing android encrypted network traffic to identify user actions. *IEEE Transactions on Information Forensics and Security*, 11(1):114–125, 2016.

[40] Danielle L Couch, Priscilla Robinson, and Paul A Komesaroff. Covid-19—extending surveillance and the panopticon. *Journal of Bioethical Inquiry*, pages 1–6, 2020.

[41] Lynne Coventry and Dawn Branley. Cybersecurity in healthcare: a narrative review of trends, threats and ways forward. *Maturitas*, 113:48–52, 2018.

[42] Helen Cowie. Cyberbullying and its impact on young people's emotional health and well-being. *The Psychiatrist*, 37(5):167–170, 2013.

[43] I.J. Cox, M.L. Miller, and J.A. Bloom. Watermarking applications and their properties. In *Proceedings of the International Conference on Information Technology: Coding and Computing*, pages 6–10. IEEE, 2000.

[44] Shirley Coyle, K-T Lau, Niall Moyna, Donal O'Gorman, Dermot Diamond, Fabio Di Francesco, Daniele Costanzo, Pietro Salvo, Maria Giovanna Trivella, Danilo Emilio De Rossi, et al. Biotex-biosensing textiles for personalised healthcare management. *Information Technology in Biomedicine, IEEE Transactions on*, 14(2):364–370, 2010.

[45] Leucio Antonio Cutillo, Mark Manulis, and Thorsten Strufe. Security and privacy in online social networks. In *Handbook of Social Network Technologies and Applications*, pages 497–522. Springer, 2010.

[46] Leucio Antonio Cutillo, Refik Molva, and Thorsten Strufe. Safebook: A privacy-preserving online social network leveraging on real-life trust. *Communications Magazine*, 47(12):94–101, 2009.

[47] Poon Carmen CY, Liu Qing, Gao Hui, Lin Wan-Hua, and Zhang Yuan-Ting. Wearable intelligent systems for e-health. *Journal of Computing Science and Engineering*, 5(3):246–256, 2011.

[48] Aaqib Bashir Dar, Auqib Hamid Lone, Saniya Zahoor, Afshan Amin Khan, and Roohie Naaz. Applicability of mobile contact tracing in fighting pandemic (covid-19): Issues, challenges and solutions. *Computer Science Review*, page 100307, 2020.

[49] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Official Journal of the European Union, L 281, 23/11/1995, pp. 31–50.

[50] Emiliano De Cristofaro, Mark Manulis, and Bertram Poettering. Private discovery of common social contacts. In Javier Lopez and Gene Tsudik, editors, *Applied Cryptography and Network Security*, volume 6715 of *Lecture Notes in Computer Science*, pages 147–165. Springer, 2011.

[51] Paul de Hert and Vagelis Papakonstantinou. The new General Data Protection Regulation: Still a sound system for the protection of individuals? *Computer Law & Security Review*, 32(2):179–194, 2016.

[52] Isabel De la Torre-Díez, Francisco Javier Díaz-Pernas, and Míriam Antón-Rodríguez. A content analysis of chronic diseases social groups on facebook and twitter. *Telemedicine and e-Health*, 18(6):404–408, 2012.

[53] Tobias Dehling, Fangjian Gao, Stephan Schneider, and Ali Sunyaev. Exploring the far side of mobile health: information security and privacy of mobile health apps on ios and android. *JMIR mHealth and uHealth*, 3(1), 2015.

[54] Pam Dixon. Medical identity theft: The information crime that can kill you. In *The world privacy forum*, volume 2006, 2006.

[55] Mari Carmen Domingo Aladrén. Managing healthcare through social networks. *Computer*, 43(7):20–25, 2010.

[56] JohnR. Douceur. The Sybil Attack. In Peter Druschel, Frans Kaashoek, and Antony Rowstron, editors, *Peer-to-Peer Systems*, volume 2429 of *Lecture Notes in Computer Science*, pages 251–260. Springer Berlin Heidelberg, 2002.

[57] Nicole B Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13:210–230, 2007.

[58] Sascha Fahl, Marian Harbach, Thomas Muders, Lars Baumgärtner, Bernd Freisleben, and Matthew Smith. Why eve and mallory love android: An analysis of android ssl (in) security. In *Proceedings of the 2012 ACM conference on Computer and communications security*, pages 50–61. ACM, 2012.

[59] Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417, 1970.

[60] Raghu K Ganti, Fan Ye, and Hui Lei. Mobile crowdsensing: current state and future challenges. *Communications Magazine, IEEE*, 49(11):32–39, 2011.

[61] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Official Journal of the European Union, L 119 (4 May 2016), pp. 1-88.

[62] Branden Ghena, William Beyer, Allen Hillaker, Jonathan Pevarnek, and J Alex Halderman. Green lights forever: analyzing the security of traffic infrastructure. In *Proceedings of the 8th USENIX conference on Offensive Technologies*, pages 7–7. USENIX Association, 2014.

[63] Victor Gradinescu, Cristian Gorgorin, Raluca Diaconescu, Valentin Cristea, and Liviu Iftode. Adaptive traffic lights using car-to-car communication. In *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, pages 21–25. IEEE, 2007.

[64] Hans Graux. Article 29 Data Protection Working Party. ec.europa.eu/newsroom/document.cfm?doc_id=44371.

[65] Jeremy A Greene, Niteesh K Choudhry, Elaine Kilabuk, and William H Shrank. Online social networking by patients with diabetes: a qualitative evaluation of communication with facebook. *Journal of general internal medicine*, 26(3):287–292, 2011.

[66] Saikat Guha, Kevin Tang, and Paul Francis. Noyb: Privacy in online social networks. In *Proceedings of the first workshop on Online social networks*, pages 49–54, 2008.

[67] Tomàs Guinovart, Amay J Bandodkar, Joshua R Windmiller, Francisco J Andrade, and Joseph Wang. A potentiometric tattoo sensor for monitoring ammonium in sweat. *Analyst*, 138(22):7031–7038, 2013.

[68] Dongjing He, Muhammad Naveed, Carl A Gunter, and Klara Nahrstedt. Security concerns in android mhealth apps. In *AMIA Annual Symposium Proceedings*, volume 2014, page 645. American Medical Informatics Association, 2014.

[69] G. Hogben. Security issues and recommendations for online social networks. *ENISA Position Paper*, 1, 2007.

[70] Jan Holler, Vlasios Tsiatsis, Catherine Mulligan, Stefan Avesand, Stamatis Karnouskos, and David Boyle. *From Machine-to-machine to the Internet of Things: Introduction to a New Age of Intelligence*. Academic Press, 2014.

[71] Xiping Hu, Terry HS Chu, Victor CM Leung, Edith C-H Ngai, Philippe Kruchten, and Henry CB Chan. A survey on mobile social networks: Applications, platforms, system architectures, and future research directions. *IEEE Communications Surveys & Tutorials*, 17(3):1557–1581, 2014.

[72] Kit Huckvale, José Tomás Prieto, Myra Tilney, Pierre-Jean Benghozi, and Josip Car. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC medicine*, 13(1):214, 2015.

[73] Eric Hughes. A cypherpunk's manifesto. http://www.activism.net/cypherpunk/manifesto.html, 1993.

[74] Jim Isaak and Mina J Hanna. User data privacy: Facebook, cambridge analytica, and privacy protection. *Computer*, 51(8):56–59, 2018.

[75] Sonia Jahid, Prateek Mittal, and Nikita Borisov. Easier: encryption-based access control in social networks with efficient revocation. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, pages 411–415. ACM, 2011.

[76] Yunhan Jack Jia, Qi Alfred Chen, Yikai Lin, Chao Kong, and Z Morley Mao. Open doors for bob and mallory: Open port usage in android apps and security implications. In *Proceedings of the 2nd IEEE European Symposium on Security and Privacy*. IEEE, 2017.

[77] Amanda Johansson, Steven Nordin, Marina Heiden, and Monica Sandström. Symptoms, personality traits, and stress in people with mobile phone-related symptoms and electromagnetic hypersensitivity. *Journal of psychosomatic research*, 68(1):37–45, 2010.

[78] Emil Jovanov and Aleksandar Milenkovic. Body area networks for ubiquitous healthcare applications: opportunities and challenges. *Journal of medical systems*, 35(5):1245–1254, 2011.

[79] Miltiadis Kandias, Lilian Mitrou, Vasilis Stavrou, and Dimitris Gritzalis. Which side are you on? A new Panopticon vs. privacy. In *Proceedings of the 10th International Conference on Security and Cryptography (SECRYPT)*, pages 98–110, 2013.

[80] Matt J Keeling, T Deirdre Hollingsworth, and Jonathan M Read. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (covid-19). *J Epidemiol Community Health*, 74(10):861–866, 2020.

[81] Darko Kirovski, Henrique Malvar, and Yacov Yacobi. A dual watermark-fingerprint system. *Multimedia*, 11(3):59–73, 2004.

[82] Leonard J Kish and Eric J Topol. Unpatients—why patients should own their medical data. *Nature biotechnology*, 33(9):921–924, 2015.

[83] Konstantin Knorr and David Aspinall. Security testing for android mhealth apps. In *Software Testing, Verification and Validation Workshops (ICSTW), 2015 IEEE Eighth International Conference on*, pages 1–8. IEEE, 2015.

[84] Konstantin Knorr, David Aspinall, and Maria Wolters. On the privacy, security and safety of blood pressure and diabetes apps. In *IFIP International Information Security Conference*, pages 571–584. Springer, 2015.

[85] Constantinos Kolias, Angelos Stavrou, Jeffrey Voas, Irena Bojanova, and Richard Kuhn. Learning internet-of-things security" hands-on". *IEEE Security & Privacy*, 14(1):37–46, 2016.

[86] B Könings, David Piendl, Florian Schaub, and Michael Weber. Privacyjudge: Effective privacy controls for online published information. In *3rd international conference on Privacy, security, risk and trust (PASSAT), and 3rd International Conference on Social Computing (SocialCom)*, pages 935–941. IEEE, 2011.

[87] Michal Kosinski, David Stillwell, and Thore Graepel. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805, 2013.

[88] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent development and applications of sumo–simulation of urban mobility. *International Journal On Advances in Systems and Measurements*, 5(3 and 4):128–138, 2012.

[89] Mirosław Kutyłowski, Anna Lauks-Dutka, and Moti Yung. Gdpr–challenges for reconciling legal rules with technical reality. In *European Symposium on Research in Computer Security*, pages 736–755. Springer, 2020.

[90] Ieng-Fat Lam, Kuan-Ta Chen, and Ling-Jyh Chen. Involuntary information leakage in social network services. In *Advances in Information and Computer Security*, pages 167–183. Springer, 2008.

[91] Ming Li, Ning Cao, Shucheng Yu, and Wenjing Lou. Findu: Privacy-preserving personal profile matching in mobile social networks. In *Proceedings of INFOCOM*, pages 2435–2443. IEEE, 2011.

[92] Matthew M Lucas and Nikita Borisov. Flybynight: mitigating the privacy risks of social networking. In *Proceedings of the 7th ACM workshop on Privacy in the Electronic Society*, pages 1–8. ACM, 2008.

[93] Wanying Luo, Qi Xie, and Urs Hengartner. Facecloak: An architecture for user privacy on social networking sites. In *International Conference on Computational Science and Engineering (CSE'09)*, volume 3, pages 26–33. IEEE, Aug 2009.

[94] Burton G Malkiel. Efficient market hypothesis. In *Finance*, pages 127–134. Springer, 1989.

[95] Michelino Mancini. Medical identity theft in the emergency department: awareness is crucial. *Western Journal of Emergency Medicine*, 15(7):899, 2014.

[96] Vincenzo Manzoni, Andrea Corti, Cristiano Spelta, and Sergio M Savaresi. A driver-to-infrastructure interaction system for motorcycles based on smartphone. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, pages 1442–1447. IEEE, 2010.

[97] Izet Masic, Suad Sivic, Selim Toromanovic, Tea Borojevic, and Haris Pandza. Social networks in improvement of health care. *Materia socio-medica*, 24(1):48, 2012.

[98] Viktor Mayer-Schönberger. *Delete: the virtue of forgetting in the digital age*. Princeton University Press, 2011.

[99] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. *arXiv preprint arXiv:1908.09635*, 2019.

[100] Faye Mishna, Mona Khoury-Kassabri, Tahany Gadalla, and Joanne Daciuk. Risk factors for involvement in cyber bullying: Victims, bullies and bully–victims. *Children and Youth Services Review*, 34(1):63–70, 2012.

[101] Masakazu Miyamae, Tsutomu Terada, Masahiko Tsukamoto, and Shojiro Nishio. An event-driven wearable system for supporting motorbike racing teams. In *Wearable Computers, 2004. ISWC 2004. Eighth International Symposium on*, volume 1, pages 70–76. IEEE.

[102] Sai T Moturu, Inas Khayal, Nadav Aharony, Wei Pan, and A Pentland. Sleep, mood and sociability in a healthy population. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 5267–5270. IEEE, 2011.

[103] Sai T Moturu, Inas Khayal, Nadav Aharony, Wei Pan, and Alex Sandy Pentland. Using social sensing to understand the links between sleep, mood, and sociability. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 208–214. IEEE, 2011.

[104] Amir Muaremi, Bert Arnrich, and Gerhard Tröster. Towards measuring stress with smartphones and wearable devices during workday and sleep. *BioNanoScience*, 3(2):172–183, 2013.

[105] Shah Nazir, Yasir Ali, Naeem Ullah, and Iván García-Magariño. Internet of things for healthcare using effects of mobile computing: A systematic literature review. *Wireless Communications and Mobile Computing*, 2019, 2019.

[106] Craig Michael Lie Njie. Technical analysis of the data practices and privacy risks of 43 popular mobile health and fitness applications. *Research Performed For: Privacy Rights Clearinghouse*, 2013.

[107] Achilleas Papageorgiou, Michael Strigkos, Eugenia Politou, Efthimios Alepis, Agusti Solanas, and Constantinos Patsakis. Security and privacy analysis of mobile health applications: The alarming state of practice. *IEEE Access*, 2018.

[108] Achilleas Papageorgiou, Athanasios Zigomitros, and Constantinos Patsakis. Personalising and crowdsourcing stress management in urban environments via s-health. In *Information, Intelligence, Systems and Applications (IISA), 2015 6th International Conference on*, pages 1–4. IEEE, 2015.

[109] Ravi B Parikh, Stephanie Teeple, and Amol S Navathe. Addressing bias in artificial intelligence in health care. *Jama*, 322(24):2377–2378, 2019.

[110] Namje Park. Secure data access control scheme using type-based re-encryption in cloud environment. In *Semantic Methods for Knowledge Management and Communication*, pages 319–327. Springer, 2011.

[111] Young Joon Park, Young June Choe, Ok Park, Shin Young Park, Young-Man Kim, Jieun Kim, Sanghui Kweon, Yeonhee Woo, Jin Gwack, Seong Sun Kim, et al. Contact tracing during coronavirus disease outbreak, south korea, 2020. *Emerging infectious diseases*, 26(10):2465–2468, 2020.

[112] Constantinos Patsakis, Alexandros Asthenidis, and Abraham Chatzidimitriou. Social networks as an attack platform: Facebook case study. In *8th International Conference on Networks*, pages 245–247. IEEE, 2009.

[113] Constantinos Patsakis, Athanasios Charemis, Achilleas Papageorgiou, Dimitrios Mermigas, and Sotirios Pirounias. The market's response toward privacy and mass surveillance: The snowden aftermath. *Computers & Security*, 73:194–206, 2018.

[114] Constantinos Patsakis, Achilleas Papageorgiou, Francisco Falcone, and Agusti Solanas. s-health as a driver towards better emergency response systems in urban environments. In *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*, pages 214–218. IEEE, 2015.

[115] Constantinos Patsakis and Agusti Solanas. Privacy as a product: A case study in the m-health sector. In *4th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pages 1–6. IEEE, July 2013.

[116] Constantinos Patsakis and Agusti Solanas. Privacy-aware event data recorders: cryptography meets the automotive industry again. *IEEE Communications Magazine*, 51(12):122–128, 2013.

[117] Constantinos Patsakis, Riccardo Venanzio, Paolo Bellavista, Agusti Solanas, and Mélanie Bouroche. Personalized medical services using smart cities' infrastructures. In *Medical Measurements and Applications (MeMeA), 2014 IEEE International Symposium on*, pages 1–5. IEEE, 2014.

[118] Constantinos Patsakis, Athanasios Zigomitros, Achilleas Papageorgiou, and Edgar Galván-López. Distributing privacy policies over multimedia content

across multiple online social networks. *Computer Networks*, 75:531–543, 2014.

[119] Constantinos Patsakis, Athanasios Zigomitros, Achilleas Papageorgiou, and Agusti Solanas. Privacy and security for multimedia content shared on osns: issues and countermeasures. *The Computer Journal*, 58(4):518–535, 2015.

[120] Constantinos Patsakis, Athanasios Zigomitros, and Agusti Solanas. Analysis of privacy and security exposure in mobile dating applications. In *International Conference on Mobile, Secure and Programmable Networking*, pages 151–162. Springer, 2015.

[121] J Peen, RA Schoevers, AT Beekman, and J Dekker. The current status of urban-rural differences in psychiatric disorders. *Acta Psychiatrica Scandinavica*, 121(2):84–93, 2010.

[122] Anna-Kaisa Pietiläinen, Earl Oliver, Jason LeBrun, George Varghese, and Christophe Diot. Mobiclique: middleware for mobile social networking. In *Proceedings of the 2nd ACM workshop on Online social networks*, pages 49–54, 2009.

[123] Sören Preibusch. Privacy behaviors after snowden. *Communications of the ACM*, 58(5):48–55, 2015.

[124] Guojun Qin, Constantinos Patsakis, and Mélanie Bouroche. Playing hide and seek with mobile dating applications. In Nora Cuppens-Boulahia, Frédéric Cuppens, Sushil Jajodia, Anas Abou El Kalam, and Thierry Sans, editors, *ICT Systems Security and Privacy Protection*, volume 428 of *IFIP Advances in Information and Communication Technology*, pages 185–196. Springer Berlin Heidelberg, 2014.

[125] Mohd Rasli, Mohd Khairul Afiq, Nina Korlina Madzhi, and Juliana Johari. Smart helmet with sensors for accident prevention. In *Electrical, Electronics and System Engineering (ICEESE), 2013 International Conference on*, pages 21–26. IEEE, 2013.

[126] Marc A Rodwin. The case for public ownership of patient data. *Jama*, 302(1):86–88, 2009.

[127] Shouq A Sadah, Moloud Shahbazi, Matthew T Wiley, and Vagelis Hristidis. A study of the demographics of web-based health-related social media users. *Journal of medical Internet research*, 17(8):e194, 2015.

[128] Pietro Salvo, Fabio Di Francesco, Daniele Costanzo, Carlo Ferrari, Maria Giovanna Trivella, and Danilo De Rossi. A wearable sensor for measuring sweat rate. *Sensors Journal, IEEE*, 10(10):1557–1558, 2010.

[129] Akane Sano and Rosalind W Picard. Stress recognition using wearable sensors and mobile phones. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 671–676. IEEE, 2013.

[130] Karen M Scott, Gastao A Gome, Deborah Richards, and Patrina HY Caldwell. How trustworthy are apps for maternal and child health? *Health and Technology*, 4(4):329–336, 2015.

[131] Dhruv R Seshadri, Ryan T Li, James E Voos, James R Rowbottom, Celeste M Alfes, Christian A Zorman, and Colin K Drummond. Wearable sensors for monitoring the physiological and biochemical profile of the athlete. *NPJ digital medicine*, 2(1):1–16, 2019.

[132] M Shelton, L Rainie, M Madden, M Anderson, M Duggan, A Perrin, and D Page. American's privacy strategies post-snowden. *Pew Research Center, Washington, USA*, 2015.

[133] Changgui Shi and Bharat Bhargava. A fast MPEG video encryption algorithm. In *Proceedings of the 6th ACM International Conference on Multimedia*, pages 81–88. ACM, 1998.

[134] Dong-Hee Shin. The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with computers*, 22(5):428–438, 2010.

[135] Laura Shipp and Jorge Blasco. How private is your period?: A systematic analysis of menstrual app privacy policies. *Proceedings on Privacy Enhancing Technologies*, 2020(4):491–510, 2020.

[136] Jake Silberg and James Manyika. Notes from the ai frontier: Tackling bias in ai (and in humans). *McKinsey Global Institute (June 2019)*, 2019.

[137] Agusti Solanas, Constantinos Patsakis, Mauro Conti, Ioannis S Vlachos, Victoria Ramos, Francisco Falcone, Octavian Postolache, Pablo A Pérez-Martínez, Roberto Di Pietro, Despina N Perrea, et al. Smart health: a context-aware health paradigm within smart cities. *IEEE Communications Magazine*, 52(8):74–81, 2014.

[138] Agusti Solanas, Constantinos Patsakis, Mauro Conti, Ioannis S. Vlachos, Victoria Ramos, Francisco Falcone, Octavian Postolache, Pablo A. Pérez-Martínez, Roberto Di Pietro, Despina N. Perrea, and Antoni Martínez-Ballesté. Smart health: A context-aware health paradigm within smart cities. *IEEE Communications Magazine*, 52(8):74–81, 2014.

[139] Agusti Solanas, Jens H. Weber, Ayse B. Bener, Frank van der Linden, and Rafael Capilla. Recent advances in healthcare software: Toward context-aware and smart solutions. *IEEE Software*, 34(6):36–40, 2017.

[140] Daniel J Solove. I've got nothing to hide and other misunderstandings of privacy. *San Diego L. Rev.*, 44:745, 2007.

[141] Anna Spagnolli, Enrico Guardigli, Valeria Orso, Alessandra Varotto, and Luciano Gamberini. Measuring user acceptance of wearable symbiotic devices: validation study across application scenarios. In *International Workshop on Symbiotic Interaction*, pages 87–98. Springer, 2014.

[142] Georgios Spanos and Lefteris Angelis. The impact of information security events to the stock market: A systematic literature review. *Computers & Security*, 58:216–229, 2016.

[143] Chad Spensky, Jeffrey Stewart, Arkady Yerukhimovich, Richard Shay, Ari Trachtenberg, Rick Housley, and Robert K Cunningham. Sok: Privacy on mobile devices–it's complicated. *Proceedings on Privacy Enhancing Technologies*, 2016(3):96–116, 2016.

[144] Anna C Squicciarini, Mohamed Shehab, and Joshua Wede. Privacy policies for shared content in social network sites. *The VLDB Journal–The International Journal on Very Large Data Bases*, 19(6):777–796, 2010.

[145] Anna Cinzia Squicciarini, Mohamed Shehab, and Federica Paci. Collective privacy management in social networks. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, pages 521–530, 2009.

[146] Stephen A Stansfeld and Mark P Matheson. Noise pollution: non-auditory effects on health. *British Medical Bulletin*, 68(1):243–257, 2003.

[147] Katherine Strater and Heather Richter. Examining privacy and disclosure in a social networking community. In *Proceedings of the 3rd symposium on Usable privacy and security*, pages 157–158. ACM, 2007.

[148] Aatif Sulleyman. Nhs cyber attack: why stolen medical information is so much more valuable than financial data. *Independent (2017).[Online*, 2017.

[149] Ali Sunyaev, Tobias Dehling, Patrick L Taylor, and Kenneth D Mandl. Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association*, pages amiajnl–2013, 2014.

[150] Melanie Swan. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. *International journal of environmental research and public health*, 6(2):492–525, 2009.

[151] Bethany Tennant, Michael Stellefson, Virginia Dodd, Beth Chaney, Don Chaney, Samantha Paige, and Julia Alber. ehealth literacy and web 2.0 health information seeking behaviors among baby boomers and older adults. *Journal of medical Internet research*, 17(3):e70, 2015.

[152] Kurt Thomas, Chris Grier, and DavidM. Nicol. unfriendly: Multi-party privacy risks in social networks. In MikhailJ. Atallah and NicholasJ. Hopper, editors, *Privacy Enhancing Technologies*, volume 6205 of *Lecture Notes in Computer Science*, pages 236–252. Springer Berlin Heidelberg, 2010.

[153] Sara Thomée, Annika Härenstam, and Mats Hagberg. Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults-a prospective cohort study. *BMC public health*, 11(1):66, 2011.

[154] Suvi Tiinanen, A Matta, Minna Silfverhuth, Kalervo Suominen, Eira Jansson-Verkasalo, and T Seppanen. Hrv and eeg based indicators of stress in children with asperger syndrome in audio-visual stimulus test. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 2021–2024. IEEE, 2011.

[155] Amin Tootoonchian, Stefan Saroiu, Yashar Ganjali, and Alec Wolman. Lockr: better privacy for social networks. In *Proceedings of the 5th International conference on Emerging networking experiments and technologies*, pages 169–180. ACM, 2009.

[156] Tom H Van De Belt, Lucien JLPG Engelen, Sivera AA Berben, and Lisette Schoonhoven. Definition of health 2.0 and medicine 2.0: a systematic review. *Journal of medical Internet research*, 12(2):e18, 2010.

[157] C Lee Ventola. Mobile devices and apps for health care professionals: uses and benefits. *Pharmacy and Therapeutics*, 39(5):356, 2014.

[158] Alexandre Viejo, Jordi Castella-Roca, and Guillem Rufián. Preserving the user's privacy in social networking sites. In *Trust, Privacy, and Security in Digital Business*, pages 62–73. Springer, 2013.

[159] Sviatolsav Voloshynovskiy, Shelby Pereira, Thierry Pun, Joachim J Eggers, and Jonathan K Su. Attacks on digital watermarks: classification, estimation based attacks, and benchmarks. *Communications Magazine*, 39(8):118–126, 2001.

[160] Mitsuhiro Wada, Tomohiro Yendo, Toshiaki Fujii, and Masayuki Tanimoto. Road-to-vehicle communication using led traffic light. In *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*, pages 601–606. IEEE, 2005.

[161] Guojun Wang, Qin Liu, and Jie Wu. Hierarchical attribute-based encryption for fine-grained access control in cloud storage services. In *Proceedings of the 17th ACM Conference on Computer and Communications Security*, pages 735–737. ACM, 2010.

[162] Roy Want, Bill N Schilit, and Scott Jenson. Enabling the internet of things. *Computer*, 48(1):28–35, 2015.

[163] Samuel D Warren and Louis D Brandeis. Right to privacy. *Harv. L. Rev.*, 4:193, 1890.

[164] Peter Wayner. *Disappearing cryptography: information hiding: steganography & watermarking*. Morgan Kaufmann, 2009.

[165] W Wen. An intelligent traffic management expert system with rfid technology. *Expert Systems with Applications*, 37(4):3024–3035, 2010.

[166] Marco Wiering, Jelle Van Veenen, Jilles Vreeken, and Arne Koopman. Intelligent traffic light control. *Institute of Information and Computing Sciences. Utrecht University*, 2004.

[167] Joshua Ray Windmiller and Joseph Wang. Wearable electrochemical sensors and biosensors: a review. *Electroanalysis*, 25(1):29–46, 2013.

[168] Ping Wah Wong and Nasir Memon. Secret and public key image watermarking schemes for image authentication and ownership verification. *Transactions on Image Processing*, 10(10):1593–1601, Oct 2001.

[169] Liang Wu, Fred Morstatter, Kathleen M Carley, and Huan Liu. Misinformation in social media: definition, manipulation, and detection. *ACM SIGKDD Explorations Newsletter*, 21(2):80–90, 2019.

[170] Heetae Yang, Jieun Yu, Hangjung Zo, and Munkee Choi. User acceptance of wearable devices: An extended perspective of perceived value. *Telematics and Informatics*, 33(2):256–269, 2016.

[171] Adrienne Yapo and Joseph Weiss. Ethical implications of bias in machine learning. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.

[172] Ali Alper Yayla and Qing Hu. The impact of information security events on the stock value of firms: The effect of contingency factors. *Journal of Information Technology*, 26(1):60–77, 2011.

[173] Shucheng Yu, Cong Wang, Kui Ren, and Wenjing Lou. Achieving secure, scalable, and fine-grained data access control in cloud computing. In *Proceedings of IEEE INFOCOM, 2010*, pages 1–9. IEEE, 2010.

[174] Kuan Zhang, Kan Yang, Xiaohui Liang, Zhou Su, Xuemin Shen, and Henry H Luo. Security and privacy for mobile healthcare networks: from a quality of protection perspective. *IEEE Wireless Communications*, 22(4):104–112, 2015.

[175] Athanasios Zigomitros, Fran Casino, Agusti Solanas, and Constantinos Patsakis. A survey on privacy properties for data publishing of relational data. *IEEE Access*, 8:51071–51099, 2020.

[176] Athanasios Zigomitros, Achilleas Papageorgiou, and Constantinos Patsakis. Social network content management through watermarking. In *IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pages 1381–1386. IEEE, 2012.