UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Course guide
230990 - DP - Data Privacy

Last modified: 26/05/2023

Unit in charge: Barcelona School of Telecommunications Engineering
Teaching unit: 744 - ENTEL - Department of Network Engineering.

Degree: MASTER'S DEGREE IN CYBERSECURITY (Syllabus 2020). (Optional subject).

Academic year: 2023 **ECTS Credits:** 5.0 **Languages:** English

LECTURER

Coordinating lecturer: Consultar aquí / See here:
https://telecos.upc.edu/ca/estudis/curs-actual/professorat-responsables-coordinadors/responsables-assignatura

Others: Consultar aquí / See here:
https://telecos.upc.edu/ca/estudis/curs-actual/professorat-responsables-coordinadors/professorat-assignat-idioma

TEACHING METHODOLOGY

- Lectures
- Laboratory practices
- Individual work (non-contact)
- Work in group
- Short answer tests (Test)
- Long answer tests (Final Exam)

LEARNING OBJECTIVES OF THE SUBJECT

Data privacy is the adaptation to the Information Society of the fundamental right to privacy and private life. This course focuses on different technological aspects to ensure information privacy, including:

- Present the main privacy metrics proposed in the literature, both qualitative and especially quantitative ones.
- Assess risks and analyze privacy technologies in the field of statistical disclosure control (SDC).
- Assess risks and analyze privacy technologies in personalized information systems.
- Assess risks and analyze privacy technologies in anonymous communications systems.
- Understand the inherent trade-off between privacy and data usability.

STUDY LOAD

Type	Hours	Percentage
Hours large group	26,0	20.80
Hours small group	13,0	10.40
Self study	86,0	68.80

Total learning time: 125 h

CONTENTS

1. Introduction

Description:
In this first topic, it is intended to give a global vision of the subject, with special emphasis on the motivation and the legal (and also ethical) requirements that support the privacy protection techniques.

Specific objectives:
- Offer an overview of the subject.
- Motivate the need to protect the privacy of users in information and communication systems.
- Introduce the legal framework and the ethical implications that require the protection of privacy.
- Introduce the fields of application of privacy technologies.
- Define basic concepts of the subject's terminology.
- Present the trust models on which the subject is based, identifying the attackers from whom we are trying to protect ourselves.

Full-or-part-time: 7h
Theory classes: 3h
Self study : 4h

2. Database anonymization

Description:
This topic deals with the anonymization of databases, possibly one of the most relevant fields of application today for students entering the labor market. Risk estimation techniques are presented, as well as different anonymity techniques and algorithms, for both categorical and numerical data. Special emphasis is placed on the compromise between privacy and usability, also considering metrics for loss of usability of anonymized data.

Specific objectives:
- Present the risks of publication of databases that include personal information of users.
- Formulate the compromise between privacy and usability inherent to disruptive techniques.
- Present the main techniques for estimating the risks of disclosure of private information in statistical databases.
- Present the anonymity algorithms for categorical data.
- Present the anonymity algorithms for numerical data.
- Present synthetic data generation techniques.
- Be able to measure the loss of usability of anonymized data.

Full-or-part-time: 24h
Theory classes: 9h
Self study : 15h

3. Privacy in personalized information systems

Description:
This topic covers privacy mechanisms based on data disturbance that are used in fields other than SDC. On the one hand, the problem of privacy in personalized information systems and recommendation systems is studied. On the other hand, the issue of privacy is addressed in applications that have access to geo-located information.

Specific objectives:
- Introduce privacy challenges in recommendation and personalized systems, including online advertising systems.
- Quantify the privacy risk inherent in the creation of user profiles by third parties.
- Present different strategies for generating false queries, including TrackmeNot.
- Present protection mechanisms in systems based on semantic annotations.
- Introduce the privacy problem in the location
- Highlight the compromise between privacy and usability, inherent in any system based on data disturbance.

Full-or-part-time: 9h
Theory classes: 3h
Self study : 6h

4. Anonymous communication systems

Description:
Even when communications are encrypted, the attacker can get sensitive by analyzing the protocol headers and inferring who is communicating with whom and when. Anonymous communication systems, which are studied in this topic, allow you to protect yourself against these traffic analysis attacks.

Specific objectives:
- Present traffic analysis techniques and justify the need for anonymous communications.
- Define the basic terminology of anonymous communication systems.
- Present and analyze anonymous communication systems based on Mix Networks.
- Present and analyze anonymous communications systems based on Onion Routing.
- Present and analyze anonymous communication systems based on the collaboration of users in P2P networks.

Full-or-part-time: 9h
Theory classes: 3h
Self study : 6h

5. Differential privacy

Description:
content english

Specific objectives:
This topic introduces the main concepts related to differential privacy

Full-or-part-time: 23h
Theory classes: 8h
Self study : 15h

6. Laboratory. Data anonymization practices

Description:
The main objective of the proposed practices is to support the Data Privacy subject, so that the student deepens through experimentation in different important aspects, especially those related to data anonymization.

Specific objectives:
- Introduction to Statistical Disclosure Control (SDC)
- Introduction to the R programming language
- Disclosure risk assessment
- Measurement of the usefulness of a data set.
- Anonymization methods for categorical variables
- Methods of anonymization of numerical variables
- Differential privacy

Full-or-part-time: 49h
Laboratory classes: 13h
Self study : 36h

GRADING SYSTEM

- This subject has theory (65%) and laboratory (35%) assessment.
- The theory grade is based on the results of two tests performed during the course. Additionally, there is a final exam for those who do not pass the tests.
- The laboratory grade is obtained by carrying out the practical exercises and delivering the corresponding reports.

BIBLIOGRAPHY

Basic:
- Torra i Reventós, Vicenç. Data privacy : foundations, new developments and the big data challenge. Skövde: Springer, 2017. ISBN 9783319573564.
- Templ, Matthias. Statistical disclosure control for microdata : methods and applications in R [on line]. Cham, Switzerland: Springer International Publishing AG, 2017 [Consultation: 05/07/2021]. Available on: https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=4855639. ISBN 9783319502724.
- Hundepool, Anco; Domingo-Ferrer, Josep. Statistical disclosure control [on line]. Chichester, West Sussex, United Kingdom: John Wiley & Sons Inc, [2012] [Consultation: 05/07/2021]. Available on: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118348239. ISBN 9781118348239.