LiTH, Linköpings tekniska högskola

IDA, Department of Computer and Information Science
Ulf Kargén

Distance exam
TDDD17 Information Security
2022-03-24

Teacher on duty
Ulf Kargén, ulf.kargen@liu.se, 013-285876

Instructions
There are 4 main questions on the exam. Your grade will depend on the total points you score. The maximum number of points is 32.

Since the number of credits for the exam have changed from 3 to 2 hp (or ECTS) in 2020, we use two different grading scales depending on which year you were admitted to the course, as shown in the table below. The grading scales are preliminary and might be adjusted during grading.

Grade	C (3)	B (4)	A (5)
Points required **TEN2 (3 hp)** Students admitted **2019 or earlier**	19	24	27
Points required **TEN3 (2 hp)** Students admitted **2020 or later**	16[1]	22	26

Answers should be submitted through Lisam as a single PDF or ASCII text (.txt) document before the end of the exam time. Figures should be integrated into the document and *not* submitted as separate files. If you use PDF format, figures could be drawn either electronically, or by hand and scanned/photographed. If you use plain text format, you can draw ASCII art. To facilitate easier grading, it is preferred that you express formulas, pseudo-code, etc., as electronic text, and not as handwritten figures.

You are allowed to use any aid, but **all kinds of collaboration with others is strictly forbidden**. Also, **copying any part of an answer from another source will be considered plagiarism**. The questions will be checked for plagiarism and sharing of answers between students using Urkund, and any suspected cheating will be reported to the university disciplinary board. **By taking the exam, you solemnly promise to abide by the above rules.**

In the event that you experience technical difficulties with Lisam that prevent you from submitting, it is allowable to submit answers via email. However, this should only be used as a last resort if Lisam for whatever reason would stop functioning.

The home exam will not be anonymous due to the exceptional COVID-19 situation.

Ulf Kargén will be available to answer questions during the exam via email and phone.

[1] And to those of you who would like to point out that 16/19 > 2/3, I say that the number of points you score on the exam is not a linear function of how many hours you studied.

1. System Security (8 points)

a) *Sandboxing* is seeing increasing use in modern systems as a means to improve security. Briefly explain what sandboxing is. (1 points)

b) Give a real-life practical example of the use of sandboxing in an existing system/technology. Briefly explain how sandboxing improves security in the particular example you choose. (2 points)

c) The concept of sandboxing relates to (at least) two of Saltzer and Schroeder's secure design principles: the principles of *Least privilege* and *Least common mechanism*. Explain how. Also, for each of the two design principles, give a practical example that shows how sandboxing relates to the principle. (3 points)

d) Consider the case where an attacker is able to compromise the BIOS/firmware, boot loader, or OS kernel. Compare how this would impact the security provided by respectively Intel SGX and ARM TrustZone. Briefly motivate your answer based on technical characteristics of the two systems. (2 points).

2. Defense against Malware (8 points)

a) Consider an attempt at creating a machine-learning based detector for *advanced persistent threats* (APTs). Explain why overfitting would likely be a significant problem in this case. (2 points)

b) Traditional computer viruses are almost unheard of today. Briefly explain why. (1 points)

c) Imagine that someone wanted to create a computer virus that targeted Android apps (i.e., if a user installed and ran an infected app, the virus would attempt to also infect other apps on the device). Would this be feasible? Explain why or why not. (1 point)

d) Consider a typical traditional antivirus engine (i.e., not using machine-learning or cloud-based detection) that is presented with an executable file. Outline and briefly explain the sequence of steps/techniques that the engine will likely go through in order to determine if the file is malicious or benign. Make sure to clearly explain the order in which the steps are taken, and how they contribute to deciding if the file is malicious or not. (4 points)

3. Network Security (10 points)

a) What is DarkNet, what technology is it using, how to open its web sites? (2 points)

b) How are CAs used to secure HTTP traffic? What does certificate transparency mean? How are certificates revoked? (4 points)

c) What new features does the WPA3 standard bring? What is its deployment status? (4 points)

4. Privacy (6 points)

a) A communication system that encrypts the content of the message between two or more parties guarantees a weak form privacy, because of the presence of meta data. Therefore, the state-of-the-art privacy preserving communication systems guarantee at the bare minimum the *anonymity* and *unlinkability* privacy properties. Explain your understanding of these two privacy properties (2 points).

b) Consider the following two tables, T and E.

Table. T

Age	Weight	Postal Code	Disease
21	70	311	Arthritis
21	71	377	Cold
20	72	483	Flu
20	72	377	Arthritis
27	69	353	Cold
24	70	377	Flu

Table. E

Name	Age	Weight
Bob	21	71
Turdy	21	71
Alice	20	70
Dave	20	72
Charlie	24	70

Suppose that the attribute Disease in table T is a sensitive attribute and all other attributes of T (i.e., Age, Weight and Postal Code) are not sensitive attributes. Further, assume that table E is a publicly available table populated with data about all persons in the postal code area 377.

Given this publicly available data, list all the quasi-identifiers of table T. Notice that there might be multiple different quasi-identifiers: if this is the case, you have to list all of them. Further, intuitively show how the table T can satisfy k-anonymity (draw your solution). (2 points)

c) To achieve the Differential Privacy definition for queries that return real-numbers, random numbers (noise) drawn from Laplace distribution are injected into the query results. In this context, explain the relationship between the amount of noise, the privacy loss parameter (epsilon) and the sensitivity of the given query. (2 points)