IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2014-2015

MEng Honours Degrees in Computing Part IV
MSc in Advanced Computing
MSc in Computing Science (Specialist)
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute

PAPER C408H

PRIVACY ENHANCING TECHNIQUES

Thursday 18 December 2014, 14:00
Duration: 60 minutes

Answer TWO questions

Paper contains 3 questions
Calculators not required

1a Briefly describe two problems from Solove's Taxonomy. The two problems must be from different categories.

b i) What are the advantages and disadvantages of having more attributes in a quasi-identifier?

ii) Consider a table that contains no sensitive attributes. For this table explain whether *k*-anonymity can be replaced by an attribute linkage model like *l*-diversity?

c Consider the following randomized response mechanism. Given the query "*Have you engaged in crime X?*" the respondent is instructed to perform the following steps:

1. Flip a coin.
2. If the result is tails, then respond truthfully.
3. If the result is heads, then flip a second coin and respond "*Yes*" if the result is heads and "*No*" if the result is tails.

For this mechanism, the response "*Yes*" can be achieved in two cases: if the person was indeed involved in crime *X* and the first coin came up tails, or if both the first and second coins came up heads.

Show that the described randomized response mechanism is ($\log_e 3$)-differentially private.

d Alice's family *may* have been infected after contacting a person that had a highly contagious disease Y. If anyone in Alice's family has been infected then everyone in Alice's family has been infected. There are 5 people in Alice's family including Alice.

Everyone in Alice's family goes for a test to check if they are infected. The results are held in a statistical database. Show that in order to preserve the privacy of the family for $\varepsilon = 0.1$ that noise $\sim$~~Lap(2)~~ has to be added to the → Lap(50) response for the query "*how many people have disease Y?*"

Hint: think about the ε-differential privacy for Alice's family of 5 individuals.

The four parts carry, respectively, 20%, 20%, 40% and 20% of the marks.

2 Three privacy campaigners are having dinner around a table at a restaurant in South Kensington. The restaurant owner tells them that their dinner has been paid for, anonymously, either by one of them, or by Bookface. One of the campaigners is unhappy with having their meal paid for by Bookface, so they decide to find out whether it is Bookface who has paid, or one of them, without exposing which one of them it is. They devise the following protocol:

1. Each campaigner flips a coin and shows it to their left neighbour, i.e. each campaigner will see the outcome of two coin flips: their own and that of the right neighbour.

2. Each campaigner then announces whether the outcomes of the two coin flips that they have seen are the "*Same*" or "*Different*". If the campaigner is the payer, the campaigner says the opposite (i.e. lies).

For this protocol show that:

i) An odd number of "*Same*" announcements means that Bookface is paying, while an even number means that one of them is paying.

ii) A non-paying campaigner cannot tell which of the other two is the payer, if Bookface is not paying.

The two parts (i) and (ii) carry, respectively, 50% and 50% of the marks.

3a Outline 4 properties that an *ideal* crypto-currency should have.

b Bitcoin mining is sometimes seen as the most anonymous method to acquire Bitcoins. Suggest two ways that an adversary might still be able to undermine anonymity at this stage.

c i) What are Bitcoin mixers?

ii) How do they operate and what features would you look for in a Bitcoin mixer?

iii) A Bitcoin mixer could be compromised. How would you mitigate against this?

The three parts carry, respectively, 40%, 10% and 50% of the marks.