Department of Computing

COMP3300 Assignment Part (a)

Assignment marks: 6% of overall unit marks.

Objective: To gain understanding of analysing and measuring information leaks in deterministic cryptography.

Please note: This assignment specification aims to provide as complete a description of this assessment task as possible. However, as with any specification, there will always be things we should have said that we have left out and areas in which we could have done a better job of explanation. As a result, you are strongly encouraged to ask any questions of clarification you might have, either by raising them during a lecture or by posting them on the iLearn discussion forum devoted to this assignment.

Deterministic Encryption and the potential for privacy breaches

Deterministic encryption enables two encrypted cypher texts to be classified as "the same" or different without having to perform the decryption step. This means encrypted databases are potentially vulnerable to inference attacks in that an adversary is able to perform a number of experiments can find out which texts are the same. This becomes a real vulnerability when the adversary also knows what the potential contents are in the database.

For example in a hospital database, specific illnesses of patients are recorded (refer to Table 1 below). This means that in the encrypted database, a preponderance of information is available to be used as prior knowledge, specifically the number of different possible cypher texts together with some statistical information about the proportion of those possibilities. This information can be put together with the experiments to determine which files are the same and which are different. All of this can then be summarised in terms of a channel and prior, so that the impact of privacy breach based on an inference attack can be measured.

In this assignment you are asked to carry out such a modelling exercise on a small case study and to report your answers using the iLearn Submission system.

Table 1: Top 10 major diagnostic categories at a large inpatient acute palliative care service [33].

Diagnosis	Percentage
Cancer	41.3%
Cardiac Disease	17.4%
Pulmonary Disease	14.0%
Stroke	9.4%
Renal Disease	3.5%
Dementia	2.4%
Liver Disease	1.7%
AIDS	0.4%
Lou Gehrig's Disease	0.2%
Other	9.6%

Databases which record the illness category of patients. In the below this will correspond to k=10.

This assignment: Modelling details and assumptions

In this assignment we will make some reasonable assumptions about the potential privacy breach, but then describe a very small scenario so that it is feasible for you to carry out the detailed modelling and calculations.

You may use Jupyter notebooks for this, but it is not compulsory. A notebook for this assignment is available to download.

Assumptions

The database consists of N cypher texts referred to as c_1, c_2, … c_N. Each cypher text refers to a particular patient's disease, eg c_1 is an encryption of patient ident 1's disease, c_2 is an encryption of patient ident 2's disease etc.

The cypher texts are all encryptions of a fixed k number of clear texts (eg the category of illness). We'll refer to this different k values as a, b, c, etc.

The general population profile of the k values is known, and this will be used to form a prior distribution about the potential secrets.

Because of the capabilities of the deterministic encryption defined above the information that it is possible for searches and experimentation performed by the adversary to observe which of the N cypher texts are the same as others. We do not take into account the time it takes the adversary to tabulate his observations, and so in this sense we are carrying out a **worst case privacy analysis**. In any case we summarise these observations in a channel using the following table, for N=3 and k=2. In this situation there are only four possible observations.

Observation	**c_1 = c_2 = c_3**	**c_2 = c_3 (but c_1 ≠ c_2, c_3)**	**c_1 = c_2 (but c_3 ≠ c_2, c_1)**	**c_1 = c_3 (but c_2 ≠ c_1, c_3)**
Short code	A	B	C	D

For N = 3 we write the possible values of the three cypher texts as a triple (x, y, z), which means that cypher text c_1 is an encryption of text x, cypher text c_2 is an encryption of text y, and cypher text c_3 is an encryption of text z. For this small study we shall also assume that k=2, so that x, y and z are all drawn from the set {a, b}. Notice that the adversary's observations cannot tell us exactly whether an encryption is a or b, only whether the encryptions are the same or different.

Using this information we can now create an information flow channel to summarise all of the information available to an adversary so that we can measure the impact of his various inference attacks.

	A	**B**	**C**	**D**
(a,a,a)	1	0	0	0
(b,a,a)	0	1	0	0
(a,b,a)				
(a,a,b)				
(a,b,b)				

	A	B	C	D
(b,a,b)				
(b,b,a)				
(b,b,b)				

The inference attacks

The adversary will carry out three different inference attacks.

(a) Bayes' attack: his intent is to infer exactly all the secret values. The gain function for this attack is **bv**.

(b) Specific patient attack: the adversary is only interested in patient id 1 (perhaps that patient is famous). The gain function for this attack is:

spa(a, (x,y,z)) = 1 if (x = a), otherwise 0

spa(b, (x,y,z)) = 1 if (x = b), otherwise 0

(c) Non-specific attack: the adversary wants to guess the illness of any of the three patients.

npa(a, (x,y,z)) = 1 if (x = a) OR (y = a) OR (z = a), otherwise 0

npa(b, (x,y,z)) = 1 if (x = b) OR (y = b) OR (z = b), otherwise 0

For each inference attack there are two scenarios:

(i) the adversary initially knows nothing, and so will use the uniform prior

(ii) the adversary knows something about the statistical profile of illnesses and so will use the prior determined by that. You will create your own prior based on your student id as follows. Observe that there are 8 possible secret values. To define the prior over those 8 secrets values take your student id which has 8 digits, i.e.
s_1s_2s_3s_4s_5s_6s_7s_8
Next, add up the 8 individual digits to obtain a total T. Then use the following prior distribution:

	(a,a,a)	(b,a,a)	(a,b,a)	(a,a,b)	(a,b, b)	(b,a,b)	(b,b,a)	(b,b, b)
Prior	s_1/T	s_2/T	s_3/T	s_4/T	s_5/T	s_6/T	s_7/T	s_8/T

For example. if your student id is 12345678 then your T= 36 and your prior is therefore

	(a,a,a)	(b,a,a)	(a,b,a)	(a,a,b)	(a,b, b)	(b,a,b)	(b,b,a)	(b,b, b)
Prior	1/36	1/18	1/12	1/9	5/36	1/6	7/36	2/9

Do not worry if your prior has some probabilities that are zero. Those correspond to the zero digits in your student id.

Your Task

1. Complete the channel above to summarise the information leaks between the inputs and outputs above. You may complete this in the Jupiter notebook of you wish.

2. For **each** attack (a),(b) and (c) above compute, for **each** scenario defined by priors (i) and (ii) and the gain function g:

    (I) The prior vulnerability, V_g(pi);

    (II) The average posterior vulnerability for the given prior and channel, V_g(pi, C)

Input all your numerical answers into the iLearn submission. Your answers should be in decimal format and only include 3 decimal places. For example if your vulnerability is 1/9 input 0.111. If your vulnerability is 5/9 input 0.556.

3. Indicate which attack on which scenario performs the best for **average posterior vulnerability.** For example if attack (a) for prior (i) is maximum amongst all scenarios then choose the (a) and (i) combination in iLearn.

What you must hand in

In the submission page on iLearn for this assignment you must input the answers to your analysis above by the due date. Some of these answers will be automatically marked and so it is important to abide by the input instructions above.

Late penalty

It is important that you hand your work on in time. Please note that a late penalty of 10% reduction per day late will be applied.