THE UNIVERSITY OF AUCKLAND

SEMESTER TWO, 2015
Campus: City

COMPUTER SCIENCE

Software Security

(Time allowed: 0.25 hours)

NOTE: Your answers on this sample examination will not affect your final grade.

Please do NOT put your name on your answer sheet.

On the last day of lectures, we will discuss how we would mark some of the answers written by you or your classmates.

We advise you to spend about 5 minutes on a 5-mark question.

Very roughly: you should write two to four sentences when answering a 5-mark question.

We advise you to reserve at least 10 minutes at the end of the examination period, so that you can review your answers for accuracy prior to submitting your answer sheet.

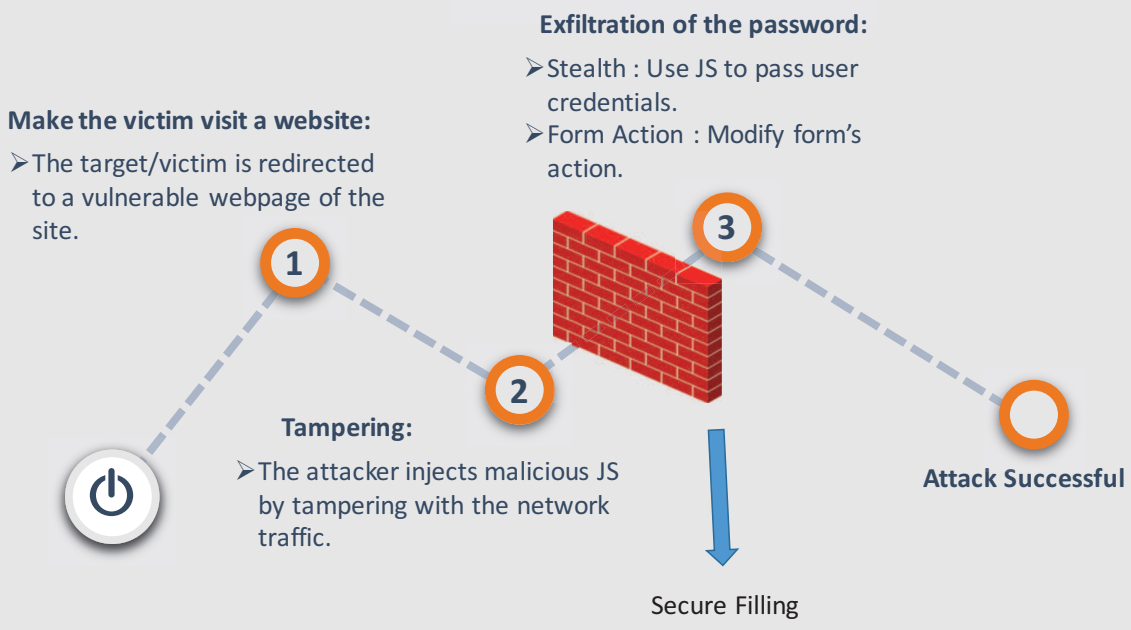
This is a **closed-book** examination.

Calculators are **not permitted** in this examination.

1. (20 marks, in total) Recall a threat scenario in [Humbert 2015]:

> ... if an adversary has access to phenotypic traits (e.g., visible traits) of an identified individual, he can use known correlations between phenotypic traits and genomic data to identify the genotype of this individual in a genomic database and to infer other sensitive information (such as predispositions to severe diseases) by using the de-anonymized genomic data... The adversary also has access to anonymized genotypes through a collaborative genome-sharing platform (such as the Personal Genome Project) and [wants] to de-anonymize them by relying upon phenotypic information gathered on online social networks (OSNs). The matching between OSNs and genomic profiles is (even) easier if, for example, the ZIP code is available with the genomic profiles, thus enabling the OSN profiles to be filtered before the matching attack.

(a) (10 marks) The following figure appeared in a student's slideshow this semester. Draw a figure, in a similar style, which depicts threat scenario in the passage quoted above. To receive full marks, your figure must show a defensive "wall", and it must show at least two steps in an attack.



- Marking rubric. Six marks for the depiction of the attack; four for the defense. We expect to see three steps (2 marks each) in the attack, as described in the first sentence of the quoted passage: 1) an attacker gains access to some phenotypic traits of an identified individual; 2) this attacker uses known correlations between phenotypic traits and genomic data to identify this individual in a genomic database; 3) this attacker infers other sensitive information about this individual. Students may incorporate some additional information from the second and third sentences, e.g. indicating that an attacker might gain phenotypic information from an OSN, and that the anonymized genomic database may have ZIP codes or other information which could be matched with data obtained from the OSN. A highly observant

student may notice that steps 2 and 3 are independent – either an identification or an inference of "other sensitive information" may follow step 1.

- Sample answer #1. 1) Access phenotypic traits on OSNs; 2) Access anonymized genotypes through a collaborative genome-sharing platform; Wall); 3) Deanonymize genotypes based on phenotypic information; 4) Filter OSN profiles based on ZIP code and match OSNs and genomic profiles; 5) Match successful.

  [We award 8/10 marks. The wall is poorly specified (3/4 marks): we didn't ask for a full description, but the sample diagram had a two-word phrase describing its wall. Steps 3 and 4 are out of order – the quoted passage says the ZIP-code filtering occurs before the (de-anonymizing) "matching attack". However steps 1-3 are very nicely constructed, so we award 5/6 marks for the attack description.]

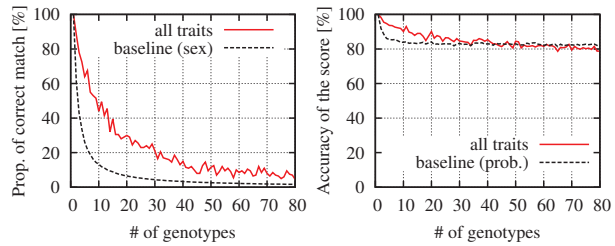
  [Note: after reviewing the article again, while marking, we now believe that the authors had made a mistake when writing the quoted passage. Their intent was (we believe) to summarise the discussion of filtering in the penultimate paragraph of Section 4. In that paragraph, the authors assert that the filtering of an anonymized genomic database on a person's ZIP code would "result [in a genomic database with] typically, a few hundreds of individuals at most".]
- Sample answer #2. The student's diagram has six numbered states, with labels on states 1 through 6. 1) Get anonymized data; 2) Process genetic data into sets of phenotypes; Wall) Block mass retrieval of image data; 3) Get data from OSNs and classify by phenotype; 4) Compare the two sets of phenotype data (potentially filter by ZIP first); 5) Guess who is who; 6) Do something evil if you want?

  [We award 6/10 marks. This figure depicts a plausible attack, with a single line of defense for that attack. It is very inaccurate but thought-provoking – so somewhat difficult to mark. The student's diagram does not depict an attack on an identified individual, as described in the direct quotation. Their "Guess who is who" step is part of a perfect-matching attack, as described in (Humbert 2015) but is not relevant in the quoted threat scenario. The student's second step, a conversion of a genomic database into "sets of phenotypes", is irrelevant to both the identification attack and to the perfect-matching attack. The student's wall is blocking "mass retrieval of image data" – this is also far afield from what is discussed in (Humboldt 2015). Even so, it is a valid defense of an anonymized genomic database so it can be depicted as a wall in this diagram. Indeed, technical and legal systems which are intended to protect privacy often include assumptions, laws, or technical restrictions on any unauthorised person's construction of a large database of personally-identifying information. However, no mass retrieval of phenotypic information is required to perform the identification attack described in the quoted passage; the student's wall would mitigate only the threat of a perfect matching attack. We resort to holistic grading in this instance, awarding a mark in the B- range, to denote a general understanding of security as taught in this course, with no demonstration of technical depth or accuracy in this instance.]
- Sample answer #3. The student's diagram has five labeled boxes, with two

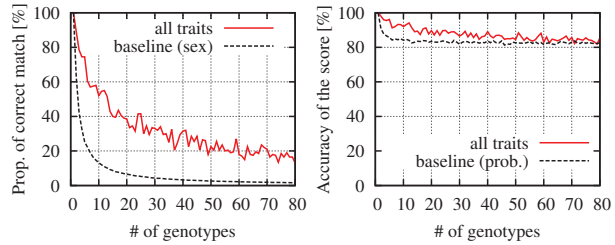
explanatory legends. The legends define the acronyms in the labelled boxes: IPTDB is an "identified phenotype database", AGTDB is an "anonymous genotype database", PT·Id is a join (or "binding") of a phenotype database with a genotype database, GT·Id is a binding of a genotype database with an identification table (from an OSN), and the fifth box is labelled "Sensitive information". The student's diagram is in an entirely different style to the sample diagram, but eventually we determined that two wavy lines (labelled "Rate limit") across arcs into IPTDB and AGTDB could be considered defensive "walls". A third wavy line, labelled "Possible confounding entries in PT/GT entries in DB's" could be considered a defensive "wall" on the arc from the region of IPTDB and AGTDB into PT·Id. The arc from PT·Id to GT·Id is labelled "biology knowledge". The arc from GT·Id to "Sensitive information" is unlabelled.

[We award 3/10 marks. This student is very creative, and is showing talent at devising a novel type of diagram to depict an attack; but the requirement was to use the same style as a diagram we had discussed in lecture. Furthermore, the depicted attack flow is technically incorrect as a summarisation of the quoted passage. The PT·Id information is not the result of a de-anonymization; de-anonymization is required to create GT·Id. Nothing in (Humbert 2015) suggests the defenses suggested by this student. Because they are valid defenses for either or both of the attacks described in that article, we are awarding 3/4 marks for the depiction of the defense; but 0/6 for the depiction of the attack.]

(b) (5 marks) What does the following figure tell you about the threat scenario of this question?



(a) unsupervised



(b) supervised

Figure 1. Performance of the identification attack for a single individual in a genomic database in the (a) unsupervised and (b) supervised cases, with respect to the proportion of correct matches (left) and the accuracy of the susceptibility score for Alzheimer's disease (right). The size of the database varies from 1 to 80.

- Marking rubric: We award 1 mark for an accurate use of each of the following concepts: identification, susceptibility to Alzheimer's disease, supervised (vs. unsupervised) learning, number of genotypes in the database, accuracy. We mark-up the student's answer with underscores to indicate our tickmarks, and brackets to indicate our questions, comments, and spelling-corrections.
- Sample answer #1: Figure 1 suggests that an adversary is attempting to attack a single individual by determining which genotype they are connected to in a database, therefore determining their susceptibility to Alzheimer's disease. Figure 1 also suggests that increased trait information increases the strength [accuracy?] of the attack.

  [3/5 marks. The student's last sentence is not obviously supported by the figure – they may be recalling an interpretation of Figure 7 in this article, rather than interpreting the reproduced figure. However sex *is* a phenotypic trait, and in each of the four sub-figures the matchings on this one trait are less accurate than the matchings on all traits.
- Sample answer #2. Although the proportion of correct matches drops noticeably as the number of genotypes increases, the accuracy of the susceptibility score

remains high. This means that although the attacker may be unable to tell exactly who you are, they can infer other things about you by looking at other people with 'similar' DNA.

[4/5 marks]

(c) (5 marks) Do you agree with the following assertion in [Humbert 2015]? "Our results demonstrate the serious de-anonymization threat currently posed to individuals sharing their SNPs in genomic databases." In either case, justify your answer.

- Marking rubric: We look for a coherent discussion of the severity of the threat, with respect to the "value, locks, punishment" framework of Lampson's "real-world security" model. The quoted passage doesn't mention punishment (either legally or socially) as a security control, so we look for "value" or "motivation" or "goal" or "cost and benefit" or "risk and reward" (2 marks), and we also look for discussions of the effectiveness or vulnerabilities of the "lock" (the anonymization) (3 marks).
- Sample answer #2. No, I do not agree. By inspecting the graphs we can see that as the number of genotypes increases, the probability of guessing correctly drops off very quickly. If you used any sizeable database you would get almost no correct results. This is particularly true if you use OSNs, which have vast numbers of people.

  [2/5 marks. The student refers to the reproduced graphs, but hasn't noticed the high probability of a correct inference of a susceptibility to Alzheimer's disease: locks 2/3. No mention of value: 0/2.]
- Sample answer #2. The phenotypes they collected are limited, and the genotype dataset was small. Thus, their results were not conclusive, but should be considered good reason to continue research into the problem. Individual's online presence will continue to increase and likely more of their phenotypic data will be stored online. With a more precise phenotipic [sic] data, and a much vaster genotypic dataset, there could be a real issue.

  [Locks 1/3 + Value 0/2 = 1/5 marks. The student hasn't answered the question. Although there's uncertainty about the effectiveness of the lock, the student might still agree that there's a serious threat; then again, the student might be saying that they believe the authors have not adequately demonstrated a serious threat because their results are "not conclusive". The discussion of future threats is irrelevant; but it is neither wildly incorrect nor wildly irrelevant, so we are ignoring it.]
- Sample answer #3. I disagree because the de-anonymization performs poorly when the sample is larger than a small number like 10. However it can be successful in identifying diseases in a cluster of genomic data if a filter can be applied to create such distinct clusters.

  [5/5 marks, showing excellent understanding of the lock, and mentioning goals.]