CPSC 457/557 // Exam 2 // December 1, 2011

Answer all of the questions. Please remember to write your name, the course number, and today's date on all blue books that you submit.

This is a closed-book exam; please do not refer to any books or notes, and please do not talk to any of the other students.

Question 1: Privacy and Online Advertising (30 points)

(a) (5 points) Recall that Goldfarb and Tucker showed that the EU Privacy and Electronic Communications Directive decreases the overall effectiveness of online advertising but that the extent of the decrease varies across websites. Give an example of a website that, according to the Goldfarb-Tucker findings, would be severely affected and an example of one that would be only modestly affected, and briefly explain why the first would be more severely affected than the second.

(b) (5 points) What is the main objection of privacy advocates to the approach embodied in the US Commercial Privacy Bill of Rights proposed by Senators Kerry and McCain?

(c) (5 points) Recall that Gmail serves ads based on "concepts" derived from the body of an email message, from files attached to the email message, and from web pages to which the message contains links. Do Google's AdSense and AdWords systems use these Gmail concepts to serve ads? If not, why not?

(d) (5 points) On its logout page, Gmail may also serve ads based on information about a user's behavior on other (non-Google) websites. Give one explanation of how, technically, Gmail can obtain that information.

(e) (5 points) Briefly explain why the NY State Bar Association believes that the use of Gmail does not violate a lawyer's duty of client confidentiality.

(f) (5 points) Recall that Gedik and Liu's "personalized *k*-anonymity" approach to location-based advertising allows a user to specify, on a message-by-message basis, the number *k-1* of other users whose locations are supposed to be indistinguishable from his. Give one advantage and one disadvantage of this approach.

Question 2: Differential Privacy (20 points)

(a) (5 points) What is differential privacy? You may give the technical definition that we covered in class or just explain in words (but precisely) the property that a differentially private data-analysis technique must have.

(b) (5 points) What is Dalenius's desideratum for privacy-preserving data analysis? Is it attainable using differential privacy? (You may just answer "yes" or "no" and need not justify your answer.)

(c) (10 points) What is the subsample-and-release paradigm for statistical data analysis, and why does it not provide differential privacy?

Question 3: Digital Currency (30 points)

(a) (8 points) For 2 points each, state four basic properties that a digital-cash scheme should have.

(b) (5 points) What are "blind signatures," and what role do they play in the signature-based digital-cash scheme that we covered in class?

(c) (9 points) In order to deposit a digital \$20 bill using the Deposit protocol that we covered in class, a vendor must present a triple (M, s, RIS), where M is the bill, s is the bank's signature on the bill, and RIS is a "random identity string." RIS is a K-tuple $(b_1, b_2, \ldots, b_K)$ of bits with the following properties:

  i. It is chosen uniformly at random by the vendor.

  ii. The bill M contains K pairs (y_1, y_1^1), (y_2, y_2^1), …, (y_K, y_K^1) for which the user, during the Withdrawal protocol, chose K pairs (x_1, x_1^1), (x_2, x_2^1), …, (x_K, x_K^1) such that, for all j, $H(x_j) = y_j$, $H(x_j^1) = y_j^1$, x_j was chosen uniformly at random, and x_j XOR x_j^1 = Username. Here, H is a one-way hash function.

  iii. For all j, the user must, during the Payment protocol, reveal x_j to the vendor if $b_j = 0$ or x_j^1 to the vendor if $b_j = 1$.

What should the bank do if he receives two triples (M, s, RIS), (M, s, RIS^1), where RIS and RIS^1 are different, from two different vendors? What should the bank do if he receives the same triple (M, s, RIS) twice from the same vendor? Why is it overwhelmingly unlikely that he will receive the same triple twice from two different vendors?

(d) (8 points) Recall that the BitCoin digital-currency scheme, like the basic HashCash scheme that we covered in class, uses one-way hash functions and proofs of work in an essential way. For 4 points each, explain two substantial differences between the BitCoin scheme and the basic HashCash scheme.

Question 4: Online identity and anonymity (20 points)

(a) (5 points) Give one argument in favor of the proposition that anonymous speech should enjoy the same ("McIntyre") protection on the web as it does in print and one argument against it.

(b) (5 points) Technically, how might a website implement a compromise between the two positions in part (a)?

(c) (5 points) What is the difference between an "identity credential" and an "attribute credential" in the National Strategy for Trusted Identities in Cyberspace (NSTIC)?

(d) (5 points) Are Mayer and Narayanan more optimistic about the NSTIC's being able to curb security breaches caused by malware or those caused by phishing? (You need not explain the basis for their optimism or pessimism.)