UNIVERSITY OF EDINBURGH

COLLEGE OF SCIENCE AND ENGINEERING

SCHOOL OF INFORMATICS

# INFR10067 COMPUTER SECURITY

**Friday 7th May 2021**

**13:00 to 15:00**

## INSTRUCTIONS TO CANDIDATES

**Answer any TWO of the three questions. If more than two questions are answered, only QUESTION 1 and QUESTION 2 will be marked.**

**All questions carry equal weight.**

**This is an OPEN BOOK examination.**

Year 3 Courses

Convener: D.Armstrong
External Examiners: J.Bowles, S.Rogers, H.Vandierendonck

THIS EXAMINATION WILL BE MARKED ANONYMOUSLY

1. **Network Security and Anonymous Communications**

(a) In lecture 2, we discussed different types of attacks on traffic as it traverses the network. List and describe the 5 different types of network attacks. Be clear on what the adversary does to achieve these attacks and the effect on the data (in terms of confidentiality, integrity, and availability). [*5 marks*]

(b) The Tor network protects both users and websites. Describe what types of protection Tor provides to both of these groups. [*5 marks*]

(c) Ruba sets up an onion service (hosted at a local ISP) to run a dissident news website. She posts news articles there (using Tor to connect to the webserver hosting the site) and people visit it to read these articles.

The government finds the URL to Ruba's onion service and wants to shut it down. They don't know Ruba's identity, but they have a strong suspicion that the ISP hosting the onion service is inside the country. What could they do—provide details of their steps—in order to confirm that the onion service is inside the country? Assume that the government can observe and record all network traffic in the country but do not collude with the ISPs or hosting companies—i.e. they can not break into the server itself. (Hint: Recall that Tor is only secure against a local adversary and weak against certain attacks.) [*6 marks*]

(d) We looked at the Three Dining Cryptographers (3DC) protocol which can be used to tell whether one of the diners has paid for the meal. For this question, we add one more party for 4DC (A, B, C, and D) in this protocol. The following coin flips are observed by each pair of parties (Heads=1, Tails=0):

AB:1, AC:1, AD:0, BC:1, BD:0, CD:1.

The following announcements are made by each party: A:0, B:0, C:0, D:1.

The completion of the protocol yields a result of 1. Using the information given in this question, is it possible to work out who paid for the meal; was it A, B, C, D, or the NSA? Show your working out. [*2 marks*]

(e) Sohail sets up a new anonymous email service called ZeroBits. This is how it works:

- A user wants to send an email to someone they know. They produce the following message:
  - To: recipient's email address
  - From: user's email address
  - Message: Hi there!
- This message is encrypted using the ZeroBit public key and is sent to the ZeroBit email server.

*QUESTION CONTINUES ON NEXT PAGE*

- The email server decrypts the email and replaces the From: address with a random address like so:
  - From: random_number@zerobit.com
- The ZeroBit email server records the original email address and the random email address in a mapping table like so:
  - user's email address: random_number@zerobit.com
- The server then forwards the email to the intended recipient, encrypted with the recipients public key. It does not copy or log the contents of the messages.
- To reply the recipient sends the encrypted reply email to the random_number@zerobit.com address and the server looks up the original email address and forwards the email to the original sender, encrypted with the sender's public key. Again the contents of the message are not recorded.

i. Who is the sender anonymous to? [*1 mark*]

ii. What is the most vulnerable part of this email service in terms of protecting sender anonymity and what would be lost if it is removed. [*2 marks*]

iii. Sohail decides to add email batching to the service to provide protection against global adversaries who can observe all network activity between the service and the senders and receivers. He decides that the server will wait until 100 emails have been received before sending them all out together. Describe what sort of attack the adversary can still mount to reveal which senders and recipients are communicating together. Assume that the adversary is able to add, drop, slow down, or change any message on the network. Provide a description of the various stage of the attack and how the adversary makes sure of the results. [*4 marks*]

2. **RSA encryption**

In class we discussed the RSA Encryption Scheme. I briefly sketch a version of this cryptoscheme here:

- **Key generation -** Bob generates two large random primes $p$ and $q$, and computes $n = pq$. He then picks a number $v$ that is relatively prime to $\phi(n)$, and computes $d$ such that $d \cdot e = 1 \mod \phi(n)$. From this point on, he can "throw away" the values $p$, $q$, and $\phi(n)$. They are no longer needed. Bob's public encryption key is $pk_B := (e, n)$. His private decryption key is $sk_B := d$. He keeps $sk_B$ secret, but publishes $pk_B$ to allow others to encrypt messages to him that he can decrypt using his decryption key.
- **Encryption algorithm -** Everyone can encrypt a message $m$ of length $\ell$ under the public encryption key $pk_B$ of Bob by prepending a random string $r$ as follows:

$$\mathsf{Enc}(pk_B, m) := \begin{cases} \textbf{sample } r \in [2^\ell, n) \\ \textbf{return } c = (r \| m)^e \mod n \end{cases}$$

  where **sample** $r \in [2^\ell, n)$ draws a random number from the specified range.
- **Decryption algorithm -** Given Bob's private decryption key $sk_B$, Bob can decrypt a ciphertext $c$ created as above by computing $r \| m = c^d \mod n$ and returning $m$. That is,

$$\mathsf{Dec}(sk_B, c) := \begin{cases} \textbf{parse } c^d \mod n \textbf{ as } r \| m \\ \textbf{return } m \end{cases}$$

The goal of this exercise is to explore whether this randomly sampled prefix $r$ to the length $\ell$ message is necessary and/or sufficient for secure RSA encryption. We will consider the simplified RSA encryption scheme which does not prepend $r$ to the message $m$ and just uses $m$ directly. So now to encrypt a message to Bob one just computes $c := m^e \mod n$ and sends $c$ to Bob.

(a) What is the decryption algorithm for the simplified RSA encryption scheme? [*1 mark*]

(b) Eve knows that Alice and Bob are both using the simplified RSA signature scheme, and wants to trick Alice. Assume that Eve has captured over the network a ciphertext $c$ from Alice to Bob. Eve is naturally jealous and suspects that Alice might be in love with Bob. What could she do to learn Alice's message? Justify your answer. [*3 marks*]

*QUESTION CONTINUES ON NEXT PAGE*

(c) Alice and Bob became suspicious that Eve might be evesdropping but they are still using the simplified RSA scheme. Now suppose that Alice and Bob pre-agreed on a secret passphrase $p$ that Alice uses to hide her message $m$. The new message becomes $m' = p \cdot m$, and Alice encrypts it to $m'^e \mod n$. Explain how Eve who knows $pk_B$ and sees Alice's ciphertexts (with the help of her friend Mallory) is able to learn the passphrase. Mallory's job is to craft a special ciphertext to Bob, while Eve does some social engineering to get Bob to reveal to her the decryption of a small number of ciphertexts that Bob decrypts, but not the passphrase itself. That is, when Bob receives a ciphertext $\tilde{c}$, he inverts the RSA trapdoor permutation to learn some value $\tilde{m}'$ and then leaks $\tilde{m}'/p$ to Eve. You can assume that $\gcd(m, n) = 1$ and $\gcd(p, n) = 1$, and that $m$ and $p$ are small enough so certain computations do not wrap around modulo $n$ during the attack. Justify your answer. [Hint: Bob will be quite confused about an obscure message he presumably received from Alice. Eve can exploit this confusion] [*8 marks*]

(d) For RSA encryption, one sometimes sets $e$ to a small prime value such as 3. For this question, let Bob's public key be $pk_B = 3$.

Bob is holding an auction. The protocol is simple. Bidders just submit signed encrypted bids to Bob. You can ignore the details of the signature here, simply assume that it is secure, but encryption uses the simplified RSA encryption scheme and Bob's public key $pk_B = 3$. The message $m$ is their bid (in pounds). Bob runs a second-price auction, the highest bidder wins, but only has to pay the price of the second highest bid. Mallory wants to mess with Alice's bidding. So, when Alice forms her bid $m$ and sends to Bob her signed encrypted bid $(c, \sigma)$ where $c = m^e \mod n$, Mallory intercepts it. Mallory would like to tamper with $c$ to form a new ciphertext that corresponds to a bid for 8 times Alices's original bid. This will allow Mallory to win the auction and pay exactly as much as Alice was willing to pay. More precisely, she'd like to find a value $c'$ such that $c'$ is a valid encryption of $8m$. She then computes her own signature $\sigma'$ and forward the result $(c', \sigma')$ onto Bob. Help Mallory out: explain to Mallory how she can compute such a $c'$. You will assume that $m$ is small enough so that $8m < n$, so that $8m$ does not wrap around modu lo $n$. Justify your answer. [*6 marks*]

(e) Are your attacks from questions 2b, 2c and 2d possible against the version of RSA encryption sketched in the beginning of the question (the one that includes the random prefix $r$)? Explain your answer. [*7 marks*]

3. **Web security - Cookies**

(a) `AcmeBank.com` is an online bank that wants to expand its activities and start offering a web hosting service. The service will allow anyone to choose their `sitename` and upload any script or HTML code. At AcmeBank they are considering two options for hosting the website creation service: either on `AcmeHosting.com/[sitename]`, or on `AcmeBank.com/sites/[sitename]`. Which of these two options is better from a security perspective? Explain you answer. [*3 marks*]

(b) Google Analytics (GA) is a web analytics service offered by Google that tracks and reports website traffic. Websites can (anonymously) register with Google to instrument their site for analytics, and gather information about who visits, and what they do when they visit. Google Analytics is implemented with "page tags", in this case, called the Google Analytics Tracking Code, which is a snippet of JavaScript code that the website owner adds to every page of the website. The tracking code runs in the client browser when the client browses the page and collects visitor data and sends it to a Google data collection server as part of a request for a web beacon. This Javascript code is not contained in any (i)frame. In addition to transmitting information to a Google server, the tracking code sets a `__utma` cookie in Javascript on each visitor's computer. This cookie stores anonymous information called the `ClientId`.

i. Assume `https://acme.com` uses Google Analytics. What is the domain of the `__utma` cookie? Explain your answer. [*3 marks*]

ii. Assume an attacker has managed to compromise Google's servers, and has managed to replace the Google Analytics script with their own malicious script. The attacker is now able to steal Acme's other cookies. Explain how they would access and send to the attacker the stolen Acme cookies, and how to prevent the attacker accessing the other Acme cookies. [*3 marks*]

iii. Does blocking third-party cookies prevent Google from tracking a single user across different domains/websites using cookies? Justify your answer. [*3 marks*]

iv. The recently proposed `Do Not Track` header and legislation aim to give users a standardised way to opt out of web tracking. A browser setting (already implemented natively in Firefox, IE, and Safari) appends a `DNT=1` header to outgoing requests, informing the receiving website that the user wishes to opt out of tracking. Does setting the `DNT` flag prevent Acme from using Google Analytics for tracking their visitors? Justify your answer. [*3 marks*]

*QUESTION CONTINUES ON NEXT PAGE*

(c) Many sites today can dynamically update a page's content via asynchronous Javascript requests that return JSON data. Sometimes, JSON can contain sensitive data. Imagine once a user has successfully logged in to her `AcmeBank.com`, a session token is stored in an `HTTP` cookie in the user's browser, and they are being taken to the page `https://AcmeBank.com/account.html` showing their balance. To fetch the user's account information, `account.html` includes a script as follows:

```
<script src="//AcmeBank.com/userdata.js">
```

This script is dynamically generated and its content is of the form:

```
display({"user": "Alice",
         "AcctNumber": 123456,
         "Balance": 100})
```

With the `display()` function being defined in `account.html`. This script is not contained in an (i)frame and is executed in the context of the page that includes it.

i. Assume that Alice has logged into her account at `AcmeBank.com`, and then visits the `https://evil.com/` website maintained by Eve. Explain how Eve can steal all of Alice accounts' information. [*7 marks*]

ii. How would you defend against this attack? Explain your answer. [*3 marks*]