MATH/CMSC 456 (Washington) Final Exam May 14, 2015
The exam is worth 140 points.

1. (15 points = 5+10) Alice uses an improvement of the Vigenère cipher: She writes all the letters in the plaintext as numbers mod 26 in the standard way (with $a = 0$ and $z = 25$) and she chooses an integer b and integers $a_1, a_2, \ldots, a_5$. She applies the affine functions $bx + a_1, bx + a_2, \ldots, bx + a_5 \pmod{26}$ as in Vigenère (so $bx + a_1$ is used on the 1st, 6th, 11th, ... letters, $bx + a_2$ is used on the 2nd, 7th, 12th, ... letters, etc.).
(a) What condition does b need to satisfy for Bob (who knows the key) to be able to decipher the message?
(b) Describe a chosen plaintext attack that will yield the key. You know the encryption method and you know the key length. You must *explicitly* say what plaintexts you use.

2. (20 points = 5+5+5+5) At the end of the semester, the professor randomly chooses and sends one of two possible messages:
$m_0 =$*YOUPASSED* and $m_1 =$*YOUFAILED*.
To add to the excitement, the professor encrypts the message using one of the following methods:
(a) Shift cipher
(b) Vigenère cipher with key length 3
(c) RSA with a public 300-digit modulus n and encryption exponent $e = 65537$ (the message is not padded with extra bits)
(d) One-time pad
You receive the ciphertext and want to decide (in less than one minute, but you have a computer) whether the professor sent m_0 or m_1. For each method (a), (b), (c), and (d), explain how to decide which message was sent or explain why this is impossible. (*Notes:* You may assume that you know which method is being used. For the Vigenère, do not use frequency analysis; the message is too short.)

3. (25 points = 5+5+5+10) Consider the ElGamal Signature Scheme: The document Alice wants to sign is m.

1. She chooses a prime p, a primitive root g, and a secret integer a.
2. She computes $h \equiv g^a \pmod{p}$. The numbers p, g, h are public, and a is secret.
3. She chooses a random integer k with $\gcd(k, p-1) = 1$ and computes $r \equiv g^k \pmod{p}$.
4. She computes $s \equiv k^{-1}(m - ar) \pmod{p-1}$.
5. The signed document is (m, r, s).
6. Bob verifies the signature by checking that $g^m \equiv h^r r^s \pmod{p}$.

The values m, r, s, p, g, h are public, and the same p, g, h are used for each signature.
(a) Why is the condition $\gcd(k, p-1) = 1$ needed?
(b) Suppose Alice uses the same value of k for two different messages m_1 and m_2.

How does Eve recognize that Alice did this?
(c) In the same situation as part (b) (that is, $k_1 = k_2 = k$), how can Eve use (m_1, r_1, s_1) and (m_2, r_2, s_2) to compute k and then a? To avoid technical problems, you may assume that $\gcd(s_1 - s_2, p-1) = 1$ and $\gcd(r_1, p-1) = \gcd(r_2, p-1) = 1$.
(d) Suppose Alice's random number generator can produce only 10^{12} different values of k (but no one knows this). After Alice signs 10^7 documents (she is a bureaucrat; that's her job), it is likely that Eve can break the system; that is, Eve can find a with high probability. Explain why.

4. (20 points = 5+5+10) **(a)** You are a contestant on a game show called *Prime Time* and you are given a 100-digit odd integer n. You have one minute to guess whether n is composite or prime. You want to make an intelligent guess, so you are given permission to use your calculator, which can do addition, subtraction, multiplication, division, exponentiation, and modular exponentiation, all of these with very large numbers. It does not have a *factor* or an *isprime* command. Since you are addicted to Twitter, you can type at most 140 characters into your calculator. What do you do in order to have a very good chance of getting the right answer? You might start with "$n = \cdots$", which uses 102 characters but allows you to avoid writing out n again.
(b) Eve thinks that she has a great strategy for breaking RSA that uses a modulus n that is the product of two 100-digit primes. She decides to make a list of all 100-digit primes and then divide each of them into n until she factors n. Why won't this strategy work? (If this is the method you gave for part (a), try again.)
(c) Suppose you know that $7961^2 \equiv 7^2 \pmod{8051}$. Use this information to factor 8051 (you *must* use this information; answers without 7 and 7961 will not receive credit).

5. (10 points) Let P and Q be points on an elliptic curve E. Peggy claims that she knows an integer k such that $kP = Q$ and she wants to convince Victor that she knows k without giving Victor any information about k. They perform a zero-knowledge protocol. The first step is the following:
1. Peggy chooses random integers r_1 and r_2 and lets $r_3 = k - r_1 - r_2$. She computes

$$X_1 = r_1 P, \quad X_2 = r_2 P, \quad X_3 = r_3 P$$

and sends them to Victor.
Give the remaining steps. Victor wants to be at least 99% sure that Peggy knows k. (There are two possible solutions; either one is ok, but one is more efficient than the other.) (*Technical note:* You may regard r_1, r_2, and r_3 as numbers mod n, where $nP = \infty$. Without congruences, Victor obtains some information about the size of k. *Non-technical note:* The "Technical note" may be ignored when solving the problem.)

6. (20 points = 10+10) **(a)** Alice and Bob are trying to implement 3-DES. They each choose keys K_1 and K_2. Alice encrypts using the formula

$$c = E_{K_1}(E_{K_2}(E_{K_1}(m))),$$

while Bob uses

$$c = E_{K_1}(E_{K_1}(E_{K_2}(m))).$$

These two methods not equally resistant to attack. Say which one is weaker, and explicitly give the steps of a method that can be used to attack the weaker system. You may assume that the attacker knows several plaintext-ciphertext pairs encrypted by the system.
(b) Let p be a 300-digit prime. Define a hash function

$$H(x) = 2^x \pmod{p}.$$

Although H can be computed quickly, it is not fast enough to be a good hash function. Give one more property of cryptographic hash functions that H does not satisfy and give one property that H satisfies. You must justify your answers.

7. (20 points = 10+5+5) (*Note:* Parts (a), (b), (c) are not related to each other.) Recall the RSA signature scheme. Alice wants to sign m. She chooses an RSA modulus $n = pq$ and encryption and decryption exponents e and d. She computes $s \equiv m^d \pmod{n}$. The signed message is (m, s). The signature is valid if $m \equiv s^e \pmod{n}$.
(a) You are given the following fact: If $n = pq$ is the product of two distinct odd primes and $\gcd(m, n) = 1$, then $m^{(p-1)(q-1)/2} \equiv 1 \pmod{n}$. Use this fact to show that if $\gcd(m, n) = 1$ and if $ed \equiv 1 \pmod{(p-1)(q-1)/2}$, and if Alice signs a message according to the above scheme, then Alice's signature is valid. You *must* explicitly say how you use the given fact.
(b) Bob wants Alice to sign m, but does not want Alice to see m (since m contains the plans for a secret invention). Bob chooses a random integer k with $\gcd(k, n) = 1$ and lets $m_1 \equiv k^e m \pmod{n}$. He asks Alice to sign m_1. Alice produces the valid signed message (m_1, s_1). What does Bob do to s_1 obtain a valid signature s for m?
(c) Alice decides to use a hash function when signing messages. How are the signing and verification congruences modified to accomplish this?

8. (10 points = 5+5) **(a)** A bank in Tokyo wants to send a terabyte of data to a bank in New York. The two banks have never communicated before. Describe how the data can be sent securely, even if Eve intercepts all of their communications. Explicitly name the algorithms that will be used.
(b) There could be transmission errors when the massive amounts of data in part (a) are transmitted. Therefore, the bank in Tokyo uses a cryptographic hash function H and computes the hash of the data. This hash value is sent to the bank in New York. The bank in New York computes the hash of the data received. If this matches the hash value sent from Tokyo, the New York bank decides that there was no transmission error. What property of cryptographic hash functions allows the bank to decide this?