Assignment 1 – Due Monday, February 3

Chapters 1–3

Note: Remember that one of the main goals of this class is to develop your skills in reasoning about security, so show your work and explain your reasoning on all problems – for homework problems, your thought process is as important as the final answer!

1. In this question, you should read about the recent security breach at Target stores, and write a summary taking into account the security goals and terminology from Chapter 1 of your textbook. For information on this security breach, you should read at least the following three articles (note: the links are live on the class web site, so you can just click) — there's certainly plenty more out there if you want to read more, but these cover the important technical points:

   - A First Look at the Target Intrusion, Malware — Krebs on Security
     http://krebsonsecurity.com/2014/01/a-first-look-at-the-target-intrusion-malware/
   - A Closer Look at the Target Malware, Part II — Krebs on Security
     http://krebsonsecurity.com/2014/01/a-closer-look-at-the-target-malware-part-ii/
   - Can hackers decrypt Target's PIN data? — by Matthew Green
     http://blog.cryptographyengineering.com/2013/12/can-hackers-decrypt-targets-pin-data.html
     Warning: This one is a little more dense to read, and uses some cryptographic terms we have not covered yet. Still, try to read it and see what you can get out of it. If the terminology is a big enough barrier where you can't get the "big picture," then ask me about things that are unclear, and I'll fill in details as needed.

   Your summary should be brief (less than a page), but be careful to use proper terminology. In particular, what *assets* were compromised in this *attack*? (Hint: Think broadly: stolen information is certainly an asset, but so is company infrastructure.) For these assets, what are *confidentiality*, *integrity*, and *availability* goals, and were they compromised? Where was a *vulnerability* used as an *attack vector*? (This is complex: there were certainly multiple attack vectors, some of which are not publicly known – just summarize the ones described in the stories referenced above.)

2. Textbook, page 56, Problem 2.1.

3. Following up on textbook Problem 2.1, how many valid keys are there in the system described in the textbook problem? If we slightly increased the alphabet size to 29 (so operations are performed mod 29), what does your discovery from Problem 2.1(c) tell you about which values of a are allowed now, and how many valid keys are there in this situation?

4. In some systems, different keys can produce the same transformation, so the keyspace seems larger than it actually is. Consider the following illustration of this: Joe decides to strengthen the

affine cipher in the preceding question by performing the affine cipher twice with two different keys: the plaintext is first transformed using key $[a_1, b_1]$, giving intermediate ciphertext C_1, and then C_1 is transformed using a different key $[a_2, b_2]$ to produce the final ciphertext C. The full key is then the set of four values a_1, b_1, a_2, b_2, where each $[a_i, b_i]$ satisfies the necessary requirements for an affine cipher.

(a) How many such keys are there?

(b) How many different transformations are possible, and how can a brute force search be performed much faster than the number of keys from part (a) would suggest?

5. The book describes the one-time pad on pages 47 and 48 using letter-based operations (mod 26), and we are discussing the exclusive-or binary version of a one-time pad in class. In this question, each bit of the key is randomly chosen, independently of all other key bits, and we explore the importance of a uniform distribution for key bits.

   (a) Consider a system in which each bit of the key is uniformly distributed, so has probability 1/2 of being 0 and 1/2 of being 1. Given a one-byte ciphertext that is hexadecimal $A7$ (binary 10100111) what is the probability that the plaintext is $4F$? What is the probability that the plaintext is 58? Is there any rational reason to prefer one of these possible plaintexts over the other?

   (b) Consider a system in which the bits of the key are biased: each bit is 0 with probability 1/4 and 1 with probability 3/4. Given a one-byte ciphertext that is hexadecimal $A7$ what is the probability that the plaintext is $4F$? What is the probability that the plaintext is 58? Is there any rational reason to prefer one of these possible plaintexts over the other?

6. Consider the Feistel network shown in the book in Figure 3.3 (page 69). Consider a round function F that satisfies the formula $F(\overline{R}, \overline{K}) = F(R, K)$ for all inputs R and K, where the $\overline{X}$ notation indicates the bitwise complement of X. In words, if you consider computing $F(R, K)$ and then flip all the bits of both R and K, the result stays the same. This seems like a strange property for the round function to satisfy, but in fact the DES round function has this property!

   Let $E(K, P)$ refer to the full 16-round Feistel cipher in Figure 3.3, where the round function F satisfies this property. What is $E(\overline{K}, \overline{P})$? Give a clear justification of your answer — the justification is more important than the formula for the answer!

7. Textbook, page 82, problem 3.6.