Intro to Cryptography

Mark Anderson
Problem 3

May 16, 2019

1. Suppose we have an encryption function given by the following table: and assume that all plaintexts and keys are equiprobable.

   - Compute the probablities $p(C = w), p(C = x), p(C = y), and p(C = z)$. Since the plaintexts and the keys are all equiprobably, and w,x,y,z must appear in every key entry, the probabilities for each ciphertext is the same. An example of the calculation is shown below for $p(C = w)$
     $p(C = w) = p(K = k_1) * p(P = a) + p(K = k_2) * p(P = c) + p(K = k_3) * p(P = b) + p(K = k_4) * p(P = b)$
     $p(C = w) = (\frac{1}{4} * \frac{1}{4}) + (\frac{1}{4} * \frac{1}{4}) + (\frac{1}{4} * \frac{1}{4}) + (\frac{1}{4} * \frac{1}{4})$
     - $p(C = w) = \frac{1}{4}$
     - $p(C = x) = \frac{1}{4}$
     - $p(C = y) = \frac{1}{4}$
     - $p(C = z) = \frac{1}{4}$
   - Compute the conditional probabilities $p(P = m|C = c)\forall$ plaintexts m, and ciphertexts c. We can compute the probabilities using the below formula
     $p(P = m|C = c) = \frac{p(P=m)p(C=c|P=m)}{p(C=c)}$
     - $C = w$:
       $p(P = a|C = w) = \frac{1}{4}$
       $p(P = b|C = w) = \frac{1}{2}$
       $p(P = c|C = w) = \frac{1}{4}$
       $p(P = d|C = w) = 0$
     - $C = x$:
       $p(P = a|C = x) = \frac{1}{4}$
       $p(P = b|C = x) = \frac{1}{4}$
       $p(P = c|C = x) = \frac{1}{4}$
       $p(P = d|C = x) = \frac{1}{4}$
     - $C = y$:
       $p(P = a|C = y) = \frac{1}{4}$
       $p(P = b|C = y) = 0$
       $p(P = c|C = y) = \frac{1}{4}$
       $p(P = d|C = y) = \frac{1}{2}$

- $C = z$:
  $p(P = a|C = z) = \frac{1}{4}$
  $p(P = b|C = z) = \frac{1}{4}$
  $p(P = c|C = z) = \frac{1}{4}$
  $p(P = d|C = z) = \frac{1}{4}$

- Argue from the above conditional probabilities whether the above encryption function is good or not.

  the encryption scheme is not very good, as there isn't an equal distribution of possible ciphertexts for (Key, Plaintext) pairs, and some pairs are unable to generate certain ciphertexts. This allows an attacker to leverage a ciphertext attack to leak more information about the key. For example, if we see the ciphertext w we know that the plaintext cannot be D, and it is more likely to be B than it is A and C.

- Compute $H(k), H(p), H(c), H(K|C)$. Does this support your prior reasoning?
  - H(K): $-4 * log_2(\frac{1}{4}) = 2$
  - H(P): $-4 * log_2(\frac{1}{4}) = 2$
  - H(C): $-4 * log_2(\frac{1}{4}) = 2$
  - H($K|C$): $-\sum_k \sum_c p(C = c) * p(K = k|C = c) * log_2 p(K = k|C = c) = 2$
  - These values support my reasoning from above, each part of this cryptosystem leaks around 2 bits of information, which is the upper bound of entropy that this system is capable of leaking.

2. Compute and approximation to the unicity distance of the Caesar Cipher and directly relate that to the ease of breaking this cipher. The Unicity distance can be computed by $U = H(k)/D$. With U representing the Unicity Distance, H(K) representing the probability of the keyspace, and D representing the plaintext redundancy.

   For the Caesar Cipher using the english language, we know that the entropy per character for english is $-1(log_2)(1/26) = 4.7$ bits per letter. This however represents the worst case scenario for the english language, where every letter is equiprobable, when in fact this is not the case with certain characters appearing more frequently than others (e, a, s...) and even moreso with bigrams and trigrams where probabilities of characters occur more frequently after other characters (TH in the). For a practical calculation of the unicity distance we need an accurate D value, a precomputed D value for the english language is $D = 3.2$ which is what we will use. Using our information we can calculate the Unicity distance

   $U = \frac{log_2(25)}{3.2} = 1.4688$ which is a very very small amount of ciphertext needed to recover the key. This shows that the caesar cipher is incredibly weak and susceptible to ciphertext attacks

3. Prove that for entropy H,

   $H(X, Y) \leq H(X) + H(Y)$

   This only holds true if both X and Y are independent variables. If X,Y are independent variables then the probability of their individual outcomes does not rely on the outcome of any other variable. Given this, we know that the Probability Mass Function for both of these variables must equal 1 (there cannot be a non existant probabilty). Because the joint entropy is defined as $-\sum_i^n \sum_j^m r_{i,j} log_2 r_{i,j}$ we can see that the individual probabilities of our variables are multiplied together. We know that the product of 2 numbers $0 \leq < x \leq 1$ will always be lower than the sum of those 2 numbers. Given this, the best case for the joint entropy to be greater than the sum of its parts would be where X and Y both have one outcome, with 1.0 probability. Even with this case that would have the best case being $(1 * 1) * log_2 * (1 * 1)$ being the joint entropy, which is still less than $(1 * log_2 * 1) + (1 * log_2 * 1)$. This property is called subadditivity.

4. All vowels and spaces have been removed from the following sentence, recover the original sentence.

   Answer: ”what would you like for lunch today”

5. H(P^3) = 2.961