The University of Newcastle
School of Electrical Engineering and Computer Science

COMP3260/COMP6360 Data Security
Week 4 Workshop – 21st and 22nd March 2019
Solutions

1) Consider the following ciphertexts:

- XXXXX
- VWXYZ
- RKTIC
- JZQAT

Which of the these ciphertexts could result from enciphering five-letter words of English using:

a) A substitution cipher, where each letter is replaced with some other letter, but the letters are not necessarily shifted as in the Caesar cipher (thus A could be replaced with K, B with W, etc).
b) Any transposition cipher.

Solution:

a) VWXYZ, RKTIC, JZQAT

b) RKTIC

Explanation:

With XXXXX there is no word in English with 5 of the same letter, so for part a) there is no letter that makes sense in a substitution. For b), no matter how we transpose we still get XXXXX.

For VWXYZ we could find a substitution that would decipher this to an English word. One example would be V=T, W=R, X=I, Y=C and Z=K. In fact, the plaintext could be any five letter word that has no repeated letters. However, for a transposition cipher, there is no word in English with these letters and no (other) vowels.

RKTIC follows same logic for substitution, and also works for a transposition (TRICK).

JZQAT, see explanation for VWXYZ.

2) Intercepted ciphertext is C=TEHAOEHIETURRNBTNIETOWDT. Single letter frequency analysis indicates that this is a transposition cipher. Find the matching plaintext.

Solution:

This is a columnar transposition with period d =4.

THRE
EINT
HEBO
ATTW
OUND
ERIT

Reading down the columns in order, we get the ciphertext. But reading across the rows, we get the plaintext; THREE IN THE BOAT TWO UNDER IT.

3) Consider a homophonic cipher that uses 26h ciphertext symbols, assigning h homophones to each letter of the English alphabet. Determine the number of possible keys (i.e., assignments of homophones), and use your result to calculate unicity distance of the cipher.

Solution:

The unicity distance is the smallest N such that HC(K) is close to 0. Can be estimated by the formula:

U = H(K)/D, where D is assumed to be 3.2.

To obtain the number of keys, we have:

NoOfKeys = $\binom{26h}{h}\binom{25h}{h}\binom{24h}{h}\ldots\binom{2h}{h}\binom{h}{h}$

Using $\binom{n}{r} = \frac{n!}{(n-r)!r!}$ we get:

NoOfKeys = $\frac{(26h)!}{(25h)!h!} \times \frac{(25h)!}{(24h)!h!} \times \frac{(24h)!}{(23h)!h!} \times \ldots \frac{(2h)!}{(h)!h!} \times \frac{h!}{0!h!}$

By definition, 0!=1

NoOfKeys = $\frac{(26h)!}{h!^{26}}$

So

$$U = \frac{log_2 \frac{(26h)!}{h!^{26}}}{3.2}$$

When h = 1:

$$U = \frac{log_2 26!}{3.2} \approx \frac{88.38}{3.2} \approx 27.62$$

This is effectively monoalphabetic.

For large h:

We use Sterling's formula $log_2(d!) \approx d\ log_2 \frac{d}{e}$

$log_2 \frac{(26h)!}{h!^{26}} = log_2(26h)! - log_2 h!^{26} = log_2(26h)! - 26\ log_2 h!$

Using Sterling's approximation

$log_2 \frac{(26h)!}{h!^{26}} = log_2(26h)! - 26\ log_2 h! \approx 26h\ log_2 \frac{26h}{e} - 26h\ log_2 \frac{h}{e}$

$= 26h\ log_2(26\text{h}) - 26h\ log_2 \text{e} - 26h\ log_2 h + 26h\ log_2 e = 26h\ log_2 26 + 26h\ log_2 \text{h} - 26h\ log_2 h$

$= 26h\ log_2 26$

Then $U \cong \frac{122.2h}{3.2} \cong 38.2h$

4) A generalization of the shift cipher, known as the affine cipher, is as follows: f(p)=(ap+b) mod 26. A requirement that any encryption function needs to satisfy is to be *one-to-one*, that is, if p≠q then f(p) ≠ f(q), otherwise the encryption would be impossible, as more than one plaintext character maps into the same ciphertext character. The affine cipher is not necessarily one-to-one; for example, for a=2 and b=3, f(0)=f(13)=3.

a) Are there any limitations on the value of b? If yes, determine which values are not allowed.
b) Are there any limitations on the value of a? If yes, determine which values are not allowed.
c) Provide a general statement of which values of a and b are not allowed.

Solution:

a) No. A change in the value of b shifts the relationship between plaintext letters and ciphertext letters to the left or right uniformly, so that if the mapping is one-to-one it remains one-to-one.
b) 2, 4, 6, 8, 10, 12, 13, 14, 16, 18, 20, 22, 24. Any value of a larger than 25 is equivalent to a mod 26.
c) The values of a and 26 must have no common integer factor other than 1. This is equivalent to saying that a and 26 are relatively prime, or that the greatest common divisor of a and 26 is 1. To see this, first note that f(p) = f(q) ($0 \le p \le q < 26$) if and only if a(p – q) is divisible by 26.
   1. Suppose that a and 26 are relatively prime. Then, a(p – q) is not divisible by 26, because there is no way to reduce the fraction a/26 and (p – q) is less than 26.
   2. Suppose that a and 26 have a common factor k > 1. Then E(a, p) = E(a, q), if $q = p + m/k \neq p$.

5) How many one-to-one affine ciphers are there?

Solution:

There are 12 allowable values of a (1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25). There are 26 allowable values of b, from 0 through 25). Thus the total number of distinct affine Caesar ciphers is $12 \times 26 = 312$.

6) A ciphertext has been generated with an affine cipher. The most frequent letter of the ciphertext is “B”, and the second most is “U”. Based on this information, try to break the cipher.

Solution:

Assume that the most frequent plaintext letter is e and the second most frequent letter is t. Note that the numerical values are e = 4; B = 1; t = 19; U = 20. Then we have the following equations:
1 = (4a + b) mod 26
20 = (19a + b) mod 26
Thus, 19 = 15a mod 26. By trial and error, we solve: a = 3. Then 1 = (12 + b) mod 26. By observation, b = 15.

7) In one of his cases, Sherlock Holmes was confronted with the following message:

534 C2 13 127 36 31 4 17 21 41 douglas 109 293 5 37 birlstone 26 birlstone 9 127 171

Although Watson was puzzled, Holmes was able immediately to deduce the type of cipher. Can you?

Solution:

The cipher refers to the words in the page of a book. The first entry, 534, refers to page 534. The second entry, C2, refers to column two. The remaining numbers are words in that column. The names DOUGLAS and BIRLSTONE are simply words that do not appear on that page. Elementary! (from The Valley of Fear, by Sir Arthur Conan Doyle)

8) For each one of the following ciphers estimate the unicity distance, assuming that all keys are equally likely:

a) Transposition cipher with period d

b) Shift cipher $c = (p + k) \bmod 26$

c) Affine cipher $c = (k_1 p + k_0) \bmod 26$, where $\gcd(k_1, n)=1$

d) General monoalphabetic substitution cipher

Solution:

The unicity distance is defined as $U = H(k)/D$, where $H(k)$ is the entropy of the key k, and D is the redundancy of the language. For English, we estimate $D=3.2$

a) There are d! possible keys so $U = \lg d! / 3.2$. For example, when $d = 27$, there are 27! possible keys so $U = \lg 27! / 3.2 \approx 27 \times \lg (27/2.7) /3.2 \approx 27 \times \lg 10 / 3.2 \approx 27 \times 3.3 / 3.2 \approx 27.84$

b) There are 26 possible keys, so $U = \lg 26 / 3.2 \approx 1.47$

c) There are 12×26 possible keys, thus $U = \lg (12 \times 26) / 3.2 = (\lg 12 + \lg 26) / 3.2 = (\lg (22 \times 3) + \lg 26) / 3.2 = (\lg 22 + \lg 3 + \lg 26) / 3.2 \approx (2 + 1.58 + 4.7) / 3.2 = 8.28 / 3.2 \approx 2.57$

d) There are 26! possible keys so $U = \lg 26! / 3.2 \approx 88.4/3.2 \approx 27.62$