Cryptology

Assignment 1: Vigenere

JH Darbyshire and Bengu Erenler

Part B

The following are the results of our decryptions:

Part C

Extended Alphabets

We believe extending the number of characters from 26 to 32 in general makes a cipher more difficult to break, particularly from a brute force point of view. Similarly would we expect the level of difficulty to diminish if the number of characters was reduced.

Indirectly, extending the number of characters may make, however, a cipher easier to break if those additions brought about strong statistical elements. For example our methods rely on letter frequencies being a signature. If an alphabet of 26 characters existed with equal probability of character appearance in the plaintext then this would likely be more difficult to break than a 32 character alphabet statistically favouring some characters more than others.

A mathematical argument for this can be made in terms of entropy: suppose a random character (X_2) from a two character alphabet has equal letter frequency then the entropy is:

$$H(X_2) = -\sum P(x_i) log_2(P(x_i)) = 1 \quad \text{where} \quad P(x_i) = 0.5 \quad \forall\, i\ .$$

The entropy for a similar character X_3in a similarly distributed set of three letters is:

$$H(X_3) = 1.58\ .$$

Whereas the entropy for a character X_{3*} three character set with probability density {0.1, 0.1, 0.8} is:

$$H(X_{3*}) = 0.92\ .$$

In this final example we would consider the unpredictability of any given character less than in the original case with two characters.

Cipher Adjustments

(a) Reversed Plaintext Key

This is not sensible and does not represent a cryptosystem: either both the key and message will be known to the recipient, or both will be unknown to the recipient. In the first case the purpose of the message is made moot, and in the second case the recipient is no better positioned than any attacker. Incidentally the ciphertext will be symmetric so this pattern can also be observed.

(b) Changed Operations

This is not sensible and does not represent a cryptosystem. A simple example shows that decryption does produce the original plaintext: let "D=3" be the plaintext, and "B=1" is the key, then the ciphertext is "C=1-3 mod M>3" and the recovered plaintext is "B=1-2 mod M>3", which is not the same as the original "D".

(c) Transforming Plaintext

This defines a cryptosystem and is likely to be more secure that the Vigenere cipher of initial keylength, but not necessarily. It usually defines a Vigenere cipher of increased keylength, via a transformed key. For example consider a message encrypted with a keylength of three:

Plaintext	a	b	c	d	e
Key	k_1	k_2	k_3	k_1	k_2
Ciphertext	$a+k_1$	$b+k_2$	$c+k_3$	$d+k_1$	$e+k_2$
Transformed Plaintext	a	$b-a$	$c-b$	$d-c$	$e-d$
Transformed Ciphertext	$a+k_1$	$b-a+k_2$	$c-b+k_3$	$d-c+k_1$	$e-d+k_2$
Transformed Key	k_1	$b+k_1+k_2$	$c+k_1+k_2+k_3$	$d+k_1+k_2+k_3+k_1$	$e+...$

Therefore this is equivalent to a traditional Vigenere cipher with transformed key = $\{\ K_1 = k_1, K_2 = k_1 + k_2, K_3 = k_1 + k_2 + k_3, K_4 = k_1 + k_2 + k_3 + k_1, ...\ \}$. For some value a, $a(k_1 + k_2 + k_3) \equiv 0 mod 3$ and the transformed key will

revert to the traditional pattern. If a is 1 then the keylength is not increased so the cipher is no more secure than its original, this will be the result for a poorly chosen original key. A carefully chosen original key can result in a transformed keylength equal to $L * M$ for L the original keylength and M the modulo of the alphabet. However the set of keys than can be generated in this way is far less than the abitrary set of keys of the same length.

For example consider the alphabet { 0,1 } and the set of original keys of length two, { 00, 01, 10, 11 }. This will generate the transformed keys, { 00[00], 0110, 1100, 10[10] }. But for arbitrary keys of length 4 there are 16 possible candidates, so although this is more difficult than the traditional Vigenere it is less difficult than a traditional Vigenere with an arbirary key taken from the same possible keylength space.

(d) Autokey Cipher

The suggested addition here is known as the autokey cipher. It is considered more difficult to break than the traditional Vigenere cipher since the non-repeating key prevents analysis such as Kasiski based on repeating elements and frequency of letters. However, since the the plaintext forms part of the key common words, bigrams and trigrams can be successively attempted and deduced from the resulting decryptions.

(e) Failed Autokey Cipher

This is not sensible. The autokey cipher works because the key is unknown. Using the ciphertext as a key leaves open the possibility of inserting it as a key and attempting to break the cipher.