HW #1

1. Ceasar Cypher with key shift 7

Vul dhf av rllw tvtluabt nvpun pz av ohcl jvuzahuasf nylhaly nvhsz.

One way to keep momentum going is to have constantly greater goals.

1. Frequency analysis works differently from language to language. In English the letters E,T, and A are used most commonly while the letters X, Q and Z appear the least.

Looking at the given text we can plainly see that some letters here are used much more often than others. Let us move on to the assignment.

FX IWBBJX PB NB PB PWX GBBD. VSP FWO, JBGX JRO, PWX GBBD? FWO IWBBJX PWUJ RJ BSA NBRK? RDL PWXO GRO FXKK RJM FWO IKUGV PWX WUNWXJP GBSDPRUD? FWO, 35 OXRAJ RNB, EKO PWX RPKRDPUI? FWO LBXJ AUIX CKRO PXQRJ? FX IWBBJX PB NB PB PWX GBBD UD PWUJ LXIRLX RDL LB PWX BPWXA PWUDNJ, DBP VXIRSJX PWXO RAX XRJO, VSP VXIRSJX PWXO RAX WRAL, VXIRSJX PWRP NBRK FUKK JXATX PB BANRDUZX RDL GXRJSAX PWX VXJP BE BSA XDXANUXJ RDL JMUKKJ, VXIRSJX PWRP IWRKKXDNX UJ BDX PWRP FX RAX FUKKUDN PB RIIXCP, BDX

Running a letter counting script on the cypher we can get a percent breakdown by each letter

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | P | B | R | W | J | D | K | U | I | O | A | F | N | S | L | G | V | C | E | M | Q | T | Z | H | Y |
| 52 | 36 | 35 | 34 | 29 | 26 | 19 | 16 | 16 | 14 | 14 | 13 | 11 | 11 | 10 | 9 | 8 | 8 | 2 | 2 | 2 | 1 | 1 | 1 | 0 | 0 |

Knowing that the 4 most commonly used letters are E,T,A,O we can use run a substitution on our encoded message to see if we can make sense of it. We get:

Fe IWaaJe ta Na ta tWe GaaD. VSt FWO, JaGe JoO, tWe GaaD? FWO IWaaJe tWUJ oJ aSA NaoK? oDL tWeO GoO FeKK oJM FWO IKUGV tWe WUNWeJt GaSDtoUD? FWO, 35 OeoAJ oNa, EKO tWe otKoDtUI? FWO LaeJ AUIe CKoO teQoJ? Fe IWaaJe ta Na ta tWe GaaD UD tWUJ LeIoLe oDL La tWe atWeA tWUDNJ, Dat VeIoSJe tWeO oAe eoJO, VSt VeIoSJe tWeO oAe WoAL, VeIoSJe tWot NaoK FUKK JeATe ta aANoDUZe oDL GeoJSAe tWe VeJt aE aSA eDeANUeJ oDL JMUKKJ, VeIoSJe tWot IWoKKeDNe UJ aDe tWot Fe oAe FUKKUDN ta oIIeCt, aDe

Right off the bat we can see a lot of 3 letter words that look like “tWe” the most common 3 letter word similar to this is “the” meaning we can safely substitute W for H. but before we run that we can take a look at a small sentence fragment “35 OeoAJ oNa”. “e” is in the second position, making me think that this is “35 years”, however “o” is out of place, since “a” and “o” are close on the use list we might have switched the two. In the next run we are substituting **W to H**, **O to Y**, **A to R** and **J to S**, also we will switch our “o” and “a” substitutions.

After this run things are starting to clear up, we are getting full and partial words that make sense:

Fe Ihoose to No to the GooD. VSt Fhy, soGe say, the GooD? Fhy Ihoose thUs as oSr NoaK? aDL they Gay FeKK asM Fhy IKUGV the hUNhest GoSDtaUD? Fhy, 35 years aNo, EKy the atKaDtUI? Fhy Loes rUIe CKay teQas? Fe Ihoose to No to the GooD UD thUs LeIaLe aDL Lo the other thUDNs, Dot VeIaSse they are easy, VSt VeIaSse they are harL, VeIaSse that NoaK FUKK serTe to orNaDUZe aDL GeasSre the Vest oE oSr eDerNUes aDL sMUKKs, VeIaSse that IhaKKeDNe Us oDe that Fe are FUKKUDN to aIIeCt, oDe

Breaking down by parts we can choose some more sentence fragments to try to find more words to decipher

“Fhy Ihoose thUs” might be “why choose this” ----- substitute **F to W**, **I to C**, **U to I**

“hUNhest GoSDtaUD?” Taking into account our changes from above, this now looks like

“hiNhest GoSDtaiD?” which in turn might make this phrase be “highest mountain”

Substitute **N to G**, **G to M**, **S to U**, **D to N**.

Now our coded message looks like this:

we choose to go to the moon. Vut why, some say, the moon? why choose this as our goaK? anL they may weKK asM why cKimV the highest mountain? why, 35 years ago, EKy the atKantic? why Loes rice CKay teQas? we choose to go to the moon in this LecaLe anL Lo the other things, not Vecause they are easy, Vut Vecause they are harL, Vecause that goaK wiKK serTe to organiZe anL measure the Vest oE our energies anL sMiKKs, Vecause that chaKKenge is one that we are wiKKing to acceCt, one

Now the message is essentially decoded and it is simply a matter of substituting the last few letters. So lets change them out.

**V to B** from “Vut why”

**K to L** and **C to P** from “wiKKing to acceCt”

**L to D** from the common occurrence of anL

**M to K** from “they may weKK asM”

With those substitution our message becomes this:

we choose to go to the moon. but why, some say, the moon? why choose this as our goal? and they may well ask why climb the highest mountain? why, 35 years ago, Ely the atlantic? why does rice play teQas? we choose to go to the moon in this decade and do the other things, not because they are easy, but because they are hard, because that goal will serTe to organiZe and measure the best oE our energies and skills, because that challenge is one that we are willing to accept, one

Last but not leas we replace **Q with X, T with V and E with F**. Interestingly enough we leave z unchanged. In the end with those final substitutions our message becomes:

we choose to go to the moon. but why, some say, the moon? why choose this as our goal? and they may well ask why climb the highest mountain? why, 35 years ago, fly the atlantic? why does rice play texas? we choose to go to the moon in this decade and do the other things, not because they are easy, but because they are hard, because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one

We are done with deciphering the message, all it took was analyzing the 4 most common letters and a few key words which could be deciphered. Here, the takeaway for me is if I ever have to encode something, I should not leave numbers in number form, instead I should write them out.

1. Since we have the choice of cyphertext, we can choose a cyphertext where all bits beside 1 are 0. This means that there is a Ci which we can XOR to produce C. To do so the attacker would do the following  
     
   m1 XOR m2 XOR m3 XOR…. Until the desired result is achieved and the attacker has decrypted the message.
2. a. AES is a Stream cipher intended to replace DES for commercial applications. False, AES is a block cypher that was developed as a result of DES becoming more susceptible to brute force attacks, however both are still in use.

b. In the AES the decryption algorithm is not identical to the encryption algorithm.

True, AES uses a symmetric encryption algorithm. Meaning that the same key is used to both encrypt and decrypt data. Main difference is it, the key, is used in reverse during the decryption process.

c. Your friend who is new to the world of computer security says, for block ciphers any random strategy for substitution and permutation will not guarantee the cipher to be strong always. Is your friend correct or not? Justify the answer.

I believe that friend is correct. In an ideal scenario a block cypher will be random, however it needs to be invertible so that a decryption could be completed. Meaning that each of the input blocks is mapped to a unique output block. Meaning that in the end the result we get is pseudo-random.

1. Given: 2GHz processor means 2\*10^9 operations per second, DES(2^56 keys), AES(2^128 keys), 100 cycles per brute force

Time to crack DES: ((2^56)\*100)/2\*10^9 = 3,602,879,701.9 seconds or about 114 years

Time to crack AES: ((2^128)\*100)/2\*10^9 = 1.7\*10^31 Seconds or about 5.39X10^23 years