Abhishek Prajapati

Cryptography

Professor Brian Garnett

Extra Credit Assignment

**Substitution Cipher**

QKQIY TIKBM YQKQI YTQKQ IFQIM HQLEI LBWFE KRCKY MLEIC MIENM

HJJTW KYYWI SFEIR EDSEF DJISF EIMIE ITIPR QLPYF FLSFE RDFDI

QYCEK QGWKQ BMIEQ FPIRQ LJWKQ GKQBK QYMIP HQWKB MYYMI QYETK QBYFC RYNMK QRWWL KEINY KFQPR YFQNI PMICK YMLEI CRSIC NRHYK

FHPPY IZPSE FDYMI YEIIR BRKQP YCMKN MPMIM RLJII QWIRQ KQBPM

INERQ ILMIE QINGY FKQAI PYKBR YIYMI ZFPPK JKWKY KIPFS RNWHD

ZFSJH PMIPY FYMIE KBMYR QLYMI QCKYM LEICS REYMI EYFFJ YRKQR

JIYYI ERQBW ISFEA KICKQ BKYPL REGEI NIPPI PYMIO HKIYC RPZEF

SFHQL IUNIZ YSFEY MIKQN IPPRQ YJHVV KQBFS KQPIN YPRQL YMIFN

NRPKF QRWNM KEEHZ FSPFD IMREL JTJKE LJERA KQBYM IDKLL RTPHQ

* First thing I did is count the number of characters in the text, which are

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **In** | I | Y | Q | K | E | M | F | R | P | L | N | B | S | J | W | C | H | D | T | Z | G | A | V | O | U |
| **Count** | 63 | 46 | 44 | 43 | 34 | 33 | 31 | 31 | 30 | 19 | 16 | 14 | 14 | 13 | 13 | 12 | 12 | 8 | 7 | 6 | 4 | 3 | 2 | 1 | 1 |
| **Result** | E | T | N | I | R | H | O | A | S | D | C | G | F | B | L | W | U | M | Y | P | K | V | Z | Q | X |

* The most repeated character in cipher text is ‘I’ and second most used character is ‘Y’ and the most used character in English language is ‘E’, and second most used character is ‘T’ according to **letterfrequency.org** So I replaced each character ‘I’ in the cipher text with ‘E’ and ‘Y’ with ‘T’.
* Then the most used three letter word in the cipher text is ‘YMI’ (given in question paper) and since Y=T and I = E, I figured M = H because “the” is used a lot in English. So, replaced each M with H.
* Now the second most used bigram is ‘MI’ in the cipher text and which we know is M=H, I=E so MI=HE and ‘MIE’ is also 3rd most used trigram so the only one word (her) that make sense is M=H, I=E, so E = R so MIE = HER. So I replaced E with R
* Second most used trigram is ‘KQB’ in the cipher text, and 2nd most used trigram in English is “AND” or “ING”. Since it was unclear for me to decide, there was one letter I can use for sure is Q=N, since N is in the middle in both. So I replace Q=N.
* Now the most used bigram in the cipher text is KQ, and we know Q=N. And one of the most used bigram according to **letterfrequency.org** is “IN”, and ‘I’ is also one of the most used letter English language. So, I choose K=I.
* Now, the first five letters in the cipher text is QKQIY, and we have solve Q=N, K=I, I=E,Y=T , SO QKQIY=NINET and the next letter in the cipher text is T, so good guess would be T=Y so QKQIYT becomes NINETY. So now we have T=Y.
* Now iterating through the text we get NINETY EI\_HT NINETY NINE, so G will go in the blank spot because it makes sense “NINETY EIGHT NINETY NINE”. So B=G.
* Now iterating one character at a time we get \_NE H\_N\_RE\_. It looks like a number “ONE HUNDRED”. This could also be a number because there are number in the front of that too. (NINETY EIGHT NINETY NINE). So, now we have F=O, H=U, L=D
* Another used bigram in cipher text is “RQ” and trigram is “RQL” and we know Q=N, L=D, so we can assume R = A, because AN and AND is used largely in English, and R is also used largely in cipher text.
* Now iterating first two line substituting the letters we know and and blank for the rest to see if I can find more words.
  + This is what we know so far

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **In** | I | Y | Q | K | E | M | F | R | P | L | N | B | S | J | W | C | H | D | T | Z | G | A | V | O | U |
| **Count** | 63 | 46 | 44 | 43 | 34 | 33 | 31 | 31 | 30 | 19 | 16 | 14 | 14 | 13 | 13 | 12 | 12 | 8 | 7 | 6 | 4 | 3 | 2 | 1 | 1 |
| **Result** | E | T | N | I | R | H | O | A |  | D |  | G |  |  |  | W | U |  | Y |  |  |  |  |  |  |

QKQIY TIKBM YQKQI YTQKQ IFQIM HQLEI LBWFE KRCKY MLEIC MIENM

HJJTW KYYWI SFEIR EDSEF DJISF EIMIE ITIPR QLPYF FLSFE RDFDI

NINET YEIGH TNINE TYNIN EONEH UNDRE DG**L**OR IA\_IT HDRE\_ HER\_H

U\_ \_Y**L**ITT**L**E \_OREA R\_ \_RO \_ \_E\_O REHER EYE\_A ND\_TO OD\_OR A\_O\_E

* After the letters ninety eight ninety nine one hundred there is G\_ORIA, that sounds like a name GLORIA, so W=L
* Now after plugging in W=L, reading in further **GLORIA** \_IT HDRE\_ **HER,** so reading this looks like “Gloria withdrew her …..” So, C=W.
* Iterating the same two lines again we see that “GLORAI WITHDREW HER \_HY\_ \_Y LITTLE **\_OREA R\_** \_E…..
  + After “LITTLE” it looks like a word “forearm” So, S=F and D=M
* Now filling in again the first two lines with what we have

NINET YEIGH TNINE TYNIN EONEH UNDRE DG**L**OR IA\_IT HDRE\_ HER\_H

U\_ \_YL ITTLE FOREA RMFRO M**\_EFO RE**HER **EYE\_**A ND\_TO ODFOR A\_O\_E

* Reading this we can see that “LITTLE FOREARM FROM \_EFORE HER EYE\_ AND”
  + So, the words that can be made from those are ‘BEFORE’ and ‘EYES’
  + So J=B, P=S
* Now plugging in everything we have in first two line
  + NINET YEIGH TNINE TYNIN EONEH UNDRE DGLOR IAWIT HDREW HER\_H
  + UBBYL ITTLE FOREA RMFRO MEEFO REHER EYESA NDSTO ODFOR AMOME
* Now looking at those two line we see that there is only one word left in between “WITHDREW HER **\_HUBBY** LITTLE FOREARM
  + This world could be “CHUBBY” (chubby little forearm?)
  + So, I plugged in N=C.
* Now plugging in what we have in the entire text

NINET YEIGH TNINE TYNIN EONEH UNDRE DGLOR IAWIT HDREW HERCH

UBBYL ITTLE FOREA RMFRO MBEFO REHER EYESA NDSTO ODFOR AMOME

NT**WRI N\_LIN G**HERN OSEAN DBLIN \_INGI NTHES UNLIG HTTHE NTRYI

NGTOW ATCHI NALLD IRECT IONSA TONCE SHEWI THDRE WAFEW CAUTI

OUSST E\_SFR OMTHE TREEA GAINS TWHIC HSHEH ADBEE NLEAN INGSH

ECRAN EDHER NEC\_T OIN\_E STIGA TETHE \_OSSI BILIT IESOF ACLUM

\_OFBU SHEST OTHER IGHTA NDTHE NWITH DREWF ARTHE RTOOB TAINA

BETTE RANGL EFOR\_ IEWIN GITSD AR\_RE CESSE STHE **\_ UIET** W AS\_RO

FOUND **E\_CE\_ T** FORT HEINC ESSAN T **BU\_ \_ IN G**OF INSEC TSAND THEOC

CASIO NALCH IRRU\_ OFSOM EHARD BYBIR D **B RA\_ ING** TH EMIDD AYSUN

* Now, I iterate through the new text and tried to fill in the blank by possible character
  + In the first highlighted word we see that “Moment wrin\_ling her” So G=K
  + Red highlighted word spells “quite” so O=Q
  + Blue highlighted word spells “except” so U=X, and Z=P
  + Green highlighted word spells “buzzing” so V=Z
  + Purple highlighted word spells “braving” so A=V

* Now putting all the characters together with proper spacing

NINETY EIGHT NINETY NINE ONE HUNDRED GLORIA WITHDREW HER CHUBBY LITTLE FOREARM FROM BEFORE HER EYES AND STOOD FOR A MOMENT WRINKLING HER NOSE AND BLINKING IN THE SUNLIGHT THEN TRYING TO WATCH IN ALL DIRECTIONS AT ONCE SHE WITHDREW A FEW CAUTIOUS STEPS FROM THE TREE AGAINST WHICH SHE HAD BEEN LEANING SHE CRANED HER NECK TO INVESTIGATE THE POSSIBILITIES OF A CLUMP OF BUSHES TO THE RIGHT AND THEN WITHDREW FARTHER TO OBTAIN A BETTER ANGLE FOR VIEWING ITS DARK RECESSES THE QUIET WAS PRO FOUND EXCEPT FOR THE INCESSANT BUZZING OF INSECTS AND THE OCCASIONAL CHIRRUP OF SOME HARD BY BIRD BRAVING THE MIDDAY SUN

**Source for looking up letter frequencies**

**http://letterfrequency.org/**