

Ministerul Educatiei, Culturii și Cercetarii al Republicii Moldova Universitatea Tehnică a Moldovei

Facultatea Calculatoare, Informatică şi Microelectronică Departamentul Ingineria Software și Automatica

# Raport

# pentru lucrarea de laborator Nr. 2

# la cursul „Metode criptografice de protecție a informației”

A efectuat:

Konjevic Alexandra, gr. FAF-213

A verificat: Mîțu Cătălin

Chișinău – 2023

**Subject**: Cryptanalysis of monoalphabetic ciphers

**Objectives:**

Given an encrypted message, decrypt it and obtain the origin message, using the frequency analysis.

**Theory:**

Frequency analysis is a foundational method in cryptanalysis used to crack encrypted messages, especially those encrypted with substitution ciphers. It relies on the observation that certain letters occur more frequently in most languages. By scrutinizing the distribution of letter frequencies in an encrypted message, analysts can make informed assumptions about their corresponding counterparts in the plaintext, ultimately leading to decryption. This technique is most effective when dealing with simple substitution ciphers, such as the Caesar cipher, where each letter in the plaintext is replaced by another letter in the ciphertext according to a fixed rule or key.

The process of frequency analysis typically involves collecting a sample of the ciphertext, counting the occurrences of each letter, comparing these frequencies with the expected distribution for the language in which the message is likely written (e.g., English), and making substitutions based on the most frequent letters in the ciphertext. As more substitutions are made, patterns emerge, making it easier to decode the entire message. However, it's essential to consider the context and potential variations in language and content that may affect the accuracy of the decryption. While frequency analysis is a potent tool for deciphering simple substitution ciphers, it becomes less effective against more advanced encryption methods designed to conceal predictable patterns in the ciphertext.

**Task Completion (Variant 22):**

We are given the cryptogram ***c***, which is an encrypted message in English:

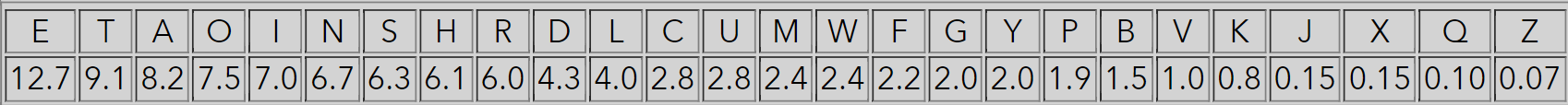
c = Rqxsv Qviaviw Ftiosvf ztf av wqv avpw lgnrg hifuwnsnjxpw,dghngwvpwtasf wqv jivtwvpw xp Rxssxtz Civovixhl Cixvoztg. Dgsxlv qxphngwvzunitif, qxp vzxgvghv xp odv znpw vzuqtwxhtssf wn rqtw qv oxo.Xgovvo, wrn zniv oxppxzxsti zvg xg t pxgjsv cxvso htg phtihvsf avxztjxgvo. Rqviv Ftiosvf rtp Itavstxpxtg, ndwjnxgj, pduvicxhxts, civv tgovtpf rxwq wqv ovwtxsp nc t jnno pwnif, tgo vkvi ivtof cni wqv ztxg hqtghv,Cixvoztg wvgovo wnrtio xgwinkvipxng, ovuwq nc pwdof, uvipngts pvhdixwf,wxzxoxwf, ovoxhtwxng, tgo thhdithf, gxhvwf, tgo ktsxoxwf nc rnil. Ovpuxwvwqv ivstwxkv oitagvpp nc wqvpv uvipngts witxwp—ni uviqtup avhtdpv ncwqvz, Cixvoztg'p wqvnivwxhts hngwixadwxngp tgo qxp uithwxhts twwtxgzvgwpvyhvvo wqnpv nc tgf nwqvi hifuwnsnjxpw. Ftiosvf'p htivvi rtp sxlv tgtztmxgj plfinhlvw wqtw vyusnovp xg ctgwtpwxh utwwvigp tjtxgpw wqvqvtkvgp. Cixvoztg'p rtp sxlv wqv pdg.Qv rtp anig Rnscv Cixvoztg ng Pvuwvzavi 24, 1891, xg Lxpqxgvk,Idppxt, wqv nsovpw png tgo pvhngo hqxso nc Civovixhl tgo Inpt Cixvoztg.Qxp ctwqvi, t Idztgxtg rqn punlv vxjqw stgjdtjvp tgo rnilvo tp tgxgwviuivwvi cni wqv Idppxtg Unpw Nccxhv, vzxjitwvo wn Tzvixht xg 1892, tw rqxhq wxzv qxp png'p gtzv rtp hqtgjvo wn Rxssxtz. Wqv ctzxsf pvwwsvoxg Uxwwpadijq, rqviv qxp ctwqvi ztgtjvo t pvrxgj zthqxgv tjvghf.Rxssxtz jitodtwvo xg 1909 tp ngv nc wqv wvg qngni pwdovgwp xg t hstpp nc300 tw Uxwwpadijq Hvgwits Qxjq Phqnns; qv wqvg rvgw wn rnil tp hqxvchsvil xg wqv Vixv Hxwf Xing Rnilp, t cxiz wqtw pnso pwvtz vgjxgvp. Tandwwqtw wxzv wqv athl-wn-wqv-ctiz znkvzvgw htssvo wn hxwf anfp, tgo xg wqvctss nc 1910, Cixvoztg tgo wqivv cixvgop vginssvo xg ZxhqxjtgTjixhdswdits Hnssvjv, rqnpv hqxvc twwithwxng rtp wqtw xw rtp wdxwxng-civv.Adw Cixvoztg pnng oxphnkvivo wqtw ctizxgj qvso sxwwsv xgwvivpw cni qxz.Qv rtp tg xgkvgwxkv fndgj cvssnr rqn sxlvo wn cxy wqxgjp tgo qto rixwwvgpnzv phxvghv cxhwxng cni qxp qxjq-phqnns utuvi; qv rtp ituxosf hnzxgj wnwqv hnghsdpxng wqtw qv sxlvo phxvghv. Tw wqv vgo nc wqv wviz qv svtigvowqtw wdxwxng rtp tspn civv xg t phxvgwxcxh cxvso tssxvo wn tjixhdswdiv—jvgvwxhp—tw ngv nc wqv Xkf Svtjdv dgxkvipxwxvp, Hnigvss. Qv aniinrvowitxg ctiv tgo tiixkvo xg Xwqtht, Gvr Fnil, xg Cvaidtif, 1911, rqviv qvjnw t ena rtxwxgj ng wtasvp. Tcwvi hnzzvghvzvgw xg Cvaidtif nc 1914, qvtwwvgovo jitodtwv phqnns, ztgtjxgj wn ctss xg snkv wrxhv, nghv rxwq taidgvwwv, nghv rxwq wqv asngov otdjqwvi nc t znkxv-qndpv nrgvi. Rqxsvqv rtp wqviv, t rvtswqf wvywxsv zvihqtgw, Jvnijv Ctaftg, rqnztxgwtxgvo stanitwnixvp xg thndpwxhp, hqvzxpwif, jvgvwxhp, tgo hifu-wnsnjf (wn wif wn uinkv wqtw Athng rinwv Pqtlvpuvtiv'p ustfp) ng qxp 500-thiv vpwtwv, Ixkviatgl, tw Jvgvkt, Xssxgnxp, ovhxovo wqtw qv gvvovo tjvgvwxhxpw wn xzuinkv wqv jitxgp tgo sxkvpwnhl ng qxp ctiz. Qv tuusxvo wnHnigvss cni t "rndso-av-vi," gnw tg "tp-xp-vi," tgo qxivo Cixvoztg, wnavjxg Edgv 1, 1915.Ctaftg rtp t ztg nc gn cnizts vodhtwxng adw nc xgwvssxjvghv tgovgvijf. Qv qto t jivtw ovpxiv wn av "pnzvanof," tgo wqtw ovpxivznwxktwvo qxp pdapxoxmxgj wqv Athngxtg pwdoxvp: uinnc nc wqxpivknsdwxngtif wqvpxp rndso hnkvi xwp utwing tp rvss tp xwp thwdtsoxphnkvivip rxwq jsnif. Qv qxzpvsc ivto sxwwsv, adw qv tapniavo vgndjqcinz wqnpv tindgo qxz wn ztlv qxp wtsl ng tsznpw tgf pdaevhw pndgoxzuivppxkv—tw svtpw pduvicxhxtssf. Qv rtp tdwnhitwxh, gvkvi tssnrxgj qxppwtcc wn oxptjivv rxwq qxz, adw nwqvirxpv gnw dgusvtptgw pn sngj tpvzusnfvvp ivhnjgxmvo wqtw qv rtp anpp. T htioxgts tiwxhsv nc ctxwq rxwqqxz rtp wqtw t rvss-vyvhdwvo ptsvp htzutxjg hndso udw thinpp tsznpwtgfwqxgj.

First of all, I found the frequencies of all the letters in the text:



*Figure 1. Frequencies of the letters in the message*

And here is the frequency of the letters in the English alphabet:



*Figure 2. Frequencies of the English alphabet*

The most frequent letter in the encrypted message is ‘V’. After counting the most frequent trigraphs, I get that “WQV” is the most frequent one, and most likely, it represents the “the” trigraph, which means that “W” -> “T” and “Q” -> “H”. After substituting these letters, the message becomes:

c = RhXSe heIAeIt FTIOSeF ZTF Ae the AePt LGNRG HIFUtNSNJXPt,DGHNGtePtTASF the JIeTtePt XP RXSSXTZ CIeOeIXHL CIXeOZTG. DGSXLe hXPHNGteZUNITIF, hXP eZXGeGHe XP ODe ZNPt eZUhTtXHTSSF tN RhTt he OXO.XGOeeO, tRN ZNIe OXPPXZXSTI ZeG XG T PXGJSe CXeSO HTG PHTIHeSF AeXZTJXGeO. RheIe …

And here we can find the combination “tN”, which most likely, represents the word “to”, and “N” -> “O”. Also, in the text there is the combination “KTSXOXtF oC RoIL.”, and “oC” is most likely the word “of”, which means that “C” -> “F”. Next, I found in the text the combination “thTt”, which means that “T” -> “A”. Also, I observed the words “to Rhat he”, and supposed that “R” -> “W”. The next word analyzed was “wheIe” and I thought that “I” -> “R”. This is how the cryptogram looked at the moment: c = whXSe herAert FarOSeF ZaF Ae the AePt LGowG HrFUtoSoJXPt,DGHoGtePtaASF the JreatePt XP wXSSXaZ freOerXHL frXeOZaG. DGSXLe hXPHoGteZUorarF, hXP eZXGeGHe XP ODe ZoPt eZUhatXHaSSF to what he OXO.XGOeeO, two Zore OXPPXZXSar ZeG XG a PXGJSe fXeSO HaG PHarHeSF AeXZaJXGeO…

The following letters deductions were like this:

“waP” means that “P” -> “S”

“eKer” means “K” -> “V”

“Hareer” means “H” -> “C”

“towarO” “O” -> “D”

“Jreatest” “J” -> “G”

“readF” “F” -> “Y”

“detaXSs” “X” -> “I”, “S” -> “L”

“cryUtologist” “U” -> “P”

“chaGce” “G” -> “N”

“sDperficial” “D” -> “U”

“Aecause” “A” -> “B”

“williaZ” “Z” -> “M”

“unliLe” “L” -> “K”

“eYplodes” “Y” -> “X”

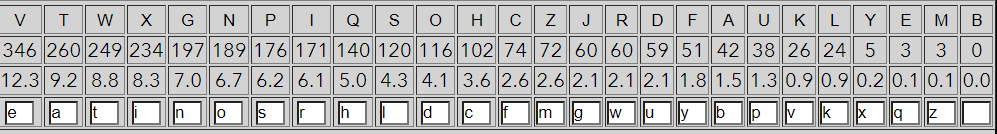
“subsidiMing” “M” -> “Z”

“subEect” “E” -> “J”

Only remaining letter “Q”, “B” -> “Q”.

**Result:**

while herbert yardley may be the best known cryptologist,uncontestably the greatest is william frederick friedman. unlike hiscontemporary, his eminence is due most emphatically to what he did.indeed, two more dissimilar men in a single field can scarcely beimagined. Where yardley was rabelaisian, outgoing, superficial, free andeasy with the details of a good story, and ever ready for the main chance,friedman tended toward introversion, depth of study, personal security,timidity, dedication, and accuracy, nicety, and validity of work. Despitethe relative drabness of these personal traits—or perhaps because ofthem, friedman's theoretical contributions and his practical attainmentsexceed those of any other cryptologist. yardley's career was like anamazing skyrocket that explodes in fantastic patterns against theheavens. friedman's was like the sun.he was born wolfe friedman on september 24, 1891, in kishinev,russia, the oldest son and second child of frederick and rosa friedman.his father, a rumanian who spoke eight languages and worked as aninterpreter for the russian post office, emigrated to america in 1892, at which time his son's name was changed to william. the family settledin pittsburgh, where his father managed a sewing machine agency.william graduated in 1909 as one of the ten honor students in a class of300 at pittsburgh central high school; he then went to work as chiefclerk in the erie city iron works, a firm that sold steam engines. aboutthat time the back-to-the-farm movement called to city boys, and in thefall of 1910, friedman and three friends enrolled in michiganagricultural college, whose chief attraction was that it was tuition-free.but friedman soon discovered that farming held little interest for him.he was an inventive young fellow who liked to fix things and had writtensome science fiction for his high-school paper; he was rapidly coming tothe conclusion that he liked science. at the end of the term he learnedthat tuition was also free in a scientific field allied to agriculture—genetics—at one of the ivy league universities, cornell. He borrowedtrain fare and arrived in ithaca, new york, in february, 1911, where hegot a job waiting on tables. after commencement in february of 1914, heattended graduate school, managing to fall in love twice, once with abrunette, once with the blonde daughter of a movie-house owner. whilehe was there, a wealthy textile merchant, george fabyan, whomaintained laboratories in acoustics, chemistry, genetics, and cryp-tology (to try to prove that bacon wrote shakespeare's plays) on his 500-acre estate, riverbank, at geneva, illinois, decided that he needed ageneticist to improve the grains and livestock on his farm. he applied tocornell for a "would-be-er," not an "as-is-er," and hired friedman, tobegin june 1, 1915.fabyan was a man of no formal education but of intelligence andenergy. he had a great desire to be "somebody," and that desiremotivated his subsidizing the baconian studies: proof of thisrevolutionary thesis would cover its patron as well as its actualdiscoverers with glory. he himself read little, but he absorbed enoughfrom those around him to make his talk on almost any subject soundimpressive—at least superficially. he was autocratic, never allowing hisstaff to disagree with him, but otherwise not unpleasant so long asemployees recognized that he was boss. a cardinal article of faith withhim was that a well-executed sales campaign could put across almostanything.

****

*Figure 3. Resulting decoded cypher*

**Conclusion:**

In conclusion, frequency analysis stands as a venerable yet powerful technique in the realm of cryptanalysis, providing a key tool for deciphering encrypted messages. Through the meticulous examination of letter frequencies within an encrypted text, this method enables us to unravel the mysteries concealed by substitution ciphers and gain access to the hidden information therein.

The successful decoding of our encrypted message using frequency analysis underscores the reliability of this technique when applied to messages encoded with simple substitution ciphers. By identifying recurring patterns and leveraging knowledge of typical letter frequency distributions in the English language, we were able to make educated guesses and progressively piece together the plaintext. However, it is important to note that the effectiveness of frequency analysis diminishes when confronted with more complex encryption methods designed to thwart predictable patterns.