UNIVERSITATEA TEHNICĂ A MOLDOVEI

MINISTRY OF EDUCATION AND RESEARCH OF REPUBLIC OF MOLDOVA

TECHNICAL UNIVERSITY OF MOLDOVA

FACULTY OF COMPUTERS, INFORMATICS AND MICROELECTRONICS

DEPARTMENT OF SOFTWARE ENGINEERING AND AUTOMATICS

ALEXANDRA BUJOR-COBILI, FAF-232

Report

Laboratory work n.2

of Cryptography and Security

Checked by: Zaica Maia, university assistant
FCIM, UTM,

Chișinău, 2025

Content

1 Theory

Monoalphabetic ciphers replace each letter of the plaintext with a fixed different letter from the alphabet. Their weakness is that the frequency of letters in the ciphertext still reflects that of the original language. This makes them vulnerable to a frequency analysis attack, where the analyst counts how often each letter appears in the ciphertext and compares these frequencies with the normal frequencies of letters in that language.

For example, in English, the letter E is the most common, followed by T and A. If the ciphertext's most frequent letter is P, we can assume P $\rightarrow$ E and test if this mapping makes sense in context. Analysts continue making substitutions based on digrams (like TH, HE, IN) and trigrams (THE, AND, ION) until meaningful words appear.

This method, first described by Al-Kindi around 850 AD, demonstrates how even simple frequency patterns can break substitution ciphers. The process often involves both computation and human reasoning since recognizing patterns and correcting guesses require intuition.

2 Objetives

The goal of this lab work is to learn how to break monoalphabetic substitution ciphers using frequency analysis.

The next text needs to be decrypted, a coded English message by comparing letter frequencies in the encrypted text with how often letters usually appear in English.

Variant 4:

ODIXGJ TSS WQVPV FVTIP, HIFUWNSNJF RTP THBDXIXGJ T WTXGW WQTW SXGJVIPVKVG WNOTF—WQV HNGKXHWXNG XG WQV ZXGOP NC ZTGF UVNUSV WQTW HIFUWNSNJFXP T ASTHL TIW, T CNIZ NC NHHDSWXPZ RQNPV UITHWXWXNGVI ZDPW, XG RXSSXTZ C.CIXVOZTG'P TUW UQITPV, "UVICNIHV HNZZDGV OTXSF RXWQ OTIL PUXIXWP WNTHHNZUSXPQ QXP CVTWP NC ZVGWTS EXD-EXWPD."XG UTIW XW XP T LXGO NC JDXSW AF TPPNHXTWXNG. CINZ WQV VTISF OTFP NC XWPVYXPWVGHV, HIFUWNSNJF QTO PVIKVO WN NAPHDIV HIXWXHTS UNIWXNGP NC RIXWXGJPOVT-SXGJ RXWQ WQV UNWVGW PDAEVHW NC ZTJXH—OXKXGTWXNGP, PUVSSP, HDIPVP,RQT HNGCVIIVO PDUVIGTWDITS UNRVIP NG XWP PNIHVIVIP. TGNWQVIXZUNIWTGW CTH-WNI RTP WQV HNGCDPXNG NC HIFUWNSNJF RXWQ WQV EVRXPQLTAATSTQ.ADW,

XZUNIWTGW TP TSS WQVPV RVIV, WQV KXVR WQTW HIFUWNSNJF XP ASTHLZTJXH XG XWPVSC PUIXGJP DSWXZTWVSF CINZ T PDUVICXHXTS IVPVZASTGHV AVWRVVGHI-FUWNSNJF TGO OXKXGTWXNG. VYWITHWXGJ TG XGWVSSXJXASV ZVPPTJV CINZHX-UQVIWVYW PVVZVO WN AV VYTHWSF WQV PTZV WQXGJ TP NAWTXGXGJ LGNRSVO-JVAF VYTZXGXGJ WQV CSXJQW NC AXIOP, WQV SNHTWXNG NC PWTIP TGO UST-GVWP, WQVSVGJWQ TGO XGWVIPVHWXNGP NC SXGVP XG WQV QTGO, WQV VG-WITXSP NC PQVVU, WQVUNPXWXNG NC OIVJP XG T WVTHDU. XG TSS NC WQVPV, WQV RXMTIO-SXLV NUVITWNIOITRP PVGPV CINZ JINWVPBDV, DGCTZXSXTI, TGO TUUTIVGWSF ZVTGXGJSVPPPXJGP. QV ZTLVP LGNRG WQV DGLGNRG.TSS WQXP PWTXGV HIFUWNSNJF PN OVVUSF RXWQ WQV OTIL QDVP NC VPNWVIXPZWQTW PNZV NC WQVZ PWXSS UVIPXPW, GNWXHVTASF HNSNIXGJ WQV UDASXH XZTJV NCHIFUWN-SNJF. UVNUSV PWXSS WQXGL HIFUWTGTSFPXP ZFPWVIXNDP. ANNL OVTSVIP PWXSSSXP HIFUWNSNJF DGOVI ”NHHDSW.” TGO XG 1940 WQV DGXWVO PWTWVP HNGCVIIVO-DUNG XWP ETUTGVPV OXUSNZTWXH HIFUWTGTSFPVP WQV HNOVGTZV ZTJXH. XG GNGV NC WQV PVHIVW RIXWXGJ WQDP CTI RTP WQVIV TGF PDPWTXGVOHIFUWT-GTSFPXP. NHHTPXNGTS HTPVP, FVP. ADW NC TGF PHXVGHV NC HIFUWTGTSFPXP,WQVIV RTP GNWQXGJ. NGSF HIFUWNJITUQF VYXPWVO. TGO WQVIVCNIV HIFUWNSNJF,RQXHQ XGKNSKVP ANWQ HIFUWNJITUQF TGO HIFUWTGTSFPXP, QTO GNW FVW HNZV XGWN AVXGJ PNCTI TP TSS WQVPV HDSWDIVP—XGHSDOXGJ WQV RVPWVIG —RVIV HNGHVIGV RTP ANIG TZNGJ WQV TITAP. WQVF RVIV WQV CXIPW WN OXPHNKVITGO RIXWV ONRG WQV ZVWQNOP NC HIFUWTGTSFPXP. WQV UVNUSV WQTW VYUSNOVONDW NC TITAXT XG WQV 600P TGO CSTZVO NKVI KTPW TIVTP NC WQV LGNRG RNISO-PRXCWSF VGJVGOVIVO NGV NC WQV QXJQVPW HXKXSXMTWXNGP WQTW QXP-WNIF -QTO FVWPVVG. PHXVGHV CSNRVIVO. TITA ZVOXHXGV TGO ZTWQVZTWXHP AVHTZV WQV AVPWXG WQV RNISO—CINZ WQV STWWVI, XG CTHW, HNZVP WQV RNIO ”HXUQVI.” UITHWXHTSTIWP CSNDIXPQVO. TOZXGXPWITWXKV WVHQGXB-DVP OVKVSNUVO. WQV VYDAVITGWHIVTWXKV VGVIJXVP NC PDHQ T HDSWDIV, VYHSDOVO AF XWP IVSXJXNG CINZ UTXGWXGJNI PHDSUWDIV, TGO XGPUXIVO AF XW WN TG VYUSXHTWXNG NC WQV QNSF LNITG,UNDIVO XGWN SXWVITIF UDIPDXWP. PWNIFWVSSXGJ, VYVZUSXCXVO AF PQVQVITMTOV’PWQNDPTGO TGO NGV GXJQWP, RNIO-IXOOSVP, IVADPVP, UDGP, TGTJITZP, TGOPXZXSTI JTZVP TANDGOVO; JITZZTI AVHTZV T ZTENI PWDOF. TGO XGHSDOVORTP PVHIVW RIXWXGJ.TCWVI VYUSTXGXGJ WQTW NGV ZTF RIXWV XG TG DGLGNRG STGJDTJV WN NAWTXGPVHIVHF, XAG TO-ODITXQXZ, THHNIOXGJ WN BTSBTPQTGOX, JTKV PVKVG PFPWVZPNC HXUQVIP. WQXP

SXPW VGHNZUTPPVO, CNI WQV CXIPW WXZV XG HIFUWNJITUQF, ANWQWITGPUNPXWXNG TGO PDAPWXWDWXNG HXUQVIP. ZNIVNKVI, NGV PFPWVZ XP WQV CXIPWLGNRG HXUQVI VKVI WN UINKXOV ZNIV WQTG NGV PDAPWXWDWV CNI T USTXGWVYWSVWWVI. IVZTILTASV TGO XZUNIWTGW TP WQXP XP, QNRVKVI, XW XP NKVIPQTONRVO AF RQTW CNSSNRP— WQV CXIPWVYUNPXWXNG NG HIFUWTGTSFPXP XG QXPWNIF.

3 Implementation

I started my decryption process by pasting the encrypted text into crypto.interactive-maps.com. Using this tool, I analyzed the frequency of each letter in the ciphertext and compared them to typical English language letter frequencies. The most frequent letter in the cipher was replaced with ”e,” since ”e” is the most common letter in English.

1: V = e

The frequencies of the English language are:

E	T	A	O	I	N	S	H	R	D	L	C	U	M	W	F	G	Y	P	B	V	K	J	X	Q	Z
12.7	9.1	8.2	7.5	7.0	6.7	6.3	6.1	6.0	4.3	4.0	2.8	2.8	2.4	2.4	2.2	2.0	2.0	1.9	1.5	1.0	0.8	0.15	0.15	0.10	0.07

The frequencies of the intercept are:

V	W	X	T	P	N	G	I	S	Q	H	O	U	F	Z	C	D	J	R	A	K	L	Y	E	B	M
299	258	219	216	202	196	191	177	120	114	105	92	82	78	66	63	60	60	46	41	23	18	14	6	5	3
10.9	9.4	8.0	7.8	7.3	7.1	6.9	6.4	4.4	4.1	3.8	3.3	3.0	2.8	2.4	2.3	2.2	2.2	1.7	1.5	0.8	0.7	0.5	0.2	0.2	0.1

Figure 3.1 - Letter Frequency

After this first replacement, I looked at the text but did not see anything recognizable yet. I decided to assume the next most frequent letter was ”t” and replaced it accordingly. After doing this and looking more carefully at the patterns, I noticed lots of repetitions of ”tQe” throughout the text. This made me think that ”Q” must be ”h,” which would give me the word ”the.”

V = e

1: W = t

2: Q = h

Looking at words around ”the,” I saw patterns like ”thePe,” ”thTt,” ”T,” and ”tN.” Based on these, I figured that ”T” is actually ”a” and ”N” is definitely ”o.” I also thought ”P” might be ”s” or ”r”, so I replaced it with ”s” to see what would happen.

V = e

W = t

Q = h

1: T = a

2: N = o

3: P = s

After making these changes, I could see the word ”as” appearing a lot, which confirmed I got the letter ”s” correct. I also noticed words like ”aSS,” ”Ras,” ”toOaF,” ”Xt,” and ”Xs.” From these patterns, I worked out that ”SS” equals ”ll” and ”R” equals ”h.” But when I looked closer at the frequency differences, I realized ”R” actually can not be ”h.” Instead, it is more likely ”R” equals ”w.” I also figured out ”O” is ”d,” ”F” is ”y,” and ”X” is ”i.”

V = e

W = t

Q = h

T = a

N = o

P = s

1: S = l

2: R = w

3: O = d

4: F = y

5: X = i

```
dDIiGJ all these yeaIs, HIyUtoloJy was aHBDiIiGJ a taiGt that
liGJeIseKeG today—the
HoGKiHtioG iG the ZiGds oC ZaGy UeoUle that HIyUtoloJyis a AlaHL
aIt, a CoIZ oC oHHDltisZ whose
UIaHtitioGeI ZDst, iG williaZ C.CIiedZaG's aUt UhIase, "UeICoIHe
HoZZDGe daily with daIL sUiIits
toaHHoZUlish his Ceats oC ZeGtal EiD-EitsD."iG UaIt it is a LiGd
oC JDilt Ay assoHiatioG.
CIoZ the eaIly days oC itseYisteGHe, HIyUtoloJy had seIKed to
oAsHDIe HIitiHal UoItioGs oC
```

Figure 3.2 - Frequent Phrase

Now I had words like ”yeaIs,” which clearly meant ”I” equals ”r.” I also saw the phrase ”ZiGDs oC ZaGy,” which was obviously ”minds of many.” This told me that ”Z” is ”m,” ”G” is ”n”, and ”C” is ”f”.

V = e

W = t

Q = h

T = a

N = o

P = s

S = l

R = w

O = d

F = y

X = i

1: I = r

2: Z = m

3: G = n

4: C = f

After these substitutions, the text was much clearer. However, I noticed some words were not quite right. When I went back to check, I had placed a double letter ”d” somewhere. After looking more carefully, I realized ”D” can not be ”d” at all, it is actually ”u.” I could tell from words like ”resemAlanHe”, ”ddrinJ”, ”Eid-Eitsd”, and ”intelliJiAle”, which meant that ”A” is ”b,” ”H” is ”c”, ”E” is ”J”, and ”J” is ”g”.

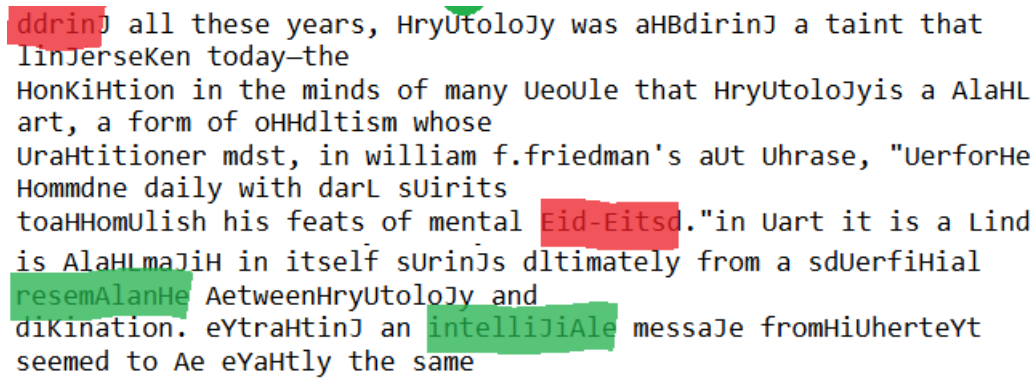

Figure 3.3 - Frequent Phrase 2

V = e

W = t

Q = h

T = a

N = o

P = s

S = l

R = w

O = d

F = y

X = i

I = r

Z = m

G = n

C = f

1: D = u

2: A = b

3: H = c

4: E = j

5: J = g

So now I had six letters left to figure out. Looking at the phrase ”cryUtology was acBuiring,” I could tell that ”U” equals ”p” and ”B” equals ”q.”

I then saw the word ”conKiction,” which showed me that ”K” is ”v.” Looking at ”itseYistence,” I realized this was actually two words ”its existence” meaning ”Y” equals ”x.”

From the word ”remarLable,” I figured out that ”L” is ”k.” The last letter left was ”m,” which had to be ”z.” The word ”ciKiliMations” confirmed this final substitution.

Final complete substitutions:

V = e

W = t

Q = h

T = a

N = o

P = s

S = l

R = w

O = d

F = y

X = i

I = r

Z = m

G = n

C = f

D = u

A = b

H = c

E = j

J = g

U = p

B = q

K = v

Y = x

L = k

W = z

And the final text is:

DURING ALL THESE YEARS, CRYPTOLOGY WAS ACQUIRING A TAINT THAT LINGERSEVEN TODAY—THE CONVICTION IN THE MINDS OF MANY PEOPLE THAT CRYPTOLOGYIS A BLACK ART, A FORM OF OCCULTISM WHOSE PRACTITIONER MUST, IN WILLIAM F.FRIEDMAN'S APT PHRASE, "PERFORCE COMMUNE DAILY WITH DARK SPIRITS TOACCOMPLISH HIS FEATS OF MENTAL JIU-JITSU."IN PART IT IS A KIND OF GUILT BY ASSOCIATION. FROM THE EARLY DAYS OF ITSEXISTENCE, CRYPTOLOGY HAD SERVED TO OBSCURE CRITICAL PORTIONS OF WRITINGSDEALING WITH THE POTENT SUBJECT OF MAGIC—DIVINATIONS, SPELLS, CURSES,WHATEVER CONFERRED SUPERNATURAL POWERS ON ITS SORCERERS. ANOTHERIMPORTANT FACTOR WAS THE CONFUSION OF CRYPTOLOGY WITH THE JEWISHKABBALAH.BUT, IMPORTANT AS ALL THESE WERE, THE VIEW THAT CRYPTOLOGY IS BLACKMAGIC IN ITSELF SPRINGS ULTIMATELY FROM A SUPERFICIAL RESEMBLANCE BETWEENCRYPTOLOGY AND DIVINATION. EXTRACTING AN INTELLIGIBLE MESSAGE FROMCIPHERTEXT SEEMED TO BE EXACTLY THE SAME THING AS OBTAINING KNOWLEDGEBY EXAMINING THE FLIGHT OF BIRDS, THE LOCATION OF STARS AND PLANETS, THELENGTH AND INTERSECTIONS OF LINES IN THE HAND, THE ENTRAILS OF SHEEP, THEPOSITION OF DREGS IN A TEACUP. IN ALL OF THESE, THE WIZARD-LIKE OPERATORDRAWS SENSE FROM GROTESQUE, UNFAMILIAR, AND APPARENTLY MEANINGLESSSIGNS. HE MAKES KNOWN THE UNKNOWN.ALL THIS STAINED CRYPTOLOGY SO DEEPLY WITH THE DARK HUES OF ESOTERISMTHAT SOME OF THEM STILL PERSIST, NOTICEABLY COLORING THE PUBLIC IMAGE OFCRYPTOLOGY. PEOPLE STILL THINK CRYPTANALYSIS MYSTERIOUS. BOOK DEALERS STILLLIST CRYPTOLOGY UNDER "OCCULT." AND IN 1940 THE UNITED STATES CONFERREDUPON ITS JAPANESE DIPLOMATIC CRYPTANALYSES THE CODENAME MAGIC. IN NONE OF THE SECRET WRITING THUS FAR WAS THERE ANY SUSTAINEDCRYPTANALYSIS. OCCASIONAL CASES, YES. BUT OF ANY SCIENCE OF CRYPTANAL-

YSIS,THERE WAS NOTHING. ONLY CRYPTOGRAPHY EXISTED. AND THEREFORE CRYPTOLOGY,WHICH INVOLVES BOTH CRYPTOGRAPHY AND CRYPTANALYSIS, HAD NOT YET COME INTO BEING SOFAR AS ALL THESE CULTURES—INCLUDING THE WESTERN —WERE CONCERNED.CRYPTOLOGY WAS BORN AMONG THE ARABS. THEY WERE THE FIRST TO DISCOVERAND WRITE DOWN THE METHODS OF CRYPTANALYSIS. THE PEOPLE THAT EXPLODEDOUT OF ARABIA IN THE 600S AND FLAMED OVER VAST AREAS OF THE KNOWN WORLDSWIFTLY ENGENDERED ONE OF THE HIGHEST CIVILIZATIONS THAT HISTORY -HAD YETSEEN. SCIENCE FLOWERED. ARAB MEDICINE AND MATHEMATICS BECAME THE BESTIN THE WORLD—FROM THE LATTER, IN FACT, COMES THE WORD "CIPHER." PRACTICALARTS FLOURISHED. ADMINISTRATIVE TECHNIQUES DEVELOPED. THE EXUBERANTCREATIVE ENERGIES OF SUCH A CULTURE, EXCLUDED BY ITS RELIGION FROM PAINTINGOR SCULPTURE, AND INSPIRED BY IT TO AN EXPLICATION OF THE HOLY KORAN,POURED INTO LITERARY PURSUITS. STORYTELLING, EXEMPLIFIED BY SHEHERAZADE'STHOUSAND AND ONE NIGHTS, WORD-RIDDLES, REBUSES, PUNS, ANAGRAMS, ANDSIMILAR GAMES ABOUNDED; GRAMMAR BECAME A MAJOR STUDY. AND INCLUDEDWAS SECRET WRITING.AFTER EXPLAINING THAT ONE MAY WRITE IN AN UNKNOWN LANGUAGE TO OBTAINSECRECY, IBN AD-DURAIHIM, ACCORDING TO QALQASHANDI, GAVE SEVEN SYSTEMSOF CIPHERS. THIS LIST ENCOMPASSED, FOR THE FIRST TIME IN CRYPTOGRAPHY, BOTHTRANSPOSITION AND SUBSTITUTION CIPHERS. MOREOVER, ONE SYSTEM IS THE FIRSTKNOWN CIPHER EVER TO PROVIDE MORE THAN ONE SUBSTITUTE FOR A PLAINTEXTLETTER. REMARKABLE AND IMPORTANT AS THIS IS, HOWEVER, IT IS OVERSHADOWED BY WHAT FOLLOWS— THE FIRSTEXPOSITION ON CRYPTANALYSIS IN HISTORY.

4 Conclusion

During this lab, I managed to decrypt a monoalphabetic cipher using the frequency analysis method. I used the online tool crypto.interactive-maths.com to check letter frequencies and match them with English patterns.

This showed that monoalphabetic ciphers are weak because they keep the same letter frequency as the original text. By matching frequent letters and testing common English words like "the" and "and," I built the full substitution key.

Overall, this lab was fun to do and showed how frequency analysis works, how guessing and checking are part of decryption, and why old ciphers like this aren't secure anymore.

Bibliography

[1] GitHub Repository For This Lab https://github.com/AlexandraB-C/Cryptography_labs.

[2] Cryptography-Analysis. https://crypto.interactive-maths.com/frequency-analysis-breaking-the-code.html.