# MAS433 Cryptography
# Tutorial 1 Classical Ciphers
# Solution

**Problem 1.** Use exhaustive key search to decrypt the following ciphertext, which is encrypted using a shift cipher (hint: the value of the encyption key is less than 7):

FEHJPEWLHVMZIGEYWIHASVWXYWQMPMXEVCFVIEGL

**Solution.** K = 4. The plaintext is:
Bad flash drive caused worst U.S. military breach

**Problem 2.** Suppose that $\pi$ is the following permutation of {1, 2, ..., 8} :

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $\pi(x)$ | 2 | 4 | 6 | 1 | 8 | 3 | 5 | 7 |

**2.1)** Compute the permutation table $\pi^{-1}$ (the inverse of $\pi$).
**Solution.**

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $\pi^{-1}(x)$ | 4 | 1 | 6 | 2 | 7 | 3 | 8 | 5 |

**2.2)** Decrypt the following ciphertext, which is encrypted using a transposition (permutation) cipher with $m = 8$, and with the key $\pi$ given above.

ETEGENLMDNTNEOORDAHATECOESAHLRMI

**Solution.** Gentleman do not read each other's mail

**Problem 3.** The ciphertext given in Appendix A is encrypted using a substitution cipher. The statistical data of the ciphertext is given in Appendix B. Try to break the cipher and decrypt the first line of the ciphertext. (In this exercise, the size of the ciphertext is a bit large so that the attack can

be relatively easy.)

**Solution Outline.** (Given in Appendix A)

Remarks: For a particular message, the distribution of the letters, digrams, trigrams do not match exactly that of English. An explanation is that in a message on a particular topic, some key words would appear frequently.

**Problem 4.** Cipher Composition
**4.1)** Two substitution ciphers, $S_1$ and $S_2$, are applied to encrypt a message as follows: $c_i = S_2(S_1(p_i))$. Discuss how to attack it.

**Solution Outline.** The composition of two substitution tables $S_2 \circ S_1$ is equivalent to one substitution table. The attack on the above cipher is equivalent to the attack on a substitution cipher.

Remarks: Suppose that each substitution table is not that random. The resulting substitution table would be more random than those two substitution tables.

**4.2)** Denote the encryption of a Vigenere cipher as $C = V_K(P)$. Two Vigenere ciphers are applied to encrypt a message as follows: $C = V_{K_2}(V_{K_1}(P))$. Discuss how to attack it. Comparing to the attack on $V_{K_1}$ or $V_{K_2}$, does the attack complexity increase? (Hint: consider the Least Common Multiple of the lengths of $K_1$ and $K_2$) Discuss the security of using more than two Vigenere ciphers.

**Solution Outline.** Suppose that two Vigenere ciphers $V_{K_1}$ and $V_{K_2}$ with key lengths $m_1$ and $m_2$, respectively. Then attacking the composition of $V_{K_2} \circ V_{K_1}$ is (almost) equivalent to attacking a Vegenere cipher with key length $\text{lcm}(m_1, m_2)$, since in a Vigenere cipher, a key is expanded by simply repeat itself.

If we use the composition of $\alpha$ Vignere ciphers, and the key lengths of these ciphers are coprime to each other (for example, each key length is a prime number), then the value of $\text{lcm}(m_1, m_2, m_3, \cdots, m_\alpha)$ would be extremely large for large value of $n$.

For example, let us consider the composition of 100 Vigenere ciphers, and their key lengths are the first 100 prime numbers. Note that the LCM of the first 100 primes is $\prod_{i \in \{\text{the first 100 prime numbers}\}} i \approx 10^{219.7} \approx 2^{729.7}$, thus it is impossible to break the resulting cipher using frequency analysis since we do not have that amount of message in the world.

**Problem 5.** If an attacker knows the ciphertext and part of the plaintext, how to attack the shift cipher, substitution cipher and Vigenere cipher, and

how to break the composition of Vigenere ciphers in Problem 4.2 efficiently?

**Solution Outline.** Breaking shift cipher, substitution cipher and Vigenere cipher with known-plaintext attack is trivial. For a substitution cipher, given 25 different ciphertext letters and their related plaintext letters, the key can be recovered. For a Vigenere cipher with length $m$, given the first plaintext letters, the key can be easily determined.

To break the composition of Vigenere ciphers efficiently, you may consider solving a system of linear equations. Consider the example that two keys $K_1$ and $K_2$ with lengths 2 and 3, respectively. We can solve the following five equations to solve for those five key letters:

$$\begin{aligned}
(m_1 + K_{1,1} + K_{2,1}) \bmod 26 &= c_1 \\
(m_2 + K_{1,2} + K_{2,2}) \bmod 26 &= c_2 \\
(m_3 + K_{1,1} + K_{2,3}) \bmod 26 &= c_3 \\
(m_4 + K_{1,2} + K_{2,1}) \bmod 26 &= c_4 \\
(m_5 + K_{1,1} + K_{2,2}) \bmod 26 &= c_5
\end{aligned}$$

This method is particularly useful when we are dealing with the composition of many Vigenere ciphers. For example, for the composition of 100 Vigenere ciphers with their lengths being the first 100 prime numbers, if we do not recover the key by solving the linear equations, then we need much more than $\prod_{i \in \{\text{the first 100 prime numbers}\}} i \approx 2^{729.7}$ plaintext-ciphertext letters to recover the key using frequency analysis (impractical). If we solve the system of linear equations, we need only $\sum_{i \in \{\text{the first 100 prime numbers}\}} i = 24133 \approx 2^{14.56}$ plaintext-ciphertext letters to recover the key (with the computational cost of solving $2^{14.56}$ linear equations. The complexity to solve a system of $n$ linear equations involving $n$ variables using Gaussian elimination is about $2^{3n}$ operations.)

# **Appendix A.** Problem 3 solution outline

## **Step 1.** Comparing single letter probabilities

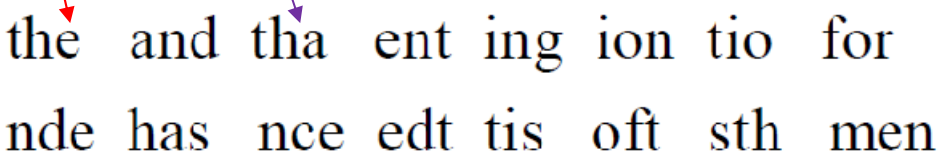
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| A | 0.0983 | B | 0.0098 | C | 0.0289 | D | 0.0120 | E | 0.0257 |
| F | 0.0175 | G | 0.0033 | H | 0.0005 | I | 0.0126 | J | 0.0705 |
| K | 0.0847 | L | 0.0016 | M | 0.0672 | N | 0.0431 | O | 0.0442 |
| P | 0.0011 | Q | 0.0115 | R | 0.0350 | S | 0.0240 | T | 0.0142 |
| U | 0.0699 | V | 0.0147 | W | 0.0584 | X | 0.0464 | Y | 0.1158 |
| Z | 0.0890 | | | | | | | | |

| letter | probability | letter | probability |
|---|---|---|---|
| *A* | .082 | *N* | .067 |
| *B* | .015 | *O* | .075 |
| *C* | .028 | *P* | .019 |
| *D* | .043 | *Q* | .001 |
| *E* | .127 | *R* | .060 |
| *F* | .022 | *S* | .063 |
| *G* | .020 | *T* | .091 |
| *H* | .061 | *U* | .028 |
| *I* | .070 | *V* | .010 |
| *J* | .002 | *W* | .023 |
| *K* | .008 | *X* | .001 |
| *L* | .040 | *Y* | .020 |
| *M* | .024 | *Z* | .001 |

Y→e

**Step 2.** Already known Y→e

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AXY | 25 | AXK | 17 | XKA | 15 | OUR | 12 | ZMD | 12 |
| AAX | 11 | YMA | 11 | MAX | 9 | KAZ | 8 | OUM | 8 |
| URE | 8 | AZU | 7 | UMA | 7 | UWA | 7 | XAX | 7 |
| YKW | 7 | YWJ | 7 | ZJA | 7 | ZUM | 7 | | |




AXY → the ⟹ AX → th

(the frequency of digrams can be used to confirm it)

AXK → tha ⟹ K → a

**Step 3.** Already known Y→e, A→t, X→h, K→a

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AX | 51 | ZM | 34 | XY | 32 | YW | 32 | KM | 30 |
| KA | 29 | AZ | 28 | MA | 27 | XK | 26 | UM | 25 |
| JA | 23 | UW | 22 | WY | 22 | OU | 21 | YJ | 21 |
| KN | 20 | RY | 19 | YA | 19 | YK | 19 | ZJ | 19 |

| | | | | | |
|---|---|---|---|---|---|
| th | 1.52% | en | 0.55% | ng | 0.18% |
| he | 1.28% | ed | 0.53% | of | 0.16% |
| in | 0.94% | to | 0.52% | al | 0.09% |
| er | 0.94% | it | 0.50% | de | 0.09% |
| an | 0.82% | ou | 0.50% | se | 0.08% |
| re | 0.68% | ea | 0.47% | le | 0.08% |
| nd | 0.63% | hi | 0.46% | sa | 0.06% |
| at | 0.59% | is | 0.46% | si | 0.05% |
| on | 0.57% | or | 0.43% | ar | 0.04% |
| nt | 0.56% | ti | 0.34% | ve | 0.04% |
| ha | 0.56% | as | 0.33% | ra | 0.04% |
| es | 0.56% | te | 0.27% | ld | 0.02% |
| st | 0.55% | et | 0.19% | ur | 0.02% |

Analyze the frequency of YW, WY to recover Y

**Step 4a.** Already known Y→e, A→t, X→h, K→a,W→r,

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AX | 51 | ZM | 34 | XY | 32 | YW | 32 | KM | 30 |
| KA | 29 | AZ | 28 | MA | 27 | XK | 26 | UM | 25 |
| JA | 23 | UW | 22 | WY | 22 | OU | 21 | YJ | 21 |
| KN | 20 | RY | 19 | YA | 19 | YK | 19 | ZJ | 19 |
| AU | 18 | KW | 18 | WZ | 17 | JZ | 16 | MU | 16 |
| OX | 16 | ZA | 16 | ZO | 16 | NZ | 15 | QY | 15 |
| UR | 15 | YO | 15 | AY | 14 | KJ | 14 | YM | 14 |
| AA | 13 | JK | 13 | JY | 13 | MD | 13 | MZ | 13 |
| OY | 13 | WJ | 13 | YC | 13 | AW | 12 | NU | 12 |
| SJ | 12 | WA | 12 | AJ | 11 | MC | 11 | NY | 11 |
| OK | 11 | RK | 11 | XZ | 11 | ZN | 11 | ZY | 11 |
| FY | 10 | NN | 10 | RE | 10 | WK | 10 | YZ | 10 |
| ZU | 10 | EY | 9 | JJ | 9 | JR | 9 | UI | 9 |
| ZK | 9 | AK | 8 | BY | 8 | CZ | 8 | JU | 8 |
| MK | 8 | MO | 8 | OZ | 8 | SA | 8 | UN | 8 |

| | | | | | |
|---|---|---|---|---|---|
| th | 1.52% | en | 0.55% | ng | 0.18% |
| he | 1.28% | ed | 0.53% | of | 0.16% |
| in | 0.94% | to | 0.52% | al | 0.09% |
| er | 0.94% | it | 0.50% | de | 0.09% |
| an | 0.82% | ou | 0.50% | se | 0.08% |
| re | 0.68% | ea | 0.47% | le | 0.08% |
| nd | 0.63% | hi | 0.46% | sa | 0.06% |
| at | 0.59% | is | 0.46% | si | 0.05% |
| on | 0.57% | or | 0.43% | ar | 0.04% |
| nt | 0.56% | ti | 0.34% | ve | 0.04% |
| ha | 0.56% | as | 0.33% | ra | 0.04% |
| es | 0.56% | te | 0.27% | ld | 0.02% |
| st | 0.55% | et | 0.19% | ur | 0.02% |

Analyze KM, MK to recover M as n or s

**Step 4b.** Already known

Y→e, A→t, X→h, K→a,W→r, M→(n,s)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AXY | 25 | AXK | 17 | XKA | 15 | OUR | 12 | ZMD | 12 |
| AAX | 11 | YMA | 11 | MAX | 9 | KAZ | 8 | OUM | 8 |
| URE | 8 | AZU | 7 | UMA | 7 | UWA | 7 | XAX | 7 |
| YKW | 7 | YWJ | 7 | ZJA | 7 | ZUM | 7 | | |

the and tha ent ing ion tio for

nde has nce edt tis oft sth men

n appears in and, ent, ing

s does not appear in the frequent trigrams

⇨ M→n

**Step 5.** Already known

Y→e, A→t, X→h,K→a, W→r, M→n

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AX | 51 | ZM | 34 | XY | 32 | YW | 32 | KM | 30 |
| KA | 29 | AZ | 28 | MA | 27 | XK | 26 | UM | 25 |
| JA | 23 | UW | 22 | WY | 22 | OU | 21 | YJ | 21 |
| KN | 20 | RY | 19 | YA | 19 | YK | 19 | ZJ | 19 |
| AU | 18 | KW | 18 | WZ | 17 | JZ | 16 | MU | 16 |
| OX | 16 | ZA | 16 | ZO | 16 | NZ | 15 | QY | 15 |

| | | |
|---|---|---|
| th 1.52% | en 0.55% | ng 0.18% |
| he 1.28% | ed 0.53% | of 0.16% |
| in 0.94% | to 0.52% | al 0.09% |
| er 0.94% | it 0.50% | de 0.09% |
| an 0.82% | ou 0.50% | se 0.08% |
| re 0.68% | ea 0.47% | le 0.08% |
| nd 0.63% | hi 0.46% | sa 0.06% |
| at 0.59% | is 0.46% | si 0.05% |
| on 0.57% | or 0.43% | ar 0.04% |
| nt 0.56% | ti 0.34% | ve 0.04% |
| ha 0.56% | as 0.33% | ra 0.04% |
| es 0.56% | te 0.27% | ld 0.02% |
| st 0.55% | et 0.19% | ur 0.02% |

Analyze AZ and ZA to recover Z

(notice the frequency of Z)

**Step 6.** Already known

Y→e, A→t, X→h, K→a, W→r, M→n, Z→i (?)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AX | 51 | ZM | 34 | XY | 32 | YW | 32 | KM | 30 |
| KA | 29 | AZ | 28 | MA | 27 | XK | 26 | UM | 25 |
| JA | 23 | UW | 22 | WY | 22 | OU | 21 | YJ | 21 |
| KN | 20 | RY | 19 | YA | 19 | YK | 19 | ZJ | 19 |
| AU | 18 | KW | 18 | WZ | 17 | JZ | 16 | MU | 16 |

| | | |
|---|---|---|
| th 1.52% | en 0.55% | ng 0.18% |
| he 1.28% | ed 0.53% | of 0.16% |
| in 0.94% | to 0.52% | al 0.09% |
| er 0.94% | it 0.50% | de 0.09% |
| an 0.82% | ou 0.50% | se 0.08% |
| re 0.68% | ea 0.47% | le 0.08% |
| nd 0.63% | hi 0.46% | sa 0.06% |
| at 0.59% | is 0.46% | si 0.05% |
| on 0.57% | or 0.43% | ar 0.04% |
| nt 0.56% | ti 0.34% | ve 0.04% |
| ha 0.56% | as 0.33% | ra 0.04% |
| es 0.56% | te 0.27% | ld 0.02% |
| st 0.55% | et 0.19% | ur 0.02% |

Analyze UM & MU => M may be decrypted as ‘o’

Comparing the frequency of ‘M’ and ‘o’, the chance is high that M → o

**Step 7.** Already known

Y→e, A→t, X→h, K→a, W→r, M→n, Z→i (?),U→o

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AX | 51 | ZM | 34 | XY | 32 | YW | 32 | KM | 30 |
| KA | 29 | AZ | 28 | MA | 27 | XK | 26 | UM | 25 |
| JA | 23 | UW | 22 | WY | 22 | OU | 21 | YJ | 21 |
| KN | 20 | RY | 19 | YA | 19 | YK | 19 | ZJ | 19 |
| AU | 18 | KW | 18 | WZ | 17 | JZ | 16 | MU | 16 |
| OX | 16 | ZA | 16 | ZO | 16 | NZ | 15 | QY | 15 |
| UR | 15 | YO | 15 | AY | 14 | KJ | 14 | YM | 14 |
| AA | 13 | JK | 13 | JY | 13 | MD | 13 | MZ | 13 |
| OY | 13 | WJ | 13 | YC | 13 | AW | 12 | NU | 12 |
| SJ | 12 | WA | 12 | AJ | 11 | MC | 11 | NY | 11 |

| | | |
|---|---|---|
| th 1.52% | en 0.55% | ng 0.18% |
| he 1.28% | ed 0.53% | of 0.16% |
| in 0.94% | to 0.52% | al 0.09% |
| er 0.94% | it 0.50% | de 0.09% |
| an 0.82% | ou 0.50% | se 0.08% |
| re 0.68% | ea 0.47% | le 0.08% |
| nd 0.63% | hi 0.46% | sa 0.06% |
| at 0.59% | is 0.46% | si 0.05% |
| on 0.57% | or 0.43% | ar 0.04% |
| nt 0.56% | ti 0.34% | ve 0.04% |
| ha 0.56% | as 0.33% | ra 0.04% |
| es 0.56% | te 0.27% | ld 0.02% |
| st 0.55% | et 0.19% | ur 0.02% |

Analyze JA and AJ, we know that J may be decrypted as ‘s’. After checking the frequency of ‘J’ and ‘s’, there is still uncertainty.

**Step 8.** Already known Y→e, A→t, X→h, K→a, W→r,M→n, Z→i (?),U→o, J→s (?) .

We now look at the ciphertext tofind out more information:

```
s-ientistsatri-e-ni-
JOZYMAZJAJKAWZOYSMZQ

ersit-an-he--ett-a--

YWJZAVKMCXYTNYAAEKOB

ar-arere-ortin-this-

KWCKWYWYEUWAZMDAXZJT
```

So we can guess that (O→c,S→u,Q→v,V→y)

**Step 9.** Already known Y→e, A→t, X→h, W→r, K→a, M→n, Z→i (?),U→o, J→s(?),O→c, S→u, Q→v,V→y

We decrypt more ciphertext:

```
scientistsatriceuniv
JOZYMAZJAJKAWZOYSMZQ

ersityan-he--ett-ac-

YWJZAVKMCXYTNYAAEKOB

ar-arere-ortin-this-

KWCKWYWYEUWAZMDAXZJT

ee-thattheycanoverco

YYBAXKAAXYVOKMUQYWOU

-ea-un-a-enta--arrie

RYKISMCKRYMAKNFKWWZY

rtothecontin-e-ra-i-

WAUAXYOUMAZMSYCWKEZC
```

(R→m,F→b,S→u)

**Step 10.** Already known Y→e, A→t, X→h, W→r, K→a, M→n, Z→i (?),U→o, J→s (?), O→c, S→u, Q→v, V→y, R→m,F→b, S→u

scientistsatriceuniv
JOZYMAZJAJKAWZOYSMZQ

ersityan-he--ett-ac-

YWJZAVKMCXYTNYAAEKOB

ar-arere-ortin-this-

KWCKWYWYEUWAZMDAXZJT

ee-thattheycanoverco

YYBAXKAAXYVOKMUQYWOU

mea-un-amenta-barrie

RYKISMCKRYMAKNFKWWZY

rtothecontinue-ra-i-

WAUAXYOUMAZMSYCWKEZC

miniaturi-ationo-com

RZMZKASWZLKAZUMUIOUR

-utermemorythathasbe

ESAYWRYRUWVAXKAXKJFY

enthebasis-orthecons

YMAXYFKJZJIUWAXYOUMJ

(L→z,E→p,I→f)

**Step 11.** Already known

Y→e, A→t, X→h, W→r,K→a, M→n, Z→i (?),U→o

J→s (?), O→c, S→u, Q→v, V→y, R→m,F→b, S→u, L→z,E→p,I→f

scientistsatriceuniv
JOZYMAZJAJKAWZOYSMZQ

ersityan-he--ettpac-

YWJZAVKMCXYTNYAAEKOB

ar-arereportin-this-

KWCKWYWYEUWAZMDAXZJT

ee-thattheycanoverco

YYBAXKAAXYVOKMUQYWOU

meafun-amenta-barrie

RYKISMCKRYMAKNFKWWZY

rtothecontinue-rapi-

WAUAXYOUMAZMSYCWKEZC

miniaturizationofcom

RZMZKASWZLKAZUMUIOUR

putermemorythathasbe

ESAYWRYRUWVAXKAXKJFY

enthebasisforthecons

YMAXYFKJZJIUWAXYOUMJ

(D→g,C→d,N→l)

**Step 12.** Already known Y→e, A→t, X→h, W→r,

K→a, M→n, Z→i (?),U→o, J→s (?), O→c, S→u, Q→v, V→y, R→m,F→b, S→u, L→z,E→p,I→f, D→g,C→d,N→l

scientistsatriceuniv
JOZYMAZJAJKAWZOYSMZQ

ersityandhe-lettpac-

YWJZAVKMCXYTNYAAEKOB

ardarereportingthis-

KWCKWYWYEUWAZMDAXZJT

ee-thattheycanoverco

YYBAXKAAXYVOKMUQYWOU

meafundamentalbarrie

RYKISMCKRYMAKNFKWWZY

rtothecontinuedrapid

WAUAXYOUMAZMSYCWKEZC

miniaturizationofcom

RZMZKASWZLKAZUMUIOUR

putermemorythathasbe

ESAYWRYRUWVAXKAXKJFY

enthebasisforthecons

YMAXYFKJZJIUWAXYOUMJ

T→? B→?

Scientists at Rice University and Hewlett-Packard are reporting this week that they can overcome a fundamental barrier to the continued rapid miniaturization of computer memory that has been the basis for the consumer electronics revolution. In recent years the limits of physics and finance faced by chip makers had loomed so large that experts feared a slowdown in the pace of miniaturization that would act like a brake on the ability to pack ever more power into ever smaller devices like laptops, smartphones and digital cameras. But the new announcements, along with competing technologies being pursued by companies like IBM and Intel, offer hope that the brake will not be applied any time soon. In one of the two new developments, Rice researchers are reporting in Nano Letters, a journal of the American Chemical Society, that they have succeeded in building reliable small digital switches — an essential part of computer memory — that could shrink to a significantly smaller scale than is possible using conventional methods. More important, the advance is based on silicon oxide, one of the basic building blocks of today's chip industry, thus easing a move toward commercialization. The scientists said that PrivaTran, a Texas startup company, has made experimental chips using the technique that can store and retrieve information."This is something that I.B.M. studied before and which is still in the research stage," said Charles Lam, an I.B.M. specialist in semiconductor memories. H.P. has for several years been making claims that its memristor technology can compete with traditional transistors, but the company will report this week that it is now more confident that its technology can compete commercially in the future. In contrast, the Rice advance must still be proved. Acknowledging that researchers must overcome skepticism because silicon oxide has been known as an insulator by the industry until now, Jim Tour, a nanomaterials specialist at Rice said he believed the industry would have to look seriously at the research team's new approach. "It's a hard sell, because at first it's obvious it won't work," he said. "But my hope is that this is so simple they will have to put it in their portfolio to explore."