实验 1-1 古典密码体制的统计分析
——Vigenere 密码

课程名称：密码学基础
课程代码：COMP130188.01
任课教师：李景涛
实验课助教：郑云涛 19210240058@fudan.edu.cn
仲崇鹏 20210240037@fudan.edu.cn

实验目的

1.了解古典密码中的加密和解密运算；
2.了解古典密码体制；
3.掌握古典密码的统计分析方法。

实验原理

Vigenere 密码

提高单字母表密码安全性的思路之一。

加密

以 FUDAN 为关键词，明文为 THEBASICOFCRYPTOGRAPHY（The Basic of Cryptography），举例说明 Vigenere 密码的加密过程：

1.构建密钥

- 密钥与明文等长，循环重复关键词。
- 明文：THEBASICOFCRYPTOGRAPHY
- 密钥：FUDANFUDANFUDANFUDANFU

2.对照字母表编写密文

- 根据密钥字母，在字母表中找到对应行；
- 根据明文字母，在字母表中找到对应列；
- 已知明文：T， 密钥：F
- 得出密文：Y

解密过程与加密过程相反。

若用 0-25 的整数与 A-Z 的 26 个字母一一对应， 为明文， 为密文， 为密钥，那么可以将加密算法写成：

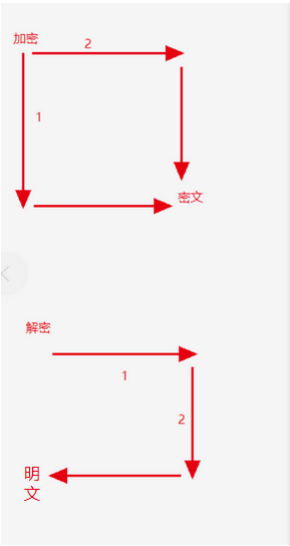
$$C_i \equiv P_i + K_i \pmod{26}$$

解密算法写成：

$$P_i \equiv C_i - K_i \pmod{26}$$

已知明文：T， 密钥：F，那么 T 为 19，F 为 5，密文即为$(19+5)\ mod\ 26 = 24$，对应字母为 Y，因此密文为 Y。与字母表得出的结果一致。 *（提示：如何在数字和字母之前转换*

可以借助 ASCII 码，也可以发挥自己的想法用字典或列表或数组存储相关关系。）



MK	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
a	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
b	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
c	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
d	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
e	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
f	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
g	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
h	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
i	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
j	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
k	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
l	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
m	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
n	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
o	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
p	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
r	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
s	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
t	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
u	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
v	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
w	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
x	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X

实验环境

运行 Windows 操作系统的计算机，具有 C/C++语言编译环境（原则上操作系统和编程语言环境不限）。

实验内容

1. 给定密钥，用 C/C++/Python 实现 Vigenere 密码的加密和解密算法；（提供测试用例帮助同学们检测代码的准确性；原则上使用的编程语言不限，要求工作量相近）

- 加密测试用例（明文为 THEBASICOFCRYPTOGRAPHY，密钥为 SECURITY）

  Input: THEBASICOFCRYPTOGRAPHY SECURITY

  Output: LLGVRABAGJELPXMMYVCJYG
- 解密测试用例（密文为 YBHBNXCFOSHLBPGTAUACMS，密钥为 FUDAN）

  Input: YBHBNXCFOSHLBPGTAUACMS FUDAN

  Output: THEBASICOFCRYPTOGRAPHY

测试代码与结果不必提交。

2. 解密文档 lab1-1_input.txt（提示：密钥为 CRYPTOGRAPHY），输出的结果保存在 lab1-1_output.txt 中，并将其中包含的信息写入报告的实验结果中。

实验提交

- 截止日期：2021 年 4 月 4 日（待定）
- 提交清单（针对实验内容 2）：
  - 实验报告 pdf 格式，文件名格式：学号_姓名_lab1-1；
  - 项目源代码，文件名格式：学号_lab1-1；
  - 可执行程序，文件名格式：姓名_lab1-1；
  - 资源文件，本实验中为 lab1-1_input.txt。
- 提交方式：

- 将提交清单中所有文件打包成一个**压缩文件**（文件名：学号_姓名_lab1-1），在 elearning 上进行提交。

评分标准

源代码可编译运行	√	√	√	√	√	
源代码风格良好	√		√			
程序运行结果正确	√	√	√	√		
实验报告规范清晰	√	√			√	√
最终得分	100	90-99	80-89	60-79	40-59	20-39

注：1. “源代码风格良好”指的是有必要的注释、合适的缩进，变量和函数命名便于理解；
2. 若出现两位同学报告或代码完全一致的情况，则双方本次实验成绩均为 0；
3. 若源码与程序无法正常运行，则成绩不高于 60 分；
4. 其他情况酌情给分。

拓展实验

有兴趣的同学可以完成，非强制性要求，且不计分。

破解 Vigenere 密码

破译 Vigenere 密码虽然不能直接使用频率分析，但由于密钥循环反复，当得知密钥长度时，可利用类似于 Caesar 密码的方法破解。
密钥长度的破解可通过以下两种方法：Kasiski 测试 & Friedman 测试。

Kasiski 测试

原理是常用单词或高频出现的单词片段，可能被同样的密钥字母进行加密。当密文足够长时，包含该信息更多，更有可能推断出密钥长度。例如：

密钥：ABCDABCDABCDABCDABCDABCDABCD
明文：CRYPTOISSHORTFORCRYPTOGRAPHY
密文：CSASTPKVSIQUTGQUCSASTPIUAQJB

相隔 16 个字符出现相同字符片段，密钥有可能是 16 的约数（16，8，4，2）。当密文长度足够长时，还能找到其他的重复片段，取其公约数，即可确定密钥长度。

Friedman 测试

定义重合指数来描述字母在频率分布上的不匀性，从而破译密码。

$$\mathbf{IC} = \frac{\sum_{i=1}^{c} n_i(n_i - 1)}{N(N-1)}$$

其中，c是指字母表的长度（英语中为 26），N指密文文本的长度，n_1到n_c到是指密文的字母频数，为整数。得到重合指数 IC 后，可用以下公式估计密钥长度：

$$L \approx \frac{\kappa_p - \kappa_r}{IC - \kappa_r} = \frac{0.027N}{IC*(N-1) - 0.0385N + 0.0655}$$

其中，L是密钥长度，k_p为目标语言中两个任意字母相同的概率（在英文中k_p约为 0.655），k_r为字母表中出现相同字母的概率（在英文字母表中，k_r=1/26=0.0385）

已知密钥长度后，可按照密钥长度重新改写密文，对于被密钥中同一个位置加密的密文，即可单独做类似于 Caesar 密码的字母频率分析破译，从而推断出密钥中的每个字母。

参考资料

- Wikipedia
  - Vigenère cipher
  - Kasiski examination
  - Index of coincidence
- 百度百科
  - 维吉尼亚密码