密码学原理作业报告

作业 1：古典密码体制的实践与分析

姓名	刘佳亮	院系	软件学院	学号	1123710401
任课教师	刘绍辉		指导教师	刘绍辉	
实验地点			实验时间		

一、实验目的

要求：综述本次实验的基本目的。

1．掌握重合指数和互重合指数的概念，并用于经典密码分析当中

2．理解密码学算法安全性分析的重要性

二、实验内容

要求：对如下内容进行详细描述。

1． 弗吉尼亚密码加密解密算法；

弗吉尼亚密码是一种恺撒密码的基础上扩展的多表密码，将 26 个凯撒密码表合成一个密码表，根据密钥来决定用哪一行的密表来进行替换，以此来对抗字频统计攻击。

加密时，选择一个关键字并重复得到密钥，将明文中的第一个字母对应密钥的第一个字母，根据密钥，选择某一行密码表，然后根据该行密码学对明文进行加密。

解密方法与加密类似，不过是加密的逆过程，根据密钥确定使用的密码表，再利用密码表和密文反推明文。

2． 重合指数计算算法；

重合指数是一种用来描述密文字母频率的不均匀性的指数，重合指数描述了在给定的文本中随机选取两个字母（或数字），这两个字母相同的概率。

弗里德曼实验使用重合指数来破译维吉尼亚密码，将密钥长度估计为 Kp - Kr / Ko - Kr，其中 Kp 是目标语言中任意字母相同的概率（英语中为 0.067），Kr 指字母表中出现这种情况的概率（英语中为 1/26 = 0.038），Ko 为观测概率，即重合指数的无差估计。

这种方法会随着文本长度的增加而更为准确。

3． Kasiski’s 计算方法；

卡西斯基实验基于类似 the 这样的常用单词有可能被同样的密钥字母进行加密，从而在密文中重复出现。

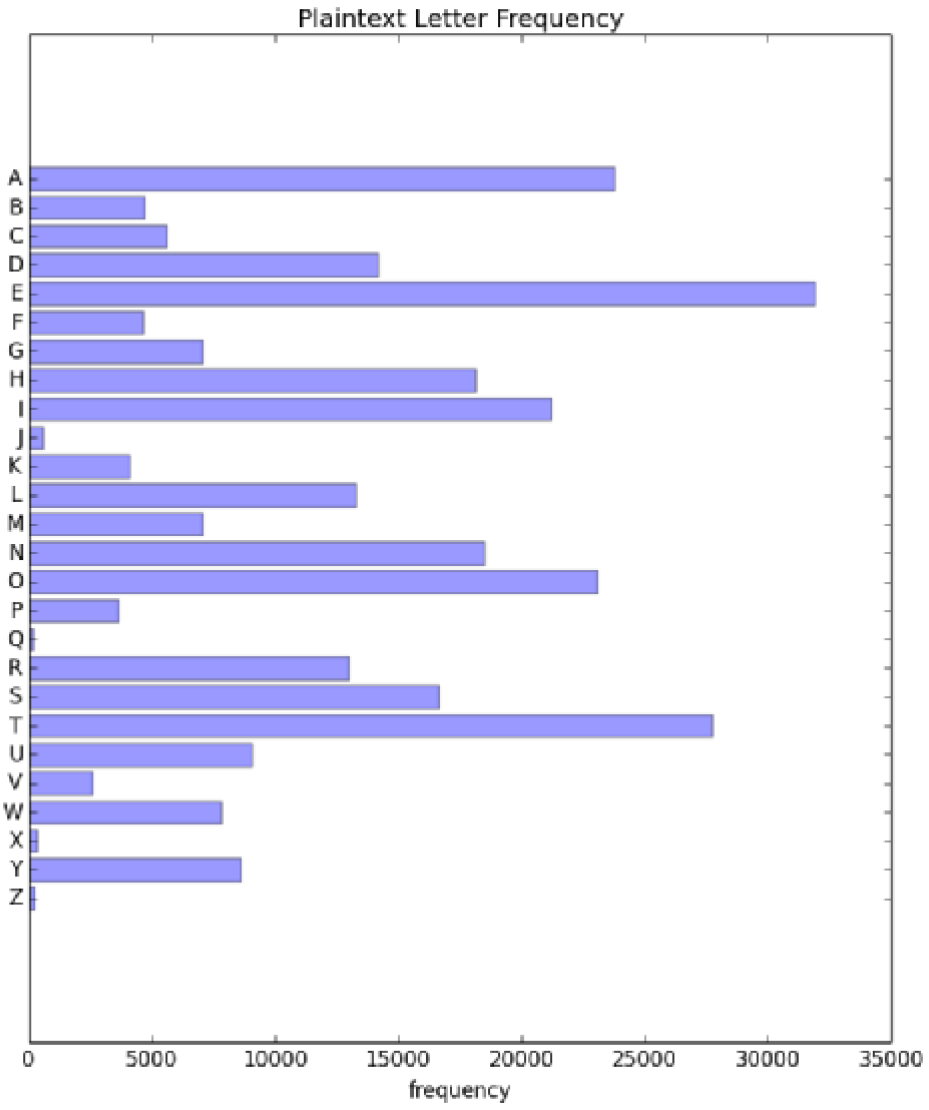
假如密文中某个字母片段重复出现，且出现间隔了 18 个字母，那么密钥的长度就可能是 18 的约数，即 18，9，6，3 和 2.。假如又有另一个字母片段也重复出现了，且出现间隔了 20 个字母，意味着密钥长度应为 20，10，5，4，和 2。两者取交集，基本可以确定密钥长度为 2。

三、实验结果及分析

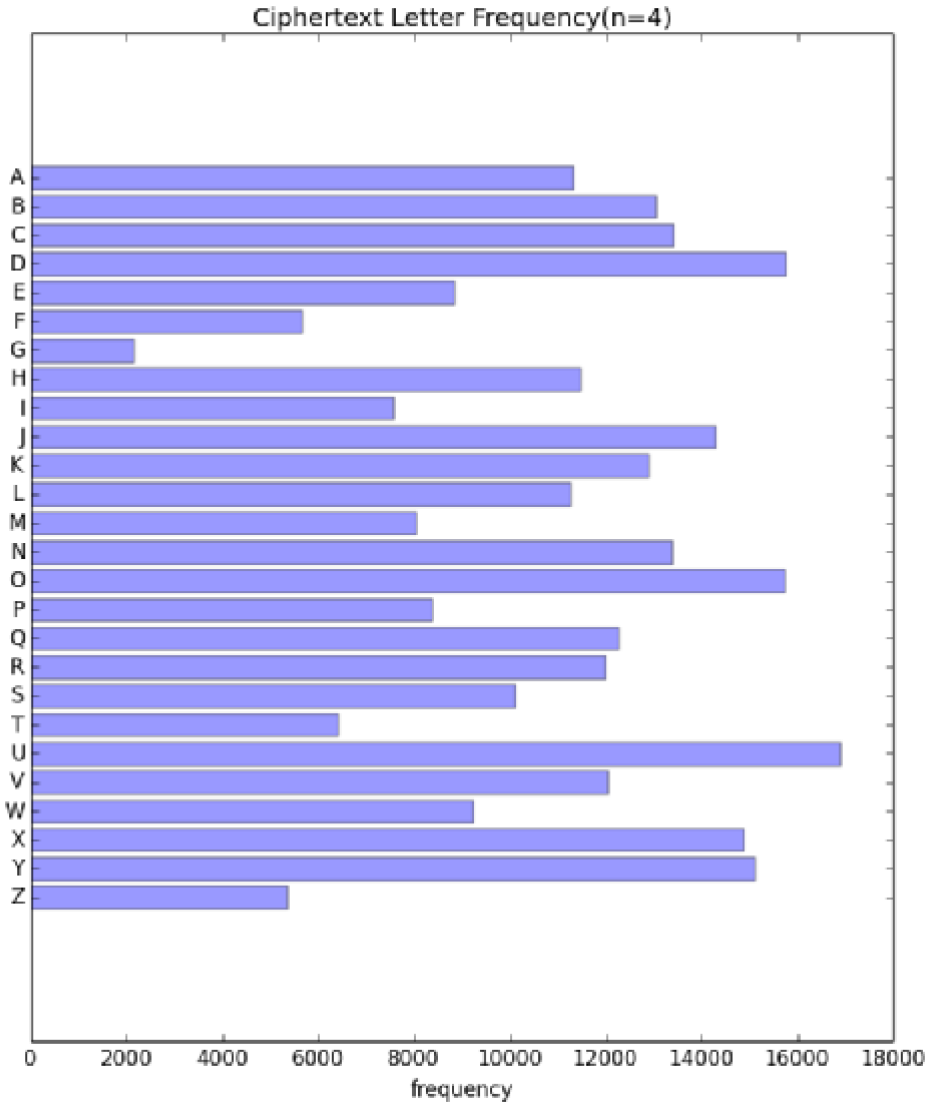
要求：将实验获得的结果进行描述，涉及不同的密钥以及密钥长度，不同密文长度情况下的 Kasiski 分析及重合指数分析得出的结果

1、频率统计

原文的字母频率如下：



当 n=4 时密文的字母频率如下：

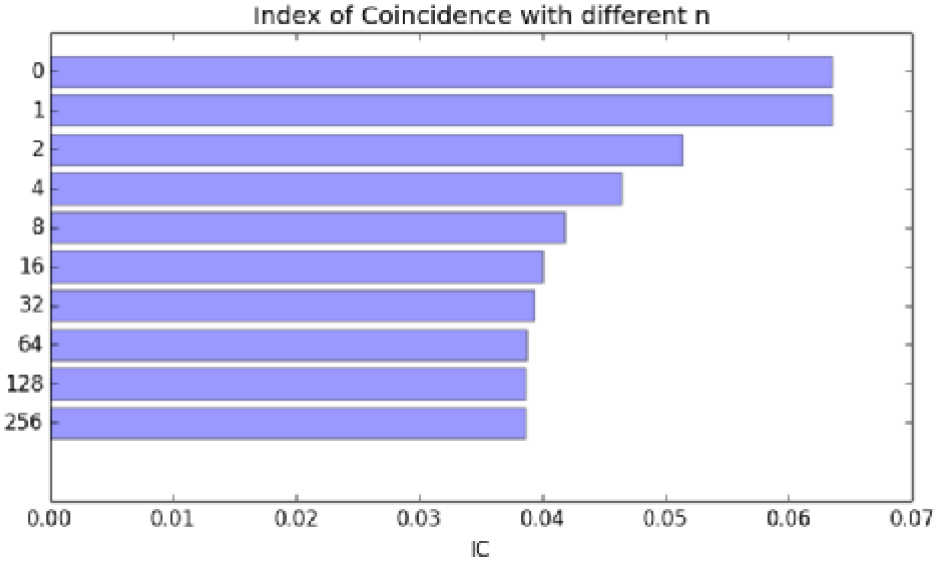


当 n=16 时密文的字母频率如下：

可见随着密钥长度的增加，密文中各个字符出现的频率越接近于随机，也就是所携带的信息越少，加密效果越好。

2、重合指数

当 n 改变时，重合指数的值如下（0 表示原文的重合指数）



可见，随着密钥长度增加，密文的重合指数越接近于 0.038，即完全随机情况下的重合指数，与上面对字母频率的观察结果也是相互吻合的。

3、密码学分析

下面对一段使用“VNEUUTU”作为密钥加密的密文进行分析。

(1) 使用 Kasiski 方法对密文进行分析

使用程序统计出密文中重合文本段的间隔以及其公约数，部分统计结果如下：

NPU 217 [1, 7, 31, 217]
DGL 77 [1, 7, 11, 77]
IFW 154 [1, 2, 7, 11, 14, 22, 77, 154]
NLI 238 [1, 2, 7, 14, 17, 34, 119, 238]
NNA 130 [1, 2, 5, 10, 13, 26, 65, 130]
NAY 413 [1, 7, 59, 413]
XBY 210 [1, 2, 3, 5, 6, 7, 10, 14, 15, 21, 30, 35, 42, 70, 105, 210]
IIA 144 [1, 2, 3, 4, 6, 8, 9, 12, 16, 18, 24, 36, 48, 72, 144]
GLU 112 [1, 2, 4, 7, 8, 14, 16, 28, 56, 112]
EHX 84 [1, 2, 3, 4, 6, 7, 12, 14, 21, 28, 42, 84]
TFG 240 [1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 16, 20, 24, 30, 40, 48, 60, 80, 120, 240]
NAC 80 [1, 2, 4, 5, 8, 10, 16, 20, 40, 80]
GRL 91 [1, 7, 13, 91]
GIB 260 [1, 2, 4, 5, 10, 13, 20, 26, 52, 65, 130, 260]
SPY 170 [1, 2, 5, 10, 17, 34, 85, 170]

SNA 203 [1, 7, 29, 203]
XBY 91 [1, 7, 13, 91]
HLZ 217 [1, 7, 31, 217]
GLU 224 [1, 2, 4, 7, 8, 14, 16, 28, 32, 56, 112, 224]
UGX 41 [1, 41]

可见，这些间距之间最常见的公约数是 7，密钥的长度也很有可能是 7。

(2) 使用重合指数法进行分析

使用弗里德曼方法的公式估算出的密钥长度为 4.90949791619，因此密钥的长度很可能在 5 到 7 范围内，对密文进行实验，得到的结果如下：

密钥长度	平均重合指数
5	0.0443678029171
6	0.0445946667181
7	**0.0626787453796**

从上表可以看出，当密钥长度为 7 时，每个字串的重合程度更接近与 0.065，结合上面使用 Kasiski 方法得到的结果，我们可以确定密钥的长度为 7。

使用改进的拟重合指数法，对第一组密文进行测试，得到的重合指数数据如下（省略部分数据）：

移位	重合指数
……	……
20	0.0361954455446
21	**0.0630162093352**
22	0.0397806647808
23	0.0316516548798
24	0.0337533521924
25	0.0438577369165

可见对应最大重合指数的移位是 21，即密钥的第一位是 V。

以此类推，使用此方法可以将密钥全部解析出来。

四、实验成绩（共 5 分）

程序设计成绩(1 分)		实验结果成绩(2 分)	
实验报告成绩(2 分)		**总成绩**	
指导教师签字		日期	