**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

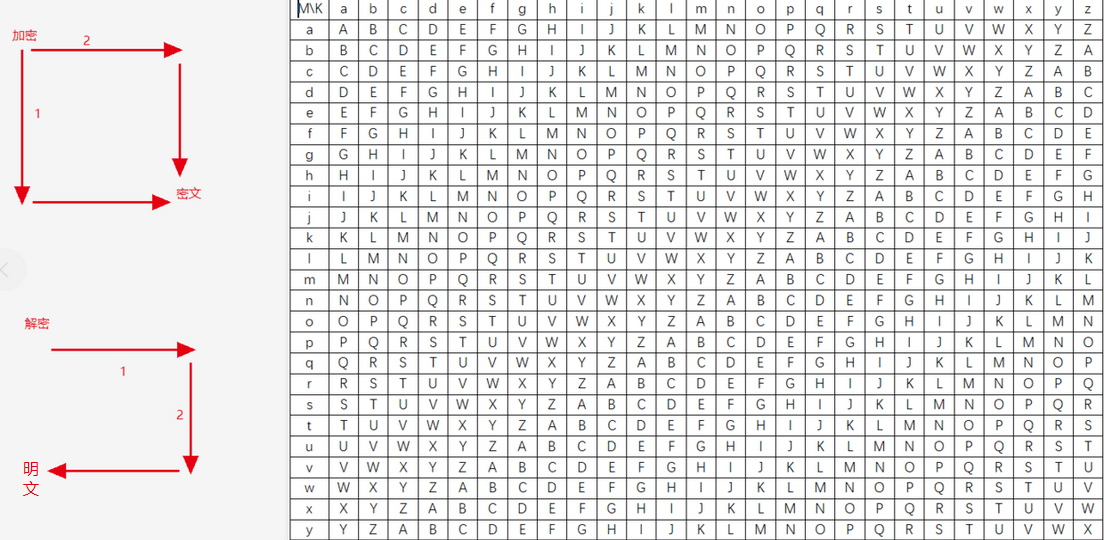
* 根据密钥字母，在字母表中找到对应行；
* 根据明文字母，在字母表中找到对应列；
* 已知明文：T， 密钥：F
* 得出密文：Y

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：

解密算法写成：

已知明文：T， 密钥：F，那么T为19，F为5，密文即为，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。）*



**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

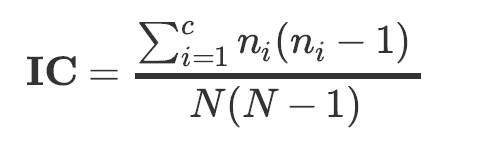
明文：CRYPTOISSHORTFORCRYPTOGRAPHY

密文：**CSASTP**KVSIQUTGQU**CSASTP**IUAQJB

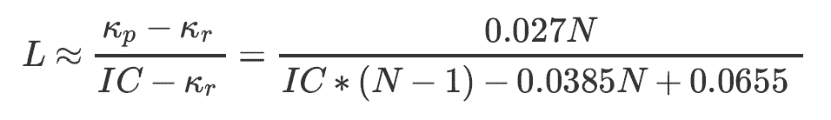
相隔16个字符出现相同字符片段，密钥有可能是16的约数（16，8，4，2）。当密文长度足够长时，还能找到其他的重复片段，取其公约数，即可确定密钥长度。

**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。



其中，是指字母表的长度（英语中为26），指密文文本的长度，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：



其中，是密钥长度，为目标语言中两个任意字母相同的概率（在英文中约为0.655），为字母表中出现相同字母的概率（在英文字母表中，=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](https://en.wikipedia.org/wiki/Vigen%C3%A8re_cipher)
  + [Kasiski examination](https://en.wikipedia.org/wiki/Kasiski_examination)
  + [Index of coincidence](https://www.wikiwand.com/en/Index_of_coincidence)
* 百度百科
  + [维吉尼亚密码](https://baike.baidu.com/item/%E7%BB%B4%E5%90%89%E5%B0%BC%E4%BA%9A%E5%AF%86%E7%A0%81/4905472?fr=aladdin)