# Homework 1

2017年3月8日 8:31

**HW1. Password Cracking**

**Description**

Crack following ciphertext:

- Caesar:
  - AWPLDPPYNCJAEJZFCYLXPHTESESPDLXPVPJLYODPYOEZESPEL
  - http://www.jianshu.com/p/c5ff37899393
- Vignere:
  - ktbueluegvitnthuexmonveggmrcgxptlyhhjaogchoemqchpdnetxupbqntietiabpsmaoncnwvoutiugtagmmqsxtvxaoniiogtagmbpsmtuvvihpstpdvcrxho kvhxotawsw quunewcgxptlcrxtevtubvewcnwwsxfsnptswtagakvoyyak
  - http://www.jianshu.com/p/8f0d23f9d64e
  - https://zh.wikipedia.org/wiki/%E7%BB%B4%E5%90%89%E5%B0%BC%E4%BA%9A%E5%AF%86%E7%A0%81
  - https://wenku.baidu.com/view/6867df07b52acfc789ebc9ad.html

  - **卡西斯基试验**

    明文中重复的元素在密文中并不重复。然而，如果密钥相同的话，结果可能便为（使用密钥ABCD）：

    密钥：ABCDAB CD ABCDA BCD ABCDABCDABCD

    明文：**CRYPTO** IS SHORT FOR **CRYPTO**GRAPHY

    密文：**CSASTP** KV SIQUT GQU **CSASTP**IUAQJB

    此时卡西斯基试验就能产生效果。对于更长的段落此方法更为有效，因为通常密文中重复的片段会更多。如通过下面的密文就能破译出密钥的长度：

    密文：**DYDUXRMH**TVDVMQ**NQD**QNW**DYDUXRM**HART JGW**NQD**

    其中，两个DYDUXRMH的出现相隔了18个字母。因此，可以假定密钥的长度是18的约数，即长度为18、9、6、3或2。而两个NQD则相距20个字母，意味着密钥长度应为20、10、5、4或2。取两者的交集，则可以基本确定密钥长度为2。

    从结果中看出，3，6，12，15，18，21，24重复字母相对较高，公因子假设为3。将密码分为3份

- Unknown:
  - MAL TIRRUEZF CR MAL RKZYIOL EX MAL OIY UAE RICF “MAL ACWALRM DYEUPLFWL CR ME DYEU MAIM UL IZL RKZZEKYFLF GH OHRMLZH”

| MAL | TIRRUEZF | CR | MAL | RKZYIOL | EX | MAL | OIY | UAE | RICF | "MAL | ACWALRM | DYEUPLFWL | CR | ME | DYEU | MAIM | UL | IZL | RKZZEKYFLF | GH | OHRMLZH |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| the | | | the | _____t | | the | | _h_ | | "the | h_he_t | _____e__e | | t_ | | th_t | _e | __e | ________e_ | | ___te__ |
| the | _a_____ | | the | ____a_t | | the | _a_ | _h_ | _a__ | "the | h_he_t | _____e__e | | t_ | | that | _e | a_e | ________e_ | | ___te__ |
| the | _a____r_ | | the | __r_a_t | | the | _a_ | _h_ | _a__ | "the | h_he_t | _____e__e | | t_ | | that | _e | are | __rr____e_ | | ___ter_ |
| the | _a__w_r_ | | the | __r_a_t | | the | _a_ | wh_ | _a__ | "the | h_he_t | ___w_e__e | | t_ | ___w | that | we | are | __rr____e_ | | ___ter_ |
| the | _a__wor_ | | the | __r_a_t | o_ | the | _a_ | who | _a__ | "the | h_he_t | __ow_e__e | | to | __ow | that | we | are | __rro___e_ | | ___ter_ |
| the | _a__wor_ | | the | __r_a_t | of | the | _a_ | who | _a__ | "the | h_he_t | __ow_e__e | | to | __ow | that | we | are | __rro___e_ | | ___ter_ |
| the | _asswor_ | is | the | s_r_a_t | of | the | _a_ | who | sai_ | "the | h_hest | __ow_e__e | is | to | __ow | that | we | are | s_rro___e_ | | __ster_ |
| the | _assword | is | the | s_r_a_t | of | the | _a_ | who | said | "the | hi_hest | __ow_ed_e | is | to | __ow | that | we | are | s_rro__ded | | __ster_ |
| the | password | is | the | s_r_a_t | of | the | _a_ | who | said | "the | hi_hest | __ow_ed_e | is | to | __ow | that | we | are | s_rro__ded | | __ster_ |
| the | password | is | the | s_rnamt | of | the | man | who | said | "the | hi_hest | _now_ed_e | is | to | _now | that | we | are | s_rro_nded | | m_ster_ |
| the | password | is | the | s_rnamt | of | the | man | who | said | "the | highest | _now_edge | is | to | _now | that | we | are | s_rro_nded | | m_ster_ |
| the | password | is | the | s_rnamt | of | the | man | who | said | "the | highest | knowledge | is | to | know | that | we | are | s_rro_nded | | m_ster_ |
| the | password | is | the | surnamt | of | the | man | who | said | "the | highest | knowledge | is | to | know | that | we | are | surrounded | | m_ster_ |
| the | password | is | the | surnamt | of | the | man | who | said | "the | highest | knowledge | is | to | know | that | we | are | surrounded | by | mystery |

    - 密文中多次出现"MAL"三字符单词，推测英文行文规律应该为” the"; 'M'->'t', 'A'->'h', 'L'->'e';
    - 密文中出现了"MAIM"由上已知为"th_t"，由于一般密码不会相同字符，因此推测'I'->'a';
    - 密文中出现"IZL"由上已知为"a_e",推测为"are",'Z'->'r';
    - 密文中出现"UL"由上已知为"_e",且其后跟着“are”，推测为"we",'U'->'w';
    - 密文中出现"ME"由上已知为"t_"二字符单词，推测为"to",'E'->'o';
    - 密文中出现"EX"由上已知为"o_” 二字符单词,推测为"of",'X'->'f';
    - 密文中出现"CR"为二字符单词，推测为"is",'C'->'i', 'R'->'s';
    - 密文中出现"RICF"由上已知为"sai_",且其后跟着“，推测为"said", 'F'->'d';
    - 密文中出现"TIRRUEZF"由上已知为"_assword",，推测为"password", 'T'->'p';
    - 密文中出现"OIY"为三字符单词，前连接“the”与“who”，为一人称，推测为"man",'O'->'m', 'Y'->'n';
    - 密文中出现"ACWALRM"由上已知为"h i _h e s t", 推测为"highest", 'W'->'g';
    - 密文中出现"DYEUPLFWLF"由上已知为"_ n o w _ e d g e",，推测为"knowledge", 'D'->'k','P;->'l';
    - 密文中出现"RKZZEKYFLFF"由上已知为"s_ r r o _ n d e d",，推测为"surrounded", 'K'->'u';
    - 密文中出现"TIRRUEZF"由上已知为"_assword",，推测为"password", 'T'->'p';
    - 最后经过统计得出未被采用的字母为b,c,d,j,q,v,x,y,z ,其中“GH”，推测为"by", 'G'->'b','H'->'y';
  - the password is the surname of the man who said the highest knowledge is to know that we are surrounded by mystery

**Deliverables**

Mail the deliverables to TA, with title: **SEC2017-HW1-ID-NAME**

- The plaintext and the substitution table / the key for each ciphertext
  - Do any thing the plaintext required you to do
- Your cryptanalysis process
- All in one PDF document.

**Deadline888**

- 23:59:59, Mar. 19th, 2017

弗里德里希·卡西斯基于1863年首先发表了完整的维吉尼亚密码的破译方法，称为卡西斯基试验（Kasiski examination）。早先的一些破译都是基于对于明文的认识、或者使用可识别的词语作为密钥。而卡西斯基的方法则没有这些限制。然而，在此之前，已经有人意识到了这一方法。1854年，查尔斯·巴贝奇受到斯维提斯（John Hall Brock Thwaites）在《艺术协会杂志》（Journal of the Society of the Arts）上声称发明了“新密码”的激励，从而破译了维吉尼亚密码。巴贝奇发现斯维提斯的密码只不过是维吉尼亚密码的一个变种而已，而斯维提斯则向其挑战，让他尝试破译用两个不同长度的密钥加密的密文。巴贝奇成功地进行了破译，得到的明文是丁尼生所写的诗《罪恶的想象》（The Vision of Sin），使用的密钥则是丁尼生妻子的名字Emily（艾米莉）。巴贝奇从未对他的方法进行过解释 。在对巴贝奇生前笔记的研究中发现，早在1846年巴贝奇就使用了这一方法，与后来卡西斯基发表的方法相同。[6]
卡西斯基试验是基于类似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。例如，明文中不同的CRYPTO可能被密钥ABCDEF加密成不同的密文：
密钥：ABCDEF AB CDEFA BCD EFABCDEFABCD
明文：**CRYPTO** IS SHORT FOR **CRYPTO**GRAPHY
密文：**CSASXT** IT UKSWT GQU **GWYQVR**KWAQJB

此时明文中重复的元素在密文中并不重复。然而，如果密钥相同的话，结果可能便为（使用密钥ABCD）：
密钥：ABCDAB CD ABCDA BCD ABCDABCDABCD
明文：**CRYPTO** IS SHORT FOR **CRYPTO**GRAPHY
密文：**CSASTP** KV SIQUT GQU **CSASTP**IUAQJB
此时卡西斯基试验就能产生效果。对于更长的段落此方法更为有效，因为通常密文中重复的片段会更多。如通过下面的密文就能破译出密钥的长度：
密文：**DYDUXRMH**TVDV**NQD**QNW**DYDUXRM**HART JGW**NQD**
其中，两个DYDUXRMH的出现相隔了18个字母。因此，可以假定密钥的长度是18的约数，即长度为18、9、6、3或2。而两个NQD则相距20个字母，意味着密钥长度应为20、10、5、4或2。取两者的交集，则可以基本确定密钥长度为2。

来自 <https://zh.wikipedia.org/wiki/%E7%BB%B4%E5%90%89%E5%B0%BC%E4%BA%9A%E5%AF%86%E7%A0%81>