后面就是秘密！
——密码漫谈

罗懋康

一大早就给床头柜上的手机闹醒。睡眼惺忪中抓起手机，摇摇头让自己清醒一些，看到底是哪个这么不知趣，太阳还没照到窗户呢，就把人吵醒，还让不让人活了！

定睛一看，一个激灵：老板催报告了！

胡乱塞了点东西到嘴里，匆匆出门，倒也还没忘将钥匙在门锁上反拧两圈。

离汽车还老远，按下车门遥控钥匙上的按钮，电磁波一下窜了出去，“咔嗒”一声轻响，打开了车门。

手中扭动点火钥匙，心里却在默念银行卡的密码——女儿早就想要的那个生日礼物只能用现金支付，待会儿还得去取款。

到了办公室，在脑中将老婆的生日排成标准 6 位数表示，逐一按下门外新装的数字密码锁按钮。

把自己沉重地扔进转椅，随手动了一下鼠标，计算机屏幕由黑变亮，却冷漠而有礼貌地要求先输入密码。暗骂一句“Shit!”——这词发音短促有力，很适宜用来发泄不大不小的郁闷——心想得赶紧装个指纹开机的玩意了，然后喃喃背诵一段名人格言，用笔杆把字头逐一戳入密码框，最后才“啪”的一声砸下回车键，总算打开了屏幕。

图 1. 密码门锁

或许你没有意识到，在一天的开头就这么半个多小时的时间里，一个都市人在意识和行为上已经不可避免地与密码至少打了 9 次交道：打进来的电话是手机基站以扩频伪码序列加密后传递过来的，回答出去的话音也一样；反拧的门锁钥匙是密码的一种等效实现形式；手中汽车门锁钥匙发出的电磁波更是加了密的；点火开关钥匙与密码的关系跟房门钥匙一样；要去银行或 ATM 机上取款，还得输入密码；办公室门锁更是以密码开闭；计算机的屏幕保护得用密码打开，要不就用指纹、眼底视网膜之类的密码等价形式来打开。

可以说，现代社会中的人，特别是都市中人，很难有哪一天能完全脱离开密码的影响，更不用说团体、机构、公司、银行、军队、国家等等利益远更重大的群体了。

不过，密码这玩意儿既没法离开，却又老是听说密码被盗、被破之类令人郁闷的事，弄得来不用不行，用又不放心，干脆，横下一条心，花上点时间，来看看——

1. 密码到底是个什么东西？

密码这个词，现代都市人已经是没几个不能随手举它十个八个应用例子的了；可真要比较全面、系统地说清楚密码到底是什么、干什么的，一时半会儿还真未必是一件容易事，至少不比试图通过拆解十把八把机械锁、电子锁来自制一把万能钥匙更容易。

> 试着回想一下我们在生活中遇到、使用密码的情形，我们首先对“密码”或“加密”、“解密”的概念给出一个通俗的界定：密码操作（或更一般的：信息加密）的本质，就是改变信息的表现方式使其令旁人难以理解的可逆过程。

第一个可能让人想到的问题是：干吗不直接就用保险箱——或者，微型保险箱？

可问题是：且不说体积问题，也暂且不考虑保险箱被 X

光透视内部结构打开甚至直接被大锤砸开、被乙炔焰割开的可能，就仅仅是一个重量问题，就没法让人忍受——总不能让人成天扛着一个上百斤的保险箱到处乱跑吧？

图 2. 微型保险箱——密码筒

OK，我们有“微型保险箱”，比如电影《达芬奇密码》中就展示过的密码筒，这种密码筒在历史上也真实存在过；但是，且不说密码筒作为容器的易损性，更关键的是，它能容纳的是载有信息的具体形态的物质，但我们需要的却仅仅是其中包含的信息而非这些物质本身；而可以选择的信息载体形式却远非物质形态这一个大类而已，比如，能量。它能收纳能量么？

更何况，别看说得那么玄，事实上可以看出：这个密码筒是可以“盲开”的！

只不过一到“以能量传输信息”，就难以避免被人中途截获；因此，将信息先行改变表现形态然后再传输，使得即使中途被他人截获，却也难以理解其中含义，这显然是一个效费比几乎最高的办法。

这就是密码（加密、解密）重大而独特的功用。

密码这种隐匿方式假定的是：就算你发现了已被加密的信息（现代密码学甚至假定你知道了加密操作的方式——加密算法），知道这里面有秘密，但在缺少密码的情况下，你仍然没法知道这些秘密是什么。

因此，我们可以给出**密码操作的几个基本要素**：

1. **明文**——不希望被未经允许的人看到的信息，可以是文字、符号、图形、图像、数据等等任何表现形式所包含的信息，相当于希望锁在保险柜里的东西；

2. **加密**——对明文的信息或搭载信息的信号进行处理，使其变得难以判读的操作过程，在现代技术条件下大多数就是某种算法，相当于将东西装进保险箱并按确定方式和步骤锁闭保险箱的过程；

3. **密钥**——加密时为保证信息只能被经过允许的人还原而设定的特定信息或信息载体，由允许的人持有，相当于打开保险箱的钥匙或在保险箱密码键盘上输入的密码；

4. **密文**——明文经过加密后所呈现的信息、信号表现形式，相当于装好东西已经锁闭的保险箱；

5. **解密**——使持有密钥的人，能通过密钥信息的输入，将已被加密的信息进行还原的方式和步骤，相当于在保险箱上插入钥匙或输入密码后打开保险箱的过程。

不过，这里要注意的是：对于任何一种加密方式来说，密钥并不是与加密方式相独立的一个要素，事实上，密钥只是这种加密方式中一些具有特定格式、可以单独改变的操作量而已；当这些操作量改变时，对信息的具体加密方式也就改变了。

我们可以用通常机械锁的结构和原理来说明：一种锁就相当于一种加密方式；如果为了使得不同使用者不能对这些锁互开，便对每个使用者都全新设计不同作用原理的锁，那任何一家锁厂干不了几天就都得关门，哪怕老板是比尔·盖茨。

因此，除了古代的锁或现代极少数特种用途的锁以外，几乎所有的机械锁无不采用“除了以吻合方式辨别特定形状钥匙的凹凸组合部分外，其余部分都相同”的通行设计方式；换成现在更一般、更流行的句式来说，就是“将识别功能和执行功能模块化”的方式。

图 3. 弹子机械锁原理及其锁芯构造

这样，锁具设计师和锁厂就可以对同一种设计，大量生产各把锁之间仅仅是锁芯中用于识别钥匙形状的弹子长度组合不同、其余所有结构都完全相同的锁具，使得不同的人买到的同一品牌、同一型号的锁，不但能保证相同的锁闭作用，

而且还由于锁芯中弹子的长度组合不同，因而所能辨别的钥匙也不同，因而不能互开。

由此可见，锁作为一个加密方式（系统），钥匙上按照锁芯的结构形式而确定的位置上的各凹凸变化点的凹凸组合，就是这个加密形式（系统）中可以改变的操作量，也就是密钥，而钥匙不过是这个密钥的物质载体而已。

因此，正如机械锁的情况一样（事实上电子锁也本质上相同），在信息加密技术中，将加密方式（系统）中一些操作量抽出来作为密钥这种做法的目的，就是使得一种加密方式能被多个持有不同密钥的人使用，但每个密钥持有人却只能解开那些以自己持有的密钥为操作量进行加密的信息；而且，这也使得同样的加密方式能在一段时间以后，仅仅需要更换密钥便可使得持有此前密钥的人不再能解密；还而且，由于密钥可由操作者自主设定，显然能使操作者对加密安全性的信心大为增强（试想想：假如钥匙的具体样式能由自己而非任何他人来任意确定，那这把锁的使用该是多么令人放心的事！）。

因此，当由于技术可行性原因或由于需求必要性原因使得这些操作量被固定在加密方式中时，密钥也就不存在了，例如人们通常使用的暗中约定的暗号。

所以，一般而言，密钥并非加密的必需，而只是对加密功能、性能的增强。

当然，对信息还有另一种隐匿的方式，那就是干脆将这些信息隐藏、隐蔽起来，压根不让别人知道有这些信息的存在。比如，以暗室技术将情报缩微成通常信件中的一个标点符号、以密写药水书写情报以及现在将信息隐藏在音乐、图片中等等。

图 4. 间谍在手表中隐藏缩微胶片

当然，大多数这类信息隐藏的方式都是直接将信息载体本身隐藏起来，比如特工将秘密文件藏在树洞里、《三国演义》中汉献帝将血书缝在玉带内让国舅董承带出宫外，等等。

这类方式，称为“信息隐藏”，不属于我们今天讨论的“密码”或“加密”的范围。

如果你还不放心，也大可先将信息加密，然后将加密后的信息再隐藏起来，便如将装有东西的保险箱再伪装成墙板、画框之类日常物品一样。

2. 很早很早以前

远古时代，一片湿地边上，两个部落的战士正围着兽皮裙、拎着大头棒，在各自头人的率领下怒目瞪视，而两边的头人，则正在伊里哇啦地争吵、威胁，要求对方退让出这片湿地——要知道，湿地可是狩猎取食维系生存的根本所在，有些类似于现在的中东石油产地。

突然，这边的头人高举手中的石斧，在空中划了两个圆圈。由于这个动作对于对方部落的人来说，既不属当时的“外族语言”、“部际语言”，又不属自己的“部落语言”，对方部落众人自然一阵莫名其妙。可战争却容不得犹豫，只一瞬间，石块和石矛已如雨点也似地从背后和两侧飞到头上。

其后的情形可想而知，自然便是后来数千年战场上不断上演的“兵败如山倒”的场面。

原来，这边的头人已经先安排了埋伏，并且，给埋伏的战士规定了“我一举起石斧划圈，你们就扔石块、石矛，接着就冲锋”的暗号。

这种暗号，除了“密钥”的功能不那么明显以外（类似于后来至今仍在采用的也是可靠性最高的“一次一密”加密方式），已经具备了前面所述的“密码”的各个基本要素。

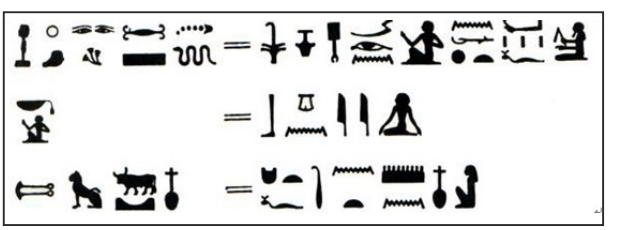

图 5. 古埃及象形文字密码：明文及对应的变体文字

图 6. 古埃及泥板文书

当然，远古时代这个依靠密码赢得战争胜利的战例是不会记入密码史的；唯一见诸记载的人类最早的密码（密文）雏形，是公元前 1900 年，古埃及一个书写员在一个描述他主人迦南·侯伯特二世的生平事迹的铭文中，使用了象形文字间的替代方法，使铭文变得难以理解和辨认。

不过，密码史上似乎并未对这个书写员的创举给予足够的肯定，原因是这段铭文的书写方式不完全符合密码的基本要义：尽可能不让未经授权的人理解；而是仍然希望后人读懂，只不过要制造一些可以克服的困难从而让后人产生神秘感和敬畏感罢了。

这样一来，中国就成了标准意义上最早发明密码的国家了：公元前 7 世纪年至公元前 4 世纪之间，也就是《孙子兵法》成书（公元前 5 世纪）的前后，有一本伪托姜太公吕望所著、后世在中国的声名与《孙子兵法》不遑多让的兵书出现，这就是《六韬》。

《六韬》中记载了殷商之际（公元前 1046 年前后）西周的姜子牙发明了最早的军队秘密通讯密码——阴符。

图 7.《六韬》竹简

武王问太公曰：“引兵深入诸侯之地，三军卒有缓急，或利或害。吾将以近通远，从中应外，以给三军之用，为之奈何？”

太公曰：“主与将有阴符凡八等：有大胜克敌之符，长一尺；破军擒将之符，长九寸；……。诸奉使行符，稽留者，若符事泄，闻者、告者皆诛之。八符者，主将秘闻。所以阴通言语不泄中外相知之术，敌虽圣智，莫之能识。”——《六韬·龙韬·阴符第二十四》

这里的意思是：按照只有我方知道的方式，以不同长度的竹片代表不同的军事用语，从而起到军事秘密通讯的作用。

由此可见阴符是一种替代密码，即对将要加密的信息以一定的方式进行分割，再对每一部分以我方特别指定的信息替代，然后保存或传递。

不过，阴符虽有其简便保密的特点，但毕竟仅仅是密码的“初级阶段”，过于简单，无法满足复杂的战场环境下军事通讯的需求。因此，姜子牙又创造出一种新的秘密通讯方法，即“阴书”。这仍然载于《六韬》之中：

武王问太公曰：“引兵深入诸侯之地，主将欲合兵，行无穷之变，图不测之利。其事繁多，符不能明，相去辽远，言语不通，为之奈何？”

太公曰：“诸有阴事大虑，当用书，不用符。主以书遗将，将以书问主。书皆一合而再离，三发而一知。再离者，分书为三部。三发而一知者，言三人，人操一分，相参而不知情也。此谓阴书。敌虽圣智，莫之能识。”——《六韬·龙韬·阴书》

这意思是说：如果有秘密而复杂的大事，则用“阴书”这种秘密军事文书，而不是用“阴符”这种简单的符号系统。

图 8. 天书

方法是：先把所要传递的机密内容完整地写在一编竹简或木简上，然后将这篇竹简或木简拆开、打乱，分成三份，即“一合而再离”；然后派三名信使各持一份，这样他们互相都不能知道具体的内容；让他们都送到同一个目的地，收件人再把三份“阴书”按顺序拼合起来，内容便一目了然了，即“三发而一知”。

可见阴书具有类似于移位密码的特性，即将原有信息的排列方式以我方特定的方式打乱，以让敌方即使截获也不能理解。

由于是分散传送，因而对于敌方截获的可能性而言，阴书的保密性不错。当然，对于我方接收的完整性而言，可靠性也就差些。

然后才是大约公元前 4 世纪时古希腊人发明的一种称为“天书”（skytale）的密码通信：发信人和收信人各持一根形状相同的特别圆棍，发信人将一张羊皮纸螺旋卷绕在他的圆棍上，然后写上情报；当取下羊皮纸时，由于先前的书写相当于每隔圆棍圆截面周长的距离写一个字母、再如此周而复始地写满所有空隙，因此这时展开的羊皮纸上就全是排列混乱不堪的字母了。信送到后，收信人将这些羊皮纸卷到同样的圆棍上，便可重新读出内容；而若没有同样的圆棍，即使卷起来也得不到原来同样的排列顺序，从而读不出原有的内容。

这显然仍是一种移位密码。

图 9. 恺撒征服高卢后凯旋

公元前 405 年，雅典和斯巴达之间的伯罗奔尼撒战争期间，斯巴达军队捕获了一名从波斯帝国回雅典送信的雅典信使，他身上除了一条布满杂乱无章的希腊字母的普通腰带外，没有任何情报。斯巴达军队统帅莱桑德研究了这些字母，最后通过在剑鞘上卷绕腰带读出了这些字母原来组成的文字——一份波斯告诉雅典他们将对斯巴达军队突袭的情报。实际上，这就是一条“天书”密文。莱桑德立即改变作战计划，回师攻击毫无防备的波斯军队，将其一举击溃。

公元前 58 年，罗马“前三巨头”之一的恺撒（Gaius Julius Caesar，公元前 102 ～ 前 44 年）发动了对高卢地区（高卢，法语：Gaule；拉丁语：Gallia；指现今西欧的法国、比利时、意大利北部、荷兰南部、瑞士西部和德国莱茵河西岸的一带）长达 8 年的征服战争，恺撒为此撰写了描述这场战争的《高卢战记》，共七卷，每年内容一卷。

恺撒在《高卢战记》中记述了他如何将密信发送给手下的事，但没有提到密码细节。好在二百余年后，苏托尼厄斯在其撰写的《恺撒传》中说明了这种密码，这就是密码史上著名的“恺撒密码”。

图 10.《高卢战记》12 世纪抄本和 18 世纪印刷版本

凯撒密码将字符表中每个明文字符都由其右边第 3 个字符代替，到结尾则接上字符表的开头进行循环：

移位前字符表：	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
移位后字符表：	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

因此这是一种简单的跨度为 3、循环周期为字符表长度的循环移位；但其开启了替代密码的先河（虽然实际上它同时也是移位密码），因而后世将凡是依某种自然的顺序进行替换的密码都称为恺撒密码。

不过，别看现在这种密码显得简单、不难破译，意大利黑手党“教父中的教父”贝尔纳多 • 普罗文扎诺在逃亡 43 年后于 2006 年 4 月 11 日落网时，从搜出的一些纸条看，他在 2002 年之前使用的都是一种与恺撒密码相似的密码，此后才因一个手下的被捕而改变了加密方式；而在他被捕的农舍中搜到的一本标有很多符号字句的《圣经》，则很可能就包含有他的新密码表，只不过一直未能破译。而且，更麻烦的是，从搜到的纸条中警方破译出一条信息：他早已用先前那种旧的密码指定了一位新教父，而这位新教父可是一个电脑高手！

图 11.《武经总要》

古代中国的军事家们似乎更喜欢类似“一次一密”的替代密码。例如，北宋仁忠时任至宰相的曾公亮，在其修撰的军事技术百科全书《武经总要》中，就提出了一种很可能是世界上保存至今最早的军用替代密码表.

他将搜集整理而得的当时军中常用的 40 个短语，以不同顺序进行排序，每一种排序构成一个不同的密码本。当部将出征时，主将发给部将一个密码本，不同部将或不同时期可用不同的密码本；然后对每个部将约好分别用某一首没有重复字的五言律诗，作为密钥。

图 12.《武经总要》中的三弓床弩图

如某部将在前线需要增拨弓、箭了，则从其持有的密码本中查出“请弓”为 1 号短语、“请箭”为 2 号短语，然后在主将与自己约定的五言诗如杜甫的《春望》中，找出第一、二字分别为“国”和“破”；然后再拟一公文，文中混编入“国”、“破”两字，并在其上加盖自己的印章以示这两个字就是密文。主将收到公文后，将其中标示的密文与《春望》和发给该部将的密码本相对照，即可得知其含义。

这个加密系统中，密码本，即以 40 个数字替代 40 个短语的方式，是可以更换的；密钥，即没有重复字的五言律诗，也是可以更换的；这就构成了一个完整的密码系统。

3. 要的就是让你头痛——古典时代的密码

按照我们在第 1 节中的介绍，我们已经知道，密码的本质可以如下描述：

假设：

P 为作为明文的信息集合；

K 为作为密钥的信息集合；

S 为作为密文的信息集合；

$\mathscr{F}$ 为作为加密方式的将两个信息集合映为另外一个信息集合的变换，$\mathscr{F}^{-1}$ 为其逆变换；

则加密、解密过程可以表示如下：

$$\text{加密：}\ \mathscr{F}(P,K)=S;$$
$$\text{解密：}\ \mathscr{F}^{-1}(S,K)=P。$$

这么说起来虽然略嫌抽象了一点，但却能将密码或加密、解密的本质可靠地概括、提炼出来。

为了有个比较具体、形象的想像，我们仍然可以用机械锁保险箱的情况来作比喻。当然，对于一般的机械锁而言，触发锁闭动作和开锁动作的钥匙都是同样的；在密码中这就是所谓的“对称密钥”。后面我们将要针对“非对称密钥”或“公开密钥”的情况设计一把锁闭钥匙和开锁钥匙不同的机械锁，但这里为简便起见，我们仍然采用通常的“对称钥匙”机械锁。

在这种情况下，明文P是保险箱中秘密文件所包含的信息；密钥K是钥匙上的凹凸组合信息；S为锁闭之后的装有秘密文

件的保险箱；$\mathscr{F}$为锁体中从插入正确的钥匙、锁芯中弹子组合识别出钥匙的正确性、锁芯按钥匙的扭力作出旋转，到推动相应执行机构进行锁闭、开锁动作的一系列特定过程的作用原理。而加密、解密就是锁体内执行锁闭、开锁动作的这一系列过程。

从加密表示式$\mathscr{F}(P,K)=S$看，就是要设计加密方式$\mathscr{F}$和加密密钥K以尽量使得旁人在仅仅获得密文S的情况下，非常难以仅由S推出加密方式$\mathscr{F}$（或其逆变换$\mathscr{F}^{-1}$）和解密密钥K，更不用说直接由密文S推出明文P了。甚至于，在现代密码学中，还进一步要求在旁人不仅获得密文S而且获得了加密方式$\mathscr{F}$的情况下，仍然不能推出解密密钥K，更不能推出明文P。

由此，一般而言，从保密的必要性看，加密方式$\mathscr{F}$和密钥K的复杂性显然越高越好（现代密码学已经不再如此笼统地要求二者的高复杂性了，而是要求$\mathscr{F}$具有“单向”性或“单向陷门”性，即“正变换不难，但反变换极难”或“在具有密钥时正变换不难，但没有密钥时反变换却极难”）；但从执行加密、解密过程的可行性看，则又是越低越好；这构成一对矛盾，通常只能在二者之间根据实际需要和实际可行性折中处理，即使是在现在海量高速计算机已经屡见不鲜的时候，依然如此。

试想想，当我们要给某人传送一封不愿让别人知道内容的密信时，我们有哪些基本的办法？

任何一种信息表示方式（如二进制、英文、中文等）均可视为一个系统，而任何一个系统均可视为由“单元集合”与“关系集合”构成；因此，任何一种信息表示方式也就均由“符号”和“语法”两个基本要素构成，符号（如 0、1、英文字符、中文单字等）表示信息的基本组成要素，语法给出以这些基本组成要素的组合来表达复杂信息的行为规则。

又由于任何信息都必须有某种载体才能表现，所以，要传送或保存任何一组信息，必须满足 3 个基本条件：载体、符号、排列。

由此可知，要让一组信息保持秘密，无非以下 5 种办法：

(1) 让人难以获知信息载体的存在（如密写药水、伪装成标点符号的缩微胶片等）；

(2) 让人即使获知信息载体的存在却难以获知在何时何处获取（如混在成千上万进出海关的人群中的秘密信使）；

(3) 让人即使获知信息载体在何时何处获取却难以获取（如首脑机关的秘密文件）；

(4) 让人即使获取信息载体却难以理解其表示符号的含义（如两河流域泥板文书上的楔形文字字符、殷墟甲骨文字符）；

(5) 让人即使理解其表示符号的含义却难以理解其组合方式的含义（例如一串不明其义的英文字符）。

前 3 种方法属于“信息隐藏”或“信息保护”，不属于“密码学”的范围；但后两种方法却正是密码学的两种基本方法：替换法与移位法。

当然，稍微复杂一点的加密方法都是这两种基本方法的结合或混合，而不单是其中某一种。

事实上，倘若单用这两种方法的一种，那么，稍不小心便可能使明文中某些字词与密文中某些符号、排列形成相对固定的对应关系，这样的情况下，利用密码破译（密码分析）中一种历史悠久的方法——频度分析法，便有可能攻破这段密码。

频度分析法，是基于这样一个事实：任何一种语言中，每个字母、单字或单词都有其基本稳定的使用频度。例如，在英语中，字母 e 出现的频率在所有字母当中最高，其次是 t，然后是 a，……；在阿拉伯语中，出现最多的字母是 a 和 l，而在汉语中频率最高的是“的”。字母或单字越少的语言，

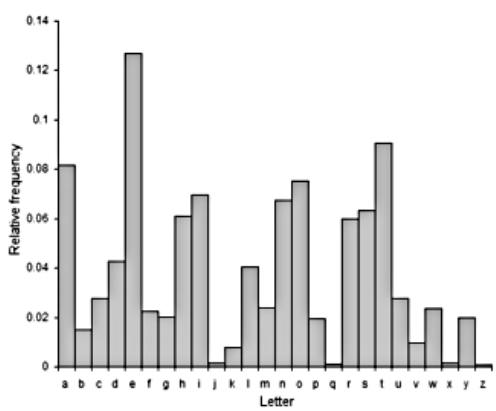

图 13. 英语字母使用频度表

长度越长的文字，这种频度的表现越是稳定。

这样，当已经估计到一段密文的明文是用哪种语言写成的时候，将其中出现频度最高的符号与该种语言使用频度最高的字母相对应，次高的符号与次高的字母相对应，……，辅以不断的分析、调整，便很有可能将其破译。

频度分析法最早是由谁提出的，已经湮不可考；但我们知道，公元 8 世纪中叶，阿拉伯阿拔斯王朝时，巴格达等地神学院中的神学家们在建立《可兰经》中穆罕默德启示录的年鉴时，就开始计算每一条启示录中各个单词的出现频率，他们甚至还研究单词的起源、变化与句子结构，来测试某篇文章是否与穆罕默德的语言模式相一致。公元 9 世纪时，同时兼为天文学家、哲学家、化学家和音乐理论家的阿拉伯人阿尔·金迪（al' Kindi，也被称为伊沙克 Ishaq，801 ～ 873 年）在他的《关于破译加密信息的手稿》中，提出了解密的频度分析方法。这是密码破译术上一次伟大的突破。

现在来看看在古典时代加密、解密曾经有哪些比较典型的方法。

(1) **“阴符”、“阴书”**：这应该分别是世界上最早的替代密码和移位密码。

(2) **“天书”**：这是密码界承认的世界最早移位密码。

(3) **“恺撒密码”**：这是密码界承认的世界最早替代密码。

(4) **“九宫格密码”**：欧洲中世纪（约公元 476 ～ 1453 年）时期，宗教势力处于高压统治地位，大量秘密结社兴起；最有名的就是影响深远的“兄弟共济会”。秘密通信的需要，使他们发明了这种替代密码：

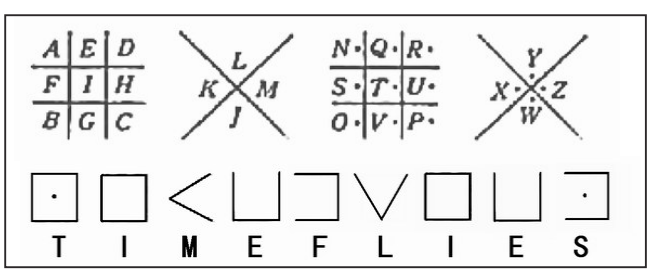

图 14. 九宫格密码

图中，上面一行以九宫格方式给出替换规则，下面一行中给出明文“TIME FLIES”（空格忽略）加密后的结果。

(5) **“书卷密码”**：以一段或一篇文章作为密钥，对其中每个单词依序编号：

01Under 02these 03circumstances, 04it 05seemed 06to 07many 08to 09define 10what 11England 12was 13fighting 14very 15hard 16for. 17Therefore, 18two 19years 20later, 21Robert 22Bridges, 23the 24Poet 25Laureate, 26asked 27Sir 28Hubert 29Parry 30to 31put 32it 33to 34music 35at 36a 37Fight 38for 39Right 40campaign 41meeting 42in 43London's 44Queen's 45Hall. 46The 47aims 48of 49this 50organisation 51were "52to 53brace 54the 55spirit 56of 57the 58nation 59that 60the 61people 62of 63Great 64Britain", 65explained 66by 67Jacob, 68a 69key 70role 71of 72it.

图 15.“书卷密码”解密表

以此作为解密表。再由此制作加密表：按字母顺序，将这段文字中每个单词的编号按相同首字母归并到一起：

a	26	35	36	47	68										
b	22	53	64	67											
c	03	40													
d	09														
e	11	65													
f	13	16	37	38											
g	63														
h	15	28	45												
i	04	32	42	72											
j	67														
k	69														
l	20	25	43												
m	07	34	41												
n	58														
o	48	50	56	62	71										
p	24	29	31	61											
q	44														
r	21	39	70												
s	05	27	55												
t	02	06	08	17	18	23	30	33	46	49	52	54	57	59	60
u	01														
v	14														
w	10	12	51												
x	75														
y	19														
z	81														

图 16.“书卷密码”加密表

注意到加密表中 x 和 z 在解密表中没有对应的编号，也就是说密钥或解密表中没有以这两个字母开头的单词，因此将它们另行单独编号为任意两个数字 75、81。

于是，加密时用加密表对照，将英文字母逐一转成对应的数字即可；当然，对于那些有不止一个数字对应的字母，可以也最好将所有对应数字都使用到。而解密时按解密表将数字转回相应的字母即可。

当作为密钥的这段文字长度够大、且很难单靠猜测和在常见文章、书籍中的逐一查找来发现时，书卷密码有非常高的强度，这从现实中一个持续至今上百年的事例——比尔密码（The Beale Ciphers）可以看出：

1822 年 1 月，美国弗吉尼亚林奇堡的华盛顿旅馆的主人莫里斯，受一个客人汤姆斯·比尔委托保管一个锁住的铁盒子；数月后，比尔给莫里斯来信，说铁盒子内保存着非常重要的

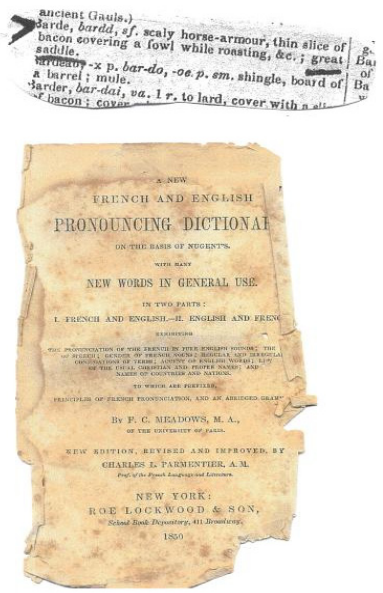

图 17. 比尔密码——曾经从中寻找密钥的词典

东西，事关他和朋友的性命；如果他和他的朋友没能来找莫里斯，就请莫里斯在 10 年之内一定保管好盒子。

莫里斯是一个很忠厚的人，一直守护了这个铁盒子 23 年，1845 年方才打开。结果里面是 4 张字条，3 张是密文，1 张是说明事情原委的明文。

原来，比尔是个冒险家，1917 年组织了一个 20 人的探险队，在一个险僻的峡谷里发现了金银矿；几年里他们聚集了大量的财宝，并将其藏匿在一个隐密的地方。比尔担心在挖掘成之前身遭不测，在遇到忠实可靠的莫里斯之后，便将各个藏宝地的位置、其中财宝数目、自己和同伴们所有亲戚的名字分别写成 3 张密文，委托莫里斯保管，以便在他们遭遇意外后，这些财宝仍能交给他们的亲人。

已经 23 年了，比尔或他的同伴仍未来认领，恐怕早已凶多吉少；莫里斯认为自己有责任找出这些宝藏来交给他们的亲人。于是他开始尝试破解这些密码，但在他余生 18 年里，却一无所获。临终前，他将此事告诉了一个朋友詹姆斯·沃德。沃德经过无数次的查找和尝试，终于破解了第 2 张密文，证实这是一个用《独立宣言》中的一段话加密的书卷密码，译出的文字是：

“我在离布法德约 4 英里处的贝德福德县里的一个离地面 6 英尺深的洞穴或地窖中贮藏了下列物品，这些物品为各队员 — 他们的名字在后面第三张纸上 — 公有。第一窖藏有 1014 磅金子，3812 磅银子，藏于 1819 年 11 月。第二窖藏有 1907 磅金子，1288 磅银子，另有在圣路易为确保运输而换得的珠宝。……”

这一破译引起轩然大波，无数的人查找、尝试了无数的文献。到了 20 世纪 60 年代，一些专门从事密码破解（密码分析）的人接受不了这个失败，专门为此组成了一个秘密协会——比尔密码协会，以便他们倾其知识和才智去攻破这个密码。计算机科学家、电脑密码统计性分析的先驱卡尔·哈默就是该协会的一位著名成员，他对比尔文件中的数字的分布做了大量统计、试验，总结得出结论：这些数字并不是随意写出的，它一定隐含着一段英文信息。

但是，虽然越来越多的数学家从事密码学研究，越来越多的巨型计算机被用来编制和破译密码，但一百七十多年前写成的比尔密码，仍然还在以冷峻的面孔冷迎世人——第 1 和第 3 张密文至今仍然未能破译。

从书卷密码的构成方式我们也可看出，书卷密码的本质是一个多对多的映射（图 18）。

假定有 A，B，C，…和 1，2，3，… 两组有序排列的符号，分别称为明文主符和密文主符；每个明文主符和密文主符又各对应着一组有序排列的符号，如符号 A 对应着 e2，g1，x4，称为明文次符，符号 1 对应着33, 47, 62, 91,称为密文次符。

在所有这些有序的明文次符 e2、g1、x4、a7、u4、……和密文次符 33、47、62、91、12、41、54、……之中，没有任何两个是相同的。

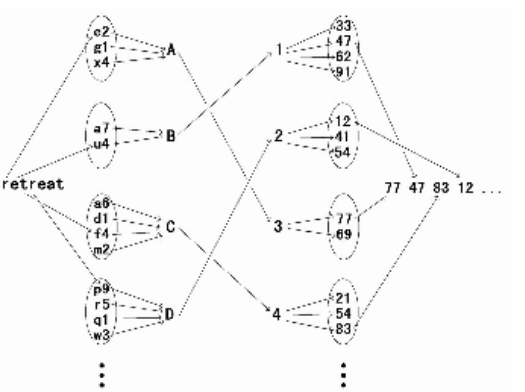

图 18.“书卷密码”基本原理图

于是，任何这样构成的明文主、次符和密文主、次符，加上明文主符与密文主符之间任何一组对应关系，就构成一个书卷密码。

比如，当明文（如图中的 retreat）中第 4 个字母 r 要加密时，先将 r 对应到第 4 个明文主符 D，再按预先确定的对应关系对应到密文主符 2，再对应到 2 的密文次符 12、41、54 中的第 4 个（待加密的字母 r 在明文中排在第 4 位）；但是，这里只有 3 个次符，因此按循环关系，取作第 1 个次符 12。最后，r 就被加密成了 12。

由于这里每个密文主符都对应着若干密文次符，因此，每个明文字符并不固定对应某一个密文次符，例如，明文 retreat 中第一个字母，虽然同样也是 r，但是，却被加密成 77。因此，只要次符的个数不要太少，便可将各个字母的使用频度很好地掩盖起来，使频度分析法失效。

事实上，多表替换之类的密码，遵循的也基本上就是这个书卷密码的基本原理。

(6) **“一次一密”**：按需要加密的明文长度需要，将图 16 的加密表扩张、用数字随机地填满、并仍然使得每个数字在表中只出现一次，就构成了“一次一密”的一次加密表；将很多各不相同且毫无规律的一次加密表装订在一起，就构成了“一次一密”的密码本；每次加解密，双方按照同样的密码本中同样的顺序使用同一张加密表（解密表），然后便将其撕掉销毁，下次再一起使用下一张表。

当使用者较多时一次一密的使用成本是很高的：必须高度可靠地分发并高度可靠地保管密码本；但是由于理论上已经证明，一次一密是唯一不存在统一的破译方法的密码，因此，直到现在，这种古老的加密方法仍被各国用来保护一些使用量小、但却具有极高密级的政治、军事机密。

(7) **“双字密码”**：我们知道任何一段信息均可表示为一段二进制数字 0100010010、10001010011100110、……等等；因此，任何有限信息也都可以用由两个不同字符组成的有限长字符串来编码表示，而且，还不一定非得按照二进制数字的进位规律来编码。在民众更多的是以那句“知识就是力量”而名垂千古的英国哲学家培根（Francis Bacon，1561-1626），就编制过这样一种隐密性很好的密码：

(i) 将 26 个英文字母中的每一个用长度为 5 的 a-b 字符串编码，如：A=aaaaa，H=aabbb，I=abaaa，K=ababa，L=ababb，M=abbaa 等等；

(ii) 将要加密的字句如“KILL HIM”去掉空格和标点符号，写一封通常的信件或文件，按“正体字母代表 a、斜体字母代表 b”的方式和 (i) 中的字母编码方式，将“KILLHIM”的每个字母表示成文件单词中由 5 个或正体、或斜体的字母组成的串，这样就完成了“KILLHIM”的 a-b 编码加密：

it seemed to many to define what England was fighting for.
ababa abaaa ababb ababb aabbb abaaa abbaa
K I L L H I M

(8) **“维吉尼亚密码”**（Vigenere Cipher）：维吉尼亚密码在密码史上名气很大，是古典密码中典型的多表替代密码，由亨利三世时法国外交官维吉尼亚（Blaise de Vigenere，1523 ～ 1596）发明。


图 19. 布莱兹·德·维吉尼亚

维吉尼亚密码的特点是用一张“维吉尼亚方表”加、解密，该表构造方法是将 26 个字母表每一行向左错一位、循环移位排列，合成一个表：

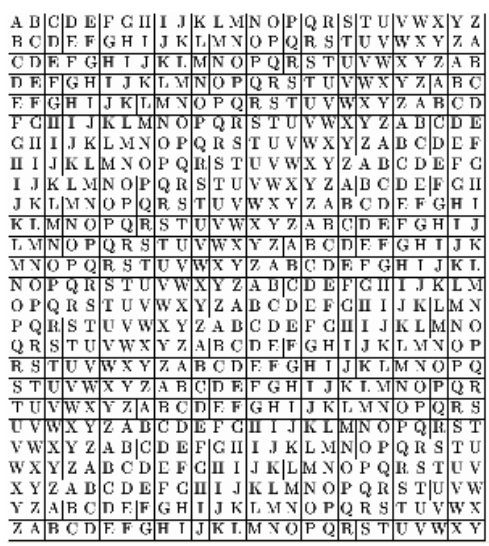
图 20. 维吉尼亚方表

要用这个表加、解密，还需要一个由字母组成的密钥，比如 KINGDOM。现设要加密的明文是“retreat and go to next city”。去掉空格，将明文变为 retreatandgotonextcity。

将密钥 KINGDOM 重复排列，直至其长度超过去掉空格后的新明文的长度，然后截掉后面比新密文多出的部分，再与新的明文上下对齐：

r	e	t	r	e	a	t	a	n	d	g	o	t	o	n	e	x	t	c	i	t	y
K	I	N	G	D	O	M	K	I	N	G	D	O	M	K	I	N	G	D	O	M	K

这样，明文第 1 个字母 r 就对应于密钥序列中的字母 K；在方表中的第 A 行（第一行）找到 R（不分大小写），顺着这一列往下找到与第 K 行相交的位置上的字母 B，这就是 r 被加密后的字母。

同样，明文第 2 个字母 e，对应密钥字母 I，在第 A 行中找到 E，往下找到第 I 行所在的字母 M，这就是 e 加密后的字母。

最后，明文“retreatandgotonextcity”用密钥“KINGDOM”加密后的结果是

BMGXHOFKVQMRHAXMKZFWFI

从维吉尼亚密码的加密表可以看出，由于明文中的同样字母只要处于不同位置就可能被加密成完全不同的字母，维吉尼亚密码很好地隐藏了字频（字母出现的频度）；这使得维吉尼亚密码对数百年来强大的频度分析攻击具有很高的抗攻击强度，因而，维吉尼亚密码在欧洲历史上纵横近 300 年，直到 1854 年，才被英国人查尔斯·巴比奇（Charles Babbage）破解。不过，由于他从未发表过这个结果，因而这个发现直到 20 世纪学者们检查巴比奇丰富的科学笔记时才被公布于世。

图 21. 巴比奇的差分机

这个巴比奇本身就是个甚是了得的人物，他发明了世界上第一台机械计算机——差分机，其基本结构原理至今仍为电子计算机沿用。

未完待续

征文启事

本刊的数学烟云栏目主要用于介绍数学学科的发展和研究内容等，欢迎广大读者投稿。来稿请寄：mc@global-sci.org

作者介绍

罗懋康，中国科技大学数学学士，四川大学数学博士，四川大学数学院教授，博士生导师，中国教育部长江学者特聘教授。曾获国家基金委杰出青年基金，为本刊编委。研究方向为不确定性理论，包括基础理论、应用理论与工程技术三个方向，涉及动力系统，计算智能，工程控制论，军事运筹学等。