Some Examples in Cryptology – The Making and Breaking of Codes

Ulrich A. Hoensch
Rocky Mountain College
Billings, MT 59102
www.rocky.edu/~hoenschu
August 2005

Abstract Examples of two classical ciphers, namely shift ciphers and the Vigenère Cipher, are given. For these ciphers, the encryption and decryption mechanisms are illustrated, and the process of breaking each of these ciphers using statistical methods is explained. Java applets that are available on the author’s web page are used for breaking the ciphers.

Caesar and Shift Ciphers

The simplest ciphers are shift ciphers; here, each letter in the text that is to be encrypted (the plain text) is replaced by the letter that occurs \( k \) places later in the alphabet. If the end of the alphabet is reached, one “wraps around”. The number \( k \) is the key for the shift cipher. If \( k = 3 \), then the cipher is also called Caesar cipher.

![Figure 1: Julius Caesar](image)

Figure 1: Julius Caesar – he invented a simple substitution cipher: replace each letter in the plain text by the 3rd letter on.

Example Let’s encrypt all gaul is divided into three parts using the Caesar cipher.

\[
\begin{array}{cccccccc}
   a & l & l & g & \ldots & t & s \\
   +3 \downarrow & +3 \downarrow & +3 \downarrow & +3 \downarrow & \ldots & +3 \downarrow & +3 \downarrow \\
   D & O & O & J & \ldots & W & V \\
\end{array}
\]

The cipher text is DOOJD XOLVG LYLGH GLQWR WKUHH SDUWV.
Cryptanalysis of Shift Ciphers  Now, we consider the situation when the cipher text is given, and we want to find the plain text, without knowing the key. This process is called *cryptanalysis*, or “breaking the cipher”.

We use the fact that the distribution of letters in a standard English text is the following:

![Frequency Distribution in Standard English](image)

Figure 2: Frequency Distribution in Standard English

If we have a long cipher text that has been encrypted using a shift cipher with key $k$, then we expect to see the same distribution, except that the picture has been shifted $k$ places to the right (and wrapped around).

**Example**  Let’s analyze Cipher Text 1 (see Appendix). Using the *Statistical Analysis Applet*, we obtain the distribution of letter frequencies in the cipher text.

![Frequency Distribution for Cipher Text 1](image)

Figure 3: Frequency Distribution for Cipher Text 1
This graph looks very much like the one in Figure 2, except everything seems to have shifted right. Since the highest bar for E now occurs for the letter L, a shift of 7 letters to the right seems to have occurred. If a shift cipher has been used, then it is highly likely that the key is \( k = 7 \).

To obtain the plain text, we must take each letter in the cipher text, and find the letter that occurs 7 places to the left in the alphabet. Since the cipher text begins with

\[ \text{OVSTL ZOHKI LLUZL HALKM VYZVT LOVBY ZPUZP SLUJL DPAOO} \ldots \]

going back 7 letters in the alphabet will give

\[ \text{HOLME SHADB EENSE ATEDF ORSOM EHOUR SINSI LENCE WITHH} \ldots \]

or, re-arranging whitespace as appropriate,

*Holmes had been seated for some hours in silence with h.*

To expedite this process, we use the Statistical Analysis Applet once again. We note that L must be decrypted to e, M must be decrypted to f, N must be decrypted to g, and so on. We enter these plain text characters in the appropriate places in the applet, press the button “Perform Decryption”, and obtain (a somewhat chopped up, but readable version of) the plaintext (see Screenshot 1 in the Appendix).

**The Vigenère Cipher**

While shift ciphers are extremely primitive, and can also be broken using a brute-force attack, i.e. by trying out all possible keys (there are only 25 of them), the Vigenère Cipher is very much more secure, and until the end of the 19th century, was considered by many people to be “unbreakable”.

The Vigenère Cipher is a *keyword cipher*; it relies on keeping the key (which is usually a word or phrase) secret. The process of encryption and decryption itself does not have to be kept secret.

Here is how it works: suppose we want to encrypt the phrase *victory is ours* using a Vigenère Cipher with keyword *cannon*. Now, we use the Vigenère Tableau (Figure 5) as follows: find the row starting with the first letter of the plain text, v, and the column headed with the first letter of the keyword, c.
The letter in this row and this column, \( X \), is the first letter of the cipher text. Then, we repeat this process with the second letters of both the plain text and the keyword, and so on. We repeat the keyword once we have reached its last letter.

<table>
<thead>
<tr>
<th>plain text</th>
<th>o</th>
<th>n</th>
<th>c</th>
<th>t</th>
<th>o</th>
<th>r</th>
<th>y</th>
<th>i</th>
<th>s</th>
<th>o</th>
<th>u</th>
<th>r</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>keyword</td>
<td>h</td>
<td>a</td>
<td>n</td>
<td>n</td>
<td>o</td>
<td>n</td>
<td>c</td>
<td>a</td>
<td>n</td>
<td>n</td>
<td>o</td>
<td>n</td>
<td>c</td>
</tr>
<tr>
<td>CIPHER TEXT</td>
<td>X</td>
<td>I</td>
<td>P</td>
<td>G</td>
<td>C</td>
<td>E</td>
<td>A</td>
<td>I</td>
<td>F</td>
<td>B</td>
<td>I</td>
<td>E</td>
<td>U</td>
</tr>
</tbody>
</table>

So the cipher text is \textbf{XIPGC EAIFB IEU}.

How do we decrypt a cipher text using a Vigenère Cipher? Well, we reverse the process described above.

**Example**  
Decrypt \texttt{LVWZL TYVO RBX} if a Vigenère Cipher with keyword \texttt{lion} has been used.

In the column headed by the first letter of the keyword, \( l \), we have to find the row in which the first letter of the cipher text, \( L \), occurs. This row starts with the letter \( a \), which is the first letter of the plain text. Then, in the column headed by the second letter of the keyword, \( i \), we have to find the row in which the second letter of the cipher text, \( V \), occurs. This row starts with the letter \( n \), which is the second letter of the plain text. Continuing this process, we obtain:

<table>
<thead>
<tr>
<th>CIPHER TEXT</th>
<th>L</th>
<th>V</th>
<th>W</th>
<th>Z</th>
<th>L</th>
<th>T</th>
<th>Y</th>
<th>O</th>
<th>R</th>
<th>B</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>keyword</td>
<td>l</td>
<td>i</td>
<td>o</td>
<td>n</td>
<td>l</td>
<td>i</td>
<td>o</td>
<td>n</td>
<td>l</td>
<td>i</td>
<td>o</td>
</tr>
<tr>
<td>plain text</td>
<td>a</td>
<td>n</td>
<td>i</td>
<td>m</td>
<td>a</td>
<td>l</td>
<td>k</td>
<td>i</td>
<td>n</td>
<td>g</td>
<td>d</td>
</tr>
</tbody>
</table>
Note To save time, you can use the Vigenere Cipher Applet to perform the encryption and decryption.

Cryptanalysis of the Vigenère Cipher Here, we want to find the keyword that was used for the encryption. Once we have found this keyword, we can decrypt the cipher text as explained above. There are two steps involved in the cryptanalysis of the Vigenère Cipher.

Step 1 The Kasiski Test is used to determine the length of the keyword. It is implemented in the Kasiski Test Applet.

Step 2 If the keyword consists of $n$ letters, divide out the cipher text into $n$ sections, and perform statistical analysis for each section. We use the Polyalphabetic Analysis Applet to do this.

Example We presume that Cipher Text 2 (see Appendix) was encrypted using a Vigenère Cipher. If we drop the cipher text into the Kasiski Test Applet, we see the following output.
In essence, what the program does is this: it find the distances between tri-graphs that occur more than once in the cipher text. A trigraph is a sequence of three letters. Then, the greatest common divisor among these distances is computed. If you press the button “Show Histogram”, the distribution of these greatest common divisors is shown. Note how the number 5 occurs most often (see Figure 7).

This leads us to believe that the keyword consists of 5 letters. The logic behind the Kasiski Test is that if a trigraph repeats in the plain text at a distance that is a multiple of the keylength, then it will also repeat at that distance in the Vigenère-encrypted cipher text.

Now that we think that the keyword has length \( n = 5 \), we drop the cipher text into the Polyalphabetic Analysis Applet and enter 5 in the field for the keyword length. What the program does now is this: it performs Statistical Analysis for the section of the cipher text that consists of the 1st, the 6th,
the 11th, . . . , the \((1 + 5 \cdot N)\)-th letter. The result is given in Figure 8.

Figure 8: The frequency distribution of the first section for Cipher Text 2 (assuming a keyword length of \(n = 5\))

This picture looks familiar — it seems to be the usual distribution of Standard English, except it is shifted. Note that where we expect the letter A to be in the Standard English distribution, there is now the letter S. This is the first letter of the keyword.

When selecting the next section using the “Next>>” button, the program shows the distribution of the section consisting of of the 2nd, the 7th, the 12th, . . . , the \((2 + 5 \cdot N)\)-th letter (see Figure 9).

Figure 9: The frequency distribution of the second section for Cipher Text 2 (assuming a keyword length of \(n = 5\))

The second letter in the keyword appears to be an \(N\). Continuing like this, the entire keyword comes out to be \texttt{SNAPE}.
Finally, enter the keyword in the applet, and click the button “Decrypt using Vigenere Cipher”. The following text results (see Screenshot 2 in the Appendix):

mrand mrsdu rsley ofnum berfo urpri vetdr ivewe repro udtos aytha tthey werep erfec tlyno
rmlalt hanky ouver ymuch theyw ereth elast peopl eyould expec ttobe invol ved anyth ingst
range ormys terio usbec auset heyju stdid nthol dwith suchn onsen semrd ursle ywast hedir
......

Experiments and Exercises

To use the Java applets and access the cipher texts, visit the web page http://www.rocky.edu/~hoenschu/HighSchoolTalks/Crypto/main.html

• Perform Statistical Analysis for Cipher Text 3. What kind of cipher has been used, and what is the key?

• Decrypt Cipher Text 3; you may do this by hand, but you can also use the Vigenere Cipher Applet. How would you use the applet? Why does it work?

• Analyze Cipher Text 4. It has been encrypted using a Vigenère Cipher. Find the length of the keyword first, then find the keyword. What is the plain text?

• Analyze Cipher Text 5. What is the plain text?

About Me: I graduated from Michigan State University in May 2003 with a Ph.D. in Mathematics, and am now an Assistant Professor of Mathematics at Rocky Mountain College. My main interests and field of expertise lie in the area of Dynamical Systems. I find great pleasure in teaching a variety of mathematics and computer science classes at Rocky, including a class on cryptology, from which this presentation was derived.

About the Math Department at Rocky Moutain College: We take great pride in our program, which offers a major and minor in both mathematics and mathematics education. As faculty at a small liberal arts college, we are committed to educating our students at the highest level. We care about our students’ academic progress and are dedicated to preparing them for their professional future. Small classes, a broad array of course offerings, accessible faculty, and a modern curriculum make the mathematics program at Rocky truly outstanding within the region. Most of our graduates enter graduate school, where we have a 100% placement rate for math majors, or become math teachers. Other career choices for math majors are to work for banks or insurance companies, in technology-related fields such as engineering and computer science, or for the Federal Government.
Cipher Text 1

OVSTL ZOHKI LLUZL HALKM VYZVT LOVBY ZPUZP SLUJL DPAOO PZSVU NAOPU IHJRJ BYCLK VCLYH JOLTP JHSC CL ZZLSP UDOPJ OOLDH ZIYLD PUNHW HYAPJ BSHYS FTHSV KYYVB ZWYVK BJAOP ZOLHK DHZBZ URBW VUOZI YLHZA HUKOL SVRVL KMYVT TFWP UAVMC PLSDP RLHZA YHUNL SHURI PYKDP AOKBS SYLFL WSSTH NLHUK HISHJ RAWVR UVAYZ DHAUV DZLKBK UVAYW VVZVL AVPCU LZAPU ZVBAO HYMPJ HUZLJ BYPAP LZZPNH CLHZA HYAVM HZAVL PZOTZ UAOHH BZAVT LTHKZP DHZAV OVSTL ZZJBY PVBBM HJBSA PLZAO PZBPK KLUPE AYBPZ VUPUA VFTPH ZAPUA PTHAL AOVBN OAZDH ZBAAL YSFPU LEWDI JHISL OVDVU LHAYO KVFBF RUVDA OHAPH ZRLKO DOLLK SLKYL BKBUK BUOPZ ZAOFI DPAOH ZALHT PUNAL ZAABI LPUOP ZOHUK HUKHN SLHTV MHTBZ LTLUA POOPZ KLLWZ LALFL ZUVDD HAZVU JVUML ZZFVZ FYLSM BAILY SFAHR LUHIN JRZHP KOLPH TPVBN OAAVT HRLFY RZPNL HWOWL YAYAO HAMDJ FILZH BZLPU MPCLT PUBL ZFVBD PSSZH FAOHA PAPZH SSZVH IZBYK SFZPT WSLPH TZWBY AOHAP ZOHHZ ZHFUV AOPUN VMAML RPUKF VBZLL TFKLH YDHAZ VUOLW YVWWL KOPZA LZAAB ILPUA OLYHJ RHUKI LNHUDA VSLJ ABLDP AOAOL HPYVM HWYVM LZSVY HKKYL ZZPUN OPJZS HZPPA PUZVA YLHSS FKPMJ PBJSA AVJVU ZAYBJ AHZLY PLZVM PUMLY LUJLZ LHJKL LWLJK LWAEP VUPAZ WYKLL JLZLV YHUKL HJOZJ TWLSP UPAPZ SNMPC MALKY VPUNZ VVULZ PTWSF RUVJR ZVBAA SSHAOL JLUAY HSPUM LLYIJ LZHUK WYLZL UAYZU LZHUB PLUJL DPAOA OLZAH YAPUN WPUPA HUKAO LJVUU SBZPV UVULT HFWYV KBZLI ZAHYA SPUNA OVBNO WZVZP ISFHFT YLHAY PJFVZ ZLMLL JAUVD PADHZ UVAYL HSSFK PMPFJ BSAIF HUPUZ WLJAP VUVMA OLYNV VCLIL ADLUL FVBYS LMAVL YLMPA NLHUY KAOBT IAVAL LSZBY LAOHA FVKVZ KUVAM YVWVZ LAVPU CLZAF BYZHT HSSJH WPASH PUAVL NVSKM PLSZK PZLIL UJAVU UCLCYF SRLPS LVAIK BAPJH UXBPJ RSFZO VFDBF HJSVZ LJVU OVJPAO HYLAO LTPZU PNSUB URZVM AOLLF YXZPT WSLJO HPUVF BOHKJ OHSRI LADDL UVFBO SLAMUR PUNLY HUKA BTIDO LUFVB YLABY ULKMY TVAOL JSBSV HZAUP NOAFV BBWAJ OHSSA OLYLQ OLVUF BWSHF IPSSH HYYZA VZLHY KFQOL JBLFY BULCL YWSNF IPSSH HYYZL EJLWA DPAOA OBYZA VUYFV ABVST LMVBE DLRZ UOND HAAOB YZAVU OHKCU VWAPZ UVUZV TLZVB AOHHY PJHUU YVWLY AFDOF JODVB SKLEW PLYLPN HTVUA OHUKD OPJOO LKLZP YLFVZ BAVZO HYLPD ADOPT FVBZJ XLBZL IVVRP ZSVJL LKPUE FYHDI LHYKU FVOH CLVU HZRLK MVYAO LRLFF VBKVU VAWYZ WZLVA VPUCF ZAFYR YTVUL FPUAU PZTHU ULYOY DHZBZ YKSFZ PTWSL PAMYPL XXBPA LZVZH PKOLH SPAAS LULAA SKLKC LHYFW VISLT ILJVT LZCILY FJOPS KPDOD OLUUV JLPAP ZLEWS HPULK AVFVB OLYLZ ZHUBU LEWSH PULKV ULZLL DOHAF VBJHU THZRL MAOHA MYPLU KIDHAI VUOLA VZZLX HZOLL AVXWH VLYBW VUOAL AHISL HUKAB YULKV UJTVY YLAVO ZJOL TPJHS HUHSF ZPZ
Cipher Text 2

EEACH EESSY JFLTC GSNJQ TRRUS MEPGM AIELI JRPGS MQTDW SLTWE LGHTC ORRTT WEFTG LYYCS
JZAAAX ZNNZC GHVTQ QZURL EUENA WEEIL WYAHX HROEP WLOJH WKPTG LQOQI AADVPS NRDXR SAYIL AAGHX
JNNVI GEMNW LRRXS MFBTG SHSTX ZRYYY KGDXH FGHDPS VSIIIL KHCWR KGRMHE MASAI QJAHX ZRDXX
WPTDV GSAUM JZCPP DRDV MANXR YFWWM UUMPH WQRPX DFNTHA SFAQM YQETJ QZACA AGHWE JQLENE FLNTG
CNLIL GHGWL WQISL SIEPZ WEYAY JTEBS MTFTPG ZRMMG VHRHP WLPW LUICE FBQBS FQEP DUEPA VUSAR WNRAC
LJIRI LUEJW MLNPQ GHNIS XAERO DURRL UNMTM FIEGC MFEDW DNSHL WFPTR LFQBY UUOUL WETXQ WPRPR
AAGDZ WEGPV VRNUI FPEHW HLICK GFTWI FRIVL TBQUG LUESY JFLTC KUAASE KBNAK KBNEE QYPWE MQLTC
SADXR LUEXV GMCIM GATWI JRWPW FBFXR WEBDC SANYL WEEIL WQUGW DRYHL SBEKI JLTWM FTTWI QJACX
WQBJX LUENF DFEWE VNSTG JRTPR VGHTM JTRTE LRSI JWRLE KGHPX KBMTF QQYLS MYDSM KPQI JVTL
WLDXH FGTWM FXTWI QPJDP VOEPV AGIUE FLOCX XBUCX GHTPH GHTIL WCOIX WESBV KCOIX WENPW EESSY
JFLTC KFIHX WEBJX LUENL SQNIQ WQFDV KRTV SYTTE JFICJ SPVBN KQUGW DRYEV WGECH WQSWI VVDCX
ZNVTE KVSII JOERF MFWEJ FJIXH WEACH ZRRVS QGFDV FBTW FTHJW TNNSA WEEPW MADJQ KYENM KUAHM
LJAHF GSXSF DRTDF WGHTE MASEI QFWSY VQEGI VQGOI AAKLL SGTWI FRIVL TBQUG OBUAH KNYXJ LUEES
LGEW SERXQ WQICX ZRSIV WRTIL WQUGW DRYHO FRIWIL SGTWI HBTII JFHPH SFMPF DFOCX GBBXJ LUENL
SQNII ZEEKI FQETR ZVMIL AFDBC ONSPR GHHTV YBOSV WNSDR XBRZI WCICK LUEES LGEW SIANX ZRYSC
VATEL GQDJH DRYBM PVMVA AGHPG ZVLSP AXEIL SGWII FZPRP VZRHJ MASEI QJQZI MCOCX ZRDJP DTRTC
LHEHH SLQJH KQQGC KGAQX KHGVX JHWAH QHHXR QNELD LQHTG DBUSC KXYDG LFSIS LBSJX QRXX NTHX
JNNVI SADBC KGEFM GHSLA AAGHA GHLSW GNQEH ZNEFZ FVNWE DOKI JQHTG GHNTV QZRSY JFLTC ZHMBI
VNSWI HVCZI VBUIL AFMDW LQOGM RTTXY XBRLS JXACX EESSY JFLTC YBHM HRDPA SLHPT HVNLNE KFHTA
JRSLP WQAHG JRABM FTDJH DRYXR LBHXW ZVGWZ ZNIGR GAEDJ LUEBR GGIWI VNLPV YRTPA FLOLP XYUIX
WEPPW LGHTA AADDAA