VIGvisual: A Visualization Tool for the Vigenère Cipher

Can Li, Jun Ma, Jun Tao,
Jean Mayo, Ching-Kuang Shene
Department of Computer Science
Michigan Technological University
Houghton, MI
{canli,junm,junt,jmayo,shene}@mtu.edu

Melissa Keranen
Department of Mathematical Sciences
Michigan Technological University
Houghton, MI
msjukuri@mtu.edu

Chaoli Wang
Department of Computer Science & Engineering
University of Notre Dame
Notre Dame, IN
chaoli.wang@nd.edu

ABSTRACT

This paper describes a visualization tool VIGvisual that helps students learn and instructors teach the Vigenère cipher. The software allows the user to visualize both encryption and decryption through a variety of cipher tools. The demo mode is useful and efficient for classroom presentation. The practice mode allows the user to practice encryption and decryption. VIGvisual is quite versatile, providing support for both beginners learning how to encrypt and decrypt, and also for the more advanced users wishing to practice cryptanalysis in the attack mode. Classroom evaluation of the tool was positive.

Categories and Subject Descriptors

K.3.2 [**Computers and Education**]: Computer and Information Science Education—*Computer science education, information systems education*

General Terms

Algorithms, Security

Keywords

Cryptography; visualization

1. INTRODUCTION

The Vigenère cipher appeared in the 1585 book *Traicté des Chiffres* by Blaise de Vigenère. It is a simple cipher, but for nearly three centuries the Vigenère cipher had not been broken until Friedrich W. Kasiski published his 1863 book [7]. Charles Babbage also broke the cipher with a similar technique in 1846, although he never published his work. Currently, the Vigenère cipher has become a standard topic in many textbooks [4, 8, 10].

Well designed pedagogical tools are very useful in helping students understand concepts and practice needed skills. While there are tools available [1, 2, 9], most of them only provide interfaces for encryption and decryption without showing the process, and very few include cryptanalysis. VIGvisual is designed to address this issue by providing an environment so that it can be used in the classroom and for self-study. It is able to animate the Vigenère cipher with a variety of cipher tools, all of which are available for students to practice encryption and decryption with error checking. Furthermore, VIGvisual also helps students learn how to break the Vigenère cipher. VIGvisual uses Kasiski's method and the Index of Coincidence method for keyword length estimation, and the χ^2 method with frequency graphs for keyword recovery. To the best of our knowledge, only [2] offers a similar capability; however, it is just an interactive cryptanalysis environment. VIGvisual goes one step further by offering a more comprehensive visualization component with tools and animation not only for cryptanalysis but also for beginners to learn and practice the Vigenère cipher.

In the following, Section 2 discusses the course in which VIGvisual was used and evaluated, Section 3 presents an overview of VIGvisual, Section 4 has our findings from a classroom evaluation, and Section 5 is our conclusion.

2. COURSE INFORMATION

VIGvisual was used in a cryptography course, MA3203 Introduction to Cryptography, that is offered out of the Department of Mathematical Sciences at Michigan Technological University. It is a junior level course that gives a basic introduction to the field of cryptography. This course covers classical cryptography, the Data Encryption Standard (DES), the Advanced Encryption Standard (AES), the RSA algorithm, discrete logarithms, hash functions, and elliptic curve cryptography. For each cryptosystem, we study how it was designed, why it works, how one may attack the system, and how it has been used in practice.

Understanding classical cryptography is essential to any introductory cryptography course, and one of the major classical cryptosystems is the Vigenère cipher. The Vigenère cipher is a generalization of the monoalphabetic shift cipher, which has a keyword length of one. This cipher is very difficult to present in class because of the long pieces of text that need to be used in order to illustrate the algorithm in a meaningful way. Because VIGvisual was used in the course, students were able to see the encryption algorithm demon-

ITiCSE'15, July 6–8, 2015, Vilnius, Lithuania.

DOI: http://dx.doi.org/10.1145/2729094.2742589.

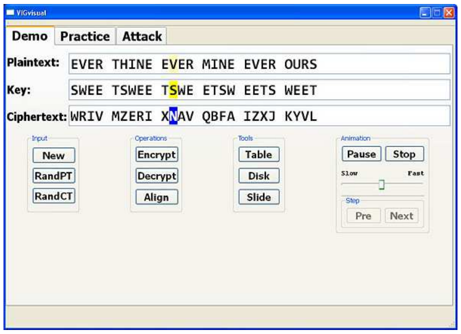
strated quickly. This allowed for a more thorough study of the most interesting aspects of the cipher, which are the decryption techniques and attacks. The attack component covers Kasiski's method, the index of coincidence method, and the χ^2 method.

3. SOFTWARE DESCRIPTION

VIGvisual supports Linux, MacOS and Windows and has three modes, `Demo`, `Practice` and `Attack`. The `Demo` mode helps the user visualize the process of encryption and decryption with animation. The `Practice` mode allows the user to practice encryption and decryption with error checking. The `Attack` mode offers a chance for the user to learn how to break the Vigenère cipher with Kasiski's and the Index of Coincidence (IOC) methods, followed by the χ^2 method to recover the unknown keyword.

3.1 The Demo Mode

VIGvisual starts with the `Demo` mode (Figure 1). The top portion has input fields for plaintext, keyword and ciphertext. The user starts a new session with `New` and enters a plaintext-keyword pair or a ciphertext-keyword pair, or uses `RandPT` and `RandCT` to automatically generate a random plaintext-keyword pair and a random ciphertext-keyword pair. This is followed by clicking `Encrypt` or `Decrypt` to encrypt or decrypt the entered message. By default, the keyword is repeated and aligned with the original word structure. Clicking `Align` switches to the view of breaking the plaintext/ciphertext to align with the keyword length.



Plaintext: EVER THINE EVER MINE EVER OURS

Key: SWEE TSWEE TSWE ETSW EETS WEET

Ciphertext: WRIV MZERI XNAV QBFA IZXJ KYVL

Figure 1: Screenshot of the Demo Mode

The user uses `Start` and `Stop` to start and stop an animation, the slider to select an animation speed, and `Pre` and `Next` to move to the previous and next position. The corresponding plaintext letter, keyword letter and ciphertext letter are shown in different colors in an animation.

The user may bring up one or more tools with buttons `Table`, `Disk` and `Slide`. Figure 2(a) shows the Vigenère table. The plaintext letter under consideration and its column use one color, the corresponding keyword letter and its row use a different one, and the ciphertext letter is at the intersection of the plaintext column and keyword row. The cipher disk has two concentric disks stacked together (Figure 2(b)). The bottom (*resp.*, top) disk, the *stationary* (*resp.*, *movable*) one, represents the plaintext (*resp.*, ciphertext) letters and is fixed (*resp.*, rotatable). The user rotates the movable disk so that the keyword letter aligns with the letter `A` of the stationary disk. Then, the corresponding plaintext and ciphertext letters align together. Figure 2(c) has the Saint Cyr Slide. The upper part is fixed while the lower part can be moved left or right. The table rows and columns, the movable disk and the bottom slide change according to the triplet of plaintext letter, keyword letter and ciphertext letter. Thus, the instructor has a demo tool for classroom use and the students have a clear view of how the Vigenère cipher performs encryption and decryption.

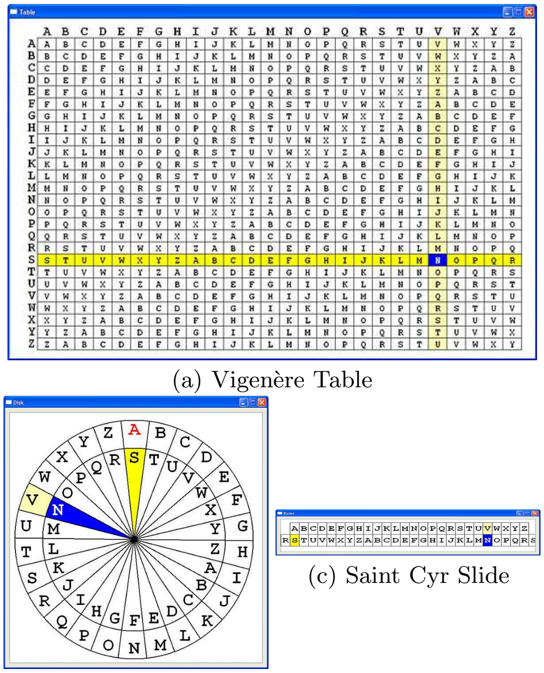


Table
Disk
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
S T U V W X Y Z A B C D E F G H I J K L M N O P Q R
Ruler

	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
A	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
B	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	A
C	C	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
D	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
E	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	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	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

(a) Vigenère Table

	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	
R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S

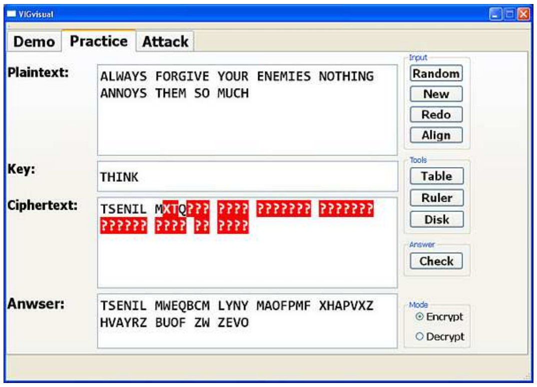
(c) Saint Cyr Slide

(b) Cipher Disk

Figure 2: Cipher Tools

3.2 The Practice Mode

Click the `Practice` tab to enter the `Practice` mode (Figure 3). All three tools are available. The user clicks `Encrypt` or `Decrypt` to start a new session, and uses `Random` to generate a random plaintext-keyword or ciphertext-keyword pair, `New` to start a new session, `Redo` to redo the current session, and `Align` to align the plaintext or ciphertext using the keyword length. Then, the user enters a keyword (if `Random` is not chosen) followed by the expected ciphertext or plaintext. The user may stop at anywhere and click the `Check` button to check the result. Incorrect letters are marked in red or with question marks if the positions are left as blank. The `Answer` field shows the correct answer.



Plaintext: ALWAYS FORGIVE YOUR ENEMIES NOTHING ANNOYS THEM SO MUCH

Key: THINK

Ciphertext: TSENIL MXTQ??? ???? ??????? ??????? ?????? ???? ?? ????

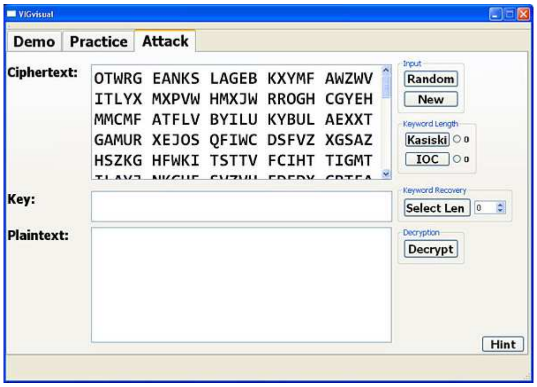
Anwser: TSENIL MWEQBCM LYNY MAOFPMF XHAPVXZ HVAYRZ BUOF ZW ZEVO

Mode: ● Encrypt ○ Decrypt

Figure 3: Screenshot of the Practice Mode

3.3 The Attack Mode

Click the `Attack` button to enter the `Attack` mode. The user clicks `Random` to randomly generate a ciphertext or `New` to start a new session and enter a new ciphertext (Figure 4). The process of breaking a message has two steps: keyword length estimation and keyword recovery. This process may not always be successful, and several iterations may be needed. VIGvisual has several `Hint` buttons in all windows under the `Attack` tab, each of which brings up a hint window to either explain what the window and/or algorithm is or provide a chance for the user to do simple exercises.



Ciphertext:

```
OTWRG EANKS LAGEB KXYMF AWZWV
ITLYX MXPVW HMXJW RROGH CGYEH
MMCMF ATFLV BYILU KYBUL AEXXT
GAMUR XEJOS QFIWC DSFVZ XGSAZ
HSZKG HFWKI TSTTV FCIHT TIGMT
```

Key:

Plaintext:

Figure 4: Screenshot of the Attack Mode

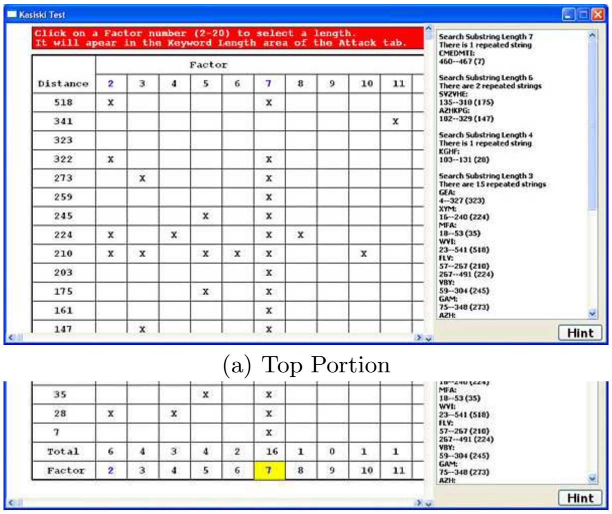
3.3.1 *Kasiski's Method*

VIGvisual provides two methods for keyword length estimation: Kasiski's method and the Index of Coincidence method. The user clicks `Kasiski` to use Kasiski's method. VIGvisual searches the given ciphertext for repeated substrings of length 3 (*i.e.*, trigraph) to 20, computes the distance between each pair of adjacent repeated substrings, finds the factors of this distance, and counts these factors.

Kasiski suggested that the factors that occur most often may be good estimates of the length of the keyword [7]. The `Kasiski` window shows a table in which each row has a distance value and the factors of this distance (Figure 5(a)). The bottom of this table (Figure 5(b)) shows the count of each found factor. The user clicks a factor, which will be shown in yellow, to select that length. This length appears to the right of the `Kasiski` button in the `Attack` tab. Repeated substrings are shown on the right panel of the `Kasiski` window along with their positions and distances. Clicking those repeated substrings can have them highlighted in the original ciphertext. Note that even though "`ab`" appears in "`abcd`" and "`abce`", only the longest repeated one "`abc`" is reported.

3.3.2 *The Index of Coincidence Method*

The concept of Index of Coincidence (IOC) was proposed by William F. Friedman in 1922 [3]. The IOC of a string is the probability of having two identical letters in that string. A typical English string without spaces and punctuation has an IOC around 0.068 while a random string of the 26 English letters has an IOC around 0.042. If a plaintext is encrypted by a single letter, the ciphertext is a shift of that letter, and its IOC is equal to that of the plaintext. Therefore, if keyword length is k, we may divide the ciphertext $C_1C_2C_3\cdots C_n$ into k cosets S_1, S_2, . . . , S_k where $S_i = C_iC_{i+k}C_{i+2k}\ldots$ $(1 \le i \le k)$, and each S_i is encrypted by the same letter with an IOC close to 0.068. To apply this idea, for each $1 \le k \le n$, we may divide the ciphertext into k cosets, calculate the IOC of each coset, and calculate the average of the k IOC values. If k is the correct length, each individual IOC is close to 0.068 and their average would also be close to 0.068. The estimated keyword length is the value of k that produces the highest average IOC.



Click on a Factor number (2-20) to select a length.
It will apear in the Keyword Length area of the Attack tab.

Distance	Factor 2	3	4	5	6	7	8	9	10	11
518	X					X				
341										X
323										
322	X					X				
273		X				X				
259						X				
245				X		X				
224	X		X			X	X			
210	X	X		X	X	X			X	
203						X				
175				X		X				
161						X				
147		X				X				

Search Substring Length 7
There is 1 repeated string
CMEDMTI:
460--467 (7)

Search Substring Length 6
There are 2 repeated strings
SVZVHE:
135--310 (175)
AZHKPG:
182--329 (147)

Search Substring Length 4
There is 1 repeated string
KGHF:
103--131 (28)

Search Substring Length 3
There are 15 repeated strings
GEA:
4--327 (323)
XYM:
16--240 (224)
MFA:
18--53 (35)
WVI:
23--541 (518)
FLV:
57--267 (210)
267--491 (224)
VBY:
59--304 (245)
GAM:
75--348 (273)
AZH:

(a) Top Portion

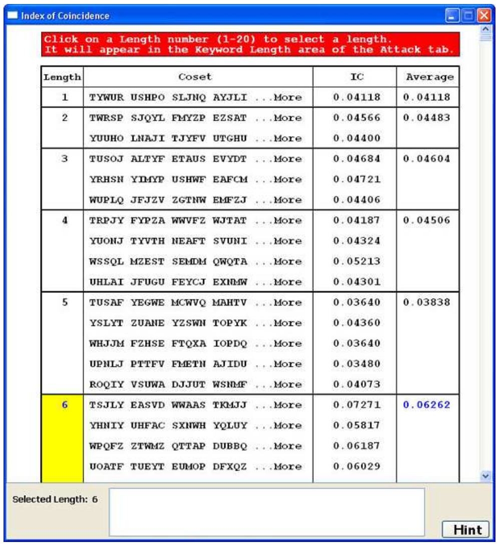
Distance	2	3	4	5	6	7	8	9	10	11
35				X		X				
28	X		X			X				
7						X				
Total	6	4	3	4	2	16	1	0	1	1
Factor	2	3	4	5	6	7	8	9	10	11

MFA:
18--53 (35)
WVI:
23--541 (518)
FLV:
57--267 (210)
267--491 (224)
VBY:
59--304 (245)
GAM:
75--348 (273)
AZH:

(b) Bottom Portion

Figure 5: Screenshot of the Kasiski Window

Click the `IOC` button to bring up the `IOC` window. VIGvisual displays a table (Figure 6). Each row corresponds to a possible keyword length in the range of 1 and 20, and the columns display the cosets, their IOC values, and the average. The highest three average IOC values are shown in blue. The user chooses a high average and clicks the length to export this value to the `Attack` window.



Click on a Length number (1-20) to select a length.
It will appear in the Keyword Length area of the Attack tab.

Length	Coset	IC	Average
1	TYWUR USHPO SLJNQ AYJLI ...More	0.04118	0.04118
2	TWRSP SJQYL FMYZP EZSAT ...More	0.04566	0.04483
	YUUHO LNAJI TJYFV UTGHU ...More	0.04400	
3	TUSOJ ALTYF ETAUS EVYDT ...More	0.04684	0.04604
	YRHSN YIMYP USHWF EAFCM ...More	0.04721	
	WUPLQ JFJZV ZGTNW EMFZJ ...More	0.04406	
4	TRPJY FYPZA WWVFZ WJTAT ...More	0.04187	0.04506
	YUONJ TYVTH NEAFT SVUNI ...More	0.04324	
	WSSQL MZEST SEMDM QWQTA ...More	0.05213	
	UHLAI JFUGU FEYCJ EXNMW ...More	0.04301	
5	TUSAF YEGWE MCWVQ MAHTV ...More	0.03640	0.03838
	YSLYT ZUANE YZSWN TOPYK ...More	0.04360	
	WHJJM FZHSE FTQXA IOPDQ ...More	0.03640	
	UPNLJ PTTFV FMETN AJIDU ...More	0.03480	
	ROQIY VSUWA DJJUT WSNMF ...More	0.04073	
6	TSJLY EASVD WWAAS TKMJJ ...More	0.07271	0.06262
	YHNIY UHFAC SXNWH YQLUY ...More	0.05817	
	WPQFZ ZTWMZ QTTAP DUBBQ ...More	0.06187	
	UOATF TUEYT EUMOP DFXQZ ...More	0.06029	

Selected Length: 6

Figure 6: Screenshot of the `IOC` Window

3.3.3 *Keyword Recovery*

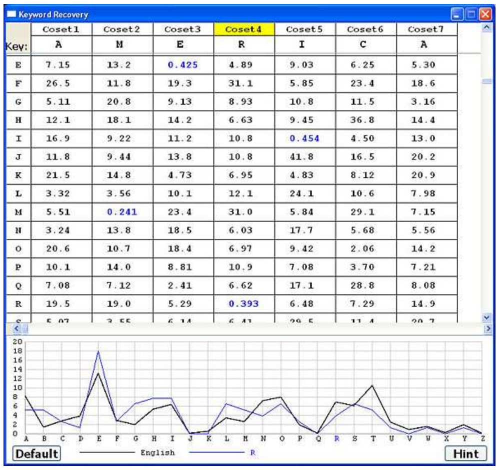
Once a length estimation is known, the user moves on to keyword recovery. The user clicks the circular button next to the estimated keyword length obtained by Kasiski's method or the IOC method to make it available in the `Select Len` field, or modifies this value to her choice. Then, the user clicks the Select Len button to recover the keyword.

VIGvisual uses the χ^2 method to recover the keyword. Assuming the estimated length k is correct, the ciphertext is divided into k cosets, each of which is encrypted by a single letter. Each coset is shifted one position to the right in a cyclic way. After each shift the letter frequency is computed and compared against the typical English letter frequency. Let F_i and f_i $(1 \leq i \leq 26)$ be the English letter frequency and calculated letter frequency of letter i, respectively. The χ^2 is defined as follows:

$$\chi^2 = \sum_{j=1}^{26} \frac{(f_i - F_i)^2}{F_i}$$

A lower χ^2 value means the letter frequency of a particular shift matches the English letter frequency better. Hence, the letter corresponding to the shift that yields the smallest χ^2 is very likely to be the correct letter in the keyword.

The Keyword Recovery window displays a table (Figure 7). Each column of this table corresponds to a keyword letter and has the χ^2 values of each shift with the smallest one in blue. The letter corresponding to the smallest χ^2 is shown in the column heading. The bottom of this window shows the English letter frequency graph in black. The user may click a coset to modify its keyword letter and a letter on the horizontal axis of the frequency graph to examine its frequency graph. The keyword changes accordingly and the frequency graph of that letter appears. The user may click on every letter and investigate the difference between the English letter frequency and the frequency of the selected letter, and pick the best match by examining all frequency graphs. This step is required as the shift corresponding to the smallest χ^2 value may not be the best choice, and shifts corresponding to other smaller χ^2 values must be examined.



	Coset1	Coset2	Coset3	Coset4	Coset5	Coset6	Coset7
Key:	A	M	E	R	I	C	A
E	7.15	13.2	0.425	4.89	9.03	6.25	5.30
F	26.5	11.8	19.3	31.1	5.85	23.4	18.6
G	5.11	20.8	9.13	8.93	10.8	11.5	3.16
H	12.1	18.1	14.2	6.63	9.45	36.8	14.4
I	16.9	9.22	11.2	10.8	0.454	4.50	13.0
J	11.8	9.44	13.8	10.8	41.8	16.5	20.2
K	21.5	14.8	4.73	6.95	4.83	8.12	20.9
L	3.32	3.56	10.1	12.1	24.1	10.6	7.98
M	5.51	0.241	23.4	31.0	5.84	29.1	7.15
N	3.24	13.8	18.5	6.03	17.7	5.68	5.56
O	20.6	10.7	18.4	6.97	9.42	2.06	14.2
P	10.1	14.0	8.81	10.9	7.08	3.70	7.21
Q	7.08	7.12	2.41	6.62	17.1	28.8	8.08
R	19.5	19.0	5.29	0.393	6.48	7.29	14.9
S	5.07	3.55	6.14	6.41	29.5	11.4	20.7

Figure 7: Screenshot of the Keyword Recovery Window

Any changes applied to the keyword will also be shown in the Keyword field in the Attack tab. Finally, the user clicks the Decrypt button to decrypt the ciphertext using the current recovered keyword and makes further changes to the keyword in the Keyword Recovery window if needed.

4. EVALUATION AND ASSESSMENT

Our survey consists of two parts, a set of 14 questions and eight write-in comments. Choices available are 5:strongly agree, 4:agree, 3:neutral, 2:disagree, and 1:strongly disagree. We collected 25 valid forms. The distribution of majors was as follows: 1 in computer network and system administration (CNSA), 5 in electrical and computer engineering (EE/CpE), 13 in computer science, 1 in mathematics (Math), 1 in chemical engineering, and 4 undeclared.

4.1 General Discussion

This paper uses $\alpha = 0.05$ as the level of significance for all statistical decisions. Our survey showed that 10 and 11 students used the table and slide for their work, and two chose to use the cipher disk. During the evaluation period, on average students used the tool for 13.2 minutes to understand the cipher with standard deviation 7.1 and confidence interval (6.1,20.3). Table 1 has the remaining questions.

Table 1: Survey Questions

Q_1	The Demo mode helped better understand
Q_2	The Practice mode helped better understand
Q_3	VIGvisual helped identify the parts that were not understood
Q_4	VIGvisual enhanced the course
Q_5	Kasiski helped understand the topic
Q_6	IOC helped understand the topic
Q_7	χ^2 helped understand the topic
Q_8	VIGvisual helped understand attack
Q_9	The Demo mode helped self-study
Q_{10}	The Practice mode helped self-study

Q_1 and Q_2 asked if the Demo mode and the Practice mode helped the students better understand the encryption and decryption processes. The means, standard deviations and confidence intervals were 3.88 and 3.6, 0.6 and 0.87, and (3.64,4.12) and (3.26,3.94). This reflected that the Practice mode was not as helpful as the Demo mode for this simple cipher. Q_3 asked the students if VIGvisual helped identify the parts of the Vigenère cipher that they did not understand. The mean, standard deviation and confidence interval were 3.8, 0.65 and (3.55,4.05). Thus, VIGvisual helped students learn the Vigenère cipher. As a result, the rating of Q_4, which asked if VIGvisual enhanced the course, is reasonably high with mean, standard deviation and confidence interval 4.24, 0.52 and (4.04,4.44).

Q_5 to Q_8 asked students to assess if the Kasiski, IOC, χ^2 and the Attack components helped them understand the topics. Table 2 has a summary. While the χ^2 question (Q_7) received a lower mean of 3.52, the answer to Q_8 has a higher mean of 4.24. This indicated that although the χ^2 component may not help students better learn than Kasiski and IOC do, the students actually understood more after using VIGvisual. Note that we cannot reject the means of Q_5 (3.96) and Q_6 (3.88) being 4 with p-values 0.832 and 0.503, respectively, and the mean of Q_7 being 3.75 with a p-value of 0.148.

4.2 Further Statistical Analysis

The ratings of questions were loosely related to each other. The highest correlation was 0.713 between Q_9 and Q_{10}, and the correlations for question pairs (Q_1, Q_2), (Q_3, Q_9), (Q_5 Q_6), (Q_6, Q_{10}), (Q_7, Q_9), (Q_7, Q_8) and (Q_8, Q_9)were all larger than 0.5. This suggested that ratings of Q_5 to Q_{10} had a positive trend. It is interesting to note that the correlations between Q_4 and other questions was mostly neutral,

Table 2: Ratings of the Kasiski, IOC and Attack Components

	Q_5 Kasiski	Q_6 IOC	Q_7 χ^2	Q_8 Attack
Mean	3.96	3.88	3.52	4.24
St Dev	0.93	0.88	0.77	0.52
CI^-	3.59	3.53	3.22	4.04
CI^+	4.33	4.23	3.82	4.44

Confidence Interval = (CI^-, CI^+)

as all the correlations were small, which indicated that the higher rating 4.24 of "if VIGvisual enhanced the course" was independent of ratings of other questions.

Since the student body consisted of several disciplines (*e.g.*, computer science, electrical and computer engineering, chemical engineering and mathematics), we would like to know if students from different disciplines reacted differently. We only grouped students into computer science (CS) and students not in computer science (non-CS). Since the questions may correlate with each other as mentioned earlier, the questions were also grouped into four groups: **(1)** Q_1, Q_2, Q_3 – the Demo, Practice modes were helpful and VIGvisual helped identify parts that were not understood, **(2)** Q_5, Q_6, Q_7 – the Kasiski, IOC and χ^2 components were helpful in general, **(3)** Q_9, Q_{10} – the Demo and Practice modes were useful in self-study, and **(4)** all questions are in a single group. MANOVA (Multivariate ANOVA) was applied to each of the question groups. In addition, we used ANOVA to investigate the difference between the student groups for Q_4 – if VIGvisual enhanced the course. While the assumptions of MANOVA (*i.e.*, normality and heteroscedasticity) are stronger than those of ANOVA, MANOVA is reasonably robust in our case [5, 6].

We used the general linear model (GLM) of R to perform all tests at $\alpha = 0.05$. The computed p-values of the four groups were 0.842, 0.805, 0.511 and 0.502. This suggested that the ratings from CS and non-CS groups did not vary significantly. The ANOVA result for Q_4 did not suggest difference at $\alpha = 0.05$ either. However, the p-value (0.068) is smaller for this question. In summary, we did not find a significant variation among disciplines and question groups.

4.3 A Test Scores Comparison

A quiz of six problems that address all aspects of the Vigenère cipher was given after the classroom lecture. Then, we discussed VIGvisual and made the software available. One week later a second quiz was given. This quiz has three questions similar to those in the first quiz and three ciphertexts for the students to practice cryptanalysis. The first ciphertext is trivial and can be broken directly using VIGvisual. The second is still easy, but VIGvisual yields a keyword with one incorrect letter. The third is not so trivial because VIGvisual yields a keyword with three incorrect letters, and the students must work a bit harder to break the ciphertext correctly. Both quizzes have a full score of 6 points (*i.e.*, one point per problem). We collected 33 papers from each quiz, and the results are shown in Table 3. The t-values of comparing the means obtained in various t-tests were all larger than 8.5 with p-values nearly 0, and Cohen's d is 2.18. This suggested that the difference between the means is significant and the effect size is large. As a result, we concluded that the software contributed to student learning significantly.

Table 3: Test Scores

	Quiz 1	Quiz 2
Mean	3.94	5.94
St Dev	1.25	0.35
CI	(3.5,4.37)	(5.8,6.0)

4.4 Student Comments

There were eight write-in questions asking students to make suggestions for further development. We focused on the following issues: whether the layout is useful, whether the Demo mode is more helpful than blackboard work, whether the Kasiski's method, IOC method and the Keyword Recovery component enhance learning, whether new features should be added, and software installation issues.

The layout was generally welcomed with comments like "*The layout is well done, the tab separation keeps everything organized and each section is clearly labeled*" and "*Everything was clearly findable. Never got confused*". Students agreed that the system functioned well and was easy to use.

The Demo mode vs. blackboard question received some interesting comments. Most students indicated VIGvisual is useful; but a few of them believed "in depth things come from the board" or the "procedures". Here is a list of student comments in various aspects: "*The program was much better than the blackboard since it was faster and easier to understand*", "*I prefer the more visual demonstration*" and "*The demo was better than the use of the chalk board since it was more organized and isolated.*".

Since Kasiski, IOC and keyword recovery components require a significant amount of information and require additional knowledge to be used properly, we expected students may encounter some difficulty in comprehending the system. However, most students were satisfied. The following are typical comments: "*It [Kasiski] was a little confusing due to the table layout, but it made sense after I figured it out*". "*The Kasiski Test tabs was my favorite function of* VIGvisual", "*It [IOC] helped me to understand cosets better*", "*Having the full [IOC] table enhanced my ability to understand because it showed more than just the answer*", "*If it is your first time seeing IOC, it will be hard. If you have a prior understanding, it makes sense*", "*The* Keyword Recovery *window made it easy to see why a keyword was likely to be correct*", and "*[Keyword Recovery] was a little hard to understand, but the* Hint *button helped a lot*".

As for new features, the most wanted one is resizable windows and more extensive hints and explanations. Very few students encountered problems installing the system. Those who had problems were mainly due to improperly installed libraries on Mac and some unknown issues on Windows 8.x. VIGvisual ran on Linux and Windows 7 and XP well.

4.5 Self-Study Investigation

We also invited students who did not take our course for a 2-stage self-study. This small scale survey was used to determine if there was a difference between classroom and self-study with our tool. There were two stages, each stage took about one week. In Stage 1, volunteers were asked to find resources to learn the Vigenère cipher, including Kasiski's and IOC methods. At the end of Stage 1, students evaluated their progress and completed six quiz problems.

In Stage 2, students were provided with VIGvisual and our web-based tutorial. At the end of this stage, students filled in the evaluation form and completed three quiz problems on cryptanalysis.

We collected seven completed survey forms from 11 volunteers. Volunteers were usually highly motivated, and, as a result, they received nearly perfect scores in both quizzes. They spent on average 14.29 minutes to understand the cipher with a confidence interval of (6.16, 22.41), which is similar to that of the students in our course. On the other hand, the median of the total time spent on the tool is about 41 minutes, which is much higher than that of the students in our course. We used median instead of average because there was a volunteer who used a very long time to practice cryptanalysis. Except for one volunteer who used the disk and slide, all others used the table.

Table 4: Self-Study Survey Results

	Q_1	Q_2	Q_3	Q'_4	Q_5	Q_6	Q_7	Q_8
Class μ	3.88	3.60	3.80	4.24	3.96	3.88	3.52	4.24
Class σ	0.60	0.87	0.65	0.52	0.93	0.88	0.77	0.52
μ	3.86	3.71	4.00	4.29	4.29	4.29	4.00	4.29
σ	0.69	0.76	1.29	0.76	0.76	0.95	0.89	1.11
CI^-	3.35	3.15	3.04	3.73	3.73	3.58	3.28	3.46
CI^+	4.37	4.27	4.96	4.85	4.85	4.99	4.72	5.11
p-value	0.95	0.75	0.71	0.88	0.36	0.33	0.23	0.97

μ: mean σ: standard deviation Confidence Interval: (CI^-,CI^+)

Since this survey was about self-study, questions in Table 1 related to classroom presentation were removed. A summary of this self-study results is given in Table 4 in which Q'_4 is the version of Q_4 for self-study. The "Class μ" and "Class σ" rows have the mean values and standard deviation obtained from our classroom survey (Section 4.1 and Table 2). No extensive hypothesis testings were performed because of small sample size. It is interesting to note that Q_1 to Q'_4 were rated similarly in both surveys. On the other hand, students involved in self-study rated our tool higher than those enrolled in our class. Regardless of the sample size issue, t-tests for comparing the means did not suggest any significant differences because the p-values in Table 4 are all larger than 0.1, suggesting no presumption against the null hypothesis (*i.e.*, the corresponding means being equal). Write-in comments were not very different from those obtained from the classroom survey. Hence, we have reasonable evidence to believe that the difference may be small.

5. CONCLUSIONS

This paper presented a visualization tool VIGvisual for teaching and learning the Vigenère cipher. With this tool, instructors are able to present all details of the cipher and a complete cryptanalysis procedure using Kasiski's and the IOC methods for keyword length estimation and the χ^2 method for keyword recovery. The animation and cipher tools help students see the "flow" of the cipher, learn the concepts and practice the cryptanalysis steps with VIGvisual. Evaluation results showed that VIGvisual was effective in the classroom presentation and for student self-study. In particular, after using the tools, the students learned cryptanalysis better and gained understanding of the cipher.

Based on the student comments, the most needed extensions are **(1)** resizable windows, **(2)** making the Vigenère table rows shaded in an alternating way so that it is more readable, **(3)** considering an extension or modification to the `Keyword Recovery` window so that the frequency graph can work alone rather than as part of the χ^2 method, **(4)** extending the error checking in the `Practice` mode so that errors can be reported on-the-fly, **(5)** adding the autocorrelation analysis for keyword estimation, and **(6)** developing a web-based version so that the system would be more "portable" as suggested by some students.

VIGvisual is a part of larger development of cryptography visualization tools supported by the National Science Foundation. In addition to VIGvisual, SHAvisual for the Secure Hash Algorithm, DESvisual for the DES cipher, AESvisual for the AES cipher, RSAvisual for RSA cipher, and ECvisual for the elliptic curve based ciphers are available online. Tools, evaluation forms, and installation and user guides for Linux, MacOS and Windows can be found at the following link, from which a complete tutorial of the Vigenère cipher, including cryptanalysis, is available:

`www.cs.mtu.edu/~shene/NSF-4`.

6. REFERENCES

[1] Cryptool. `http://www.cryptool.org`.

[2] M. E. Dalkilic and C. Gungor. An Interactive Cryptanalysis Algorithm for the Vigenere Cipher. In *Proceedings of the First International Conference on Advances in Information Systems*, pages 341–351, 2000.

[3] W. F. Friedman. *The Index of Coincidence and Its Applications in Cryptanalysis*. Riverbank Laboratories, 1922.

[4] J. Hoffstein, J. Pipher, and J. H. Silverman. *An Introduction to Mathematical Cryptography*. Springer, 2008.

[5] K. Ito. On the Effect of Heteroscedasticity and Nonnormality Upon Some Multivariate Test Procedures. In P. R. Krishnaiah, editor, *Multivariate Analysis II*. Academic Press, 1969.

[6] K. Ito and W. Schull. On the Robustness of the T^2 Test in Multivariate Analysis of Variance When Variance–Covariance Matrices Are Not Equal. *Biometrika*, 51:71–82, 1964.

[7] F. W. Kasiski. *Die Geheimschriften und die Dechiffrirkunst*. Mittler und Sohn, 1863.

[8] D. Salomon. *Data Privacy and Security*. Springer, 2003.

[9] D. Schweitzer and L. Baird. The Design and Use of Interactive Visualization Applets for Teaching Ciphers. In *Proceedings of IEEE Information Assurance Workshop*, pages 69–75, 2006.

[10] W. Trappe and L. C. Washington. *Introduction to Cryptography with Code Theory*. Prentice-Hall, 2002.

Acknowledgements

The authors are supported by the National Science Foundation under grants DUE-1140512, DUE-1245310 and IIS-1456763.