Spring 2015
Chris Christensen
Cryptology notes

Fractionation

Fractionation is a technique that can be used to diffuse plaintext frequencies over ciphertext.

Classically fractionation begins with a polygraphic substitution – a substitution of more than one ciphertext letter for a plaintext letter. Many fractionating ciphers as based upon the Polybius square.

Polybius Square

Polybius was a Greek historian and cryptographer of the second century BC. Polybius used a 5×5 square into which he inserted the 24 letters of the Greek alphabet. If we use the English alphabet of 26 letters, we must combine 2 of the letters into one cell – say, `i` and `j` (alternatively, in older versions of the square, `k` and `q` also seemed to be common cell-mates.)

	1	2	3	4	5
1	a	b	c	d	e
2	f	g	h	i/j	k
3	l	m	n	o	p
4	q	r	s	t	u
5	v	w	x	y	z

The square is then used for substitution -- to convert letters to numbers; e.g., `h` would be 23.

Polybius used a torch system by which an observer could determine the row and column of the letter being transmitted. The sender would stand on a

high location and hold a torch in each hand. To send an h (23), the sender would raise the right-hand torch twice followed by the left-hand torch three times.

World War I German ciphers

The most famous field cipher in all cryptology is a fractionating cipher -- the ADFGVX cipher. Fritz Nebel (1891 – 1967), a German radio staff officer, invented the cipher, and the German army began using an earlier version of it, the ADFGX cipher, on March 5, 1918, on the Western Front. The ADFGVX cipher involves both a polygraphic ubstitution and a transposition.

Fritz Nebel (1891 – 1967)
German cryptographer

ADFGX Cipher

Before the ADFGVX cipher, there was the - very similar -- ADFGX cipher. Like the ADFGVX cipher, it combined polygraphic substitution and transposition. The substitution portion of the ADFGX cipher was based upon a 5×5 Polybius square. For example:

	A	D	F	G	X
A	a	m	r	e	t
D	q	d	n	f	l
F	i/j	o	b	u	y
G	c	k	h	w	z
X	p	v	s	g	x

For example, GF would be substituted for h.

It is thought that the letters ADFGX (and later V) were chosen because their Morse code equivalents were dissimilar and, therefore, unlikely to be confused during transmission. And, it would, of course, be much easier and quicker for the Germans to train radio operators if the operators only needed to know the Morse code equivalents of six letters of the alphabet.

International Morse Code

A ●−−
D −●●
F ●●−●
G −−●
V ●●●−
X −●●−

Georges Painvin (1886 – 1980)
French Cryptanalyst

David Kahn describes the first encounter that the French had with the ADFGX cipher:

> When the first ADFGX messages were brought to [29-year-old French artillery captain and cryptanalyst Georges] Painvin, the best cryptanalyst in the Beau du Chiffre, he stared at them, ran a hand through his thick black hair with the air of perplexity, and then set to work. The presence of only five letters immediately suggested a checkerboard. Without much hope, he tried the messages as [simple substitution ciphers]: The tests were, as he expected, negative. He …

> was left with the hypothesis that the checkerboard substitution had been subjected to a transposition. On this basis he began work.
>
> Nothing happened. The traffic was too light for him to even determine by frequency counts whether the checkerboard key changed each day, and without this information he did not dare to amalgamate the cryptograms of successive days for a concerted assault. [Captain Francois] Cartier [the chief of the French military cryptologic bureau in World War I] looked on over his shoulder as he braided and unbraided letters and mused sadly, "Poor Painvin. This time I don't think you'll get it." Painvin, goaded, worked harder than before. …
>
> At 4:30 a.m. March 21, 6,000 guns suddenly fired upon the Allied line at the Somme in the most furious artillery cannonade of the war. Five hours later, 62 German divisions rolled forward on a 40-mile front. The surprise was complete and its success was over whelming.

The rapid German advance led to the French intercepting a large number of messages encrypted with the ADFGX cipher. Painvin had enough messages to find a pattern. Kahn in *The Codbreakers* (chapter *A War of Intercepts, II*) describes Painvin's cryptanalysis of the transposition and the checkerboard in some detail.

ADFGVX Cipher

By April and May, Painvin was having success against the ADFGX cipher. The transmission of May 29 took only two days to break, and the transmission on May 30 were broken the next day. Then, on June 1, 1918, the Germans replaced the ADFGX cipher by the ADFGVX cipher. This was just before German General Eric Ludendorff's (1865 – 1937) 1918 spring offensive. Again, from Kahn:

> … Painvin suddenly saw, on June 1, the ADFGX message complicated by the addition of a sixth letter, V. Probably the Germans expanded their checkerboard to 6×6. But why? … Painvin did not know.
>
> "In short," he said, "I had a moment of discouragement. The last two keys of the 28th and the 30th of May had been discovered under

> conditions of such rapidity that their exploitation was of the greatest usefulness. The offensive and the German advance still continued. It was of the greatest importance not to lose [cryptographic] contact and in my heart I did not want to brusquely shut off this source of information to the interested services of the armies which had become accustomed to counting on its latest results."

Painvin immediately spotted two messages that were almost identical, and, by tweezing those messages, he was able to solve enough of the ADFGVX dispatches that the German spring offensive of June 9 was halted several days later by the French.

ADFGVX cipher cryptography

As Painvin correctly assumed, the substitution portion of the ADFGVX cipher was based upon a 6×6 square, which permits inclusion of all 26 letters of the English alphabet and the ten digits 0, 1, … 9. Here is an example of an ADFGVX square:

	A	D	F	G	V	X
A	a	i	2	o	0	d
D	1(one)	b	h	6	m	s
F	t	n	w	c	q	4
G	l(el)	g	7	v	y	r
V	f	5	e	3	x	z
X	9	p	j	k	8	u

Typically, the symbols were randomly arranged in the square; so, it would be necessary to have a written key to encrypt and decrypt.

Just as for the ADFGX cipher, each letter and number of the plaintext message was substituted polygraphically -- with the two letters designating row and column of its positions in the square. For example,

```
t   h   e
FA  DF  VF
```

The message

```
This cipher features substitution and
transposition.
```

becomes

```
FADFADDXFGADXDDFVFGXVAVFAAFAXXGXVFDXDDAGFADFDXX
XDDDXFAADFAXXFAADAGFDAAFDAXFAGXAAFDDXXDAGDXADFA
ADAGFD
```

Notice that in the substitution portion of the encryption, the length of the message is doubled. (This can be a serious problem for long messages.)

After the substitution portion of the encryption, the message is encrypted again by a columnar transposition.

To do the transposition, a keyword is needed. Let us use *Sinkov* (one of William Friedman's first junior cryptanalysts). The message is re-written in six columns underneath the six letters of the keyword: “In by rows.”

S	I	N	K	O	V
F	A	D	F	A	D
D	X	F	G	A	D
X	D	D	F	V	F
G	X	V	A	V	F
A	A	F	A	X	X
G	X	V	F	D	X
D	D	A	G	F	A
D	F	D	X	X	X
D	D	D	X	F	A
A	D	F	A	X	X
F	A	A	D	A	G
F	D	A	A	F	D
A	X	F	A	G	X
A	A	F	D	D	X
X	D	A	G	D	X
A	D	F	A	A	D
A	G	F	D	V	X

There are 50 letters in the message; this doubles to 100. Because there are 6 columns and 17 rows in the table ($6 \times 17 = 102$) two nulls (a V and an X) have been added to the end of the message to complete the columns.

Now the columns of the table are rearranged by alphabetizing the letters of the keyword.

I	K	N	O	S	V
A	F	D	A	F	D
X	G	F	A	D	D
D	F	D	V	X	F
X	A	V	V	G	F
A	A	F	X	A	X
X	F	V	D	G	X
D	G	A	F	D	A
F	X	D	X	D	X
D	X	D	F	D	A
D	A	F	X	A	X
A	D	A	A	F	G
D	A	A	F	F	D
X	A	F	G	A	X
A	D	F	D	A	X
D	G	A	D	X	V
D	A	F	A	A	D
G	D	F	V	A	X

Now the string of ciphertext is created by going down columns: "Out by columns."

AXDXAXDFDDADXADDGFGFAAFGXXADAADGADDFDVFVADDFAAFFAFF
AAVVXDFXFXAFGDDAVFDXGAGDDDAFFAAXAADDFFXXAXAXGDXXVDX

Typically the message was transmitted in five-letter blocks.

AXDXA XDFDD ADXAD DGFGF AAFGX XADAA DGADD FDVFV
ADDFA AFFAF FAAVV XDFXF XAFGD DAVFD XGAGD DDAFF
AAXAA DDFFX XAXAX GDXXV DX

Notice that each ciphertext letter depends on 2 plaintext letters.

Cryptanalysis

The ADFGVX cipher is not hard to spot, but this is a difficult cipher to break. Let us consider the problems that are faced when cryptanalyzing the ADFGVX cipher.

First, consider a message that was encrypted once with a Caesar cipher and then again with columnar transposition. Frequency analysis will show a shifted alphabet, but, after shifting the ciphertext letters back by the key that was determined from frequency analysis, the message will still not be plaintext. Frequency analysis would indicate plaintext, but the message would not be plaintext. That indicates a transposition cipher, and the remaining ciphertext would be attacked as a transposition cipher.

Similarly, consider a message that was encrypted once with a simple substitution cipher with a randomly generated permutation and then again with a columnar transposition. Frequency analysis will show a simple substitution cipher. The cryptanalyst might begin by assuming that the most frequent ciphertext letter corresponds to plaintext e, etc., but the transposition would complicate this analysis.

The ADFGVX cipher is even worse. If only the substitution portion of the cipher were done, it would not be difficult to break. Even if the cryptanalyst knew nothing about the ADFGVX cipher, the fact that only six letters appear in the ciphertext and every ciphertext message has even length would suggest that a 6×6 checkerboard was used for substitution. Suspecting this, the message could be broken into digraphs and the frequency of the digraphs could be analyzed to determine the corresponding plaintext letters as is done for any simple substitution cipher. The devil is the transposition. Every plaintext letter is substituted by a digraph. When placed into the rectangular array for columnar transposition, the first letter of the digraph will lie in one column and the second letter of the digraph will lie in another column. After transposition, these letters will be separated – the plaintext frequencies will be "fractionated." This is the strength of the ADFGVX cipher and similar ciphers. Single-letter characteristics are scattered; single letter frequencies are diffused over the ciphertext.

In *The Codebreakers*, Kahn describes Painvin's difficult solution to a few of the ADFGVX messages. Painvin's partial (but sufficient) success was

followed by "a long leave of convalescence." (And, subsequently by an "immensely successful business career.") But, Painvin said of his solution of the ADFGVX ciphers that they left "an indelible mark on my spirit, and remain for me one of the brightest and most outstanding memories of my existence."

An excellent discussion of cryptanalysis of the ADFGX cipher appears in Craig Bauer's *Secret History*, p. 190 *f.*

Bifid and Trifid ciphers

Two famous fractionating ciphers were constructed around 1901 by the French cryptographer Felix Delastelle.

Bifid cipher

Like the ADFGX cipher, the bifid cipher uses a Polybius square for polygraphic substitution. A Polybius square is constructed with numbered rows and columns. The key is a 5×5 square with a mixed alphabet.

	1	2	3	4	5
1	n	o	r	s	e
2	a	b	c	d	f
3	g	h	i/j	k	l
4	m	p	q	t	u
5	v	w	x	y	z

The plaintext message `polybiussquare` is converted to its coordinates, but the coordinates are written vertically.

p	o	l	y	b	i	u	s	s	q	u	a	r	e
4	1	3	5	2	3	4	1	1	4	4	2	1	1
2	2	5	4	2	3	5	4	4	3	5	1	3	5

The substitutions are then read off in rows – the first row followed by the second row.

4135234114421122542354435135

The string is divided into pairs.

41 35 23 41 14 42 11 22 54 23 54 43 51 35

And, the pairs are converted back to letters using the square.

41	35	23	41	14	42	11	22	54	23	54	43	51	35
M	L	C	M	S	P	N	B	Y	C	Y	Q	V	L

The ciphertext is

MLCMSPNBYCYQVL

Notice that each ciphertext letter depends upon 2 plaintext letters; so, input frequencies are diffused in the ciphertext.

Trifid cipher

Delastelle's trifid cipher is similar except that it is "three-dimensional." The key consists of three 3×3 tables with a mixed alphabet distributed among the table. Because there are 27 cells, an additional symbol, in this case the period, can be included in the key,

Table 1

	1	2	3
1	D	A	M
2	X	T	N
3	O	G	I

Table 2

	1	2	3
1	P	R	.
2	W	U	F
3	C	Y	Q

Table 3

	1	2	3
1	H	Z	S
2	B	V	L
3	E	J	K

The plaintext message `polybiussquare` is converted to its coordinates, e.g., Table – Row - Column, and the coordinates are written vertically.

p	o	l	y	b	i	u	s	s	q	u	a	r	e
2	1	3	2	3	1	2	3	3	2	2	1	2	3
1	3	2	3	2	3	2	1	1	3	2	1	1	3
1	1	3	2	1	3	2	3	3	3	2	2	2	1

The coordinates are then read out in rows – the top row, followed by the middle row, followed by the bottom row.

213231233221231323232113211311321323332221

The string is divided into triples.

213 231 233 221 231 323 232 113 211 311 321 323 332 221

And the triples are converted back to letters using the table.

213	231	233	221	231	323	232	113	211	311	321	323	332	221
.	C	Q	W	C	L	Y	M	P	H	B	L	J	W

The ciphertext is

`.CQWCLYMPHBLJW`

Each ciphertext letter depends on 3 plaintext letters.

Exercises

1. Use the Polybius square

	1	2	3	4	5
1	a	b	c	d	e
2	f	g	h	i/j	k
3	l	m	n	o	p
4	q	r	s	t	u
5	v	w	x	y	z

to convert the message `Northern Kentucky University` to a string of numbers.

2. Use the ADFGVX square

	A	D	F	G	V	X
A	a	i	2	o	0	d
D	1(one)	b	h	6	m	s
F	t	n	w	c	q	4
G	l(el)	g	7	v	y	r
V	f	5	e	3	x	z
X	9	p	j	k	8	u

and the keyword *mathematician* to encrypt the message

```
The First World War marks the great turning point
in the history of cryptology.
```

3. Decrypt the following message that was enciphered using the ADFGVX square given in exercise 2 and the keyword *Sinkov*.

GGDFD VXGDF DAGXG VAAAA FVGVF FFDFD ADAFD D

4. Decrypt the following message that was encrypted using the ADFGVX square given in exercise 2 and the keyword *Kentucky*.

AGVXF GXGAX XDFFD XXXGD FAGAF DAXGA AFGDV DGDVV