Fall 2019
Chris Christensen
MAT/CSC 483
Section 12

# ADFGVX Cipher

The most famous field cipher in all cryptology is 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 composes a substitution and a transposition. But it introduces a new method – fractionation.

## Polybius Square

The **substitution** portion of the cipher is based on the Polybius square.

Polybius was a Greek historian and cryptographer of the second century BC. Polybius used a $5 \times 5$ 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.

An audible version of the square, called the "knock cipher," has also been used by prisoners over the centuries to transmit messages from cell to cell. For example, to transmit the letter `h`, the prisoner would

knock knock – knock knock knock

on the walls of the cell. In Czarist Russia, the knock cipher was based upon a $6 \times 6$ square containing the 33 letters of the Russian alphabet. Alexander Solzhenitsyn refers to prisoners using the knock-knock cipher in *The Gulag Archipelago*.

> … back in the twenties, prison authorities had been very lenient toward prisoners communicating with each other by knocking on the walls: this was a carry-over from the stupid tradition in the Tsarist prisons that if the prisoners were deprived of knocking, they would have no way to occupy their time. Part I, page 189.
>
> (… we didn't feel up to knocking out a message on the wall. This was punished very severely.) Part I, page 205.
>
> What kind of struggle is it over the question of whether cells are kept locked and whether prisoners, to exercise their right to communicate, can openly spell out messages to each other by knocking … ? Part I, p. 461.

Russian anarchists brought the knock-knock cipher to Western Europe.

ADFGX Cipher

Before the ADFGVX cipher, there was the ADFGX cipher. Like the ADFGVX cipher, it composed substitution and transposition. The substitution portion of the ADFGX cipher was based upon a $5 \times 5$ 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` (row coordinate-column coordinate) 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 | –●●– |

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 Codebreakers* ("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 $6 \times 6$. 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 Cryptography

As Painvin correctly assumed, the substitution portion of the ADFGVX cipher was based upon a $6 \times 6$ 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 with the two letters designating row and column coordinates of its positions in the square. For example,

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

The message

This ciph**e**r features substitution and transposition.

becomes

FADFADDXFGADXDDF**VF**GXVAVFAAFAXXGXVFDXDDAGFADFDXX
XDDDXFAADFAXXFAADAGFDAAFDAXFAGXAAFDDXXDAGDXADFA
ADAGFD

Notice the bolded **VF**. This substitution corresponds to the bolded **e** in the plaintext. We will track this digraph through the encryption process.

Notice that in the substitution portion of the encryption, the length of the message is doubled.

After the substitution portion of the encryption, the message is re-encrypted using 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 ciphertext is removed "out by columns" in the order that results from considering the letters of the keyword in alphabetical order.

Here are the columns arranged in the order in which they are removed:

| 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 |

Notice that the digraph that we are tracking **VF** (the digraph that is the substitution for e) is split – fractionated.

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

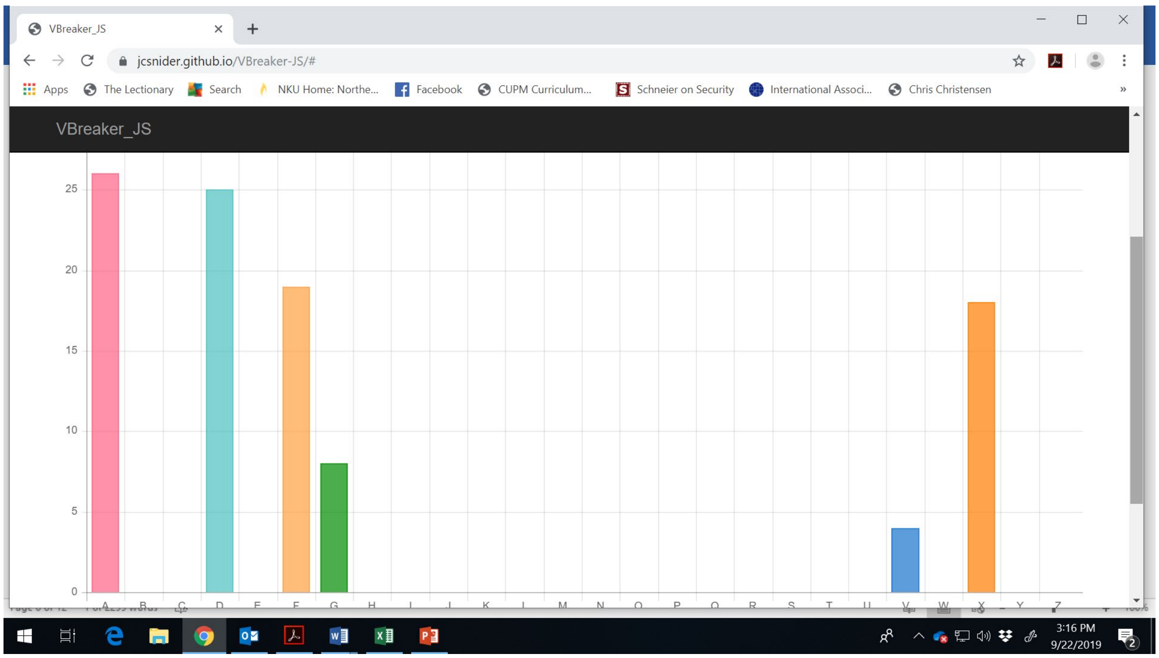
AXDXAXDFDDADXADDGFGFAAFGXXADAADGADDFDVFVADDFAAFFAFF
AAV**V**XDFXFXAFGDDAVFDXGAGDDDAFFAAXAADDF**F**XXAXAXGDXXVDX

Typically the message was transmitted in five-letter blocks.

AXDXA XDFDD ADXAD DGFGF AAFGX XADAA DGADD FDVFV
ADDFA AFFAF FAAV**V** XDFXF XAFGD DAVFD XGAGD DDAFF
AAXAA DDF**F**X XAXAX GDXXV DX

The digraph **VF** that was substituted for **e** has been split – fractionated. To break the message, the digraph must be reassembled.

Here are the frequencies:



## Cryptanalysis

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

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 $6 \times 6$ 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. This is a strength of the ADFGVX cipher and similar ciphers. Single-letter characteristics are split and scattered. Ciphertext letters cannot be associated with particular plaintext letters.

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 summary of one method to break ADFGVX is given in *Introduction to Cryptography with Coding Theory* by Wade Trappe and Lawrence C. Washington.

> Suppose two different ciphertexts intercepted at approximately the same time agree for the first several characters. A reasonable guess is that the two plaintexts agree for several words. That means that the top few entries of the columns for one are the same as for the other. Search through the ciphertexts and find places where they agree.

These possible represent the beginnings of the columns. If this is correct, we know the column lengths. Divide the ciphertexts into columns using these lengths. For the first ciphertext, some columns will have one length and others will be one longer. The longer ones represent columns that should be near the beginning; the other columns should be near the end. Repeat for the second ciphertext. If a column is long for both ciphertexts, it is very near the beginning. If a column is long for one ciphertext and not for the other, it goes in the middle. If it is short for both, it is near the end. At this point, try various orderings of the columns, subject to these restrictions. Each ordering corresponds to a potential substitution cipher. Use frequency analysis to solve these. One should yield the plaintext and the initial encryption matrix.

Frank Rowlett

Frank Rowlett (1908 – 1998) was one of the first cryptanalysts hired by William Friedman for the Signal Intelligence Service (SIS) of the Army Signal Corp in 1930. SIS was the successor to Herbert Yardley's (1889 – 1958) Black Chamber.

Frank Rowlett
National Cryptological Musuem
Hall of Honor

Rowlett, in his book *The Story of Magic*, [Magic was the name for decryptions produced by the Purple analog. Friedman called his staff "magicians."] describes his encounter with ADFGVX.

> [Friedman] sketched out for us the work that had been done in France during the last war on the German field army cipher known as the ADFGVX cipher. He explained that during the war no satisfactory general solution to the system had been devised. Only occasionally could the Allied cryptanalysts recover one of the keys used by the Germans, and this could only be achieved under very special circumstances. He identified this type of solution as the "Painvain Solution," named after it inventor Captain Georges Painvain, who was a French Army cryptanalyst [Friedman] had worked with during World War I, a man whose cryptanalytic ability was held in high esteem by both the French and the Americans.
>
> The initial break into the system is described, starting on page 215, in *Precis de Cryptographie Moderne* by the French cryptographer Charles Eyraud published in 1953. According to Eyraud, this break

> came during early April 1918 when the French intercepted two messages, each of three parts, which they soon determined to be retransmissions of the same basic information, since there were only slight differences in the ciphertexts of the two messages. One of the French cryptographers, M. Painvain, was able to exploit this case of retransmission and recovered both the transposition key and the digraphic substitution employed in the messages. With the keys recovered from these two messages, the French cryptanalysts were able to read all the other messages sent during the same period.
>
> Although the French cryptanalysts had recovered the full details of the system from this initial break, they were unable to solve the keys for other days, except when they were fortunate enough to intercept two messages in which substantial repetitions occurred.

Craig Bauer in *Secret History* describes in detail the cryptanalysis of an ADFGX cipher in section 6.3. William Friedman describes how to solve an ADFGVX cipher in "General Solution for the ADFGVX Cipher System."

## Fractionation

In the case of the ADFGVX cipher, the substitution portion of the cipher doubles the length of the text; that multiliteral substitution permits fractionation to occur. In digital communications, typically plaintext characters are substituted by 8-bit ASCII codes, and fractionation of the bits then can occur.

| N | K | U |
|---|---|---|
| 01001110 | 01001011 | 01010101 |