Secret Key Cryptography (Spring 2006)

Instructor: Adi Shamir Teaching assistant: Eran Tromer

Appendix to Lecture 2:
Rejewski's method for recovering the Enigma message password

The example in Garlinksi's book (partially shown in class) is partial and inaccurate. The following works out the missing details. For another account (containing all the notation but also missing details of the analysis), see Rejewski's paper on the course webpage.

Preparation

By observing how the 1st ciphertext letter is mapped to the 4th ciphertext letter in the day's transmissions, we obtain the permutation AD and observe that it has the following cycle structure:

$$AD = (\mathtt{A})(\mathtt{S})(\mathtt{BC})(\mathtt{RW})\cdots$$

Similarly, looking at the 2nd and 5th ciphertext letters we find the permutation BE:

$$BE = (\mathtt{D})(\mathtt{K})(\mathtt{AXT})(\mathtt{CGY})(\mathtt{BLFQVEOUM})(\mathtt{HJPSWIZRN})$$

And looking at the 3nd and 6th ciphertext letters we can find the permutation CF:

$$CF = (\mathtt{ABVIKTJGFCQNY})(\mathtt{DUZREHLXWPSMO})$$

Looking at the captured ciphertext prefixes, we note that the prefix `SYXSCW` appears several times in adjacent captured transmissions, implying a lazy Enigma operator that used the same message password. We guess (based on past experience) that this message password is the "cilly" password `AAA`. Hence, by definition of A, B, C, D, E, F:

$$A(\mathtt{A}) = \mathtt{S},\ B(\mathtt{A}) = \mathtt{Y},\ C(\mathtt{A}) = \mathtt{X},\ D(\mathtt{A}) = \mathtt{S},\ E(\mathtt{A}) = \mathtt{C},\ F(\mathtt{A}) = \mathtt{W}$$

Analyzing AD

Since the only cycles of size 1 in AD are $(\mathtt{A})$ and $(\mathtt{S})$, it must hold that both A and D contain the cycle $(\mathtt{AS})$:

$$A = (\mathtt{AS})\cdots\ , \quad D = (\mathtt{AS})\cdots$$

This is consistent with the cilly ($A(\mathtt{A}) = \mathtt{S}$ and $D(\mathtt{A}) = \mathtt{S}$), and thus lends some credence to our guess of the cilly. But we don't learn anything else yet.

Analyzing BE

We know $B(\mathtt{A}) = \mathtt{Y}$ and $BE(\mathtt{A}) = \mathtt{X}$, hence $E(\mathtt{Y}) = E(B(\mathtt{A})) = BE(\mathtt{A}) = \mathtt{X}$.[1] From this and $BE(\mathtt{G}) = \mathtt{Y}$ we deduce $B(\mathtt{X}) = B(E(\mathtt{Y})) = EB(\mathtt{Y}) = E^{-1}B^{-1}(\mathtt{Y}) = (BE)^{-1}(\mathtt{Y}) = \mathtt{G}$. From this and $BE(\mathtt{X}) = \mathtt{T}$ we deduce $E(\mathtt{G}) = E(B(\mathtt{X})) = BE(\mathtt{X}) = T$. From the latter and $BE(\mathtt{C}) = \mathtt{G}$ we deduce $B(\mathtt{T}) = \mathtt{C}$, and finally using $BE(\mathtt{T}) = \mathtt{A}$ we deduce $E(\mathtt{C}) = \mathtt{A}$. Overall, we have deduced the following cycle of mappings:

$$\rightsquigarrow \mathtt{A} \xrightarrow{B} \mathtt{Y} \xrightarrow{E} \mathtt{X} \xrightarrow{B} \mathtt{G} \xrightarrow{E} \mathtt{T} \xrightarrow{B} \mathtt{C} \xrightarrow{E} \rightsquigarrow$$

Hence, we've fully derived the values of B and E on all letters involved in the 3-cycles of BE. If you've solved Exercise 1 by direct argument, you probably see that this is not a coincidence — what we've done amounts to the following.

[1] Our notation is $PQ(\alpha) = Q(P(\alpha))$.

The only cycles of size 3 in BE are (AXT) and (CYS). By a side effect of the (natural) proof for Exercise 1, we know that B and E transpose letters in the first cycle with letters in the second cycle, forming a cyclic mapping chain of the above form. Moreover, the letters of the two cycles alternate along the chain, with the letters of one cycle of BE appearing in the same order and the letters of the other cycle appearing in reversed order — up to cycle shift . The only uncertainty is due to this cyclic shift, and the cilly gave us the datum $B(\mathtt{A}) = \mathtt{Y}$ which determined the shift. Hence we got all relevant values of B and E.

Analyzing CF

Given $C(\mathtt{A}) = \mathtt{X}$ we can work out the deductions as done above for BE. However, it is much easier to note that CF consists of two known cycles of size 13, giving a cyclic chain of mappings in which the two cycles alternative; the relative shift of the cycles in the chain is determined by $C(\mathtt{A}) = \mathtt{X}$. The resulting cyclic chain of mappings is as follows:

$$\begin{aligned}\rightsquigarrow \mathtt{A} \xrightarrow{B} \mathtt{X} \xrightarrow{E} \mathtt{B} \xrightarrow{B} \mathtt{L} \xrightarrow{E} \mathtt{V} \xrightarrow{B} \mathtt{H} \xrightarrow{E} \mathtt{I} \xrightarrow{B} \mathtt{E} \xrightarrow{E} \mathtt{K} \xrightarrow{B} \mathtt{R} \xrightarrow{E} \mathtt{T} \xrightarrow{B} \mathtt{Z} \xrightarrow{E} \mathtt{J} \xrightarrow{B} \\ \mathtt{U} \xrightarrow{E} \mathtt{G} \xrightarrow{B} \mathtt{D} \xrightarrow{E} \mathtt{F} \xrightarrow{B} \mathtt{O} \xrightarrow{E} \mathtt{C} \xrightarrow{B} \mathtt{M} \xrightarrow{E} \mathtt{Q} \xrightarrow{B} \mathtt{S} \xrightarrow{E} \mathtt{N} \xrightarrow{B} \mathtt{P} \xrightarrow{E} \mathtt{Y} \xrightarrow{B} \mathtt{W} \xrightarrow{E} \rightsquigarrow\end{aligned}$$

Completing the cryptanalysis

Using one good guessed cilly we have fully recovered the C and F permutations, and also some of of the B and E permutations. We can thus decrypt the 3rd message key letter, and sometimes also the 2nd letter, throughout the day. Going over the received ciphertexts, we find one in which the partial decryption of the first 6 letters is ?CC?CC. We guess that this is the cilly CCC, and this lets us analyze the cycles (BC)(RW) in AD and thereby deduce the values of A and D on $\mathtt{B}, \mathtt{C}, \mathtt{R}, \mathtt{W}$. Given this new knowledge we keep looking for cillies and making deductions about A, B, C, D, E, F. If we ever get inconsistent results, it means we mis-guessed a cilly, so we backtrack and try a different guess. Eventually we fully learn the permutations, which lets us decrypt the message key of *any* message received in the day analyzed.