Cryptologia
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ucry20

Rejewski's Catalog

Alex Kuhl

Available online: 05 Oct 2007

To cite this article: Alex Kuhl (2007): Rejewski's Catalog, Cryptologia, 31:4, 326-331

To link to this article: http://dx.doi.org/10.1080/01611190701299487

Cryptologia, 31:326–331, 2007

ISSN: 0161-1194 print
DOI: 10.1080/01611190701299487

Taylor & Francis
Taylor & Francis Group

Rejewski's Catalog

ALEX KUHL

Abstract When attacking the German Enigma cipher machine during the 1930s, the Polish mathematician Marian Rejewski developed a catalog of disjoint cycles of permutations generated by Enigma indicators. By comparing patterns that resulted from message indicators with his catalog, Rejewski was able to determine the ground settings. Well, not quite—the mapping from the disjoint cycles to the ground settings is not one-to-one. Rejewski's catalog no longer exists. This article reports on the output of a program that "recreates" the catalog and answers the question "How far from being one-to-one is the mapping?"

Keywords enigma, Marian Rejewski, permutation groups

1. Introduction

In 1924, in anticipation of yet another war with Germany, the Polish cipher bureau—the Biuro Szyfów—conducted a cryptology course at Poznan University. Among the students who were recruited for the course was the 23-year-old mathematics student Marian Rejewski (1905–1980). The goal of the course was to "help Polish radio intelligence against its difficult German adversary [3]." In 1932, Rejewski solved the German Enigma machine—the primary German encryption-decryption device.

At its heart, Enigma is a sequence of turning rotors and a fixed reflector that changes plaintext to ciphertext. Each Enigma cipher is a permutation of the letters of the alphabet. Each plaintext letter is encrypted with a different cipher than the previous one because of the motion of the rotors. The rotors turn in a regular fashion: each time a letter is pressed on the keyboard the right-most rotor turns forward one position; the middle rotor turns forward one position during one complete revolution of the right rotor; and the left rotor turns forward one position during one complete revolution of the middle rotor. The motion of the rotors is similar to the dials of an odometer of an automobile.

Part of the message key is the order and the position of the rotors. In 1932, the Enigma that Rejewski attacked had three rotors that were placed into three positions in the rotor system. The three rotors can be placed into the rotor system in any of the $3! = 6$ possible orders. Around the circumference of each rotor are the 26 letters of the alphabet (or sometimes 26 numbers). Each rotor can be set at any of these 26 positions. Thus, Rejewski faced $6 \times 26^3 = 105{,}456$ possible settings of the rotor system.

An additional factor in the encryption is a plugboard. On the front of Enigma is

Address correspondence to Alex Kuhl, Computer Science Department, North Carolina State University, 890 Oval Drive, Box 8206, Raleigh, NC 27695-8206, USA. E-mail: arkuhl@ncsu.edu

a plugboard that looks like an old telephone switchboard. There are 26 sockets—one for each letter of the keyboard. In 1932, the Enigma had six plugs that were used to connect six pairs of letters. The effect of the plugboard is to swap six pairs of letters and let the remaining 14 letters pass through unchanged. There are 100,391,791,500 ways to connect six pairs of letters. The remaining part of the Enigma key is the setting of the plugboard. There are $105{,}456 \times 100{,}391{,}791{,}500$ ways to set up Enigma to encrypt a message. The sheer number of settings that a cryptanalyst might have to try is what gives Enigma its strength.

The rotor system and the plugboard combine to determine the Enigma permutation. Every time a letter is pressed on the keyboard, the right rotor turns. Therefore, each plaintext letter is encrypted by a different permutation—almost. After $26^3 = 17{,}576$ letters are encrypted, Enigma would return to its starting position, and the ciphertext message would begin to show depth.

To avoid depth, messages were generally quite short and the operator would "randomly" choose a message setting—one of the 26 possible starting positions for each of the three rotors—for each message. The plugboard set up was determined by a codebook. In addition, the codebook specified a "ground setting" for the rotor system. The ground setting was in effect for a day. The sending operator needed to inform the receiving operator of the three-letter message setting. The message setting was encrypted by the ground setting and transmitted at the beginning of the message. Because of the possibility of interference during transmission, the message setting was sent twice. The three-letter message setting that was encrypted twice using the ground setting appeared at the beginning of the Enigma message.

2. Rejewski's Attack

Rejewski discovered a pattern in the twice-encrypted message setting. He exploited this pattern to determine the ground settings of the rotor system. Rejewski labeled the six ciphers that were used to encrypt the message setting A, B, C, D, E, and F. Permutations A and D would both encrypt the first letter of the message setting, B and D the second letter, and C and F the third letter. Rejewski discovered that the disjoint cycle decompositions of the composed permutations AD, BE, and CF (composition from left to right) could be used to determine the ground settings.

Each Enigma permutation, for example A, B, C, D, E, or F, is the product of 13 transpositions—two cycles, swaps of two letters. It is this fact that permits Enigma to encrypt and decrypt a message using the same setting. If, for example, n changes to K for a particular setting of Enigma, then K changes to n at the same setting. Rejewski noted that this fact implied that the permutations, AD, BE, and CF consist of "disjunctive cycles of the same length in even numbers" [Marian Rejewski, [3] Appendix E]. Without this restriction, the theoretically possible number of disjoint cycle structures for each of AD, BE, and CF would be the number of partitions of 26. However, because the number of disjoint cycles of the same length is even, the theoretically possible number of disjoint cycle structures is the number of partitions of 13, namely 101. The triple of composed permutations AD, BE, and CF could theoretically have $101^3 = 1{,}030{,}301$ possible sets of disjoint cycles. The Enigma rotor system could be set in 105,465 ways. Each ground setting resulted in AD, BE, and CF that had one of 1,030,301 possible disjoint cycle structures. It is possible that each ground setting corresponds to a unique disjoint cycles structure.

If that were true, knowing the disjoint cycle structure would determine the ground setting, but is it true?

> If we have a sufficient number of messages (about eighty) for a given [ground setting], then, in general, all the letters of the alphabet will occur in all six places [of the message indicators] at the opening of the messages. [This allows the permutations AD, BE, and CF to be determined.] In each place they form a mutually unique transformation of the set of letters into themselves, that is, they are permutations. These permutations, designated respectively by the letters A through F, are not known to the cryptologist, but the transitions from the first letters to the fourth, from the second to the fifth, and from the third to the sixth likewise form permutations, and these are known to the cryptologist. They are products AD, BE, CF of the previous permutations. They may be represented as disjunctive products of cycles and then assume a very characteristic form, generally different for each [ground setting]...[Marian Rejewski, [3] Appendix E].

In Appendix D of [3], Rejewski claimed that "what was needed was a sufficient number of messages from the same [ground setting], about sixty, to make possible the formation of the characteristic set of AD, BE, and CF." Rejewski clearly believed that the disjoint cycle structure could determine the ground setting; however, in another place, he implied that the mapping from ground settings to disjoint cycles is not unique.

> ...permutations AD, BE, and CF have a characteristic form, and a set of three permutations with the same configuration of cycles recurs infrequently. [Marian Rejewski, [3] Appendix E].

What was needed was a catalog of the ground settings and the corresponding disjoint cycle structures.

> ...if it were possible to design a device that gave the length and number of cycles in the characteristic for each position of the rotors, and if next the lengths and number of cycles were catalogued, then it would suffice to compare AD, BE, CF for a given day with the products with the same configuration in the catalog to at once obtain the order of the rotors...[Marian Rejewski, [3] Appendix E].

The Polish codebreakers designed a device called a cyclometer [see, for example, Appendix E] to catalog the disjoint cycle structures that correspond to the ground settings.

> [Cataloging the disjoint cycles] took a long time, over a year, since we carried it out along with our normal work.... Once the six catalogs [one for each rotor order] were ready, though, obtaining a daily key was usually a matter of twenty minutes. The card told the drum positions, the box from which the card had been taken told the drum sequence....

> Unfortunately, on 2 November 1937, when the card catalog was ready, the Germans exchanged the reversing drum from the one they had been using, which was designated by the letter A, for another drum, a B drum, and, consequently, we had to do the whole job over again . . . [Marian Rejewski [3] Appendix D].

3. Recreating the Catalog

Apparently, no copies of the Polish catalog exist. Therefore, it is not known how the Polish mathematicians ordered the characteristics—the disjoint cycle structures.

Frank Carter, a mathematician who is now a Bletchley Park guide, has published the disjoint cycle structure for one order of the rotors and reflector B [1]. A simulation of the cyclometer developed by Tony Sale is available on the internet [5]. Sale's simulation also uses reflector B.

I have written a program that simulates the process that the cyclometer went through to create the catalog of disjoint cycles for the 105,465 possible rotor settings. The program uses reflector A, which was used by Rejewski and the Polish codebreakers in making the first catalog. The wiring of reflector A was determined in 2000 [4] (Tables 1 and 2).

The results are similar to Carter's results for the one rotor order and reflector B. The majority of the time, the disjoint cycle structure corresponds to either a unique ground setting or to a small number of ground settings, which would have to be checked by hand. 21,230 different disjoint cycle structures occur. Of these, 11,466 disjoint cycle structures (54.40 percent) correspond to unique ground settings. 20,433 correspond to 10 or fewer. Together, these account for 92.34 percent of the possibilities. Rejewski was generally correct that there are very few ground settings that correspond to a given disjoint cycle structure, but there are some bad cases. The disjoint cycle structure (13 13)(13 13)(13 13), for example, corresponds to 1,771 possible ground settings.

Table 1. Most frequent cycle structures

Disjoint cycle structure (AD)(BE)(CF)	Number of occurrences
(13 13)(13 13)(13 13)	1771
(12 12 1 1)(13 13)(13 13)	898
(13 13)(13 13)(12 12 1 1)	866
(13 13)(12 12 1 1)(13 13)	854
(11 11 2 2)(13 13)(13 13)	509
(13 13)(12 12 1 1)(12 12 1 1)	494
(13 13)(13 13)(11 11 2 2)	480
(12 12 1 1)(13 13)(12 12 1 1)	479
(13 13)(11 11 2 2)(13 13)	469
(12 12 1 1)(12 12 1 1)(13 13)	466
(13 13)(10 10 3 3)(13 13)	370
(13 13)(13 13)(10 10 3 3)	360
(10 10 3 3)(13 13)(13 13)	358
(13 13)(13 13)(9 9 4 4)	315
(9 9 4 4)(13 13)(13 13)	307

Table 2. Low frequencies

Number of disjoint cycle structures	Number of occurrences
11466	1
3381	2
1658	3
958	4
660	5
456	6
343	7
265	8
234	9
183	10
146	11
118	12
103	13
95	14
74	16

4. The Plugboard

What about the plugboard? Rejewski noticed that it was possible to find the ground settings from the disjoint cycle structure of AD, BE, and CF without considering the effect of the permutation caused by the plugboard. The theorem that he used is commonly learned in elementary permutation theory. Let G represent the permutation that occurs because of the rotor system and P represent the permutation that occurs because of the plugboard. After a letter is pressed on the Enigma keyboard, the right rotor turns forward one place and an electrical charge passes from the key on the keyboard to the plugboard, where it is changed by P; then, the charge passes through the rotor system where it is changed by G; then, the charge passes backwards through the plugboard where it is changed by P^{-1}. The Enigma permutation (composing from left to right) is PGP^{-1}. A theorem in elementary permutation theory says that G and PGP^{-1} have the same disjoint cycle structure. Therefore, Rejewski realized that when determining the rotor setting, the effect of the plugboard could be ignored.

> Polish penetration into the secrets of the Enigma began in earnest when Rejewski realized the application of a simple property of permutations—namely, that if G and P are permutations, then the permutation defined by PGP^{-1} has the same cycle structure as the permutation G. No doubt practitioners of group theory should introduce this property of permutations to students as "the theorem that won World War II" [2].

5. Conclusion

Rejewski's claim that "a set of three permutations [AD, BE, and CF] with the same configuration of [disjoint] cycles recurs infrequently" [3] is essentially correct.

For most days, given enough message indicators to determine the disjoint cycle structure of AD, BE, and CF; the Polish codebreakers could use Rejewski's catalog to determine the Enigma ground setting or, at least, to reduce the number of possible settings to a small number that could be checked. However, on some days, Rejewski's catalog could determine that the disjoint cycle structure for AD, BE, and CF corresponded to many ground settings. It is not known how the Polish codebreakers proceeded in those situations.

Acknowledgments

Dr. Chris Christensen for inspiration, sanity checks, and long hours going through disjoint cycle creations.

Frank Carter for his original work on the subject, editing, and encouragement.

About the Author

Alex Kuhl invented the "Lucky Guess" algorithm while at Northern Kentucky University as an undergraduate. Today he is found wandering around North Carolina State University acting like a graduate student in computer science. He is a spirited individual that values the finer things in life, such as kidnapping jayhawks, playing disc golf, and marathon evenings of Mario Kart. In his spare time he operates his own site, http://www.alexkuhl.com, and solves the latest mathematical challenges over breakfast.

References

1. Carter, F. 1999. "The First Breaking of Enigma: Some of the Pioneering Techniques Developed by the Polish Cipher Bureau," *The Bletchley Park Trust Reports*, 10.
2. Deavours, C. A. July 1981. "Afterward to How Polish Mathematicians Deciphered the Enigma by Marian Rejewski," *Annals of the History of Computing*, 3(3):229–232.
3. Kozaczuk, W. 1984. *Enigma: How the German Machine Cipher was Broken, and How It was Read by the Allies in World War Two*, translated by Christopher Kasparek, University Publications of America.[1]
4. Marks, P. and Weierud, F. January 2000. "Recovering the Wiring of Enigma's Umkehrwalze A," *Cryptologia*, 24(1):55–66.
5. Sale, T. Codes and ciphers in the second world war. http://www.codesandciphers.co.uk/. (accessed September 6, 2007).

[1]Appendix D, How the Polish mathematicians broke Enigma by Marian Rejewski. Appendix E, The mathematical solution of the Engima cipher by Marian Rejewski. The paper that is Appendix E also appears in January 1989 *Cryptologia*, 6(1):1–18; and in Deavours, C. Kahn, D. Kruh, L. Mellen, G., and Winkel, B. Editors, 1989, *Cryptology: Machines, History, & Methods* Artech House, 310–327. Rejewski, M. 1980. An Application of the Theory of Permutations in Breaking the Enigma Cipher, *Applicaciones Mathematicae* 16(4) is similar to the material in the two appendices and may be found on the web.