Copy 8 of 8 copies

SECRET

IL 3639
Report #F-68
29 July 1944

SUBJECT: Fish Hotes
TO : CO, SSA, War Dept.

1. This report will be devoted to miscellaneous observations resulting from studies of the permutation Fish machine of the latest type. The problem of solving current traffic seems completely hopeless. With the addition of the auto-key element eliminated whereas the only feasible method of solving messages enciphered on the machine described inReport #F-46 seems to be through depths. As was mentioned inReport #F-46 the motor action is sometimes switched off and this gives rise to several possible techniques which are described below. For the most part, however, the problems which seem capable of solution are comparatively trivial. The fundamental difficulty of the general problem arises from the fact that a crib does not yield key. Even with a crib of several thousand letters there is no known method of determining wheel order and settings. In the discussion which follows the symbols and nomenclature are those ordinarily used in Fish problems and those defined in Report #F-46. The only new symbol is δ which means simply P + Σ.

2. The simplest method of representing the relations between P, Z, Σ and π is through a 32 x 32 x 32 cube. Either P, Z or Σ can be placed within the cube and the other 3 elements become the co-ordinates. π could not be placed inside because it is not, in most cases, uniquely determined by P, Z and Σ. Actually the π co-ordinate need be only 30 elements long because / and Z produce identical permutations and so do T and E. For actual use the cube is cut by planes along any axis and represented by 32 squares each 32 x 32. The choice of mode of representation depends on the problem under consideration. For the co-ordinates the teleprinter alphabet is most conveniently written in the order set forth below and spaced as indicated:-

/ E493T ASDZIRLNHO UJWFYBCPGM KQ+XV 8

As a result the body of the square is divided into blocks of 1, 5, 10, 25, 50 and 100.

3. For reading in depth it has been found useful to prepare a rather extensive catalogue of the alphabets. A separate page represents each plain-cipher pair - actually only 512 pages are needed because of complementary relationships. On the top margin of the page is the cipher alphabet (arranged as indicated in paragraph 2) and the left-hand margin of each page shows, in the form of a digraph, the 32 possible combinations of Σ and π keys which can yield the P-A relation which the page represents. The body of the page contains 32 rows, each of which shows, for the key at its left, the plain values corresponding to the cipher values in the top margin. The columns of the page will inevitably contain many repeats and a plain text assumption in a second

SECRET

message does not, therefore, uniquely determine the key. Ordinarily the key will be determined from plain text in 4 or 5 messages. A unique determination from 2 messages is very rare. Actually the catalogue is nothing more than a sort of the 1024 possible alphabets (really only 960) arranged in 512 groups of 32. After assuming a plain text letter in one message, the cryptanalyst selects the appropriate catalogue page and scans the columns under the cipher text letters of the other legs of the depth. This shows the range of possible plain text letters in the other messages. His task is to find a group of acceptable plain values all of which appear in a single row.

4. The simplest conceivable problem that can arise with the auto-key machine is the determination of the wheel order when all settings are known. If a crib is available this can readily be done. Key is derived beginning from the known setting. There are 10 key symbols at each position and they must first be divided into Σ and π and then assigned to their correct impulses. If at a given point the key contains only 1 cross, P contains none and Z contains 1, we can identify a Σ wheel and also determine the impulse it controls. Instances where such definitive conclusions can be reached will be rare. Usually the evidence will merely limit the possibilities but by combining data solution can be achieved. There is no answer to the question as to how long a crib is needed because the same type of procedure can be used with no crib whatever. Clearly, the longer the crib, the quicker the solution.

5. Without a crib the problem is tedious and difficult. The best approach seems to be to select a section of key where there are only 1 or 2 crosses (or only 1 or 2 dots) in several positions in sequence or near sequence. An ideal stretch would consist of 3 successive positions in each of which the ratio of dots and crosses is 9 to 1 or 1 to 9 with the lone element always in the same position. Then only 10 assumptions need be made. Each of these will yield a trigraph and the correct assumption will produce plain text. After 1 or 2 of the wheels have been identified, or after some plain text has been found, the process is much simplified. However, the initial entry is most difficult and involves a tremendous number of assumptions unless, by good fortune, a particularly favorable stretch of key is available. (But see paragraph 12).

6. No ideas have been developed as to solution of the general problem when the wheel order is known. Even with a crib of great length there seems to be no feasible method of determining the starting points of the wheels (but see paragraph 9).

7. The limitations on the permuting possibilities of the machine seem to present a possible point of attack. In the square set forth below, the left-hand co-ordinates represent the positions of the P (or rather the 5) impulses, the top co-ordinates

the positions of the Z impulses and the figures within the square the number of times out of 32 that each δ position goes to each Z position. The figures are obtained simply by enumerating the results of all permutations which the machine can produce and by considering all π keys as equally likely.

	1	2	3	4	5
1	9	9	2	4	8
2	8	8	16	0	0
3	4	4	8	16	0
4	2	2	4	8	16
5	9	9	2	4	8

8. The fact that δ, position 2, must go to Z, position 1, 2 or 3 makes it possible to derive δ_2 from those Z letters which have identical symbols in these 3 positions (i.e. ¼ of the entire cipher text). If we have a crib we can then derive a fragmentary ε_2. If there were no motor action the fragment thus obtained could be matched against each of the 10 wheel patterns and a start made on determining settings and wheel order. About 40 letters of crib (yielding about 10 ε_2 characters) would be sufficient to identify and set the ε_2 wheel. A similar process would next be used for ε_3. If we started with ε_3 we could use only 1/8 of the text instead of ¼, but with ε_2 (and consequently δ_2) fully known, δ_3 (and consequently ε_3) can be derived wherever δ_2 is cross and Z is E, Z, 4, L, 9, H, K or 8 and wherever δ_2 is dot and Z is the complement of any of the foregoing. This seems to be 8 timds out of 32, as before, but is actually 8 out of 28 because there are 4 values of Z (/, 3, T and O) that cannot be associated with a cross in δ_2 (and 4 that cannot be associated with a dot). Therefore, the same length crib should enable the identity and setting of ε_3 to be determined. With ε_2 and ε_3 identified and set, any other ε impulse can be derived in the same manner despite the absence of any absolute limitation with respect to the Z positions at which δ_1, δ_4 and δ_5 can emerge. Actual enumeration shows that after ε_2, and ε_3 are set, any other ε can be deduced from 7/24 of the Z values so the same length of crib should suffice. The setting of the last 2 ε wheels should be progressively easier. The final step of identifying and setting the π wheels is simple. With P, Z and ε all known, some of the elements of π can be determined in nearly all positions. π is uniquely determinable from Z and δ in only 24 cases out of the 1024 but it seems apparent that sufficient π symbols could be derived so as to produce unique matches with the remaining wheel patterns. No calculations have been made but the procedure seems simple, particularly since identification and setting of any one of the π wheels leads to further derivation of π symbols.

9. The trouble with the foregoing procedure is that it fails completely when the machine is motorized. All that can now be derived is a fragmentary portion of extended Σ_2 and it seems as though it would be a virtual impossibility to match this against the wheel patterns. Conceivably with a very long crib sequences such as .x.x or x.x. might appear which, of course, would be unextended Σ_2. The probability of obtaining such a sequence at any point is the product of (a) the probability of such a sequence appearing in a Σ pattern - approximately 1/8 - by actual enumeration slightly less - (b) the probability of 3 successive motor crosses - approximately $(3/4)^3$ - and (c) the probability of being able to derive 4 successive Σ_2 symbols - $(\frac{1}{2})^4$. This is about .000206. It might also be possible to find fragments which could be identified as repeating on a subsequent cycle. The extended period would average about 1/3 more than the unextended period and it seems barely possible that fragments might be slowly and laboriously pieced together. (The approximations $(3/4)^3$ and 1/3 are based on the machine described in Report #F-46 whose wheels step forward about 3 times out of 4. It is not known whether the auto-key element changes this ratio.) With the wheel order known this procedure would seem to be substantially easier but still too difficult to be successful. If Σ_2 and Σ_3 were both in the tetrad consisting of the 67, 69, 71 and 73 wheels and were consequently motorized identically, some aid might be obtained by an independent derivation of Σ_3. If .x or x. in Σ_3 coincided with .. or xx in Σ_2 it would establish that .. or xx is a true fragment of Σ_2 pattern. The chance of obtaining such a relation at any point is, however, only 3/1024. There are other possibilities of the same kind but all rather remote. If, for example, / and 8 appeared in sequence in the cipher text (or //, 8/ or 88) 2 successive elements of δ (and consequently of Σ) could be derived on all 5 levels. If both elements of Σ_2 were the same and the 2 elements of some other Σ in the tetrad with Σ_2 were different, the Σ_2 fragment would be established as true pattern. (The discussion assumes that all inferences are applied only to build up the Σ_2 pattern, whereas many of them might be used to develop the patterns of other wheels. However, Σ_2 is clearly the most likely point of initial entry.) Thus far we have considered only deductions that can be made with mathematical certainty. There are others which can be made on a probability basis but there seems no way to combine probabilities at successive cycles and their use, therefore, introduces an additional element of uncertainty. It is, perhaps dangerous to draw conclusions with respect to a technique that has never been actually tried but with the motorized machine it seems almost hopeless to set the wheels by this method even with a very long crib and with wheel order known. In an exceptional case, which fortuitously yielded more data than might normally be expected, success might be achieved. The foregoing discussion disregards completely the problem of locating the crib and assumes that it can be correctly set in the message.

10. A statistical method has been developed for the motorless machine but it is completely inapplicable when it is motorized. Its basis is plain language characteristics plus the limitations on the machine permutations. The method can be used on single impulses or on pairs. I shall describe the method as applied to a pair of impulses, first, because plain text characteristics are more pronounced when pairs of impulses are considered and second, because it is slightly more involved. If the 2 wheel theory is understood, its adaptation to a 1 wheel is simple whereas the converse might not be true.

11. In the following diagram the left-hand marginal symbols are ō characters and the top margins are Z characters. The figures within the diagram show the number of times out of 32 that each ō goes to each Z. It is clear that the outer left-hand margin must always go with the upper of the top margins and that the inner left-hand margin is always associated with the lower of the top margins. It will be noted that the 5 x 5 portion of the diagram is identical with the square of paragraph 7.

Z

			/	E	4	9	3	T	A	S	D	Z	I	R	L	N	H	O
			8	~~C~~ V	X	+	Q	K	M	G	P	C	B	Y	F	W	J	U
	/	8	32															
	E	V		9	9	2	4	8										
	4	X		8	8	16	0	0										
	9	+		4	4	8	16	0										
	3	Q		2	2	4	8	16										
	T	K		9	9	2	4	8										
	A	M							8	6	1	2	6	1	2	2	4	0
ō	S	G							4	2	6	1	2	6	1	4	2	4
	D	P							2	1	2	6	1	2	6	0	4	8
	Z	C							4	2	4	8	2	4	8	0	0	0
	I	B							0	8	4	0	8	4	0	8	0	0
	R	Y							0	4	2	4	4	2	4	4	8	0
	L	F							8	6	1	2	6	1	2	2	4	0
	N	W							0	0	4	2	0	4	2	8	4	8
	H	J							4	2	6	1	2	6	1	4	2	4
	O	U							2	1	2	6	1	2	6	0	4	8

If we consider each entry as the numerator of a fraction whose denominator is 32 and if we assume that all π keys are equally likely the diagram is a probability table. The figures in the rows opposite each ō represent the a priori probabilities that it will go to the various Z's indicated; the figures in the column under each Z represent the a posteriori probabilities that it derived from the various ō's indicated. The last portion of this statement assumes that all ō's are equally likely which is true if either all P's or all Σ's are equally likely, i.e. ō is random if either of its constituents is random. P is obviously not random but ε can be considered random with a negligible error. (Neither ε nor π is truly random because the wheel patterns contain a slight preponderance of crosses.) Consequently the columns can be assumed to represent true inverse probabilities. We must next tabulate the probability that a given Z will correspond to a δ the sum of whose i^{th} and j^{th} impulses is dot. This is determined by adding together those entries in the Z columns where δ_{ij} is dot and dividing by 32. The following the the result for ij = 45.

Cipher	Prob. x 32
/ 8	32
E V	21
4 X	21
9 +	26
3 Q	20
T K	8
A M	14
S G	17
D P	13
Z C	9
I B	17
R Y	13
L F	9
N W	14
H J	10
O U	12

We now select the most favorable Z's and make a tabulation as follows:-

Z	δ_{45}	Prob.
/ 8	•	1
9 +	•	26/32
T K	x	24/32
Z C	x	23/32
L F	x	23/32
H J	x	22/32

The number of signs derived from these 12 values of Z which will agree with the true δ_{45} will be the average of the probabilities which is .781. If P_{45} is dot 67% of the time, the number of signs which will agree with the true ε_{45} is (.67 x .781) + (.33 x .219) = 59.5%. On a message of 5000 letters we would actually use about 1875 letters so that the bulge would be about 178 or 8σ. The derived ε_{45} must be compared with the full 2 wheel development of all possible pairs of patterns. The average development is about 3600 and there are 45 pairs so about 162,000 trials are necessary but these can be handled rapidly with IC equipment or tape comparators. Actually 5σ is about enough so that messages considerably shorter than 5000 letters could be set. After 2 ε wheels are set other pairs can be chosen involving one of the patterns already set so as to shorten the runs, just as is done in ordinary Fish procedure. The determination of the optimum number of Z letters to be used is quite simple but since it depends on the distribution of probabilities for the particular ij it must be separately determined for each pair. For ij = 45 the optimum selection seems to be the 12 letters tabulated above. The determination of the best ij pair to start with depends, of course, on plain language characteristics as well as on the probability distribution. After the Σ key is completely determined the π key should not be difficult. One method would be to try to read plain text on generatrices. In 1/16 of the positions the plain text is unambiguously determined, in 5/16 it would have to be found on a maximum of 5 generatrices and in the remainder on a maximum of 10. The exact number of generatrices for each cipher letter can be found from the diagram at the beginning of this paragraph by counting the number of entries, other than zero, in the column under it. The same diagram shows the probabilities to be assigned to the respective generatrices. Identifying and setting the π wheels should be possible after a short stretch of plain language has been reconstructed.

12. The foregoing statistical method can be adapted to solve the problem of paragraph 5. We derive δ_{ij} signs for favorable cipher letters and compare the results with the key development, trying all 45 possible pairings. The illustrative figures of paragraph 11 will produce a bulge of just under 3σ on a message of 600 letters. This is ample since no sliding is involved. The subsequent steps require no elaboration.

13. Another statistical method of solving the paragraph 5 problem seems inferior but will be outlined because there may be other applications. All cipher letters corresponding to the same key or to a complementary key are collected together producing 512 groups. In 10 of these groups the key will have a 9 to 1 ratio of dots and crosses. Select the group with the greatest number of members and make 10 assumptions as to the Σ or π impulses corresponding to the lone element. Each assumption leads to a series of plain text letters and the right one should yield a distribution corresponding to plain language. With a very long message it might be possible to reach a solution from these 10 groups alone. But we could, if necessary, use some of the 45groups with 8 to 2 ratios, either for original assumptions or to check tentative conclusions. Obviously it is necessary to collect cipher letters only in those groups we are planning to use.

14. Examination of the wheel patterns set forth in Report #F-46 shows, curiously enough, that each of them starts with dot cross. These are known to be the true starting points of th e wheels, that is, the points marked 0 and 1 (or possibly 1 and 2) on the actual machine. A theory as to the reason for this peculiarity is that it facilitates testing the machine. If all wheels are set at 0 no Σ key is operative but all permuting elements are effective so that it is easy to tell,

probably from parallel plain-cipher lists, whether the permuting mechanism is functioning properly. When all the wheels are set at 1 there is no permuting effect and each plain letter must be converted into its complement.

15. The zeroizing mechanism mentioned in paragraph 22, Report #F-46, enables the operator to return quickly to the point at which he started his message - not to the true zero points of the wheels. Presumably he must turn them by hand to test his machine. The dangers of the zeroizing device were apparently realized a long time ago because in February 1943 the operators were issued instructions (which were read in E traffic) forbidding them to use this mechanism.

Walter J. Fried
Capt. Signal Corps