JNA-20: An Alternative I.C. Plate Attack

C.H.O’D. Alexander

Editor's Preface

This document is one of seven papers in a series of documents written by the late Conel Hugh O'Donel Alexander while he worked as a cryptanalyst at Bletchley Park during the Second World War. The documents have been obtained from the US National Archives and permission has been obtained from the Controller of Her Majesty's Stationary Office to publish the papers on the editor's personal Web Page. The documents have been faithfully retyped by the editor with help from Ralph Erskine. The original documents were typed and had the style and layout of a typewritten document of that period. To make the re-edited presentation more pleasing the documents have been both left and right justified and a more modern type font has been used. Apart from these modifications to the layout the documents have the appearance of the original.

Where there are obvious typing errors these have been corrected in square brackets. In this particular text where the editor's copy is quite faint at places there are some numbers that can not be determined with certainty. These numbers have been marked in the text with an asterisk enclosed in square brackets, e.g. 42[*].A further modification is the adoption of the German spelling of the word Stecker with a capital, however, the constructed verb 'steckered' has been left in its original form.

The Editor,

Frode Weierud, © June 1998

Source:

Addendum to Captain Walter J. Fried's report No. 78 of 19 August 1944. National Archives and Records Administration (NARA), Record Group 457, NSA Historical Collection, Box 880, Nr. 2612.

COPYRIGHT

Crown copyright is reproduced with the permission of the Controller of Her Majesty's Stationery Office.

JNA-20: An alternative I.C. Plate attack.

Introduction

1. This is a description of an alternative method of attack on JNA-20 to break into the new period. It can only be used on motion F M S or S F M but as these are the motions not covered by the X method (which is probably better than the 'crash' attack anyway) they are the motions on which an I.C. Place method is most likely to be used.

General Theory

2. The idea of the method is to work within a slow wheel turnover stretch, absorbing the slow wheel in the input or output Stecker. Suppose the motion to be F M S (exactly the same method applies if it is S F M). Then have plates made for the decode of each of the letters A, B, C… Z through F and M wheels. Now take any letter, say P and assume P/a through the slow wheel (i.e. P steckered to a if the slow wheel is absorbed in the Stecker). Make a plate with a hole at the position of each P in the message and compare it with each of the master plates. Suppose that in fact a large number of the P's happen to come opposite clear text A's, say, and that the true input Stecker of A is Q and the output Stecker, through slow wheel, of P is J then we should get a large number of clicks between the Q master plate and the J 'P plate'.

3. Comparing this with the O crash plate we may make the following comments.

(1) The fundamental principle is the same in that a large number of one particular clear text letter is going to have a large number of clicks between the correct message and master plates.

(2) The amount of work is of the same order; a factor of 26 is gained on the second method by absorbing the slow wheel in the Stecker but this factor is lost again since correct input and output Steckers have to be assumed whereas in the crash method one assumption O/a say covers clear and cipher.

(3) We can only use a stretch of 25 alphabets at most on the new method which is a decided drawback.

4. There are however 3 very great advantages of the new method.

(4) If we take a true Stecker, say Pa, then if <u>any</u> letter (it need not be specified in advance) does unusually well it will show up by a large number of coincidences.

(5) We can try as many cipher letters as we like and thus have opportunities of confirming the true position by further good scores.

Crown Copyright

(6) Given a really favourable case such as a triple level start we can make practically certain of success by trying a large number of cipher text letters and comparing results, i.e. we can 'flog' a really good message or set of messages.

Another advantage –

(7) is that by the comparison of results from runs on different cipher letters we can eliminate or greatly reduce the use of Mike.

5. My view is that on a triple level start this method is superior to the crash method, on an ordinary level start also better if the messages are over 30 alphabets long so that we are certain of a stretch of 15 or 20 alphabets without turnover, while on a single long message I am not sure. A good deal depends on the amount of I.C. Plate machine time available as this method makes more use of the Plate machine and much less use of Mike.

Details of Procedure.

6. For convenience I shall refer to Steckers on clear text side as clear Steckers and to Steckers through slow wheel on the cipher text side as cipher Steckers.

7. If we are making assumptions clear Stecker O/K and cipher Stecker P/L then we are only interested in comparing the positions of P with the positions at which K decodes as L. To collect such positions together the following is suggested.

8. Decode K through F and M wheels and then produce a second 'decode' from the first as follows. Suppose H (the 8th letter of the alphabet) comes in alphabet 4 at fast position 11. Then replace this by K (11th letter) in alphabet 8 (corresponding to the H) in position 4. If we do this for all the letters we shall collect in separate groups all the A's, B's, C's etc. in the decode and (just as with the crash plates in the other method) the position within a group will correspond to the alphabet and the letter at that position to the first wheel position within the alphabet.

9. The 24 positions in which K decodes as X say will now appear as 24 letters in a stretch of 25 places. To allow for the message being more than 25 alphabets long allow 60 places. This gives 10 sets per plates i.e. 68 plates in all or 78 if we allot three plates per letter which might be the most sensible thing. (Note that if it were any advantage we could arrange that a set of 3 plates would contain not all the decodes of a particular letter but all the encodes of a particular letter).

10. Now – just as with the crash plate method – we shall have to make 25 assumptions for the T.O. position of the fast wheel. If the message is 40 alphabets long, say, each assumption will need 40 columns on the plate, the position within the column corresponding to the fast wheel position of the letter in question for that hypothesis. If we put 7 assumptions on one plate (using 'marks' as we did for the crash method) then we shall need 4 message plates per cipher text letter used (i.e. 4 plates for all the cipher text A's, 4 for all the B's if we decide to work on these as well and so on).

Crown Copyright

11. The comparison is made in exactly the same way as in the crash method and will take about the same sort of time. However we can work on 26 letters instead of on one and get good scores on any clear text letters (See Para 4).

12. The chief new problem that arises is this. Suppose we have a triple level start of 25 alphabets and that the S.W.T.O. in fact comes between alphabets 15 and 16. The 'bulge' in clicks in the right position is sufficient to differentiate the right position from the wrong one but we are in the dilemma that if we put our target low enough to pick up a true score with overlap of 15 alphabets between message and master plate we shall get masses of results we do not want when the overlap is 25 alphabets. In fact suppose we had (to take convenient numbers) 33 A's in the first 11 alphabets, 45 in 15 alphabets, 57 in 19 alphabets, 69 in 25 alphabets. Then we might wish to accept scores of 7/33, 9/45, 8/57, 10/69. Suppose now on the master plates we have holes in position 27 in columns 4, 8, 12, 16 … of the 35 columns between two successive sections of 25 (See Para 9), there being no holes in position 27 within any section of 25. On the message plates have a hole in every 27th position. Now set the 'target' at a score of 10. Suppose we have only an overlap of 15 alphabets between message plate and a section of the master plate: then there is a stretch of 10 places where the message plate overlaps the interval in between two sections of the master plate and this will give 2 repeats in position 27. If therefore we succeed in scoring 8 genuine repeats the 2 bogus ones will bring the score up to 10. I think this will provide a fairly satisfactory solution of this difficulty. If we do not wish to make special allowance for overlap of less than 21 alphabets, say, we can do this by having only '27 holes' in positions 1 to 4 and 22 to 25 on the message plate so that no extra bonus will be received if the overlap is under 21 alphabets.

Statistical analysis of method in various cases.

13. This is an attempt to estimate the chance of success in various cases and to compare this method with the original plate method, considering a few typical cases.

14. Suppose first that we have a triple level start, 25 alphabets long. On the crash method we could get (assuming 65 cipher text O's – expected number is 72 but quite a number of these will be cipher text clicks and therefore the number of distinct O's will be less) something like 50[*]% chance for about 2000 stops, all of which have to be tested individually.

15. The chance of success on the new method depends very much on where the turnover happens to come. Let us suppose that it comes between alphabets 17 and 18 (not a very favourable case) i.e. that we have a stretch of 17 alphabets without turnover. Then an average number of any particular cipher text letters excluding repeats would be about 46: if we took the dozen most frequent letters they would average about 50 I should say. If our target is 8 repeats for this overlap (see Para 12 for method of adjusting target for different overlaps) then we get about 200 stops per run with a 20% chance of picking up the right answer on the most frequent letter (and a further chance – perhaps about 10% at a vague guess – of picking it up on another letter which happens to do very well against this particular cipher text letter). Suppose we run a series of cipher text letters and insist on 2 stops at the same position and for the same clear Stecker. Then when we have run 12

cipher text letters we shall have had about a 75% chance of success and will have approximately $\frac{12 \cdot 11}{2} \times 200 \times \frac{200}{625 \times 26} \cong 162$ stops to test i.e. for much more plate time than on the other method we should get a very considerably higher chance of success and much less subsequent testing.

[15]. Alternatively, and probably better, we could set the target higher (at 9 repeats in this case) and examine all stops. This would give us (in this example) 40 stops with a 1/9th chance on O alone (and therefore a better chance including I, A, U) for each cipher letter tried ∴ A dozen shots would give 480 stops with a 75% chance on O only and probably therefore up to 90% altogether.

16. Several possible refinements of method come up at once.

(1) We can insist on a minimum of 13 alphabets of overlap since we must get at least that. We shall – on scheme of Para 12 – achieve this more or less automatically since the smaller the overlap the harder it is to get to the target score e.g. if our target is 10 it is far easier to get this by scoring 10/69 on the full 25 alphabet overlap than with 9 alphabet overlap to score 6/21 plus 4 from the '27 hole' clicks.

(2) We can try to allow for clicks. This will almost certainly mean something like recording everything twice on the master plate and clicks only twice on the message plate. I am rather doubtful whether this is worthwhile. An alternative would be to make special message plates with clicks only on them, e.g. suppose there were 6 cipher text clicks on Q the Q click plate will have six holes in it, and then make special runs on these. Insisting on 3 clicks on the right position would give about 500 stops and about a 12% chance of success and I rather doubt its being worthwhile. (Marshall Hall's method of attack was based on using clicks but not), I think, confining himself to a T.O. stretch. It might be possible to look for clicks between the results of his method and this one).

(3) It will always pay to take as our first choices cipher text letters with numerous clicks since each click has a much better than normal chance of being a common letter and therefore our chance of a high score is far better. It might very well pay to set the stop criterion lower and then reject all positions not containing at least one click or something of that sort e.g. one might require a minimum for further examination of a score of 7 with 2 clicks, 8 with 1 click or 9 with 0 clicks. The policy of accepting a large number of stops on one or two cipher text letters with numerous clicks instead of working on a large number of cipher text letters and restricting stops is one which needs careful consideration and probably the best thing to do would be to amass a fair amount of empirical material and decide from that.

(4) We might look for good scores on different clear Steckers for the same cipher Stecker. On the whole I don't think this will be nearly as profitable as looking for good scores on the same clear Stecker for different cipher letters. Every cipher letter has a high chance of doing well as the best clear Stecker, but a true cipher

Crown Copyright

Stecker can only do well on one or two clear Steckers and its chance falls off rapidly with fall in frequency of the remaining common clear text letters.

17. Next suppose we have an ordinary level start, 30 alphabets long and say 57 cipher text O's. We could probably get about a one third chance for 2000 stops on the crash method (I have not got figures for 57 O's and this is a rough estimate). Suppose that the T.O. comes between alphabets 20 and 21 (a fairly unfavourable case – there is a 2/3ds chance of doing better than this. Then we should get an average of about 38 occurrences of each cipher text letter so that we could presumably pick a dozen with an average of say 42[*]. By running these 12 letters we could get a 2/3ds chance on the commonest letter alone (I should say 75% chance in all, at least, allowing for a lucky break on another common letter) for a total of about 600 stops.

18. Suppose we have a single message of 40 alphabets. I should be inclined to assume that this contained a stretch of 25 alphabets without turnover as 25 alphabets is very near the bone on a single message and not only is the chance small but the difficulty of failing stops would be very considerable on less than 600 letters. Given 25 alphabets we get about 24 occurrences of each cipher text letter. So we could get a dozen letters averaging 27 probably. From these we could get, for about 15 stops per letter, a 2½% chance on O (insisting on 7/27) which could give just 25% chance on O and therefore I should think 30-40% chance in all for a total of 180 stops.

Use with a Pattern Repeat

19. A Pattern repeat would be very valuable.

(1) It restricts turnover and therefore reduces the amount to be run and the number of stops.

(2) It means that if P throws onto Q then for any given fast wheel position (i.e. any given message plate section) Stecker of P implies a definite Stecker of Q and therefore if we use the first method suggested in Para. 15, stops are further restricted.

Summary

20. This is a better method than the original crash plate method. I think it is better – because considerably more powerful and applicable to a much wider ranger of messages - than the Marshall Hall method except on a quadruple level start for which the latter might be quicker. There is scope for a good deal of statistical work to determine the chance of success and best variant of the method to be used. A collection of a 100 or so frequency dotteries would enable this to be cleared up. The technique of the method is pretty well identical with that used on the crash method so there should be no doubt as to its feasibility.

C. H. O'D. Alexander

Editor’s Notes

1. The document is not dated but it is known to have been prepared some time before 19 August 1944. There are also some indications that this is not the latest version of the document.
2. JNA-20 is the designator for the traffic generated by Coral, the cipher machine used by the Japanese Naval Attachés. The machine is made out of 3 banks of 26 position telephone selectors which move with cyclometric stepping. The machine is equipped with plugboards at each end and was apparently used with the same sequence at both ends. Only three of the six possible “wheel” orders were used. The machine is cryptographically stronger than Purple but the use of the machine was relatively weak. For more details on Coral and Purple see Cipher A. Deavours and Louis Kruh, *Machine Cryptography and Modern Cryptanalysis*; Norwood MA: Artech House, 1985.
3. The I.C. Plate machine, belonging to the RAM (Rapid Analytical Machinery) family, was performing optical I.C. (Index of Coincidence) measurements on two specially prepared photographic plates. A later model was the I.C. Film Projector or HYPO which used two photographic films instead of plates and which was more automated than the I.C. Plate machine.
4. F M S = Fast, Medium, Slow refers to the order of the “wheels” in the machine. In the Enigma the wheels are in the order S M F (Slow, Medium, Fast).
5. W.O. = Wheel Order. The order of the three moving wheels in the machine.
6. T.O. = Turn Over. The turnover of a wheel governed by its right-hand neighbouring wheel. In this case the stepping of a neighbouring telephone selector.
7. S.W.T.O. = Slow Wheel Turn Over.
8. Mike was a cryptanalytical aid made by the National Cash Register Company (NCR). It was used to count digraphs and consisted of a great number of electromechanical counters which are said to have filled a whole wall. For further details about Mike and the RAM see Colin Burke, *Information and Secrecy: Vannevar Bush, Ultra and the Other Memex*; Metuchen, N.J. & London: The Scarecrow Press Inc., 1994.
9. Dottery was a hand method for determining the Stecker plugging when wheel order, Ringstellung and message key had been determined. The name is derived from the dots that were used as tallies next to each letter on a frequency counting sheet.