REF ID:A66757

MEMO ROUTING SLIP | NEVER USE FOR APPROVALS, DISAPPROVALS, CONCURRENCES, OR SIMILAR ACTIONS

1 NAME OR TITLE	INITIALS		CIRCULATE
ORGANIZATION AND LOCATION	DATE		COORDINATION
2			FILE
			INFORMATION
3			NECESSARY ACTION
			NOTE AND RETURN
4			SEE ME
			SIGNATURE

Suspense

~~to [illegible]~~ ~~to Joe~~ ~~to Joan~~ ~~to [illegible]~~ ~~to [illegible]~~

REMARKS

Capt M. — ~~(he'll be back → 1500)~~

Pse tell him:

1) I think the paper is well worth putting into the 34— series of ~~technical~~ papers

2) I still think it should be classified at least conf because it does contain certain good ideas as to construction of 1-time keys.

3) ~~We ought to send copy to~~ ~~[illegible]~~.

FROM NAME OR TITLE | DATE

ORGANIZATION AND LOCATION DONE [illegible] | TELEPHONE

DD FORM 1 FEB 50 95 Replaces DA AGO Form 895, 1 Apr 48, and AFHQ Form 12, 10 Nov 47, which may be used. 16—48487-4 GPO

Suspense

1 May

(Send to Nyquist?)

MEMO ROUTING SLIP

NEVER USE FOR APPROVALS, DISAPPROVALS, CONCURRENCES, OR SIMILAR ACTIONS

1 NAME OR TITLE	INITIALS		CIRCULATE
ORGANIZATION AND LOCATION	DATE		COORDINATION
2			FILE
			INFORMATION
3			NECESSARY ACTION
			NOTE AND RETURN
4			SEE ME
			SIGNATURE

REMARKS

Dr. Campaigne says he wondered about the classification but decided the "U.S. systems" reference is not "commissive" enough to warrant classification. He says that you and himself are the only ones who have seen the paper and as far the need for classifying is concerned, you can keep it between yourselves.

SRM

FROM NAME OR TITLE	DATE 2 Oct.
ORGANIZATION AND LOCATION	TELEPHONE

DD FORM 1 FEB 50 95 Replaces DA AGO Form 895, 1 Apr 48, and AFHQ Form 12, 10 Nov 47, which may be used.

16—48487-4 GPO

REF ID:A66757

MEMO ROUTING SLIP	NEVER USE FOR APPROVALS, DISAPPROVALS, CONCURRENCES, OR SIMILAR ACTIONS		
1 NAME OR TITLE Mr Friedman S/ASST	INITIALS		CIRCULATE
ORGANIZATION AND LOCATION	DATE		COORDINATION
2		✓	FILE
		✓	INFORMATION
3			NECESSARY ACTION
			NOTE AND RETURN
4			SEE ME
			SIGNATURE
REMARKS As per our Conversation			
FROM NAME OR TITLE H Campagno		DATE 5 Oct	
ORGANIZATION AND LOCATION		TELEPHONE	

DD FORM 1 FEB 50 94 REPLACES NME FORM 94 1 FEB 49 WHICH MAY BE USED 16—50303-8 GPO ☆

Excursus on Random

28 May 1954
by H. Campaigne

Many cipher systems depend on a stream of "random" information, called "key". This concept of random is a very elusive one. Random can be defined to a certain approximation in mathematical terms, and this approximation can be tightened to any extent one is willing to undertake, but it is still an approximation to our elusive concept.

A dictionary gives for "at random: without aim, direction, rule, of method; haphazard, aimless; irregularly". It lists as antonyms, "planned, designed, considered, deliberate". To the cryptologist it means "unpredictable". Mathematically it is a process, not a finite sample, which is random. By a process I mean a procedure which on demand will produce digits indefinitely such that each digit occurs $\frac{1}{10}$ of the time, each possible pair of digits occurs $\frac{1}{10^2}$ of the time, and so forth, until for some n each of the 10^n combinations of n letters occur $\frac{1}{10^n}$ of the time. Clearly in a practical case we can test only a finite sample, and must limit our testing to some maximum n. It is the finiteness of the number of statements which makes this definition fall short of the ideal of random. For example, suppose we take $n = 3$, and have each digit occur $\frac{1}{10}$ of the time, each pair $\frac{1}{100}$, and each triple $\frac{1}{1000}$ of the time. Now if we take the following example we see that it satisfies the three criteria approximately but is perfectly predictable from any 4 consecutive digits. This process is defined by induction. Let $d_n = d_{n-4} + d_{n-3}$ (10), and select d_1, d_2, d_3, and d_4 arbitrarily. This process almost but not quite fits the 3 criteria.

The mathematical rule given above is too rigid, of course. One can change it to read thus: In any sample of M digits the number of occurrences of a specific n-nome will differ from $\frac{M}{10^n}$ by $S\sqrt{\frac{M}{10^n}}$ or more in only a small number of cases. Here S is a suitably chosen number, say 3, and the number of exceptions to the rule is chosen accordingly, say once in 10,000. This definition is harder to apply. If we apply it to our counter-example we find that it passes for n = 3 or less. Here is another sequence

9 9 9 8 8 7 6 5 3 1 8 4 9 2 3 1 5 4 6 9 0 ---

which is also uniquely determined by any 4 adjacent digits (and the rule of generation).

The important property for a cryptanalyst is predictability. If from a small stretch he can predict the rest, then no other property is of any significance; the cryptanalyst can guess at plain-text, derive key, predict more key, and verify his guess. Therefore the problem of reading any message is solved, in theory at least. Even if the prediction is only statistical the cryptanalyst can use the information. He can do this in two ways: He can search for "clichés" or long repeats in the key. Or he can use the fact, if it is a fact, that digits tend to be alike.

Now this last is usable even if the key came from a stochastic process which is unpredictable except that the digits come with various probabilities. Thus we see that the mathematical definition given above is a necessary condition for secure key. Whether it is sufficient is another question.

On the other hand key can be predictable and yet be secure in the following sense. If the law of generation of the key (assumed to be reproducible by the legitimate recipient) is unknown to the cryptanalyst, and if he is unable to reconstruct it from the data at hand, then so far as he is concerned it is unpredictable. Thus "predictable" is a subjective term. It is this theory which has been widely used by cryptographers, who attempt to design laws of generation so complex that they believe reconstruction is virtually impossible. Many U. S. systems are based on this theory, and the key is frequently referred to as "random".

To return to the mathematical definition of random. This definition has two drawbacks from the viewpoint of the cryptographer. One is that it says nothing about predictability, and the other is that for true random the definition states an infinite number of conditions. It is impossible of course for any periodic process to satisfy all these conditions, and it is equally impossible to set up a law of generation (reproducible) which will satisfy all of them. It is not practicable for a law of generation to satisfy more than a few of the conditions. A standard procedure for a cryptanalyst is to check his material against these conditions one at a time, the simplest first. Experience has shown that usually some of the simpler conditions are not satisfied, and the way these conditions are violated generally gives clues to the law of generation. Only the work of checking the conditions, and the small size of the sample of key, prevents this from being a general solution.

~~CONFIDENTIAL~~

We see then that the cryptographer is on the defensive, trying to protect his material on each of an infinite number of sides with only finite, even severely limited, resources. The cryptanalyst, being on the offensive, has only to find a side which is undefended.

~~CONFIDENTIAL~~