Non - Responsive

The Unbreakable Cipher

BY H. CAMPAIGNE

Unclassified

Does there exist an unbreakable cipher?

Edgar Allan Poe said "it may be roundly asserted that human ingenuity cannot concoct a cipher which human ingenuity cannot resolve."

Professor Littlewood, in *A Mathematician's Miscellany* (Methuen London 1957, p. 23) states: "The legend that every cipher is breakable is of course absurd, though still widespread among people who should know better. I give a sufficient example, without troubling about its precise degree of practicability. Suppose we have a 5-figure number N. Starting at a place N in a 7-figure log-table take a succession of pairs of digits d_1d_1', d_2d_2', . . . from the last figures of the entries take the remainder of the 2-figure number d_nd_n' after division by 26. This gives a 'shift' s_n, and the code is to shift the successive letters of the message by s_1, s_2, . . . respectively.

"It is sufficiently obvious that a *single* message cannot be unscrambled, and this even if all were known except the key number N (indeed the triply random character of s_n is needlessly elaborate). If the same code is used for a number of messages it could be broken, but all we need do is to vary N. It can be made to depend on a date, given in clear; the key might, e.g., be that N is the first 5 figures of the 'tangent' of the date (read as degrees, minutes, seconds: 28°12′52″ for Dec. 28, 1952). This rule could be carried in the head, with nothing on paper to be stolen or betrayed. If any one thinks there is a possibility of the entire scheme being guessed he could modify 26 to 21 and use a date one week earlier than the one given in clear."

I think it is clear to all of you that this system would not stand very many messages per day. But let us look at some key from a ten-place table (page 40). Reduction mod 26 was omitted, but the redundancy you note here is only mildly disguised by that step. The width of seven was selected because of the interesting vertical differences.

David Kahn, "Lyen Otuu Wllwgh Wl Etjown" pp. 71, 83, 84, 86, 88 and 90 of the *New York Times Magazine* November 13, 1960 says that an umbreakable cipher system can be made from one time key "that is absolutely random and never repeats." He suggests that key can be derived from an almanac by taking a table, such as the populations of towns. There are 1570 towns with populations more than 2500 listed in the Almanac, which gives enough key for 3 short (100 group) messages.

Ten-Place Logarithm Table

Vol. Two

Berlin 1919 Ungar 1957

79	09	74	92	56	53	05	91	22
06	25	89	97	39	35	66	41	60
34	42	95	91	32	18	37	01	10
72	79	11	96	26	91	19	62	69
11	07	37	02	20	84	81	33	29
50	34	53	17	15	67	63	94	79
98	62	70	22	19	50	35	65	29
47	90	96	38	13	43	17	26	88
86	27	13	43	17	26	89	86	38
24	54	29	48	01	19			

What is the truth? Any fool can plainly see that either every cipher system is breakable, or there must be at least one which is unbreakable. Nevertheless, neither is true. The reason is that although the sentences parse correctly and have all the form of making sense, the words are not used properly. It is like the paradox of self-descriptive adjectives. "Short" is a short word. "Polysyllabic" has many syllables. Then there are words which do not describe themselves, such as "long" which is a short word, and "colorful" which has no color of its own, and these are non-self-descriptive. Does the adjective "non-self-descriptive" describe itself? If it does then it doesn't and if it doesn't then it must. Although the question has the form of decent language it has no content. In much the same way the question of the existence of an unbreakable cipher is without meaning.

In order to see what the truth is we must examine very carefully the meaning of each word we use. I am going to start by defining a "communications system." It is a system for sending messages from one point to another on demand. Each message is of finite length. The point to be stressed is that the messages to be sent are not under the control of the cryptographer; he must be prepared to encipher, but he cannot put limitations on the traffic. For our purposes we can suppose the messages come from a stochastic process with some known distribution.* The cryptographer might like to require that there be a limit on the redundancy of each message, but he cannot do this. He must provide in advance for all contingencies. Unlikely as it is to occur, he must be prepared for the message with 10,000 consecutive repetitions of the letter "a". Of course the cryptographer does have some control in that he can prescribe changes under certain conditions. Long messages can be required to go in parts. When the volume of traffic accumulates to a certain level, or the number of successive "a's" reaches a certain bound, resort to a reserve cipher may be mandatory. So one cipher can be protected by exposing another in its stead.

*In fact the distribution is not known to me, and it might be interesting to know more about it (or them, as there must be many).

Next I define a "cipher system" as a message. Let me tell you why I say this. For two friends, A and B, to communicate in the presence of an enemy without that enemy understanding, even given time to study, it is necessary that the friends have some information, some prearrangement, not available to the enemy. If the enemy has exactly the same information as B then their positions are symmetrical, and a message understood by one must be understandable to the other. The only other way conceivable would be if A and B were sufficiently more intelligent than the enemy, as though they were talking in front of a malignant child. Even this will not do, for a persistent even though stupid enemy will eventually make out the meaning. The child grows up. For our purposes here we will assume that the enemy has intelligence. Now this information shared by *A* and *B* must be communicated in some way, by a courier, by some other communications system, or perhaps a trip, and this communication is a message. As a message a cipher system is seen to be subject to certain restrictions which have been stated elsewhere under information theory. Whether the cipher system is given by a five-letter group preceding the text or by a shipload of one-time pads these restrictions hold.

A consequence of this definition is that each cipher system must be describable in a finite number of words. This does not imply that the system must be bounded or must repeat, for consider a cipher such as this: the key is the decimal digits of the number π, which are infinite in number. There are however only a limited number of irrationals which can be named as we name π with a short designation. Therefore the enemy can try these one by one. If there was a way to designate an arbitrary irrational with a small number of symbols then this system would be very secure. But information theory tells us this is impossible. The system can only be secure for a finite amount of traffic.

We can put that into a formal statement. For each cipher system there is an upper bound to the amount of traffic it can protect against cryptanalytic attack. What is "cryptanalytic attack"? It is a process applied to cipher text in order to extract information, especially information contained in the messages and intended to be kept secret. If some of the information is gotten by other means and this results in more being extracted from the cipher, this is (at least partially) a successful attack. If certain phrases can be recognized when they are present, this is successful cryptanalysis. If a priori probabilities on possible contents are altered by examination of the cipher, this is cryptanalytic progress. If in making trial decipherments it is possible to pick out the correct one then cryptanalysis is successful. Some of these procedures may be possible but impracticable because of time or expense; in this paper we will ignore such considerations and imagine that both cryptographer and cryptanalyst have all the computing facilities they require.

Consider for example the system just described, the digits of π. It would not take a very long crib to reveal a recognizable piece of the number. "3.14159" would give it away, only six digits. The work of preparing digits far out in the expansion would be large, but we will pretend that it can be done. Therefore an upper bound in this system is established by the danger of a crib long enough to betray the origin of the key, say 100 words.

Another example is that of Mr. Kahn, one-time key. Here the limit is quite clear; it is the amount of key on hand. The key arrives in finite "messages," so there is only a finite amount on hand at any one time, and this limits the amount of traffic which can be sent securely. Of course another shipment of key raises this bound, but technically another cipher system is now in effect, for by my definition a cipher system is a message. A sequence of messages is a sequence of cipher systems, related perhaps, but not the same.

Any other system has an upper bound which can be estimated in the same way. CSEC is well-practised in making just such estimates. Finding the least upper bound is another matter; it cannot be done I think without sharpening definitions more than I have. Finding a least upper bound—even in a very specific case— would be very interesting.

My answer to the question, "Does there exist an unbreakable cipher" would be this, "Every cipher is breakable, given enough traffic, and every cipher is unbreakable, if the traffic volume is restricted enough."

Non - Responsive