THE REAL REASON FOR RIVEST'S PHENOMENON

Don Coppersmith
IBM Research
Yorktown Heights, NY 10598

Burt Kaliski, Ronald Rivest and Alan Sherman [Crypto 85] noticed a short cycle in their experiments with weak keys in DES. We explain this in terms of fixed points (messages which are left unchanged by encipherment). We predict similar short cycles using semi-weak keys. We indicate how Rivest *et al*'s experimental setup can be used to show that the group of permutations of message space, generated by DES encryptions, is a large group.

Notation: Let $E_K X$ denote the ciphertext resulting from DES-encrypting the cleartext X under the key K. Similarly let $D_K X$ represent decryption. Let 0 denote the key of all 0's, and 1 the key of all 1's. Let the input to DES (the cleartext) be broken into halves M_0, M_1. On round i, $1 \le i \le 16$, we compute some function f of the 48 key bits K_i and the 32 message bits M_i, add this 32-bit quantity to M_{i-1} bitwise, and obtain M_{i+1}. So $M_{i+1} = M_{i-1} + f(K_i, M_i)$, and $M_{i-1} = M_{i+1} + f(K_i, M_i)$. The ciphertext is the pair M_{17}, M_{16}. (Notice the order of indices, which is correct.)

The keys 0 and 1 (and two others) are known to be *weak keys* [Davies, Crypto 82] in the sense that the 48 key bits K_i entering into the computation on round i are the same for each round i: $K_i = K_j$. One consequence of this is that E_0 is an involution: $E_0 X = D_0 X$, so that $E_0^2 X = X$.

A new consequence of being a weak key is that E_0 has 2^{32} fixed points, i.e. messages Y for which $E_0 Y = Y$. Indeed, for some message Y, suppose that $M_8 = M_9$. (There are 2^{32} such values of Y.) Then

$$M_7 = M_9 + f(K_8, M_8) = M_8 + f(K_9, M_9) = M_{10}.$$

Continuing, we find $M_6 = M_{11}, \ldots, M_1 = M_{16}, M_0 = M_{17}$, and $Y = (M_0, M_1) = (M_{17}, M_{16}) = E_0 Y$. In fact this is the only way a fixed point can arise for any weak key.

Now pick a random starting message X, and alternately apply E_0 and E_1. Continue until you return to the starting point X: $(E_1 E_0)^N X = X$. In Rivest *et al*'s experiment, N turned out to be around 2^{32}. Indeed, suppose that for some $I < N$, $(E_1 E_0)^I X = Y$ and Y is a fixed point of E_0. Then the next application of E_0 leaves Y unchanged, so that $E_0 (E_1 E_0)^I X = Y$. On the next application of E_1, we find

$$(E_1 E_0)^{I+1} X = E_1 E_0 (E_1 E_0)^I X = E_1 Y = D_1 Y = D_1 E_1 E_0 (E_1 E_0)^{I-1} X = E_0 (E_1 E_0)^{I-1} X.$$

Continuing, for $J \le I$, $(E_1 E_0)^{I+J} X = E_0 (E_1 E_0)^{I-J} X$, and we are just retracing our steps. This is because both E_0 and E_1 are involutions. We continue until, for some $J > I$, $(E_1 E_0)^{I+J} X = Z$ and Z is another fixed point of E_0. (We could also find fixed points of E_1.) We again end up retracing our steps, until we return to the starting value X.

H.C. Williams (Ed.): Advances in Cryptology - CRYPTO '85, LNCS 218, pp. 535-536, 1986.

The cycle length N is approximately the number of trials we needed to find two fixed points (of either E_0 or E_1.) Since these fixed points are plentiful (2^{32} out of 2^{64}, or 1 out of 2^{32}), the expected value of N is 2^{32}, in close agreement with Rivest *et al*'s results.

In a similar vein, suppose K is alternating 010101... or 101010... in each key half (a special case of the "semi-weak" keys [Davies]). Let $\overline{Z}$ denote the complement of Z. Then we find that $E_K = D_{\overline{K}}$; there are 2^{32} values Y for which $E_K Y = \overline{Y}$ (namely those for which $M_8 = \overline{M}_9$); and for a random X we expect to find that $E_K^N X = X$ for $N \simeq 2^{33}$.

Finally, for different starting values X_i, we expect to find different cycle lengths N_i. Consider the subgroup G of the group of permutations on message space $S_2{}^{64}$ generated by the DES encryptions E_K, $K \in Z_2^{56}$. Each N_i divides the size of G. Run either of the above experiments several times, finding different values N_i corresponding to $E_1 E_0$ or to E_K for one of the four alternating semiweak keys K. Each experiment takes a few days. Then the least common multiple $lcm(N_1, N_2, \ldots, N_i)$ divides the order of the group, and thus provides a lower bound. So the experiments, which were designed to detect a small group size ($|G| < 2^{70}$?) might be used to show a large group size ($|G| > 2^{300}$?).