ON BREAKING THE ITERATED MERKLE-HELLMAN PUBLIC-KEY CRYPTOSYSTEM

Leonard M. Adleman*

University of Southern California and
Massachusetts Institute of Technology

I. Introduction

In 1976 Diffie and Hellman introduced the concept of a public-key cryptosystem [1]. In 1977 Rivest, Shamir, and Adleman discovered the first incarnation of such a system [4], and soon afterwards Merkle and Hellman produced a second one [3]. Despite the widespread interest in the area, the years have produced no other public-key cryptosystems which have attracted widespread interest.

The Merkle-Hellman system is based on the knapsack problem, and in the original paper on the topic, both a basic method and an iterated method were presented. The iterated method was introduced "for improving the security and utility of the basic method." In April of 1982, Adi Shamir demonstrated that the basic knapsack cryptosystem was insecure [5]. In addition, Shamir states that the most important remaining open is the cryptographic security of the iterated systems. In this paper, we build upon Shamir's results to establish the insecurity of the iterated systems as well.

Our method of attack uses recent results of Lenstra and Lovacz [2]. We treat the cryptographic problem as a lattice problem, rather than a linear programming problem as in Shamir's result. Like Shamir, we are unable to present a rigorous proof that the algorithm works. However, an analysis of the algorithm in the presence of "reasonable

*Research sponsored by National Science Foundation, Grant #MCS-8022533

assumptions" will be presented. We will be particularly concerned with defining a broad class of iterated systems where the algorithm is virtually guaranteed to work. Our Analysis involves consideration of the Lovacz-Lenstra algorithm for finding "almost" the smallest vector in a latice.

In addition we will apply a varient of the algorithm to the Graham-Shamir public-key cryptosystem -- a variant of the Merkle Hellman system. Finally, since the algorithm has actually been implmented, examples of is performance will be given.

II. Iterated Knapsack Systems

Public-key cryptosystems require the generation of a "mated pair" of keys. One key is kept secret, the other is made public. It is crucial that the problem of computing the secret key from the public key be intractable In the iterated knapsack systems this is apparently not the case. Below is a description of the procedure used to generate a mated pair of keys for such a system. How these keys are used for encryption and decryption will not concern us.

Step 0

Generate a sequence of natural numbers $a_{0,1}, a_{0,2}, \ldots, a_{0,n}$ such that

$$a_{o,i} \geq \sum_{j=1}^{i-1} a_{0,j} \qquad i = 2, 3, \ldots, n$$

(such a sequence is said to be "super-increasing").

Step 1

Generate numbers W_1, M_1 such that

a) $M_1 \geq \sum_{i=1}^{n} a_{0,i}$

b) $(W_{1,} M_1) = 1$.

Compute

$$a_{1,i} \equiv W_1 a_{0,i} \text{MOD}(M_1)$$

$$\vdots$$

Step z

Generate numbers W_z, M_z such that

a) $M_z \geq \sum_{i=1}^{n} a_{z-1,i}$

b) $(W_z, M_z) = 1$

Compute

$$a_{z,i} \equiv W_z a_{z-1,i} \text{MOD}(W_z)$$

The Secret Key is

$$<<a_{0,1}, \ldots, a_{0,n}>, M_1, W_1, M_2, W_2, \ldots, M_z, W_z>$$

The Public Key is

$$<a_{z,1}, a_{z,2}, \ldots, a_{z,n}>$$

(In fact, the public key is a permutation of the above sequence. This is not a serious problem for the breaking algorithm, but will be ignored in this abstract for the sake of a clearer exposition.)

III. Breaking the Iterated Knapsack

Below we outline the method for breaking the iterated knapsack. Many details are omitted for the sake of clarity.

We will show how to recover $W = W_z$ and $M = M_z$. By iterating the process, the entire secret key can be obtained.

Let $L = L_z$ denote the inverse of W MOD(M). Clearly it is enough to recover L, M.

We know the following system of congruences holds.

$$[\text{SI}] \quad La_{z,i} \equiv a_{z-1,i} \text{MOD}(M) \qquad i = 1, 2, \ldots, b$$

(b depends on certain parameters used in constructing the system under attack. For a typical real system $b \simeq 6$.)

Rewriting SI as a system of equalities we get

$$[\text{SII}] \quad La_{z,i} - K_iM = a_{z-1,i}$$

for some natural numbers K_i.

The $a_{z,i}$'s are known, since they are part of the key, but all other quantities are unknown. We could solve SII for L and M and the K_i's except that,.

Problem 1

The system is underdetermined and has infinitely many sets of solutions and we need a way of distinguishing the correct one.

Problem 2

The system is non-linear (K_iM terms) and so we have no polynomial time algorithm to solve it anyway.

Curiously Problem 1 will provide a solution to Problem 2.

However, before solving Problem 2, we will simplify SII a bit, by removing some of the unknowns. Note that M is larger than the $a_{z-1,i}$'s by construction. The larger M is with respect to the $a_{z-1,i}$'s the better for the breaking algorithm. Just how much larger is enough will be analyzed in the paper. For now we will assume there is a fixed d (known to us) such that the $a_{z-1,i}$'s $\leq M/d$. This allows us to use SIII below.

$$[\text{SIII}] \quad La_{z,i} - K_iM \leq M/d$$

Even with SIII, however, Problems 1 and 2 remain.

Problem 1 says that in addition to the desired solution with L and M, SIII has infinitely many solutions which are undesirable. In fact, solutions are so abundant that we can be sure there will be a solution when

2^c (for a c easily computable from the public key) is substituted for M. That is we know

$$[\text{SIV}] \quad La_{z,i} - K_i 2^c \leq 2^c/d$$

is solvable.

SIV has some good properties and some bad ones.

1. (Good) Since the $a_{z,i}$'s (and c and d) are known, SIV is linear and can be solved (in many cases) using Lovacz-Lenstra.

2. (Good) It can be heuristically argued that the solution for L and the K_i's is unique.

3. (Bad) When solved, the value for L obtained has no useful relationship to the L we want (because the wrong M was used).

4. (Good) The K_i's which are obtained are correct. That is, we can show that they are the same K_i's that we would get if the correct M had been used!

   Therefore we can substitute these K_i's into SIII and remove the non-linearity.

Eliminating Problem 1 requires transformation of the problem in a way which will not be described in detail here. The end effect will be that we will be able to augment system SIII with new linear inequalities which state, in essence, that the values $La_{z,i} - K_i M$ ($= a_{z-1,i}$) are not only small but "special" in that they are either super-increasing ($z = 2$) or the iterates of some super increasing sequences. Finally it will be argued that the augmented system has a unique solution -- the desired L and M.

References

[1] Diffie, W., and Hellman, M. E., "New directions in cryptography," IEEE Trans. Information Theory, IT-22, Nov. 1976, pp. 644-654.

[2] Lenstra, A. K., Lenstra, H. W., and Lovacz, L., "Factoring polynomials with rational coefficients," Report 82-05, March 1982, Department of Math., Univ. of Amsterdam.

[3] Merkle, R. C., and Hellman, M. E., "Hiding information and signatures in trapdoor knapsacks," *IEEE Trans. Information Theory*, IT-24, Sept. 1978, pp. 525-530.

[4] Rivest, R., Shamir, A., and Adleman, L., "A method for obtaining digital signatures and public-key cryptosystems," *CACM* 22-2, Feb. 1978.

[5] Shamir, A., "A polynomial time algorithm for breaking th basic Merkle-Hellman cryptosystem," *Proc. 23rd Annual Foundations of Computer Science*, 1982.

Rump Session: Impromptu Talks by Conference Attendees