MysteryTwister C3

THE CRYPTO CHALLENGE CONTEST

# Merkle-Hellman Knapsack Challenge – Part 1

Author: Peter Uelkes

October 2022

# The Merkle-Hellman Knapsack (1/6)

In the 1970s, Ralph Merkle and Martin Hellman played a decisive role in developing the idea of public-key encryption (PKE).
In symmetric encryption, the key must remain secret, and thus a secure exchange of keys is a critical security aspect.
In PKE, a part of the key is public and can be used by all parties [2]. Well-known examples of a PKE system are RSA, Diffie-Hellman and ElGamal.

# The Merkle-Hellman Knapsack (2/6)

An early example of *public-key* encryption is the *Merkle-Hellman knapsack* [1]. This is used relatively little in practise as Shamir and others have found attack vectors and the method is therefore considered insecure [3].

In this challenge, we first consider a simple example of the elementary Merkle-Hellman knapsack in a form that is only suitable for symmetric encryption, i.e., a secret key. In Part 2 of the series, we will then move on to the public-key variant.

# The Merkle-Hellman Knapsack (3/6)

Our simplified Merkle-Hellman knapsack works as follows: Alice wants to send the message P="MTC3" encrypted to Bob. She converts this into a decimal number $M$ by concatenating the corresponding ASCII codes.
Alice now knows $M = 77846751$ (77=M, 84=T, ...).

Now she converts $M$ to binary and gets
$B = 100101000111101100011011111$.
The binary number in this example has a length $\lambda = 27$.

Alice then chooses the "knapsack" $K$ as a list of $\lambda$ disjoint natural numbers, e.g.:

$$K = [593, 13, 6252407, 1327, 958200, 2568118886, 51234, 24023, 6, 3470, 54136509, 160, 930178262, 3, 27, 102299137, 82, 15267027, 395469607, 117674, 2064250, 5494084005, 18363310574, 337256, 2, 38825629419, 7920]$$

# The Merkle-Hellman Knapsack (4/6)

In the final step of the encryption, Alice now adds up all those values in K to the ciphertext C whose corresponding element in B is a 1. These values are highlighted in red:

$$\mathrm{K} = [593, 13, 6252407, 1327, 958200, 2568118886, 51234, 24023, 6, 3470, 54136509, 160, 930178262, 3, 27, 102299137, 82, 15267027, 395469607, 117674, 2064250, 5494084005, 18363310574, 337256, 2, 38825629419, 7920]$$

Thus:

$$\mathrm{C} = 593 + 1327 + \ldots + 7920 = 60846205466$$

C is the ciphertext Alice sends to Bob.

# The Merkle-Hellman Knapsack (5/6)

Bob has received the ciphertext $C = 60846205466$ from Alice and wants to convert it back to plaintext. He knows the knapsack $K$ (we are considering here the symmetric case) and must now find a selection of the elements of $K$ whose sum gives $C$. This is known in mathematics as *subset-sum problem* (SSP). The SSP is NP-complete and therefore cannot be solved efficiently for sufficiently large $\lambda$ [4].

However, there are instances of the subset-sum problem that have a special property and are therefore very easy to solve. This is the case with the example knapsack $K$ on the previous pages.

Once Bob has solved the subset-sum problem, he knows the binary number $B$, converts it back to the decimal number $M$ and obtains from it the plaintext $P$ via the ASCII encoding described.

# The Merkle-Hellman Knapsack (6/6)

Public-key cryptosystems are based on a *trapdoor function*, i.e., a mathematical function that is easy to calculate but whose inverse (without knowledge of an additional hint) is not efficiently possible to find for large enough numbers:

- **RSA**: The multiplication of primes $p$ and $q$ to $N = p \cdot q$ is trivial; the factorization of $N$ is hard.
- **Diffie-Hellman, El Gamal**: Exponentiation in residue classes (i.e., calculation of $a^b \pmod{n}$ with $a, b, n \in \mathbb{N}$) is easy; the inversion (discrete logarithm) is difficult.
- **Merkle-Hellman knapsack**: The summation of the selected knapsack elements is trivial; the solution of the subset-sum problem is not.

# Challenge

An English plaintext was encrypted with the described symmetrical knapsack method. The knapsack $K$ used has a special property, that makes the subset-sum problem easily solvable.

$K$ and the ciphertext $C$ are in the file `knapsack_part1_add.txt`. The length of $B$ (binary representation of $M$) and thus also of $K$ is $\lambda = 179$.

Decrypt the ciphertext. The solution to this challenge is the plaintext in the exact notation resulting from the ASCII encoding (including any spaces and punctuation).

# Resources


1. Wikipedia article: en.wikipedia.org/wiki/Merkle-Hellman_knapsack_cryptosystem
2. David Kahn: "The Code-Breakers", revised version from 1996
3. M. Stamp, R. Low: "Applied Cryptanalysis - Breaking Ciphers in the Real World", 2007
4. Wikipedia article on the subset-sum problem: en.wikipedia.org/wiki/Subset_sum_problem

# Additional files

→ `knapsack_part1_add.txt`: The ciphertext C and the knapsack K of the challenge.