The Subset Sum Problem

The subset sum problem is the following: you are given a target number, and a modulus like this:

TARGET = 40541043 mod 100000000

and told that the target number is the sum of some number of elements from a *public* list of numbers:

LIST_SIZE = 24

NUM_SUMMED = 12

THE_LIST = [46989681, 12273567, 87570662, 32941824, 9215565, 92034381, 36189443, 6283813, 98906035, 26382035, 32838757, 52550259, 90663083, 22659704, 78749464, 51642241, 44717509, 59893817, 58843431, 49911223, 89019133, 77713592, 43514002, 94045560]

The problem is to identify which of the listed numbers were summed, i.e. create a list of indices of size NUM_SUMMED for which

$$\sum_{j=1}^{m} L_{i_j} = T \mod N$$

where T, N, L, m are the target, modulus, list and num_summed, respectively.

Notes

The subset sum problem has a rich history. Some observations are:

- If the modulus is small, then there could be multiple solutions. We'll mostly be interested in the case when the elements of the list cannot be easily distinguished from random, and where the modulus is large enough for it to be very likely that there is only one solution.
- When NUM_SUMMED = LIST_SIZE/2 the number of possible subsets is maximized
- If NUM_SUMMED > LIST_SIZE/2, one would consider the items that are *not* in the sum, since the sum of all the numbers is public.
- If NUM_SUMMED is not provided, it is not too hard to enumerate over every possible NUM_SUMMED <= LIST_SIZE
- The subset sum problem was shown to be NP-complete by Karp [1], building on Cook's foundational work on NP-completeness [2]. See also [3].

References

[1] R. M. Karp, "Reducibility among combinatorial problems," in *Complexity of computer computations*, R. E. Miller and J. W. Thatcher, Eds., New York: Plenum Press, 1972, pp. 85–103.

[2] S. A. Cook, "The complexity of theorem-proving procedures," in *Proceedings of the 3rd annual ACM symposium on theory of computing*, ACM, 1971, pp. 151–158.

[3] M. R. Garey and D. S. Johnson, *Computers and intractability: A guide to the theory of NP-completeness*. New York: W. H. Freeman, 1979.