Faster Construction of Carmichael Numbers and Other Pseudoprimes

Andrew Shallue *Illinois Wesleyan University*

Introduction

The construction of large Carmichael numbers, as well as many other pseudoprimes, has traditionally relied on solving the subset product problem in some abelian group. We show how new algorithms for the subset-product problem, most notably [6] and [3], lead to improvements in many different constructions.

Erdös construction

Choose L to be a positive integer with many prime factors. Choosing the right L is an art and depends on the number sought, examples include products over primes such as

$$L = \prod_{p \leq B} p \qquad \text{and} \qquad L = \prod_{p \leq B} p^{\lfloor \log_p(B) \rfloor} \ .$$

With L chosen, the algorithm to construct a Carmichael number is as follows.

1. Construct the set P of all primes such that $p-1$ divides L, but $p \nmid L$.
2. Find a subset S of P that products to the identity in $(\mathbf{Z}/L\mathbf{Z})^\times$.

With $n = \prod_{p \in S} p$, L dividing $n-1$, and $p-1$ dividing L for all $p \mid n$, n satisfies the Korselt condition and hence is Carmichael.

To construct a Carmichael with many prime factors, let $\prod_{p \in P} p \equiv a \mod L$ and find a small subset S with $\prod_{p \in S} p \equiv a \mod L$. Then the product of the primes in $P \setminus S$ will be a Carmichael number.

Other pseudoprimes are constructed by modifying the definition of P, and/or changing the target of the subset product problem.

Subset product problem

Let G be an abelian group, and let $a_1, \ldots, a_n$ be elements of G. The *subset product problem* is to find a subset of the a_i that product to the identity in G (more generally, any element of G).

Definition 1 *The* density *of a subset product problem is given by*

$$\frac{n}{\log_2(|G|)} \ .$$

Solutions will be rare unless density is greater than 1. Problems with density 1 are the most difficult.

Until recently, algorithms for the subset product problem have been deterministic. The new algorithms to be discussed are randomized, and for correctness require the a_i to be independent and uniformly random elements of G. One can generally prove most choices of $(a_1, \ldots, a_n)$ act sufficiently like uniformly random elements. Here we simplify things by making the following assumption.

Heuristic Assumption The primes in P act like independent, uniformly random elements of $(\mathbf{Z}/L\mathbf{Z})^\times$.

Wagner's k-tree algorithm

This algorithm [6] for the subset product problem starts with k lists of uniformly random elements of G. Lists are combined in the shape of a binary tree, so that an element a of a child list is a product of two elements, one from each parent list so that a is in a subgroup of G. An element in the root list is a solution to the problem.

Assumptions: Problem density is at least $\frac{k}{\log_2(k)}$ (this is a new result of the author, was previously k). G has subgroups $G = H_0, H_1, \ldots, H_{\log_2(k)}$ such that $|H_i/H_{i+1}| \approx |G|^{1/(\log_2(k)+1)}$ and computable homomorphisms $\phi_i : H_i \to H_{i+1}$.

Theorem 1 (S, 2010) *Given assumptions above, there is a randomized algorithm for the subset-product problem (namely, Wagner's k-tree algorithm) that takes time and space $\widetilde{O}(k \cdot |G|^{1/(\log_2(k)+1)})$ and succeeds with probability exponentially close to $1/2$.*

The proof is based heavily on [5].

Problems with density close to one

Where the Wagner algorithm only works for problems of sufficiently high density, there is a new result that works for almost all problems of density greater than one. It is written for the subset sum problem, but easily carries over to the subset product problem.

Theorem 2 ([3]) *There is a randomized algorithm for the subset product problem that expects to find a solution using time and space $\widetilde{O}(2^{0.3113 \cdot n})$*

Carmichael with a billion prime factors

The current record is a Carmichael with 1101518 prime factors [4]. Here we outline a new technique and give a rough idea of the computation required by focusing on the main exponential term.

From [4], let L be the 168-bit number

$$L = 2^{15} \cdot 3^8 \cdot 5^5 \cdot 7^4 \cdot 11^3 \cdot 13^2 \cdot 17^2 \cdot 19^2 \cdot 23^2 \cdot 29 \cdot 31 \cdot 37 \cdot 41 \cdot 43 \cdot 47 \cdot 53 \cdot 59 \cdot 61 \cdot 67 \cdot 71 \cdot 73 \cdot 79$$

We expect the size of P to be at least 2^{30} (a 2^{35} calculation). Let the product of all elements of P be a. Pick out $168 \cdot 32 = 5376$ of the elements of P and solve the subset product problem corresponding to a. With density over 32, apply the 32-tree algorithm at a cost of roughly $32 * 2^{168/6} = 2^{33}$ time and space.

Other pseudoprimes

The same technique can be used to improve upon searches for other pseudoprimes found in [1].

Williams numbers: Let $\epsilon(N)$ be $\left(\frac{\Delta}{\mathrm{N}}\right)$, the Jacobi symbol applied to some fixed integer Δ. To construct a Δ-Lucas pseudoprime build the set

$$P_\Delta(L) = \{p \text{ prime}, p - \epsilon(p) \mid L, \ p \nmid L\} \ .$$

Let a be the product of all the primes in $P_\Delta(L)$. We seek a subset that products to $\pm a$ modulo L that also satisfies an additional condition on the Jacobi symbol. Using the k-tree algorithm is asymptotically faster than the method used in [1] ($\widetilde{O}(k \cdot \phi(L)^{1/(\log_2(k)+1)})$ versus $\widetilde{O}(\phi(L)^{1/2})$).

Elliptic pseudoprimes: In this case build the set

$$P_D(L) = \{p \text{ prime}, \left(\frac{\text{-D}}{\text{p}}\right) = -1, \ p+1 \mid L, \ p \nmid L\} \ .$$

Now use a k-tree algorithm to find a subset product with an odd number of elements that is congruent to -1 modulo L.

Can use similar work for strong pseudoprimes and strong Fibonacci pseudoprimes as well.

Higher order Carmichael numbers

See [2] for a definition and the following equivalent condition.

Theorem 3 *Composite n is a Carmichael of order m if and only if n is square free and for every prime divisor p of n and for every $1 \leq r \leq m$, there exists $i \geq 0$ with $n \equiv p^i \mod (p^r - 1)$.*

Build the set

$$S_m(L) = \{p \text{ prime}, p^r - 1 \mid L \text{ for all } 1 \leq r \leq m\}$$

and find a subset that products to 1 modulo L.

In this case the resulting problem has density close to 1. We apply the algorithm from Theorem 2, which improves upon the technique in [2] ($\widetilde{O}(2^{.3113 \cdot n})$ versus $\widetilde{O}(2^{.5 \cdot n})$).

References

[1] Dominique Guillaume and François Morain, *Building pseudoprimes with a large number of prime factors*, Appl. Algebra Engrg. Comm. Comput. **7** (1996), no. 4, 263–277.

[2] Everett W. Howe, *Higher-order Carmichael numbers*, Math. Comp. **69** (2000), no. 232, 1711–1719.

[3] Nick Howgrave-Graham and Antoine Joux, *New generic algorithms for hard knapsacks*, Proceedings of Eurocrypt, 2010.

[4] Günter Löh and Wolfgang Niebuhr, *A new algorithm for constructing large Carmichael numbers*, Math. Comp. **65** (1996), no. 214, 823–836.

[5] Lorenz Minder and Alistair Sinclair, *The extended k-tree algorithm*, SODA '09: Proceedings of the Nineteenth Annual ACM -SIAM Symposium on Discrete Algorithms (Philadelphia, PA, USA), Society for Industrial and Applied Mathematics, 2009, pp. 586–595.

[6] David Wagner, *A generalized birthday problem (extended abstract)*, Advances in Cryptology – CRYPTO 2002, Lecture Notes in Comput. Sci., vol. 2442, Springer, Berlin, 2002, pp. 288 – 303.