

Solving Subset Sum, Integer relations, Univariate Modular polynomial Equations Using LLL Algorithm

Mr.R.Santosh Kumar
Dept. of Information Technology
MVGR college of Engineering
Vizianagaram
santu_hcunitk@yahoo.co.in

Prof.C.Narasimham
Dept. of Information Technology
Siddhartha Engineering college
Vijayawada
narasimham_c@yahoo.com

Prof.S.P.Pallam Setty
Dept. of Computer Science & Systems Engineering
Andhra University
Vishakapatnam
drspsetty@yahoo.com
March 03, 2011

Abstract ---Lattices are regular arrangements of points in n -dimensional space, whose study appeared in the 19th century in number theory. Since the appearance of the celebrated Lenstra-Lovasz lattice reduction algorithm twenty years ago, lattices have had surprising applications in cryptology. In this paper we present some applications of LLL algorithm, which can be used to break some cryptosystems.

keywords: Lattice reduction, subset sum, coppersmith.

I. INTRODUCTION

The lattices have played an extremely important role in cryptology. In this introduction we state only the results about lattices that we will need later, and refer the reader elsewhere [COH95] for a comprehensive introduction.

Lattice Let $(v_1, v_2, v_3, \dots, v_n) \in \mathbb{Z}^n$, $m \geq n$ be linearly independent vectors. A lattice L spanned by $v_1, v_2, v_3, \dots, v_n$ is the set of all integer linear combinations of $v_1, v_2, v_3, \dots, v_n$

$$L = \{ V \in \mathbb{Z}^n | V = \sum a_i v_i \text{ with } a_i \in \mathbb{Z} \}.$$

If $m = n$, the lattice is called a full rank lattice. The set of vectors $B = (v_1, v_2, v_3, \dots, v_n)$ is called basis for L . We also say that L is spanned by the vectors of the basis B . We call $\dim(L) = n$, the dimension of L . The determinant of L is given by $d(L) = |\det (b_1, b_2, b_3, \dots, b_n)|$ the b_i being written as

row vectors. Reduced basis for a given lattice means basis vectors have small lengths and they span the lattice.

Lattice Reduction The goal of lattice basis reduction is this: given an integer lattice basis as input, to find a basis with short, nearly orthogonal vectors. Lattice reduction algorithms are used in a number of modern number theoretical applications. Although determining the shortest basis is possibly an NP-complete problem, algorithms such as the LLL algorithm can find a short basis in polynomial time with guaranteed worst-case performance. LLL is widely used in the cryptanalysis of public key cryptosystems. Several definitions of a reduced basis can be found in the literature for which there is no polynomial time algorithm. However for the definition given in [LLL82] there is a polynomial time, called LLL algorithm. Several definitions of a reduced basis can be found in the literature for which there is no polynomial time algorithm. However for the definition given in [LLL82] there is a polynomial time, called LLL algorithm.

LLL Reduced Let $b_1, b_2, b_3, \dots, b_n$ be a basis of lattice L , and $b_1^*, b_2^*, b_3^*, \dots, b_n^*$ be an orthogonal basis derived from the given vectors. The basis $b_1, b_2, b_3, \dots, b_n$ is called LLL reduced if

$$|\mu_{i,j}| \leq \frac{1}{2} \text{ for } 1 \leq j < i \leq n \quad (1)$$

$$|b_i^* + \mu_{i,j} b_j^*|^2 \geq \frac{3}{4} |b_{i-1}^*|^2 \text{ for } 1 < i \leq n \quad (2)$$

where $\|$ denotes the ordinary Euclidean length. The constant $\frac{3}{4}$ in second equation is arbitrarily chosen and may be replaced by any fixed real number y with $\frac{1}{4} < y < 1$.

LLL Algorithm Let L be a lattice spanned by $(u_1, u_2, u_3, \dots, u_w)$. The LLL algorithm given $u_1, u_2, u_3, \dots, u_w$ produces a new basis $b_1, b_2, b_3, \dots, b_w$ of L satisfying

$$(\|b_i^*\|)^2 \leq 2 (\|b_i^*\|)^2 \text{ for all } 1 \leq i < w \quad (3)$$

$$\text{If } b_i = b_i^* + \sum_{j=1}^{i-1} \mu_{i,j} b_j^*$$

$$\text{then } |\mu_j| \leq \frac{1}{2} \forall i \quad (4)$$

The algorithm performs $O(w^4 l)$ arithmetic operations, where $l = \log(\max_{i,j} \mu_{i,j})$ when $b_1, b_2, b_3, \dots, b_w$ is the output of the LLL algorithm on a basis for a lattice L , we say that it is an LLL reduced basis.

Properties of LLL reduced basis Let $b_1, b_2, b_3, \dots, b_n$ be a reduced basis for a lattice L in $\mathbb{R}^n$, and let $b_1^*, b_2^*, b_3^*, \dots, b_n^*$ be orthogonal vectors obtained from Gram-Schmidt orthogonalization process. Then we have

$$|b_j|^2 \leq 2^{i-1} |b_i^*|^2, \quad 1 \leq j \leq i \leq n \quad (5)$$

$$d(L) \leq \prod_{i=1}^n |b_i| \leq 2^{\frac{n(n-1)}{4}} d(L) \quad (6)$$

$$|b_1| \leq 2^{\frac{(n-1)}{4}} d(L)^{\frac{1}{n}} \quad (7)$$

In this paper we will use the property stated in [7].

II. APPLICATION-1

Solving subset sum problem of low density.

Given a set $\{a_1, a_2, a_3, \dots, a_n\}$ of positive integers called a knapsack set, and positive integers, determine whether or not there is a subset of the

a_j that sum to s . Equivalently, determine whether or not there exist $x_i \in \{0,1\}, 1 \leq i \leq n$ such that $\sum a_i x_i = s$. The density of S is defined to be $d = \frac{n}{\max\{\log a_i | 1 \leq i \leq n\}}$. The algorithm given below, reduces the subset sum problem to one finding a particular short vector in a lattice. Already we know that, the reduced basis produced by LLL algorithm includes a vector of length which is guaranteed to be within a factor of $2^{\frac{n-1}{2}}$ of the shortest non-zero vector of the lattice. In practice, however, the LLL algorithm usually finds a vector which is much shorter than what is guaranteed. Hence the LLL algorithm can be expected to find the short vector which yields a solution to the subset sum problem provided that this vector is shorter than most of the non zero vectors in the lattice.

Algorithm

1. Let $m = \frac{\sqrt{n}}{2}$.
2. Let L be a lattice generated by the rows of the matrix of order $(n+1) \times (n+2)$.

$$\begin{bmatrix} 2 & 0 & 0 & \dots & 0 & ma_1 & 0 \\ 0 & 2 & 0 & \dots & 0 & ma_2 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 & ma_n & 0 \\ 1 & 1 & 1 & \dots & 1 & ms & 1 \end{bmatrix}$$

3. Find a reduced basis B of L .
4. For each vector $y = (y_1, y_2, y_3, \dots, y_n, y_{n+1}, y_{n+2})$ in B , if $y_{n+1} = 0$, $y_{n+2} = 1$ and $y_i \in \{-1, 1\} \forall i = 1, 2, 3, \dots, n$ then do the following:
For those $y_i = 1$, if $\sum_{i=1}^n a_i x_i = s$ then return the solution $(x_1, x_2, x_3, \dots, x_n)$
For those $y_i = -1$, if $\sum_{i=1}^n a_i x_i = s$ then return the solution $(x_1, x_2, x_3, \dots, x_n)$
5. Return(failure) (Either no solution, or the algorithm has failed to find one).

Justification Let the rows of the matrix B be $b_1, b_2, b_3, \dots, b_n, b_{n+1}$ and L be the lattice generated by these vectors. If $x_1, x_2, x_3, \dots, x_n$ is a solution to the subset sum problem, Consider

$$y = \sum_{i=1}^n x_i b_i - b_{n+1}$$

$$= (x_1 b_1 + x_2 b_2 + x_3 b_3 + \dots + x_n b_n - b_{n+1})$$

$$(2x_1 - 1, 2x_2 - 1, \dots, 2x_n - 1, m(a_1 x_1 + a_2 x_2 + \dots + a_n x_n - s), 1)$$

Since $(x_1, x_2, x_3, \dots, x_n)$ is a solution and each x_i ($1 \leq i \leq n$) is either 0 or 1, we have $y_i \in \{-1, 1\}$ and $y_{n+1} = 0$. Since $\|y\| = \sqrt{y_1^2 + y_2^2 + \dots + y_{n+1}^2}$, the vector y is a vector of short length in L : If the density of the knapsack set is small, i.e. the a_i are large, then most vectors in L will have relatively large lengths, and hence y may be unique shortest non zero vector in L : If this is indeed the case then there is a good possibility of the algorithm finding a basis which includes this vector. Above algorithm is not guaranteed to succeed. Assuming that the LLL algorithm always produces a basis which includes the shortest non zero lattice vector, then algorithm succeeds with high probability if the density of the knapsack set is less than 0.9408.

Example Given integers are 3,5,7,9,11,13,20. The sum we want is 20. Here $m=2$. Consider the lattice L , which is spanned by the rows of the following matrix.

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 10 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 14 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 22 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 26 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 40 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 136 & 1 \end{pmatrix}$$

Apply LLL algorithm to above matrix, we get reduced matrix

$$\begin{pmatrix} 1 & 1 & -1 & 1 & 1 & -1 & 1 & 0 & 1 \\ 1 & 1 & 1 & -1 & -1 & 1 & 1 & 0 & 1 \\ 2 & -2 & 2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 3 & -1 & 1 & -1 & 1 & 0 & 1 \\ -1 & 1 & -1 & -3 & 1 & 1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 1 & 0 & -1 \\ -1 & -3 & -1 & 1 & 1 & 1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 & 1 & -1 & 1 & 7 & 1 \end{pmatrix}$$

Now observe that the rows 1,2,6 have desired property. Corresponding to those rows we can get possible positions for given problem are 3,6;4,5;7. We can apply this technique to cryptanalyze Merkle-Hellman cryptosystem

III. APPLICATION-2

Integer relations We say that there exists an integer relation among real numbers $x_1, x_2, x_3, \dots, x_n$ if there exists $a_1, a_2, a_3, \dots, a_n$ not all zero, such that $\sum_{i=1}^n a_i x_i = 0$. Although the Euclidean and continued fraction algorithms solve the problem of finding integer relations for the vector $x_1, x_2, x_3, \dots, x_n$ when $n = 2$. Until recently there were no known polynomial-time algorithms that solved the problem for $n \geq 3$, and it is likely that no algorithm exists in general.

A breakthrough was made in 1977 with the generalized Euclidean algorithm of Ferguson and Forcade, a recursive algorithm that is guaranteed to find an integer relation when one exists. Following this, a number of non recursive algorithms were developed, including the PSLQ algorithm, the HJLS algorithm and a method based on the LLL algorithm. One use of the LLL algorithm is to find small integer relations among nonzero values $x_1, x_2, x_3, \dots, x_n$. For the LLL algorithm to find integer relation for

given non zero values, define $(n + 1) * n$ lower trapezoidal matrix B as

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ Nx_1 & Nx_2 & \dots & Nx_{n-1} & Nx_n \end{pmatrix}$$

(where N is a large number) and let be the column vectors of B . If we now consider the vectors in the

lattice L spanned by $b_1, b_2, b_3, \dots, b_n$, we see they of the form $M = \sum_{i=1}^n m_i b_i = [m_1, m_2, m_3, \dots, m_n, N \sum_{i=1}^n m_i x_i]^T$.

We may view the last term in M , as a penalty term. If the vector $m = [m_1, m_2, m_3, \dots, m_n]$ is an integer relation for x , then this term will be large, provided that N is large enough. The penalty for not being an integer relation depends on the choice of N . If N is taken large enough and m is a short integer relation for x , then M will be one of the shortest vectors in L . With this in mind, to find an integer relation for x we choose a suitably large value of N and run the LLL algorithm on the vectors $b_1, b_2, b_3, \dots, b_n$. The first vector in the returned basis will be one of the smallest vectors in L . If N is large and if an integer relation exists, then LLL will succeed.

Algorithm This algorithm finds small integer relation among given numbers.

- 1) Let the input vectors $x_1, x_2, x_3, \dots, x_n$.
- 2) Let L be the lattice spanned by columns of a matrix $(n+1) \times n$ B as

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ Nx_1 & Nx_2 & \dots & Nx_{n-1} & Nx_n \end{pmatrix}$$

Where N is a predefined number

- 3) Apply LLL algorithm to B .
- 4) Return first vector of the reduced basis.

Justification Let M be the norm of a smallest integer relation for x and consider the finite set of vectors $\{y \in \mathbb{Z}^n : \|y\| < 2^{\frac{n}{2}} M, y \cdot x = 0\}$. From this non empty set, choose a vector y with the property $|y \cdot x| = |\sum x_i y_i|$ is minimal. For this y , choose N so that $N |\sum y_i x_i| > 2^{\frac{n}{2}} M$. Now for any $m' \in L$, if $\sum_{i=1}^n m_i x_i \neq 0$ then

$$\|m'\| > 2^{\frac{n}{2}} M. \text{ If } m = [m_1, m_2, m_3, \dots, m_n, N \sum_{i=1}^n m_i x_i]^T$$

is greater than $2^{\frac{n}{2}}$ times the norm of a shortest non zero vector in L and hence cannot be the first vector in LLL reduced basis.

Example Given numbers are 56, 78, 89, 21, 45. Construct lattice L which is spanned by the rows of the matrix.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 16184 \\ 0 & 1 & 0 & 0 & 0 & 22542 \\ 0 & 0 & 1 & 0 & 0 & 25721 \\ 0 & 0 & 0 & 1 & 0 & 6069 \\ 0 & 0 & 0 & 0 & 1 & 13005 \end{pmatrix}$$

Apply LLL algorithm to above matrix, we get a reduced basis

$$\begin{pmatrix} -1 & -1 & 1 & 0 & 1 & 0 \\ -2 & 2 & -1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 & 2 & 0 \\ 2 & 1 & -2 & -7 & 3 & 0 \\ -1 & 1 & 0 & -1 & 0 & 289 \end{pmatrix}$$

The first 4 rows of the reduced basis gives the desired result.

Example We are aware of Machin's formula

$a \arctan(1) + b \arctan\left(\frac{1}{5}\right) + c \arctan\left(\frac{1}{239}\right) = 0$ where a, b, c are small numbers, but we do not know their values. We apply LLL to the matrix

$$\begin{bmatrix} \arctan(1)A & \arctan\left(\frac{1}{5}\right)A & \arctan\left(\frac{1}{239}\right)A \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

With a large value of A . If A is not large enough, LLL suggests that $\arctan\left(\frac{1}{239}\right) \approx 0$. It is clearly true, but not exactly 0. If $A > 1000$, LLL suggests the relation

$$\arctan(1) - 4 \arctan\left(\frac{1}{5}\right) + \arctan\left(\frac{1}{239}\right) = 0$$

IV. APPLICATION

Finding small solutions to univariate polynomial congruences. Suppose we are given a polynomial $f(x)$ and a real number M , and we wish to find a value $x_0 \in \mathbb{Z}$ for which $f(x_0) \equiv 0 \pmod{M}$. The main tool we use is stated in the following simple fact.

Let $h(x) = \sum a_i x^i$ then $\|h(xX)\|^2 = \sum_i X^i a_i^2$.

Fact Let $h(x) \in R[x]$ be a polynomial of degree w , and let $X \in \mathbb{R}$ be given. Suppose there is some $|x_0| \in X$ such that

1. $h(x_0) \in \mathbb{Z}$, and
2. $\|h(xX)\| < \frac{1}{\sqrt{w}}$. Then $h(x_0) = 0$.

Consider,

$$\begin{aligned} |h(x_0)| &= |\sum a_i x_i| \\ &= \left| \sum a_i X^i \left(\frac{x_0}{X}\right)^i \right| \\ &\leq \sum \left| a_i X^i \left(\frac{x_0}{X}\right)^i \right| \\ &\leq \sum |a_i X^i| \\ &\leq \sqrt{w} \|h(xX)\| \\ &< 1 \end{aligned}$$

But since $h(x_0) \in \mathbb{Z}$ we must have $h(x_0) = 0$.

Above fact suggests that we should look for a polynomial $h(x)$ of small weighted norm satisfying $h(x_0) \in \mathbb{Z}$. To do this we will build a lattice of polynomials related to f and use LLL to look for short vectors in that lattice. We have $f(x_0)/M \in \mathbb{Z}$ because $f(x_0) \equiv 0 \pmod{M}$.

Define $g_{i,k}(x) = x^i (f(x)/M)^k$. And we have $g_{i,k}(x_0) = x_0^i (f(x_0)/M)^k \in \mathbb{Z} \forall i, k \geq 0$. Furthermore, this is true for all integer linear combinations of the $g_{i,k}(x)$. The idea behind the coppersmith technique is to build a lattice L from $g_{i,k}(xX)$. The first vector b_1 The first vector b_1 returned by the LLL algorithm will be a low norm polynomial $h(xX)$ also satisfying $h(x_0) \in \mathbb{Z}$. If its norm is small enough, by above fact we have $h(x_0) = 0$. Traditional root finding methods can find x_0 . By above fact we have

$\|h(xX)\| < \frac{1}{\sqrt{w}}$. Fortunately, LLL algorithm allows us to compute a good bound on the norm of the first vector in an LLL reduced basis. We have $\det(L) < 2^{\frac{-w(w-1)}{4}} w^{\frac{-w}{2}}$. From this we can obtain $\|h(xX)\| < \frac{1}{\sqrt{w}}$. The determinant of L depends on the choice of polynomials $g_{i,k}(xX)$ defining lattice. To compute determinant of a lattice one should choose basis polynomials carefully. We will be able to choose a basis so that the matrix whose rows are the coefficient vectors of $g_{i,k}(xX)$ is full-rank and diagonal, with an explicit formula for the entries.

Example Suppose we wish to find a root x_0 of the polynomial

$$x^2 - 2849x + 5324 \equiv 0 \pmod{10001} \text{ satisfying } |x_0| \leq 17.$$

Define $f(x) = (x^2 - 2849x + 5324)/10001$. We build a lattice with basis polynomials $(1, 17x, f(17x), 17xf(17x), f^2(17x))$. The determinant of the lattice of matrix formed by using above basis of polynomials we get

$$\det(L) = 17^{10} 10001^{-4} \approx 2.0 \times 10^{-4}.$$

We require

$$\det(L) < \gamma \text{ where } \gamma = 2^{\frac{-5(5-1)}{4}} 5^{\frac{-5}{2}} \approx 5.6 \times 10^{-4}$$

Hence condition is satisfied. We find that the LLL algorithm returns polynomial

$$h(17x) = \frac{(-417605x^4 - 7433369x^3 + 1970691x^2 - 2625174x + 7250016)}{10001^2}.$$

This leads to

$$h(x) = \frac{(-5x^4 - 1513x^3 + 6819x^2 - 154422x + 7250016)}{10001^2}.$$

The roots of $h(x)$ over the reals are $16, -307.413$. We find the only integer solution x_0 for given equation satisfying $|x_0| \leq 17$ is x_0 is 16.

V. RESULT ANALYSIS

In this paper we have presented some applications of LLL algorithm. The first application is subset sum

problem. It is a well known NP-hard problem, but it was solved in low density by using LLL algorithm. Based on this, the well known knapsack cryptosystem namely Merkle-Hellman cryptosystem was broken. The second application is integer relations. To find integer relation among given real numbers is a difficult problem, but one can solve by using LLL algorithm. The other good solutions for this problem are PSLQ, HJLS. But no clarity to decide which one is best. The third application is to find small solutions to univariate modular polynomial equation. In this paper we have presented simple version of Howgrave-Graham. By using this problem we can solve the RSA problem in special settings.

VI. CONCLUSION

Lattice basis reduction is a powerful and general tool that has had a large impact on computational algebra, and is becoming increasingly important in cryptanalysis and theoretical computer science. These latter fields present an interesting tension. The cryptography community demonstrates that basis reduction has more application than its theoretical limits. Twenty years of lattice reduction yielded surprising applications in cryptology. We hope next twenty years will prove as exciting.

REFERENCES

- [1][COH95] Henri Cohen. *A Course in Computational Algebraic Number Theory*, Springer, 1996.
- [2][SHP05] Victor Shoup *A Computational Introduction to Number Theory and Algebra*, Cambridge, 2005.
- [3][KOB94] Neal Koblitz. *A Course in Number Theory and Cryptography*. Springer, 1994.
- [4][VAN96] Alfred J. Menezes, Paul C. Van Oorschot, and Scott A. Vanstone. *Hand Book of Applied Cryptography*. CRC Press, 1996.
- [5][LLL82] A.K. Lenstra, H.W. Lenstra, and L. Lovasz. *Factoring polynomials with rational coefficients*. Math. Ann., 261: 515-534, 1982.
- [6][NTLwww] Victor Shoup. *Ntl: A library for doing number theory*. Website: <http://www.shoup.net/ntl/>
- [7][COP96] D. Coppersmith. *Finding a small root of a univariate modular equation*. In Eurocrypt'96, 1996.
- [8][BRU06] Bruce Schneier. *Applied Cryptography*. Wiley, 2006.
- [9][POH97] M. Pohst. *A modification of LLL Reduction algorithm*. J. Symbolic Computation, 4:123-127, 1987.
- [10][STI06] D. Stinson. *Cryptography Theory and Practice*. Chapman & Hall/CRC, 2006.
- [11][SCH94] Schnorr and M. Euchner. *Lattice basis reduction: Improved practical algorithms and solving subset sum problems*. Math. Prog., 66: 181-199, 1994.
- [12][LAG83] J.C. Lagarias and A.M. Odlyzko. *Solving low-*

- density subset sum problems*. In Proc. 24th IEEE Symp. on the Found. of Comp. Sci., pages 1-10, 1983.
- [13][MER78] R. Merkle and M. Hellman. *Hiding information and signatures in trapdoor knapsacks*. IEEE Trans. Inform. Theory, IT-24:525-530, September, 1978.
- [14][HOW97] N.A. Howgrave-Graham. *Finding small solutions of univariate modular equations revisited*. In Cryptography and Coding, volume 1355 of LNCS, pages 131-142. Springer-Verlag, 1997.
- [15][SCH95] C.P. Schnorr and M. Euchner. *Lattice basis reduction: Improved practical algorithms and solving subset sum problems*. Math. Prog., 66:181-199, 1994.
- [16][SCH87] C.P. Schnorr. *A hierarchy of polynomial time lattice basis reduction algorithms*. Theoretical Computer Science, Vol. 53, pp. 201-224, 1987.