Generalization of Boneh- Durfee's Attack for Arbitrary Public Exponent RSA

R. Santosh Kumar
Department of IT
MVGR College of Eng.,
INDIA

C.Narasimam
Department of CSE
V.R.SIDDARTHA ENGG COLLEGE
INDIA

S.Pallam Setty
Department of CS&SE
ANDHRA UNIVERSITY
INDIA

ABSTRACT
In 2000, Boneh-Durfee extended the bound for low private exponent from 0.25 (provided by wiener) to 0.292 with public exponent size is same as modulus size. They have used powerful lattice reduction algorithm (LLL) with coppersmith's theory of polynomials. In this paper we generalize their attack to arbitrary public exponent.

Keywords: Lattices, LLL, RSA, Cryptanalysis, Boneh-Durfee

1. INTRODUCTION
Lattice is a discrete subset of $\mathbb{R}^n$. It has found many applications in various fields like the geometry of numbers, integer relations and Diophantine approximations and notably in cryptology. The main problem in lattices is the lattice reduction deals with finding the good representation of a lattice. For thi good representations many versions exist, but the one given by Lenstra-Lenstra-Lovasz is a famous one, because the polynomial time algorithm exists for this version, called LLL algorithm. The algorithm not only solves good representation for the lattice, It also solves the problem called shortest vector problem (SVP) in some extent. In section II, we state the algorithm and its complexity issues. In section III, we provide some inequalities used in this paper for RSA cryptosystem with balanced primes. In section IV, we state the attack and given the justification, which is a generalization of Boneh-Durfee attack .

2. TERMINOLOGY
2.1 Lattices
A lattice is a discrete subgroup of $\mathbb{R}^n$. Equivalently, given $m \le n$ linearly independent vectors$b_1, b_2, b_3, \dots, b_m \in \mathbb{R}^n$, the set$\mathcal{L} = \mathcal{L}(b_1, b_2, b_3, \cdots, b_m) = \{\sum_{i=1}^{m} \alpha_i b_i \,|\alpha_i \in \mathbb{Z}\}$, is a lattice. The $b_i$are called basis vectors of $\mathcal{L}$ and $\mathcal{B} = \{b_1, b_2, \cdots, b_m\}$ is called a lattice basis for $\mathcal{L}$. Thus, the lattice generated by a basis $\mathcal{B}$ is the set of all integer linear combinations of the basis vectors in $\mathcal{B}$. The determinant of a lattice, denoted by $vol(\mathcal{L})$ is the square root of the gramian determinant$det_{1\le i,j\le m}\langle b_i, b_j\rangle$, which is independent of particular choice of basis. A general treatment of this topic see[1].

2.2 Lattice reduction:
Lattice reduction is a problem of finding the basis of vectors which are short in terms of norm. There are numerous algorithms exists in the literature, but we use LLL algorithm here. Because there is a polynomial time algorithm exists and vectors are near orthogonal and the first vector solves the approximate SVP problem.

2.3 LLL reduced
The following LLL reduced version given by Lenstra, Lenstra, Lovasz[1],[2],[3].

LLL reduced: A basis $b_1, b_2, b_3, \cdots, b_n$ of a lattice is said to be Lovasz-reduced or LLL-reduced if

$|\mu_{i,j}| \le \frac{1}{2}$ for $1 \le j < i \le n$

$|b_i^* + \mu_{i,i-1} b_{i-1}^*|^2 \ge \frac{3}{4}|b_{i-1}^*|^2$ for $1 < i \le n$. where the b_i^* and $\mu_{i,j}$ are defined by the Gram-Schimdt orthogonalization process acting on the b_i. Above in place of ¾ one can replace any quantity $\frac{1}{4} < \delta < 1$.

2.4 LLL Algorithm
The Lenstra –Lenstra -Lov´asz (LLL) algorithm [1][2][3] is an iterative algorithm that transforms a given lattice basis into an LLL-reduced one. Since the definition of LLL-reduced uses Gram-Schmidt process, the LLL algorithm performs the Gram-Schmidt method as subroutine. Here we listed some of the propertied of LLL reduced basis.

Let $b_1, b_2, b_3, \cdots, b_m$ be an LLL reduced basis for a lattice$\mathcal{L} \subset \mathbb{R}^m$. Then

1)$d(\mathcal{L}) \le \prod_{i=1}^{n} |b_i| \le 2^{\frac{n(n-1)}{4}} d(\mathcal{L})$,

2) $|b_j| \le 2^{\frac{i-1}{2}} |b_i^*|$, if $1 \le j \le i \le n$,

3) $|b_1| \le 2^{\frac{n-1}{4}} d(\mathcal{L})^{\frac{1}{n}}$,

4) For every $x \in \mathcal{L}$ with $x \ne 0$ we have$|b_1| \le 2^{\frac{n-1}{2}}|x|$.

2.5 RSA cryptosystem with balanced primes:
RSA cryptosystem [4] is well known cryptosystem and using widely for encryption and signature purposes. In the literature, so many versions exist but in this paper, we consider only balanced primes, which mean that the two RSA primes are roughly the same size. In particular we have $4 < \frac{1}{2}N^{\frac{1}{2}} < p < N^{\frac{1}{2}} < q < 2N^{\frac{1}{2}}$, or equivalently, we assume that $p < q < 2p$. It follows that when the RSA primes are balanced, Euler's totient function $\varphi(N) = (p-1)(q-1)$ satisfies$|N - \varphi(N)| < 3N^{\frac{1}{2}}$.

2.6 Resultant of two bivariate polynomials:
The resultant of two polynomials $f(x,y)$ and $g(x,y)$with respect to the variable$y$, is defined as the determinant of Sylvester matrix of $f(x,y)$ and $g(x,y)$ when considered as polynomials in the single indeterminate y. The resultant is non-zero if and only if the two polynomials are algebraically independent. When the polynomials are algebraically independent, the resultant yields a new polynomial $h(x)$ such that if (x_0, y_0) is a root of both $f(x,y)$ and $g(x,y)$ then $h(x_0) = 0$.

Assumption: We assume that the two polynomials return by LLL algorithm are algebraically independent. There is no theoretical proof for this one, but in practice most of the times achieved.

2.7 Howgrave-Graham Result for Bivariate Integer Polynomials:

Let $h(x,y) \in \mathbb{Z}[x,y]$ be a polynomial in 2 variables with at most w monomials and let m be a positive integer. Suppose in addition that

1) $h(x_0,y_0) \equiv 0 (mod\ e^m)$ where $|x_0| < X$ and $|y_0| < Y$, and
2) $\|h(xX,yY)\| \leq \frac{e^m}{\sqrt{w}}$,

Then $h(x_0,y_0) = 0$ holds over the integers.

Here we state the attack and we follow the ideas of Boneh-Durfee[8].

3. THE GENERALIZED ATTACK OF BONEH AND DURFEE

3.1 Attack:

For every $\epsilon > 0$, there exists an n_0 such that for every $n > n_0$the following holds: Let $N = pq$be an n-bit RSA modulus with balanced primes, let (e,N) be a valid public key and d be its corresponding private exponent defined modulo $\varphi(N)$. Let $e = N^{\alpha}$ and $d = N^{\delta}$. Given the public key, if the private exponent satisfies$\delta < \frac{7}{6} - \frac{1}{3}\sqrt{1+6\alpha} - \epsilon$, then the modulus N can be factored in time polynomial in log (N).

3.2 Justification

Consider the key equation$ed = 1 + k(N-s)$. Modulo e gives the key equation as $kN - ks + 1 \equiv 0 (mod\ e)$, where k,s are unknown. From this, consider the bivariate polynomial $f_e(x,y) \in \mathbb{Z}[x,y]$ given by $f_e(x,y) = Nx + xy + 1$, since $(x_0,y_0) = (k,-s)$ is a root of $f_e(x,y)$ modulo e. Also we have the upper bounds for x_0,y_0 are $X = 2N^{\alpha+\delta-1}$ and $Y = 3N^{1/2}$. Using the above polynomial and the bounds X and Y, we construct a lattice whose every element corresponds to a polynomial with root (x_0,y_0) modulo some power of e. For some fixed integer $m > 0$, define the polynomials $g_{i,k}(x,y) = x^i f_e^k(x,y) e^{m-k}$, $h_{j,k}(x,y) = y^j f_e^k(x,y) e^{m-k}$. Here we will g as x-shift polynomials and h as y shift polynomials. With this construction, notice that $i,j \geq 0$ and $0 \leq k \leq m$, the root (x_0,y_0) of $f_e(x,y)$ modulo e is also a root of $g_{i,k}(x,y)$ and $h_{j,k}(x,y)$ modulo e^m. Now we construct the basis matrix that Boneh and Durfee use, which we will denote by B, consists of coefficient vectors of $\{g_{i,k}(xX,yY) | 0 \leq k \leq m, 0 \leq i \leq m-k)\}, \{h_{j,k}(xX,yY) | 0 \leq k \leq m, 1 \leq j \leq t\}$, for some integer $t > 0$, we will determine later.These are the basis vectors for the lattice L. Next one should arrange these basis vectors such that the matrix is lower triangular, we followed the Boneh and Dufee strategy here. With that ordering of the basis matrix the diagonal elements are given by $X^{i+k}Y^k e^{m-k}$ for x-shift polynomials and $X^k Y^{j+k} e^{m-k}$ for the y-shift polynomials. The example for a basis matrix for m=2 and t=1 is provided in fig1.

Example for $m = 2$ and $t = 1$:

	1	x	Xy	x^2	x^2y	x^2y^2	Y	xy^2	x^2y^3
e^2	e^2								
xe^2		e^2X							
Xe	E	eNX	e X Y						
x^2e^2				e^2X^2					
Xfe		eX		eNX^2	eX^2Y				
f^2	1	2NX	2XY	N^2X^2	2NX^2Y	X^2Y^2			
ye^2							e^2Y		
Yfe			eNXY				eY	eXY^2	
yf^2			2NXY		N^2X^2Y	2NX^2Y^2	Y	2XY^2	X^2Y^3

Fig1. Boneh and durfee lattice for $m = 2$ and $t = 1$

The empty places are filled with zeros. The determinant of B is simply the product of its diagonal elements, the volume of L is

$$vol(L) = \left(\prod_{k=0}^{m}\prod_{i=0}^{m-k} X^{i+k}Y^k e^{m-k}\right)\left(\prod_{k=0}^{m}\prod_{j=1}^{t} X^k Y^{j+k} e^{m-k}\right)$$

$$= (eX)^{m(m+1)(m+2)/3 + \frac{tm(m+1)}{2}} * Y^{m(m+1)(m+2)/6 + tm(m+1)(m+t+1)/2}.$$

Computing an LLL-reduced basis for the lattice L, the two smallest reduced basis vectors corresponding to two linearly independent polynomials $p_1(x,y)$ and $p_2(x,y)$ satisfying $\|p_1(xX,yY)\| \leq \|p_2(x,y)\| \leq 2^{\omega/4} vol(L)^{1/(\omega-1)}$. If both polynomials bounded by $\omega^{-1/2} e^m$, then then one can apply Howgrave-Graham's result. But that can be achieved if $vol(L) < \gamma e^{m(\omega-1)}$, where $\gamma = 2^{-\omega(\omega-1)/4}\omega^{-(\omega-1)/2}$ is a constant. If the polynomials obtained above are algebraically independent, one can compute the resultant of the polynomials $p_1$and p_2. We can get s from the resultant polynomial, in turn we can compute the factorization of N. To derive more general condition, substitute the , Y , e values into the above condition and let $t = \tau m$ for some real $\tau > 0$. After some tricky calculations we get, $N^{\left(\tau^2/4 + \left(\alpha+\delta/2 - 1/4\right)\tau + 2\alpha/3 + \delta/3 - 1/4\right) + O(m^3)} < N^{(\alpha\tau+\alpha/2)m^3 + o(m^3)}$, Where we consider the m^3 terms, the condition on the exponents be simplified as $\frac{1}{4}\tau^2 + \left(\frac{\delta}{2} - \frac{1}{4}\right)\tau + \frac{\alpha}{6} + \frac{\delta}{3} - \frac{1}{4} < 0$. This inequality is minimized when τ is equal to $\frac{1}{2} - \delta$. Substituting the value τ back into the inequality, we get

$-\frac{1}{4}\delta^2+\frac{7}{12}\delta+\frac{\alpha}{6}-\frac{5}{16}<0$. Solving the inequality for δ then yields the new condition$\delta<\frac{7}{6}-\frac{1}{3}\sqrt{1+6\alpha}-\epsilon$, where $\epsilon>0$ has been added for neglecting low order terms. Thus, for sufficiently large N, if the private exponent $d=N^{\delta}$ satisfies the above condition, then two polynomials with root $(x_0,y_0)=(k,-s)$ can be found. If these polynomials are algebraically independent, then $y_0=-s$ can be computed and used to factor the modulus. Since all computations can be done in time polynomial in $\log(N)$, the result follows.

Observe that whenever $\alpha\approx 1$, the above condition reduces that $0.2847-\epsilon$, which is Boneh-Durfee initial result. They have improved this bound by using geometric progressive matrices. We use the same technique here to improve the above condition to get optimal result. The motivation for the improvement is the observation that some basis vectors in B contribute more to volume to the lattice than others. That is, the diagonal elements of some rows are much larger than others. If some of the basis vectors with diagonal elements exceeding e^m re removed from the basis matrix, then the volume of the resulting lattice will be decreased and the bounds on δ in the final condition will increase. But the resulting lattice is no more full rank lattice, Boneh-Durfee used special class of matrices to compute volume of the lattice called geometrical progressive matrices. Now we will state their attack, generalized for arbitrary public exponent.

4. IMPROVING THE BOUND

One can improve the above bound by using geometrical progressive matrices as Boneh did in his paper[8].

4.1 Attack:

For every $\epsilon>0$, there exists an N_0 such that for every $N>N_0$ the following holds: Let $N=pq$ be an RSA modulus with balanced primes, let $e=N^{\alpha}$ be a valid public exponent and let $d=N^{\delta}$ be its corresponding private exponent defined modulo $\varphi(N)$. Given (e,N), if the private exponent satisfies $\delta<\frac{2-\sqrt{2\alpha}}{2}-\epsilon$, then the modulus N can be factored in time polynomial in $\log(N)$.

4.2 Justification:

Here also we construct the basis matrix B for some fixed $m>0$ and $t=(1-2\delta)m$ using the same above bounds. Construct a new basis matrix by removing every basis vector that corresponds to a y-shift polynomial with diagonal element greater than e^m. Let L be a basis generated by the new basis matrix. Now compute LLL-reduced basis for L, we can find again two linearly independent polynomials bounded by $2^{\omega'/4}vol(L)^{1/\omega'-1}$, where ω' is the dimension of the lattice L. If these polynomials satisfy the Howgrave-Graham's bound $\omega^{-1/2}e^m$, then we can get our desired result. From the two upper bounds above, we can derive

$vol(L)<2^{-\omega'(\omega'-1)/4}\omega^{-(\omega'-1)/2}e^{m(\omega'-1)}$. (1)

If this condition is satisfied and if the two polynomials obtained are also algebraically independent then we can efficiently solve for $y_0$and factor the modulus. To derive the general condition, construct the matrix of blocks as $B=\begin{bmatrix}A & 0\\ C & D\end{bmatrix}$, where A is the lower triangular sub matrix of B of order ω_x corresponding to the x-shift polynomials, C is the sub mtrix corresponding to the first ω_x columns of the y shift polynomials, and D is the lower triangular sub matrix of order ω_y, corresponding to last ω_y columns of the y shifts. Reconstruct above matrix by removing the y shift polynomials from the blocks C and D, we get $B'=\begin{bmatrix}A & 0\\ C' & D'\end{bmatrix}$ where C' and D' are obtained from C and D by removing the appropriate rows. Since A is full rank, there exists an unitary matrix U such that $UB'=\begin{bmatrix}S & 0\\ 0 & M\end{bmatrix}$ where S is a diagonal matrix and M is an integer linear combination of only the rows in D'. The determinant of above matrix is$\left|\det(B')\right|=\left|\det(UB')\right|=|\det(S)\det(M)|=|\det(A)\det(D)|$. Notice that already we compute the determinant of A in the previous section and we have $\det(A)=\prod_{k=0}^{m}\prod_{i=0}^{m-k}X^{i+k}Y^{k}e^{m-k}$. Substituting the upper bounds of X,Y and $e=N^{\alpha}$ into the above yields $\det(A)=N^{\left(\frac{2\alpha}{3}+\frac{\delta}{3}-\frac{1}{4}\right)m^3+o(m^3)}$, where we ignored the constant factors of X and Y. In other case, computing $\det(D)$ is non trivial since the matrix D is not a square matrix. In this case Boneh-Durfee used the geometrical progressive matrices to compute the determinant of D. We used the same approach and we can get the upper bound for $\det(D)$ as $N^{\frac{(1-2\delta)(2\delta+6\alpha-1)m^3}{12}+o(m^3)}$, where we ignored all constant factors not depending on N and the dimension of the lattice also we can derived as $(1-\delta)m^2+o(m^2)$. From this, we have $vol(L)\le N^{\left(-\frac{\delta^2}{3}+\left(\frac{2}{3}-\alpha\right)\delta+\frac{7\alpha}{6}-\frac{1}{3}\right)m^3+o(m^3)}$, where we have ignored all the constant factors not depending on N. From this, we can reduce the above condition as $N^{(-\frac{\delta^2}{3}-\alpha\delta+\frac{2\delta}{3}-\frac{1}{3}+\frac{7\alpha}{6})m^3}<N^{\alpha(1-\delta)m^3+o(m^3)}$, where we have ignored all constant factors that do not depend on N. Looking only at the exponents of N and ignoring high order terms, this is simplified to$-2\delta^2+4\delta-2+\alpha<0$. This implies $\delta<\frac{2-\sqrt{2\alpha}}{2}-\epsilon$, where $\epsilon>0$ for ignored constsnts. Since all the computations are done in polynomial time, the modulus N can be factored in time polynomial in $\log(N)$.

In a typical instance of small private exponent RSA the public exponent will be roughly the same size as modulus. Using the approximation $\alpha\approx 1$, we find that a sufficient condition for the attack becomes $\delta<\frac{2-\sqrt{2}}{2}-\epsilon$ which is Boneh-Durfee's original result.

5. EXPERIMENTS

We have done experiment to test our results when $d>N^{0.25}$ and $\alpha\approx 1$. We used the LLL code from the Victor Shoup's NTL package, which is freely available.

The prime numbers which have length 512 bits each:

The first prime number is

10114792273660656874618568712406420344176220457790563178092222929337786916374923318745284718351487926620784106195715878875311958793629905453919697155685507

The second prime number is

10843221374140991753173625949764386011485161421520044246309105053489500519257941272796681417497061734054081478280518835582353321569961722963922828311576983

The encryption exponent is

49446678600051379228760906286031155509742239832659705731559249988210578539211813543612425990507831160407165259046991194935262200565953842567148786053040450198919753834397378188932524599840027093290217612285214105791999673535556558448523448336314401414644879827127064929878383237432895170442176211946286617205

The decryption exponent is

217803521555886180205636419713373442439073919698997648777906738918315273011 37

The number of bits for public exponent is 1023.
The number of bits for private exponent is 254.
After applying LLL algorithm by taking lattice parameters m=3, t=1; LLL returns the reduction matrix, in which we apply the resultant methodology for first two vectors, we achieve the polynomial in one variable. For LLL reduction, for above said parameters takes 70 seconds time under Intel core i5 CPU of 2.53GHz. The lattice basis reduction is done using shoup's NTL[18].

6. CONCLUSION

By using low private exponent for RSA cryptosystem, one can speed up the decryption process or signatures. But one should take care to use low private exponent, because Boneh-Durfee proved that if low private exponent is less than 0.292, then the system is insecure. It is an open problem of the security of RSA cryptosystem, if private exponent is greater than 0.292, whenever the size of the public exponent is same as modulus size.

7. REFERENCES

[1] Cohen, H. 1995. A Course in Computational Algebraic Number Theory. Springer-Verlag. Second edition.

[2] Menezes, A.J, Van Oorschot P.C, and Vanstone. 1997. Hand book of Applied Cryptography. CRC Press.

[3] Lenstra A.K, Lenstra Jr. H.W, Lovasz L. 1982. "Factoring polynomials with rational coefficients". Mathematische A1nnalen, volume 261(4): pages 515-534.

[4] Rivest R.L, Shamir A, Adleman L. 1978. "A method for obtaining digital signatures and public key cryptosystems". Commun.of the ACM, 21: 120-126.

[5] Coppersmith D. 1997. "Small solution to polynomial equations, and low exponent RSA vulnerabilities. Journal of Cryptology", 10(4):233-260.

[6] Howgrave-Graham N. 1997. Finding small roots of univariate modular equations revisited. Proceedings of Cryptography and Coding, Springer-LNCS, vol. 1355, Springer-Verlag, pp.13-142.

[7] Wiener M.J. 1990. Cryptanalysis of short RSA secret exponents. IEEE Trans. In formation Theory, 36(3):553:559.

[8] Boneh.D, and Durfee,G. 2000. Cryptanalysis of RSA with private key d less than N^0.292.IEEE Transactions on information Theory, 46(4):1339-1349.

[9] D.Boneh, G.Durfee,Y.Frankel. 1998 An attack on RSA given fraction of the private key bits. Proceeding of Asiacypt'98. Springer-Verlag, LNCS 1514:25-34.

[10] M. Ernst, E. Jochemsz, A,May, and D.Weger 2005. Partial key exposure attacks on RSA upto full size exponents. Advances in Cryptology –Eurocrypt 2005. Springer-Verlag, LNCS 3494:371-386.

[11] P. Schnorr and M. Euchner. 1994. Lattice basis reduction: Improved practical algorithms and solving subset sum problems Math.Prog. 66: 181- 199.

[12] Santosh kumar R, Narasimham C, Pallam setty S.2012 Lattice based tools for cryptanalysis in various applications. Springer-LNICST, 84:530-537.

[13] Boneh.,D. 1999. Twenty Years of Attacks on the RSA Cryptosystem. Notices the AMS 46(2), 203-213.

[14] Durfee, G, Nguyen, P.Q.2000. Crtptanalysis of the RSA schemes with short exponent from Asiacrypt '99. Proceedings of cryptography-ASIACRYPT, LNCS 1976, Springer-Verlag, pp 1-11.

[15] H.M. Sun, W.C Yang, C.S. Laih.1999. On the design of RSA with short secret exponent. Proceedings of Cryptology –ASIACRYPT'99,LNCS 1716, Springer-Verlag, pp.120-126,1978.

[16] Verhaul, E., van Tilborg.1997. "Cryptanalysis of less short RSA secret exponents". Applicable Algebra of Engineering, Communication and Computing, Vol.8, Springer-Verlag, pp.425-435.

[17] Aono,Y. 2009. Simplification of the lattice based attack of Boneh and Durfee for RSA cryptoanalysis. Proceedings of joint conference of ASCM and MACS.

[18] Victor Shoup. NTL: A library for doing Number Theory, online available at http://shoup.net/ntl.