A Nonuniform Algorithm for the Hidden Number Problem in Subgroups

Igor E. Shparlinski¹ and Arne Winterhof²

 ¹ Department of Computing, Macquarie University Sydney, NSW 2109, Australia igor@ics.mq.edu.au
² Johann Radon Institute for Computational and Applied Mathematics Austrian Academy of Sciences Altenbergerstraße 69, 4040 Linz, Austria arne.winterhof@oeaw.ac.at

Abstract. Boneh and Venkatesan have proposed a polynomial time algorithm in a non-uniform model for recovering a "hidden" element $\alpha \in \mathbb{F}_p$, where p is prime, from very short strings of the most significant bits of the residue of αt modulo p for several randomly chosen $t \in \mathbb{F}_p$. Here we modify the scheme and amplify the uniformity of distribution of the 'multipliers' t and thus extend this result to subgroups of \mathbb{F}_p^* , which are more relevant to practical usage. As in the work of Boneh and Venkatesan, our result can be applied to the bit security of Diffie–Hellman related encryption schemes starting with subgroups of very small size, including all cryptographically interesting subgroups.

Keywords: Hidden number problem, Diffie-Hellman key exchange, Lattice reduction, Exponential sums, Waring problem in finite fields, Nonuniform algorithm

1 Introduction

For a prime p, denote by \mathbb{F}_p the field of p elements and always assume that it is represented by the set $\{0, 1, \ldots, p-1\}$. Accordingly, sometimes, where obvious, we treat elements of \mathbb{F}_p as integer numbers in the above range.

For a real $\eta > 0$ and $t \in \mathbb{F}_p$ we denote by $MSB_{\eta}(t)$ any integer which satisfies the inequality

$$|t - MSB_{\eta}(t)| < p2^{-\eta - 1}.$$
 (1)

Roughly speaking, $MSB_{\eta}(t)$ is an integer having about η most significant bits as t. However, this definition is more flexible and better suited to our purposes. In particular we remark that η in the inequality (1) need not be an integer.

Given a subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$ we consider the following *hidden number problem* over \mathcal{G} :

Recover a number $\alpha \in \mathbb{F}_p$ such that for k elements $t_1, \ldots, t_d \in \mathcal{G}$, chosen independently and uniformly at random from \mathcal{G} , we are given k pairs

$$(t_h, \mathrm{MSB}_\eta(\alpha t_h)), \qquad h = 1, \dots, d,$$

for some $\eta > 0$.

For $\mathcal{G} = \mathbb{F}_p^*$ this problem has been introduced and studied by Boneh and Venkatesan [3, 4]. In [3] a polynomial time algorithm is designed which recovers α for some $\eta \sim (\log p)^{1/2}$ and $k = O(\log^{1/2} p)$. The algorithm of [3] has been extended in several directions. In particular, in [8] it is generalised to all sufficiently large subgroups $\mathcal{G} \subseteq \mathbb{F}_p^*$. This and other generalisations have led to a number of cryptographic applications, see [20–22]. Using bounds of exponential sums from [9, 11] it has been shown that the algorithm of [3] works for subgroups $\mathcal{G} \subseteq \mathbb{F}_p^*$ of order $\#\mathcal{G} \ge p^{\nu+\varepsilon}$ where for any $\varepsilon > 0$ and sufficiently large p one can take

 $-\nu = 1/3$ for all primes,

 $-\nu = 0$ for almost all primes p.

Using a recent improvement of [5] of the bounds of exponential sums over small subgroups of \mathbb{F}_p^* one can obtain the same result with $\nu = 0$ for all primes p and thus extend the results of [3, 8] to subgroups of order $\#\mathcal{G} \ge p^{\varepsilon}$.

For $\mathcal{G} = \mathbb{F}_p^*$ in [4] an algorithm is constructed which works with much smaller values $\eta \sim \log \log p$, however this algorithm is non-uniform. This means that if the points $t_1, \ldots, t_k \in \mathcal{G}$ are known in advance, one can design (in exponential time) a certain data structure, that now given k values $\text{MSB}_{\eta}(\alpha t_i)$, $i = 1, \ldots, k$, the hidden number α can be found in polynomial time. In the present paper we extend the algorithm of [4] to essentially arbitrary subgroups of \mathbb{F}_p^* . As in [4] we discuss possible applications of our algorithm to proving bit security results for several exponentiation based cryptographic schemes.

As in [3,4], the method is based on some properties of lattices, but also makes use of exponential sums, however not in such a direct way as in [8]. Namely, we introduce certain new arguments allowing to amplify the uniformity of distribution properties of small subgroups \mathcal{G} . This allows us to use the bound of exponential sums from [10] with elements of \mathcal{G} , which is very moderate in strength (and does not imply any uniformity of distribution properties of \mathcal{G} which would be the crucial argument of the method of [8]). The bound of [10] has however the very important advantage over the bounds of [5,9,11] that it applies to subgroups of order

$$\#\mathcal{G} \ge \frac{\log p}{(\log \log p)^{1-\varepsilon}}.$$

It is interesting to note that our approach has links with the famous *Waring* problem which has been studied in number theory for several hundred years. In fact, the Waring problem in finite fields has been the main motivation of the bound of exponential sums of [10] which we use in this paper. For surveys of recent results on this problem see [6, 10, 25]. We also remark that a uniform algorithm, which is also based on a similar use of the bound of [10] and which improves the results of [8], has recently been proposed in [23].

Throughout the paper $\log x$ always denotes the binary logarithm of x > 0and the constants in the 'O'-symbols may occasionally, where obvious, depend on a small positive parameter ε and are absolute otherwise. We always assume that p is a prime number with $p \ge 5$, thus the expression $\log \log p$ is defined (and positive).

Acknowledgements: The first author was supported in part by ARC grant DP0211459. The second author was supported in part by DSTA grant R-394-000-011-422, by the Austrian Academy of Sciences, and by FWF grant S8313.

2 Exponential sums and distribution of short sums of elements of subgroups

For a complex z we put $\mathbf{e}_p(z) = \exp(2\pi i z/p)$.

Let $T = \#\mathcal{G}, T|(p-1)$, be the cardinality of a subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$. If we put n = (p-1)/T then each element $r \in \mathcal{G}$ has exactly *n* representations $r = x^n$ with $x \in \mathbb{F}_p^*$. Therefore, for any $\lambda \in \mathbb{F}_p$,

$$\sum_{r \in \mathcal{G}} \mathbf{e}_{p}\left(\lambda r\right) = \frac{T}{p-1} \sum_{x \in \mathrm{IF}_{p}^{*}} \mathbf{e}_{p}\left(\lambda x^{n}\right).$$

Now by Theorem 1 of [10] we have the following bound, see also [6, 11].

Lemma 1. For any $1 > \varepsilon > 0$ there exists a constant $c(\varepsilon) > 0$ such that for any subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$ of order

$$T \geq \frac{\log p}{(\log \log p)^{1-\varepsilon}}$$

the bound

$$\max_{\gcd(\lambda,p)=1} \left| \sum_{r \in \mathcal{G}} \mathbf{e}_p(\lambda r) \right| \le T \left(1 - \frac{c(\varepsilon)}{(\log p)^{1+\varepsilon}} \right)$$

holds.

For an integer $k \geq 1$, a subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$ and $t \in \mathbb{F}_p$ we denote by $N_k(\mathcal{G}, t)$ the number of solutions of the equation

$$r_1 + \ldots + r_k \equiv t \pmod{p}, \qquad r_1, \ldots, r_k \in \mathcal{G}.$$

Recalling the relation between the set of *n*th powers, where n = (p-1)/T, we see that studying the above congruence is equivalent to studying the congruence

$$x_1^n + \ldots + x_k^n \equiv t \pmod{p}, \qquad x_1, \ldots, x_k \in \mathbb{F}_p^*$$

The problem of finding the smallest possible value of k for which the congruence (or in more traditional settings the corresponding equation over \mathbb{Z}) has a solution for any t is known as the Waring problem. However for our purposes just a solvability is not enough. Rather we need an asymptotic formula for the number of solutions.

We show that Lemma 1 can be used to prove that for reasonably small k, $N_k(\mathcal{G}, t)$ is close to its expected value.

Lemma 2. For any $1 > \varepsilon > 0$ there exists a constant $C(\varepsilon) > 0$ such that for any subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$, of order

$$T \geq \frac{\log p}{(\log \log p)^{1-\varepsilon}}$$

 $the \ bound$

$$\max_{t \in \mathbb{F}_p} \left| N_k(\mathcal{G}, t) - \frac{T^k}{p} \right| \le \frac{T^k}{p^2}$$

holds for any integer $k \ge C(\varepsilon)(\log p)^{2+\varepsilon}$.

Proof. The well-known identity (see for example [14, Chapter 5.1])

$$\sum_{\lambda=0}^{p-1} \mathbf{e}_p(\lambda u) = \begin{cases} 0, & \text{if } u \not\equiv 0 \pmod{p}, \\ p, & \text{if } u \equiv 0 \pmod{p}, \end{cases}$$

implies that

$$N_{k}(\mathcal{G}, a) = \sum_{r_{1}, \dots, r_{k} \in \mathcal{G}} \frac{1}{p} \sum_{\lambda=0}^{p-1} \mathbf{e}_{p} \left(\lambda(r_{1} + \dots + r_{k} - t) \right)$$
$$= \frac{1}{p} \sum_{\lambda=0}^{p-1} \mathbf{e}_{p} \left(-\lambda t \right) \left(\sum_{r \in \mathcal{G}} \mathbf{e}_{p} \left(\lambda r \right) \right)^{k}.$$

Separating the term T^k/p , corresponding to $\lambda = 0$, and applying Lemma 1 to other terms, we obtain

$$\max_{t \in \mathbb{F}_p} \left| N_k(\mathcal{G}, t) - \frac{T^k}{p} \right| \le T^k \left(1 - \frac{c(\varepsilon)}{(\log p)^{1+\varepsilon}} \right)^k = T^k \exp\left(O\left(k(\log p)^{-1-\varepsilon}\right)\right)$$

and the desired result follows.

3 Rounding in lattices

Let $\mathcal{B} = (\mathbf{b}_1, \dots, \mathbf{b}_s)^T \in \mathbb{R}^{s \times s}$ be a nonsingular $s \times s$ matrix over the set of real numbers \mathbb{R} with rows $\mathbf{b}_1, \dots, \mathbf{b}_s$. The set of vectors

$$\mathcal{L} = \left\{ \sum_{i=1}^{s} n_i \mathbf{b}_i \mid n_i \in \mathbb{Z} \right\},\$$

spanned by the rows of \mathcal{B} , is called an *s*-dimensional full rank lattice associated with the matrix \mathcal{B} . The set $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ is called a *basis* of \mathcal{L} .

One of the most fundamental problems in this area is the *closest vector* problem. This problem can be defined with respect to any vector norm ||w|| as

follows: given a basis of a lattice \mathcal{L} in \mathbb{R}^s and a target vector $\mathbf{u} \in \mathbb{R}^s$, find a lattice vector $\mathbf{v} \in \mathcal{L}$ with

$$\|\mathbf{u} - \mathbf{v}\| = \operatorname{dist}(\mathbf{u}, \mathcal{L})$$

where

dist
$$(\mathbf{u}, \mathcal{L}) = \min \{ \|\mathbf{u} - \mathbf{z}\| \mid \mathbf{z} \in \mathcal{L} \}$$
.

It is well known that the closest vector problem in the Euclidean norm is **NP**hard (see [16, 17] for references). However, its approximate version [2] admits a polynomial time algorithm which goes back to the lattice basis reduction algorithm of Lenstra, Lenstra and Lovász [12], see also [1] for more recent developments.

However, it has been noticed in [4] that for some special class of lattices a simple rounding technique gives an exact solution to the closest vector problem. Here we summarise several results from [4] which underlie this technique and its applications to the hidden number problem.

For our purposes the L_1 -norm is most relevant thus from now on we always assume that $\|\mathbf{w}\| = \sum_{i=1}^{s} |w_i|$ is the L_1 -norm of $\mathbf{w} = (w_1, \ldots, w_s) \in \mathbb{R}^s$, in particular dist $(\mathbf{u}, \mathcal{L})$ is always assumed to be defined with respect to the L_1 norm.

Given a target vector $\mathbf{u} \in \mathbb{R}^s$, using standard linear algebra tools, we find its representation in the basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$

$$\mathbf{u} = \sum_{i=1}^{s} w_i \mathbf{b}_i$$

and then put

$$\lfloor \mathbf{u} \rfloor = \sum_{i=1}^{s} \lfloor w_i \rfloor \mathbf{b}_i$$

where for $w \in \mathbb{R}$, $\lfloor w \rfloor$ denotes the closest integer (in the case of $2w \in \mathbb{Z}$ we put $\lfloor w \rceil = \lfloor w \rfloor$). Clearly, $\lfloor u \rceil \in \mathcal{L}$ but certainly it is not the closest (or even just a close) vector.

Now, for a matrix $\mathcal{C} \in \mathbb{R}^{s \times s}$ with columns $\mathbf{c}_1^T, \ldots, \mathbf{c}_s^T$, we introduce the following measure

$$\boldsymbol{\omega}(\mathcal{C}) = \max_{1 \le j \le s} \|\mathbf{c}_j\|.$$

The following statement, which is essentially [4, Lemma 2.1], gives a sufficient condition under which $|\mathbf{u}|$ is a solution to the closest vector problem for \mathbf{u} .

Lemma 3. If

$$\rho\left(\mathcal{B}^{-1}\right) < \frac{1}{2\operatorname{dist}\left(\mathbf{u},\mathcal{L}\right)}$$

then

$$\|\mathbf{u} - [\mathbf{u}]\| = \operatorname{dist}(\mathbf{u}, \mathcal{L}).$$

We consider the lattice $\mathcal{L}(t_1, \ldots, t_d)$ spanned by the rows of the matrix

$$\mathcal{B}(t_1, \dots, t_d) = \begin{pmatrix} p & 0 & \dots & 0 & 0 \\ 0 & p & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \dots & p & 0 \\ t_1 & t_2 & \dots & t_d & 1 \end{pmatrix}.$$

The next statement follows from [4, Theorem 2.2].

Lemma 4. Let p be a prime and $d > 4 + \log p + \log \log p$. Let $t_1, \ldots, t_d \in \{0, 1, \ldots, p-1\}$ be integers chosen uniformly and independently at random. Then with probability at least 1/2 there exists a basis of the lattice $\mathcal{L}(t_1, \ldots, t_d)$ spanned by rows of a certain matrix \mathcal{C} with entries of polynomial size $(\log p)^{O(1)}$ and with

$$\rho(\mathcal{C}^{-1}) < \frac{3d\log p}{p}$$

4 Nonuniform algorithm

For an integer w we denote by $\lfloor w \rfloor_p$ the remainder of w on division by p.

Assume that for $\alpha \in \mathbb{F}_p^*$ and a subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$ of order T, generated by $g \in \mathbb{F}_p^*$, we are given an *oracle* \mathcal{HNP}_{μ} such that for every $x \in \{0, 1, \ldots, T-1\}$, it returns $\text{MSB}_{\mu}(\lfloor \alpha g^x \rfloor_p)$.

Theorem 1. For any $1 > \varepsilon > 0$ there exists a constant $a(\varepsilon) > 0$ such that, for $\mu = a(\varepsilon) \log \log p$, for any $g \in \mathbb{F}_p^*$ of order

$$T \ge \frac{\log p}{(\log \log p)^{1-\varepsilon}}$$

after taking a polynomial number $(\log p)^{O(1)}$ of advice bits depending only on p and \mathcal{G} but independent on α , one can design a deterministic algorithm which makes $O\left((\log p)^{3+\varepsilon}\right)$ calls of the oracle \mathcal{HNP}_{μ} and then recovers α in polynomial time.

Proof. Put

$$d = 5 + \lfloor \log p + \log \log p \rfloor, \qquad k = \lceil C(\varepsilon) (\log p)^{2+\varepsilon} \rceil,$$

where $C(\varepsilon)$ is given by Lemma 2.

The advice bits which we request describe:

- the values of $t_1, \ldots, t_d \in \mathbb{F}_p$ for which the lattice $\mathcal{L}(t_1, \ldots, t_d)$ is spanned by a matrix \mathcal{C} with

$$\rho(\mathcal{C}^{-1}) < \frac{3d\log p}{p}$$

which exist by Lemma 4, and the above matrix C;

- the exponents x_{hj} , $h = 1, \ldots, d$, $j = 1, \ldots, k$ with

$$t_h \equiv \sum_{j=1}^k g^{x_{hj}} \pmod{p}, \qquad h = 1, \dots, d,$$

which exist by Lemma 2.

We call the oracle with x = 0 getting an approximation $u_0 = \text{MSB}_{\mu}(\alpha)$. Now we call the oracle \mathcal{HNP}_{μ} for the dk integers

$$r_{hj} = g^{x_{hj}} \in \mathcal{G}, \qquad j = 1, \dots, k, \ h = 1, \dots, d,$$

and get integers u_{hj} with

$$|\lfloor \alpha r_{hj} \rfloor_p - u_{hj}| < p/2^{\mu+1}, \qquad h = 1, \dots, d, \ j = 1, \dots, k.$$

For $h = 1, 2, \ldots, d$ we put

$$v_h = \sum_{j=1}^k \lfloor \alpha r_{hj} \rfloor_p, \qquad t_h = \left\lfloor \sum_{j=1}^k r_{hj} \right\rfloor_p, \qquad u_h = \sum_{j=1}^k u_{hj},$$

where all the additions are over \mathbb{Z} .

Note that for sufficiently large p,

$$|v_h - u_h| < kp/2^{\mu+1} \le p/2^{\eta+1},$$

where

$$\eta = \mu - \log k \ge \log \left(3d(d+1)\log p \right).$$

for an appropriate value of $a(\varepsilon)$ and sufficiently large p. Letting $\mathbf{u} = (u_1, \ldots, u_d, u_0)$, we obtain

dist
$$(\mathbf{u}, \mathcal{L}(t_1, ..., t_d)) \le (d+1)p/2^{\eta+1}$$
.

Therefore,

$$\rho(\mathcal{C}^{-1}) < \frac{3d\log p}{p} \le \frac{2^{\eta}}{(d+1)p} \le \frac{1}{2\operatorname{dist}\left(\mathbf{u}, \mathcal{L}(t_1, \dots, t_d)\right)}$$

and the result follows by Lemma 3.

5 Application to Diffie-Hellman related schemes

Our result applies to the establishing bit security of the same exponentiation based cryptographic schemes as those of [4]. Such schemes include, but are not limited to, the Okamoto conference sharing scheme and a certain modification of the ElGamal scheme, see [4] for more details.

The main distinction between our result and that of [4] is that we do not need anymore assume that the generating element is a primitive root, which is a rather impractical assumption. Indeed, in practical applications of the Diffie-Hellman and other related schemes, one would probably choose a subgroup of \mathbb{F}_p^* of prime order T. Moreover, it is quite reasonable to choose T of size about $\exp(c(\log p)^{1/3}(\log \log p)^{2/3})$ for some constant c > 0, in order to balance time complexities of the number field sieve based attacks and Pollard's rho-method based attacks, see [7, 15, 18, 19, 24]. Thus our result closes the gap between the settings of [4] and settings more relevant to practical usage of the above schemes.

It also seems to be plausible that one can obtain similar, albeit slightly weaker, results for other cryptographically interesting subgroups in finite fields and rings, for which relevant bounds of character sums are available. For example, such bounds are known for XTR subgroups, see [13].

References

- M. Ajtai, R. Kumar, and D. Sivakumar, 'A sieve algorithm for the shortest vector problem', Proc. 33rd ACM Symp. on Theory of Comput., Crete, Greece, 2001, 601–610.
- L. Babai, 'On Lovasz' lattice reduction and the nearest lattice point problem', Combinatorica, 6 (1986), 1–13.
- D. Boneh and R. Venkatesan, 'Hardness of computing the most significant bits of secret keys in Diffie-Hellman and related schemes', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1109** (1996), 129–142.
- D. Boneh and R. Venkatesan, 'Rounding in lattices and its cryptographic applications', Proc. 8th Annual ACM-SIAM Symp. on Discr. Algorithms, SIAM, 1997, 675–681.
- J. Bourgain and S. V. Konyagin, 'Estimates for the number of sums and products and for exponential sums over subgroups in fields of prime order', *Comptes Rendus Mathematique*, **337** (2003), 75–80.
- T. Cochrane, C. Pinner, and J. Rosenhouse, 'Bounds on exponential sums and the polynomial Waring's problem mod p', Proc. Lond. Math. Soc., 67 (2003), 319–336.
- R. Crandall and C. Pomerance, Prime numbers: A Computational perspective, Springer-Verlag, Berlin, 2001.
- M. I. González Vasco and I. E. Shparlinski, 'On the security of Diffie-Hellman bits', Proc. Workshop on Cryptography and Computational Number Theory, Singapore 1999, Birkhäuser, 2001, 257–268.
- D. R. Heath-Brown and S. V. Konyagin, 'New bounds for Gauss sums derived from kth powers, and for Heilbronn's exponential sum', *Quart. J. Math.*, **51** (2000), 221– 235.
- S. V. Konyagin, 'On estimates of Gaussian sums and the Waring problem modulo a prime', *Trudy Matem. Inst. Acad. Nauk USSR*, Moscow, **198** (1992), 111–124 (in Russian); translation in *Proc. Steklov Inst. Math.*, **1** (1994), 105–117.
- 11. S. V. Konyagin and I. Shparlinski, *Character sums with exponential functions and their applications*, Cambridge Univ. Press, Cambridge, (1999).
- A. K. Lenstra, H. W. Lenstra, and L. Lovász, 'Factoring polynomials with rational coefficients', *Mathematische Annalen*, 261 (1982), 515–534.

- W.-C. W. Li, M. Näslund, and I. E. Shparlinski, 'The hidden number problem with the trace and bit security of XTR and LUC', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2442** (2002), 433–448.
- R. Lidl and H. Niederreiter, *Finite fields*, Cambridge University Press, Cambridge, (1997).
- A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, Handbook of applied cryptography, CRC Press, Boca Raton, FL, 1996.
- P. Q. Nguyen and J. Stern, 'Lattice reduction in cryptology: An update', *Lect. Notes Comp. Sci.*, Springer-Verlag, Berlin, **1838** (2000), 85–112.
- P. Q. Nguyen and J. Stern, 'The two faces of lattices in cryptology', *Lect. Notes Comp. Sci.*, Springer-Verlag, Berlin, **2146** (2001), 146–180.
- O. Schirokauer, 'Discrete logarithms and local units', Philos. Trans. Roy. Soc. London, Ser. A, 345 (1993), 409–423.
- O. Schirokauer, D. Weber, and T. Denny, 'Discrete logarithms: The effectiveness of the index calculus method', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, 1122 (1996), 337–362.
- I. E. Shparlinski, 'Playing "Hide-and-Seek" in finite fields: Hidden number problem and its applications', Proc. 7th Spanish Meeting on Cryptology and Information Security, Vol.1, Univ. of Oviedo, 2002, 49–72.
- I. E. Shparlinski, 'Exponential sums and lattice reduction: Applications to cryptography', *Finite Fields with Applications to Coding Theory, Cryptography and Related Areas*, Springer-Verlag, Berlin, 2002, 286–298.
- 22. I. E. Shparlinski, Cryptographic applications of analytic number theory, Birkhäuser, 2003.
- 23. I. E. Shparlinski and A. Winterhof, 'Hidden number problem in small subgroups', *Preprint*, 2003.
- 24. D. R. Stinson, Cryptography: Theory and practice, CRC Press, Boca Raton, FL, 2002.
- A. Winterhof, 'On Waring's problem in finite fields', Acta Arith., 87 (1998), 171– 177.