# Aitken and Neville Inverse Interpolation Methods over Finite Fields 

E.C. Laskari ${ }^{* 1,3}$, G.C. Meletiou ${ }^{* * 2,3}$, and M.N. Vrahatis ${ }^{* * * 1,3}$<br>${ }^{1}$ Computational Intelligence Laboratory, Dept. of Mathematics, University of Patras, GR-26110 Patras, Greece<br>${ }^{2}$ A.T.E.I. of Epirus, Arta, Greece, P.O. Box 110, GR-47100 Arta, Greece<br>${ }^{3}$ University of Patras Artificial Intelligence Research Center (UPAIRC), University of Patras, GR-26110 Patras, Greece

Received 31 October 2004, revised 21 March 2005, accepted 21 March 2005
Published online 22 April 2005

In this contribution the application of two inverse interpolation methods over finite fields is studied. More specifically, we consider the Aitken and Neville inverse interpolation methods for a "shifted" discrete exponential function. The results indicate that the computational cost of finding the discrete logarithm through this approach remains high, however interesting features regarding the degree of the resulting interpolation polynomials are reported.
(c) 2005 WILEY-VCH Verlag GmbH \& Co. KGaA, Weinheim

## 1 Introduction

Public key cryptography is intimately related to a number of hard and complex mathematical problems from the fields of computational algebra, number theory, probability theory, mathematical logic, Diophantine's complexity and algebraic geometry. Such problems are the factorization [1], the discrete logarithm [2, 3, 4] and others [5]. Cryptosystems rely on the assumption that these problems are computationally intractable, in the sense that their computation cannot be completed in polynomial time.

The Discrete Logarithm Problem (DLP) [2, 3, 4] amounts to the development of an efficient algorithm for the computation of an integer $x$ that satisfies the relation:

$$
\alpha^{x}=\beta,
$$

where $\alpha$ is a fixed primitive element of a finite field $\mathbb{F}_{q}$ (i.e., $\alpha$ is a generator of the multiplicative group $\mathbb{F}_{q}^{*}$ of $\mathbb{F}_{q}$ ) and $\beta$ is a non-zero element of the field. We assume that $x$ is the smallest nonnegative integer with $\alpha^{x}=\beta$. Then, $x$ is called the index or the discrete logarithm of $\beta$. In the special case of a finite field $\mathbb{Z}_{p}$ of prime order $p$, a primitive root $g$ modulo $p$ is selected. If $u$ is the smallest nonnegative integer with $g^{u} \equiv h(\bmod p)$, then $u$ is called the index or the discrete logarithm of $h$.

The security of various public and private key cryptosystems [ $2,3,4,6,7,8,9,10,11,12,13$ ] relies on the assumption that DLP is computationally intractable. Specifically, we refer to:
(1) the Diffie-Hellman exchange protocol [14],
(2) the El Gamal public key cryptosystem as well as the El Gamal digital signature scheme [13].

The Diffie-Hellman key Problem (DHP) $[6,10,15]$ is defined as follows. Let $\alpha$ be a fixed primitive element of a finite field $\mathbb{F}_{q} ; x, y$, satisfying, $0 \leqslant x, y \leqslant q-2$, denote the private keys of two users; and $\beta=\alpha^{x}, \gamma=\alpha^{y}$ represent the corresponding public keys. Then, the problem amounts to computing $\alpha^{x y}$ from $\beta$ and $\gamma$, where $\alpha^{x y}$ is the symmetric key for secret communication between the two users.

[^0]Consider the special case of the DHP, where $\beta=\gamma$. The term Diffie-Hellman Mapping refers to the mapping:

$$
\beta=\alpha^{x} \longmapsto \alpha^{x^{2}} .
$$

The definition of the Diffie-Hellman Mapping Problem (DHMP) follows naturally from the aforementioned definition of DHP. The two problems, DHMP and DHP, are computationally equivalent, since the following relation holds:

$$
\alpha^{x^{2}} \alpha^{y^{2}} \alpha^{2 x y}=\alpha^{(x+y)^{2}},
$$

and the computation of $\alpha^{x y}$ from $\alpha^{2 x y}$ is feasible (square roots over finite fields).
The factorization problem is related to the RSA cryptosystem [1]. The security of this cryptosystem relies on the computational intractability of the factorization of a positive integer $N=p q$, where $p$ and $q$ are two distinct odd primes [16]. The factorization of $N$ is equivalent to determining $\phi(N)$ from $N$, where $\phi(N)=$ $(p-1)(q-1)[1]$.

Numerous techniques, including algebraic, number theoretic, soft computing and interpolation methods, have been proposed to tackle the aforementioned problems $[2,3,4,6,7,8,9,10,11,12,15,17,18,19]$. We focus on interpolation techniques over finite fields. In particular, we consider the Aitken and the Neville inverse polynomial interpolation methods applied on a "shifted" discrete exponential function.

The rest of the paper is organized as follows. In Section 2, we briefly present previous work on interpolation methods over finite fields and describe the Aitken and Neville interpolation methods. In Section 3, the proposed approach for inverse interpolation over finite fields is described and results for a "shifted" exponential function are reported. The paper closes with a synopsis in Section 4.

## 2 Background Material

This section is devoted to a brief description of interpolation methods over finite fields and also presents the Aitken and Neville interpolation methods.

### 2.1 Interpolation Methods over Finite Fields

In a finite field $\mathbb{F}_{q}$, where $q=p^{n}$, with $p$ prime and $n$ a positive integer, every function can be represented as a polynomial through Lagrangian interpolation. For every function, $f: \mathbb{F}_{q} \rightarrow \mathbb{F}_{q}$, there exists a unique polynomial $p(x)$ of degree at most $(q-1)$ that coincides with $f$. Interpolation is computationally attractive only in the case of a polynomial with small number of non-zero coefficients (low sparsity). Since encryption and decryption functions are defined as functions over finite fields, it is natural to attempt to express them as polynomials.

Regarding the discrete logarithm function, the well-known formula due to Wells [20] exists:

$$
\log _{a}(x)=\sum_{i=1}^{p-2} \frac{x^{i}}{1-a^{i}},
$$

where $x \neq 0, a, x \in \mathbb{Z}_{p}$, and $a$ is a generator of $\mathbb{Z}_{p}^{*}$. This formula can be generalized for the case of a field $\mathbb{F}_{q}$ of prime power order and moreover, for the case of $a$ not being a generator of the multiplicative group of the field [7, 8, 11]. A discrete Fourier transformation can also be used [7]:

$$
\log _{a}(x)=(1,2, \ldots, p-1) \underset{1 \leqslant i, j \leqslant p-1}{\left(a^{-i j}\right)}\left(\begin{array}{c}
x \\
x^{2} \\
\vdots \\
x^{p-1}
\end{array}\right) .
$$

For the Diffie-Hellman mapping, the following formula has been proposed [10]:

$$
K(x, y)=-\sum_{1 \leqslant i, j \leqslant p-1} x^{i} y^{j} a^{-i j}
$$

This expression can be represented through a discrete Fourier transformation as follows:

$$
K(x, y)=\left(y, y^{2}, \ldots, y^{p-1}\right) \underset{1 \leqslant i, j \leqslant p-1}{\left(-a^{-i j}\right)}\left(\begin{array}{c}
x \\
x^{2} \\
\vdots \\
x^{p-1}
\end{array}\right)
$$

The Diffie-Hellman key exchange is based on the fact that there is no simple representation of the Diffie-Hellman mapping $F\left(g^{u}, g^{v}\right)=g^{u v}$, with $0<u, v<d$. It can be easily verified that the polynomial [10]:

$$
F(x, y)=e \sum_{i, j=0}^{d-1} g^{-i j} x^{i} y^{j}, \quad e d \equiv 1(\bmod p)
$$

represents the Diffie-Hellman key function in the field $\mathbb{F}_{q}$, where $q=p^{n}$.
The polynomial $F$ can be of degree at most $2(d-1)$, while the largest number of non-zero coefficients of $F$ is $d^{2}$. Recently, lower bounds for the degree of a two-variable polynomial have been identified [10].

### 2.2 The Aitken and Neville Interpolation Methods

A Lagrangian type interpolation achieves the desirable accuracy using the smallest possible number of interpolation points. However, the proper number of interpolation points is not a priori known. Typically, a number of interpolation points are initially considered and, if they are not sufficient, new points are added iteratively until the desired accuracy is achieved. The addition of a new interpolation point in the case of Lagrangian interpolation cannot be performed directly, because the interpolation polynomial changes and, thus, all computations need to be performed from the beginning.

The Aitken and Neville interpolation methods are well-known and they are considered as the state-of-the-art for the interpolation of functions over real numbers [21]. Furthermore, Aitken and Neville interpolation methods are constructive in a way that permits the addition of a new interpolation point with low computational cost. This advantage of the aforementioned methods over the Lagrangian interpolation method and the fact that their interpolation formulae can be applied in any field, has motivated our investigation of their performance over finite fields.

Let $f(x)$ be a function defined on a field $\mathbb{F}, x_{i} \in \mathbb{F}, i=0, \ldots, n$, be mutually different interpolation points, and $f_{i}=f\left(x_{i}\right)$. Then, the Aitken polynomial is defined as:

$$
P_{0,1, \ldots, m, i}(x)=\frac{1}{\left(x_{i}-x_{m}\right)}\left|\begin{array}{ll}
P_{0,1, \ldots, m}(x) & x_{m}-x \\
P_{0,1, \ldots, m-1, i}(x) & x_{i}-x
\end{array}\right|, \quad \text { for }\left\{\begin{array}{l}
m=0, \ldots, n-1 \\
i=(m+1), \ldots, n
\end{array}\right.
$$

where, in general, $P_{0,1, \ldots, k}$ denotes the Aitken polynomial that interpolates all points $x_{0}, x_{1}, \ldots, x_{k}$.
The corresponding Neville interpolation formula is:

$$
P_{i, i+1, \ldots, i+m}(x)=\frac{1}{\left(x_{i+m}-x_{i}\right)}\left|\begin{array}{ll}
P_{i, i+1, \ldots, i+m-1}(x) & x_{i}-x \\
P_{i+1, i+2, \ldots, i+m}(x) & x_{i+m}-x
\end{array}\right|, \quad \text { for }\left\{\begin{array}{l}
m=1, \ldots, n \\
i=0, \ldots,(n-m)
\end{array}\right.
$$

where, in general, $P_{i, i+1, \ldots, i+k}$ denotes the Neville polynomial that interpolates all points $x_{i}, x_{i+1}, \ldots, x_{i+k}$.
In several applications, the values $f_{i}, i=0, \ldots, n$, of a function $y=f(x)$ at the corresponding points $x_{i}, i=$ $0, \ldots, n$, are given, and the point $x^{*}$, such that $f\left(x^{*}\right)=y^{*}$, is required. This is the inverse interpolation problem for the function $f$. Both Aitken and Neville interpolation methods can be applied for the inverse interpolation problem by simply exchanging $x_{i}$ and $y_{i}=f_{i}$ in the corresponding formulae [21]. Thus, the formula of the inverse Aitken interpolation method is:

$$
P_{0,1, \ldots, m, i}(y)=\frac{1}{\left(y_{i}-y_{m}\right)}\left|\begin{array}{ll}
P_{0,1, \ldots, m}(y) & y_{m}-y \\
P_{0,1, \ldots, m-1, i}(y) & y_{i}-y
\end{array}\right|, \quad \text { for }\left\{\begin{array}{l}
m=0, \ldots, n-1 \\
i=(m+1), \ldots, n
\end{array}\right.
$$

while the corresponding inverse Neville interpolation formula is:

$$
P_{i, i+1, \ldots, i+m}(y)=\frac{1}{\left(y_{i+m}-y_{i}\right)}\left|\begin{array}{l}
P_{i, i+1, \ldots, i+m-1}(y) \\
P_{i+1, i+2, \ldots, i+m}(y) \\
y_{i+m}-y
\end{array}\right|, \quad \text { for }\left\{\begin{array}{l}
m=1, \ldots, n \\
i=0, \ldots,(n-m)
\end{array}\right.
$$

Table 1 Instances of test problems used.

| \# Problem | $p$ | $\alpha$ | $b$ |
| :---: | ---: | ---: | ---: |
| 1 | 101 | 2 | 30 |
| 2 | 599 | 7 | 200 |
| 3 | 1759 | 6 | 500 |
| 4 | 2003 | 5 | 100 |
| 5 | 2411 | 6 | 700 |
| 6 | 2801 | 3 | 300 |
| 7 | 3001 | 14 | 100 |
| 8 | 3313 | 10 | 400 |
| 9 | 3631 | 15 | 3600 |
| 10 | 4001 | 3 | 1000 |
| 11 | 4441 | 21 | 500 |
| 12 | 4751 | 19 | 4500 |
| 13 | 5003 | 2 | 1000 |
| 14 | 5209 | 17 | 300 |
| 15 | 5881 | 31 | 3000 |
| 16 | 6007 | 3 | 2000 |
| 17 | 6841 | 22 | 4000 |
| 18 | 7001 | 3 | 500 |
| 19 | 7841 | 12 | 3500 |
| 20 | 8011 | 14 | 40 |
| 21 | 8761 | 23 | 6000 |
| 22 | 8929 | 11 | 3000 |
| 23 | 9001 | 7 | 200 |
| 24 | 10007 | 5 | 1000 |

## 3 Inverse Interpolation over Finite Fields

We study the inverse Aitken and the inverse Neville interpolation methods over finite fields. In particular, we investigate the inverse interpolation approach on the values of the "shifted" discrete exponential function:

$$
f(x)=\alpha^{x}-b \quad(\bmod p),
$$

over $\mathbb{Z}_{p}$ with $p$ a prime number and $\alpha$ a primitive element of $\mathbb{Z}_{p}$. This function is a bijection, since $\alpha$ is a primitive element.

Since the considered methods are constructive, not all the elements of the range of the function are used as interpolation points, but rather, as many points as necessary to construct a polynomial that interpolates the function value $f\left(x^{*}\right)=0(\bmod p)$. The procedure begins by interpolating two function values of the function $f(x)$ for two random values of $x$. The resulting polynomial is evaluated at zero and, unless the obtained value is the discrete logarithm of $b$ over $\alpha$ modulo $p$, its function value becomes a new interpolation point.

The capability and the characteristics of the considered inverse interpolation methods over finite fields of prime order were tested for different primes, $p$, and values of $b$. The instances of the test problems are reported in Table 1. The results over 100 independent experiments for these problems are reported in Table 2. This table reports the number of verifications (\# Verifications), the number of points used (\# Points Used), and the degree of the resulting polynomials (Pol. Degree). Verifications refers to the number of times that a polynomial evaluation at zero is performed, to verify whether the outcome equals the value of the discrete logarithm. Each verification requires one evaluation of the "shifted" exponential function, $f$, for a specific value of $x$. Points Used refers to the number of interpolation points, $y=f(x)$, employed by the procedure (and, thus, evaluated) until the proper interpolation polynomial is found. Note that not all of these points are necessarily interpolated by the final polynomial. Finally, the third quantity (Pol. Degree) corresponds to the degree of the resulting polynomial that interpolates the value of the discrete logarithm. The number of interpolation points that are used to construct this polynomial is (Pol. Degree +1 ). For each experiment, the two initial interpolation points are chosen randomly
in the range $[1, p-1]$. For each of the three quantities, the mean value (Mean), the median (Med), the standard deviation (StD), as well as, the minimum (Min) and maximum (Max) values are reported.

The results indicate that the computational cost, with respect to the required verifications, for tackling the discrete logarithm problem is high for both methods. However, the relatively low degree of the resulting polynomials is an interesting finding as it indicates that in most cases a polynomial with low degree that interpolates the discrete logarithm exists. Therefore, finding the proper interpolation points is critical for reducing the computational cost.

Another attractive finding is the high frequency with which, low degree polynomials appear in the conducted experiments for each instance of the problem. This conclusion is derived by the low median values of the degree of the obtained polynomial reported in Table 2, in contrast to the corresponding maximum values that are rarely observed. The frequency of appearance of different degrees of polynomials over all tested problems for both methods is depicted in Figure 1. Regarding the ability of each method to find a proper polynomial, Aitken's method was slightly better than the Neville's method in most cases.

The performance of the methods was also investigated for several instances of $b$, while keeping the value of $p$ constant. Table 3 reports results for this setting. Overall, the behavior of both methods is quite stable with respect to this parameter.

## 4 Synopsis

The performance of two inverse interpolation methods, namely the Aitken and Neville interpolation methods, was studied on a "shifted" discrete exponential function over finite fields. The results indicate that the computational cost for tackling the problem of the discrete logarithm through both methods is high. Overall, Aitken's method proved slightly better than the Neville's method. It is important to observe that the resulting polynomials were most often of low degree. This finding implies that in most cases there exists a low degree polynomial that interpolates the discrete logarithm. Therefore, finding the proper interpolation points is crucial to reduce the computational cost of this approach and it will be the subject of future work.

Acknowledgements We acknowledge the partial support by the "Archimedes" research programme awarded by the Greek Ministry of Education and Religious Affairs and the European Union.

## References

[1] R. Rivest, A. Shamir, L. Adlemann, A method for obtaining digital signatures and public-key cryptosystems, Commun. ACM, 21, 120-126, (1978).
[2] L. Adleman, A subexponential algorithm for discrete logarithm problem with applications to cryptography, 20th FOCS, 55-60, (1979).
[3] A. M. Odlyzko, Discrete logarithms: the past and the future, Des. Codes Cryptog., 19, 129-145, (2000).
[4] S. C. Pohlig, M. Hellman, An improved algorithm for computing logarithms over $\operatorname{GF}(p)$ and its cryptographic significance, IEEE Trans. Inf. Theory, 24, 106-110, (1978).
[5] R. C. Merkle, M. E. Hellman, Hiding information and signatures in trapdoor knapsacks, IEEE Trans. Inf. Theory, 24, 525-530, (1978).
[6] D. Coppersmith, I. Shparlinski, On polynomial approximation of the discrete logarithm and the Diffie-Hellman mapping, J. Cryptology, 13, 339-360, (2000).
[7] G. C. Meletiou, G. L. Mullen, A note on discrete logarithms in finite fields, Appl. Algebra Engrg. Comm. Comput., 3, No. 1, 75-79, 1992.
[8] G. L. Mullen, D. White, A polynomial representation for logarithms in GF (q), Acta Arith., 47, 255-261, (1986).
[9] H. Niederreiter, A short proof for explicit formulas for discrete logarithms in finite fields, Appl. Algebra Engrg. Comm. Comput., 1, 55-57, (1990).
[10] A. Winterhof, A note on the interpolation of the Diffie-Hellman mapping, Bull. Australian Math. Soc., 64, No. 3, 475477, (2001).
[11] G. C. Meletiou, Explicit form for the discrete logarithm over the field $\operatorname{GF}(p, k)$, Arch. Math. (Brno), 29, No. 1-2, 25-28, (1993).
[12] A. Winterhof, Polynomial interpolation of the discrete logarithm, Des. Codes Cryptogr., 25, No. 1, 63-72, (2002).
[13] T. El Gamal, A public key cryptosystem and a signature scheme based on discrete logarithms, IEEE Trans. Inf. Theory, 31, 469-472, (1985).

Table 2 Results of the inverse Aitken and inverse Neville methods over 100 independent experiments.



Fig. 1 Frequency of appearance of polynomial degrees over all problems with (a) Aitken method and (b) Neville method.

Table 3 Results over 100 independent experiments for $p=2003, \alpha=5$ and several instances of $b$.

| Problem | measure (\%) | \# Points Used | Pol.Degree |
| :---: | :---: | :---: | :---: |
|  | Mean | 61.07 | 29.25 |
| $p=2003$ | Median | 60.00 | 25.00 |
| $a=5$ | Min | 5.00 | 1.00 |
| $b=9$ | Max | 142.00 | 112.00 |
|  | Mean | 58.86 | 26.87 |
| $p=2003$ | Median | 57.00 | 19.00 |
| $a=5$ | Min | 13.00 | 1.00 |
| $b=100$ | Max | 180.00 | 140.00 |
|  | Mean | 57.41 | 29.52 |
| $p=2003$ | Median | 51.00 | 21.00 |
| $a=5$ | Min | 10.00 | 1.00 |
| $b=600$ | Max | 153.00 | 108.00 |
|  | Mean | 54.07 | 27.46 |
| $p=2003$ | Median | 51.00 | 21.00 |
| $a=5$ | Min | 8.00 | 2.00 |
| $b=900$ | Max | 110.00 | 78.00 |
|  | Mean | 55.89 | 28.86 |
| $p=2003$ | Median | 49.00 | 20.00 |
| $a=5$ | Min | 8.00 | 1.00 |
| $b=1300$ | Max | 173.00 | 133.00 |
|  | Mean | 55.48 | 28.09 |
| $p=2003$ | Median | 50.00 | 23.00 |
| $a=5$ | Min | 5.00 | 2.00 |
| $b=1550$ | Max | 145.00 | 123.00 |
|  | Mean | 55.11 | 27.88 |
| $p=2003$ | Median | 51.00 | 23.00 |
| $a=5$ | Min | 9.00 | 2.00 |
| $b=1900$ | Max | 162.00 | 121.00 |
|  | Mean | 52.23 | 23.82 |
| $p=2003$ | Median | 48.00 | 20.00 |
| $a=5$ | Min | 3.00 | 1.00 |
| $b=2000$ | Max | 138.00 | 100.00 |

[14] W. Diffie, M. Hellman, New directions in cryptography, IEEE Trans. Inf. Theory, 22, 644-654, (1976).
[15] U. Maurer, S. Wolf, The relationship between breaking the Diffie-Hellman protocol and computing discrete logarithms, SIAM J. Computing, 28, 1689-1721, (1999).
[16] A. Menezes, P. van Oorschot, S. Vanstone, Handbook of applied cryptography, CRC Press, (1996).
[17] I. Shparlinski, Cryptographic applications of analytic number theory: complexity lower bounds and pseudorandomness, Progress in Computer Science and Applied Logic, 22, Birkhaüser Verlag, Basel, (2003).
[18] T. Lange, A. Winterhof, Incomplete character sums over finite fields and their application to the interpolation of the discrete logarithm by Boolean functions, Acta Arith., 101, No. 3 223-229, (2002).
[19] G. C. Meletiou, A polynomial representation for exponents in $\mathbb{Z}_{p}$, Bull. Greek Math. Soc., 34, 59-63, (1992).
[20] A. L. Wells, A polynomial form for logarithms modulo a prime, IEEE Trans. Inform. Theory, IT-30, 845-846, (1985).
[21] R. L. Burden, J. D. Faires, Numerical Analysis, Brooks/Cole Publishing Company, (1997).


[^0]:    * e-mail: elena@math.upatras.gr, Phone: +302610997348
    ** e-mail: gmelet@teiep.gr, Phone: +30 2681076941
    *** e-mail: vrahatis@math.upatras.gr, Phone: +302610997374

