CSE 533: Error-Correcting Codes **(Autumn 2006)**

Lecture 17: Reed Solomon List Decoding

1 Dec '06

Lecturer: Venkatesan Guruswami *Scribe: Gyanit Singh*

In this lecture we discuss list decoding for Reed Solomon codes. RS-Decoding a given message m = ($m_1, \ldots, m_n$) means, finding a degree k polynomial p(x), which satisfies the message at more than error correction bound ((D - 1)/2; D = min distance) number of places. For the scribes define Q(x,y) = $\sum q_{ij}x^iy^j$.

1 RS List Decoding Problem

For RS codes list decoding a message m with parameters t is to find all the codewords(polynomials) which satisfy the message at atleast t places. This problem can be stated as:

List Decoding Problem
Given n distinct pairs (α_i, y_i) $\in \mathbb{F} \times \mathbb{F}$, a degree parameter k and an agreement parameter t, find all degree k polynomial p(x) such that p(α_i) = y_i for atleast t values of i $\in 1, 2, \ldots, n$.
Goal: Solve for t $> \sqrt{kn}$, decoding a 1 - $\sqrt{R}$ fraction of errors.

---

[1] has the following lemma.

Lemma 1.1. *Given any n points (*α_i, y_i*) $\in \mathbb{F} \times \mathbb{F}$, $\exists$ nonzero Q(X,Y) with* $deg_X(Q) \leq \frac{n}{l}$ *and* $deg_Y(Q) \leq l$*, s.t. Q(*α_i, y_i*) = 0, $\forall$ i.*

Proof. Note that Q(x,y) = $\sum_{0\leq i\leq d_x, 0\leq j\leq d_y} q_{ij}x^iy^j$, and we get there are ($\frac{n}{l}$ + 1)(l + 1) $\geq$ n variables (variables being q_{ij}). So we have a system of homogeneous equation with n constraints and more than n variable. Hence a non-zero solution exists. $\square$

2 Algorithm Schema

1. Find non-zero Q(X,Y) (with some degree restrictions), s.t. Q explains all the points.

2. Factor Q(X,Y) and for each factor of form y - p(x) with deg(p) $\leq$ k; check if p(α_i) = y_i for atleast t values of i. If so output p(X).

Why the above algorithm runs in polynomial time?
Step 1 is solving a system of homogeneous linear equation. Which can be done in polynomial time.
Step 2: Step can also be done in the polynomial time. For details see ([2])

Lemma 2.1. *For a polynomial p(x), s.t. deg(p(x)) $\leq$ k, p(α_i) = y_i for atleast t values and t $> \frac{n}{l} + lk$ then y - p(x) is a factor of Q(x,y).*

Proof. We show this by showing that R(x) = Q(x,p(x)) is a 0 polynomial. For this, we will show that number of roots are greater than the degree of R(x). Note that, deg(R) $\leq$ n/l + lk; because y is replaced by a (atmost) k degree polynomial and $deg_y < l$. If p(α_i) = y_i, then R(α_i) = Q(α_i,P(α_i)) = Q(α_i, y_i) = 0. Number of roots is atleast t. Now t $> \frac{n}{l} + lk$, R is a 0 polynomial. $\square$

We can try to optimize for t by choosing l appropiately. Now n/l + lk $\geq 2\sqrt{nk}$, (AM-GM). For l = $\sqrt{n/k}$, n/l + lk = $2\sqrt{nk}$. Hence this choice of l optimizes for t, which now has to follow t $> 2\sqrt{kn}$.

2.1 Improvement using (1,k)-weighted deg)

Definition 2.2. *For a polynomial Q(x,y) = $\sum_{i\geq 0, j\geq 0} q_{ij}x^i y^j$, define (1,k)-weighted degree of Q(x,y) as maximum (i + kj).*

Lemma 2.3. *Given any n points (α_i, y_i) $\in \mathbb{F} \times \mathbb{F}$, $\exists$ nonzero Q(X,Y) with (1,k)-weighted degree D, s.t. Q(α_i, y_i) = 0, $\forall$ i, for D = $\lfloor\sqrt{2kn}\rfloor$.*

Proof. Let us count the number of coefficient q_{ij} for i $\geq$ 0, j $\geq$ 0 and i+kj le D let there be N.

$$N = \sum_{j=0}^{\lfloor \frac{d}{k} \rfloor} \sum_{i=0}^{D-kj} 1 = \sum_{j=0}^{\lfloor \frac{d}{k} \rfloor} (D - kj + 1)$$

$$= (D+1)(\lfloor \frac{d}{k} \rfloor + 1) - \frac{k\lfloor \frac{d}{k} \rfloor (\lfloor \frac{d}{k} \rfloor + 1)}{2}$$

$$= \frac{(\lfloor \frac{d}{k} \rfloor + 1)}{2}(2D + 2 - k\lfloor \frac{d}{k} \rfloor)$$

$$\geq \frac{(\lfloor \frac{d}{k} \rfloor + 1)}{2}(D+2) \geq \frac{D(D+2)}{2k}$$

For D = $\lfloor\sqrt{2kn}\rfloor$, N $\geq \frac{2kn}{2k} = n$. And hence the system of equation has non-zero solution. $\square$

Now Q(x,y) be the polynomial with (1,k)-weighted degree D = $\lfloor\sqrt{2kn}\rfloor$.

Theorem 2.4. *For a polynomial p(x), s.t. deg(p(x)) $\leq$ k, if p(α_i) = y_i for atleast t values and t $> \sqrt{2kn}$ then y - p(x) is a factor of Q(x,y).*

Proof. Again consider R(x) = Q(x,p(x)). We will show that number of roots of R(x) are greater than the degree of R(x). Note that, deg(R) $\leq$. If p(α_i) = y_i, then R(α_i) = Q(α_i,P(α_i)) = Q(α_i, y_i) = 0. Number of roots is atleast t. Now t $> \sqrt{2kn}$, R is a 0 polynomial or y - p(x) is a factor of Q(x,y). $\square$

So this will give us a decoding fraction of p = 1 - $\sqrt{2R}$. Note that as R $\longrightarrow$ 0, p $\longrightarrow$ 1.

3 Improvements to match Johnson Bound

Now, we will consider improvements to match Johnson bound, t $> \sqrt{kn}$. The main idea here is weighted polynomial reconstruction. For each pair (α_i, y_i) we are also given an integer weight w_i as input.
Let Q(x,y) be polynomial such that Q(α_i, y_i) = 0, $\forall$ i. We impose a stronger condition for the points (α_i, y_i) with higher w_i; i.e. Q(x,y) has a root of multiplicity w_i at (α_i, y_i).

Definition 3.1. *Given a polynomial Q(x,y), define Q^i(x,y) as the polynomial, s.t. Q(α_i, y_i) = Q^i(0,0). In general Q^i(x,y) = Q(x+α_i,y+y_i).*

Given Q(x,y) and a pair (α_i, y_i), Q^i(x,y) = $\sum q^i_{rs} x^r y^s$. To see how q^i_{rs} is related to coefficients of Q(x,y), note that

$$Q^i(x,y) = \sum_{r,s} q_r s(x+\alpha_i)^r (y+y_i)^s \qquad This\ \ gives$$

$$q^i_{rs} = \sum_{r' \geq r, s' \geq s} q_{r's'} (\begin{pmatrix} r' \\ r \end{pmatrix} \alpha_i^{r'-r} \begin{pmatrix} s' \\ s \end{pmatrix} y_i^{s'-s})$$

The w_i multiplicity of root implies that partial derivaties upto total of w_i order are all zero at that point. More precisely

$$[\frac{\partial}{\partial x^r} \frac{\partial}{\partial y^s} Q(x,y)](\alpha_i, y_i) = 0 \qquad \forall r, s, \quad s.t. \quad r+s < w_i$$

or

$$[\frac{\partial}{\partial x^r} \frac{\partial}{\partial y^s} Q^i(x,y)](0,0) = 0 \qquad \forall r, s, \quad s.t. \quad r+s < w_i$$

i.e. q^i_{rs} = 0 whenever r + s $<$ 0.

Let N_i = Number of constraints introduced to impose the w_i multiplicity of root (α_i, y_i) for Q(x,y).

$$N_i = \sum_{r=0}^{w_i-1} \sum_{s=0}^{w_i-r-1} 1 \quad = \quad \sum_{r=0}^{w_i-1} w_i - r \quad = \quad w_i * w_i - \frac{w_i * (w_i - 1)}{2}$$

$$= w_i * \frac{w_i + 1}{2} \quad = \quad \begin{pmatrix} w_i + 1 \\ 2 \end{pmatrix}$$

Lemma 3.2. *Given any n points $(\alpha_i, y_i) \in \mathbb{F} \times \mathbb{F}$ and corresponding integer weights w_i, $\exists$ nonzero Q(X,Y) with (1,k)-weighted degree D, s.t. Q(x,y) has (α_i, y_i) as a root with w_i multiplicity, $\forall$ i, for D =* $\lfloor \sqrt{2k \sum \begin{pmatrix} w_i + 1 \\ 2 \end{pmatrix}} \rfloor$.

Proof. Let us count, the number of constraints. Total constraints = $\sum N_i$ = $\sum \binom{w_i+1}{2}$. Now let us count,the number of variables. As in the proof of 2.3 we know that number of variables $>$ $\frac{D^2)}{2k}$ For D = $\lfloor\sqrt{2k\sum \binom{w_i+1}{2}}\rfloor$, number of variables $> \sum \binom{w_i+1}{2}$. Number of variables are greater than the number of constraints, hence the system of equation has non-zero solution. □

Lemma 3.3. *If p(α_i) = y_i and Q(x,y) has w_i roots at (α_i, w_i) then R(x), defined as Q(x,p(x)) is divisible by* $(x-\alpha_i)^{w_i}$.

Given n distinct pairs (α_i, y_i) $\in \mathbb{F} \times \mathbb{F}$, with associated integer weights $w_i \geq 1$, find all degree k polynomial p(x) such that $\sum_{i:p(\alpha_i)=y_i} w_i > W$, for some weighted arguement parameter W.

We will solve this for W = $\sqrt{2k\sum \binom{w_i+1}{2}}$.

Lemma 3.4. *For a polynomial p(x), s.t deg(p(x)) $\leq$ k, if* $\sum_{i:p(\alpha_i)=y_i} w_i > W$ *and W =* $\sqrt{2k\sum \binom{w_i+1}{2}}$, *then y - p(x) is a factor of Q(x,y).*

Proof. Consider R(x) = Q(x,p(x)). Degree of R(x) = D, as Q(x,y) is of (1,k)-weighted degree D. Now Lemma 3.3 says that if $p(\alpha_i) = w_i$ then α_i is a root of multiplicatiy w_i of R(x). Number of roots of R(x) = $\sum_{i:p(\alpha_i)=y_i} w_i$. Now number of roots ¿ W = D. Hence R(x) is a 0 polynomial or y - p(x) divides Q(x,y). □

Now for

$$w_i = 1, \qquad t > \sqrt{2kn}$$
$$w_i = 2, \qquad 2t > \sqrt{6kn}$$

$\sqrt{3kn/2}$ is an improvement from $\sqrt{2kn}$. We can use this approach to get better results. If we pick

$$w_i = w = 2kn, \qquad t > \sqrt{2kn\frac{w+1}{2w}} \quad = \quad \sqrt{kn + kn/w} \quad = \quad \sqrt{kn + 1/2}$$

Assuming Lemma 3.3, we can obtain our goal of solving list decoding problem defined earlier for t $> \sqrt{kn}$, i.e. decoding a 1 - $\sqrt{R}$ fraction of errors.

References

[1] Madhu Sudan. Decoding of Reed-Solomon codes beyond the error-correction bound. *Journal of Complexity*, 13(1):180-193, 1997.

[2] E Kalfoten. Polynomial Factorization. *LATIN*, 92.