Error Correcting Codes: Combinatorics, Algorithms and Applications **(Fall 2007)**

Lecture 12: Reed-Solomon Codes

September 28, 2007

Lecturer: Atri Rudra *Scribe: Michel Kulhandjian*

Last lecture we saw the proof of the Singleton bound which claims that for any $(n, k, d)_q$ code, $k \leq n - d + 1$. In today's lecture we will study Reed-Solomon codes. These codes meet the Singleton bound, i.e. satisfy $k = n - d + 1$ (but have the unfortunate property that $q \geq n$). Note that this implies that the Singleton bound is tight, at least for $q \geq n$.

1 Reed-Solomon Codes

We begin with the definition of Reed-Solomon codes.

Definition 1.1 (Reed-Solomon code)**.** *Let $\mathbb{F}_q$ be a finite field and $\mathbb{F}_q[x]$ denote the $\mathbb{F}_q$-space of univariate polynomials where all the coefficients of x are from $\mathbb{F}_q$. Pick $\{\alpha_1, \alpha_2, ...\alpha_n\}$ distinct elements (also called* evaluation points*) of $\mathbb{F}_q$ and choose n and k such that $k \leq n \leq q$. We define an encoding function for Reed-Solomon code as $RS : \mathbb{F}_q^k \to \mathbb{F}_q^n$ as follows. A message $\mathbf{m} = (m_0, m_1, ..., m_{k-1})$ with $m_i \in \mathbb{F}_q$ is mapped to a degree $k - 1$ polynomial.*

$$\mathbf{m} \mapsto f_{\mathbf{m}}(x),$$

where

$$f_{\mathbf{m}}(x) = \sum_{i=0}^{k-1} m_i x^i. \tag{1}$$

Note that $f_{\mathbf{m}}(x) \in \mathbb{F}_q[x]$ is a polynomial of degree $\leq k - 1$. The encoding of $\mathbf{m}$ is the evaluation of $f_{\mathbf{m}}(x)$ at all the α_i's :

$$RS(\mathbf{m}) = \langle f_{\mathbf{m}}(\alpha_1), f_{\mathbf{m}}(\alpha_2), ..., f_{\mathbf{m}}(\alpha_n)\rangle$$

We call this image Reed-Solomon code or RS code after two inventors Irving Reed *and* Gus Solomon *of this code [1]. A common special case is $n = q - 1$ with the set of evaluation points being $\mathbb{F}^* \triangleq \mathbb{F} \setminus \{0\}$.*

Notice that by definition, the entries in $\{\alpha_1, ..., \alpha_n\}$ are distinct and thus, must have $n \leq q$.
We now turn to some properties of Reed-Solomon codes.

Claim 1.2. *RS codes are linear codes.*

Proof. The proof follows from the fact that if $a \in \mathbb{F}_q$ and $f(x), g(x) \in \mathbb{F}_q[x]$ are polynomials of degree $\leq k-1$, then $af(x)$ and $f(x)+g(x)$ are also polynomials of degree $\leq k-1$. In particular, let messages $\mathbf{m}_1$ and $\mathbf{m}_2$ be mapped to $f_{\mathbf{m}_1}(x)$ and $f_{\mathbf{m}_2}(x)$ where $f_{\mathbf{m}_1}(x), f_{\mathbf{m}_2}(x) \in \mathbb{F}_q[x]$ are polynomials of degree $\leq k-1$ and because of the mapping defined in (1), it is easy to verify that:

$$f_{\mathbf{m}_1}(x) + f_{\mathbf{m}_2}(x) = f_{\mathbf{m}_1+\mathbf{m}_2}(x),$$

and

$$af_{\mathbf{m}_1}(x) = f_{a\mathbf{m}_1}(x).$$

Therefore,

$$RS(\mathbf{m}_1) + RS(\mathbf{m}_2) = RS(\mathbf{m}_1 + \mathbf{m}_2)$$
$$aRS(\mathbf{m}_1) = RS(a\mathbf{m}_1)$$

Therefore RS is a $[n,k]_q$ linear code. □

The second and more interesting claim is the following:

Claim 1.3. *RS is a $[n,k,n-k+1]_q$ code. That is, it matches the Singleton bound.*

The claim on the distance follows from the fact that every polynomial of degree $k-1$ over $\mathbb{F}_q[x]$ has at most $k-1$ (not necessarily distinct) roots, and that if two polynomials agree on more than $k-1$ places then they must be the same polynomial.

Proposition 1.4 ("Degree Mantra")**.** *A nonzero polynomial $f(x)$ of degree t over a field $\mathbb{F}_q$ has at most t roots in $\mathbb{F}_q$*

Proof. We will prove the theorem by induction on t. If $t = 0$, we are done. Now, consider $f(x)$ of a degree $t > 0$. Let $\alpha \in \mathbb{F}_q$ be a root such that $f(\alpha) = 0$. If no such root α exists, we are done. If there is a root α, then we can write

$$f(x) = (x-\alpha)g(x)$$

where $\deg(g) = \deg(f) - 1$ (i.e. $x-\alpha$ divides $f(x)$). This is because by the fundamental rule of division of polynomials:

$$f(x) = (x-\alpha)g(x) + R(x)$$

where $\deg(R) \leq 0$ (as the degree cannot be negative this in turn implies that $\deg(R) = 0$) and since $f(\alpha) = 0$,

$$f(\alpha) = 0 + R(\alpha),$$

which implies that $R(x) = 0$. By induction, $g(x)$ has at most $t-1$ roots, which implies that $f(x)$ has at most t roots. □

We are now ready to prove Claim 1.3

Proof of Claim 1.3 We start by proving the claim on the distance. Fix arbitrary $\mathbf{m}_1 \neq \mathbf{m}_2 \in \mathbb{F}_q^k$. Note that $f_{\mathbf{m}_1}(x), f_{\mathbf{m}_2}(x) \in \mathbb{F}_q[x]$ are distinct polynomials of degree $\leq k-1$ since $\mathbf{m}_1 \neq \mathbf{m}_2 \in \mathbb{F}_q^k$. Then $f_{\mathbf{m}_1}(x) - f_{\mathbf{m}_2}(x) \neq 0$ also has degree $\leq k-1$. Note that $w(RS(\mathbf{m}_2) - RS(\mathbf{m}_1)) = \Delta(RS(\mathbf{m}_1), RS(\mathbf{m}_2))$. The weight of $RS(\mathbf{m}_2) - RS(\mathbf{m}_1)$ is n minus the number of 0's in $RS(\mathbf{m}_2) - RS(\mathbf{m}_1)$ which is equal to n minus the number of roots that $f_{\mathbf{m}_1}(x) - f_{\mathbf{m}_2}(x)$ has among $\{\alpha_1, ..., \alpha_n\}$. That is,

$$\Delta(RS(\mathbf{m}_1), RS(\mathbf{m}_2)) = n - |\{\alpha \mid f_{\mathbf{m}_1}(\alpha) = f_{\mathbf{m}_2}(\alpha)\}|$$

By Proposition 1.4, $f_{\mathbf{m}_1}(x) - f_{\mathbf{m}_2}(x)$ has at most $k-1$ roots. Thus, the weight of $RS(\mathbf{m}_2) - RS(\mathbf{m}_1)$ is at least $n-(k-1) = n-k+1$. Therefore $d \geq n-k+1$, and since the Singleton bound implies that $d \leq n-k+1$, we have $d = n-k+1$.[1] The argument above also shows that distinct polynomials $f_{\mathbf{m}_1}(x), f_{\mathbf{m}_2}(x) \in \mathbb{F}_q[x]$ are mapped to distinct codewords. Therefore, the code contains $[q]^k$ codewords and has dimension k. The claim in linearity of the code follows from Claim 1.2. □

Definition 1.5 (MDS codes)**.** *An $(n,k,d)_q$ code is called* Maximum Distance Separable (MDS) *if* $d = n-k+1$.

Thus, Reed-Solomon codes are MDS codes.

Let us now find a generator matrix for RS codes (which exists by Claim 1.2). By Definition 1.1, any basis $f_{\mathbf{m}_1}, ..., f_{\mathbf{m}_k}$ of polynomial of degree at most $k-1$ gives rise to a basis $RS(\mathbf{m}_1), ..., RS(\mathbf{m}_k)$ of the code. A particularly nice polynomial basis is the set of monomials $1, x, ..., x^i, ..., x^{k-1}$. The corresponding generator matrix, whose ith row (numbering rows from 0 to $k-1$) is

$$(\alpha_1^i, \alpha_2^i, ..., \alpha_j^i, ..., \alpha_n^i)$$

and this generator matrix is called the *VanDerMonde* matrix with $k \times n$ size

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_j & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_j^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \alpha_1^i & \alpha_2^i & \cdots & \alpha_j^i & \cdots & \alpha_n^i \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_j^{k-1} & \cdots & \alpha_n^{k-1} \end{pmatrix}$$

RS codes are used in storage of information in CD's because they are robust against burst-errors that come in contiguous manner, unlike the random error method studied by Shanon. The drawback of Reed-Solomon codes is the condition that $q \geq n$. The problem is that we need each coordinate of a codeword to correspond to a distinct element of $\mathbb{F}_q$

[1] Alternatively, consider the distance between the all zero codeword and the codeword corresponding to the polynomial $\prod_{i=1}^{k-1}(x-\alpha_i)$.

Remark 1.6. *One might ask does q have to vary as a function of n to satisfy the Singleton bound? The answer is yes. We can show this by the Plotkin bound, which we will prove in a couple of lectures.*

2 Hamming versus Shannon

Let us compare Hamming and Shannon theories in terms of the asymptotic bounds we have seen so far (recall rate $R = \frac{k}{n}$ and relative distance $\delta = \frac{d}{n}$).

- Hamming theory: Can correct $\leq \frac{\delta}{2}$ fraction of worse case errors for codes of distance δ. By the Singleton bound,
$$\delta \leq 1 - R,$$
which implies that p fraction of errors can be corrected, where
$$p \leq \frac{1-R}{2}$$
The above can be achieved via efficient decoding algorithms for RS codes.

- Shannon theory: In $qSCp$, we can have reliable communication with $R < 1 - H_q(p)$. It can be shown that
  1. $1 - H_q(p) \leq 1 - p$
  2. $1 - H_q(p) \geq 1 - p - \varepsilon$, iff $q = 2^{\Omega(1/\varepsilon)}$ for large q.

Thus we can have reliable communication with $p \sim 1 - R$ on qSC_p for large enough q.

Remark 2.1. *There is a gap between Shannon and Hamming world: one can correct twice as many errors as in the Shannon world. One natural question to ask is whether we can somehow "bridge" this gap.*

We will now re-visit the the bad example for unique decoding and consider an extension of the bad example as shown in Figure 1.

Recall that $\bar{\mathbf{y}}$ and the codewords $\mathbf{c}_1$ and $\mathbf{c}_2$ form the bad example for unique decoding that we have already seen before. Recall that for this particular received word we can not do error recovery by unique decoding since there are two codewords $\mathbf{c_1}$ and $\mathbf{c_2}$ having the same distance $\frac{\delta}{2}$ from vector $\bar{\mathbf{y}}$. On the other hand, the received word $\mathbf{z}$ has an unique codeword $\mathbf{c_1}$ with distance $p > \frac{\delta}{2}$. However, unique decoding does not allow for error recovery form $\mathbf{z}$. This is because by definition of unique decoding, the decoded codeword cannot have Hamming distance larger than $\delta/2$ from the received word. In this example, there is no codeword within Hamming distance $\delta/2$ of $\mathbf{z}$ and thus, it can not correct the received word $\mathbf{z}$. In this example is because of the received word $\bar{\mathbf{y}}$.

Let us consider the example in Figure 1 for the binary case. It can be shown that the number of vectors in dotted lines is insignificant compared to volume of shaded area (for large enough block

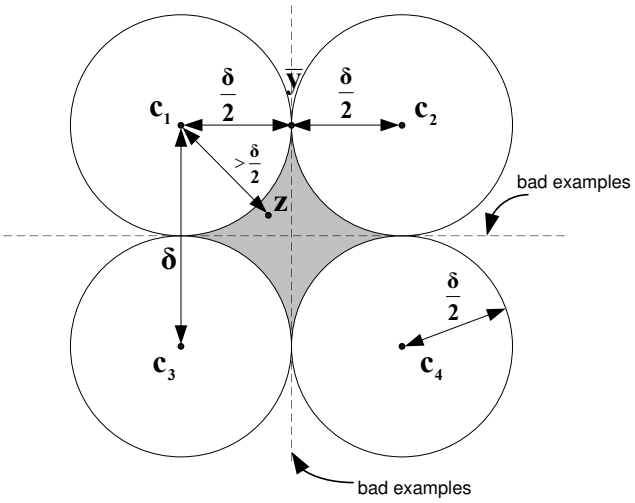

Figure 1: In this example vectors are embedded into Euclidean space such that the Euclidean distance between two mapped points is the same as the Hamming distance between vectors. The $\mathbf{c_1}, \mathbf{c_2}, \mathbf{c_3}, \mathbf{c_4}$ are codewords. The dotted lines contain the "bad examples," that is, the received words for which unique decoding is not possible.

length of the code). The volume of all Hamming balls radius of $\frac{\delta}{2}$ around all the codewords is roughly equal to:

$$2^k 2^{nH(\frac{\delta}{2})},$$

which implies that the volume of the shaded area (without the dotted lines) is approximately equal to:

$$2^n - 2^k 2^{nH(\frac{\delta}{2})}.$$

In other words, the volume when expressed as a fraction of the volume of the ambient space is roughly:

$$1 - 2^{-n(1-H(\frac{\delta}{2})-R)}, \tag{2}$$

where $k = Rn$ and $R \leq 1 - H(\frac{\delta}{2})$. If $R < 1 - H(\frac{\delta}{2})$ then second term of (2) is very small. Therefore the number of vectors in shaded area (without the bad examples) is almost all of the ambient space. Note that by the stringent condition on unique decoding none of these received words can be decoded (even though for such received words there is a unique closest codeword). Thus, in order to be able to decode such received vectors, we need to relax the notion of unique decoding. We will consider such a relaxation called *list decoding* in the next lecture.

References

[1] Irving S. Reed and Gustav Solomon. Polynomial codes over certain finite fields. *SIAM Journal on Applied Mathematics*, 8(2):300–304, 1960.