

COMPUTATION OF THE WEIGHT DISTRIBUTION OF CRC CODES.

FELICE MANGANIELLO

1. INTRODUCTION

This poster illustrates an algorithm for the computation of the weight distribution of CRC codes.

Cyclic Redundancy Check (CRC) codes are an important class of error detecting codes. These codes are used a lot nowadays in computer communication networks because of their properties, e.g. their easy and fast encoder and decoder implementation, and their considerable burst-error detection capability ([Wic95]).

To measure the degree of goodness of error detecting codes we have to investigate about two properties: the *minimum distance* of the code and its *undetected error probability* (P_{ue}). To investigate this two properties it is important to know the *weight distribution* of the code.

A way to compute this distribution is to list all the words of the code and compute their Hamming weight. There is no other way to compute the weight distribution independently from the structure of the code. The recursive structure of CRC codes offers the opportunity to construct an *ad-hoc* algorithm that has lower computational cost; see [CBH93].

This work is a generalization of the paper [CBH93], where the algorithm is shown in the binary case. This poster illustrates the extension of the algorithm to CRC codes over any finite field.

2. PRELIMINARIES

We give a definition of this codes from [Ros01] to then go back to the more operational one.

Definition 1. Let $g(x) \in \mathbb{F}_q[x]$ be a monic polynomial over the finite field \mathbb{F}_q , whose characteristic is p . Let us consider the encoding map

$$\begin{aligned} \phi : \mathbb{F}_q[x] &\rightarrow \mathbb{F}_q[x] \\ m(x) &\mapsto c(x) = m(x)g(x). \end{aligned}$$

A (CRC) code is the ideal

$$(g(x)) = \text{Im } \phi.$$

This definition gives the basic property of CRC code, i.e. the fact that they are generated by a *generator polynomial* $g(x)$.

Such a definition of CRC codes is appropriate from a theoretical point of view, but this definition is not enough in applications. The resulting code is not *observable*, see [Ros01].

The only way to correct this problem in real applications is to set an *a priori* bound to the length of the message. This way, once the message $c(x)$ has been received and the long division

$$c(x) = \tilde{m}(x)g(x) + r(x)$$

has been computed, if $r(x) = 0$ then the receiver can conclude that $\tilde{m}(x)$ is the original message $m(x)$. Otherwise a retransmission will be requested.

Hence we conclude that every CRC code can be represented also in the following way.

Definition 2. Let n, r belong to \mathbb{N} with $n > r > 0$, let $q \in \mathbb{N}$ be some power of a prime number p and $g(x) \in \mathbb{F}_q[x]$ a monic polynomial such that $\deg g(x) = r$ and $g(0) \neq 0$.

A CRC code C of length n is the set

$$C = \{c(x) \in \mathbb{F}_q[x] \mid c(x) = g(x)m(x), \deg m(x) < n - r\}.$$

Such a set has the structure of a linear code with regard to the natural isomorphism onto the ring $\mathbb{F}_q[x]/(x^n - 1)$.

Remark. A CRC code is a cyclic code if and only if the generator polynomial $g(x)$ divides $x^n - 1$.

CRC codes are close to cyclic codes. In fact a set of generators of a CRC code is the same as the *canonical* one of a cyclic code.

To turn our interest to the dual code of a CRC is a first way to decrease the computational weight in order to compute the weight distribution of the code. In a real-world use of a CRC code, the length of the code is greater than the degree of the generator polynomial. In terms of cardinality this difference becomes even bigger, as the cardinality of the CRC code is q^{n-r} whereas that of the dual code is q^r . This step is possible thanks to MacWilliams' theory [MS88].

Remark. The dual code of a CRC code has an interesting structure. Given a polynomial $g(x)$ over \mathbb{F}_q , the dual code of a CRC code of any length generated by $g(x)$ is isomorphic as a vector space to the ring $\mathbb{F}_q[x]/(g(x))$.

The words of the dual of a CRC code can be characterized in the following way.

Remark. Let C be a CRC code over \mathbb{F}_q^n and $g(x) = g_0 + g_1x + \cdots + g_{r-1}x^{r-1} + x^r$ its generator polynomial. Then for $i = r, \dots, n$, the following equivalence holds:

$$c = (c_0, \dots, c_{n-1}) \in C^\perp \iff c_i = -g_0c_{i-r} - \cdots - g_{r-1}c_{i-1}.$$

2.1. Notations. During the whole work we will use the following notations:

- p will be the prime number that is the characteristic of the ring \mathbb{F}_q ; then q is a power of p , i.e. $q = p^\delta$ for some $\delta \in \mathbb{N}_+$;
- $n \in \mathbb{N}$ will be the length of the CRC code;
- $g(x) \in \mathbb{F}_q[x]$ will be the monic generator polynomial of CRC code, with $g(0) \neq 0$, $\deg g(x) = r$ and $0 < r < n$;
- $g(x) = \prod_{i=1}^m g_i(x)^{e_i}$ will be the irreducible decomposition of $g(x)$.

3. FUNDAMENTAL STEP OF THE ALGORITHM

For the sake of clearness we begin to explain the step that characterizes the algorithm. First of all we state some results related to this algorithmic step.

Lemma 3. *Let $C \subset \mathbb{F}_q^n$ be a CRC code with generator polynomial $g(x)$, and let $(c_i)_{i \in \mathbb{N}} \subset \mathbb{F}_q$ be a Linear Recurring Sequence (LRS) whose characteristic polynomial is $g(x)$; then*

$$(c_k, \dots, c_{k+n-1}) \in C^\perp \quad \forall k \in \mathbb{N}$$

The previous Lemma gives us a way to “extract” words of the dual code using only a LRS and the length of the code.

The next Lemma gives a bound on the number of different words that can be extracted from a fixed LRS.

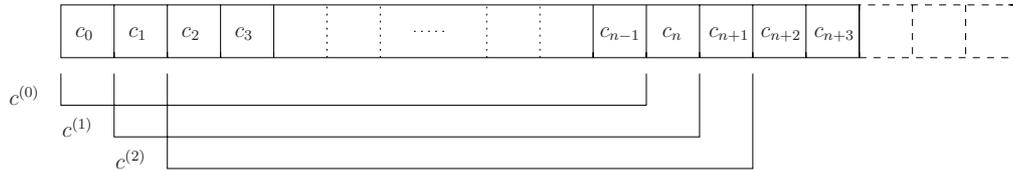
Lemma 4. *Let $(c_i)_{i \in \mathbb{N}} \subset \mathbb{F}_q$ be a LRS with generator polynomial $g(x)$ and $u(x) \in \mathbb{F}_q[x]$ be the polynomial related to $(c_i)_{i \in \mathbb{N}}$ (see Theorem 5).*

Let $C_u^\perp \subset C^\perp$ be the set of words of the dual code extracted from the $(c_i)_{i \in \mathbb{N}}$. Then the cardinality of C_u^\perp is

$$|C_u^\perp| = \text{ord} \left(\frac{g(x)}{\text{gcd}(g(x), u(x))} \right).$$

The next section is devoted to the fundamental step of the algorithm.

3.1. The Fundamental Step. Let us now consider a LRS $(c_i)_{i \in \mathbb{N}}$ with characteristic polynomial $g(x)$. We will make the way to extract words of the dual code of a $(n, n - r)$ CRC code explicit. The following figure will depict the idea of the mechanism of the algorithm. This scheme follows from Lemma 3.



In the figure above, $c^{(k)} \subset \mathbb{F}_q^n$ denotes the k -th word of the dual code extracted from the above sequence.

The figure leads directly to relations between the weight distribution of the words thus extracted. These relations are resumed in the next Remark.

Remark. (**Weight relations between words**)

- if $c_{k-1} \neq 0$ and $c_{k+n-1} = 0$, then $wt(c^{(k)}) = wt(c^{(k-1)}) - 1$;
- if $c_{k-1} = 0$ and $c_{k+n-1} \neq 0$, then $wt(c^{(k)}) = wt(c^{(k-1)}) + 1$;
- $wt(c^{(k)}) = wt(c^{(k-1)})$ otherwise.

Once the weight of the first “extracted” word is computed, the weight of the others is quickly resumable.

Now we are able to “extract” words of the dual code from LRS’s efficiently, but some questions arise:

- how many LRS’s are enough to determine the weight distribution of the dual code?
- how can we be sure that we are not considering the same word more than once?
- which LRS’s should we use, and which ones should we not?

4. LRS’S AND x -ORBITS OF \mathcal{R}_q^g

This section gives relations between the three structures of the title.

Theorem 5. *Let $u(x) \in \mathbb{F}_q[x]$ be a polynomial with $\deg u(x) < \deg g(x)$. Then there exists exactly one sequence $(c_i)_{i \in \mathbb{N}} \subset \mathbb{F}_q$ such that*

$$(1) \quad \frac{u(x)}{g(x)} = \sum_{i=0}^{\infty} \frac{c_i}{x^{i+1}} =: c(1/x).$$

Moreover the sequence $(c_i)_{i \in \mathbb{N}}$ satisfies the linear relation

$$(2) \quad c_i = -g_0 c_{i-r} - \cdots - g_{r-1} c_{i-1}, \quad i \geq r.$$

Remark. We can compute the series (1) explicitly in the following way.

Let $u(x) \in \mathbb{F}_q[x]$ be a polynomial satisfying the hypothesis of Theorem 5. Then the following relation holds

$$(3) \quad \frac{u(x)}{g(x)} = \frac{u_{r-1}}{x} + \frac{u'(x)}{xg(x)},$$

where $u'(x) = xu(x) - u_{r-1}g(x) \equiv xu(x) \pmod{g(x)}$ is a polynomial satisfying the same hypothesis as $u(x)$. We obtain the series recursively via the above relation.

As an immediate application of the Theorem, one obtains the following Corollary.

Corollary 6. *There exists a bijection between the ring $\mathbb{F}_q[x]/(g(x))$ and the set of all LRS's whose characteristic polynomial is $g(x)$.*

Now we define another algebraic structure.

Definition 7. *The x -orbits of the ring $\mathbb{F}_q[x]/(g(x))$ result by the action of the subgroup $\langle x \rangle$ on the ring. We denote with \mathfrak{C}_u^\perp the x -orbit of $u(x) \in \mathbb{F}_q[x]/(g(x))$.*

The next Theorem states the relation between x -orbits, LRS's and words of the dual code.

Lemma 8. *Let $u_1(x), u_2(x)$ be two distinct elements of $\mathbb{F}_q[x]/(g(x))$. Then the following relation holds:*

$$u_2(x) \in \mathfrak{C}_{u_1}^\perp \iff C_{u_1}^\perp = C_{u_2}^\perp.$$

In order to avoid repetitions of words we have to characterize the x -orbits of $\mathbb{F}_q[x]/(g(x))$. To do this we have to find representatives of these orbits.

5. REPRESENTATIVES OF $\mathbb{F}_q[x]/(g(x))$

In this section we reduce the problem of the determination of the x -orbits of a ring to the determination of the x -orbits of some subsets.

5.1. Chinese Remainder Theorem. Thanks to the properties of the actions of cyclic groups, to [CBH93], and to the Chinese Remainder Theorem we can turn our attention back to the case of x -orbits of quotient rings of powers of irreducible polynomials.

$$\mathbb{F}_q[x]/(g(x)) \xrightarrow{\text{reduces}} \mathbb{F}_q[x]/(g_l(x)^{e_l})$$

with $g_l(x) \in \mathbb{F}_q[x]$ an irreducible polynomial in the irreducible decomposition of $g(x)$.

5.2. $\mathbb{F}_q[x]/(g(x)^t)$ **as Disjoint Union.** We now consider an irreducible polynomial $g(x) \in \mathbb{F}_q[x]$. In studying the elements of $\mathbb{F}_q[x]/(g(x)^t)$, we obtain the next corollary.

Corollary 9. *The ring $\mathbb{F}_q[x]/(g(x)^t)$ can be decomposed as follows:*

$$\mathbb{F}_q[x]/(g(x)^t) = \{0\} \cup \bigsqcup_{l=1}^{t-1} \left\{ g(x)^l \cdot M_{g^{t-l}}^q \right\}$$

where $M_{g^{t-l}}^q := (\mathbb{F}_q[x]/(g(x)^{t-l}))^*$ and the sets

$$g(x)^l \cdot M_{g^{t-l}}^q = \left\{ u \in \mathbb{F}_q[x]/(g(x)^t) \mid u = [g(x)^l \bar{u}(x)], \bar{u} \in M_{g^{t-l}} \right\}$$

are stable under x -multiplication.

And then we obtain

$$\mathbb{F}_q[x]/(g(x)^t) \xrightarrow{\text{reduces}} M_{g^t}^q,$$

with $0 \leq l < t$.

5.3. **Structure of $M_{g^l}^q$ and its x -orbits' Representative.** From now on we will make use of the following notation

$$(4) \quad a_{i,j,k}(x) = 1 + \alpha^i x^j g(x)^k,$$

where $0 \leq i < \delta$, $0 \leq j < r$ and $1 \leq k < l$.

The multiplicative group $M_{g^l}^q$ is finite, hence it can be represented as a product of cyclic groups. The next Theorem gives a set of generators of these cyclic groups.

Theorem 10. *The order of the group $M_{g^l}^q$ is $(q^r - 1)q^{(l-1)r}$. Moreover,*

$$M_{g^l}^q \approx M_g^p \times S_p,$$

where S_p is the p -Sylow subgroup of $M_{g^l}^q$.

Moreover S_p can be represented as

$$S_p \approx \prod_{i,j,k} (a_{i,j,k}(x)),$$

where $(a_{i,j,k}(x)) \subset S_p$ are the cyclic groups generated by $a_{i,j,k}(x)$ respectively, with parameters satisfying conditions (4), and where the parameter k satisfies also the condition $p \nmid k$.

After some calculations we are able to obtain a Theorem that gives the representatives of the groups $M_{g^l}^q$.

Theorem 11. *Let $g(x) \in \mathbb{F}_q[x]$ be a degree- r irreducible polynomial such that $g(0) \neq 0$ and $l \geq 2$. There exist $0 \leq i_0 < \delta - 1$ and $0 \leq j_0 < \deg g(x)$ such that the set*

$$\left\{ h(x)^t \cdot \prod_{\substack{(i,j,k) \neq (i_0,j_0,1) \\ 0 \leq i < \delta, 0 \leq j < r \\ 1 \leq k \leq t, p \nmid k}} (1 + \alpha^i x^j g(x)^k)^{c_{(ijk)}} \pmod{g(x)^l} \right\},$$

where $h(x)$ is a primitive element of $\mathbb{F}_q[x]/(g(x))$ and $\alpha \in \mathbb{F}_q$ is an algebraic element of degree δ over \mathbb{F}_p , is a representative family of orbits in $M_{g^l}^q$, for t and c such that

- $0 \leq t < \frac{q^r - 1}{\text{ord}(g(x))}$,
- $0 \leq c_{(ijk)} \leq p^{\lceil \log_p l/k \rceil}$.

REFERENCES

- [CBH93] G. Castagnoli, S. Bräuer, and M. Hermann. *Optimization of Cyclic Redundancy-Check Codes with 24 and 32 Parity Bits*. IEEE Trans. on Communication, Vol. 41(No. 6):883–892, June 1993.
- [MS88] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North Holland, 1988.
- [Ros01] J. J. Rosenthal. *Connection between Linear Systems and Convolutional Codes*, 2001.
- [Wic95] S. B. Wicker. *Error Control Systems for Digital Communication and Storage*. Practice Hall, 1995.

MATHEMATICS INSTITUTE, WINTERTHURERSTR. 190, CH - 8057 ZÜRICH
E-mail address: felice.manganiello@math.unizh.ch