Correction Capabilities of the Reed-Solomon Codes Decoded with the Guruswami-Sudan Algorithm

Zsolt Polgar, Vasile Bota, Mihaly Varga

Technical University of Cluj-Napoca
Communications Department
Data Transmissions Laboratory

26, G.Baritiu St. 400027 CLUJ-NAPOCA, ROMANIA

Zsolt.Polgar@com.utcluj.ro

The purpose of Guruswami-Sudan decoding algorithm

- The Guruswami-Sudan (GS) decoding algorithm is a list type decoding algorithm intended for Reed-Solomon codes.
- The purpose of this algorithm is to ensure a correction capability beyond the limit of $(d_{min}-1)/2$ of the classical decoding algorithms.
- The GS decoding algorithm is a maximum likelihood-type hard decoding algorithm.
- Exceeding the above decoding limit imposes multiple solutions for the decoded codeword.
- Exceeding the decoding limit is possible practically due the employment of syndromes that are not considered by classical algorithms.

Data Transmission Laboratory

Definition of Reed-Solomon (RS) Codes

- Let F_q denote a field of size q and let $F_q^k[x]$ denote the vector space of polynomial of degree at most k over F_q.
- Let α be a primitive element of the field F_q and let $\{\alpha^0, \alpha^1, \ldots, \alpha^{n-2}, \alpha^{n-1}\} \in F_q$ n distinct elements of the field F_q.
- The length of the RS code is n=q-1 and k designate the dimension of the code.
- **RS codes as evaluation codes**
- The code can be defined as the evaluation of all polynomial f of degree at most k-1 at the points $\{\alpha^0, \alpha^1, \ldots, \alpha^{n-2}, \alpha^{n-1}\}$ [1].

$$f(x) = \sum_{j=0}^{k-1} v_j \cdot x^j \qquad (1)$$

$$RS_G(k) = \left\{\left(f(\alpha^0), f(\alpha^1), \ldots, f(\alpha^{n-1})\right), \deg\ f < k\right\};\ f \in F_q[x] \qquad (2)$$

- The generator matrix, G, of the code has the entries $G_{ij} = \alpha^{ij}$ for i=0,…,k-1 and j=0,…,n-1.

$$G = \begin{bmatrix} 1 & 1 & 1 & 1 & \ldots & 1 \\ 1 & \alpha & \alpha^2 & \alpha^3 & \ldots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & \ldots & \alpha^{n-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \alpha^{k-1} & \alpha^{2k-2} & \alpha^{3k-3} & \ldots & \alpha^{n-k+1} \end{bmatrix} \qquad (3)$$

Data Transmission Laboratory

Definition of Reed-Solomon (RS) Codes

- **RS codes as duals of evaluation codes**
- The check matrix, H, is constructed by evaluating the polynomials x^i for i=0,…,k-1 at the field elements $\{\alpha^0, \alpha^1, \ldots, \alpha^{n-2}, \alpha^{n-1}\}$ [1].

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & \ldots & 1 \\ \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \ldots & \alpha^{n-k} \\ \alpha^2 & \alpha^4 & \alpha^6 & \alpha^8 & \ldots & \alpha^{2n-2k} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha^{n-1} & \alpha^{n-2} & \alpha^{n-3} & \alpha^{n-4} & \ldots & \alpha^k \end{bmatrix} \quad (4)$$

- **RS codes as cyclic codes**
- This description uses an association between vectors $(v_0, v_1, \ldots, v_{n-1})$ and polynomials $v_0+v_1 \cdot x+v_2 \cdot x^2+\ldots+v_{n-1} \cdot x^{n-1}$ [1].
- If a code word $c \in C$ and τ is the cyclic shift operator, then $\tau(c) \in C$.
- The code word can be expressed as:

$$RS_C(k) = \left\{ (c_0, c_1, \ldots, c_{n-1}) : \sum_{j=0}^{n-1} c_j x^j = 0 \text{ for } \alpha, \alpha^2, \ldots, \alpha^{n-k} \right\} \quad (5)$$

Classical decoding algorithms

•Are based on the ***Key Equation*** [2] $\sigma(x)\cdot S(x)=\omega(x)\left(\mathrm{mod}\ x^{r}\right);\ r=n-k$ (6)

– $\sigma(x)$ denotes the error locator polynomial, $\omega(x)$ the error evaluator polynomial, and $S(x)$, the syndrome polynomial.

•If $c=(c_1, c_2, \ldots, c_n)$ is the transmitted code word, $p=(p_1, p_2, \ldots, p_n)$ the received code word and $e=(e_1, e_2, \ldots, e_n)$ is the error vector ($p=c+e$) we have the following relations [2]:

$$S_p(x)=\sum_{i=1}^{n}\frac{p_i}{1-\alpha^{i-1}\cdot x}\left(\mathrm{mod}\ x^{r}\right) \quad (7)$$

$$S(x)=S_p(x)=\left(S_c(x)+S_e(x)\right)\left(\mathrm{mod}\ x^{r}\right)=S_e(x)\left(\mathrm{mod}\ x^{r}\right) \quad (8)$$

$$S_e(x)=\sum_{b\in B}\frac{e_i}{1-\alpha^{b-1}\cdot x}\left(\mathrm{mod}\ x^{r}\right);\ B=\{i\in(1,n)\,|\,e_i\neq 0\} \quad (9)$$

$$\omega(x)=\sum_{b\in B}e_b\cdot\left(\prod_{a\in B,a\neq b}\left(1-\alpha^{a-1}\cdot x\right)\right) \quad (10) \qquad \sigma(x)=\prod_{b\in B}\left(1-\alpha^{b-1}\cdot x\right) \quad (11)$$

$$e_b=\frac{-\alpha^{b-1}\cdot\omega\left(\alpha^{-(b-1)}\right)}{\sigma'\left(\alpha^{-(b-1)}\right)};\ B=\left\{b\in(1,n)\,|\,\sigma\left(\alpha^{-(b-1)}\right)=0\right\} \quad (12)$$

- Complete performance evaluation of the GS algorithm requires the comparison with a classic RS decoding algorithm.
- A variant of classic algorithm based on the Berlekamp-Massey computation of the error locator polynomial was used for comparison with the GS algorithm.

Data Transmission Laboratory

The Berlekamp-Massey algorithm

- The classical RS decoding algorithm used for comparison with GS algorithm has some modifications that allows a more efficient implementation [1].

- **The steps of the used algorithm:**
  - Syndrome vector computation by s=w·H ; w – the received word ; the H check matrix and the G generator matrix of the considered systematic codes are :

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \alpha & \alpha^2 & \alpha^4 & \dots & \alpha^{n-k-1} \\ 1 & \alpha^2 & \alpha^4 & \alpha^8 & \dots & \alpha^{2n-2k-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \alpha^{n-1} & \alpha^{n-2} & \alpha^{n-3} & \dots & \alpha^{k+1} \end{bmatrix} \quad (13)$$

$$G(x) = \prod_{i=0}^{n-k-1} \left(x - \alpha^i\right) \quad (14)$$

  - If w is interpreted as a polynomial, then the product s=w·H is the row vector {w(1), w(α), w(α²),…, w(α^{k-1})} ; each polynomial can be easy computed using Horner's method.
  - The Berlekamp-Massey algorithm generates the error locator polynomial, σ(x), and the error evaluator polynomial, ω(x), of the error vector e.
  - The error locations, e_{li}, are computed as :

$$e_{li} = \begin{cases} 1 \text{ if } \sigma\left(\alpha^i\right) = 0 \\ 0 \text{ otherwise} \end{cases} \quad (15)$$

  - The error values, e_{vi}, are computed as : if $e_{li} \neq 0$ then $e_{vi} = -\left(\sigma'\left(\alpha^i\right) \cdot \omega\left(\alpha^i\right)\right)^{-1}$ (16)

Data Transmission Laboratory

Formulations of the decoding problem

- Let be [n, k, d] a generalized RS code defined over Galois Field F_q (1)-(2) ; n is the number of symbols of the code word, k is the number of information symbols of the code word and d is the code distance, d=n-k+1.

- **The τ-error correction problem for an [n, k, d] code** [5] is defined for $2\tau+1 \le d$ as follows : given a string s with length n symbols find a codeword c which is within Hamming distance τ of s, if one such exists – if the answer exists is unique.

- **The maximum-likelihood decoding problem** [5] is set in a model where a larger number of errors is considered less likely and is defined as follows : given a string s with length n symbols find a codeword c which is nearest to s considering Hamming distance – the solution may not be unique ; problems with the solution selection if the answer is not unique – referred sometimes as the nearest codeword problem.

- **The τ-reconstruction problem** [5]: given a string s with length n symbols find all code words c that are within Hamming distance τ of s – referred also as the list decoding problem.

- **General formulation of the list decoding problem** [4] :
  - Input : a field F_q; n distinct pairs of elements $\{(x_i , y_i)\}_{i=1}^{n}$ from $F_q \times F_q$; integers h and t
  - Output : a list of al functions $f: F_q \to F_q$ satisfying : f(x) is a polynomial of degree at most h with

$$\left|\{i \mid f(x_i) = y_i\}\right| \ge t \qquad (17)$$

  - h=k-1 for generalized RS codes and the number of corrected errors is n-t.

Data Transmission Laboratory

Guruswami-Sudan decoding algorithm

- **GS-I decoding algorithm** [4] [8]
- Inputs : n, h=k-1, t; $\{(x_1, y_1), \ldots (x_n, y_n)\}$

  1. find any function $Q:F_q^2 \rightarrow F_q$ satisfying :
     - Q(x,y) has (1,d)-weigted degree at most m+lh ; l, m parameters
     - $\forall i \in \{1, 2, \ldots, n\}$, $Q(x_i, y_i)=0$
     - Q is not identical zero

  2. Factor the polynomial Q into irreducible factors

  3. Output all the polynomials f such that (y-f(x)) is a factor of Q and $f(x_i)=y_i$ for at least t values of i from $\{1, 2, \ldots, n\}$

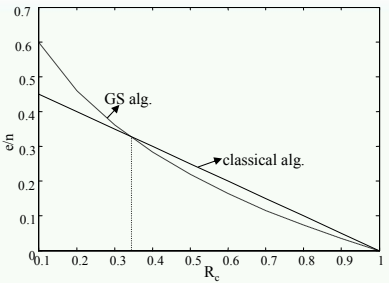


Fig. 1 Correction rate e/n function of the code rate $R_c=k/n$ for classical and GS-I algorithm

- Optimum values for m and l : $l=\sqrt{\frac{2\cdot(n+1)}{k-1}}-1 \quad ; \quad m \geq \frac{k-1}{2}-1$ (18)
- An algorithm which runs in time polynomial in n exists if : $t \geq (k-1)\cdot\left[\sqrt{\frac{2\cdot(n+1)}{k-1}}\right]-\left[\frac{k-1}{2}\right]$ (19)
- The $e/n=\tau_{max}/n$ ratio for classical and GS decoding algorithms are given by :

$$\frac{e}{n_{|classical}}=\frac{1}{2}\cdot(1-R_c) \quad ; \quad \frac{e}{n_{|GS-I}} \cong 1-\sqrt{2+\frac{R_c}{4}}\cdot\sqrt{R_c}+\frac{R_c}{2} \qquad (20)$$

- **Conclusion :** GS-I algorithm is worth using only for $R_c \leq 1/3$.

Guruswami-Sudan decoding algorithm

- **GS-II decoding algorithm** [5] [6] [8]
- Inputs : n, k, t; $\{(x_1, y_1), \ldots (x_n, y_n)\}$; $x_i, y_i \in F$
- Parameters : m, l : $m = 1 + \left\lfloor \frac{k \cdot n + \sqrt{k^2 \cdot n^2 + 4 \cdot \left(t^2 - k \cdot n\right)}}{2 \cdot \left(t^2 - k \cdot n\right)} \right\rfloor$; $l = m \cdot t - 1$ (21)

1. find a polynomial Q(x,y) such that (1, k)-weighted degree≤l – find values for its coefficients $\{q_{j1j2}\}_{j1,j2 \geq 0: j1 + k \cdot j2 \leq l}$ such that the following conditions hold:
   - At least one $q_{j1,j2}$ is non-zero.
   - For every $i \in \{1,\ldots,n\}$, if $Q^{(i)}$ is the shift of Q to (x_i , y_i), then all coefficients of $Q^{(i)}$ of total degree less than m are 0 – specifically [5]:

$$\forall i \in \{1,\ldots,n\}, \forall j_1, j_2 \geq 0,\ j_1 + j_2 < m,$$
$$q_{j_1 j_2}^{(i)} = \sum_{j_1' \geq j_1} \sum_{j_2' \geq j_2} \binom{j_1'}{j_1} \cdot \binom{j_2'}{j_2} \cdot q_{j_1', j_2'} \cdot x_i^{j_1' - j_1} \cdot y_i^{j_2' - j_2} = 0 \quad (22)$$

2. Find all polynomial $f(x) \in F_q[x]$ of degree at most k-1 such that (y-p(x)) is a factor of Q(x,y) ; for each polynomial f(x) check if $f(x_i)=y_i$ for at least t values of $i \in \{1,\ldots,n\}$; if so, include f(x) in the output list.

Data Transmission Laboratory

Guruswami-Sudan decoding algorithm

- **Conditions for GS-II decoding algorithm** [5] [6]
- If $n \cdot \binom{m+1}{2} < \frac{1 \cdot (l+2)}{2k}$ (23) then a polynomial Q(x,y), as presented above, does exist and can be found in polynomial time by solving a linear system.
- If n, k, t satisfy: $t^2 > k \cdot n$ (24) then for m and l given by (21) relation (23) holds
- Relation (21) can be rewritten as:

$$m = 1 + \left\lceil \frac{R_c + \sqrt{R_c^2 + 4 \cdot \left(\frac{t^2}{n^2} - \frac{R_c}{n^2} \right)}}{2 \cdot \left(\frac{t^2}{n^2} - R_c \right)} \right\rceil \cong 1 + \left\lceil \frac{R_c}{\left(1 - \frac{\tau}{n} \right)^2 - R_c} \right\rceil \quad (25)$$

- From relation (24) results the $e/n = \tau_{max}/n$ ratio :

$$\frac{e}{n}_{|GS-II} < 1 - \sqrt{R_c} \quad (26)$$

- The same ratio for I GS algorithm is [4] [5] :

$$\frac{e}{n}_{|GS-I} < 1 - \sqrt{2R_c} \quad (27)$$

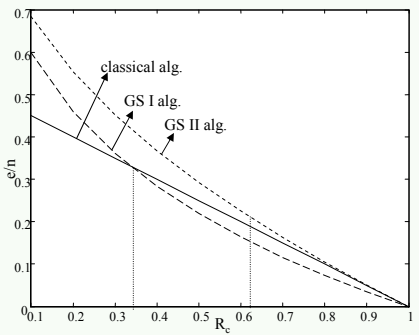


Fig. 2 Correction rate e/n vs. the code rate R_c=k/n for the classical, GS-I and GS-II algorithms

Data Transmission Laboratory

Guruswami-Sudan decoding algorithm

- **Conclusion :** II GS decoding algorithm is worth to be used for coding rates $R_c<0.65$.

- zero-multiplicities, m, computed according to relation (25) exhibit very high values ; algorithms employing such values of parameter m are difficult cu implement and require very long decoding times.

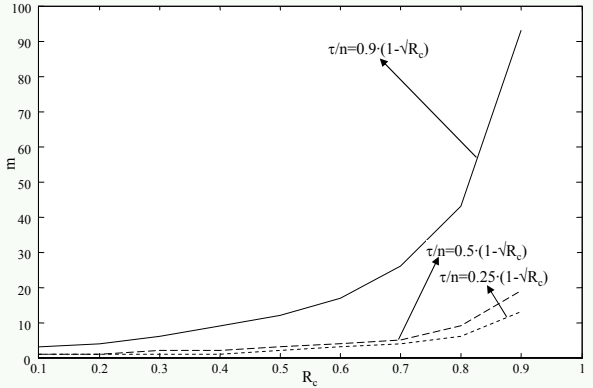


Fig. 3 The (computed) zero multiplicities, m, of the GS-II algorithm function of the coding rate, R_c, for different correction rates τ/n

Data Transmission Laboratory

Guruswami-Sudan decoding algorithm

- **Another approach for GS-II decoding algorithm** [3]

- **The Interpolation Theorem.** [3] Let $\{m(\alpha,\beta);\ (\alpha,\beta)\in F^2\}$ be a multiplicity function and let $\phi_0 < \phi_1 < \ldots < \phi_n$ be an arbitrary monomial order. Then, there exists a nonzero polynomial Q(x,y) of the form :

$$Q(x,y)=\sum_{i=0}^{C} a_i\cdot\phi_i(x,y) \quad \text{where} \quad C=\sum_{\alpha,\beta}\binom{m(\alpha,\beta)+1}{2} \tag{28}$$

  which has a zero multiplicity $m(\alpha,\beta)$, at $(x,y)=(\alpha,\beta)$, for all $(\alpha,\beta)\in F^2$.

- **Definition :** [3] If $Q(x,y)\in F[x,y]$, and $f(x)\in F[x]$, define the Q-score of f(x) as :

$$S_Q(f)=\sum_{\alpha\in F}\mathrm{ord}(Q:\alpha,f(\alpha)) \tag{29}$$

- **The Factorization Theorem,** [3]: Suppose $f(x)\in F_v[x]$, $Q(x,y)\in F[x,y]$ and $S_Q(f)>\deg_{1,v}Q$. Then y-f(x) is a factor of Q(x,y).

- Limits for the decoding radius, r_m, for II GS algorithm [3] :

$$n-\left\lfloor\sqrt{n\cdot(k-1)\cdot\frac{m+1}{m}}\right\rfloor=r_{min}\le r_m\le n-1-\left\lfloor\sqrt{n\cdot(k-1)\cdot\frac{m+1}{m}}-\frac{k-1}{2m}\right\rfloor=r_{max} \tag{30}$$

- The average size, L_m, of the decoding list [3]:

$$L_m=\left\lfloor\sqrt{\frac{n}{v}\cdot m\cdot(m+1)+\left(\frac{k+1}{2\cdot(k-1)}\right)^2}-\frac{k+1}{2\cdot(k-1)}\right\rfloor<\left(m+\frac{1}{2}\right)\cdot\sqrt{\frac{n}{k-1}} \tag{31}$$

Data Transmission Laboratory

Guruswami-Sudan decoding algorithm

- **Interpolation and factorization algorithms of two variable polynomials**
- **The Interpolation problem :** construct a bivariate polynomial Q(x,y) with minimal (1,v) degree, which interpolates the points $\{(x_1, y_1), \ldots (x_n, y_n)\}$; $x_i, y_i \in F$ and has an imposed $m(x_i,y_i)$ order zero--multiplicity in every point ; the m-th order zero multiplicity in point (x_i,y_i) is defined in (22).
- One of the best solution for interpolation : **Kotter Interpolation Algorithm** [3]:

input data: L – number of code words in the list, $(\alpha_i\ \beta_i)_{i=1}^{n}$ – interpolation points, $(m_i)_{i=1}^{n}$ – zero's multi-plicity order, (1,k-1) – monomials weighted degree.

1. FOR $j=0$ to L
$g_j=y_j$
2. FOR $i=1$ to n DO
2. FOR $(r,s)=(0,0)$ to $(m_i-1,0)$ DO /* lex order
4. . FOR $j=0$ to L DO
5. $\Delta_j=D_{r,s}g_j(\alpha_i,\beta_j)$
6. $J=\{j: \Delta_j \neq 0\}$
7. IF $J\neq\Phi$
8. $j^*=min\_rank\ \{g_j: j\in J\}$
9. $f=g_{j^*}$; $\Delta=\Delta_{j^*}$
10. FOR $j\in J$ DO
11. IF $(j\neq j^*)$
12. $g_j=\Delta\cdot g_j+\Delta_j f$
13. ELSE IF $(j=j^*)$
14. $g_j=\Delta\cdot(x+\alpha_i)\ f$
15. $Q_0(x,y)=min\_rank\{g_j(x,y)\}$ /* the interpolation polynomial

- One of the best solution for factorization : **Roth-Ruckenstein Algorithm** [3]: - takes as input a bivariate polynomial Q(x,y) and a positive integer D, and returns as output the set of all (y-f(x)) roots of Q(x,y) with the degree of $f(x)\leq D$.

Data Transmission Laboratory

The goals of the study

- The main goal of this study is to compare, by computer simulations, the correction capability and required processing time of the GS and BM (representative for the classical algorithms) RS decoding algorithms.
- The analysis is intended to establish the optimum values of the zero order multiplicity of the GS algorithm for which a maximum ratio correction capability/decoding time is accomplished, and to elaborate some "thumb rules" for adapting these parameters, so that shorter decoding times would be attained.
- Study conditions : interpolation with constant zero-multiplicity; hard decoding; packet errors with imposed characteristics.
- Comparison between the performances of the GS and BM decoding algorithms is made considering the ratio, R_d, between the number of error code words after the GS decoding and the number of code words which have more than t_b errors before the GS decoding.
  - t_b is number of errors that could be corrected by classical decoding algorithms for a given code.

Data Transmission Laboratory

The implemented software simulator

- The implemented software simulator can operate in the Galois fields GF(2^3), GF(2^4), GF(2^5), GF(2^6) and GF(2^8) and performs the following functions:
  - o generation of a symbol-sequence represented on the number of bits corresponding to the employed Galois field.
  - o RS encoding (cyclic code for the BM or evaluation code for the GS), depending on the decoding algorithm employed.
  - o serialization of the coded bits.
  - o generation of the packet-errors.
  - o superposing the errors on the coded data flow.
  - o GS or BM decoding.
  - o computation of parameters of the simulated transmission, namely:
    - bit and symbol error rates.
    - the ratio of the correction capability of the GS algorithm versus the correction capability of the BM algorithm.
    - the numbers of words in the decoding list and erasures, both for the GS algorithm.

Data Transmission Laboratory

The implemented software simulator

- The generation of the packet-errors is performed according to the impulse noise models employed for the xDSL transmissions [9].

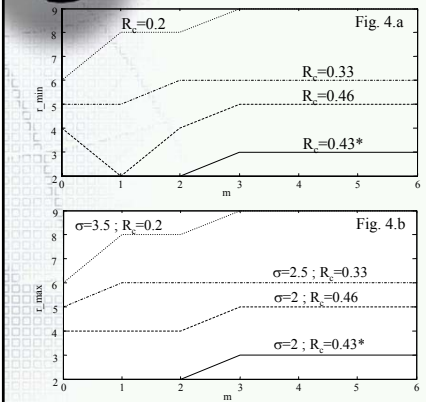
  - model representative for transmission systems employing RS codes as outer codes, adopted with several simplifications.

- The main features of algorithm that generates the packet-errors are:

  - the distance in symbols between two packet-errors has a Poisson distribution [9], with a modifiable average value λ ; in the simulations performed, the value of λ equaled the number of symbols of two code words, for each GF.
  - the packet-error length, in bits, has a Gaussian distribution [9], defined by the average value τ and variance σ.
    - the value of τ equaled $t_b \cdot g$, t_b denoting the number of error-symbols that could be corrected by the classical decoding algorithms (e.g. BM) and g denoting the number of bits/character of the GF employed.
    - the value of σ was set according to the estimated correction capability of the GS algorithm.
  - the positions of the errors inside the packet are random, being distributed according to a uniform law.

DTL Data Transmission Laboratory

Computer simulation results. Conclusions



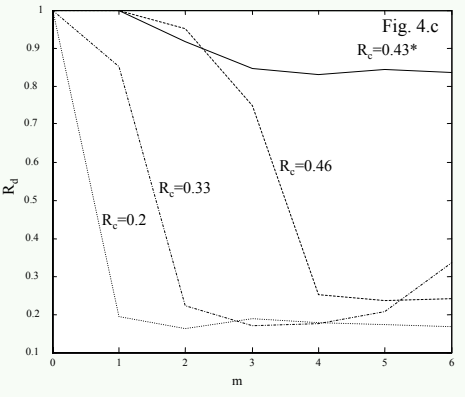
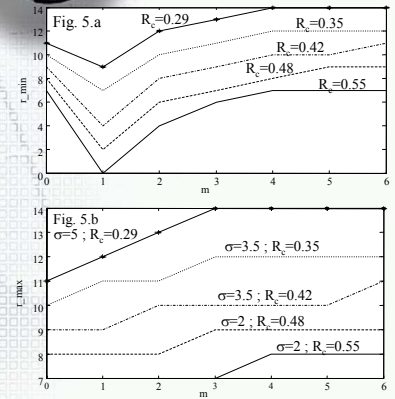


Fig.4 Minimum, r_{min} (4.a), maximum decoding radius r_{max} (4.b), correction rate R_d (4.c) in terms of m ; RS codes in Galois GF(2^3) and GF(2^4) ; * denotes codes defined in GF(2^3);

- for RS codes defined over GF(2^3) and GF(2^4) $r_{max}=r_{min}$, at optimum values of parameter m.
- for RS codes defined in GF(2^3) and GF(2^4), m has to be set to 3 or 4, for R_c close to 0.5, and to 1 or 2 for R_c close (or smaller) than 0.3.
- the increase of m above a certain limit does not bring a performance improvement, but it might lead to a decrease of performances (see R_c = 0.33).

Data Transmission Laboratory

Computer simulation results. Conclusions



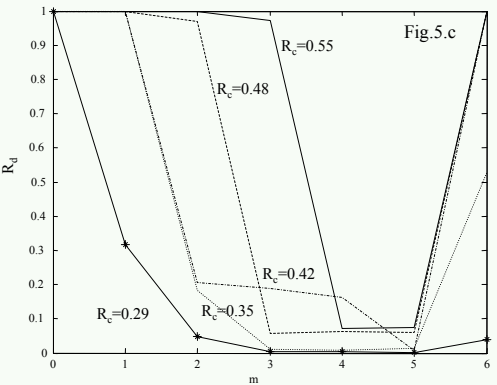
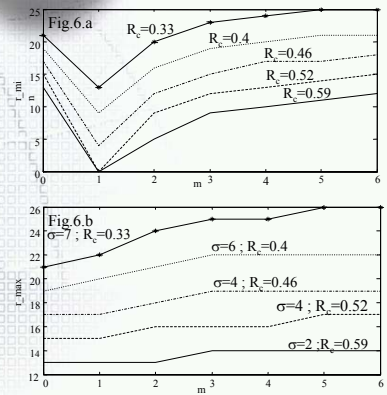


Fig.5 Minimum, r_{min} (5.a), maximum decoding radius r_{max} (5.b), correction rate R_d (5.c) in terms of m ; RS codes in Galois GF(2^5);

- for RS codes defined in GF(2^5) having the mentioned R_c , $r_{max} = r_{min}+1$ or $r_{min}+2$ at optimum values of parameter m.
- the optimum values of m are m = 3 - 4 for R_c close to 0.5, m = 2 - 3 for R_c around 0.3 and m = 4 - 5 for R_c around 0.4.
- there should be noticed that for m=6, the performances of the GS decoder exhibit a significant decrease, especially for high values of the coding rate R_c.

Data Transmission Laboratory

Computer simulation results. Conclusions



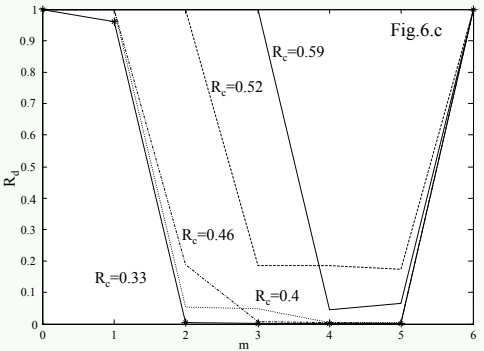
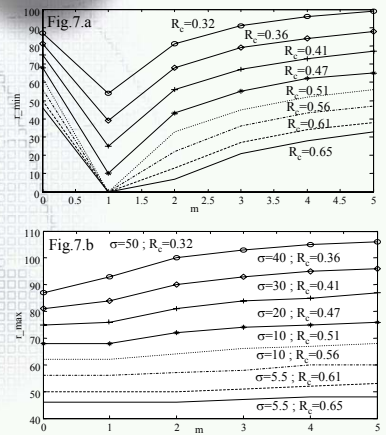


Fig.6 Minimum, r_{min} (6.a), maximum decoding radius r_{max} (6.b), correction rate R_d (6.c) in terms of m ; RS codes in Galois GF(2^6);

- for codes defined in GF(2^6) at optimal values of m the difference r_{max} - r_{min} takes values between 3 and 6 and the difference r_{max} - t_b takes values between 0 and 3.
- the optimum values of m exhibit a clear separation in terms of the coding rate R_c; for $R_c \geq 0.5$, optimum m equals 3 or 4, but for $Rc \leq 0.45$, optimum m equals 2 or 3.
- the codes defined in GF(2^6) exhibit the same decrease of performance for higher values of m (e.g. m=6), as the ones defined in GF(2^5); for the considered values of R_c, the performances secured by the GS become equal to the ones of the BM.

DTL Data Transmission Laboratory

Computer simulation results. Conclusions



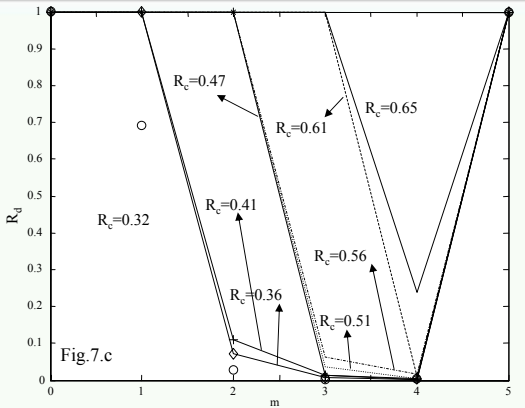


Fig.7 Minimum, r_{min} (7.a), maximum decoding radius r_{max} (7.b), correction rate R_d (7.c) in terms of m ; RS codes in Galois GF(2^8).

- for codes defined in GF(2^8) at optimal values of m the difference r_{max} - r_{min} takes values higher or equal with 20 and the difference r_{max} - t_b takes values between 2 and 13.
- for coding rates higher or equal to 0.6 the optimum value of m is 4, for $R_c \in (0.6, 0.45)$ the optimum value of m is 3, and for coding rates ranging between 0.3 and 0.45, the optimal m equals 2.
- the maximum limit of m decreases to 4 for the coding rates considered.

Evaluation of the decoding time. Conclusions

- The evaluation of the decoding time implies:

  o the measurement, for a certain number of code words, of the simulation time t_{sim}.

  o the measurement of the time required for encoding and error-pattern insertion t_{aux}.

- The ratio between the average decoding times, t_{dec}, of the two algorithms is expressed by:

$$t_d = \frac{t_{decGS}}{t_{decBM}} = \frac{t_{simGS} - t_{auxGS}}{t_{simBM} - t_{auxBM}} \approx \frac{t_{simGS}}{t_{simBM}} \quad (32)$$

- The decoding time required by the GS algorithm is much larger than the one required by the BM algorithm.

- The t_d ratio increases significantly with the increase of m and with the increase of the dimension of the Galois field employed.

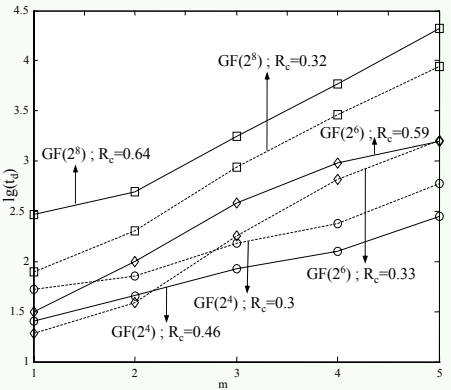


Fig.8 $\lg(t_d)$ ratio between the average decoding time of the GS and BM algorithms, in terms of m, for various coding rates and for RS codes defined in several Galois fields

- The increase of the coding rate for codes defined over GF higher than $GF(2^4)$ also increases the value of t_d ratio - changing the coding rate from 0.3 to 0.6 for these codes the t_d ratio increases by a factor of 2 or 3.

- The results displayed in fig. 8 underline the importance of establishing optimal values for the parameter m and the necessity of finding some GS decoding strategies with a decoding time as small as possible.

Data Transmission Laboratory

Conclusions

- the optimum values of the zero multiplicity m of GS algorithm can't be obtained only considering the maximum, r_{max}, and minimum, r_{min}, decoding radius (especially for codes defined over GF(2^n) n>5) – however this decoding radius gives an idea about the correction capability of the GS algorithm.
  - for RS codes defined over GF(2^3), GF(2^4), GF(2^5) $r_{max}=r_{min}$, $r_{max}=r_{min}+1$ or $r_{max}=r_{min}+2$ at optimum values of parameter m.
  - for RS codes defined over GF(2^6), GF(2^8) and for low values of parameter m the computed r_{min} (30)<t_b (in reality $r_{min}>t_b$) ; at optimum values of parameter m the computed r_{min} (30)>t_b and the difference $d_r=r_{max}-r_{min}$ increases ; $d_r \in [3,6]$ for codes defined over GF (2^6) and $d_r \in [2,13]$ for codes defined over GF (2^8).
- the maximum coding rate where the GS algorithm with constant zero multiplicity and relatively low m parameter has better performances than classical decoding algorithms is about 0.45 - 0.55 for codes defined over GF(2^3), GF(2^4), GF(2^5) and 0.6 – 0.65 for codes defined over GF(2^6), GF(2^8).
- the performance loss exhibited by the GS algorithm for high values of m, regardless the coding rate, could be explained by the fact that are not fulfilled the conditions required for factorization (see (29) and the factorization theorem) ; these performance loss might be eliminated by the use of other interpolation and factorization algorithms.

Data Transmission Laboratory

Conclusions

- a primary analysis of the used interpolation algorithm (Koetter algorithm) and the properties of the two-variable polynomials [3] leads to the following:

o the number of iterations, n_{it}, performed by the interpolation algorithm for n-symbol code words and multiplicity order of zeros equaling m, is :

$$n_{it} = n \cdot \frac{m \cdot (m+1)}{2} \tag{33}$$

o the initial polynomials of Koetter interpolation algorithm, for maximum L words in the final decoding list, are: $p_0(x,y) = 1, p_1(x,y) = y, p_2(x,y) = y^2, \ldots, p_2(x,y) = y^L$ (34)

o supposing that the values of Δ never equal zero then after L·(L+1)/2·(k-1) iterations all polynomials will have the same degree L·(k-1) and the leading monomials of these polynomials are:

$$lp_0(x,y) = x^{L\cdot(k-1)}, lp_1(x,y) = x^{(L-1)\cdot(k-1)} \cdot y, lp_2(x,y) = x^{(L-2)\cdot(k-1)} \cdot y^2, \ldots, lp_L(x,y) = y^L \tag{35}$$

o the increase of the degree of each polynomial will require L+1 iterations ; by the end of the algorithm the degree, deg_{min}, of the minimum-degree polynomial would be:

$$\deg_{min} = \left| \frac{n_{it} - L \cdot (L+1)/2 \cdot (k-1)}{L+1} + L \cdot (k-1) \right| \tag{36}$$

o the minimum value of the S_Q parameter (29) of an interpolation polynomial associated to a n-symbol code word and to a multiplicity order of zeros equaling m and to a decoding radius r, is:

$$S_{Q\,min} = (n-r) \cdot m \tag{37}$$

Conclusions

- from the factorization requirements we have:

$$\frac{S_{Q\,min}}{deg_{\,min}} \cong \frac{\left(1-\frac{r}{n}\right)\cdot\frac{(L+1)}{2}}{(m+1)+\frac{L\cdot(L+1)\cdot R}{m}} > 1 \qquad (38)$$

- the values of the ratio defined in (38), for the considered codes and for various values of m, are smaller than 1 (approximately equal, but smaller);
  - there should be noted that the evolution of polynomials degrees within the interpolation algorithm would be different, mostly because of the fact that Δ might equal zero quite often decreasing significantly the values of deg_{min}.
  - the value of S_Q might be higher than the value computed by (37).
  - the considerations above show that, for some interpolation algorithms, for different coding rates and various values of m, there is a possibility that the GS algorithm would not be effective, even for a high decoding radius.
- the performed simulations show the following aspects regarding the number of words in the decoding list:
  - for the RS codes defined in $GF(2^5)$ and in the higher fields, the number of the words in the list equals 1, with very few exceptions.
  - for the codes defined in $GF(2^3)$ and $GF(2^4)$, there are more cases when the decoding list contains more than one code word, but their percentage is still small, about 1%.
- **general conclusion** : if the GS decoding algorithm can not correct a code word, this fact is owned to an unsuccessful interpolation or factorization and, quite seldom, to the presence of more than one code words in the decoding list.

Conclusions

- **Decoding strategies for GS algorithm**

▪ Employing the same value of m for the decoding of every code word ; this variant would require a very large average decoding time, even for the correctly received code words.

▪ The successive increase of the value of m, from 1 to a maximum optimal value ; the decoding is stopped when the decoding list contains at least one code word.

o this approach would require a smaller average decoding time for packet-errors with relatively small lengths, compared to the maximum packet length for which a successful GS decoding is accomplished.

▪ The employment of two values for parameter m, namely 1 and an optimum value m_{opt} ; the correct code words and the ones affected by a small number of errors would be decoded using m=1, and the code words with more errors would be decoded with $m = m_{opt}$.

o this last option should be employed if the decoding with m=1 generates no code word in the decoding list.

o this variant provides a smaller average decoding time for long error-packets, compared the maximum packet length for which a successful GS decoding is accomplished, see results in fig.8.

Data Transmission Laboratory

References

[1] M.E.O'Sullivan, *Coding Theory*, http://wwwroham.sdsu.edu/~mosulliv/Courses/coding02.html.

[2] J.I.Hall, *Notes on Coding Theory*, http://www.mth.msu.edu/~jhall/classes/codenotes/coding-notes.html.

[3] R.J. McEliece, "The Guruswami-Sudan Decoding Algorithm for Reed-Solomon Codes", *IPN Progress Report 42-153,*15 May, 2003.

[4] M. Sudan, "Decoding of Reed Solomon codes beyond the error-correcting bound*", Journal of Complexity*, vol. 13, 1997.

[5] V. Guruswami, M. Sudan, "Improved Decoding of Reed-Solomon Codes and Algebraic Geometry Codes", *IEEE Trans. Inform. Theory*, vol. 45, no. 6, September 1999.

[6] V. Guruswami, M. Sudan, "Reflections on Improved Decoding of Reed-Solomon Codes and Algebraic Geometry Codes", http://citeseer.ist.psu.edu/guruswami02reflections.html

[7] W. Gross, Fr. Kschischang, R. Koetter, P.G. Gulak, "Applications of Soft-Decision Decoding of Reed-Solomon Codes", submitted to *IEEE Trans. Comm*. July 2003, http://www.macs.ece.mcgill.ca/~wjgross/papers/gkkg_tc.pdf.

[8] M. K. Cheng, P. H. Siegel, X. Wu, "List Decoding Reed Solomon Codes", NSIC Annual Meeting, June 28, 2000.

[9] W. Henkel, T. Kessler, "A wideband impulsive noise survey in the German telephone network-Statistical description and modeling", AEU, vol. 48, no. 6, Nov./Dec. 1994.