M a t h e m a t i c a
B a l k a n i c a

New Series Vol. 21, 2007, Fasc. 1-2

Bounds on the Size of Equidistant Codes over an Alphabet of Five and Six Elements [1]

Galina Bogdanova, Todor Todorov

Presented by St. Dodunekov

In this paper we consider the problem of finding upper and lower bounds and exact values on the size of equidistant codes over an alphabet of five and six elements. We present a computer realization of an algorithm for solving the maximum clique problem. The exact values of the maximum size of equidistant codes over an alphabet of 5 and 6 elements for $n \leq 10$ are given.

Key Words: Equidistant codes, Bounds on codes, Computer search, Maximum clique problem

1. Error-correcting code basics

The theory of error-detecting and error-correcting codes is a branch of engineering and mathematics which deals with the reliable transmission and storage of data. Information media are not 100% reliable in practice, in the sense that noise (any form of interference) frequently causes data to be distorted. To deal with this undesirable but inevitable situation, some form of redundancy is incorporated in the original data. With this redundancy, even if errors are introduced (up to some tolerance level), the original information can be recovered, or at least the presence of errors can be detected. Adding to the original message the parity bit or the arithmetic sum allows the detection of a (certain type of) error for example. However, that kind of redundancy doesn't allow for the correction of the error. Error-correcting codes do exactly this: they add redundancy to the original message in such a way that it is possible for the receiver to detect the error and correct it, recovering the original message. This is crucial for certain applications where the re-sending of the message is not

[1]Supported partially by the Bulgarian National Science Fund under Grant IO-03-02/2006

possible (for example, for interplanetary communications and storage of data). The crucial problem to be resolved then is how to add this redundancy in order to detect and correct as many errors as possible in the most efficient way.

The object of an error-correcting code is to encode the data, by adding a certain amount of redundancy to the message, so that the original message can be recovered if not too many errors have occurred.
Let Z_q^n be the set of all n-tuples over $Z_q = \{0, 1, 2, ..., q-1\}$. A q-ary code of length n is a non empty subset of Z_q^n.
The set Z_q is called the alphabet. If q is a prime power we often take the alphabet Z_q to be the finite field of order q.

Definition 1. The Hamming weight of a vectors x of Z_q^n is the number of non-zero entries of x. It is denoted by $w(x)$.

Definition 2. The Hamming distance between two vectors x and y of Z_q^n is the number of places in which they differ. It is denoted by $d(x, y)$.

Definition 3. The minimum distance of a code C is the smallest of the distances between distinct codewords. It is denoted by $d(C)$.

$$d(C) = \min\{d(x, y)|x, y \in C, x \neq y\}$$

Definition 4. An (n, M, d)-code is a code of length n, containing M codewords and having minimum distance d.

Definition 5. We call an (n, M, d)-code optimal if for fixed n, d it has the largest possible M.

2. Introduction to equidistant codes

A code is called *equidistant* if all the distances between distinct codewords are equal to d. An $E_q(n, M, d)$ equidistant code (EC) is a code over Z_q of length n, cardinality M and distance d. Let $B_q(n, d) = max\{M : \exists E_q(n, M, d)\}$ denote the largest possible value of M when the other parameters are fixed. Codes with parameters $(n, B_q(n, d), d)$ are called optimal. An $E_q(n, M, d, w)$ code is called *equidistant constant weight* code (ECWC) if all its codewords have the same weight equal to w. Let $B_q(n, d, w)$ denote the largest possible value of M in an ECWC when the other parameters are fixed. Codes with parameters $(n, B_q(n, d, w), d, w)$ are called optimal.

One of the main open problem of the algebraic and combinatorial coding theory is the construction of optimal equidistant codes. Some works published on this topic are [8],[9],[14],[16],[17], [18],[7],[2],[3],[4],[5],[11], etc.

In the present paper we consider the problem of finding upper and lower bounds on the size of equidistant codes for $q = 5$ and $q = 6$. We use both combinatorial and computer methods for their construction. The best known upper and lower bounds for q-ary EC are presented in Section 3. To find equidistant codes we can apply computer search for the cases that we were not able to solve using the combinatorial constructions. Computer methods and new results for equidistant codes over alphabet of 5 and 6 elements are given in Section 4.

3. Preliminaries

Some bounds for EC and ECWC are given by the following theorems:

Theorem 1. *(Trivial values) The maximum number of codewords in a q-ary EC and ECWC satisfy the inequalities:*

$B_q(n, n) = q$

$B_q(n, n, w) \leq q$

$B_q(n, d, n) = B_{q-1}(n, d)$

$B_q(n+1, d, w) \geq B_q(n, d, w)$

$B_q(n+1, d, w+1) \geq B_q(n, d, w)$.

The EC and ECWC are closely related as shown by following theorem:

Theorem 2. [7] *It is true that* $B_q(n, d) = 1 + B_q(n, d, d)$.

Theorem 3. [6] $B_q(n, d) \leq (q-1)n + 1$.

It is easy to prove the Johnson bounds [12] for ECWC:

Theorem 4. *The maximum number of codewords in a q-ary ECWC satisfy the inequalities:*

$$B_q(n, d, w) \leq \frac{n}{n-w} B_q(n-1, d, w)$$

$$B_q(n, d, w) \leq \frac{n(q-1)}{w} B_q(n-1, d, w-1).$$

Theorem 5. *[7] For $k = 1, 2, ..., n$, if* $P_k^2(w) > P_k(d) P_k(0)$, *then*

$$B_q(n, d, w) \leq \frac{P_k^2(0) - P_k(d) P_k(0)}{P_k^2(w) - P_k(d) P_k(0)}.$$

Here $P_k(x)$ is the Krawtchouk polynomial defined by

$$P_k(x) = \sum_{i=0}^{k} \binom{x}{i} \binom{n-x}{k-i} (-1)^i (q-1)^{k-i}$$

and

$$P_k(0) = \binom{n}{k} (q-1)^k .$$

Theorem 6. *[11]* $B_q(q+1, q, q-1) \leq (q^2+q)/2$.

Theorem 7. *[16] The optimal equidistant $E_q(n, qt, d)$ codes and resolvable balanced incomplete block designs(RBIBD) $(v = qk, b, k, r, \lambda)$ are equivalent to one another and their parameters are connected by the conditions $v = M$, $b = nq$, $k = t$, $r = n$, $\lambda = n - d$.*

4. Computer Search and Results

The main problem we are solving in this paper is the code construction of some $E_q(n, M, d)$ EC. This is a maximal clique problem. A fast algorithm for solving this problem is given in [15]. The variant of this method is considered in [19] (for constant weight composition codes). In this work we present a variant of these algorithms for EC and ECWC.

A simple graph $G = (V; E)$ is a set of vertices V and set of unordered pairs of distinct elements of V called edges.

A clique of a graph is a set of vertices, any two of which are adjacent. Maximal clique is a clique which vertices are not a subset of the vertices of a larger clique. Maximum clique is the largest clique in the graph. It this paper we are interested of a maximum clique in a graph.

Finding the maximum clique in a graph is an NP-hard problem − no polynomial time algorithms [1, 10, 13]

We assume some order for the vertices $V = v_1, v_2, ..., v_n$. Let $S_i = \{v_1, v_2, ..., v_i\} \subseteq V$. We define the function $c(i)$ to be the size of the maximum clique in the subgraph induced by S_i. Obviously, for every $i = 1, ..., n-1$ we have either $c(i+1) = c(i)$ or $c(i+1) = c(i)+1$. Moreover, $c(i+1) = c(i)+1$ iff there exists a clique in S_{i+1} of size $c(i)+1$ that includes vertex v_{i+1}.

Then we calculate the values of $c(i)$ starting from $c(1) = 1$ up, and stores the values found. The algorithm then is searching for a clique of size $c(i)+1$ within S_{i+1}, it has formed a clique W and is considering adding vertex v_j, when it can prune the search if $\mid W \mid + c(j) \leq c(i)$. As j is chosen to be the largest index in the set of vertices to be considered, it follows that a clique of size

$c(i)+1$ that contains W cannot exist in S_{i+1}. Trivially, if it finds a clique of size $c(i)+1$, it can prune the whole search and start calculating $c(i+2)$. Table $c[i]$ gives the largest clique that includes the vertex v_i. When searching for all maximum cliques, we first determine the size of the maximum cliques, and then starts the search again at the suitable position.

Let C be an $E_q(n, M, d, w)$ ECWC for $w = d$ and $C_0 = C \bigcup \{0\}$ be an $E_q(n, M_0, d)$ equidistant code. The EC and ECWC are closely related as shown by Theorem 2. Our approach is based on this relation and on the observation that an $E_q(n, M, d, d)$ code C can be shortened to an $E_q(n-1, M, d, d)$ code C'. Conversely, if we want to construct an $E_q(n, M, d, d)$ code C, we only need to consider lengthening of the $E_q(n-1, M, d, d)$ code C'.

The following theorem is derived from Theorem 4:

Theorem 8. *Any* $E_q(n, M, d, w)$ *ECWC code* C *contains* $E_q(n-1, M', d, w)$ *codes with* $M' \geq \left\lceil M\dfrac{n-w}{n} \right\rceil$ *codewords.*

The search space will only be the vectors which are at a distance exactly equal to d from the code C_0 and have exactly weight w. We will only have to care about the distance between codewords and for their weights. In the case $q = 5, 6$ we can construct the graph whose vertices represent vectors of length n over Z_q. We join two vertices by an edge if and only if the Hamming distance between the vectors is exactly equal to d and their weight is exactly equal to w. Then what we are interested in is the quantity $B_q(n, d, w)$, the size of the largest clique in this graph.

The results for EC are obtained by a computer program based on these methods. We made our own realization for EC. The upper bounds for EC which we used for our research are obtained from theorems presented in Section 3. The obtained bounds on the size of EC over an alphabet of five and six elements of length $n \leq 10$ are displayed in Table 1 and Table 2.

Some codes in the tables are obtained by the following construction:

Construction *I*: From the $E_q(n, M, d)$ code A we construct the code $E_q(n+k, M, d)$ in the following way: $\{(\underbrace{0...0}_{k}, a) | a \in A\}$.

The exact values for $n = d$ follow from Theorem 1. The exact values for $d = 3$ in Table 1 and Table 2 are obtained in [2]. The codewords of the $E_5(4, 9, 3)$ code are (up to equivalence): {0000, 0111, 0222, 1012, 1120, 1201, 2021, 2102, 2210}. There exists a family of unique optimal equidistant codes with parameters $E_3(n, 9, 3)$ for $n \geq 4$.

The values in Table 1 and Table 2 are described by the next Theorems.

Theorem 9. *There exist optimal equidistant codes with parameters:*

a) $E_5(n, 16, 4)$ *for* $5 \le n \le 10$

$E_5(5, 16, 4)$: {00000, 01111, 02222, 03333, 10123, 11032, 12301, 13210, 20231, 21320, 22013, 23102, 30312, 31203, 32130, 33021}.

b) $E_5(n, 25, 5)$ *for* $6 \le n \le 10$

$E_5(6, 25, 5)$: {000000, 011111, 022222, 033333, 044444, 101234, 112340, 123401, 134012, 140123, 202413, 213024, 224130, 230241, 241302, 303142, 314203, 320314, 331420, 342031, 404321, 410432, 421043, 432104, 443210.}

c) $E_5(7, 15, 6)$

$E_5(7, 15, 6)$: {0000000, 0111111, 0222222, 1012333, 1103244, 1234401, 2041442, 2323031, 2430213, 3144023, 3332140, 3421304, 4240134, 4304312, 4413420}.

d) $E_5(8, 10, 7)$

$E_5(8, 10, 7)$: {00000000, 01111111, 02222222, 10123333, 11032444, 13344012, 22404134, 24240341, 33410423, 34331230}.

e) $E_5(9, 10, 8)$ It follows from [16].

f) $E_5(10, 7, 9)$

$E_5(10, 7, 9)$: {0000000000, 0111111111, 0222222222, 1012333333, 2103234444, 3330441234, 4444042143}.

Theorem 10. *There exist optimal equidistant codes with parameters:*

a) $E_6(n, 16, 4)$ *for* $5 \le n \le 10$

$E_6(5, 16, 4)$: {00000, 01111, 02222, 03333, 10123, 11032, 12301, 13210, 20231, 21320, 22013, 23102, 30312, 31203, 32130, 33021}.

b) $E_6(n, 25, 5)$ *for* $6 \le n \le 10$

$E_6(6, 25, 5)$: {000000, 011111, 022222, 033333, 044444, 101234, 112340, 123401, 134012, 140123, 202413, 213024, 224130, 230241, 241302, 303142, 314203, 320314, 331420, 342031, 404321, 410432, 421043, 432104, 443210}.

c) $E_6(8, 15, 7)$

$E_6(8, 15, 7)$: {00000000,01111111,02222222, 03333333, 04444444, 10123455, 11035524, 12550143, 20254531, 21542305, 25310254, 32415035, 34051352, 43245150, 45421503}.

d) $E_6(9, 13, 8)$

$E_6(9, 13, 8)$: {000000000, 011111111, 022222222, 033333333, 044444444, 101234555, 110555234, 225015345, 252150453, 335451502, 353542015, 445523150, 454305521}.

e) $E_6(10, 8, 9)$

$E_6(10, 8, 9)$: {0000000000, 0111111111, 0222222222, 0333333333, 0444444444, 1012345555, 2105552345, 3550153254}.

Theorem 11. *There exist equidistant codes with parameters:*

a) $E_5(9, 12, 7)$

$E_5(9, 12, 7)$: {000000000, 001111111, 002222222, 003333333, 010123444, 011444023, 024014234, 034241340, 100244431, 201340244, 204413420, 304134042}.

b) $E_5(10, 16, 8)$

$E_5(10, 16, 8)$: {0000000000, 0011111111, 0022222222, 0033333333, 1100112233, 1111003322,1122330011, 1133221100, 2200223311, 2211332200, 2222001133,2233110022, 3300331122, 3311220033, 3322113300, 3333002211}.

Theorem 12. *There exist equidistant codes with parameters:*

a) $E_6(7, 18, 6)$

$E_6(7, 18, 6)$: {0000000, 0111111, 0222222, 0333333, 0444444, 0555555, 1012345, 1103254, 1234501, 1325410, 1450132, 1541023, 2021534, 2210453, 2304125, 3053421, 3402513, 3514230}.

b) $E_6(9, 15, 7)$

$E_6(9, 15, 7)$: {000000000, 001111111, 002222222, 003333333, 004444444, 010123455, 011035524, 012550143, 020254531, 021542305, 025310254, 032415035, 034051352, 043245150, 045421503}.

c) $E_6(10, 16, 8)$

$E_6(10, 16, 8)$: {0000000000, 0011111111, 0022222222, 0033333333, 1100112233, 1111003322,1122330011, 1133221100, 2200223311,2211332200, 2222001133, 2233110022, 3300331122, 3311220033, 3322113300, 3333002211}.

Remark : All $E_q(n + k, M, d)$ codes in the four previous theorems, which codewords are not explicitly listed are obtained from $E_q(n, M, d)$ by construction I.

Bounds on the size of equidistant codes

Table 1. $q = 5$

d	n	M
3	$4-10$	9
4	4	5
4	$5-10$	16
5	5	5
5	$6-10$	25
6	6	5
6	7	15
6	8	$15-33$
6	9	$15-36$
6	10	$15-36$
7	7	5
7	8	10
7	9	$12-19$
7	10	$12-38$
8	8	5
8	9	10
8	10	$16-21$
9	9	5
9	10	7
10	10	5

Table 2. $q = 6$

d	n	M
3	$4-10$	9
4	4	6
4	$5-10$	16
5	5	6
5	$6-10$	25
6	6	6
6	7	$18-36$
6	8	$18-41$
6	9	$18-46$
6	10	$18-51$
7	7	6
7	8	15
7	9	$15-46$
7	10	$15-51$
8	8	6
8	9	13
8	10	$16-51$
9	9	6
9	10	8
10	10	6

Acknowledgment. The authors wish to thanks Prof. V. Zinoviev for helpful discussions.

References

[1] J. L. Balcazar, J. Diaz, and J. Gabarro, Structural complexity theory I, *EATCS Monographs in Theoretical Computer Science* no. **11**, Springer-Verlag; Berlin, (1988).

[2] G. Bogdanova, Ternary Equididistant Codes and Maximum Clique Problem, *Proceedings of the EWM International Workshop on Groups and Graphs*, Varna, September, (2002), 15-18.

[3] G.T.Bogdanova and T.A.Yorgova, Bounds for Ternary Equidistant Constant Weight Codes, *Mathematics and Education in Mathematics 31*, Borovec, April 3-6, (2002), 131-135.

[4] G. Bogdanova and T. Yorgova, Bounds for Quaternary Equidistant Constant Weight Codes, *In proceedings ACCT'2002*, Tsarskoe selo, Russia, 7-14 September (2002), 46-49.

[5] G. Bogdanova and T. Todorov, Some bounds for equidistant codes,In: *Proceedings International Workshop on Algebraic and Combinatorial Coding Theory*, Bulgaria, June, (2004).

[6] P. Delsarte, Bounds for unrestricted codes, by linear programming, *Philips Res.*, Rep. **27** (1972), 47-64.

[7] F.W. Fu, T. Klove, Y. Luo, and V.K. Wei, On Equidistant Constant Weight codes,In: *Proceedings WCC'2001 Workshop on Coding and Cryptograpy*, Paris, France, Jan (2001), 225-232.

[8] J.I. Hall, Bounds for equidistant codes and partial projective planes, *Discrete Math.*, **17** (1977), 85-94.

[9] J.I. Hall, A.J.E.M. Jansen, A.W.J. Kolen, and J.H. van Lint, Equidistant codes with distance 12, *Discrete Math.*, **17** (1977), 71-83.

[10] J. Hartmanis, Computational complexity theory, *Proceedings of Symposia in Applied Mathematics* no. **38**. American Mathematical Society; Providence, RI, 1989.

[11] W. Heise and Th. Honold, Some Equidistant Constant Weight Codes, http://fatman.mathematik.tu-muenchen.de/heise/MAT/code_oval.html

[12] S.M. Johnson, A new upper bound for error-correcting codes, *IRE Transaction on Information Theory*, IT-8(3), April 1962, 203–207.

[13] J.van Leeuwen, *Handbook of Theoretical Computer Science*, part A. MIT Press. Chapter 2: (1990), 69-161.

[14] J.H. van Lint, A theorem on equidistant codes, *Discrete Math.*, **6** (1973), 353-358.

[15] P.R.J. Ostergard, A fast algorithm for the maximum clique problem, *Discrete Appl. Math.*, 120 (2002), 195-205.

[16] N.V. Semakov and V.A. Zinoviev, Equidistant q-ary codes with maximal distance and resolvable balanced incomplete block designs, *Problems of Information Transmission*, **4** (1968), 2, 3-10.

[17] N.V. Semakov, G.V. Zaitsev and V.A. Zinoviev, Class of maximal equidistant codes, *Problems of Information Transmission*, **5** (1969), 2, 84-87.

[18] D.R. Stinson and G.H.J. van Rees, The equivalence of certain equidistant binary codes and symmetric BIBDs, *Combinatorica*, **4** (1984), 357-362.

[19] M. Svanstrom, P.R.J. Ostergard and G.T. Bogdanova, Bounds and Constructions for Ternary Constant-Composition Codes, *IEEE Trans. Inform. Theory*, **48** (2002), 1, 101-111.

Institute of Mathematics and Informatics,
Bulgarian Academy of Sciences
E-mail: galina@moi.math.bas.bg
E-mail: todor@moi.math.bas.bg

Received 17.02.2006