Upper Bounds on the Rate of Superimposed (s, ℓ)-Codes Based on Engel's Inequality [1]

A.G. D'yachkov, P.A. Vilenkin, S.A. Yekhanin
Moscow State University, Faculty of Mechanics and Mathematics,
Department of Probability Theory, Moscow, 119992, Russia
dyachkov@mech.math.msu.su, paul@vilenkin.dnttm.ru, yekhanin@cityline.ru

Abstract

Applying an important combinatorial result of K. Engel [2], we improve upper bounds on the rate of superimposed (s, ℓ) - codes obtained in [3, 4].

1 Definitions and Formulations of Results

We use the symbol $\triangleq$ to denote definitional equalities.

Let $N \geq 1$, $t \geq 1$, $s \geq 1$ and $\ell \geq 1$, where $s + \ell \leq t$, be arbitrary integers. A family of t binary codewords of length N is called a *superimposed* (s, ℓ)*-code* [3, 4] of size t and length N if for any two non-intersecting subsets of codewords $\mathcal{S}$ of size $|\mathcal{S}| = s$ and $\mathcal{L}$, $|\mathcal{L}| = \ell$, there exists a coordinate k, $k = 1, 2, \ldots, N$, in which all codewords from set $\mathcal{S}$ have 0's and all codewords from set $\mathcal{L}$ have 1's.

Let $N(t, \ell, s) = N(t, s, \ell)$ denote the minimal possible length of superimposed (s, ℓ) - code of size t. For fixed s and ℓ, the number

$$R(\ell, s) = R(s, \ell) \triangleq \overline{\lim_{t \to \infty}} \ \frac{\log_2 t}{N(t, \ell, s)}$$

is called [3, 4] a *rate* of superimposed (s, ℓ) - code.

Let $h(u) \triangleq -u \log_2 u - (1-u) \log_2 (1-u)$, $0 < u < 1$, be the binary entropy. To formulate the upper bound on the rate $R(s, \ell)$, $s \geq \ell \geq 1$, we introduce the function [1]

$$\mathsf{f}_s(v) \triangleq h(v/s) - v h(1/s), \quad s = 1, 2, \ldots,$$

of argument v, $0 < v < 1$. The following three statements are true.

Theorem 1. 1. *If $s = 1, 2, \ldots$, then the rate $R(s, 1) \leq \overline{R}(s, 1)$, where*

$$\overline{R}(1,1) = R(1,1) = 1, \quad \overline{R}(2,1) \triangleq \max_{0<v<1} \ \mathsf{f}_2(v) \ = \ 0.321928 \tag{1}$$

and sequence $\overline{R}(s, 1)$, $s = 3, 4, \ldots$, is defined recurrently as the unique solution of the equation

$$\overline{R}(s,1) = \mathsf{f}_s \left(1 - \frac{\overline{R}(s,1)}{\overline{R}(s-1,1)} \right). \tag{2}$$

2. *The rate*

$$R(2,2) \leq \overline{R}(2,2) \triangleq \overline{R}(2,1)/2 = 0.160964. \tag{3}$$

[1] Work was supported by the Russian Foundation of Basic Research, grant N. 01-01-00495 and the INTAS, grant N. 00-738

3. *If $s > \ell \geq 2$ or $s \geq \ell \geq 3$, then the rate $R(s,\ell)$ satisfies the inequality*

$$R(s,\ell) \leq \overline{R}(s,\ell) \triangleq \min_{x=0,1,\ldots,s-1} \min_{y=0,1,\ldots,\ell-1} \left\{ \overline{R}(s-x,\ell-y) \cdot \frac{x^x \cdot y^y}{(x+y)^{x+y}} \right\}, \tag{4}$$

where sequence $\overline{R}(s,1)$, $s = 1,2,\ldots$, and the number $\overline{R}(2,2)$ are defined by (1)-(3).

The first statement was proved in [1] (see, also [4]). The second statement was proved in [4]. The third statement is an evident consequence of the following result obtained by K. Engel [2].

Theorem 2. (Engel's inequality [2].) *If $s \geq \ell \geq 2$, then for any $x = 0,1,\ldots,s-1$ and any $y = 0,1,\ldots,\ell-1$, the rate*

$$R(s,\ell) \leq \ R(s-x,\ell-y) \cdot \frac{x^x \cdot y^y}{(x+y)^{x+y}}. \tag{5}$$

In section 3, we briefly present the proof of Theorem 2 from paper [2]. The numerical values of upper bound $\overline{R}(s,\ell)$, $1 \leq \ell \leq s \leq 4$, are:

$$\overline{R}(2,1) = .32193, \quad \overline{R}(3,1) = .19928, \quad \overline{R}(4,1) = .14046, \quad \overline{R}(2,2) = .16096,$$

$$\overline{R}(3,2) = .08048, \quad \overline{R}(4,2) = .04769, \quad \overline{R}(3,3) = .04024, \quad \overline{R}(4,3) = .02012$$

and $\overline{R}(4,4) = .01006$.

2 Asymptotics of $\overline{R}(s,\ell)$

If $s \to \infty$ and $\ell \geq 2$ is fixed, then the optimal values of x and y in definition (4) of $\overline{R}(s,\ell)$ are $y = \ell - 1$, $x \sim ps$, $0 < p < 1$, and

$$\overline{R}(s,\ell) \sim \min_{0<p<1} \left\{ \overline{R}(s(1-p),1) \cdot \frac{(ps)^{ps} \cdot (\ell-1)^{\ell-1}}{(ps+\ell-1)^{ps+\ell-1}} \right\}.$$

Using the asymptotic ($s \to \infty$) form [1, 4] of upper bound $\overline{R}(s,1) \sim 2\log s/s^2$, we get

$$\overline{R}(s,\ell) \sim \min_{0<p<1} \left\{ \frac{2\log[s(1-p)]}{s^2(1-p)^2} \cdot \frac{(ps)^{ps} \cdot (\ell-1)^{\ell-1}}{(ps+\ell-1)^{ps+\ell-1}} \right\} \sim \frac{(\ell+1)^{\ell+1}}{2\mathsf{e}^{\ell-1}} \cdot \frac{\log s}{s^{\ell+1}}, \tag{6}$$

where $\mathsf{e} = 2.71828$ is the base of natural logarithm and we took into account that

$$\max_{0<p<1} \{(1-p)^2 p^{\ell-1}\} = (\ell-1)^{\ell-1} \frac{4}{(\ell+1)^{\ell+1}}$$

with the optimal value $p = \frac{\ell-1}{\ell+1}$. For $\ell \geq 2$, upper bound (6) is better than the similar upper bound

$$\overline{R}_{old}(s,\ell) \sim (\ell+1)! \cdot \frac{\log s}{s^{\ell+1}},$$

which was obtained in [4].

3 Proof of Engel's inequality

Let $1 \le \ell < t$, $1 \le s \le t$, where $2 \le s + \ell \le t$, be arbitrary integers, $[t] \triangleq \{1, 2, \dots, t\}$ and the set B_t of size $|B_t| = 2^t$ be the *Boolean lattice* constituted of all subsets of $[t]$.

Introduce the set $P = P(t, \ell, t-s) \subseteq B_t$, $|P| = |P(t, \ell, t-s)| = \sum\limits_{n=\ell}^{t-s} \binom{t}{n}$, whose elements are n-subsets of $[t]$, where $\ell \le n \le t-s$. Let $Z \subseteq Y \subseteq [t]$ be arbitrary subsets of $[t]$. Denote by $J = J(t, \ell, t-s)$ the set of all *intervals* $I = I(Z, Y)$:

$$I = I(Z, Y) \triangleq \{X \,:\, X \in P,\ Z \subseteq X \subseteq Y\}, \quad \text{where} \quad |Z| = \ell,\ |Y| = t-s,\ |[t] \setminus Y| = s.$$

Obviously, each interval $I \in J$ is isomorphic to $B_{t-s-\ell}$ and $|I| = 2^{t-s-\ell}$. In addition, any element $X \in P$ is contained in $\binom{|X|}{\ell}\binom{t-|X|}{t-s-|X|}$ intervals of J. Taking all X with $|X| = \ell$ (resp. all X with $|X| = t-s$) we obtain

$$|J| = \binom{t}{\ell}\binom{t-\ell}{t-s-\ell} = \binom{t}{t-s}\binom{t-s}{\ell}. \tag{7}$$

A set $T \subseteq P$ is called a *point cover* of J if for any interval $I \in J$, the intersection $T \cap I \ne \varnothing$. The minimal size of point cover T is denoted by $\tau(t, \ell, t-s)$.

Lemma 1. *The minimal length of superimposed* (s, ℓ)*-code* $N(t, \ell, s) = \tau(t, \ell, t-s)$.

Proof of Lemma 1. Let C be a superimposed (s, ℓ)- code of length N and size t. Fix an order over codewords of $C = \{c_1, c_2, \dots, c_t\}$. Introduce the following *correspondence* between coordinates of codewords $c_1, c_2, \dots, c_t$ and subsets of $[t]$: *a set* $X_k \subseteq [t]$ *corresponding to a coordinate* k, $k = 1, 2, \dots, N$, *contains the numbers* i *of codewords* c_i *having* 1*'s in the* k*-th coordinate.* Without loss of generality, $\ell \le |X_k| \le t-s$. Consider the set $T \triangleq \{X_1, X_2, \dots, X_N\} \subseteq P = P(t, \ell, t-s)$. Take an arbitrary interval $I = I(Z, Y) \in J = J(t, \ell, t-s)$. By definition of the superimposed (s, ℓ) - code C, there exists a coordinate k such that all codewords with numbers in Z have 1's in the k-th coordinate and all codewords with numbers in $[t] \setminus Y$ have 0's in the k-th coordinate, i.e., $Z \subseteq X_k \subseteq Y$. Hence, $X_k \in I$ and $T \cap I \ne \varnothing$. Therefore, T is a point cover of J. Thus, we have proved that $N \ge \tau(t, \ell, t-s)$, i.e., $N(t, \ell, s) \ge \tau(t, \ell, t-s)$. To prove $N(t, \ell, s) \le \tau(t, \ell, t-s)$ one needs to check that superimposed (s, ℓ)-code can be constructed from a point cover using the correspondence described above.

We introduce several additional definitions. A *fractional matching* of $P = P(t, \ell, t-s)$ is a function $\quad f = f(I) \ge 0, \quad I \in J = J(t, \ell, t-s) \quad$ such that

$$\forall X \in P: \qquad \sum_{I \ni X} f(I) \le 1.$$

A *fractional point cover* of J is a function $\quad g = g(X) \ge 0, \quad X \in P \quad$ such that

$$\forall I \in J: \qquad \sum_{X \in I} g(X) \ge 1.$$

The *fractional matching number* $\nu^*(t, \ell, t-s)$ and *fractional covering number* $\tau^*(t, \ell, t-s)$ are defined by

$$\nu^*(t, \ell, t-s) \triangleq \max\left\{\sum_{I \in J} f(I) : f \text{ is a fractional matching of } P\right\},$$

$$\tau^*(t,\ell,t-s) \triangleq \min\left\{\sum_{X\in P} g(X) : g \text{ is a fractional point cover of } J\right\}.$$

Lemma 2. *We have* $\nu^*(t,\ell,t-s) = \tau^*(t,\ell,t-s) = \min_{\ell\le m\le t-s} \binom{t}{m}/\binom{t-s-\ell}{m-\ell}$.

Proof of Lemma 2. The first equality follows from the Duality Theorem of linear programming. Suppose that the minimum in the right-hand side is attained at $m=m_0$. To prove the second equality, it is enough to find a fractional matching f and a fractional point cover g such that

$$\sum_{I\in J} f(I) = \frac{\binom{t}{m_0}}{\binom{t-s-\ell}{m_0-\ell}} = \sum_{X\in P} g(X). \tag{8}$$

We choose

$$f(I) \triangleq \frac{1}{\binom{m_0}{\ell}\binom{t-m_0}{t-s-m_0}} \qquad \text{for all } I\in J$$

and

$$g(X) \triangleq \begin{cases} 0, & \text{if } |X|\ne m_0;\\ \frac{1}{\binom{t-s-\ell}{m_0-\ell}}, & \text{if } |X| = m_0. \end{cases}$$

The function f is a fractional matching since

$$\sum_{I\ni X} f(I) = \frac{\binom{|X|}{\ell}\binom{t-|X|}{t-s-|X|}}{\binom{m_0}{\ell}\binom{t-m_0}{t-s-m_0}} = \frac{\binom{t}{m_0}/\binom{t-s-\ell}{m_0-\ell}}{\binom{t}{|X|}/\binom{t-s-\ell}{|X|-\ell}} \le 1 \qquad \text{for all } X\in P,$$

and g is a fractional point cover since

$$\sum_{X\in I} g(X) = \frac{\binom{t-s-\ell}{m_0-\ell}}{\binom{t-s-\ell}{m_0-\ell}} = 1 \quad \text{for all} \quad I\in J.$$

The equality (8) can be verified by straightforward computation using equality (7).

Lemma 3. *For fixed* ℓ, s *and* $t\to\infty$, *the number* $\tau^*(t,\ell,t-s) \sim \frac{(s+l)^{s+l}}{s^s l^l}$.

Proof of Lemma 3. Let ℓ, s and u, $0<u<1$, be fixed. If $t\to\infty$ and $m\sim ut$, then

$$\frac{\binom{t}{m}}{\binom{t-s-\ell}{m-\ell}} = \frac{t(t-1)\cdots[t-(s+\ell-1)]}{\{[m-(\ell-1)]\cdots m\}\cdot\{[(t-m)-(s-1)]\cdots(t-m)\}} \sim \left[u^\ell\cdot(1-u)^s\right]^{-1}.$$

Using the definition of $\tau^*(t,\ell,t-s)$ in Lemma 2, we have

$$\tau^*(t,\ell,t-s) \sim \left\{\max_{0<u<1}\left[u^\ell\cdot(1-u)^s\right]\right\}^{-1} = \left\{\frac{s^s l^\ell}{(s+\ell)^{s+\ell}}\right\}^{-1},$$

where the maximum is achieved at $u=\frac{\ell}{\ell+s}$.

Lemma 4. *For any* $x=0,1,\ldots,s-1$ *and* $y=0,1,\ldots,\ell-1$,

$$\frac{\tau(t,\ell,t-s)}{\tau(t-x-y,\ell-y,t-s-y)} \ge \tau^*(t,y,t-x).$$

Proof of Lemma 4. Let T, $|T| = \tau(t, \ell, t-s)$ be an optimal point cover of $J(t, \ell, t-s)$. We have $0 \le y < \ell < t-s < t-x$ and $T \subset P(t, \ell, t-s) \subset P(t, y, t-x)$. For $X \in P(t, y, t-x)$, we define the function

$$g(X) \triangleq \begin{cases} 1/\tau(t-x-y, \ell-y, t-s-y), & \text{if } X \in T, \\ 0, & \text{otherwise.} \end{cases}$$

It is enough to show that g is a fractional point cover of $J(t, y, t-x)$. Consider an arbitrary interval $I \in J(t, y, t-x)$ which is isomorphic to the Boolean lattice B_{t-x-y}. Moreover, the part of I which lies between levels ℓ and $t-s$ is isomorphic to $P(t-x-y, \ell-y, t-s-y)$. Since the considered set T is a point cover of $J(t, \ell, t-s)$ the intersection $T \cap I$ must be a point cover of the corresponding set of intervals $J(t-x-y, \ell-y, t-s-y)$. Thus,

$$\sum_{X \in I} g(X) \ \ge \ \frac{|T \cap I|}{\tau(t-x-y, \ell-y, t-s-y)} \ \ge 1.$$

Proof of Theorem 2. If $t' \triangleq t-x-y$, then $t-s-y = t'-(s-x)$. Using Lemma 1, we have $\tau(t-x-y, \ell-y, t-s-y) = N(t-x-y, \ell-y, s-x)$. Therefore, we can rewrite the inequality from Lemma 4 in the form

$$N(t, \ell, s) \ge \tau^*(t, y, t-x) \cdot N(t-x-y, \ell-y, s-x).$$

For s, ℓ, x, y fixed and $t \to \infty$, the application of Lemma 3 yields

$$N(t, \ell, s) \ \ge \ \frac{(x+y)^{x+y}}{x^x y^y} \cdot N(t, \ell-y, s-x)(1+o(1)). \tag{9}$$

If we multiply by $\log_2 t$ the opposite inequality for reciprocals in (9) and pass to the limit, then we obtain inequality (5).

Theorem 2 is proved.

References

[1] A.G. D'yachkov, V.V. Rykov, "Bounds on the Length of Disjunct Codes", *Problemy Peredachi Informatsii*, **17**, 1982, n. 3, pp. 7-13, (in Russian).

[2] K. Engel, "Interval Packing and Covering in the Boolean Lattice", *Combinatorics Prob. and Computing*, **5**, 1996, pp. 373-384.

[3] A.G. D'yachkov, A.J. Macula, D.C. Torney, P.A. Vilenkin, S.M. Yekhanin, "New Results in the Theory of Superimposed Codes", *Proc. of the 7-th Inter. Workshop "Algebraic and Combinatorial Coding Theory", ACCT-7,* Bansko, Bulgaria, June 17-24, 2000, pp. 126-136.

[4] A.G. D'yachkov, A.J. Macula, D.C. Torney, P.A. Vilenkin, "Families of Finite Sets in which No Intersection of ℓ Sets Is Covered by the Union of s Others", *J. Comb. Theory, ser. A*, **99**, 2002, pp. 195-218.