Recent results on superimposed codes

Hyun Kwang Kim[1]

Abstract

A (w, r) *cover-free family* is a family of subsets of a finite set such that no intersection of w members of the family is covered by a union of r others. A (w, r) *superimposed code* is the incidence matrix of such a family. Such a family also arises in cryptography as the concept of a key distribution pattern. In the present paper, we survey some recent results on superimposed codes. We first briefly review basic properties of superimposed codes. Next we survey some results on the construction of superimposed codes, the optimal superimposed codes, and the uniqueness of some optimal superimposed codes.

1 Definition and basic properties

We start with the definition of (w, r) superimposed codes.

Definition 1. An $N \times T$ $(0, 1)$-matrix C is called a (w, r) *superimposed code of size* $N \times T$, if for any pair of subsets $I, J \subset [T] = \{1, 2, \cdots, T\}$ such that $|I| = w$, $|J| = r$ and $I \cap J = \varnothing$ there exists a coordinate $x \in [N] = \{1, 2, \cdots, N\}$ such that $c_{xp} = 1$ for all $p \in I$ and $c_{xq} = 0$ for all $q \in J$.

[1]Hyun Kwang Kim is with the Department of Mathematics, Pohang University of Science and Technology, Pohang 790–784, The Republic of Korea (hkkim@postech.ac.kr). His work was supported by Com[2]MaC-KOSEF and POSTECH BSRI research fund.

Let C be a (w,r) superimposed code of size $N \times T$. The elements of $[T]$ (resp. $[N]$) are referred to as points (resp. blocks) of C. We conventionally use letters $p_1, p_2, \ldots$ (resp. $x, y, \ldots$) to denote points (resp. blocks). The number T of points of C is called the *cardinality* of C and the number N of blocks is called the *length* of C. Sometimes we need to view C as a code. In this circumstance, we take the columns of C as codewords and regard C as a code of length N with cardinality T.

Let S_p denote the characteristic set of p-th column and L_x denote the characteristic set of x-th row of C. We say that p is incident with x (resp. x is incident with p) if $p \in L_x$ (resp. $x \in S_p$). It follows from the definition of a (w,r) superimposed code that the family $\mathcal{A} = \{S_1, \ldots, S_T\}$ of subsets of $[N] = \{1, 2, \ldots, N\}$ becomes a family in which no intersection of w members is covered by the union of r others. It is called a (w,r) *cover-free family*. Such a family also arises in cryptography as the concept of a key distribution pattern. Superimposed codes have wide applications in combinatorial science such as group testing, perfect hash families, key storage in secure network, and tracing traitors.

There are simple examples of (w,r) superimposed codes. If we take a matrix whose rows are all possible binary vectors of weight w, then this matrix becomes a (w,r) superimposed code for any r. We call this matrix the trivial (w,r) superimposed code of size $N \times T$, where $N = \binom{T}{w}$. Obviously we are interested in superimposed codes which have better performance than trivial superimposed codes. The main problem in the study of superimposed codes is to find the minimal length $N(T;w,r)$ of a (w,r) superimposed code for a given cardinality T, or the maximal cardinality $T(N;w,r)$ of a (w,r) superimposed code for a given length N. It is clear that $N(T;w,r) = N(T;r,w)$; thus we may only consider the case $w \le r$. Let C be a (w,r) superimposed code of size $N \times T$. By deleting one column of C, we obtain a (w,r) superimposed code of size $N \times (T-1)$. Therefore we have $N(T-1;w,r) \le N(T;w,r)$.

For fixed T and (w,r), a (w,r) superimposed code of size $N \times T$ is called an *optimal superimposed code* if $N = N(T;w,r)$. There are many cases in which the trivial superimposed code becomes an optimal superimposed code. For example, it is known [4] that $N(T;w,r) = \binom{T}{w}$ when $T \leq w + r + \frac{r}{w}$. We remark that there is a case which is not covered by Engel's result and the trivial superimposed code is still an optimal superimposed code. For example, it will be prove that $N(7;2,3) = 21$. Moreover we can prove that $N(T;w,r) = \binom{T}{w}$ whenever $T \leq \frac{(w+1)r}{w-1} - \sqrt{\frac{36r}{w-1}}$ (see [8]). When the trivial superimposed code is not optimal, the computation of the value $N(T;w,r)$ has become a serious research problem. We refer to the appendix of [7] for the known values of $N(T;w,r)$.

The idea of using some recurrent methods to obtain an upper bound of asymptotic rate was considered in earlier works. We will formulate two recurrent methods in Lemmas 1 and 2. We will combine these methods with classical bounds in coding theory to prove the nonexistence of superimposed code of certain size.

Definition 2. Denote by

$$J_1 = \min_{1 \leq i \leq T} |S_{p_i}|$$

the minimal column weight of a superimposed code C.

Definition 3. Denote by

$$J_0 = N - \max_{1 \leq i \leq T} |S_{p_i}|$$

the minimal number of 0's in the columns of a superimposed code C.

Lemma 1 [L.Bassalygo]. *If there is a (w,r) superimposed code of size $N \times T$, then there are $(w-1,r)$ superimposed code of size $J_1 \times (T-1)$ and $(w,r-1)$ superimposed code of size $J_0 \times (T-1)$.*

Lemma 2 [G.Kabatianski]. *If there is a (w,r) superimposed code, say C, of size $N \times T$ then there is a $(w-1,r-1)$ superimposed code of size $[d/2] \times (T-2)$, where d denotes the minimum distance of the code C.*

Proof. Consider the code C. We may assume that the first two columns c^1, c^2 are a pair of codewords for which $d(c^1,c^2) = d$. By symmetry we may assume that

$$|\{i : c_i^1 = 1; c_i^2 = 0\}| \leq |\{i : c_i^1 = 0; c_i^2 = 1\}|.$$

Let

$$U_C = \{i : c_i^1 = 1; c_i^2 = 0\}.$$

Then, $|U_C| \leq d/2$.

Consider the submatrix C_1 of C consists of i-th row of C, where i runs through U_C. We claim that the code $C_1 \setminus \{c^1; c^2\}$, obtained from C_1 by deleting first two columns, is a $(w-1, r-1)$ superimposed code of size $|U_C| \times (T-2)$. Take arbitrary subsets X, Y of $\{3, 4, \ldots, T\}$ with $|X| = w-1$, $|Y| = r-1$, and $X \cap Y = \emptyset$. Put $\tilde{X} = X \cup \{1\}$ and $\tilde{Y} = Y \cup \{2\}$. From the superimposedness of C we have a coordinate i such that $c_{ij} = 1$ for $j \in \tilde{X}$ and $c_{ij} = 0$ for $j \in \tilde{Y}$. Since $c_{i1} = 1$ and $c_{i2} = 0$, we should have $i \in U_C$. Note that, for this i, we have $c_{ij} = 1$ for $j \in X$ and $c_{ij} = 0$ for $j \in Y$. This proves the lemma.

2 Construction via combinatorial designs

In this section we consider constructions of superimposed codes of small size. We first construct superimposed codes using t-designs. Next we introduce the notion of super-simple designs. Using super-simple designs, we finally construct some new superimposed codes of small size.

Proposition 1. [13] *A $(t+1)-(v,k,1)$ design is a (w,r) superimposed code of size $N \times v$, where*

$$w = t, N = \frac{\binom{v}{t+1}}{\binom{k}{t+1}} = \frac{(v-t)\binom{v}{t}}{(k-t)\binom{k}{t}} \text{ and } r < \frac{v-t}{k-t}.$$

It is known that a $3-(Q^2+1, Q+1, 1)$ design exists when Q is a prime power. Applying Proposition 1, we obtain

Proposition 2. *Suppose Q is a prime power. Then there exists a $(2,Q)$ superimposed code of size $Q(Q^2+1) \times (Q^2+1)$.*

Corollary 1. There is a $(2,3)$ superimposed code of size 30×10.

A super-simple t-design was introduced in [12] (the terminology 'super-simple design' is due to Gronau and Mullin [6]). A super-simple t-design is defined to be a $t-(v,k,\lambda)$ design with $\lambda > 1$ in which the intersection of any two blocks has at most t elements.

We are interested in super-simple designs because of the following theorem which is a generalization of some constructions used in [11] and [13].

Theorem 1. *A super-simple $t-(v,k,\lambda)$ design is a $(t,\lambda-1)$ superimposed code of size $N \times v$, where $N = \frac{\lambda\binom{v}{t}}{\binom{k}{t}}$.*

Proof. Consider any t points $p_1, p_2, \ldots, p_t$. There are exactly λ blocks that contain these points ,i.e., $|S_{p_1} \bigcap S_{p_2} \bigcap \ldots \bigcap S_{p_t}| = \lambda$. Consider any other r points $h_1, h_2, \ldots, h_r$, where $r = \lambda - 1$. Since no two (or more) blocks of a super-simple t design can have more than t common points, for any l with $1 \leqslant l \leqslant r$, we have

$$|S_{p_1} \bigcap S_{p_2} \bigcap \ldots \bigcap S_{p_t} \bigcap S_{h_l}| \leq 1.$$

So

$$|S_{p_1} \bigcap S_{p_2} \bigcap \ldots \bigcap S_{p_t} \bigcap \{ \bigcup_{l=1,\ldots,r} S_{h_l} \}| \leq r < \lambda.$$

If it were

$$S_{p_1} \bigcap S_{p_2} \bigcap \ldots \bigcap S_{p_t} \subseteq \bigcup_{l=1,\ldots,r} S_{h_l},$$

$$|S_{p_1} \bigcap S_{p_2} \bigcap \ldots \bigcap S_{p_t} \bigcap \{ \bigcup_{l=1,\ldots,r} S_{h_l} \}| = |S_{p_1} \bigcap S_{p_2} \bigcap \ldots \bigcap S_{p_t}| = \lambda.$$

So we have

$$S_{p_1} \bigcap S_{p_2} \bigcap \ldots \bigcap S_{p_t} \nsubseteq \bigcup_{l=1,\ldots,r} S_{h_l}.$$

This proves the theorem.

It is easy to see that every $(t+1)-(v,k,1)$ design is a super-simple $t-(v,k,(v-t)/(k-t))$ design. We note that, in this case, Theorem 1 is equivalent to Proposition 1. So we are interested in super-simple designs which do not arise from Steiner systems. In

[1] and [2] Chen proved that there exists a super-simple $2-(v,4,3)$ design if and only if $v \equiv 0$ or $1 \pmod 4$, $v \geq 8$, and there exists a super-simple $2-(v,4,4)$ design if and only if $v \equiv 1 \pmod 3$, $v \geq 10$. So we have

Corollary 2. There is a $(2,2)$ superimposed code of size 14×8.

Corollary 3. There is a $(2,2)$ superimposed code of size 18×9.

3 Optimality of certain superimposed codes

In this section we combine the idea of recurrence methods and the methods in coding theory to prove that some superimposed codes constructed in the previous section are optimal. We only give proofs for Theorems 2 and 3, and refer to [7] for proofs of remaining parts.

Proposition 3. *The optimal* $(2,2)$ *superimposed code with cardinality 8 has length 14.*

Theorem 2. *The optimal* $(2,2)$ *superimposed code with cardinality* 9 *has length* 18.

Proof. Let us first prove that $N(9;2,2) \geq 18$. Suppose that there is a $(2,2)$ superimposed code of size 17×9. From the Plotkin bound we have $d(17,9) \leq 9$, where $d(N,T) = \max_{|C|=T} d(C)$ denotes the maximal possible code distance for codes of length N with cardinality T. Then, by Lemma 2, there is a $(1,1)$ superimposed code of size 4×7. Since $T(4;1,1) = 6$, we have $18 \leq N(9;2,2)$. On the other hand, by Corollary 3, there is a $(2,2)$ superimposed code of size 18×9. This proves that $N(9;2,2) = 18$.

Theorem 3. *The optimal* $(2,3)$ *superimposed code with cardinality 7 has length 21.*

Proof. Let us first prove that $N(7;2,3) \geq 21$. Assume that there is a $(2,3)$ superimposed code, say C, of size 20×7. We claim that the weight of any codeword is 6,i.e., $|S_{p_i}| = 6$ for all $i = 1,2,\ldots,T$. If there is a codeword of weight 7 or more, then by Lemma 1, we may have a $(2,2)$ superimposed code of size 13×6. However, Table 3 in the appendix of [7] shows that $N(6;2,2) = 14$. On the other hand, if there is a codeword of weight 5 or less, we may have a $(1,3)$ superimposed code of size 5×6. However, Table 2 in the appendix of [7] gives us $N(6;1,3) = 6$. So there is neither a codeword of weight 7 or more nor a codeword of weight 5 or less in the code C. This proves our claim.

By our claim the code C is a constant weight code. Due to the Johnson bound (see [10] pp. 525), every constant weight code of weight g and cardinality T satisfies

$$T \leq [(n \cdot d/2)/(g^2 - gn + nd/2)]$$

provided that the denominator is positive. So we have $d \leq 9$. Therefore, by Lemma 2, there is a $(1,2)$ superimposed code of size 4×5. However, Table 1 in the appendix of [7] gives us that $N(5;1,2) = 5$. So we have $N(7;2,3) \geq 21$. On the other hand, there is a trivial $(2,3)$ superimposed code of size 21×7. This proves that $N(7;2,3) = 21$.

Theorem 4. *The optimal* $(2,3)$ *superimposed code with cardinality 10 has length 30.*

Theorem 5. *The optimal* $(3,3)$ *superimposed code with cardinality* 11 *has length* 66.

4 Uniqueness of some optimal superimposed codes

In this section we prove that some optimal superimposed codes constructed in [5], [7], and [11] are unique. For this purpose we need information on the values of $N(T;1,1)$, $N(T;1,2)$, $N(T;1,3)$, and $N(T;2,2)$ for small values of T. It is well known (the Sperner theorem) that $N(T;1,1)$ is the smallest N such that $\binom{N}{\lfloor \frac{N}{2} \rfloor}$ is greater than or equal to T. For the remaining values, we refer to the appendix of [7].

It is easy to see that if we permute rows or columns of a superimposed code then new matrix is also a superimposed code. This motivates the following definition.

Definition 4. Two superimposed codes C and C' are equivalent if one can be transformed into the other by a finite number of row or column permutations.

We have the following three uniqueness theorem for optimal superimposed codes. We only give a proof for Theorem 7 and refer to [9] for proofs of remaining theorems.

Theorem 6. *A binary matrix C is a $(1,2)$ superimposed code of size 9×12 if and only if it is the transpose of the incidence matrix of a $2-(9,3,1)$ design.*

The following lemma will be useful in proving Theorem 7 and 8.

Lemma 3. *Let C be a (w,r) superimposed code. Suppose that, for some w points $p_1,p_2,\ldots,p_w$ of C, we have $|L_x|>w$ for all $x\in S_{p_1}\bigcap S_{p_2}\bigcap\ldots\bigcap S_{p_w}$. Then we have*

$$\left|S_{p_1}\bigcap S_{p_2}\bigcap\ldots\bigcap S_{p_w}\right|\geq r+1.$$

Proof. Suppose that there are only r blocks in the set $S_{p_1}\bigcap S_{p_2}\bigcap\ldots\bigcap S_{p_w}$. For each block $x_j\in S_{p_1}\bigcap S_{p_2}\bigcap\ldots\bigcap S_{p_w}$ there exists at least one point q_j incident with x_j and distinct from $p_1,p_2,\ldots,p_w$. Consider the set J consists of such q_j's. Then,

$$S_{p_1}\bigcap S_{p_2}\bigcap\ldots\bigcap S_{p_w}\subset\bigcup_{q_j\in J}S_{q_j}.$$

From the superimposedness of C it follows that $J\cap\{p_1,p_2,\ldots,p_w\}\neq\emptyset$ for the set J with $|J|\leq r$. However it is clear from our construction that

$$J\cap\{p_1,p_2,\ldots,p_w\}=\emptyset.$$

This proves the lemma.

Theorem 7. *A binary matrix C is a $(2,2)$ superimposed code of size 14×8 if and only if it is the incidence matrix of a $3-(8,4,1)$ design.*

Proof. Let C be the incidence matrix of a $3-(8,4,1)$ design. It is easy to see that every $3-(8,4,1)$ design is a $2-(8,4,3)$ super-simple design, and vice versa. It follows from Theorem 1 that C is a $(2,2)$ superimposed code of size 14×8.

Conversely, let C be a $(2,2)$ superimposed code of size 14×8. If $|L_x|=2$ or 3 for some $x\in\{1,2,\cdots 14\}$ then there exists a $(2,2)$ superimposed code of size 14×6 with $|L_x|$=0 or 1. So there exists a $(2,2)$ superimposed code of size 13×6. However it is well known that $N(6;2,2)=14$ (see [7]). Similarly we can prove that there is no $x\in\{1,2,\cdots,14\}$

such that $|L_x| = 5$ or 6. So we should have that $|L_x| = 4$ for all $x \in \{1, 2, \ldots 14\}$. It follows from Lemma 1 that $|S_{p_1} \bigcap S_{p_2}| \geq 3$ for any two distinct p_1, p_2 in $\{1, 2, \cdots, 8\}$. By counting the number of $(1,1)$'s in C, we have that $|S_{p_1} \bigcap S_{p_2}| = 3$ for any two distinct p_1, p_2. This proves that C is a $2-(8,4,3)$ design. Now suppose that there are three distinct points p_1, p_2, and p_3 such that $|S_{p_1} \bigcap S_{p_2} \bigcap S_{p_3}| \geq 2$. It follows from the definition of superimposed code that there is at least one row x which is incident with p_1, p_2, but not with p_3. Let p_4 be a point distinct from p_1, p_2, p_3 which is incidence with x (since $|L_x| = 4$, we can always choose such a p_4). By our construction, we cannot find a row y such that $c_{yp_1} = c_{yp_2} = 1$ and $c_{yp_3} = c_{yp_4} = 0$, a contradiction. This proves that C is a $2-(8,4,3)$ super-simple design, hence it is a $3-(8,4,1)$ design. □

Theorem 8. *A binary matrix C is a $(2,3)$ superimposed code of size 30×10 if and only if it is the incidence matrix of a $3-(10,4,1)$ design.*

Remark. It is known [3] that 2-(9,3,1), 3-(8,4,1), and 3-(10,4,1) designs are unique. Hence the superimposed codes mentioned in Theorems 6,7, and 8 are unique.

References

[1] K. Chen, "On the existence of super-simple $(v,4,4)$ BIBDs", Journal of Statistical Planning and Inference 51 (1996), 339-350.

[2] K. Chen, "On the existence of super-simple $(v,4,3)$ BIBDs", J.Combin. Math. Combin. Comput., 149-159, 1995.

[3] C.J. Colbourn and J.H. Dinitz, *CRC Handbook of Combinatorial Designs*, CRC Press, Inc., 1996.

[4] K. Engel, "Interval packing and covering in the Boolean lattice", *Combinatorics Prob. and Computing*, **5**(1996), 373-384.

[5] P. Erdos, F. Frankl, and F. Furedi, "Families of finite sets in which no set is covered by the union of r others", *Israel Journal of Math.*, **51**(1985), 75-89.

[6] H.-D.O.F Gronau and R.S.Mullin, "On super-simple $2-(v,4,\lambda)$ designs", *J. Combin. Math. Combin. Comput.*, **11**(1992), 113-121.

[7] H.K. Kim and V. Lebedev, "On optimal superimposed codes", *Journal of Combinatorial Designs*, **12**(2004), 79-91.

[8] H.K. Kim and V. Lebedev, "On optimality of trivial (w,r) cover-free codes", *submitted to Problems of Information Transmission.*

[9] H.K. Kim, V. Lebedev and D.Y. Oh, "Some new results on superimposed codes", submitted.

[10] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*, North Holland, 1983.

[11] C.J. Mitchell and F.C. Piper, "Key storage in secure networks", *Discrete Applied Mathematics,* **21**(1988), 215-228.

[12] N.V. Semakov and V.A. Zinov'ev, "Constant weight codes and tactical configurations", *Problemy Peredachi Informatsii*, **5**(1969), 28-36 (in Russian).

[13] D.R. Stinson, "On some methods for unconditionally secure key distribution and broadcast encryption", *Designs, Codes and Cryptography*, **12**(1997), 215-243.

[14] D.R. Stinson, R.Wei, and L. Zhu, "New constructions for perfect hash families and related structures using combinatorial designs and codes", *J. Combinatorial Designs,* **8**(2000), 189-200.

[15] D.R. Stinson, R.Wei, and L. Zhu, "Some new bounds for cover-free families", *J. Combin. Theory Ser. A,* **90**(2000), 224-234.