Single user tracing superimposed codes

Miklós Csűrös[1]
Dép. Informatique et recherche opérationnelle, Université de Montréal
CP 6128 succ. Centre-Ville,
Montréal, Qué. H3C 3J7, Canada
Email: csuros@iro.umontreal.ca.

Miklós Ruszinkó[2]
Computer and Automation Research Institute, Hungarian Academy of Sciences
Budapest, P. O. Box 63, Hungary-1518
Email: ruszinko@lutra.sztaki.hu.

Abstract — **We investigate the zero-error capacity region of r-out of T user multiple access *OR* channel. A family $\mathcal{F}$ subsets of $[n] = \{1, \ldots, n\}$ is an *r-single-user-tracing superimposed code* (r-SUT) if there exists such a *single-user-tracing function* $\phi: 2^{[n]} \mapsto \mathcal{F}$ that for all $\mathcal{F}' \subseteq \mathcal{F}$ with $|\mathcal{F}'| \le r$, $\phi(\cup_{A\in\mathcal{F}'} A) \in \mathcal{F}'$. In this paper we introduce the concept of these codes and investigate bounds on their rate. We also consider disjointly r-superimposed codes.**

I. Single-user-tracing superimposed codes

Suppose that T users share a common channel. To each of them a binary vector of length n is associated. The i^{th} user transmits its vector $x_i = (x_i^1, \ldots, x_i^n)$ $(i = 1, 2, \ldots, T)$ if it is active, otherwise not. It is assumed that the transmission is bit and block synchronized. The destination of the messages is a single receiver, which gets the bitwise OR vector of the vectors associated to the active users. Moreover, suppose that at most r users are active simultaneously. In the classical framework of superimposed coding, the receiver has to be able to identify the set of all active users from the output vector $\mathbf{y}$ of the channel. That is, the code must satisfy the property that for all choices of $x_1, \ldots, x_k$ and $z_1, \ldots, z_\ell$ of codewords with $1 \le k, \ell \le r$ and $\{x_1, \ldots, x_k\} \ne \{z_1, \ldots, z_\ell\}$, we have $\vee_{i=1}^k x_i \ne \vee_{j=1}^\ell z_j$. The rate of these codes have been studied extensively in e.g., [2, 3, 4].

Here we investigate the case when the receiver has to be able to identify *just one* user out of at most r active ones. Clearly, if a code is superimposed in the classical sense then it satisfies this requirement. A practical motivation for our study of r-SUT families rises from a novel application of combinatorial designs in genomics [1].

As the question is rather of a combinatorial nature, we switch to that terminology. Codewords are characteristic vectors of subsets of the set $[n] = \{1, \ldots, n\}$ with $n > 0$.

Definition I.1 *A family $\mathcal{F} \subseteq 2^{[n]}$ is r-superimposed if $\bigcup_{i=1}^k A_i \ne \bigcup_{j=1}^\ell B_i$ for any $\{A_1, \ldots, A_k\} \ne \{B_1, \ldots, B_\ell\}$, $1 \le k, \ell \le r$; $A_1, \ldots, A_k, B_1, \ldots, B_\ell \in \mathcal{F}$.*

We are interested in *r-single-user-tracing families* (r-SUT), defined as follows.

Definition I.2 *A family $\mathcal{F}$ is r-SUT if for all choices of $\mathcal{F}_1, \ldots, \mathcal{F}_k \subseteq \mathcal{F}$ with $|\mathcal{F}_i| \le r$, $\bigcup_{A\in\mathcal{F}_1} A = \bigcup_{A\in\mathcal{F}_2} A = \ldots = \bigcup_{A\in\mathcal{F}_k} A$ implies $\cap_{i=1}^k \mathcal{F}_k \ne \emptyset$. Equivalently, there exists such a* single-user-tracing function *$\phi: 2^{[n]} \mapsto \mathcal{F}$ that for all $\mathcal{F}' \subseteq \mathcal{F}$ with $|\mathcal{F}'| \le r$, $\phi(\cup_{A\in\mathcal{F}'} A) \in \mathcal{F}'$.*

For every base set size n and r, let $f(n, r)$ denote the maximum size of an r-superimposed family, and $g(n, r)$ denote the maximum size of an r-SUT family. We give bounds on the rate of r-SUT families, which is $R_g(r) = \limsup_{n\to\infty} \frac{\log g(n,r)}{n}$.

Theorem I.3 *$\exists\ c_1, c_2 > 0$ such that $\frac{c_1}{r^2} \le R_g(r) \le \frac{c_2}{r}$.*

II. Disjointly r-superimposed codes

Another important case implicated in the multiple access model of Section 1 is when the receiver must distinguish only between *disjoint sets of active users*. The following definition captures this notion.

Definition II.1 *$\mathcal{F} \subseteq 2^{[n]}$ is disjointly r-superimposed if $\bigcup_{i=1}^k A_i \ne \bigcup_{j=1}^\ell B_i$ for $\{A_1, A_2, \ldots, A_k\} \cap \{B_1, B_2, \ldots, B_\ell\} = \emptyset$; $1 \le k, \ell \le r$; $A_1, \ldots, A_k, B_1, \ldots, B_\ell \in \mathcal{F}$.*

Let $h(n, r)$ be the maximum size of disjointly r-superimposed families.

Lemma II.2 *If $\mathcal{F}$ is r-superimposed then it is r-SUT. If $\mathcal{F}$ is r-SUT, then it is disjointly r-superimposed. Hence, $f(n, r) \le g(n, r) \le h(n, r)$.*

Despite the seemingly slight difference between Definitions II.1 and I.1, the extremal properties of disjointly r-superimposed codes (families) and r-superimposed ones differ substantially.

Theorem II.3 *$\exists\ c_1, c_2 > 0$ such that $\frac{c_1}{r} \le R_h(r) \le \frac{c_2 \log r}{r}$.*

Acknowledgments

We would like to thank Noga Alon and László Györfi for fruitful discussions.

References

[1] M. Csűrös and A. Milosavljevic. Pooled Genomic Indexing (PGI): mathematical analysis and experiment design. In *Proc. Workshop on Algorithms in Bioinformatics*, pp. 10–28, NY: Springer-Verlag 2002.

[2] A. G. D'yachkov and V. V. Rykov. Bounds on the length of disjunctive codes. *Problemy Peredachi Informatsii* vol. 18, no. 3, pp. 7-13, 1982.

[3] W. H. Kautz and R. C. Singleton. Nonrandom binary superimposed codes. *IEEE Trans. Inform. Theory,* vol. IT-10, pp. 363-377, October 1964.

[4] M. Ruszinkó. On the upper bound of the size of the r-cover-free families. *Journal of Combinatorial Theory, Series A* vol. 66, no. 2, pp. 302-310, May 1994.

[1]Research supported by NSERC Grant 250391-02.

[2]Research supported in part by OTKA Grants T038198, T046234. This work relates to Department of the Navy Grant N00014-04-1-4034 issued by the Office of Naval Research International Field Office. The United States Government has a royalty-free license throughout the world in all copyrightable material contained herein.