ISSSTA2004, Sydney, Australia, 30 Aug. - 2 Sep. 2004

On the Design of Complex-valued Spreading Sequences Using a Genetic Algorithm

Hai Huyen Dam*, Hans-Jürgen Zepernick*, and Helge Lüders†
*Western Australian Telecommunications Research Institute
39 Fairway, Nedlands, WA 6907, Australia
E-mail: {dam, hans}@watri.org.au

†Aachen University of Technology
Templergraben 55, 52056 Aachen, Germany
E-mail: helge@fueg.de

***Abstract*— In direct sequence code division multiple access (DS-CDMA) applications, system performance depends on the auto-correlation and cross-correlation properties of the used spreading sequences. Since these two measures come at the expense of each other, their properties and limits need to be investigated when designing DS-CDMA systems. In this paper, we study the trade-off between the out-of-phase average mean-square aperiodic auto-correlation and the average mean-square aperiodic cross-correlation. The problem of minimizing the maximum out-of-phase auto-correlation, the maximum cross-correlation and the maximum nontrivial correlation values are also investigated. A genetic algorithm is proposed to solve these optimization problems. The spreading sequences studied are based on two classes of complex-valued sequences, namely the Oppermann and the modified Walsh-Hadamard sequences. It can be seen from these applications that the genetic algorithm is well suited to efficiently design complex-valued sequences especially when the number of parameters for the optimization problem is large.**

I. Introduction

Direct sequence code division multiple access (DS-CDMA) represents a spread spectrum technique that is used in many wireless communication systems [1]. With DS-CDMA, each active user is assigned a unique sequence or signature, which distinguishes it from other users. The sequences in a spreading code should have low cross-correlation (CC) values to suppress multiple access interference (MAI). The auto-correlation (AC) function of the sequences, on the other hand, should have a narrow peak to avoid inter-symbol interference (ISI) and to enable proper synchronization. Since good AC comes at the expense of the CC properties and vise versa, a trade-off between these two performance characteristics has to be accepted and derived by using efficient optimization techniques.

In this paper, we investigate the design of complex-valued spreading sequences with respect to a combination of different correlation properties. In particular, the trade-off between the out-of-phase average mean-square aperiodic auto-correlation and the average mean-square aperiodic cross-correlation is studied. The problem of designing the sequences with low maximum out-of-phase auto-correlation, low maximum cross-correlation and low maximum nontrivial correlation values are also investigated. The two classes of complex sequences considered are the Oppermann sequences [2], which offer a wide range of correlation properties and the modified Walsh-Hadamard sequences, which have been shown in [3] to offer good correlation properties. Since the number of parameters for the optimization problem is large, especially for the modified Walsh-Hadamard sequences, it is difficult if not impossible to use global optimization methods for solving such problems. Thus, we propose to transform the problems with continuous variables into problems with discrete variables. These problems can then be solved efficiently using a genetic algorithm [4], [5], [6].

The paper is organized as follows. Section II defines the various correlation values for the spreading sequences. Section III describes the Oppermann sequences and the modified Walsh-Hadamard sequences. The genetic algorithm for the design of these sequences with different cost functions is given in Section IV. Section V shows the design examples and Section VI concludes the paper.

II. Correlation Measures for Spreading Sequences

Denote N as the length of each spreading sequence. It has been shown in e.g. [2] that the correlation measures have been focused on aperiodic instead of periodic correlation when performance analysis of wireless communication systems is concerned. The aperiodic correlation between a pair of sequences $\mathbf{u}_k = [u_k(0), \ldots, u_k(N-1)]$ and $\mathbf{u}_l = [u_l(0), \ldots, u_l(N-1)]$, is defined as

$$c_{k,l}(i) = \begin{cases} \frac{1}{N} \sum_{v=0}^{N-1-i} u_k(v) u_l^*(v+i), & 0 \le i \le N-1 \\ \frac{1}{N} \sum_{v=0}^{N-1+i} u_k(v-i) u_l^*(v), & 1-N \le i < 0 \\ 0, & |i| \ge N \end{cases} \quad (1)$$

where '*' denotes the complex conjugate of a complex variable.

For a set of S sequences $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_S\}$, the out-of-phase average mean-square aperiodic auto-correlation for $\mathcal{U}$ is defined as

$$R_{ac}(\mathcal{U}) = \frac{1}{S} \sum_{k=1}^{S} \sum_{i=1-N, i \ne 0}^{N-1} |c_{k,k}(i)|^2 \quad (2)$$

which measures the energy in the side lobes of the auto-correlation for lags different to zero. Similarly, the average mean-square aperiodic cross-correlation for the set of sequences is given by

$$R_{cc}(\mathcal{U}) = \frac{1}{S(S-1)} \sum_{k=1}^{S} \sum_{l=1, l\neq k}^{S} \sum_{i=1-N}^{N-1} |c_{k,l}(i)|^2. \quad (3)$$

In this paper, we study the trade–off between $R_{ac}(\mathcal{U})$ and $R_{cc}(\mathcal{U})$. The associated optimization problem can be formulated as

$$\min_{\mathcal{U}} \left\{ R_{ac}(\mathcal{U}) + \alpha R_{cc}(\mathcal{U}) \right\}, \quad (4)$$

where α is a weighting factor.

Furthermore, we investigate the performance in terms of the maximum out-of-phase auto-correlation value $c_{am}(\mathcal{U})$ and the maximum cross-correlation value $c_{cm}(\mathcal{U})$. These correlation values are defined as

$$\begin{aligned} c_{am}(\mathcal{U}) &= \max_{\substack{1 \le k \le S \\ 1 \le i \le N-1}} |c_{k,k}(i)|, \\ c_{cm}(\mathcal{U}) &= \max_{\substack{1 \le k,l \le S,\ k \neq l \\ 0 \le i \le N-1}} |c_{k,l}(i)|. \end{aligned} \quad (5)$$

The maximum nontrivial correlation value is defined as

$$c_{max}(\mathcal{U}) = \max\{c_{am}(\mathcal{U}), c_{cm}(\mathcal{U})\}. \quad (6)$$

The optimization problems can be formulated as

$$\begin{aligned} &\min_{\mathcal{U}} c_{cm}(\mathcal{U}), \\ &\min_{\mathcal{U}} c_{am}(\mathcal{U}), \\ &\min_{\mathcal{U}} c_{max}(\mathcal{U}). \end{aligned} \quad (7)$$

III. Classes of Complex-valued Spreading Sequences

A. Oppermann Sequences

A family of complex spreading sequences with a wide range of correlation properties is proposed in [2]. For $1 \le k < N$, the $(v-1)^{th}$ element $u_k^o(v-1)$ of an Oppermann sequence $\mathbf{u}_k^o$ is defined by

$$u_k^o(v-1) = (-1)^{k\,v} \exp\left[\frac{j\pi(k^m v^p + v^n)}{N}\right],\ 1 \le v \le N, \quad (8)$$

where m, n and p take real values. For a fixed combination of m, n and p, all the sequences have the same auto-correlation function magnitude. This magnitude depends only on n if $p = 1$. In view of this, we only consider the case $p = 1$. The optimization problem (4) for the set of $N-1$ sequences $\mathcal{U}^o = \{\mathbf{u}_1^o, \cdots, \mathbf{u}_{N-1}^o\}$ can be posed as the following problem with two variables m and n:

$$\min_{m\in[m_1,m_2), n\in[n_1,n_2)} \left\{ R_{ac}(\mathcal{U}^o) + \alpha R_{cc}(\mathcal{U}^o) \right\}, \quad (9)$$

where $[m_1, m_2)$ and $[n_1, n_2)$ are the search regions for m and n, respectively, and N is a prime number. Similarly, the optimization problems in (7) can be posed as

$$\begin{aligned} &\min_{m\in[m_1,m_2), n\in[n_1,n_2)} c_{am}(\mathcal{U}^o), \\ &\min_{m\in[m_1,m_2), n\in[n_1,n_2)} c_{cm}(\mathcal{U}^o), \\ &\min_{m\in[m_1,m_2), n\in[n_1,n_2)} c_{max}(\mathcal{U}^o). \end{aligned} \quad (10)$$

The problems (9)-(10) can be approximated by an integer discrete optimization problem by restricting the continuous parameters m and n within b binary bits. These parameters are transformed to discrete variables as follows:

$$\begin{aligned} m &= (m_2 - m_1)\mathbf{x}_1^T \mathbf{g}(b) + m_1, \\ n &= (n_2 - n_1)\mathbf{x}_2^T \mathbf{g}(b) + n_1 \end{aligned} \quad (11)$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are $b \times 1$ binary vectors and

$$\mathbf{g}(b) = [2^{-1}, \cdots, 2^{-b}]^T.$$

The grid size of the search region for m and n are

$$\Delta_m = (m_2 - m_1)/2^b \text{ and } \Delta_n = (n_2 - n_1)/2^b, \quad (12)$$

respectively, which are small for large b. Thus, the problem (9) can be approximated as the following integer discrete optimization problem:

$$\begin{cases} \min_{\mathbf{x}_1, \mathbf{x}_2} \left\{ R_{ac}(\mathcal{U}^o) + \alpha R_{cc}(\mathcal{U}^o) \right\}, \\ \mathbf{x}_1, \mathbf{x}_2 \text{ are } b \times 1 \text{ binary vectors.} \end{cases} \quad (13)$$

Similarly, (10) can be reduced to the following discrete optimization problems:

$$\begin{aligned} &\begin{cases} \min_{\mathbf{x}_1, \mathbf{x}_2} c_{am}(\mathcal{U}^o) \\ \mathbf{x}_1, \mathbf{x}_2 \text{ are } b \times 1 \text{ binary vectors,} \end{cases} \\ &\begin{cases} \min_{\mathbf{x}_1, \mathbf{x}_2} c_{cm}(\mathcal{U}^o) \\ \mathbf{x}_1, \mathbf{x}_2 \text{ are } b \times 1 \text{ binary vectors,} \end{cases} \\ &\begin{cases} \min_{\mathbf{x}_1, \mathbf{x}_2} c_{max}(\mathcal{U}^o) \\ \mathbf{x}_1, \mathbf{x}_2 \text{ are } b \times 1 \text{ binary vectors.} \end{cases} \end{aligned} \quad (14)$$

B. Modified Walsh-Hadamard Sequences

The second considered class of sequences is based on the modified Walsh-Hadamard matrix $\tilde{\mathbf{H}}_N$ of size $N \times N$, where N is a power of two. This matrix can be obtained from the Hadamard matrix $\mathbf{H}_N$ with

$$\tilde{\mathbf{H}}_N = \mathbf{H}_N \mathbf{D}_N \quad (15)$$

where $\mathbf{D}_N = \mathrm{diag}\{[\exp(j2\pi d_1), \cdots, \exp(j2\pi d_N)]\}$ and $d_k \in [0,1)$, for all $1 \le k \le N$. The N spreading sequences $\mathcal{U}^M = \{\mathbf{u}_1^M, \cdots, \mathbf{u}_N^M\}$ form the rows of the matrix $\tilde{\mathbf{H}}_N$. Similar to the design of the Oppermann sequences, the search for the minimum correlation properties is restricted to q bits for each parameter d_k in the interval [0,1). These parameters can be expressed as

$$d_k = \mathbf{z}_k^T \mathbf{g}(q),\ 1 \le k \le N \quad (16)$$

where $\mathbf{z}_k$ is a $q \times 1$ binary vector and $\mathbf{g}(q) = [2^{-1}, \cdots, 2^{-q}]^T$. The design of the modified Walsh-Hadamard sequences in terms of a minimum of a weighted combination of $R_{ac}(\mathcal{U}^M)$

and $R_{cc}(\mathcal{U}^M)$ can be transformed to the following discrete optimization problem

$$\begin{cases} \min\limits_{\mathbf{z}_1,\cdots,\mathbf{z}_N} \left\{R_{ac}(\mathcal{U}^M) + \alpha R_{cc}(\mathcal{U}^M)\right\}, \\ \mathbf{z}_k \text{ are } q \times 1 \text{ binary vectors}, \ 1 \leq k \leq N. \end{cases} \tag{17}$$

The design of the modified Walsh-Hadamard sequences in terms of maximum correlation characteristics can be formulated similar to (14).

IV. Genetic Algorithm

The problems (13), (14) and (17) are integer discrete optimization problems with binary variables. These problems can be solved efficiently by using the genetic algorithm [4], [5], [6] which is a stochastic search method that mimics the metaphor of natural biological evolution. The advantage of the method is its significant computational saving over other discrete optimization methods. It is noted that the presented genetic algorithm can be easily modified to include performance characteristics other than those suggested above.

The idea of the genetic algorithm can be summarized as follows. In the first step, the algorithm begins with a population of random chromosomes. This population generally has a fixed size L that does not change over the generations. Each member of the population represents a possible solution to the discrete optimization problem. In the second step, the cost functions for all the population members are evaluated and each member is assigned a fitness value for reproduction. Examples of the fitness functions for a population of size L, $\{\mathcal{U}^{o,1}, \cdots, \mathcal{U}^{o,L}\}$ and the two optimization problems (13) and (14) are given as:

$$\begin{aligned} f(\mathcal{U}^{o,l}) &= \frac{1/[R_{ac}(\mathcal{U}^{o,l})+\alpha R_{cc}(\mathcal{U}^{o,l})]}{\frac{1}{L}\sum\limits_{i=1}^{L} 1/[R_{ac}(\mathcal{U}^{o,i})+\alpha R_{cc}(\mathcal{U}^{o,i})]}, \ 1 \leq l \leq L \\ f_{am}(\mathcal{U}^{o,l}) &= \frac{1/c_{am}(\mathcal{U}^{o,l})}{\frac{1}{L}\sum\limits_{i=1}^{L} 1/c_{am}(\mathcal{U}^{o,i})}, \ 1 \leq l \leq L. \end{aligned} \tag{18}$$

The fitness functions $f_{cm}(\mathcal{U}^{o,l})$, $f_{max}(\mathcal{U}^{o,l})$ and the functions for the modified Walsh-Hadamard sequences can be defined similarly.

In the third step, an intermediate population is selected based on the fitness functions (18) by employing the stochastic universal sampling algorithm [6], [8]. Genetic recombination, e.g. crossover and mutation operations is randomly applied to pairs of parents to create offsprings. A new population is obtained by combining the offspring population with the intermediate population. The natural selection process is then applied to select the L strongest members from the new population. The second and third steps are repeated until the variance of the cost function for the generation is less than a small specified value or when certain stoping criteria are met.

V. Design Examples

Consider the design of the Oppermann sequences and the modified Walsh-Hadamard sequences with minimum correlation properties as formulated in equations (13), (14) and (17). The optimum values of (13) and (14) for the Oppermann sequences with length $N = 31$ using the genetic algorithm are shown in Tables I and II. The number of binary bits is $b = 16$ and the number of discrete variables for these optimization problems is $16 \times 2 = 32$. The weighting factor α for Table I is taken as 0, 29.9 and 60. It can be seen from the table that depending on the application, the Oppermann sequences can be designed for a large range of trade–offs between R_{ac} and R_{cc}. The results of the algorithm for the cases of minimizing $c_{am}(\mathcal{U}^o)$, $c_{cm}(\mathcal{U}^o)$ and $c_{max}(\mathcal{U}^o)$ are given in Table II.

TABLE I

Optimized Oppermann sequences with average mean square correlation criteria.

Minimize $R_{ac}(\mathcal{U}^o) + \alpha R_{cc}(\mathcal{U}^o)$	$R_{ac}(\mathcal{U}^o)$	$R_{cc}(\mathcal{U}^o)$
$\alpha = 0$	0.1107	0.9963
$\alpha = 29.9$	7.4847	0.7493
$\alpha = 60$	19.6774	0.3418

TABLE II

Optimized Oppermann sequences with maximum correlation criteria.

Minimize	$c_{max}(\mathcal{U}^o)$	$c_{am}(\mathcal{U}^o)$	$c_{cm}(\mathcal{U}^o)$
$c_{am}(\mathcal{U}^o)$	0.9997	0.0853	0.9997
$c_{cm}(\mathcal{U}^o)$	0.2889	0.1876	0.2889
$c_{max}(\mathcal{U}^o)$	0.2889	0.1876	0.2889

Tables III and IV show the results for the modified Walsh-Hadamard sequences of length $N = 32$ where the number of bits q is chosen as 2 and 3. The number of binary variables for these cases are 64 and 96. It can be seen from the tables that the cost functions reduce slightly with an increase in the number of bits q. In fact, these values cannot be significantly improved for $q \geq 4$. It follows from Table III that the family of modified Walsh-Hadamard sequences does not offer large options for a trade-off between R_{ac} and R_{cc}. Furthermore, the maximum nontrivial correlation value is higher than that for the Oppermann sequences given in Table II.

TABLE III

Modified Walsh-Hadamard sequences with average mean square correlation criteria.

q	Minimize	$R_{ac}(\mathcal{U}^M)$	$R_{cc}(\mathcal{U}^M)$
2	$R_{ac}(\mathcal{U}^M)$ $(\alpha = 0)$	0.5625	0.9819
	$R_{cc}(\mathcal{U}^M)$ $(\alpha = 60)$	2.5625	0.9173
3	$R_{ac}(\mathcal{U}^M)$ $(\alpha = 0)$	0.5489	0.9823
	$R_{cc}(\mathcal{U}^M)$ $(\alpha = 60)$	4.9726	0.8396

Fig. 1 shows the convergence of the genetic algorithm for the case of minimizing $R_{ac}(\mathcal{U}^o)$ applied to the Oppermann sequences with a population size of 400. The algorithm converges relatively fast, within 6 iterations. The plot also shows the magnitude of the auto-correlation and the frequency

TABLE IV

MODIFIED WALSH-HADAMARD SEQUENCES WITH MAXIMUM CORRELATION CRITERIA.

q	Minimize	$c_{max}(\mathcal{U}^M)$	$c_{am}(\mathcal{U}^M)$	$c_{cm}(\mathcal{U}^M)$
2	$c_{am}(\mathcal{U}^M)$	0.4169	0.2380	0.4169
	$c_{cm}(\mathcal{U}^M)$	0.3497	0.3078	0.3497
	$c_{max}(\mathcal{U}^M)$	0.3494	0.3263	0.3494
3	$c_{am}(\mathcal{U}^M)$	0.4436	0.2224	0.4436
	$c_{cm}(\mathcal{U}^M)$	0.3460	0.3102	0.3460
	$c_{max}(\mathcal{U}^M)$	0.3463	0.3330	0.3463

spectra for three sequences taken from the spreading code. It can be seen that the auto-correlation has a distinct peak while the overlapping spectra are high among users.

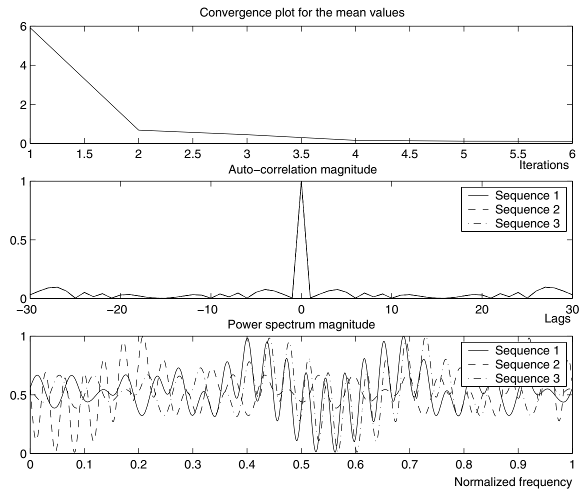

Fig. 1. Convergence and correlation plots for the Oppermann sequences with lowest $R_{ac}(\mathcal{U}^o)$.

Fig. 2 shows the convergence for the mean value of the cost function for the modified Walsh-Hadamard sequences for the case with $q = 3$. The population size is the same as for the Oppermann sequences, $L = 400$. The algorithm converges slower than with the Oppermann sequence since the number of binary variables increased to 96. Since the algorithm converges around 40 iterations, the number of search combinations in this case is significantly smaller than the total number of combinations $2^{96} = 7.9228 \times 10^{28}$. The figure also shows auto-correlation magnitude and frequency spectra for the first three sequences taken from the spreading code with lowest $R_{ac}(\mathcal{U}^M)$. The first two sequences have the same auto-correlation magnitude while the third one has a different magnitude.

VI. CONCLUSIONS

In this paper, a genetic algorithm is proposed for designing complex-valued sequences with optimized correlation characteristics. The trade-off between the out-of-phase average mean-square aperiodic auto-correlation and the average mean-square aperiodic cross-correlation as well as the problem of minimizing the maximum out-of-phase auto-correlation, the maximum cross-correlation and the maximum nontrivial correlation values are investigated. Since the number of parameters for the optimization problem can be large, it is difficult if not impossible to use global optimization methods for solving the problems whereas the genetic algorithm can cope very well with this type of scenarios. The genetic algorithm is applied to the design of Oppermann sequences and modified Walsh-Hadamard sequences. It can be seen from these design examples that the genetic algorithm is well suited to efficiently design complex-valued sequences especially when the number of parameters for the optimization problem is large.

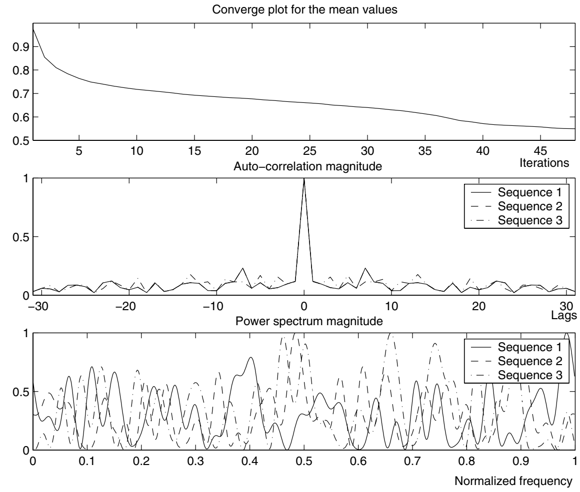

Fig. 2. Convergence and correlation plots for the modified Walsh-Hadamard sequences with lowest $R_{ac}(\mathcal{U}^M)$.

ACKNOWLEDGEMENT

This work was partly supported by the Commonwealth of Australia through the Cooperative Research Centre program.

REFERENCES

[1] R. Prasad and T. Ojanpera, "A Survey on CDMA: Evolution towards Wideband CDMA,"*IEEE International Symposium on Spread Spectrum Techniques and Applications*, vol. 1, pp. 323-331, Sept. 1998.

[2] I. Oppermann and B. S. Vucetic, "Complex Spreading Sequences with a Wide Range of Correlation Properties," *IEEE Transactions on Commununications*, vol. 45, pp. 365–375, 1997.

[3] B. J. Wysocki and T. A. Wysocki, "Modified Walsh-Hadamard Sequences for DS CDMA Wireless Systems," *International Journal of Adaptive Control and Signal Processing*, vol. 16, pp. 589–602, 2002.

[4] N. Ansari and E. Hou, "Computational Intelligence for Optimization," Kluwer Academic Publishers, 1997.

[5] D. E. Goldberg, "Genetic Algorithm in Search Optimization and Machine Learning," Reading, MA: Addison-Wesley, 1989.

[6] H. H. Dam, H.-J. Zepernick, and H. Lüders, "Polyphase Sequence Design using a Genetic Algorithm," *Proceedings of IEEE Vehicular Technology Conference*, 2004.

[7] L. Guo, D.-S. Huang, and W. Zhao, "The Optimization of Radial Basis Probabilistic Neural Networks Based on Genetic Algorithms," *Proceedings of the International Joint Conference on Neural Networks,* vol. 4, pp. 3213-3217, Jul. 2003.

[8] J. Baker, "Reducing Bias and Inefficiency in the Selection Algorithm," *Proceedings of the 2nd International Conference on Genetic Algorithms*, pp. 14-21, Jul. 1987.