

Landscape Properties and Hybrid Evolutionary Algorithm for Optimum Multiuser Detection Problem

Shaowei Wang¹, Qiuping Zhu¹, and Lishan Kang²

¹ School of Electronic Information, Wuhan University, Wuhan, Hubei, 430079, P.R. China

² State Key Laboratory of Software Engineering, Wuhan University, Wuhan, Hubei, 430072, P.R. China.

{shwwang, qpzhu, kang}@whu.edu.cn

Abstract. Optimum multiuser detection (OMD) for CDMA systems is an NP-complete problem. Fitness landscape has been proven to be very useful for understanding the behavior of combinatorial optimization algorithms and can help in predicting their performance. This paper analyzes the statistic properties of the fitness landscape of the OMD problem by performing autocorrelation analysis, fitness distance correlation test and epistasis measure. The analysis results, including epistasis variance, correlation length and fitness distance correlation coefficient in different instances, explain why some random search algorithms are effective methods for the OMD problem and give hints how to design more efficient randomized search heuristic algorithms for it. Based on these results, a multi-start greedy algorithm is proposed for multiuser detection and simulation results show it can provide rather good performance for cases where other suboptimum algorithms perform poorly.

1 Introduction

Multiple access interference (MAI) is the main factor limiting performance in CDMA systems. While optimum multiuser detection (OMD) [1], which is based on the maximum likelihood sequence estimation rule, is the most promising technique for mitigating MAI, its computational complexity increases exponentially with the number of active users, which leads to its implementation impractical.

From a combinatorial optimization viewpoint, the OMD is an NP-complete problem [2]. Randomized search heuristics (RSH) are effective methods for such kinds of problems, so many RSH based multiuser detectors have been studied and exhibit better performance than that of the other linear or nonlinear detectors. Earlier works on applying RSH to OMD problem can be found in [3][4][5][6][7]. The essence of optimum multiuser detection is to search for possible combinations of the users' entire transmitted bit sequence that maximizes the logarithm likelihood function (LLF) derived from the maximum likelihood sequence estimation rule [1], which is called fitness function or objective function in the RSH multiuser detectors. Comparing with so much emphasis on the implementation details and the performance analysis of these algorithms, little attention has been paid on the analysis of statistic characteristics of the OMD problem in terms of combinatorial optimization.

Fitness landscape has been proven to be a power concept in combinatorial optimization theory [8][9][10]. Moreover, the concept can be used to understand the behavior of heuristic algorithms for combinatorial optimization problems and to predict their performance. In this paper, we formulate the fitness landscapes of OMD problem by taking the LLF as the objective function and analyze their local and global characteristics as well as gene interaction properties [11][12]. With the analysis results, we propose a multi-start greedy (MSG) multiuser detector and compare its performance with other detectors.

The remainder of this paper is organized as follows. In Sect.2, we state the OMD problem and construct its fitness landscape. Statistic properties of the fitness landscape, including autocorrelation function, fitness distance correlation and epistasis variance, are described in Sect.3 with analysis results. In Sect.4, we propose the MSG algorithm and compare its performance with others. A short conclusion is given in Sect.5.

2 Optimum Multiuser Detection Problem and Fitness Landscape

2.1 Optimum Multiuser Detection Problem

Assume a binary phase shift keying (BPSK) transmission through an additive-white-Gaussian-noise (AWGN) channel shared by K active users with packet size M in an asynchronous DS-CDMA system (synchronous system is the special case of the asynchronous one). The sufficient statistics for demodulation of the transmitted bits \mathbf{b} are given by MK matched filter outputs \mathbf{y} [13]

$$\mathbf{y} = \mathbf{R}\mathbf{A}\mathbf{b} + \mathbf{n} \quad (1)$$

here \mathbf{A} is the $MK \times MK$ diagonal matrix whose $k + iK$ diagonal element is the k th user's signal amplitude A_k and $i = 1, 2, \dots, M$. $\mathbf{R} \in \mathbb{R}^{MK \times MK}$ is the signature correlation matrix and can be written as

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}[0] & \mathbf{R}^T[1] & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{R}[1] & \mathbf{R}[0] & \mathbf{R}^T[1] & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}[1] & \mathbf{R}[0] & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{R}[1] & \mathbf{R}[0] \end{bmatrix} \quad (2)$$

where $\mathbf{R}[0]$ and $\mathbf{R}[1]$ are $K \times K$ matrices defined by

$$R_{jk}[0] = \begin{cases} 1, & \text{if } j = k; \\ \rho_{jk}, & \text{if } j < k; \\ \rho_{kj}, & \text{if } j > k; \end{cases} \quad (3)$$

$$R_{jk} [1] = \begin{cases} 0, & \text{if } j \geq k; \\ \rho_{jk}, & \text{if } j < k. \end{cases} \quad (4)$$

ρ_{jk} denotes the partial crosscorrelation coefficient between the j th user and the k th user.

The optimum multiuser detection problem is to generate an estimation sequence $\hat{\mathbf{b}} = [\hat{b}_1(1), \hat{b}_2(1), \dots, \hat{b}_K(1), \dots, \hat{b}_1(M), \hat{b}_2(M), \dots, \hat{b}_K(M)]^T$ to maximum the objective function

$$f(\mathbf{b}) = 2\mathbf{y}^T \mathbf{A}\mathbf{b} - \mathbf{b}^T \mathbf{A}\mathbf{R}\mathbf{A}\mathbf{b} \quad (5)$$

which means to search 2^{MK} possible bit sequence exhaustively. It is a combinatorial optimization problem and is proven to be NP-complete [2].

2.2 Fitness Landscape of OMD Problem

The notion of fitness landscape comes from biology, where it is used as a framework for thinking about evolution. Landscape theory has emerged as an attempt to devise suitable mathematical structures for describing the static properties of landscape as well as their influence on the dynamics of adaptation [10].

In formal terms, a fitness landscape consists of three ingredients: A set \mathbf{S} of possible solutions, the notion of element distance d in \mathbf{S} and a fitness function $f : \mathbf{S} \rightarrow \mathbb{R}$. So the fitness landscape of OMD problem can be defined as follows.

1. The solution space $\mathbf{S} : \{-1, 1\}^{MK}$, where K is the number of active users and M is the packet size.
2. If \mathbf{b}_i and \mathbf{b}_j are two elements in \mathbf{S} , $i, j \in \mathbb{N}$, the distance between them is the

$$\text{Hamming distance of the two binary vectors: } d(\mathbf{b}_i, \mathbf{b}_j) = \sum_{n=1}^{MK} b_i^n \oplus b_j^n.$$

3. The fitness function is the LLF defined in equation (5).

3 Statistic Properties Analysis of Fitness Landscape

3.1 Random Walk Correlation, Fitness Distance Correlation and Epistasis Correlation

3.1.1 The Random Walk Correlation Function

To measure the ruggedness of a fitness landscape, Weinberger [8] suggests the use of random walk correlation function

$$r(s) = \frac{\langle f(x_t)f(x_{t+s}) \rangle - \langle f \rangle^2}{\langle f^2 \rangle - \langle f \rangle^2} \quad (6)$$

where $\langle \cdot \rangle$ is expectation and f is fitness. x_t is a time series and $f(x_t)$ defines the correlation of two points s steps away along random walk through the fitness landscape, $-1 \leq r(s) \leq 1$. Then the normalized correlation length l is defined as

$$l = -N \times \ln(|r(1)|) \quad (7)$$

for $r(1) \neq 0$, N is the dimension of \mathbf{S} .

3.1.2 The Fitness Distance Correlation

Fitness distance correlation (FDC) coefficient is another measure for problem difficulty for heuristic algorithms such as evolutionary algorithms [9]. Denote the shortest distance between a possible solution and the global optimum solution as d , the FDC coefficient ρ is defined as

$$\rho = \frac{\langle fd \rangle - \langle f \rangle \langle d \rangle}{\sqrt{(\langle f^2 \rangle - \langle f \rangle^2)(\langle d^2 \rangle - \langle d \rangle^2)}} \quad (8)$$

where $-1 \leq \rho \leq 1$.

3.1.3 The Epistasis Correlation

Denote \hat{f} as the estimation fitness value based on gene decomposition [11], epistasis variance $EpiN$ and epistasis correlation $EpiC$ are defined as [12]

$$EpiN = \frac{\sqrt{\sum_{x \in \mathbf{S}} (f(x) - \hat{f}(x))^2}}{\sqrt{\sum_{x \in \mathbf{S}} f^2(x)}} \quad (9)$$

$$EpiC = \frac{\sum_{x \in \mathbf{S}} (f(x) - \bar{f})(\hat{f}(x) - \bar{\hat{f}})}{\sqrt{\sum_{x \in \mathbf{S}} (f(x) - \bar{f})^2} \sqrt{\sum_{x \in \mathbf{S}} (\hat{f}(x) - \bar{\hat{f}})^2}} \quad (10)$$

where $\bar{f} = \langle f \rangle$, $\bar{\hat{f}} = \langle \hat{f} \rangle$. $0 \leq EpiN \leq 1$, $0 \leq EpiC \leq 1$. Generally, the lower the value of $EpiN$, the higher the value of $EpiC$, and the weaker the epistasis of the fitness function.

3.2 Numerical Results and Discussions

Experiments have been conducted to estimate the coefficients discussed above. E_b is the bit energy, N_0 is the two-sided power spectrum density of Gaussian noise, and K is the number of active users. All results are obtained by performing 10^6 Monte Carlo runs. Table 1 shows the numerical result of 9 different cases.

Table 1. Numerical results of different cases

Case	K	E_b/N_0	l	ρ	$EpiN$	$EpiC$
1	10	3	2.313	-0.911	0.109	0.982
2	10	6	2.317	-0.910	0.110	0.980
3	10	9	2.314	-0.930	0.111	0.979
4	20	3	2.351	-0.870	0.123	0.962
5	20	6	2.225	-0.869	0.124	0.962
6	20	9	2.327	-0.869	0.124	0.958
7	30	3	2.249	-0.846	0.128	0.951
8	30	6	2.310	-0.849	0.128	0.945
9	30	9	2.355	-0.834	0.128	0.944

Though the fitness landscape of OMD problem is dynamic because of the changes of E_b/N_0 and K , from Table 1, we can see the statistics of the fitness landscape (l , ρ , $EpiN$ and $EpiC$) vary within a narrow range. The FDC coefficient ρ keeps negative value close to -1 , which means the LLF is appropriate as the cost function for heuristics based multiuser detector. The normalized correlation length l is almost the same in all cases, which gives hints how to select the neighborhood size in heuristics. For example, $k = 2$ is appropriate for the $k - opt$ multiuser detector [6] and may give the best tradeoff between performance and computational complexity, because l is close to 2.

Table 1 also shows that $EpiN$ and $EpiC$ change very little when E_b/N_0 and K vary within a wide range and the value of $EpiN$ is relatively small, which means the epistasis of the LLF is very weaker and the fitness value of a possible solution is almost determined independently by each bit of the solution vector. Therefore greedy heuristic method may be very efficient local search algorithm for OMD problem [3][5].

4 Multi-Start Greedy Algorithm for Multiuser Detection

4.1 Multi-Start Greedy Algorithm

The MSG algorithm consists of three phases. Phase 1, generate n initial solutions distributed in the whole solution space randomly. This phase is consider as multi-start process which imitates the population of evolutionary algorithms (EAs). Phase 2, perform greedy local search algorithm on each solution and produce n local optima.

This phase is a local search process. Phase 3, choose the best one (based on their fitness) from these local optima as the global optimum solution.

The greedy local search algorithm in phase 2 can be described as follows. Without loss the generality, for a current solution vector $\mathbf{b}^{t-1} = [b_1^{t-1}, b_2^{t-1}, \dots, b_N^{t-1}]^T$, where N is the length of the solution vector, denote the vector with only the j th bit different from \mathbf{b}^{t-1} by $\mathbf{b}_{\text{diff}}^{t-1} = [b_1^{t-1}, \dots, -b_j^{t-1}, \dots, b_N^{t-1}]^T$. Calculate the gain g_j^{t-1}

$$g_j^{t-1} = f(\mathbf{b}_{\text{diff}}^{t-1}) - f(\mathbf{b}^{t-1}) \quad (11)$$

for $j = 1, 2, \dots, N$, here f is defined in equation (5). Then each bit of the next solution vector \mathbf{b}^t can be updated as

$$b_k^t = \begin{cases} -b_k^{t-1} & \text{if } k = \arg\{\max_j g_j^{t-1}\} \text{ and } g_j^{t-1} > 0 \\ b_k^{t-1} & \text{otherwise} \end{cases} \quad (12)$$

This process is repeated until a local optimal is obtained (there are no better solutions that are in its neighborhood).

4.2 Simulation Results

The BER performance of the conventional detector (CD), evolutionary programming detector (EP) [4], gradient guided detector (GGD) [5], parallel interference cancellation (PIC) detector [14] and the proposed MSG is illustrated in Fig.1 by the curves of BER versus E_b/N_0 . The number of users is 10 and 20 ($K = 10, 20$) in Fig.1(a) and (b) respectively and the packet size is 3 ($M = 3$). It is obvious that MSG detector outperforms other detectors.

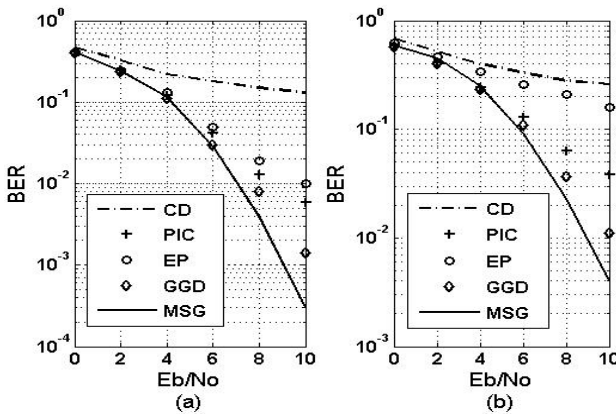


Fig. 1. BER against E_b/N_0 performance of MSG, GGD, EP, PIC and CD for $M = 3$ (a) $K = 10$ and (b) $K = 20$

The computational complexity of MSG can be determined easily. The initialization is to generate n random vector and requires computations of $O(n)$. Each local optimal requires $n \times (n-1)/2$ fitness value computations. So the computational complexity of a local optimal is $O(n^2)$ and the computational complexity of MSG is $O(n^3)$. A fast greedy heuristic is proposed in [15]. With this method, a local optimal can be obtained in $O(n)$. Then the computational complexity of MSG can be reduced to $O(n^2)$.

5 Conclusions

This paper studies the important statistic properties of the fitness landscape of the OMD problem and proposes a multi-start greedy algorithm based multiuser detector. The analysis results about the fitness landscape explain why some heuristics multiuser detector are effective and give hints how to design more efficient randomized search heuristics for the OMD problem. The multi-start greedy algorithm which is founded on these analysis results can provide good performance for cases where other suboptimum algorithms perform poorly.

Acknowledgement. This work is supported by the National Natural Science Foundation of China (No. 60473081).

References

1. S. Verdu: Minimum Probability of Error for Asynchronous Gaussian Multiple-access Channels. IEEE Transactions on Information Theory, Vol.32 (1986) 85-96.
2. S. Verdu: Computational Complexity of Optimal Multiuser Detection. Algorithmica, vol.4 (1989) 303-312.
3. A. AlRustamani, B.R Vojcic: A New Approach to Greedy Multiuser Detection. IEEE Transactions on Communications, Vol.50 (2002) 1326-1336.
4. H.S.Lim, M.V.C.Rao, Alan W.C. Tan, H.T.Chuah: Mulatiuser Detection for DS-CDMA Systems Using Evolutionary Programming. IEEE Communications Letters, Vol.7 (2003) 101-103.
5. Jun Hu, Rick S.Blum: A Gradient Guided Search Algorithm for Multiuser Detection. IEEE Communications Letters, Vol.4 (2000) 340-342.
6. H.S.Lim, B. Venkatesh: An Efficient Local Search Heuristics for Asynchronous Multiuser Detection. IEEE Communications Letters, Vol.7 (2003) 299-30.
7. Peng Hui Tan, Lars K. Rasmussen: Multiuser Detection in CDMA-A Comparison of Relaxations, Exact, and Heuristic Search Methods. IEEE Transactions on Wireless Communicatons, Vol.3 (2004) 1802-1809.
8. E.D. Weinberger: Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference. Biological Cybernetics, Vol.63 (1990) 325-336.
9. Peter Merz, Bernd Freisleben: Fitness Landscape Analysis and Memetic Algorithms for the Quadratic Assignment Problem. IEEE Transactions on Evolutionary Computation, Vol.4 (2000) 337-352.

10. P.F. Stadler: Fitness Landscape. Lecture Notes in Physics, Springer-Verlag Berlin Heidelberg, (2002) 183-204.
11. C. Reeves, C. Wright: An experimental design perspective on genetic algorithms. Foundations of Genetic Algorithms 3, L.D. Whitley and M. D. Vose, Eds. San Mateo, CA: Morgan Kaufmann, (1995) 7–22.
12. Bart Naudts, Leila Kallel: A Comparison of Predictive Measure of Problem Difficulty in Evolutionary Algorithms. IEEE Transactions on Evolutionary Computation, Vol.4 (2000) 1-15.
13. S. Verdu: Multiuser Detection. Cambridge University Press, Cambridge, U.K. (1998).
14. Dariush Divsalar, Marvin K. Simon and Dan Raphaeli: Improved Parallel Interference Cancellation for CDMA. IEEE Transactions on Communications, Vol.46 (1998) 258-268.
15. P. Merz and B. Freisleben: Greedy and local search heuristics for unconstrained binary quadratic programming. Journal of Heuristics, vol. 8 (2002)197–213.