

PERFORMANCE ANALYSIS OF LINKAGE LEARNING TECHNIQUES IN GENETIC ALGORITHMS

R. Lakshmi¹, K. Vivekanandan²

¹ Assistant Professor, Department of Computer Science, Pondicherry University, Puducherry, India, rlakshmiselva@yahoo.co.in

² Professor, Department of Computer Science, Pondicherry Engineering College, Puducherry, India, k.vivekanandan@pec.edu

Abstract

One variance of Genetic Algorithms is a Linkage Learning Genetic Algorithm (LLGA) enhances the efficiencies of Simple Genetic Algorithm (SGA) while solving NP hard Problems. Discovery of Linkage Learning Technique is an important task in GA. Almost all existing Linkage Learning Techniques follow either random approach or probabilistic approaches. This makes repeated passes over the population to determine the relationship between individuals. SGA with random linkage technique is simple but may take long time to converge to the optimal solutions. This paper uses a linkage learning operator called Gene Silencing which is an inspired mechanism from biological systems. The Gene Silencing mechanism is used to improve the linkages by preserving the building blocks in an individual from the disruption of recombination processes such as Crossover and Mutation. It converges quickly to the optimal solution without compromising the diversification on search spaces. To prove this phenomenon, the Travelling Sales Person problem (TSP) has been chosen to retain the order of cities in a tour. Experiments carried out on different TSP benchmark instances taken from TSPLIB which is a standard library for TSP problems. These benchmark instances have also been applied on various linkage learning techniques and analyses the performance of these techniques with Gene Silencing (GS) mechanism. The performance analysis has been made on experimental results with respect to optimal solution and convergence speed.

Index Terms: Linkage Learning, Gene Silencing, Building Blocks, Genetic Algorithm, TSPLIB, Performance Analysis

-----***-----

1. INTRODUCTION

Genetic algorithm is an adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. Genetic Algorithms have proved to be efficient in handling complex NP-Hard problems in the recent past. A wide variety of Genetic Algorithms have been devised to solve a range of problems. Genetic algorithms differ from each other in terms of solution representation, genetic operators used, termination criteria, etc. These existing algorithms, in no way, mark the end of the search for a better Genetic Algorithm, there is always a scope for further improvement of these Genetic Algorithms.

Genetic algorithms work very well if the genes belonging to the same building block are tightly linked together within the chromosome. Otherwise, if these genes are spread all over the chromosome, building blocks are very hard to be created and easy to be destroyed by the recombination operator. Genetic algorithms cannot perform well under such circumstances. Traditional genetic algorithms suffer from the linkage problem and they are unable to learn the linkage among genes. Traditional genetic algorithms require the users to possess prior domain knowledge of the problem so that the genes on chromosomes can be correctly arranged

with respect to the chosen operators. When problem domain knowledge is available, the problem can be solved easily using traditional genetic algorithms with an appropriate chromosome representation. However, when that knowledge is not available, one way to handle the problem is to make the GA capable of adapting and learning the linkages that exist within the chromosome.

The performance of any GA is mainly based on four parameters namely optimal solutions, convergence rate, diversification and error rate. The performance of SGA is enriched by adapting the linkage learning methods inside the genetic algorithms which learns the relationships among genes in chromosomes. There is several linkage learning techniques available for real and binary coded genetic algorithms. Though many linkage learning techniques like Bayesian Optimization Algorithms (BOA) [23], Adaptive Linkage Crossover (ADX) [1], Dependency Structure Matrix (DSM) [10], Estimation of Distribution Algorithm (EDA) [2], Masked Operator (MO) [17], Linkage Identification by Non Linearity Check (LINC) for binary chromosomes [5], Gene Silencing (GS) [9] [20] [21] is also used to improve the linkages exist within the chromosome of any coded representation. So far there is no any

performance comparison made on various linkage learning techniques. The objective of this paper is to compare the efficiency of different linkage learning techniques with the Gene Silencing method. The results proved effective with Gene Silencing and are quite comparable.

In this paper linkage learning has been explored on TSP problem. There are more than 256 solutions of various approaches including mathematical approaches, dynamic programming methods and evolutionary algorithms exist to solve TSP problem. From the literature survey, it has been already proved that solving TSP with enormous data using mathematical models and dynamic programming approaches are more time consuming processes. These approaches may take years of years to return the optimal path when the data set is huge. When genetic algorithm [24] [25] is applied, it converges fast to return the optimal path. The linkage learning genetic algorithm works even better when compared to simple GA. This paper uses gene silencing mechanism to identify the linkages in a chromosome. In the experiments, TSP benchmark instances from TSPLIB library have been used on various linkage learning techniques.

The organization of this paper is as follows: Section 2 discusses the brief study of various linkage learning techniques. Section 3 discusses the experimental results and the performance analysis and Section 4 concludes the paper.

2. LINKAGE LEARNING TECHNIQUES

Linkage learning techniques can be classified according to different facets and aspects like how it can be employed in the algorithm, representation of chromosomes etc. Linkage problem is an ordering problem of the chromosome and addressed to the same issue of building-block (BB) identification or linkage learning. The success of a simple genetic algorithm relies upon good coding scheme that puts genes belonging to the same building blocks together on the chromosome to provide tight linkage of building blocks. Based on that the linkage learning techniques are classified into the following categories:

2.1 Random Linkage Learning Techniques

Random linkage discovery is a costless, effective linkage recognition technique which performs randomly on chromosomes to identify the linkages. In this technique few cities are randomly chosen from a parent and change their loci, we get Offspring. It computes their fitness values and compare to find the linkages on both offspring. Offspring with higher fitness value will contain BBs and have more chances to survive in the next generation. Selection of highly fit individuals to the next generation is performed by

one of the selection techniques. Several linkage learning techniques have been existing in this category such as Inversion Operator [1], Dynamic Linkage Discovery [2] Algorithm and so on.

2.2 Algorithmic Approaches

This is multi metric approach model, other than fitness value it also employs extra measurements to determine the quality (BBs) of individuals in a population. Estimation of Distribution Algorithms (EDA), Bayesian Optimization Algorithm (BOA), Extended Compact Genetic Algorithm (ECGA) and the like [10] [3] [4] [7] are classified as multi metric approaches

2.3 Binary Approaches

These approaches are more suitable for binary coded representation. It does bitwise perturbations in individuals and detects linkage by checking nonlinearity in the individual caused by perturbations. Linkage Identification by Nonlinearity Check (LINC) [5], Linkage Identification by Monotonicity Detection (LIMD) [6], Gene Expression Messy GA (GEMA) [4] and the like are [11] [16] more suitable for binary coded representation.

2.4 Operator Techniques

There are several operator techniques such as Harick's Linkage Learning [19], Linkage Evolving Operator [2], Local Search Method (LSM) [18], and Neighbor Affinity Value method (NAV) [17] of uni metric approaches available in GAs to find linkages exist within the chromosome.

2.5 Gene Silencing Mechanism

The principles of genetic algorithm mimic the process of natural evolution. The concept of gene silencing [9] could be adopted in genetic algorithm as an operator where other natural operators like crossover and mutation are already being applied for any standard GA. Normally, crossover operator disrupts the good building blocks, even though they are considered to be useful to evolve good solutions. Since gene silencing is heritable, the required genes are preserved from the disruptions of crossover or mutation and carried over through successive generations till the optimal solution is reached. In TSP, Gene silencing [15] is used to preserve the order of the cities in a tour to obtain optimal distance and also it makes the GA to converge quickly.

3. EXPERIMENTATION AND ANALYSIS

This paper conducts an analysis on performance of Gene Silencing and compares it with the existing Linkage learning Techniques. The various Linkage Learning methods

including gene silencing mechanism are implemented in MATLAB tool on an i5 system with TSP benchmark instances gr17, wi29, dj38, swis42, att48, eil101, qa194, uy734, d1291, nu3496, eg7146, gr9882, d15112, d18512 obtained from the standard library TSPLIB. The following table is one example of many which illustrate the distance matrix for the benchmark instance gr17.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0	633	257	91	412	150	80	134	259	505	353	324	70	211	268	246	121
1	633	0	390	661	227	488	572	530	555	289	282	638	567	466	420	745	518
2	257	390	0	228	169	112	196	154	372	262	110	437	191	74	53	472	142
3	91	661	228	0	383	120	77	105	175	476	324	240	27	182	239	237	84
4	412	227	169	383	0	267	351	309	338	196	61	421	346	243	199	528	297
5	150	488	112	120	267	0	63	34	264	360	208	329	83	105	123	364	35
6	80	572	196	77	351	63	0	29	232	444	292	297	47	150	207	332	29
7	134	530	154	105	309	34	29	0	249	402	250	314	68	108	165	349	36
8	259	555	372	175	338	264	232	249	0	495	352	95	189	326	383	202	236
9	505	289	262	476	196	360	444	402	495	0	154	578	439	336	240	685	390
10	353	282	110	324	61	208	292	250	352	154	0	435	287	184	140	542	238
11	324	638	437	240	421	329	297	314	95	578	435	0	254	391	448	157	301
12	70	567	191	27	346	83	47	68	189	439	287	254	0	145	202	289	55
13	211	466	74	182	243	105	150	108	326	336	184	391	145	0	57	426	96
14	268	420	53	239	199	123	207	165	383	240	140	448	202	57	0	483	153
15	246	745	472	237	528	364	332	349	202	685	542	157	289	426	483	0	336
16	121	518	142	84	297	35	29	36	236	390	238	301	55	96	153	336	0

Fig -1: Intercity distance table for TSP

Table -1: 1 Experimentation Setup

Genetic Algorithm (GA)		
Sl. No.	Parameters	Values
1	Population Size	100
2	Selection	Tournament Selection with 10%
3	Elitism%	10%
4	Crossover	Two Point Crossover and Partially Matched Crossover
5	Mutation	Swap Mutation
6	Mutation%	5%
7	Termination	Fixed Number (10000) generations.

Genetic Algorithm with its parameters listed above have been applied to different linkage learning techniques such as Local Search method (LSM), Random Linkage Method (RLM) and Harik's Crossover technique. Their optimized solutions are compared with the optimized solutions of Gene Silencing mechanism which is shown in chart 1.

For each linkage learning technique, the obtained optimal distance of different TSP benchmark instances have been collected from 10 runs. For various techniques, the average case values are used in performance analysis.

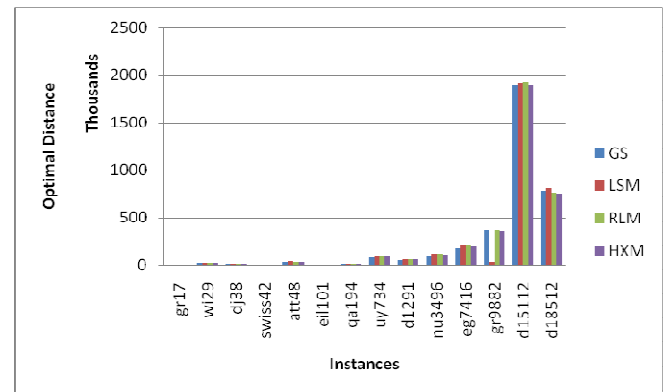


Chart -1: Optimal Distances of TSP instances on Different Linkage Learning Techniques

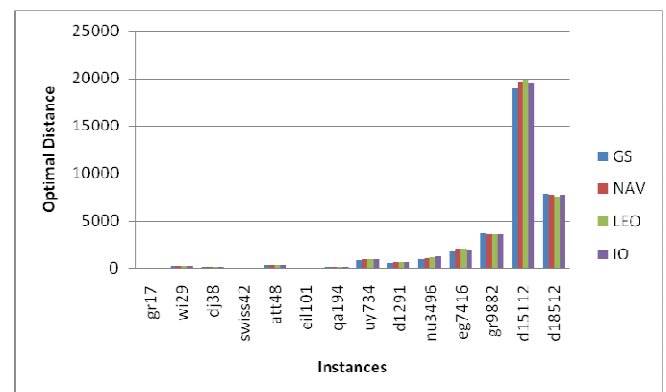


Chart -2: Optimal Distances of TSP instances on Different Linkage Learning Techniques

The two graphs show the experimental results of different linkage learning techniques. The chart 1 shows the shortest distance found by Random method, local search method, harik's crossover method and gene silencing. The chart 2 shows the distance generated by Neighbor Affinity Value, Linkage Evolving Operator, and Inversion Operator and Gene Silencing operator. From the graph it is clearly proved that the gene silencing mechanism outperforms the RLM, LSM, HXM, NAV, LEO and IO.

CONCLUSIONS

The experimental results show that the distance (optimal solution) for all benchmark instances obtained from gene silencing is minimum as compared to the distance obtained from other linkage learning techniques. In this paper, the linkage learning techniques in the context of the traveling salesman problem is discussed. Since crossover and mutation will disrupt the good building blocks in subsequent generations, a new biologically inspired operator, Gene Silencing was applied to preserve those good building blocks which improve the linkages in a chromosome. A computation was made with all linkage learning techniques for the Travelling Salesman problem. From the results it was found that the Gene Silencing in linkage Learning Genetic algorithm performed better than the existing algorithms.

REFERENCES

- [1]. R Ayed A. Salman, Kishan Mehrotra, and Chilukuri K. Mohan, "Adaptive Linkage Crossover"
- [2]. R Ying-ping Chen, A Survey of Linkage Learning Techniques in Genetic and Evolutionary Algorithm
- [3]. R H. Kargupta, "Search polynomial complexity, and the fast messy genetic algorithm. Technical Report 95008, University of Illinois at Urbana Champaign, Urbana, IL, October 1995.
- [4]. R Kargupta, H, "The Gene Expression Messy Genetic Algorithm", Evolutionary Computation, Proceedings of IEEE International Conference on Volume, Issue, and 20-22 Pages: 814 – 819, May 2009.
- [5]. M. Munetomo and D. E. Goldberg, "Identifying linkage by non-linearity check", IlliGAL Report No. 98012, University of Illinois at Urbana-Champaign, Urbana, IL, 2010.
- [6]. M. Munetomo and D. E. Goldberg, "Linkage Identification by Non – Monotonicity Detection for overlapping functions", IlliGAL Report No.99005, University of Illinois at Urbana-Champaign, Urbana, IL, Genetic and Evolutionary Computation Conference, January 2011.
- [7]. Chen, Y.-p., Peng, W.-C., & Jian, M.-c, "Particle swarm optimization with recombination and dynamic linkage discovery", IEEE Transactions on Systems, Vol. 37(6): pp.1460–1470, 2007.
- [8]. David R. Newman, "The Use of Linkage Learning in Genetic Algorithm".
- [9]. S.Siva Sathya, S.Kuppuswami, Department of Computer Science, Pondicherry University, "Gene Silencing for Course Time-Tabling with Genetic Algorithm".
- [10]. Yu, T.-L., Goldberg, D. E., Yassine, A. A., & Chen, Y.-p, "Genetic algorithm design inspired by organizational theory: Pilot study of a dependency structure matrix driven genetic algorithm", in Proceedings of Artificial Neural Networks in Engineering, ANNIE, pp. 327–332, 2003
- [11]. Jian, M.-C., & Chen, Y.-p, "Introducing recombination with dynamic linkage discovery to particle swarm optimization", in Proceedings of ACM SIGEVO Genetic and Evolutionary Computation Conference (GECCO-2006), pp. 85–86, EI, 2006
- [12]. Oliver Kramer, Patrick Koch, "Self-Adaptive Partially Mapped Crossover", proceedings of Genetic and Evolution Computation Conference, pp. 593 – 697, ACM, July 2007.
- [13]. Chen., "A Text book of Fundamentals of Microbiology".
- [14]. Gene Silencing Technique offers new strategy for treating, curing disease, Science Daily, 2005.
- [15]. Lakshmi R, Vivekanandan K. Gene Silencing in Linkage Learning Genetic Algorithm, page 2, conference proceeding of NCFC, 2012.
- [16]. Chen Y-P, Goldberg DE., "An analysis of a reordering operator with tournament selection on a GA-hard problem", Lecture Notes in Computer Science (LNCS) 2003; 2723: 825–836.
- [17]. Corcoran AL, Roger L, Wainwright, "Reducing disruption of superior building blocks in genetic algorithms", Proceedings of the, ACM SIGAPP Symposium on Applied Computing, Feb 2003.
- [18]. Ray SS, Bandyopadhyay S, Pal SK. New Operators of Genetic Algorithms for Traveling Salesman Problem, Evolutionary Computation CEC2004;2(1):1457-1464.
- [18]. Murata T, Miyata S. Gene Linkage Identification in Permutation Problems for Local Search and Genetic Local Search', Systems, Man and Cybernetics. IEEE International Conference 2005; 2(10-12): 1920 – 1924.
- [19]. G. R. Harik, "Learning gene linkage to efficiently solve problems of bounded difficulty using Genetic Algorithms", University of Michigan, Ann Arbor, IlliGAL Report No. 97005, 1997
- [20]. R.Lakshmi et.l , "A New Biological Operator in Genetic Algorithm for Class Scheduling Problem", International Journal of Computer Applications (0975 – 8887), Volume 60– No.12, December 2012.
- [21]. R.Lakshmi , K.Vivekanandan, "Interference Induced Silencing in Travelling Sales Person Problem using Linkage Learning Genetic Algorithm", IJESR, March 2013/ Volume-3/Issue-3/Article No-12/2670-2679 ISSN 2277-2685.
- [22]. Domingo Ortiz-Boyer, Cesar Hervas Martnez CIXL2: A Crossover Operator for Evolutionary Algorithms Based on Population Features.
- [23]. Martin pelican et.al,"BOA: The Bayesian Optimization Algorithm", Illigal Report No.99003, jan 2009.
- [24]. Kalyanmoy Deb, Amarendra Kumar, "Real coded Genetic Algorithms with Simulated Binary Crossover : Studies on Multimodal and Multi objective Problems
- [25]. Andrej Kazakov, "Travelling Salesman Problem: Local Search and Divide and Conquer working together", Independent Research Review, University of Southampton, March 2009