Improve and Compare the Performance of Genetic Algorithm for PSP

Chandra Akash Kiran, Deepak Xaxa

M.Tech Student, Dept. of CSE, School of Engineering and IT, MATS University, Raipur, India,

Assistant Professor, Dept. of CSE, School of Engineering and IT, MATS University, Raipur, India.

ABSTRACT: Genetic Algorithm (GA) is very helpful for Predicting the protein structure for small as well as larger protein sequence. In this research work it is attempted to improve and compare the performance of Genetic Algorithm by using modified GA for large protein sequence of PSP problem. The performance of genetic algorithm basically depends on its various specifications like initialization, selection, crossover, mutation and replacement etc. In this paper applied three selection methods are rank selection, elitism selection, tournament selection and three crossover methods are single point crossover, double point crossover and uniform crossover. The experimental result shows that proposed modified Genetic Algorithm, combination of Rank Selection with Uniform Crossover provides better performance in most of the cases then other combination of GA methods.

KEYWORDS: Genetic algorithm, Protein structure prediction, selection method, crossover method

I. INTRODUCTION

Proteins assume a critical part in life forms as they perform numerous vital undertakings in their cells. As needs be, the better see how to proteins are fabricated, the better you can manage numerous regular sicknesses. Specifically, data on basic property of proteins can give close into the way they work and how changes, for example, might influence their operability. Such information can thus help and impact present day prescription and medication improvement. A protein is a succession of amino corrosive atoms. Gatherings of amino corrosive succession that are identified with a 3D structure are particular as kept structure. To pick this subset, we anticipated use of a straightforward hereditary calculation.

II. RELATED WORK

Luis C. Sheep et.al.[1] proposed a half and half GA to escape from nearby minima which is joined with an organized populace and hybridized with a way relinking strategy which makes a difference. Mahmood A Rashid, Md. Tamjidul Hoque, M.A. Hakim Newton, Duc Nghia Pham, and Abdul Sattar [2] proposed another hereditary calculation for foreseeing F. Enakshi Sar is with the West Bengal University of Technology, protein structure utilizing HP Lattice model. In 2008, Xiaolong Zhang, Wen Cheng [3] proposed an enhanced tabu pursuit calculation in off-cross section AB model. R. F. Mansour[4] proposed a half and half calculation which is mix of hereditary calculation in light of network coding and tabu hunt calculation in off-cross section AB model. Xin Chen, Mingwei Lv, Lihui Zhao and Xudong Zhang [5] consolidate demand flight and molecule swarm improvement (PSO) to unravel protein structure forecast taking into account 3D AB off-grid model. Anshul Sharma, Anuj Mehta [6], proposed various choice strategy for hereditary calculation and portrayed different determination techniques. Poonam Sharma, Amit Wadhwa, Komal [7], proposed relative investigation of choice plans for tackling a streamlining issue in hereditary calculation and assesses their execution. Omar Al Jadaan, Lakishmi Rajamani, C. R. Rao [8], proposed change of roulette wheel choice strategy to expand the increase of assets, unwavering quality and differences and abatement the instability in determination process. Shalini Singh, Ejaz Aslam lodhi [9], proposed variety in voyaging sales representatives issue utilizing hereditary calculation procedure and look at the administrator of sought after methodology which give the best result for finding the most brief way in a briefest time for moving toward the objective. They acquired the ideal separation with the visit length in a more powerful manner. Rakesh kumar, jyotishree [10], proposed looked at two choice technique mixing roulette wheel choice and rank choice with various issue size. Xunbo shuai, Xiangguang Zhou

[11], proposed two hereditary administrators double administrator and opposite administrator for nonlinear advancement issues.

III. PROPOSED ALGORITHM

Essential Genetic Algorithm (GA) is for the most part made out of two procedures. The principal procedure is determination of people for the creation of the cutting edge and the second process is control of the chose people to frame the cutting edge by hybrid and transformation strategies. The determination component figures out which people are decided for mating (generation) and what number of posterity each chose singular produces. The fundamental standard of choice system is "the better is an individual; the higher is its shot of being guardian." Generally, hybrid and transformation investigate the hunt space, though determination diminishes the pursuit territory inside of the populace by tossing poor arrangements.
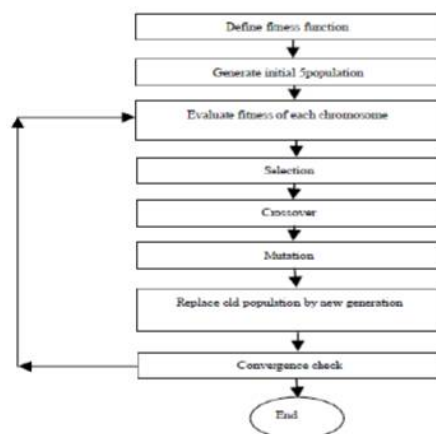

Figure 1 Flowchart for genetic algorithm

a. Population

Cleary depicted the populace as an aftereffect of a solitary cycle of hereditary calculation. Emphasis can make another populace. Populace contains an arrangement of chromosomes; every chromosome is one finished conceivable answer for the issue to be tackled utilizing hereditary calculation. Parameters utilized as a part of this examinations is populace size, hybrid likelihood is 0.9 and transformation likelihood is 0.1

b. Evaluations

For every chromosome there is a wellness capacity used to assess the wellness of every chromosome. Wellness' worth mirrors the nature of every chromosome.

c. Encoding

The quality is an issue parameter; it can be encoded as a parallel, whole number, or buoy number.

d. Selections

It is the procedure of selecting the chromosomes to apply Steady State Genetic Algorithm. Sorts of choice are:

I. Rank Selection

Rank Selection positions the populace and each chromosome gets wellness from the positioning. The most noticeably awful has wellness 1 and the best has wellness N. It results in moderate union yet avoids too fast merging. It additionally keeps up determination weight when the wellness fluctuation is low. It jam differences and subsequently prompts an effective inquiry. In Linear Rank determination, people are doled out subjective wellness in light of the rank inside of the populace. The people in the populace are sorted from best to most noticeably awful as per their

wellness values. Every person in the populace is alloted a numerical rank in view of wellness, and choice depends on this positioning as opposed to contrasts in wellness.

II.   Elitism Selection

The thought here is to organize the chromosomes in the diminishing request as indicated by their wellness values. At that point apply the choice with every two chromosomes in the masterminded set. Thusly, Genetic Algorithm will be connected between solid chromosomes or between powerless chromosomes. This implies there is no opportunity to apply Genetic Algorithm in the middle of feeble and solid chromosomes. Elitism is a sort of determination in which the best individual went to the cutting edge all things considered with no adjustment. Elitism keeps the best individual to experience the proliferation prepare to pass them with no change into cutting edge.

III.   Tournament Selection

GAs utilizes a choice instrument to choose people from the populace to embed into a mating pool. People from the mating pool are utilized to produce new posterity, with the subsequent posterity shaping the premise of the cutting edge. A determination instrument in GA is just a procedure that supports the choice of better people in the populace for the mating pool. The determination weight is the extent to which the better people are favoured: the higher the select particle weight, the more the better people are favoured. This determination weight drives the GA to enhance the populace wellness over succeeding eras. The merging rate of a GA is to a great extent dictated by the determination weight, with higher choice weights bringing about higher union rates. In any case, if the determination weight is too low, the joining rate will be moderate, and the GA will pointlessly take more time to locate the ideal arrangement. In the event that the determination weight is too high, there is an expanded shot of the GA rashly merging to a wrong (imperfect) arrangement. Competition determination gives choice weight by holding a competition among s contenders, with s being the competition size. The champ of the competition is the person with the most noteworthy wellness of the s competition contenders. The champ is then embedded into the mating pool. The mating pool, being included competition victors, has a higher normal wellness than the normal populace wellness. This wellness distinction gives the determination weight, which drives the GA to enhance the wellness of each succeeding era. Expanded determination weight can be given by just expanding the competition size s, as the victor from a bigger competition will, by and large, have a higher wellness than the champ of a littler competition.

e.   Crossover

This procedure is utilized to trade qualities between chromosomes to make posterity. Sorts of hybrid are:

i.   Single Point

Select the hybrid point inside of a chromosome haphazardly and trade the two guardian chromosomes right now to create two new offspring's.

ii.   Two Points

Select two focuses haphazardly and trade the two guardian qualities between these focuses.

iii.   Uniform

By likelihood, hybrid will choose the guardian commitment in the posterity chromosome. In the event that the blending proportion is equivalent to 0.5 this implies half of qualities in the posterity will originate from guardian 1 and the other will originate from guardian 2.

f.   Mutation

This procedure will change the estimation of haphazardly chose quality.

g.   Replacements

This procedure will contrast between a few chromosomes with pick the best.

h.   Stopping Criterions

Beginning with a starting populace, the advancement procedure is rehashed until the fulfillment of the end condition. Some basic ending conditions, for example,

- The discovered arrangement fulfills the base measure.
- An altered number of eras came to.

- Allocating spending plan (ex: time, cash) came to.
- Successive emphasess no more create better

IV. PSEUDO CODE

Step 1: [Start] Generate random population of N chromosomes (suitable solutions for the problem) and set initial parameters.

Step 2: [Fitness] Evaluate the fitness f(x) of each chromosome x in the population.

Step 3: [Selection] Select two parent chromosomes from a population according to their fitness (the better fitness, the bigger chance to be selected)

Step 4: [Crossover] with a crossover probability cross over the parents to form a new offspring (children). If no crossover was performed, offspring is an exact copy of parents.

Step 5: [Mutation] with a mutation probability mutate new offspring at each locus (position in chromosome).

Step 6: [Loop] Go to step 2

V. SIMULATION RESULTS

No. of amino acid	Genetic algorithm operators	single point crossover		double point crossover		uniform crossover	
	Selection method	Time	Average energy	Time	Average energy	Time	Average energy
30	Rank	9.8410	1.1464	13.4580	3.1916	8.6247	1.3764
	Elitism	9.5362	2.6768	17.7999	1.1552	8.7135	2.9334
	Tournament	12.1820	4.5226	16.3078	4.7551	7.5675	5.1527
34	Rank	9.8020	1.5153	15.7102	3.6220	6.9342	1.5109
	Elitism	16.3517	2.2033	11.2185	2.1366	7.0721	3.0355
	Tournament	10.5675	9.2449	24.3950	9.2607	6.5341	1.7463
38	Rank	12.5047	2.9326	14.3270	1.8315	7.1178	1.7337
	Elitism	10.1421	5.7930	11.5063	1.4711	7.5468	6.1320
	Tournament	10.1212	7.5459	13.2597	2.4477	7.1687	4.6972
44	Rank	9.3380	1.3651	12.0752	5.1961	6.7818	1.0548
	Elitism	9.6351	9.4127	12.1516	1.1357	7.2592	2.3896
	Tournament	9.2870	2.0261	11.0731	9.9995	6.4008	3.5759

Table 1 Comparative study selection method with crossover method of GA

In table 1 show that comparative study between various selection method and crossover method of genetic algorithm. Here clearly show that rank selection with uniform crossover gives best result in most of the cases.

VI. CONCLUSION AND FUTURE WORK

The aim of this work was to show that performance of Genetic Algorithm in the field of protein structure prediction problem. According to study in table 1 observe that rank selection with single point crossover gives the best result for

all cases. In table 2 observe that Elitism selection with double point crossover gives the best result for all cases. In table 3 observe that rank selection with uniform crossover gives the best result for all cases. So result shows that rank selection with uniform crossover gives the best result for protein sequence size 34 and 44 and rank selection with single point crossover gives the best result for protein sequence size 30 and Elitism selection with double point crossover gives the best result for protein sequence size 38. When compare the energy values, which have been obtained from various combinations of genetic algorithm methods. It is observing that Genetic algorithm with combination of rank selection and uniform crossover outperforms than other method in minimizing energy. In future work we can use more effective GA operators.

REFERENCES

1. M. Dorn, L. S. Buriol, L. C. Lamb, "A Hybrid Genetic Algorithm for the 3D Protein Structure Prediction Problem using a Path-ReI inking Strategy, "IEEE Congress on Evolution Computation, 2011.
2. Md. A. Rashid , Md. T. Hoque, M.A. H. Newton, D. N. Pham, and A. Sattar, "A New Genetic Algorithm for Simplified Protein Structure Prediction, " Australasian Conference on Artificial intelligence, vol. 7691, pp.107-119, 2012.
3. X. Zhang, W. Cheng, "Protein 3D Structure Prediction by improved Tabu Search in Off-Lattice AB model, " iCBBE, 2nd international Conference, pp.11 04-11 09, 2008.
4. R. F. Mansour, "Applying an Evolutionary Algorithm for protein Structure prediction, " American Journal of Bioinformatics Research, pp. 18-23, 2011.
5. X. Chen, M. Lv, L. Zhao and X. Zhang, "An Improved Particle Swarm Optimization for Protein Folding Prediction, " MECS Press, vol. I, pp. 1-8, February 2011.
6. Anshul Sharma, Anuj Mehta, "Review Paper of Various Selection Methods in Genetic Algorithm", *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 3, Issue 7, July 2013
7. Poonam Sharma, Amit Wadhwa, Komal, "Analysis of Selection Schemes for Solving an Optimization Problem in Genetic Algorithm", *International Journal of Computer Applications (0975 – 8887)* Volume 93 – No.11, May 2014
8. Omar Al Jadaan, Lakishmi Rajamani, C. R. Rao, "IMPROVED SELECTION OPERATOR FOR GA", *Journal of Theoretical and Applied Information Technology*
9. Shalini Singh, Ejaz Aslam Lodhi, "Study of Variation in TSP using Genetic Algorithm and Its Operator Comparison", *International Journal of Soft Computing and Engineering (IJSCE)* ISSN: 2231-2307, Volume-3, Issue-2, May 2013
10. Rakesh Kumar, Jyotishree, "Blending Roulette Wheel Selection & Rank Selection in Genetic Algorithms", *International Journal of Machine Learning and Computing*, Vol. 2, No. 4, August 2012
11. Xunbo Shuai, Xiangguang Zhou, "A Genetic Algorithm Based on Combination Operators", *Procedia Environmental Sciences* 11346 – 350, 2011.
12. K.F.Lau,K.A.Dill.A "Lattice Statistical Mechanics Model of the Conformation and Sequence Space of Proteins [J]".Macromolecules,22(10):3968-3997,1992.
13. K.A.Dill,H.S.Chan. From Levinthal to pathways to funnels [J]. "Nature Structural & Molecular Biology". 4(1):10- 19, 1997.

BIOGRAPHY

Chandra Akash Kiran, is a M.Tech CSE, School of Engineering and IT, MATS University, Raipur, India,

Mr. Deepak Xaxa is a Assistant professor, Dept of CSE, School of Engineering and IT, MATS University, Raipur, India.