Tournament Selection by Segmentation

Egbenimi Beredugo Eskca

Abstract— Several attempts to find a suitable approximation to the travelling sales person(TSP) problem have resorted to the use of meta heuristic optimization techniques to obtained a practical solution within an acceptable time frame. The Genetic algorithm(GA) is one of the most frequently used evolutionary algorithms in application of meta heuristics. Central to the efficacy of the genetic algorithm is the idea of selection- The process of choosing individuals within a confined population set to produce a new generation. Several selection methods have been proposed and used in the application of GA; Tournament, Roulette, and Rank selection are some of the common selection mechanism employed in GA. This paper presents a Novel idea that broadens the Tournament selection technique. Whereas, this paper employs some aspect of the classic Tournament selection, the technique proposed by this paper introduces a novel tournament criterion in the selection of suitable population members for crossing. This paper is focused on the application of this new approach called the Grouped Attribute Tournament(GAT) or Tournament by Segmentation (TBS)Selection in solving the TSP problem. This new approach considers every node as an attribute of the tour. Tour attributes are grouped to form tour property or character component(CC). In TBS selection, tournament among population members, is based on the average value of attributes set collection(ASC), or the average value of the individual's character component. This paper proposes a CC based tournament selection in which the selection process targets the weakest node group of the population member.

Indexed Terms—Genetic algorithm, Tournament selection, character component, grouped attribute, segmentation, weak link

1. INTRODUCTION

In formulating the GAT selection, this paper considered the natural tendency of animal species to seek out reproductive partners possessing traits and attributes that they presumed can enhance the quality and survival of their offspring. For descriptive convenience, this paper will henceforth refer to animals and other species with selective and discriminatory tendencies in their observed reproductive behaviors simply as species. With some specie colonies, some members with height disadvantage may consciously seek reproductive partners with more than average height with the intention of producing offspring without or with a milder height disadvantage [1]. Scientifically backed testing and experimentations in more complex populations like human colonies are not able to ascertain an absolute fitness value of an individual. Most species are not concerned with nor are interested in an absolute fitness value of their potential reproductive partners. Since species do not have their entire life time to seek out mating partners, and because they are primarily concerned with their survival and the continuity of their kind, they seek mating partners that possess quality they consider complimentary to theirs in ensuring the survival of their offspring [1]. Most variations of the tournament selection method of the GA had resorted to the use of the overall fitness of the population members as the basis of the tournament. Although this has produced rather significant results, it does not take into cognizance the fact that in natural habitations, species have no way of evaluating the complete fitness of their fellow individuals [1]. The classic tournament selection consists of a random sampling of n population members from an active population of P members. A tournament ensues amongst the n selection of population member with the absolute fitness as the yardstick of the tournament. The member with the adjudge best fitness is selected as the winner of the tournament and is placed in the mating pool awaiting its mating partner to emerge from another random tournament selection. The partners so selected are crossed to produce the offspring for the next generation of the population. Another importer component of the tournament selection scheme is the number of n members of the population considered for the tournament. This number simply referred to as the tournament size plays a significant role in the outcome of the tournament selection process. By adjusting the tournament size, the selection pressure- the probability of the selection process to be biased in favor of population members with better fitness; can be influenced. Because an initial sorting of the population is not required for tournament selection, the time complexity for a tournament of size n is O(n) [2], [3], [4], [5], [6], [7], [8]. In proposing the GAT selection, this paper aims to maintain the time complexity of the classic tournament selection while smoothening the selection pressure, giving every member of the tournament an equally likely probability of being selected. The GAT selection considers every single trait that constitute the population member's character component as significant enough to influence the member's absolute fitness. In doing so, the GAT selection, allows for a small subset of the individual's total collection of traits to determine the individual's suitability for mating. Whereas the classic tournament selection requires a potential mating partner to be randomly selected from the population to participate in a mate tournament selection based on the absolute fitness of the participating members; the GAT selection randomly select tournament members from the active population, in a bid to select not the individual with the best absolute fitness, but the individual whose specific character component value meets the tournament preference. This paper had experimented with tournament winners whose weak character components are the least among contestants, and with tournament winners with the strongest weak character component.

The presentation of the rest of this paper is structured into the following sections: Section 2 outlines the Goals and Basis of the Grouped Attribute Selection. Section 3 gives a detailed description of the GAT selection. Section 4 discusses some experimental data of GAT selection and classic tournament selection. Section 5 provides conclusion narrative and recommendation for future research.

2. Goal

It is the goal of this paper to demonstrate the significance of a more granular selection criteria for the tournament selection process. The classic tournament selection is based on the absolute fitness of the population member. This approach to the tournament selection has served the classic tournament as well as several of its offshoots well over the years. However, this paper takes a closer look at the selection criteria, revealing the absolute fitness as a collection of multiple attribute. In considering the cellular significance of each attribute, this paper shows that the suitability of an individual for selection can be determined by a subset of attributes, rather than a consideration of the entire gamut of attributes of the population member. Because a subset of attributes rather than the absolute fitness of the population member is used as the tournament criteria, every contestant in the tournament selection process is equally likely to be selected as a potential mating partner. this technique has the added advantage of smoothening out the selection pressure. This paper experimented with different values of attribute set collections and recommends the most appropriate character component value for the crossing strategy adopted by this paper. A description of the grouped attribute or character component tournament selection process is given in the next section.

3. TOURNAMENT SELECTION BY SEGMENTATION

Important Definitions

A. Character Component(CC): Is the collection or group of nodes that defines a segment of fitness in a tour. In Figure 1, A, B, and C are examples of CCs with the smaller arrows in the group representing the nodes.

B. Average Component Fitness: This is the average of the individual fitness values of the nodes that constitutes a CC. In figure 1, tour X, CC A, it is the average of (5, 7, 9).

C. Component Fitness: The sum of the node values that constitute a character component.

D. Tuning: A variation of selection in which the contesting population member with the least average component fitness value is preferred for crossing.

E. Enhancement: A variation of selection in which the contesting member with the highest component fitness average becomes the winner of the tournament.

From this point onwards, this paper will be using the terms Grouped Attribute Tournament and Tournament by Segmentation (TBS) interchangeably. In the grouped attribute tournament selection process, emphasis is laid on the significance of the value of a node that constitutes a tour (in the case of the TSP problem). The consideration of an attribute (single tour node) in the determination of the victor in the tournament process is based on the recognition that each attribute uniquely contributes towards the overall fitness of the individual. This instantly give credence to this paper's argument that: a population member with a large and unsuitable fitness value might not necessarily be unsuitable for selection in consideration for crossing. In large collection of attributes, a small subset of attribute collection might add up to significantly influence the overall fitness of the attribute collection. This paper approaches the problem by segmenting the tour into subsets of nodes called character components. The value of each character component is evaluated by adding the value of the nodes that forms the segment. The sum of the nodes values that constitute a character component is called the component fitness. In the tournament by segmentation, the average of the component fitness is the basis of the tournament. This paper presents two variations of the tournament by segmentation technique: in the first variation, the winner of the tournament, is the contesting population member whose weakest CC has the least average component fitness among the contestants (tuning);

For instance, if two population members A and B each containing four CCs are in a tournament for selection, if A's weakest CC (CC with the greatest sum of nodes) is Y and B's weakest CC is X, the segment (between X and Y) with the smallest average component fitness will be the winner of the tournament. while the second variation selects the member whose weakest CC has the highest average component fitness as the winner of the tournament (enhancement).

1. Formation of Segments

The entire tour length is divided into three, four or five equal segments. If equal segments are not possible, the last segment gets the extra node. This paper experimented with three and four segments. There are no criteria in segment formation other than the preference for segments with equal nodes which is the case with a tour with even number of nodes.

The value of the nodes in the segment so formed are summed and the average of each segment evaluated. The average of nodes in a segment are used during the selection.

While tuning and enhancement both aims at strengthening the weakest CC of the selected winner during the crossing phase, the strategies of the two methods are uniquely dissimilar. The objective of tuning is to select a member with a relatively good

ISSN 2229-5518

fitness with the aim of further strengthening the already strong weak CC with the promise of producing offspring that are better than the parents. Enhancement on the other hand is aimed at selecting members with one remarkably weak CC in course of the tournament, aiming to enhance the fitness of the member by strengthening the weakest CC. This has the effect of substituting the links of the weak CC in the crossing phase, with the probability of producing an offspring whose weakest CC is better than that of its parent.

The effect of tuning on the general population is such that it results in a population with an increasing polarization between very fit members, and the weak members of the population. Enhancement on the hand, has the effect of smoothening the fitness of members across the population; resulting in a population that tends towards members with fitness falling within a similar range.

Algorithm- Tournament by Segmentation(TBS)

- Segment the tour by attempting to divide it into four equal parts, if all parts can't contain equal number of nodes, add extra node to last segment.
- Evaluate segments sum, by add the value of each node that forms the segment, and store the sum as a property of the type representing the CC. Do this four all population members.
- Evaluate the average component character by diving the segment total by the total number nodes that constitute the segment and store separate for each member.

Tournament by Segmentation(Tuning)

Set tournament size (number of tournament selections to be made).

- Randomly select any two population members.
- select the worst CC of the selected members and extract the evaluated average component character of each member and compare.
- The winner of this comparison will be the member with the smallest average component character. Select the member with the smallest average component character as the winner of the tournament.
- Continue comparison until tournament count is exhausted.
- The winner of the tournament is the member with the **least** average CC in the tournament.

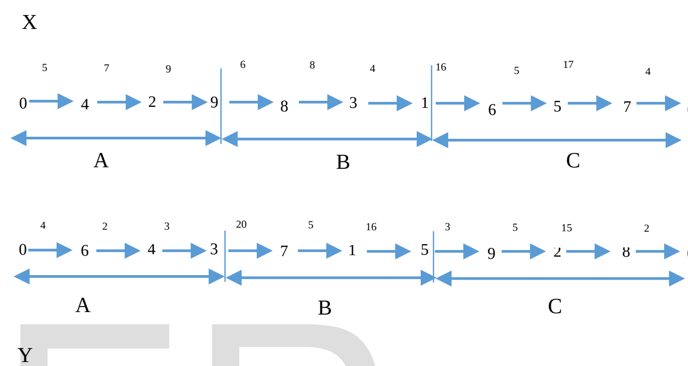
Tournament by Segmentation(Enhancement)

Set tournament size (number of tournament selections to be made).

- Randomly select any two population members.
- Compare the worst CC of the selected members and extract evaluated average component character.
- The winner of this comparison will be the member with the **largest** average component character. Select winner.
- Continue comparison until tournament count is exhausted.
- The winner of the tournament is the member with the **highest** average CC in the tournament.

Return selected winners for crossing.

Figure 1.



In figure 1 above, tour X and Y are representations of two population members taking part in a tournament by segmentation. Tour X's weakest CC is segment C while tour Y's weakest CC is segment B. The figures between the arrowed lines are the nodes of the tour, while the figures on the lines represents the node values. Because the average component fitness of tour X (10.5) is smaller than that of Y (13.7), in the course of a tournament by segmentation between X and Y, X will emerge as the winner if the tuning strategy is employed. This enhances tour X's probability of strengthening its weak CC during the crossing phase and producing a fitter offspring in the next generation. A classic tournament selection would have selected tour Y as the winner of the tournament since tour Y has smaller overall fitness (75) as compared to X (85) overall fitness. A tournament by segmentation employing the enhancement strategy will have Y as the winner, since Y's weakest CC has an average component fitness larger than that of X. This signals Y as a weaker member of the population whose fitness needs to be enhanced by enhancing the component fitness of its weak CC. These strategies allow all population members to participate in the tournament selection irrespective of their overall fitness value. The next section discusses the experimental performances of these two strategies and a comparison between these strategies and the classic tournament selection.

4. EXPERIMENTAL DATA

This section shows the experimental results of first twenty runs out of five hundred runs for the purpose conserving space in this report, for each of:

classic tournament selection, tournament by segmentation(tuning), tournament by segmentation(enhancement). The experiment was based on a GA algorithm that utilizes the island model [9]. Three islands where run in parallel, each using a different tournament size.

Tournament sizes of four, five and six were randomly used for each island. At every iteration of an island instance, a random number generator was used to produce either four, five or six as the tournament size. Every island did five hundred iterations, at the end, the best member of each island is crossed with other islands and used to replace the weakest member of each island before the beginning of the next iteration. The island communicated their results to a central controller which does the work of crossing the results of one island to another and flags off the commencement of the next iteration. The tabulations below show columns for the selection type, crossover algorithm, best fit from each run, time to converge. Runs where carried out on 52, 93, 700 and 1250 cities models. The experiment was carried out using the partially matched crossover(PMX).

Table 1.

Runs	SelectionType-TBS	Cross-Over Type	Best fit	Converge Time(Min)
1	Enhancement	PMX	10041	3
2	Enhancement	PMX	8817	2
3	Enhancement	PMX	9560	2
4	Enhancement	PMX	9862	3
5	Enhancement	PMX	10486	3
6	Enhancement	PMX	8986	3
7	Enhancement	PMX	9591	3
8	Enhancement	PMX	9648	3
9	Enhancement	PMX	8581	2
10	Enhancement	PMX	9427	3
11	Enhancement	PMX	9100	3
12	Enhancement	PMX	9138	3
13	Enhancement	PMX	8551	3
14	Enhancement	PMX	9905	3
15	Enhancement	PMX	8353	3
16	Enhancement	PMX	9036	3
17	Enhancement	PMX	9823	3
18	Enhancement	PMX	8330	3
19	Enhancement	PMX	9183	3
20	Enhancement	PMX	8466	3

Table 2.

Runs	Selection Type-TBS	Cross Over Type	Best fit	Converge Time(Min)
1	Tuning	PMX	10900	3
2	Tuning	PMX	9898	3
3	Tuning	PMX	9197	3
4	Tuning	PMX	9293	3
5	Tuning	PMX	9745	3
6	Tuning	PMX	8109	3
7	Tuning	PMX	8907	2
8	Tuning	PMX	9907	3
9	Tuning	PMX	9536	2
10	Tuning	PMX	8159	3
11	Tuning	PMX	5529	3
12	Tuning	PMX	8322	3
13	Tuning	PMX	9470	3
14	Tuning	PMX	9212	3
15	Tuning	PMX	8115	2
16	Tuning	PMX	8053	2
17	Tuning	PMX	8083	3
18	Tuning	PMX	10114	3
19	Tuning	PMX	8808	3
20	Tuning	PMX	9645	3

Table 3.

Runs	Selection Type Classical	Cross Over Type	Best fit	Converge Time(Min)
1	Tournament	PMX	9571	3
2	Tournament	PMX	9867	3
3	Tournament	PMX	9418	3
4	Tournament	PMX	8926	3
5	Tournament	PMX	9496	3
6	Tournament	PMX	9624	3
7	Tournament	PMX	9636	3
8	Tournament	PMX	9810	3
9	Tournament	PMX	9138	3
10	Tournament	PMX	10060	3
11	Tournament	PMX	8846	2
12	Tournament	PMX	8624	3
13	Tournament	PMX	8425	2
14	Tournament	PMX	8679	2
15	Tournament	PMX	9322	3
16	Tournament	PMX	9177	3
17	Tournament	PMX	9524	2
18	Tournament	PMX	9596	3
19	Tournament	PMX	9292	3
20	Tournament	PMX	10307	3

Table 4.

Selection Type	Cross over Type	Best fitness	Converge Time(min)	Number of cities
Enhancement	PMX	8330	3	52
Tuning	PMX	5529	3	52
Classic Tournament	PMX	8425	2	52

Table 5.

Selection Type	Cross over Type	Best fitness	Converge Time(min)	Number of cities
Enhancement	PMX	10222	4	93
Tuning	PMX	8654	3	93
Classic Tournament	PMX	10621	4	93

Table 6.

Selection Type	Cross over Type	Best fitness	Converge Time(min)	Number of cities
Enhancement	PMX	13926	8	700
Tuning	PMX	10201	5	700
Classic Tournament	PMX	14701	9	700

Table 7.

Selection Type	Cross over Type	Best fitness	Converge Time(min)	Number of cities
Enhancement	PMX	15315	10	1250
Tuning	PMX	12300	8	1250
Classic Tournament	PMX	18108	12	1250

Table 1 shows the results of selection by tournament by segmentation(enhancement) while table 2 and table 3 shows the results of tuning and the classic tournament selection respectively. Tables 4, 5, 6, and 7 shows converge time by number of cities as well as the best fitness for all 500 runs per route.

The overall result of the experimentation shows the both variation of the tournament by segmentation performing better than the classic tournament selection. Although there seem to be marked difference in the pattern of the final best fitness produced by the three selection methods, there isn't a marked difference in the time taken to converge to the best fitness in each run except for the tuning variation. In all the selection by segmentation(tuning) out performs its enhancement counterpart as well as the classic tournament by selection.

5. CONCLUSION AND RECOMMENDATION

Because of the consideration of the significance of all node values in contributing to the final fitness value of the route, this paper has been able to demonstrate that a selection process based on a targeted segment of the individual's trait collection is a significant factor in determining the individual's suitability for a crossover process. This paper has experimented on a new tournament selection process based on the weakest segment of the individual's character component; by selecting(tuning) population members whose weakest contiguous section has the smallest average CC and selecting(enhancement) members whose CC has the strongest average CC, this paper has been able to produce a selection process that consistently outperform the classic tournament selection by more than a tight margin using the partially mapped crossover technique. By exploiting the fact that every population member is only as fit as its weakest component character, the selection by segmentation has proven to be an efficient selection strategy.

Although this paper acknowledges the significance of the segment sum which is a consequence of the segment size, the number of segments that are most appropriate for any given number of nodes was not determined by this paper. However, this paper recognizes that smaller segment sizes that amounts to greater segment sum would serve to mask the inherent weak contiguous section of the individual CC, as each segment tends to be stronger (smaller in fitness value) as the segment size decreases. Future research could be directed to determining a definitive relationship between the node value and the appropriate segment size.

REFERENCES

[1] D. S. Wilson, "What is wrong with absolute individual fitness?" *Trends in Ecology and Evolution* Vol.19 No.5 May 2004.

[2] K. Di Lu et al., "Comparison of Binary Coded Genetic Algorithms with Different Selection Strategies for Continuous
Optimization Problems", Wenzhou University, Wenzhou, 325035, China, 2013 IEEE.

[3] H. Xie, and M. Zhang, "Parent Selection Pressure Auto-Tuning for Tournament Selection in Genetic Programming", *IEEE transactions on evolutionary computation*, Vol. 17, no. 1, Feb. 2013.

[4] J. Zhong, X. Hu, M. Gu and J. Zhang, "Comparison of Performance between Different Selection Strategies on SimpleGenetic Algorithms", IEEE *international conference on computational intelligence for modelling, control and automation*, 2005.

[5] M. Chakraborty, and U. K. Chakraborty, "Branching process analysis of linear ranking and binary tournament selection in genetic algorithms", *Journal of computing and information technology*, 1999, pp. 107-113.

[6] T. Blickle and L. Thiele, "A comparison of selection schemes used in genetic algorithms", Swiss federal institute of technology, Dec. 1995.

[7] J. S. Velazco, J. A. Bullinaria, "Gendered selection strategies in genetic algorithms for optimization", University of Birmingham, Birmingham, B15 2TT UK.

[8] N. Saini, "Review of Selection Methods in Genetic Algorithms", *International journal of engineering and computer science,* vol. 6 Dec. 2017, pp. 22261-22263.

[9] D. Whitley, S. Rana and R. B. Heckendorn, "The island model genetic algorithm: on separability, population size and convergence", Journal *of computing and information technology*, 1999, pp. 33-47.