GENETIC ALGORITHM WITH DIFFERENT SELECTION STRATEGIES FOR ROUGH SET ATTRIBUTE REDUCTION PROBLEMS

GADEER MAHMOOD ALATHAMNEH

PROJECT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF COMPUTER SCIENCE

FACULTY OF INFORMATION SCIENCE AND TECHNOLOGY
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2018

ALGORITMA GENETIC MENGGUNAKAN STRATEGI PEMILIHAN YANG BERBEZA DENGAN SET KASAR UNTUK MASALAH PENGURANGAN ATRIBUT

GADEER MAHMOOD ALATHAMNEH

PROJEK YANG DIKEMUKAKAN UNTUK MEMENUHI SEBAHAGIAN DARIPADA SYARAT MEMPEROLEHI IJAZAH SARJANA SAINS KOMPUTER

FACULTI TEKNOLOGI DAN SAINS MAKLUMAT
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2018

DECLARATION

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged.

04 July 2018 GADEER MAHMOOD ALATHAMNEH

GP05855

ACKNOWLEDGEMENT

First of all, praise is to Allah the Almighty, who is recognized without being seen and who creates without trouble, I have managed to finish my master thesis.

My heartfelt gratitude first goes to my supervisor Prof. Dr. Salwani Abdullah, for her constructive criticisms and advice in the writing of this thesis. I truly appreciate all the time she spent reading my chapters and offering suggestions for revisions that greatly improved the quality of this thesis.

I would like also to thank all the lecturers and staff of FTSM who gave me a great deal of knowledge and helped me in need. I am grateful for the support of my parents, my sister and brother for their patience through my study all over the way to finish my work.

Special thanks go to unforgettable friends for their encouragement and support never stopped in many ways. Thank you for believing in me; you have shown me kindness and made me Du'aa, even at the end.

ABSTRACT

The Attribute reduction is considered as one of the vital topics that have been the attention for the studies that consider the actual data intricacy. The process of the attribute reduction is aiming at finding a minimum attribute set from an actual set of attributes. It is denoted as an NP-Hard optimization problem in ordinary. In the optimization field, optimization algorithms are established to efficiently address NP-Hard problems. In information systems, optimization algorithms attempt to find the minimum attributes from a large attribute set. This approach is famed by virtue of its utility in knowledge discovery and data mining. The plentiful studies managed to utilize meta-heuristic methods to address the attribute reduction problems has promoted this research to suggest an improved one population meta-heuristic method. Genetic Algorithm (GA) holds sundry genetic operators modifiable for improving certain implementations' performance. These operators comprise of selection, crossover, and mutation. In this proposed method, a comparison of the performance of GA in Attribute Reduction based Rough Set utilizing different selection strategies is held i.ie. Roulette Wheel, Tournament, and Rank selection. The experiment on different selection strategies was performed on eighteen datasets from the public domain available in UCI repository. The results demonstrated that the tournament selection strategy performed better than the roulette-wheel and rank-based selection strategies. And other published meta-heuristic algorithms.

ABSTRAK

Pengurangan Atribut dianggap sebagai salah satu topik penting yang menjadi perhatian untuk kajian yang mempertimbangkan kerumitan data sebenar. Proses pengurangan atribut bertujuan untuk mencari atribut minimum yang ditetapkan dari set asal atribut yang sebenar. Ia disifatkan sebagai masalah pengoptimuman NP-Sukar. Di dalam bidang pengoptimuman, algoritma pengoptimuman ditubuhkan untuk menangani masalah NP-Sukar dengan cekap. Di dalam sistem maklumat, algoritma pengoptimuman cuba mencari atribut minima dari set atribut yang besar. Pendekatan ini terkenal dengan kebolehgunaannya dalam penemuan pengetahuan dan perlombongan data. Terdapat kajian berleluasa yang berjaya menggunakan kaedah meta-heuristik untuk menangani masalah pengurangan atribut telah mempromosikan penyelidikan ini untuk mencadangkan satu kaedah meta-heuristic populasi yang lebih baik. Algoritma genetik (GA) berupaya sebagai pengendali genetik yang boleh diubah suai untuk meningkatkan prestasi pelaksanaan tertentu. Pengendali ini terdiri daripada pemilihan dan mutasi. Dalam proses GA, pemilihan dianggap sebagai salah satu operasi yang penting. Proses pemilihan memainkan peranan penting dalam mencari penyelesaian kepada penumpuan awal yang berlaku akibat kekurangan kepelbagaian populasi. Akibatnya, proses pemilihan penduduk di setiap generasi sangat penting. Dalam kaedah yang dicadangkan ini, perbandingan prestasi GA dalam pengurangan atribut berdasarkan Set kasar menggunakan strategi pemilihan yang berbeza diadakan. Eksperimen pada strategi pemilihan yang berlainan dilakukan pada 18 dataset dari domain awam yang terdapat di repositori UCI untuk pembelajaran mesin. Keputusan ujikaji menunjukkan pemilihan strategi secara T*ournament* menghasilkan kualiti keputusan yang lebih baik berbanding dengan strategi pemilihan secara *Roulette wheel* dan *Rank*, serta lebih daripada kaedah meta-heuristik yang sedia ada. kejohanan mencapai hasil yang lebih baik berbanding dengan hasil lain dalam tinjauan literatur.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Great Deluge	GD
Genetic Algorithm	GA
Simulated Annealing	SA
Attribute Reduction	AR
Tabu Search	TS
Whale Optimization Algorithm	WOA
Ant Colony Optimization	ACO

CHAPTER I

INTRODUCTION

1.1 RESEARCH BACKGROUND AND MOTIVATION

Most datasets start to store a big number of attributes due to the internet's fast growth and information technology. In this case, at data pre-processing stage, attribute reduction step is considered to be taken. Decreasing the quantity of attributes in datasets is useful in many areas such as knowledge discovery. However, attribute reduction is a process of selection attempts to explore small subset of original set of attributes with least loss of information. The selected subset of attributes not only should be necessary and sufficient (avoiding redundancy) to express concepts of the target but also keep the representation of the original attributes (Mining 2013).

In fields such as data mining and machine learning, their databases established with a large number of attributes oftentimes come across. The existence of irrelevant and redundant attributes leads to exhausting the computing resources and seriously affecting the process of decision making. For these reasons, eliminating the redundant or irrelevant information and making the data set short are considered to be truly important. Besides, attribute reduction is called feature selection, that is performed for an information system refinement, has been widely researched (Zheng et al. 2014).

In the Rough Set Theory, attribute reduction is considered to be one of the most significant subjects. It is an approach of creating an optimum subset from a system to represent a particular dataset efficiently. It carries out a significant job in shrinking the size a problem for classification and clustering problems. Finding all minimum attribute reductions is deemed an NP_ hard problem due to the intricacy of real-life data. Over the past years, attribute reduction domain was a hot research area, researchers had a great attention to apply meta-heuristic algorithms to locate the

optimum solution and exhibit some successful signs such as vague simulated annealing and genetic algorithm, ant colony, scatter search, tabu search, hyper-heuristic, composite neighbourhood structure and great deluge algorithm (Arajy et al. 2014).

Rough Set Theory, for the first time, was introduced by (Zdzislaw Pawlak & Sets 1991), (Zdzisław Pawlak 1982) for approximating a vague set of concepts by a twosome of precise concepts known as the upper and lower approximations. Recently, immense technical evolution in storage capability, computer technology's interconnectivity and processing power creates tremendous amounts of digital data. Thus, a new important field in computer science known as data mining has emerged. Data mining is the process of data analysis to extract useful knowledge.

Genetic algorithm is one among the countless algorithms used this problem e.g., tabu search (TSAR), scatter search (SSAR), simulated annealing (SimRSAR), ant colony (ACORA and AntRSAR). To address this problem, many versions of GA have been proposed. The invention of GA optimization algorithm was inspired by the evolution process and the natural selection. The basic idea of the GA lies in the method of encoding, crossover, fitness function, selection, and mutation operations(Jaddi & Abdullah 2013).

Recently, to solve the problem of feature selection, researchers have been proposed many meta-heuristic methods. Each algorithm has an appointed specification with diverse settings of parameters. For example, methods applied to feature selection problem can be located in (Tabakhi et al. 2014) who proposed an algorithm of unsupervised ant colony based feature selection algorithm. A matrix-based approach for computing set of reducts and approximations of a covering decision information system is proposed by Tan et al. (2015). In 2016, Jing et al , presented a method of an incremental attribute reduction for feature selection by utilizing the granularity of information in decision systems possessing attribute variation. In 2017, Ge et al presented two general algorithms for reduction using proportional discernibility in discordant decision tables . Another example is (Pacheco et al. 2017), where the authors proposed a new feature selection algorithm based on attribute clustering and

Rough Set Theory for unsupervised data.

1.2 PROBLEM STATEMENT AND RESEARCH QUESTIONS

The attribute reduction process is concerned with locating minimum reducts from an information system. It demands a production of all reducts, and chooses the best with minimum cardinality. It is considered as an NP-hard problem (Polkowski & Skowron 1998). When handling real-world data that can have losses of information and errors, the problem becomes more complicated. Moreover, when the data has a big size, for finding a solution to the problem, a longer time is demanded.

To address this problem, several meta-heuristic approaches have been utilized, such as genetic algorithm, tabu search and simulated annealing. Nevertheless, the available techniques here are not able to solve all data sizes. Some techniques work well when applied on some datasets but perform worse when applied on other datasets.

Genetic Algorithm (GA) is one of the popular algorithms among the different EA. GA uses both crossover and mutation operators makes its population more diverse and thus move immune to be trapped in a local optimum. In theory the diversity also helps the algorithm to be faster in reaching the global optima since it will allow the algorithm to explore the solution space faster.

In GA, the step of indentifying the suitable selection technique is ticklish. The selection process is significant in settling premature convergence that takes place due to the diversity deficiency in the population. hence, population selection is substantial in each generation . The different selection strategies utilised in the process of GA will substantially impact the algorithm’s performance in a different way.

GA has diverse genetic operators that might be adapted to enhance the performance of certain implementations. These operators comprise crossover, selection of parents and mutation. In the GA process, selection is a significant operation. There are many strategies for selection. However, the classical selection strategy proposed in the original GA work by Holland is the Roulette wheel selection,

as the predictable value (the estimated number of times of being selected) of an individual is proportionate to its fitness. However, roulette wheel selection is bias in selecting good solution, thus can cause a poor diversity of solution.

In the case of theoretically infinite population in Roulette Wheel selection strategy, each expected value of each individual in proportion will be allocated to its fitness. However, if the real application of GA has a comparatively small population size, the selected factual number of an individual could be totally different from its anticipated value. In the case of worst scenario, an excessive spin series of the roulette wheel can “allocate all the offspring to the worst individual in the population” (Mitchell 1998). One more limitation of roulette wheel strategy is that whole fitness values of the objective function should be positive. The minimal value of the function is substantial, namely: adding a constant to all the fitness values, a scaling technique that has to be harmless will modify the expected individuals' values.

To cope with the disadvantage of roulette wheel selection, this research is concerned with employing more than one selection strategy. Rank-based and Tournament selection strategies were used to get more efficient solution. This research is aiming at the comparison of GA performance in solving attribute reduction process using different parent strategies of selection such as Tournament selection, Roulette wheel selection and Rank-based selection.

Research Questions

Based on the previous discussion in the problem statement, the aim of this study is to answer the following research questions:

i. How can the GA algorithm with different strategies assist in finding a feature subset more appropriate than the original one?

ii. How can the GA algorithm with different strategies avoid easily reaching the local optimum?

1.3 RESEARCH OBJECTIVES

This research is aiming at the investigation of the effect of the selection strategies in Genetic algorithm towards the quality of the final solution in the rough set attribute reduction problem. In order to achieve this aim, the objectives are outlined as follows:

i. To develop the rank based and tournament selection strategies in GA in order to maintain the diversity of chosen solution and avoid a premature convergence.
ii. To compare the performance of the proposed selection strategy with the standard selection strategy (i.e. roulette wheel).

1.4 RESEARCH SCOPE

Attribute reduction is known as a search technique for an adequate subset of attributes that are strongly corresponded with a decision attribute, referred as minimum reducts sometimes. The minimum reducts might be defined in terms of redundancy and relevancy. In rough set theory, a relevant attribute is the one that predicts the decision attribute. on the other hand, the one that does not predict the decision attribute is irrelevant, and the highly correlated attribute with the other attributes is said to be redundant.

This research focuses on obtaining near-optimal reducts by the iterative algorithm for improvement to solve attribute reduction problems. For evaluating the proposed approach's performance, a standard 18 benchmark datasets that are downloaded from UCI (https://archive.ics.uci.edu/ml/datasets.html) repository. The datasets are different with regards the number of objects and attributes. The results of the experiments are compared among the approaches that are proposed in this research work and compared with other approaches addressed the same problems in the literature. The proposed approaches' performance is evaluated in terms of the minimal reducts.

The relevancy between the selected attributes is based on the dependency degree calculated using the RST. The dependency degree with value 1 shows that the

selected attributes are at the highest relevancy.

1.5 APPROACH OF RESEARCH

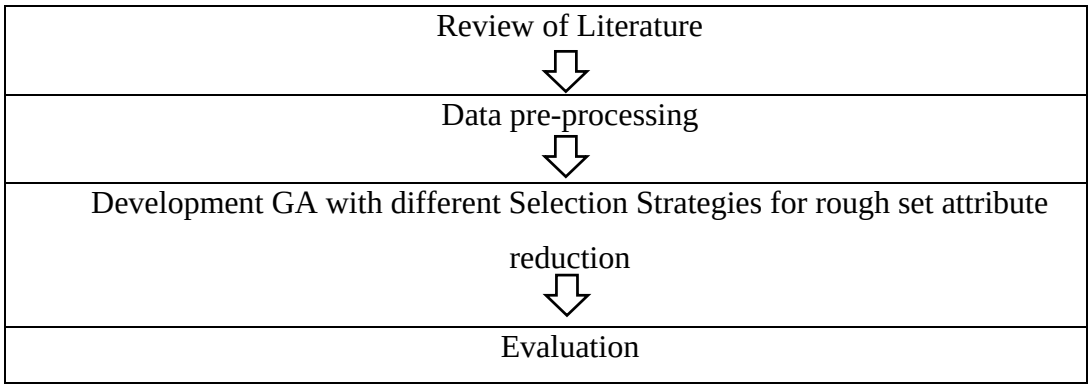

Figure 1.1 Research methodology

The current research is conducted in five phases as follows:

Phase 1: Review of literature

In the literature review, the primary studies of rough set attribute reduction problem, namely: GA approach, were comprehensively reviewed and written in order to understand the background of the problem. The problem has been formulated and compared with the existing approaches.

Phase 2: Data Pre-processing

The phase of pre-processing data focuses on downloading the required datasets from UCI repository, and then reforming them into a structured format - after assigning random values (based on the datatype of the dataset) to the missing data. The determination of a formal model of “Rough set theory” is then set. This structured format will be used in the next phase.

Phase 3: Development Genetic Algorithm with different Selection Strategies for rough set attribute reduction

In this phase, improvement algorithm structure is proposed for rough set attribute reduction. three selection strategies, rank based, roulette wheel and tournament selection strategy were developed and applied on the GA. This study uses C++ programming which is applied together with GA. This algorithm starts with a random initial population.

Phase 4: Evaluation

The evaluation phase is aiming at the comparison of the performance of the different selection strategies used and to compare the proposed approach with the state-of-art approaches.

1.6 THESIS OVERVIEW

This thesis consists of V chapters. This chapter starts with the research background and motivation, problem statement and research questions, objective of the research, research scope and approach of research that presents the steps to be pursued to carry out the current research. The rest of the thesis is organized as the following:

Chapter II introduces a review of the attribute reduction and approaches that are available in the literature. A description of the rough set theory as a measurement tool for calculating the degree of the dependency is presented in chapter 2 section 2.3 as well.

In Chapter III, the proposed approach is presented in detail. There are three selection strategies used to solve the attribute reduction problem i.e. Rank-based selection, Roulette wheel selection, and Tournament selection.

The experimental results attained from the proposed approaches are presented in Chapter IV, where the experiment is performed on 18 renowned UCI datasets with various numbers of objects and attributes. The comparison with the state-of-art methods is presented in this chapter as well.

Lastly, Chapter V concludes overall the work and the future work’s direction.

CHAPTER II

LITERATURE REVIEW

2.1 INTRODUCTION

Over the years, many researchers around the world, investigated and studied the problem of attribute reduction, and widely explored diverse methodologies. Many recent techniques and ideas, which were successfully performed to tackle various NP-hard problems, are presently being utilized to solve the problems in attribute reduction.

This chapter is contained of five sections. Section 2.2 presents a brief explication of the attribute reduction problem. Section 2.3 presents in detail, the rough set theory, which is utilized as a tool of measurement for evaluating the attained minimum reducts. Selection Strategies are presented in Section 2.4. The overview that summaries the published approaches utilized in dealing with the problem of attribute reduction is presented in Section 2.5. Section 2.6 illustrates diversity and selection strategies in the Literature. Finally, section 2.7 provides a summary to this chapter.

2.2 ROUGH SET ATTRIBUTE REDUCTION PROBLEM

Among the problems in rough set theory (RST), attribute reduction problem is a major one. However, Attribute reduction is considered as a tool for extracting the beneficial information from a domain without deforming the right meaning of the included knowledge. This filter has a mechanism that works on detecting subsets that have the minimum number of related attributes existing in the original datasets (known as minimum reducts) where the residual attributes are able to be removed with most tenuous loss of information. This process can deal with different formats of the

attribute values i.e. real or symbolic-values. Finding minimum reducts depending on attribute reduction is beneficial for filtering the datasets by eliminating noisy and vague data that can be utilized later in other areas of application. Still, these filtered datasets are able to be utilized in the process of data mining. Moreover, this helps in enhancing the performance of the process of data mining and produce results of better quality. One of the particular usages in the rough set theory, is attribute reduction in datasets. A main advantage of utilizing RST as stated in (Jensen & Shen 2004) is that:

"Rough set analysis requires no additional parameters to operate other than the supplied data. It works by making use of the granularity structure of the data only".

2.3 CONCEPT OF ROUGH SET ATTRIBUTE REDUCTION

RST is a mathematical method to analyse ambiguity, uncertainty and vagueness in a big dataset. During the decision making process, RST uses sets' approximation, called *upper* and *lower* set's *approximation* (Pawlak 1982, 1991).

Table 2.1 Example of Dataset

$x \in U$	$f1$	$f2$	$f3$	$f4$	d
1	2	1	1	0	3
2	0	1	1	0	4
3	0	1	1	0	4
4	0	1	0	0	4
5	0	1	0	4	5
6	0	1	0	4	5
7	0	1	1	0	4
8	0	1	1	0	4

An information system consists of a pair $S = (U, F)$, where a non-empty finite set of objects U is denoted as the universe, and F is a non-empty finite set of attributes such that $f: U \rightarrow Vf$, for every $f \in F$. The set Vf is called the domain. An information system in RST is similar to a dataset in the tasks of clustering and unsupervised

machine learning. An information system of the form $S = (U, F, d)$, where d is the decision attribute is called a decision system. In a supervised learning and classification, a dataset can be deemed as a decision system where the instances are the objects of universe, and attributes are the elements of F and labels that represent values of decision attribute (Eskandari & Javidi 2016).

For any set $B \subseteq F \cup \{d\}$, we define the B-indiscernibility relation as:

$$INDIS(B) = \{(x, y) \in U \times U | \forall f \in B, f(x) = f(y)\} \quad (2.1)$$

For the dataset of Table 2.1, if $B = \{f3, f4\}$, then objects 4 is indiscernible; as are objects 1,2,3,7,8 and 5,6 are indiscernible. U/B is as follows: U/B = {{4}, {1,2,3,7,8}, {5,6}}.

Two essential concepts of rough sets are the upper and lower approximations of sets. Let $X \subseteq U$ and $B \subseteq F$, the B-upper and B-lower approximations of X are defined as follows:

$$\underline{B}\,X = \{x | [x]B \subseteq X\} \quad (2.2)$$

$$\overline{BX} = \{x | [x]B \cap X = \emptyset\} \quad (2.3)$$

The $\overline{BX}$ and $\underline{B}X$ approximations define information contained in B. If $x \in \underline{B}X$, it particularly belongs to X but if $x \in BX$, it may or may not belong to X. For example, let $B = \{f3, f4\}$ and $X = \{1, 2, 5, 4, 6\}$, then

$\underline{B}\,X = \{4, 5, 6\}$

$\overline{BX} = \{1, 2, 3, 4, 5, 6, 7, 8\}$

By the definition of $\overline{BX}$ and $\underline{B}X$, the objects in U can be compartmentalized into three parts, called the negative and positive regions.

$$POSB(X) = \underline{B}\,X \quad (2.4)$$

$$NEGB(X)=U-BX \tag{2.5}$$

In the example, the two regions for $B=\{f3, f4\}$ and $X=\{1, 2, 5, ,4, 6\}$ are as follow:

$POSB\ (X) = \{4,5,6\}$

$NEGB\ (X)= \{1,2,3,7,8\}$

In data analysis, discovering dependencies among attributes is an important issue. Let D and C be subsets of $F\cup\{d\}$. For $0 \leq k \leq 1$, it is said that D depends on C in the kth degree (denoted $C \Rightarrow kD$), if

$$k = \gamma\ (C, D)\ + \frac{|POSC\ (D)|}{|U|} \tag{2.6}$$

Where

$$POSC\ (D) = \cup \underline{C}X$$
$$X \in U/D$$

Is called a positive region of the partition U/D with regard to C. This region is the set of all elements of U that can be uniquely classified into blocks of the partition U/D, by means of C. In the example, if $C = \{f\,3, f\,4\}$ then:

$POSC\ (d) = \cup\ (\underline{C}\ \{1, 2, 3, 8, 7\}, \underline{C}\ \{4\}, \underline{C}\ \{5,6\}) = \{4, 5, 6\}$.

The degree of dependency of attribute d on attributes $\{f\,3, f\,4\}$ is:

$$\gamma\ (\{f3, f4\}, d) = \frac{|POS\{f3, f4\}(d)|}{|U|} = \frac{3}{8}$$

The functional dependency of D and C ($C \Rightarrow D$) is a special case of dependency where $\gamma\ (C, D) = 1$. In this case it is said that all attributes' values from D are uniquely specified by the values of attributes from C.

A reduct is defined as a subset of minimum cardinality of the conditional attribute set C such $\gamma\ R\ (D) = \gamma\ C\ (D)$

$$R = \{X : X \subseteq C, \gamma\, x(D) = \gamma\, c(D)\} \quad (2.7)$$

$$\text{Rmin} = \{X : X \in R, \forall Y \in R, |X| \leq |Y|\} \quad (2.8)$$

The Core is defined as an intersection of all the sets in R_{min}

$$Core\ (R) = \cap\, X_{\,X \in R} \quad (2.9)$$

The core elements are those attributes that are impossible to omit without introducing more contradictions to the data set.

Utilizing the dataset in Table 2.1 and the degree of dependency D = {d} on all possible subsets of C can be calculated as:

$$\gamma\{1\} = \frac{1}{8} \quad \gamma\{2\} = 0 \quad \gamma\{3\} = 0 \quad \gamma\{4\} = \frac{2}{8} \quad \gamma\{1,2\} = \frac{1}{8}$$

$$\gamma\{1,3\} = \frac{5}{8} \quad \gamma\{1,4\} = 1 \quad \gamma\{2,3\} = 0 \quad \gamma\{2,4\} = \frac{2}{8} \quad \gamma\{3,4\} = \frac{3}{8}$$

$$\gamma\{1,2,3\} = \frac{5}{8} \quad \gamma\{1,2,4\} = 1 \quad \gamma\{1,3,4\} = 1 \quad \gamma\{2,3,4\} = \frac{3}{8}$$

The minimal results obtained in this example are: $R_{min} = \{f_1, f_4\}$.

Minimum reducts finding process is labelled as an NP-hard problem. Calculating all the potential reducts process (Core(R)) is a time exhausting. Thence, the researchers attempt to conform sundry heuristic algorithms for finding approximate solutions to this problem.

Table 2.2 shows the dataset after reduction where the dependency value of attributes equals to 1.

Table 2.2 Dataset after reduction

$x \in U$	$f1$	$f4$	d
1	2	0	3
2	0	0	4
3	0	0	4
4	0	0	4
5	0	4	5
6	0	4	5
7	0	0	4
8	0	0	4

Reduct Computation

The measurement of the solution's quality depends on the degree of the dependency, denoted as γ. Given two solutions i.e. trial solution *x*' and current solution *x*, the trial solution *x*' is accepted in the case of an increase in the degree of the dependency (i.e. if γ (*x*') > γ (*x*)). In the case of the dependency degree for is the same the both solutions (i.e. γ (*x*') = γ (*x*)), then the solution with the least attributes' number (denoted as #) will be accepted.

2.4 SELECTION STRATEGY

The selection phase identifies the individuals that are chosen for mating (reproduction) and the number of offspring's produced by each chosen individual. The selection strategy's main principle is "the better is an solution; the higher is its chance of being a parent (Blickle & Thiele 1995). It is the process that decides which solutions are to be conserved and allowed to reproduce and which one is merit to disappear. The main target of the selection operator is to assure the valid solutions and eliminate the invalid ones in a population whilst preserving a constant the size of the population (Shukla et al. 2015). This research describes the selection strategy required to be implemented. The three selection strategies employed in this research are discussed in this section.

2.4.1 Roulette Wheel Selection

Roulette Wheel selection, which is one of the classical selection strategies in GA, is the most straightforward selection strategy. In this strategy, all the solutions in the population are positioned on the Roulette Wheel depending on their fitness values. Each solution is allocated in a section of the Roulette Wheel, which its size is proportionate to the solution's fitness value (Razali et al. 2011). The size of the segment is directly proportionate to the fitness value, as the greater the fitness value is, the larger the section is. Thence, the Roulette Wheel is spinned. The solution located in the section on which Roulette Wheel stops is then selected. This process is reiterated till the coveted number of solutions is selected. Solutions with higher fitness value have a higher probability to be selected. However, at certain times, it is possible that the best solutions of a population can be missed. It is not guaranteed that valid solutions will be selected for the next generation (Kumar & Jyotishree 2012). In Roulette Wheel selection, solutions are selected with a probability that is directly proportionate to their fitness values i.e. a solution's selection corresponds to a section of a Roulette Wheel. The probabilities of selecting a parent can be considered an as spinning a Roulette Wheel with the size of the section for each parent being proportionate to its fitness as shown in Table 2.3 (Lin 2017).

The fitness calculation for each solution is based on a formula discussed in Chapter III (Section 3.2.4) as detail.

Table 2.3 Ratio between Population and Fitness

Solution	Fitness value
1	25.0
2	5.0
3	40.0
4	10.0
5	20.0

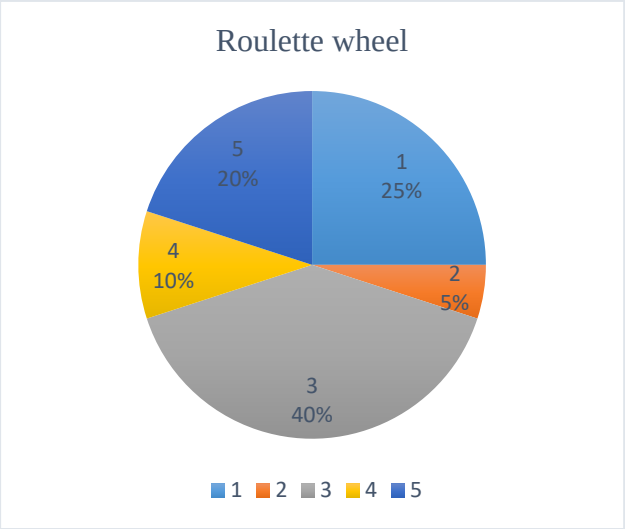

Figure 2.1 Roulette Wheel selection strategy

In Figure 2.1 Since the third solution has a higher fitness value than others, it is expected that the third solution will be chosen by roulette-wheel selection more than any other solutions.

2.4.2 Rank-Based Selection

Rank-Based Selection ranks and sorts the population based on the fitness value. The best solution holds rank N while the worst solution holds rank 1. Then, every solution is allocated to probability for selection with regard to its rank (Baker, 1985). solutions are selected depending on their probability of being selected. Rank-based selection, which is a preliminary selection strategy, stops too quick convergence and vary from Roulette Wheel selection with regarding to the pressure of the selection. Similarly, Rank-based selection dominates the scaling problems such as premature convergence or stagnation. Ranking has the control of the selective pressure *S* through uniform scaling technique across the population. Figure.2.2. Rank-based selection strategy (Talbi 2009).

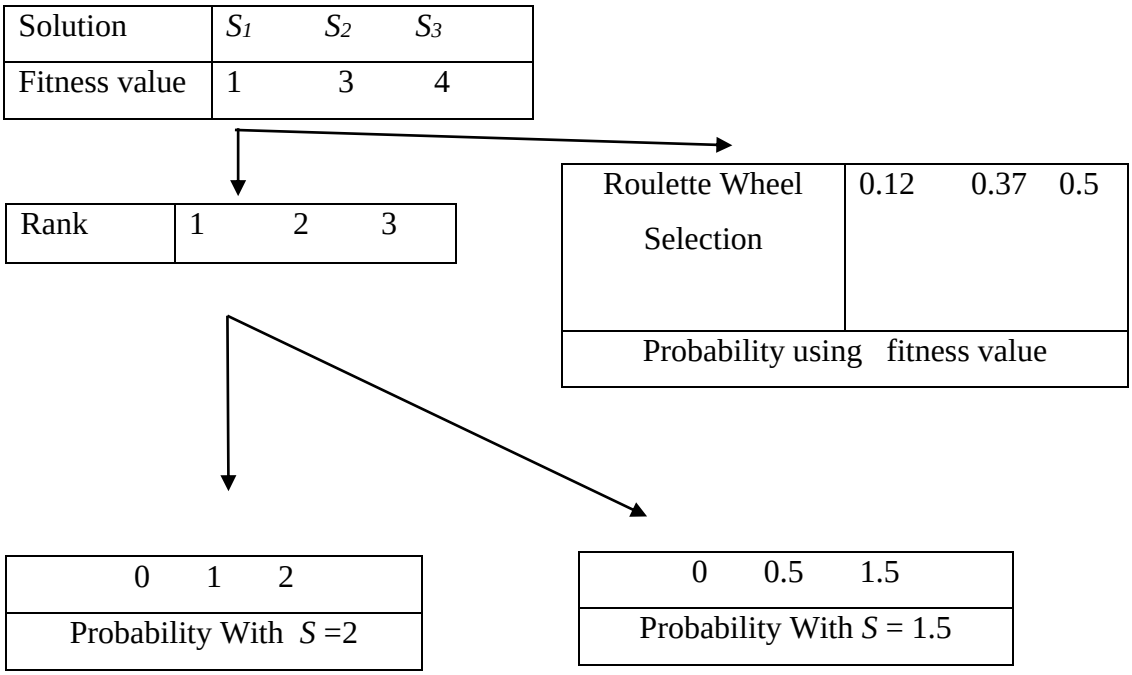

Figure 2.2 Rank-based selection strategy

2.4.3 Tournament Selection

The strategy of tournament Selection has been proclaimed by (J H Holland & Goldberg 1989). In Tournament Selection, a number of individuals *k*, are selected at random from the population. At this point, *k* is the size of the tournament that refers to the number of individuals who are selected randomly. In Tournament Selection strategy, the individual that is being selected from this group (Tournament) as parent is the best one. Tournament Selection provides seclectic pressure by convening a Tournament competition between k individuals. The individual with the highest fitness value among *k* individuals is denoted as the best individual. Then, the mating pool will accommodate the winner of the competition in the tournament. The tournament competition is reiterated until the mating pool for producing a new offspring have no more space (Blickle & Thiele 1995).

The mating pool contains the tournament winner with higher average fitness among the population. The difference of the fitness provides the selection pressure, which pushes the GA to ameliorate the succeeding generations' fitness. Tournament Selection strategy is an effective method and more docile to parallel implementation.