

# Implementation of Genetic Algorithm for School Monitoring System for Matiari District, Sindh, Pakistan

Zeeshan Qureshi<sup>1</sup>, Arjumand Bano Soomro<sup>1\*</sup>, Arifa Bhutto<sup>1</sup>, Muhammad Memon<sup>2</sup>, Azhar Ali Shah<sup>1</sup> and Imdad Ali Ismaili<sup>1</sup>

<sup>1</sup>Institute of Information and Communication Technology, University of Sindh, Jamshoro, Pakistan; qureshizeeshan1982@gmail.com, arjumand@usindh.edu.pk

<sup>2</sup>Institute of Business Administration, University of Sindh, Jamshoro, Pakistan

## Abstract

**Background:** This study focuses on Optimization field of computer science which works well when human beings are unable to solve the problem of finding an efficient path. **Methods:** The in-depth analysis of the problem has followed by the calculation of the school distances. The detailed route formation has been done so that the distances can be actualized in simulation-based optimization. The optimization is based on Genetic Algorithm and for this purpose a graphical user interface has been developed. **Findings:** The school monitoring teams were given an optimized path for their route to visit where they can easily save the time and the fuel. Results show that the genetic algorithm has produced an efficient and accurate path for the monitoring team so that the fuel and time can be saved. **Improvements/Applications:** The current study can be extended by comparing genetic algorithm output to any other optimization algorithms, so that the efficiency and time can be measured.

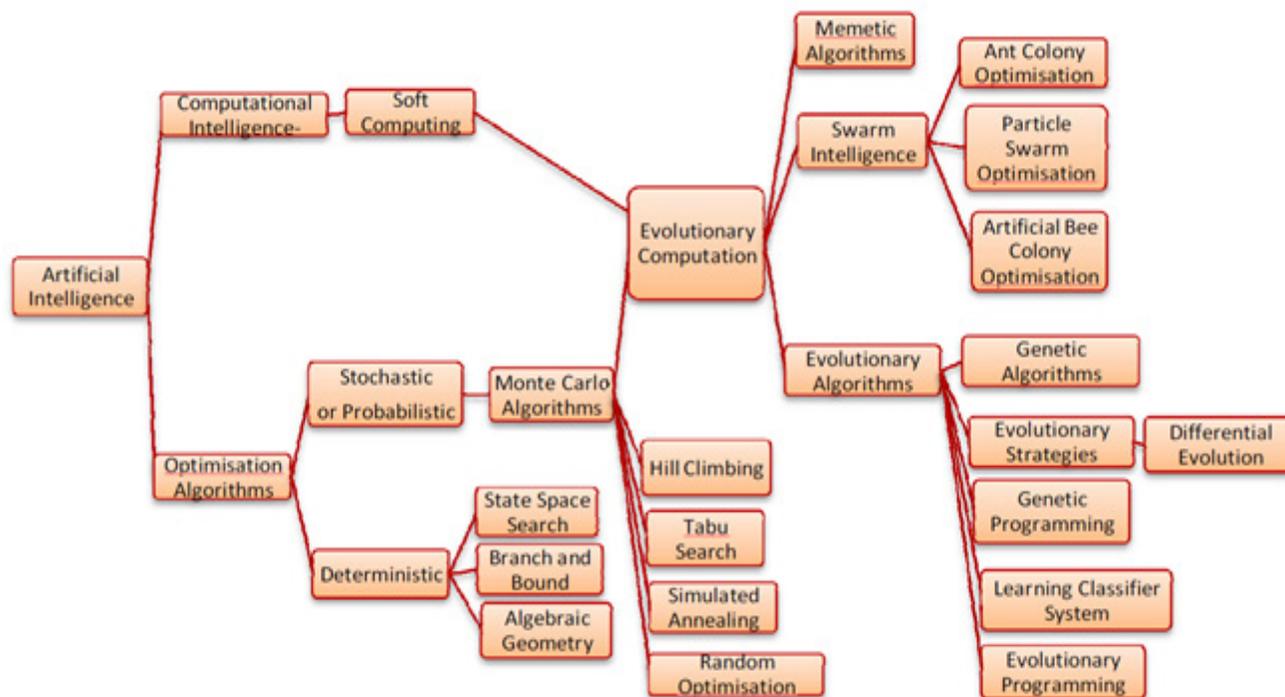
**Keywords:** Genetic Algorithm, Optimization, School Monitoring

## 1. Introduction

There are number of the jobs which are tedious for human beings whereas it is easy for computers and machines, especially the situations where much of the calculations are needed<sup>1</sup>. Optimization algorithms are one of the examples which can cover the calculations that are nearly impossible for human beings. Many of the schools in Province Sindh of Pakistan are operating and many of the teachers remain absent from their duties. One of the major reasons behind it is the political recruitment of staff. The teachers at primary and middle schools across the Sindh province remain absent mostly (as per local newspapers) from their schools and the students; future is at its risk. Recently

government of Sindh literacy department has initiated a monitoring program where the schools are monitored at a district base in which the teaching staff attendance has been checked. Various span of controls has been started for the monitoring purpose and many of the monitoring supervisors are working under the supervision of district monitoring officer entitled as Chief Monitoring Officer (CMO). This study has focused district Matiari district in Province Sindh as the case study. The district monitoring teams are formed to visit the schools located at various union councils at district Matiari. The monitoring teams visit the union councils or schools for the sake of checking attendance of teachers, this procedure cost them in term of time, fuel and other expenses.

\*Author for correspondence



**Figure 1.** Artificial intelligence algorithms

The Optimization algorithms<sup>2,3</sup> can perform optimization for these types of problems and can save sufficient resources and human efforts. This can expedite the performance of monitoring teams and their supervisors. Genetic Algorithm (GA)<sup>3-5</sup> is one of the successful algorithms in providing efficient optimization problems from Medical<sup>6</sup> to Energy<sup>7</sup>, science<sup>8</sup> to telecommunication<sup>9</sup> and many others. Various algorithms under the umbrella of Artificial Intelligence are shown in Figure 1.

## 1.1 Genetic Algorithm (GA)

Professor John Holland from USA proposed the Genetic Algorithm (GA) in computer science and Artificial Intelligence (AI) in 1975. This algorithm mimics the process of natural selection of biology into computer science. GA finds out the optimized solutions of bigger classes of evolutionary classes of algorithms<sup>10</sup>. Various fields are enriched with GA including medicine, plan-

ning, technology scheduling, and biology, engineering and others<sup>11,12</sup>. Many of the variations of GA are employed in real time problems and other intelligent systems<sup>13,14</sup>. The GA works on the genetic operators namely fitness function, mutation, crossover, selection and population. A simple and working diagram of GA is illustrated in Figure 2.

### 1.1.1 Population

This operation considers each and every state, node or any other identifier as a member of population. All of the possible solutions are inside the set of a population. Population is the foremost operator which considers each and every node (all sensor nodes). The population is the set of all possible solutions. Every solution can be termed as the individuals /chromosomes. It is usually randomly selected chromosomes. The population represents in the form of array of genes/alleles or string. It is worth to mention that the length of each population or string should



Figure 2. Working diagram of GA

be the same. A fitness value is assigned to each and every individual. The fitness value is the result of fitness function of GA; furthermore a fitness value is the indefinite specific value for each individual.

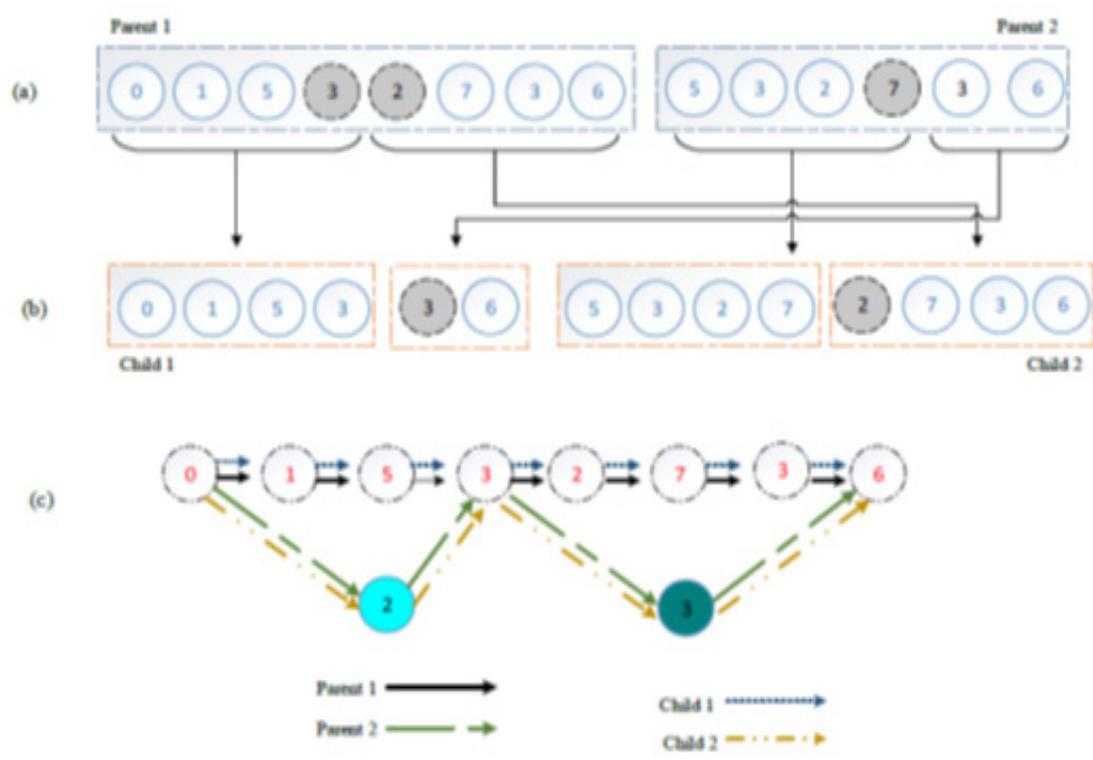
### 1.1.2 Selection

The other operator is the selection used with GA and the basic job is to reproduce the next generation node with energy. It selects the desired place or the next node which is more desirable whereas avoiding the rest of the nodes or places having less interest. The next desirable or favourite nodes or place is selected and the method is also called reproduction. The most desirable node or place has the better chance to go for the next level. A good selection strategy or a good selection may pro-

vide fitness evaluation more efficiently and will go for the optimized solution. The other side of the coin of selecting improper strategy may lead to perform GA in a wrong direction and may resist the success of overall algorithm.

### 1.1.3 Fitness function

The optimization and its suitability of the solution of GA is checked and for the understanding it is used. Every place or location is checked against the other ones and checked for the optimality of the solution and that is the result of the fitness function. The more suitable function has the tendency to produce an optimal path during the next pathways.



**Figure 3.** Crossover of individuals. In this figure, (a) Illustrates pairs of parent individual, (b) represents child candidate resulted from (a) & (c) shows the optimal route selection options

### 1.1.4 Crossover

The reproduction or crossover is one of the important steps of GA in which an offspring is produced from two parents and let them changeover their data between each other and producing succeeding children<sup>15,16</sup>. The crossover operator helps to create a better type of the offspring from the selected candidates of parents. A good implication can generate a better individual. The generation of the new child is shown in Figure 3. The child is inherited or created from the Parent1 and Parent2.

### 1.1.5 Mutation operator

This operator or stage can be used after the crossover operator<sup>17</sup>. This operator is also used for premature merging of the Genetic Algorithm. This operator is used to search or the new optimized local solution or possible solution from available places or nodes or solutions rather

than to be search for the existing one. It checks for the new solution from the entire place.

## 2. Research Methodology

The current study has been performed on MATLAB software and the results are produced based on the simulation containing various scenarios. The Genetic Algorithm (GA) has been employed to obtain the optimization. The focus will be on the optimized route selection for the efficient path of the visiting team.

In proposed study, the multiple union counsels are placed on the artificial map and the schools are then monitored. The schools are drawn on the map as a try to approximate the locations of the schools. The monitoring teams are shown as the links between the two schools. The calculation is made on the distance matrix calculated on a real data collected from government if permissible.

## 2.1 Proposed Research Framework

The main consideration started with the understanding of problem and its possible solutions. The thorough examination of the problem has been made at the start. Then the problem was formulated on the basis of original data. The original map and other data have been achieved. Then the locations were plotted on the field or simulation graph. The software has been used for the simulation is MATLAB 2017. There is another option available for the manual placement of schools on the ground or graph where the locations of the schools can be plotted manually on anywhere on the map. The number of schools or locations can be understood as the visiting locations during the monitoring process for the employees. The plotted locations are then given as input to the GA to select any location for the start and then with the help of GA, the optimized path or efficient path is provided so that the fuel and time of the monitoring staff can be saved. The GA finds next path with an additional quality that the next path is the best path or most optimized state or location from the available solutions. The GA will

provide the best path as a local and as a global and calculate the distances from every location to other location and finally calculate the minimum distance of overall tour for the objective of the efficient monitoring by saving the fuel and time. The proposed model has following steps (see Figure 4).

### Phase 1: In-depth Analysis of School Monitoring System

During this phase and in-depth analysis of system has been provided based on the school monitoring system formulated by Government of Sindh, education and literacy department for the monitoring of their schools and the formation of monitoring officers and teams. The monitoring system rules and regulations, monitoring mechanisms, devices used and some other parameters has been presented in this study.

### Phase 2: Study Based on Calculations of Schools and Distances

During this phase the requirements have been defined and analysis of schools located and their distances have been calculated against given data and district data in

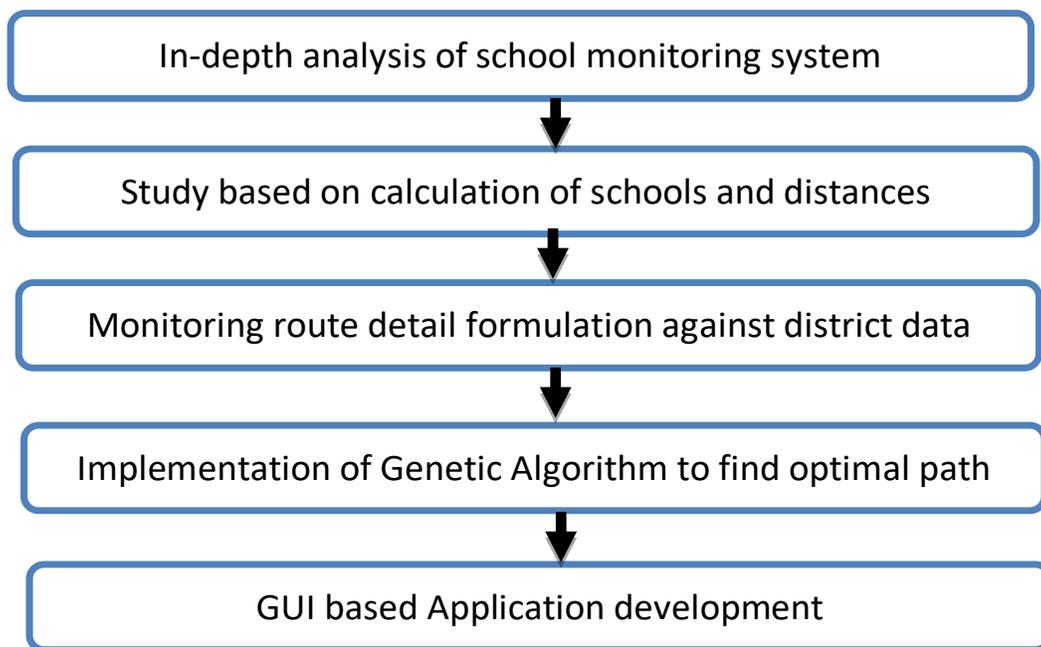


Figure 4. Research Frame Work

order to formulate the problem and to provide optimal solution for the shortest path so that the school monitoring officers and teams have to visit in order to save fuel and time.

### Phase 3: Monitoring Route Detail Formulation against District Data

In this stage, the routes have been identified, the distances have been analyzed and the timing of start and end of the tour will be calculated. The visiting mechanism, the repetition avoidance and many other factors have been analyzed and defined in this stage.

### Phase 4: Implementation of GA to Find Optimal Path

During this crucial phase GA have been applied with its all stages, step by step in order to achieve an optimal path among the schools located at various places available in the district data. This optimal path (shortest route) have been proved to be the shortest path during a particular day of school monitoring officers or teams visit in order to save their fuel and time. The visits can also be scheduled efficiently.

### Phase 5: GUI based Application development

At this final stage the application is developed for the above stages and graphical representation of total distance and optimal path has been given so that the problem can be formulated and solved visually. For the implementation of GA MATLAB 2017 has been used.

## 2.2 Proposed Research Scenario

The proposed model states the locations as a dot, circle or star or node on the plane, a total of 12 schools of Matiari district have been added on the ground. This is for the example purpose as there are various combinations of schools are placed on the ground and their optimal paths have been calculated. The ground has been set as the 100 x 100 points and the location of the schools is placed as circle for the sake of understanding. The x-coordinates and y-coordinates are labelled to understand the plane of the experiments. The visit of monitoring team may start from any point or location of the school and may

end to cover all locations in the district or points on the plane.

Another scenario has been created for the testing purpose of the algorithm that if the proposed framework is working or not. For the purpose of testing various planes along with varying number of schools plotted so that the efficiency of GA can be tested. The proposed algorithm has been tested and proved efficient for the small number of schools as well as increasing number of schools as shown in Figure 5.

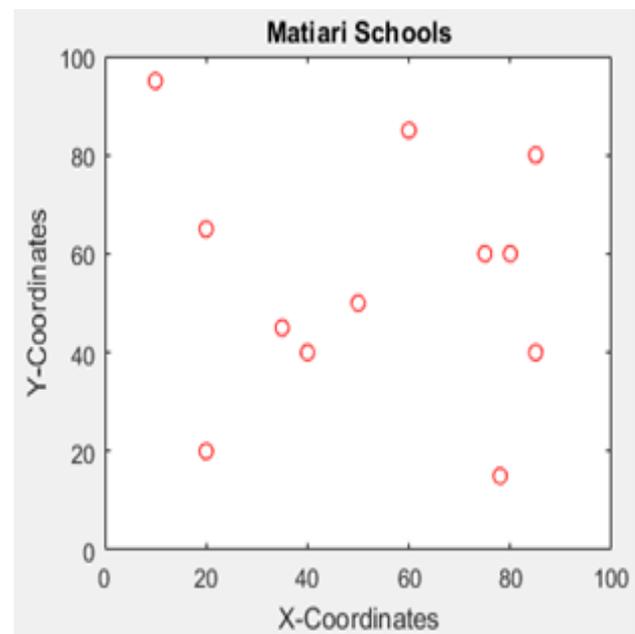


Figure 5. Proposed Scenario.

## 2.3 Proposed Optimization Algorithm

The GA has been modified to suit the conditions so that the optimization can be achieved applying starting parameters with the N number of schools or circles on the plane. Number of repetitions has been represented by  $\gamma$  and  $\varphi$  initial used the initial state of the tour. The roads or the number of schools are thought of a network and called as fully connected network.





**Figure 6.** Matiari District and schools located at various places

perform the job within short time by saving the fuel. The data of the schools collected from internet and the old district profile for Matiari district which helped in getting the exact estimation about the number of schools and locations of schools in Matiari district. The distance has been measured and estimated on the information collected from internet. The deployment of school identifiers has been done by placing circles as a sign on 10 x 10 Km<sup>2</sup> area. The area of plotting can be changed accordingly if there is a variation in the size of the experimental area or the location of the schools. The proposed approach is

equally useful for the both scenarios. The starting point maybe the different and the monitoring visit can be started from any point. The minimum distance is explored by GA which is a well-known and popular approach for finding the optimized paths in number of problems. For the purpose of simulation, the paths have been determined by GA and the schools have been located with the help of google maps and other available sources so that the schools can be places on the locations where they are located as shown in Figure 6 and gender wise distribution of schools can be seen in Figure 7. The simulations have

District Summary: All Schools and Teachers

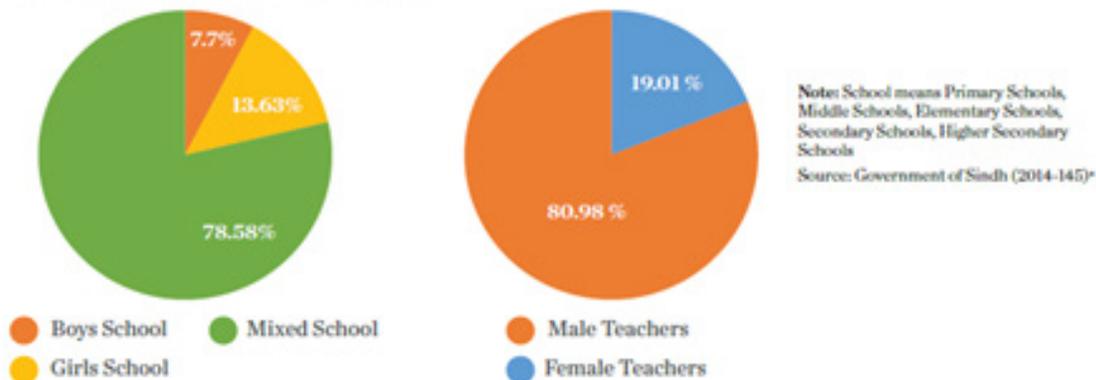


Figure 7. Matiari District school statistics

been tested also changing the locations of the schools and to check whether the algorithms perform well.

According to census 2014-2015, the total number of schools available in Matiari is 926 out of 98 schools are

Table 1. Matiari district school statistics

School Gender	School Level	Total
Boys	Primary	61
	Middle	4
	Elementary	---
	Secondary	31
	Higher Secondary	2
	Total	98
Girls	Primary	118
	Middle	4
	Elementary	---
	Secondary	12
	Higher Secondary	---
	Total	134
Mixed	Primary	678
	Middle	11
	Elementary	2
	Secondary	2
	Higher Secondary	1
	Total	694
Total	Primary	857
	Middle	19
	Elementary	2
	Secondary	45
	Higher Secondary	3
Grand Total		926

Source: SEMIS Census 2014-2015

reserved for boys and 134 as girls' schools. The total number of schools for both (mixed) for boys and girls are the 694. The overall total number of primary schools is 857 in which 61 schools are for boys, 118 for girls and mixed as 678. There are 4 schools available for each boys and girls whereas the number of middle schools for both genders is 11 totaling 19 as overall schools. The number of elementary schools is only two which are mixed schools, in this way there is no any single elementary reserved for boys only or girls only. There are 37 secondary schools reserved for boy's and 12 for girls. Only two secondary schools are mixed for both boys and girls. In this connection, the total number of secondary schools available in Matiari is 45. There is no any single higher secondary school in Matiari reserved for girls especially. Only 2 higher secondary schools are available in Matiari district for boys and one school is available for both or mixed. There are only 3 higher secondary schools available in Matiari district as shown in Table 1.

## 3. Results and Discussion

### 3.1 Simulation System Model

The system contains various scenarios, the various situations are encompassing to form different types of monitoring teams for various types of schools such as one team is responsible for primary schools for boys and the other is reserved for primary schools for girls. The other teams have been formed to monitor mixed schools as the number of schools is too many so there is a need of forming more teams. Following sections are the proposed scenarios for various teams and schools.

### 3.2 Simulation Parameters

The main parameters have been extracted and utilized in our study and the GA parameters are utilized from original available GA. The modified and selected parameters are shown in Table 2. The Original parameters have

**Table 2.** Parameters used in GA flow chart

Parameters	Symbol
Starting with initial search for current population	$h$
Initial Population	$\Pi 0$
A satisfied number of Optimized Paths	$\rho Optim$
Strongest candidates	$Sh$
The selection on the basis of binary tournament	$\eta Selection -1$ and $\eta Selection -2$
Selected strong candidates, which are ready for Crossover	$\beta h$
The candidate nodes Selected by Crossover	$\partial h + \partial' h = Uh$
Mutation operation	$\mu h$
Loops removing	$Lh$
New population by Elitism method	$Sh + Lh = \Pi h + 1$

been merged with the selected parameters. The mutation and the cross over properties are also illustrated in Table 2. The selection of an individual has been based on the energy levels which show the fitness level of the value. The solution is based on the fitness function which has been already defined.

### 3.3 Boys School Distance Optimization Results

The system bifurcates various school systems and then optimize for the routes for various teams as discussed

earlier there are more than 900 schools and it is near to impossible to form one team for such a number in this connection, the number of teams have been formed so that the monitoring process can be done efficiently. The system first introduces the plotting of the number of schools and then optimizes the routes for the sake of fuel saves and time saving.

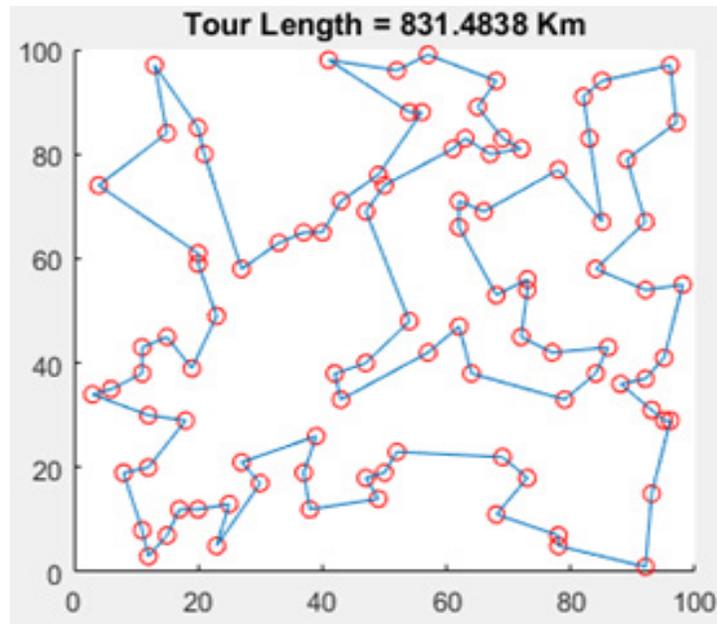
Various experiments have been performed and the optimized distance has been obtained. The distances and the number of schools for boys' locations have been represented in Tables 3,4. There is no any elementary school for

**Table 3.** Overall distance optimization for schools (boys)

S.No:	School Level	No: of schools	Optimized Distance
1.	Primary	61	633.32
2.	Middle	04	144.91
3.	Elementary	-----	-----
4.	Secondary	31	442.64
5	Higher Secondary	02	-----
	<b>Total</b>	<b>98</b>	<b>831.48</b>

**Table 4.** Overall distance optimization (girls)

S.No:	School Level	No: of schools	Optimized Distance
1.	Primary	118	895.91
2.	Middle	4	130.52
3.	Elementary	-----	-----
4.	Secondary	12	286.57
5.	Higher Secondary	-----	-----
6.	<b>Total</b>	<b>134</b>	<b>908.71</b>



**Figure 8.** Final results as optimized path for overall schools for boys

boys in Matiari district so there was no any experiment set for this type of data. There are two higher secondary schools for boys and there is no need for the optimization as there is only one distance between these two.

Figure 8 shows the final optimized path for the monitoring teams, a total of 831.48 kilometres has been the final distance to visit 98 schools for boys in Matiari district.

### 3.4 Girls School Distance Optimization Results

There are 134 schools for girls in district Matiari out of which 118 are the primary schools, 4 middle schools and 12 secondary schools for girls. There is no elementary or higher secondary school for girls in Matiari district. The following sections introduce the plotting of the number of schools and then optimize the routes for the sake of fuel saves and time saving.

For the sake of testing mechanism of algorithm that it is working on large amount of data the GA has been tested on overall school. The overall path has been optimized and the results show that the GA works fine with the same

proportion even in the case when the nodes or locations or number of points are increased. Figure 9 illustrates the final result of optimized result of 134 overall schools of girls in Matiari district.

### 3.5 Overall Boys and Girls Schools Distance Optimization Results

The experiments were performed to form a new set of experiments on the basis of mixed or total number of schools located in Matiari district. To make experiments more versatile some of the new experiments were added to the study.

For the sake of testing mechanism of algorithm that it is working on large amount of data the GA has been tested on overall school. The overall path has been optimized and the results show that the GA works fine with the same proportion even in the case when the nodes or locations or number of points are increased. Figure 10 illustrates the final result of optimized result of 232 overall schools of both boys and girls in Matiari district.

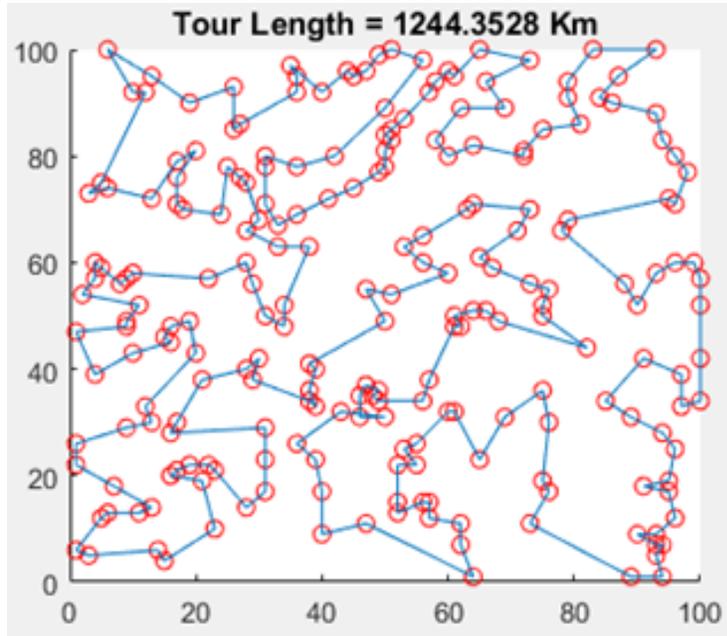


Figure 9. Optimized path for Overall girls school (134)

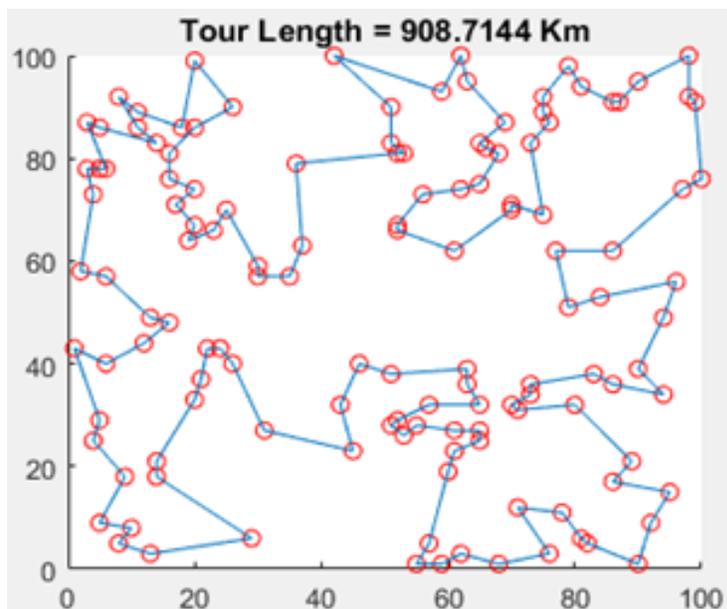


Figure 10. Optimized path for total schools for boys and girls

## 4. Conclusion

Optimization is the field of computer science which works well when a human being is unable to solve the problem of finding an efficient path of a particular problem. Many

of the schools in Sindh province face the absenteeism of teachers and they remain absent in their schools. For the sake of this problem Government of Sindh education and literacy department has recently started a monitoring program to monitor the teacher's attendance so that the

education of the children of the Sindh can be assured. For the purpose of the process of monitoring, various teams have been formed to stand arise the process of monitoring.

School monitoring officers or teams have to visit certain schools located at various villages and union councils of district Matiari. These schools have differentiating distances among them, so to optimize the easiest path, a short route is provided so that the time and fuel can be saved. The solution is very difficult and time consuming for human being if calculated manually, so for the sake solution, an integrated application is developed which will provide the optimal path or shortest path for the school monitoring officers or teams of Sindh education and literacy department so that school monitoring process can be made easy and a much fuel can be saved.

For the sake of solution of the above problem the in-depth analysis of the problem has been the first phase followed by the calculation of the school distances. The detailed route formation has been done so that the distances can be actualized in simulation-based optimization. The optimization is based on GA and for this purpose a graphical user interface has been developed.

The optimization process has been performed with the help of GA by placing the school as a point of optimization in a x and y plane. The school system along with classification of the schools and the facts has been extracted or collected from the Sindh Education Management Information System (2013-2014). A total of 926 schools are under the consideration of the optimization. The single team is not possible to monitor for such number of schools. So, the monitoring teams have been divided according to the number of schools. Various experiments have been performed so that the actual idea of the monitoring team for the saving of fuel and time can be formulated.

The schools were bifurcated as schools for boys and school for girls. The schools of girls as primary and secondary and the same schools for boys. The school monitoring teams were given an optimized path for their route to visit where they can easily save the time and the fuel. Results show that the GA has produced an efficient

and accurate path for the monitoring team so that the fuel and time can be saved.

## 5. Future Work

The current study can be extended by comparing GA output to any other optimization algorithms, so that the efficiency can be measured. The study can also be employed on the data of other 21 districts of the Sindh province or even wider scope can be achieved.

## 6. References

1. Durán RM, Rojas LP, Daza VP. A genetic algorithm for the Flexible Job Shop problem. *Ingeniare*. 2011; 19(1): 53–61. <https://doi.org/10.4067/S0718-33052011000100006>.
2. Deaven DM, Ho KM. Molecular geometry optimization with a genetic algorithm. *Physical Review Letters*. 1995; 75(2): 288–291. <https://doi.org/10.1103/PhysRevLett.75.288>.
3. Deb K, Agrawal S, Pratap A, Meyarivan T. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2000; 849–858. [https://doi.org/10.1007/3-540-45356-3\\_83](https://doi.org/10.1007/3-540-45356-3_83).
4. Chu P. A genetic algorithm for the generalised assignment problem. *Elsevier*. 1997; 24(1): 17–23. [https://doi.org/10.1016/S0305-0548\(96\)00032-9](https://doi.org/10.1016/S0305-0548(96)00032-9).
5. Mühlenbein H, Schomisch M, Born J. The parallel genetic algorithm as function optimizer. *Parallel Computing*. 1991; 17(6–7): 619–632. [https://doi.org/10.1016/S0167-8191\(05\)80052-3](https://doi.org/10.1016/S0167-8191(05)80052-3).
6. Gupta SK, Kuila P, Jana PK. Genetic algorithm for k-connected relay node placement in wireless sensor networks. *Advances in Intelligent Systems and Computing*. 2016; 379: 721–729. [https://doi.org/10.1007/978-81-322-2517-1\\_69](https://doi.org/10.1007/978-81-322-2517-1_69).
7. Haapola J, Shelby Z, Pomalaza-Ráez C, Mähönen P. Cross-layer energy analysis of multi-hop wireless sensor networks. *Proceedings of the Second European Workshop on Wireless Sensor Networks, EWSN*. 2005; 2005; 33–44. <https://doi.org/10.1155/WCN.2005.523>.
8. Badi A. *Reliable, energy-aware cross-layer protocol for wireless sensor networks*. Florida Atlantic: Florida Atlantic University. 2009.
9. Tavakkoli-Moghaddam R, Safari J, Sassani F. Reliability optimization of series-parallel systems with a choice of redundancy strategies using a genetic algorithm. *Reliability*

- Engineering & System Safety. 2008; 93(4): 550–556. <https://doi.org/10.1016/j.res.2007.02.009>.
10. Kalaycı TE, Yıldırım KS, Uğur A. Maximizing Coverage in a Connected and K-Covered Wireless Sensor Network Using Genetic Algorithms. *Applied mathematics*. 2007; 1(3): 123–130.
  11. Melanie M. *An introduction to genetic algorithms*. MIT press. 1998.
  12. Liu JL, Ravishankar CV. LEACH-GA: Genetic Algorithm-Based Energy-Efficient Adaptive Clustering Protocol for Wireless Sensor Networks. *IJMLC*. 2011; 1(1): 79–85. <https://doi.org/10.7763/IJMLC.2011.V1.12>.
  13. Gunasekaran M, Periakaruppan S. GA-DoSLD: Genetic algorithm based denial-of-sleep attack detection in WSN. *Security and Communication Networks*. 2017; 10. <https://doi.org/10.1155/2017/9863032>.
  14. Pannu A. Artificial Intelligence and its Application in Different Areas. *Int J Innov Res Sci Eng Technol*. 2008.
  15. Safa H, El-Hajj W, Zoubian H. A robust topology control solution for the sink placement problem in WSNs. *J NETW COMPUT APPL*. 2014; 39(1): 70–82. <https://doi.org/10.1016/j.jnca.2013.04.009>.
  16. Yetgin H, Cheung KTK, El-Hajjar M, Hanzo L. Network-Lifetime Maximization of Wireless Sensor Networks. *IEEE Access*. 2015; 3: 2191–2226. <https://doi.org/10.1109/ACCESS.2015.2493779>.
  17. Cherifa B, Saidane LA, Boucetta C, Idoudi H. Ant Colony Optimization Based Hierarchical Data Dissemination in WSN. *International Wireless Communications and Mobile Computing Conference*. 2015.