

# Genetic Algorithm Crypt-Analysis

Dylan Souvage  
Department of Computer Science  
Brock University  
St. Catharines, Canada  
ds16bz@brocku.ca

**Abstract**—A genetic algorithm to solve the Vigenere Cipher was developed, then run on two separate encrypted texts. 100 total runs across several parameters/crossovers were then run and compared to produce the best parameters for solving the Vigenere Cipher, thus also producing the optimal genetic algorithm. In this case the most successful algorithm was one with a high crossover rate, but not 100, and a low mutation rate, but not 0. This is likely because without mutation the algorithm gets stuck at local maxima, and with high crossover there is too much inbreeding.

## I. INTRODUCTION

A genetic algorithm (GA) is a metaheuristic inspired by the process of natural selection. Genetic algorithms are commonly used to generate high-quality solutions to optimization and search problems by relying on bio-inspired operators such as mutation, crossover and selection. [1]

The Vigenere Cipher is a method of encrypting alphabetic text. It uses a simple form of polyalphabetic substitution. A polyalphabetic cipher is any cipher based on substitution, using multiple substitution alphabets. The encryption of the original text is done using the Vigenere square or Vigenere table.

The table consists of the alphabets written out 26 times in different rows, each alphabet shifted cyclically to the left compared to the previous alphabet, corresponding to the 26 possible Caesar Ciphers. At different points in the encryption process, the cipher uses a different alphabet from one of the rows. The alphabet used at each point depends on a repeating keyword. [2]

The goal of this project was to optimize a genetic algorithm, by solving the Vigenere Cipher, or at least produce closer to optimal results, and to compare different parameters and crossover strategies. The purpose was not to necessarily solve the Vigenere Cipher, but rather focus entirely on the genetic algorithm's effectiveness.

## II. BACKGROUND

A genetic algorithm was constructed, with the following variables being parameters:

- Population Size
- Chromosome Size
- Number of Generations to run the algorithm for
- Length of time to run the algorithm
- Percentage of population to crossover
- Percentage of population to mutate

Within the main program (Genetic Algorithm), there were several other algorithms used:

- Randomize/create population
- Randomize/create chromosome
- K-Tournament Selection
- Crossover
- Selection

Tournament selection is best explained with a concrete example. Suppose you want to pick 20 individuals from 100. Randomly choose (with uniform probability) a small number of individuals (typically fewer than 10) from the 100 (with replacement). Keep the fittest one. Do this again and again until you have got your 20. Tournament selection, while slower and more complicated, can create more diverse populations[3]

In crossover, portions of two parents from the current generation are combined to create two offspring: a random subpart of the father's bit string is swapped with a random subpart of the mother's bit string. In this case, both single point crossover and double point crossover was used.

In mutation, two purely random genes are selected from a chromosome, then they are swapped. Then, a third random gene from the chromosome is selected and randomized.

Then, for n number of generations, this entire process is repeated and the population is updated. Thus producing a better population's "fitness" score by favouring better performing chromosomes (strings).

## III. EXPERIMENTAL SETUP

A variety of parameters were used in this algorithm.

Per each parameter, per each crossover strategy, and per two different encrypted passwords, one of size 26 and one of size 40, 5 runs were attempted. In total, 100 runs were attempted.

For each generation, the average fitness and best fitness were collected.

The values were exported to a CSV and then analyzed using Excel.

Population Size : 250

Maximum Generations : 250

Max Chromosome Size (with two separate encrypted strings) : 26 & 40

Crossover percentage : 100(%), 90(%)

Mutate Percentage : 0(%), 10(%)

And a final program was run with the following unique parameters: Mutate: 15(%) and Crossover: 80(%)

#### IV. PSEUDO CODE FOR ALGORITHMS USED

---

##### Algorithm 1 General Genetic Algorithm Pseudo code

---

**Input:** encrypted string, GA parameters in

**Output:** GA results out

*Initialisation* : set GA parameters

- 1: generate random initial population, POP, of size popSize
  - LOOP Process*
  - 2: **for**  $gen = l$  to  $MAXGEN$  **do**
  - 3:    $bestFitness$  = evaluate fitness of each chromosome in POP
  - 4:   select a new population based on K-Tournament selection
  - 5:   apply crossover function to POP
  - 6:   apply mutation function to POP
  - 7:   **if** ( $bestFitnessOfCurrPop < bestFitness$ ) **then**
  - 8:      $bestFitness = bestFitnessOfCurrPop$
  - 9:   **end if**
  - 10:   update  $finalResults$  with  $currentResults$
  - 11: **end for**
  - 12: **return**  $finalResults$
- 

---

##### Algorithm 2 K-Tournament Selection

---

**Input:** population, k in

**Output:** return best out

*Initialisation* : best = null

*LOOP Process*

- 1: **for** population size **do**
  - 2:   **for**  $i = l$  to  $k$  **do**
  - 3:      $bestFitness$  = evaluate fitness of each chromosome in POP[i to k]
  - 4:     add best  $bestFitness$  to new kPop list
  - 5:   **end for**
  - 6: **end for**
  - 7: **return**  $kPop$
- 

---

##### Algorithm 3 Crossover Algorithm

---

**Input:** population, percentToCross in

**Output:** return newPopulation out

*Initialisation* : best = null

*LOOP Process*

- 1: **for** population size \* percentToCross **do**
  - 2:    $randomPoint = 0 < randomPoint < chromosome$  length
  - 3:    $parentOne = population[i]$
  - 4:    $parentTwo = population[i+1]$
  - 5:    $childOne = parentOne$  split from start to  $randomPoint$  and  $parentTwo$  from  $randomPoint$  to end
  - 6:    $childTwo = parentTwo$  split from start to  $randomPoint$  and  $parentOne$  from  $randomPoint$  to end
  - 7:   update  $newPopulation$
  - 8: **end for**
  - 9: **return**  $newPopulation$
- 

---

##### Algorithm 4 Mutation Algorithm

---

**Input:** population, percentToMutate in

**Output:** return newPopulation out

*Initialisation* : best = null

*LOOP Process*

- 1: **for** population size \* percentToMutate **do**
  - 2:    $chromosome = population[i]$
  - 3:    $geneOne = chromosome$  random gene
  - 4:    $geneTwo = chromosome$  random gene
  - 5:    $chromosome[geneOne] = geneTwo$
  - 6:    $chromosome[geneTwo] = geneOne$
  - 7:    $chromosome[random] = randomGene$
  - 8: **end for**
  - 9: **return**  $newPopulation$
- 

#### V. RESULTS

##### VI. SINGLE POINT Crossover

###### A. One point crossover, 80 Crossover 15 Mutation 26 Chromosome

There is a clear maxima reached by about 50 generations. Further generations fail to find a better maxima.

The final best fitness achieved: 0.092123695

The final key deciphered: thisisasucsecuretaassord

Overall: this is a very successful set of parameters: the deciphered text is relatively legible to a native English speaker.

Mean 0.195334544

Standard Error 0.00364948

Median 0.148424904

Mode 0.116148321

Standard Deviation 0.129286403

Sample Variance 0.016714974

Kurtosis 4.760302851

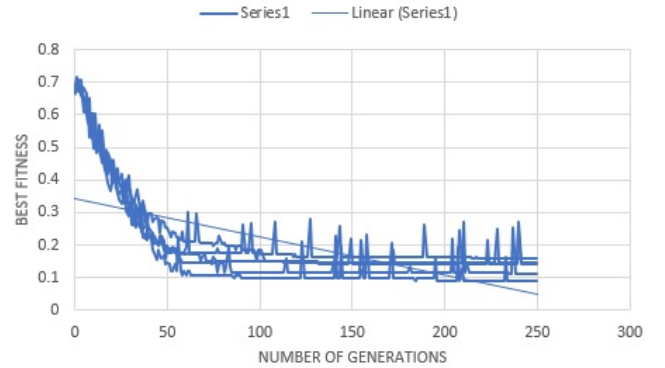


Fig. 1. One point crossover 80 Crossover 15 Mutation 26 Chromosome best fit per generation.

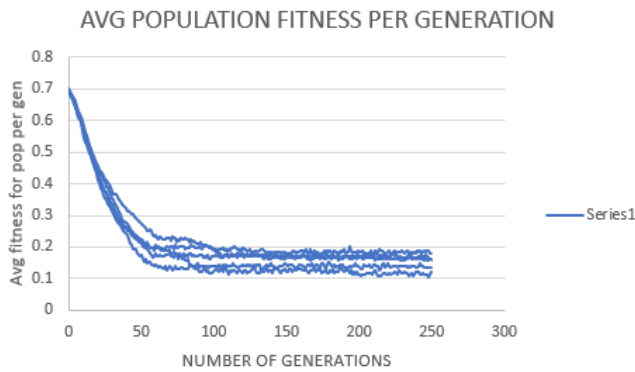


Fig. 2. One point crossover 80 Crossover 15 Mutation 26 Chromosome avg pop fit per generation.

Skewness 2.291954613  
 Range 0.624709502  
 Minimum 0.092123695  
 Maximum 0.716833197  
 Sum 245.144853  
 Count 1255  
 Largest(1) 0.716833197  
 Smallest(1) 0.092123695  
 Confidence Level(95.0%) 0.00715976  
 Figure 1: Best fit per generation  
 Figure 2: Avg fit per generation

#### B. One point crossover, 80 Crossover 15 Mutation 40 Chromosome

There is a clear maxima reached by about 75 generations. Further generations fail to find a better maxima.

The final best fitness achieved: 0.484381336

The final key deciphered: tcptititijkciitixapivotqgphippy-ilealtd

Mean 0.529112276  
 Standard Error 0.001344138  
 Median 0.514306655  
 Mode 0.484381336  
 Standard Deviation 0.047617418  
 Sample Variance 0.002267418  
 Kurtosis 3.299324444  
 Skewness 1.985794519  
 Range 0.232717513  
 Minimum 0.484381336  
 Maximum 0.717098849  
 Sum 664.0359066  
 Count 1255  
 Largest(1) 0.717098849  
 Smallest(1) 0.484381336  
 Confidence Level(95.0%) 0.002637008  
 Figure 3: Best fit per generation  
 Figure 4: Avg fit per generation

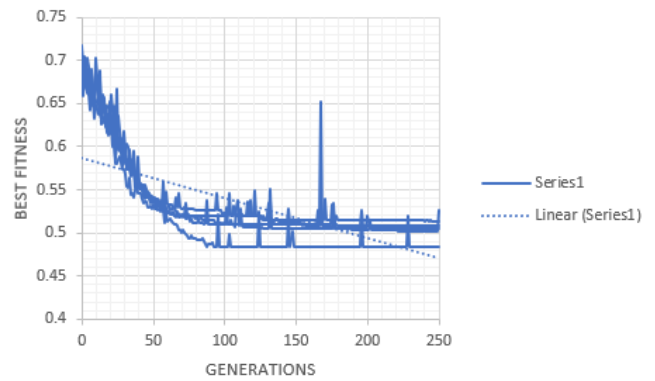


Fig. 3. One point crossover 80 Crossover 15 Mutation 40 Chromosome best fit per generation.

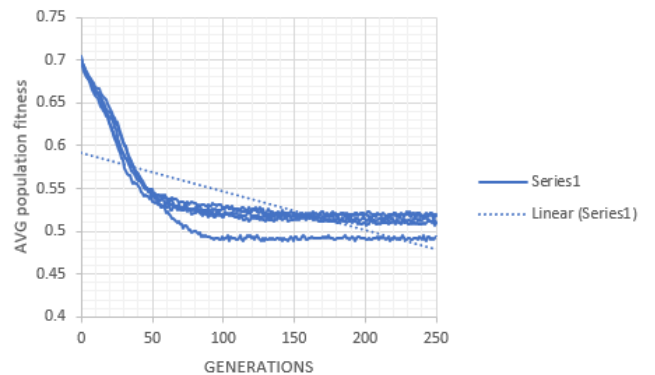


Fig. 4. One point crossover 80 Crossover 15 Mutation 40 Chromosome avg pop fit per generation.

#### C. One point crossover, 90 Crossover 0 Mutation 26 Chromosome

There is a clear maxima reached by about 75 generations. Further generations fail to find a better maxima.

The final best fitness achieved: 0.151820079

The final key deciphered: thistsayupeestpuretassoro

Mean 0.244512554  
 Standard Error 0.002999052  
 Median 0.210293988  
 Mode 0.27883828  
 Standard Deviation 0.10624434  
 Sample Variance 0.01128786  
 Kurtosis 6.816170188  
 Skewness 2.537652448  
 Range 0.59168032  
 Minimum 0.151820079  
 Maximum 0.743500399  
 Sum 306.8632553  
 Count 1255  
 Largest(1) 0.743500399  
 Smallest(1) 0.151820079  
 Confidence Level(95.0%) 0.005883712  
 Figure 5: Best fit per generation

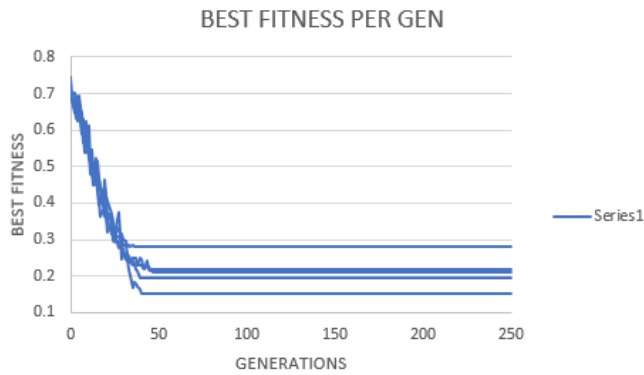


Fig. 5. One point crossover 90 Crossover 0 Mutation 26 Chromosome best fit per generation.

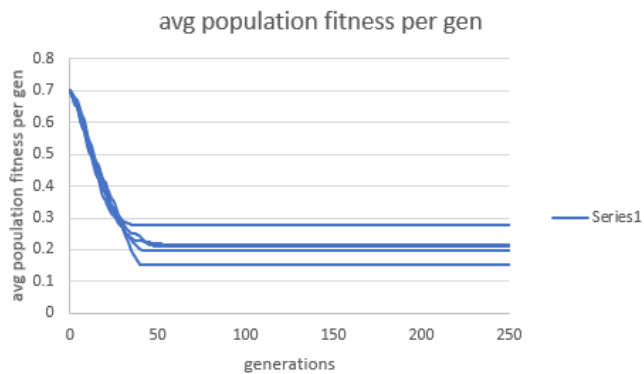


Fig. 6. One point crossover 90 Crossover 0 Mutation 26 Chromosome avg pop fit per generation.

Figure 6: Avg fit per generation

#### D. One point crossover, 90 Crossover 0 Mutation 40 Chromosome

There is a clear maxima reached by about 50 generations. Further generations fail to find a better maxima.

The final best fitness achieved: 0.522456463

The final key deciphered: knafptadpitiditityvfeaoslqcgbf-vairaeaktc

Mean 0.541693606

Standard Error 0.001055047

Median 0.527189749

Mode 0.550180875

Standard Deviation 0.037376078

Sample Variance 0.001396971

Kurtosis 6.275780701

Skewness 2.526324345

Range 0.203395722

Minimum 0.515109563

Maximum 0.718505285

Sum 679.8254761

Count 1255

Largest(1) 0.718505285

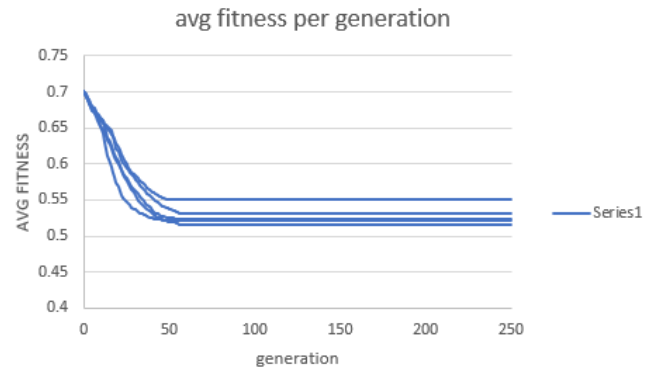
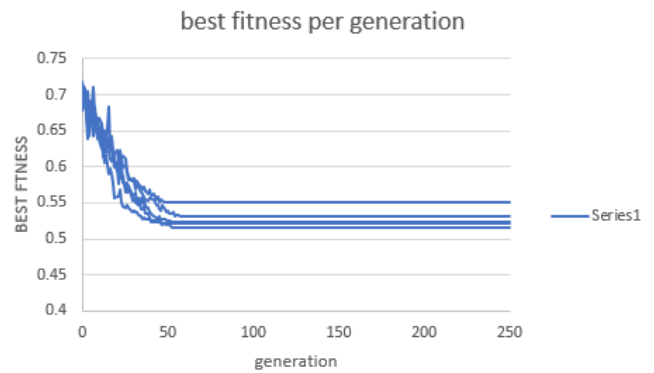


Fig. 7. One point crossover 90 Crossover 0 Mutation 40 Chromosome best fit and avg fit per generation.

Smallest(1) 0.515109563

Confidence Level(95.0%) 0.002069852

Figure 7: Best fit and avg fit per generation

#### E. One point crossover, 90 Crossover 10 Mutation 26 Chromosome

There is a clear maxima reached by about 50 generations. Further generations continue to slowly optimize to a better solution.

The final best fitness achieved: 0.112913257

The final key deciphered: ihisisasupersecureaassjarh

Mean 0.198656842

Standard Error 0.003591854

Median 0.150904682

Mode 0.13661365

Standard Deviation 0.127244943

Sample Variance 0.016191275

Kurtosis 5.608473614

Skewness 2.538263659

Range 0.606597319

Minimum 0.112913257

Maximum 0.719510576

Sum 249.3143364

Count 1255

Largest(1) 0.719510576

Smallest(1) 0.112913257

Confidence Level(95.0%) 0.007046706

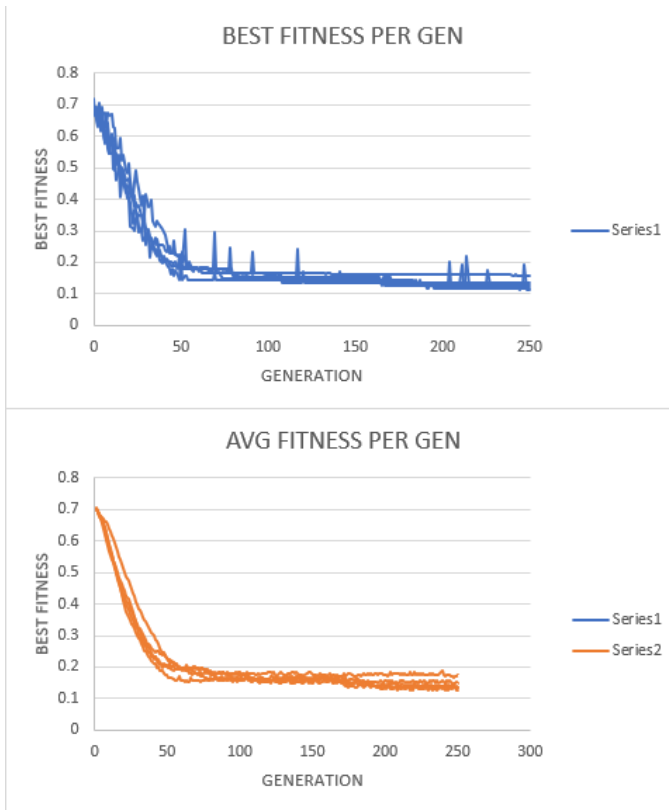


Fig. 8. One point crossover 90 Crossover 10 Mutation 26 Chromosome best fit and avg fit per generation.

Figure 8: Best fit and avg fit per generation

#### F. One point crossover, 90 Crossover 10 Mutation 40 Chromosome

There is a clear maxima reached by about 70 generations. Further generations continue to very slowly optimize to a better solution.

The final best fitness achieved: 0.487864886

The final key deciphered: kcpeitipiittcpitimiaippiluiyhretlr-ixitti

Mean 0.524132422

Standard Error 0.001293077

Median 0.508099767

Mode 0.503094417

Standard Deviation 0.045808534

Sample Variance 0.002098422

Kurtosis 4.383517124

Skewness 2.265131135

Range 0.219147052

Minimum 0.487202453

Maximum 0.706349505

Sum 657.7861898

Count 1255

Largest(1) 0.706349505

Smallest(1) 0.487202453

Confidence Level(95.0%) 0.002536834

Figure 9: Best fit and avg fit per generation

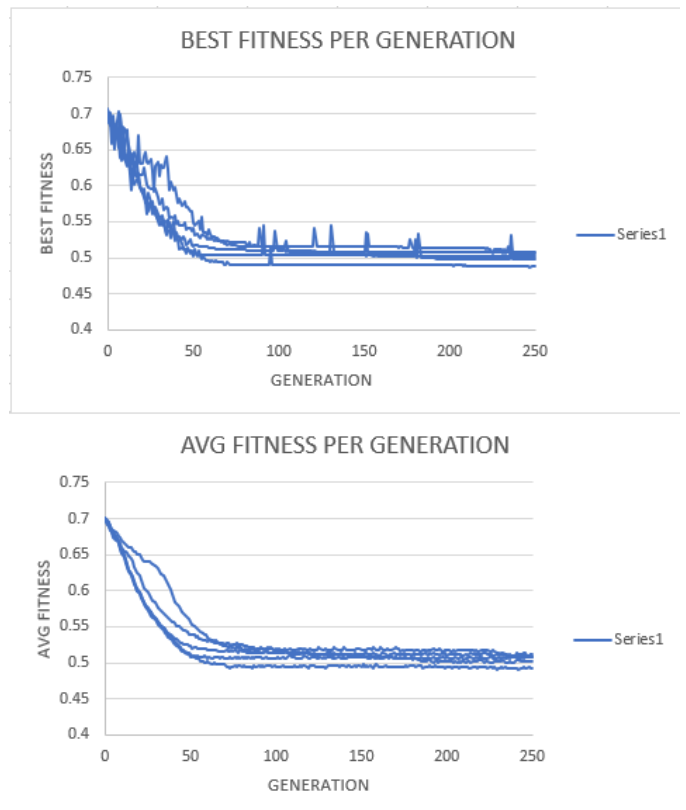


Fig. 9. One point crossover 90 Crossover 10 Mutation 40 Chromosome best fit and avg fit per generation.

#### G. One point crossover 100 Crossover 0 Mutation 26 Chromosome

There is a very clear maxima reached by about 50 generations. Further generations make no progress, and no further states are explored.

The final best fitness achieved: 0.155491558

The final key deciphered: xhesiadupvesecuretaasmord

Mean 0.228948645

Standard Error 0.003067431

Median 0.186678448

Mode 0.186678448

Standard Deviation 0.108666747

Sample Variance 0.011808462

Kurtosis 6.394266252

Skewness 2.490696425

Range 0.560077933

Minimum 0.155416589

Maximum 0.715494522

Sum 287.3305498

Count 1255

Largest(1) 0.715494522

Smallest(1) 0.155416589

Confidence Level(95.0%) 0.006017863

Figure 10: Best fit and avg fit per generation

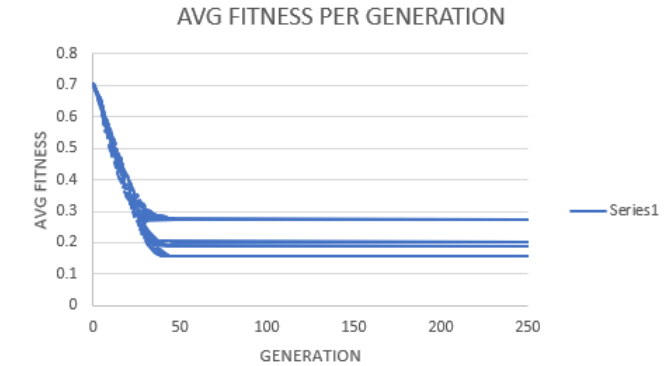
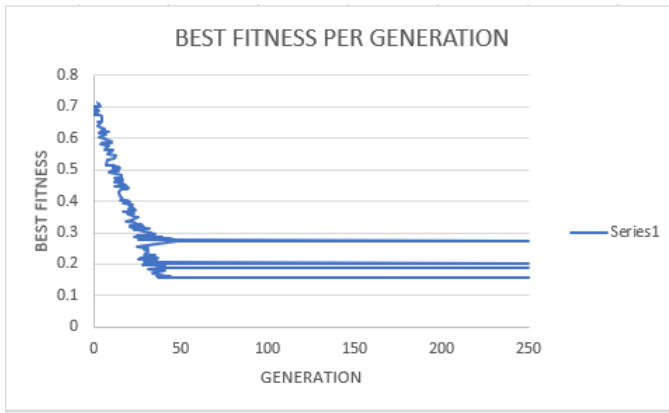


Fig. 10. One point crossover 100 Crossover 0 Mutation 26 Chromosome best fit and avg fit per generation.

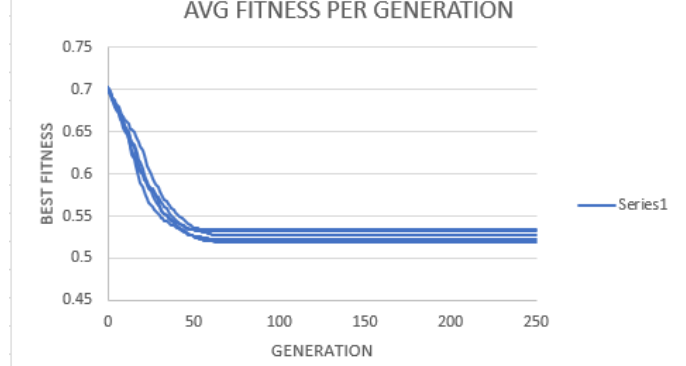
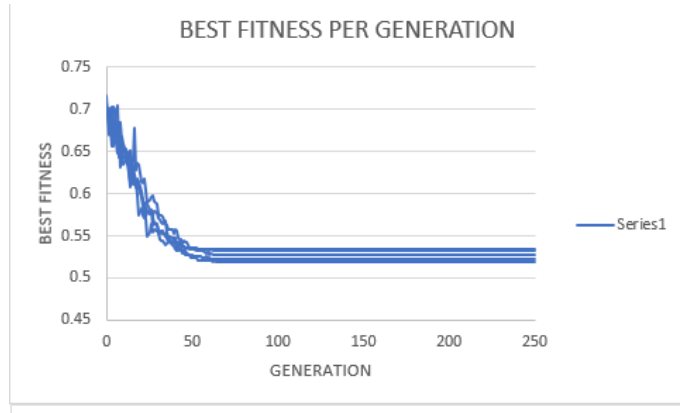


Fig. 11. One point crossover 100 Crossover 0 Mutation 40 Chromosome best fit and avg fit per generation.

#### H. One point crossover 100 Crossover 0 Mutation 40 Chromosome

There is a very clear maxima reached by about 50 generations. Further generations make no progress, and no further states are explored.

The final best fitness achieved: 0.52272287

The final key deciphered: tiltdujtiljrmiiitnibavfldibm-vadmesuakic

Mean 0.540851402

Standard Error 0.00103179

Median 0.532669321

Mode 0.534340007

Standard Deviation 0.036552181

Sample Variance 0.001336062

Kurtosis 7.476165765

Skewness 2.834122774

Range 0.197436429

Minimum 0.518897318

Maximum 0.716333747

Sum 678.7685094

Count 1255

Largest(1) 0.716333747

Smallest(1) 0.518897318

Confidence Level(95.0%) 0.002024226

Figure 11: Best fit and avg fit per generation

#### I. One point crossover 100 Crossover 10 Mutation 26 Chromosome

There is a maxima reached by about 50 generations. Further generations make slow progress, but do work towards finding a better solution.

The final best fitness achieved: 0.110443373

The final key deciphered: tbtisisasupersegurepasssoed

Mean 0.190763416

Standard Error 0.003444158

Median 0.146974026

Mode 0.146974026

Standard Deviation 0.122012679

Sample Variance 0.014887094

Kurtosis 6.598107526

Skewness 2.680597633

Range 0.599452476

Minimum 0.110443373

Maximum 0.709895849

Sum 239.4080873

Count 1255

Largest(1) 0.709895849

Smallest(1) 0.110443373

Confidence Level(95.0%) 0.006756948

Figure 12: Best fit and avg fit per generation

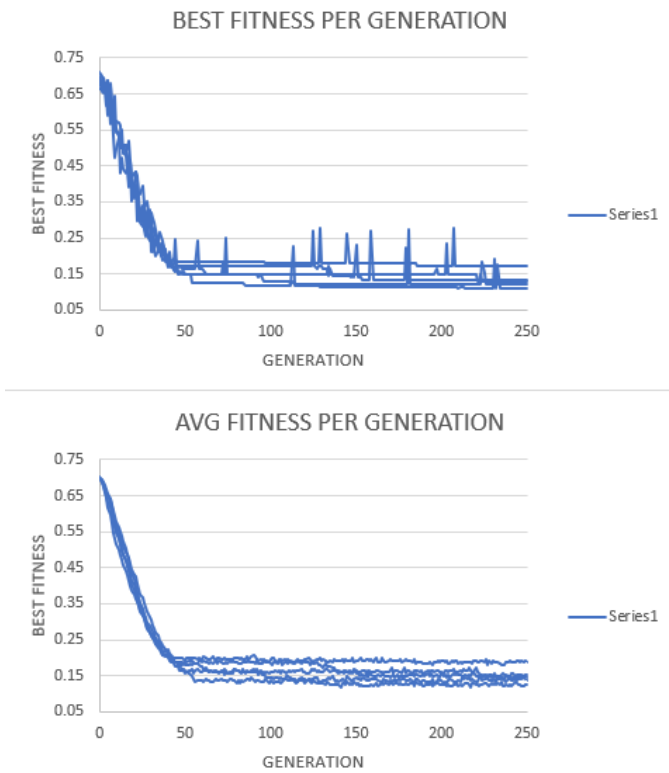


Fig. 12. One point crossover 100 Crossover 10 Mutation 26 Chromosome best fit and avg fit per generation.

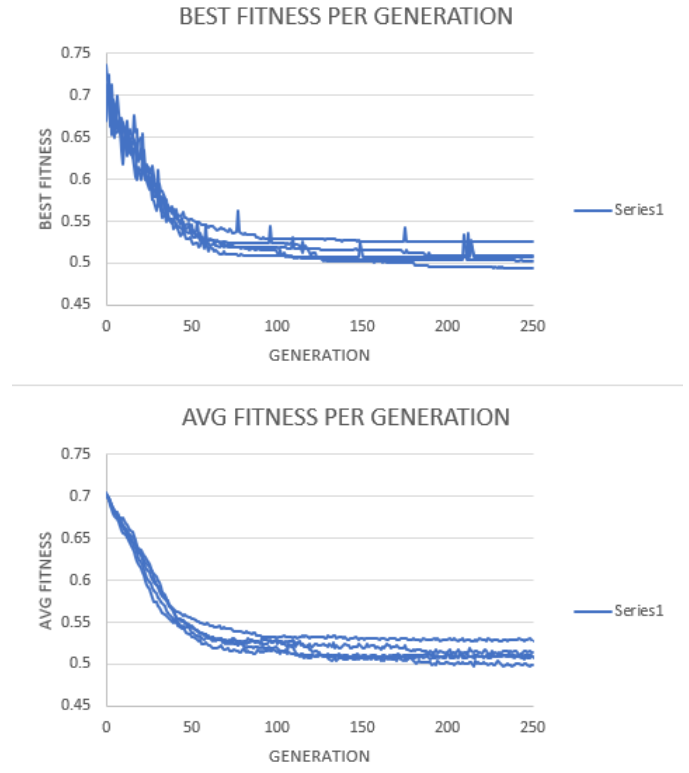


Fig. 13. One point crossover 100 Crossover 10 Mutation 40 Chromosome best fit and avg fit per generation.

#### J. One point crossover 100 Crossover 10 Mutation 40 Chromosome

There is a maxima reached by about 75 generations. Further generations clearly make progress, but they move much slower.

The final best fitness achieved: 0.494413777

The final key deciphered: rpndpppiiytmptgtpbrptuott-pivrpfunipeaiit

Mean 0.532630255

Standard Error 0.001260933

Median 0.517367095

Mode 0.506634419

Standard Deviation 0.044669781

Sample Variance 0.001995389

Kurtosis 4.100303559

Skewness 2.163895971

Range 0.241651258

Minimum 0.494413777

Maximum 0.736065035

Sum 668.4509701

Count 1255

Largest(1) 0.736065035

Smallest(1) 0.494413777

Confidence Level(95.0%) 0.002473771

Figure 13: Best fit and avg fit per generation

#### VII. RESULTS FROM SINGLE POINT CROSSOVER

It is clear, that different parameters have vastly different outcomes.

Of note:

- without a mutation, it is clear that the algorithm will stop improving fitness quickly. Mutation/randomness is needed to avoid local maxima.
- similarly, with too much crossover we run into more inbreeding, and again lack the randomness needed to avoid local maxima.
- larger chromosome size clearly is much harder for the algorithm to solve, and requires many more generations to even come close to the smaller 26 size chromosome fitness level

#### VIII. TWO POINT CROSSOVER

##### A. Two Point Crossover 80 Crossover 15 Mutation 26 Chromosome

There is a maxima reached by about 150 generations. Further generations slowly make progress.

The final best fitness achieved: 0.152567334

The final key deciphered: uhisesashpeesecureatassshro

Mean 0.267277841

Standard Error 0.003190311

Median 0.229153679

Mode 0.152567334



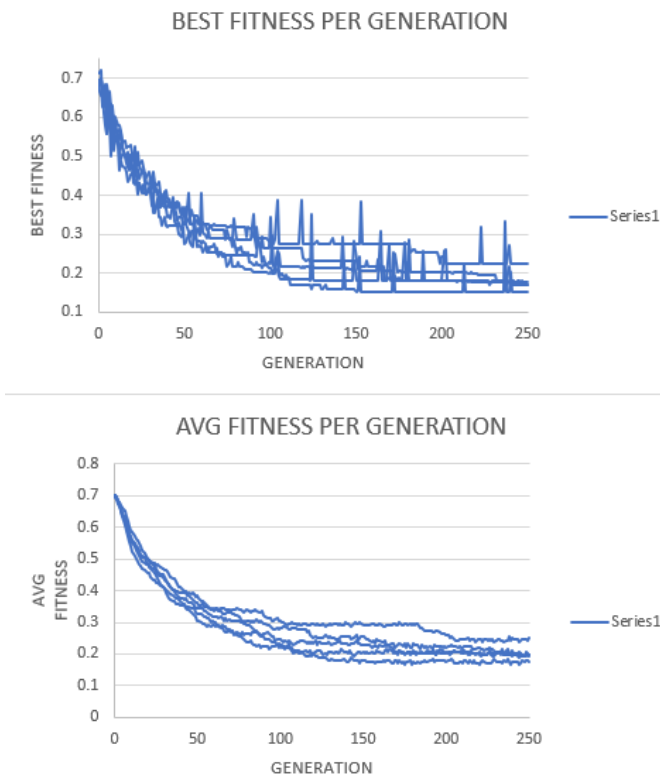


Fig. 14. Two Point Crossover 80 Crossover 15 Mutation 26 Chromosome best fit and avg fit per generation.

Standard Deviation 0.112115701  
 Sample Variance 0.01256993  
 Kurtosis 2.557761177  
 Skewness 1.618587888  
 Range 0.567787685  
 Minimum 0.152567334  
 Maximum 0.720355019  
 Sum 330.0881333  
 Count 1235  
 Largest(1) 0.720355019  
 Smallest(1) 0.152567334  
 Confidence Level(95.0%) 0.006259033

Figure 14: Best fit and avg fit per generation

#### B. Two Point Crossover 80 Crossover 15 Mutation 40 Chromosome

There is a local maxima reached by about 50 generations. Further generations slowly make progress.

The final best fitness achieved: 0.517741997

The final key deciphered: lnlitagiectiityiauploltigmzxi-itixryv

Mean 0.564738331  
 Standard Error 0.000967212  
 Median 0.555622645  
 Mode 0.551797096  
 Standard Deviation 0.034264424  
 Sample Variance 0.001174051

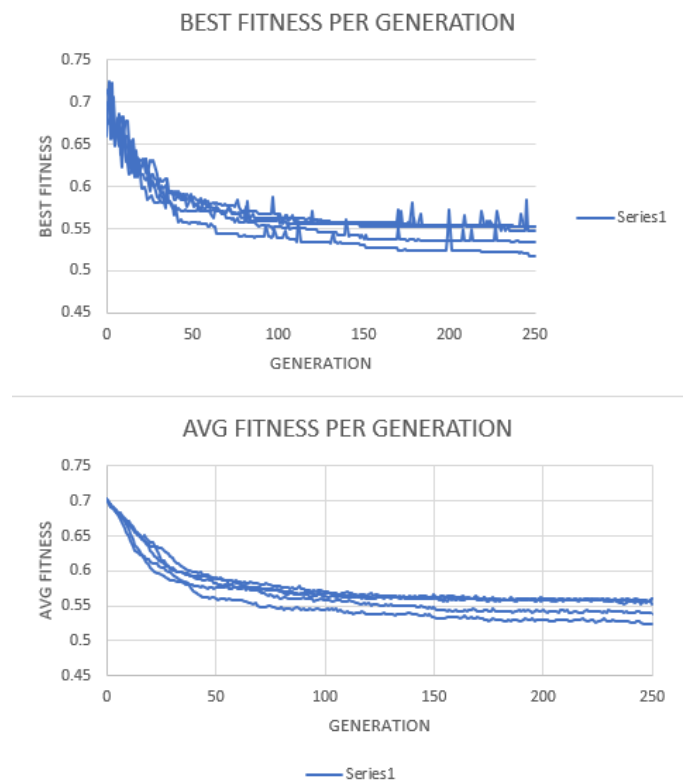


Fig. 15. Two Point Crossover 80 Crossover 15 Mutation 40 Chromosome best fit and avg fit per generation.

Kurtosis 4.009075079  
 Skewness 1.879997456  
 Range 0.207007965  
 Minimum 0.517741997  
 Maximum 0.724749962  
 Sum 708.7466049  
 Count 1255  
 Largest(1) 0.724749962  
 Smallest(1) 0.517741997  
 Confidence Level(95.0%) 0.001897532

Figure 15: Best fit and avg fit per generation

#### C. Two Point Crossover 90 Crossover 0 Mutation 26 Chromosome

There is a local maxima reached by about 25 generations. Further generations make no progress.

The final best fitness achieved: 0.479896931

The final key deciphered: ueesvjesvpevaosueptkkfmrnn

Mean 0.50853232  
 Standard Error 0.000912623  
 Median 0.500836418  
 Mode 0.500836418  
 Standard Deviation 0.032330554  
 Sample Variance 0.001045265  
 Kurtosis 17.42387397  
 Skewness 3.696532948  
 Range 0.247113769



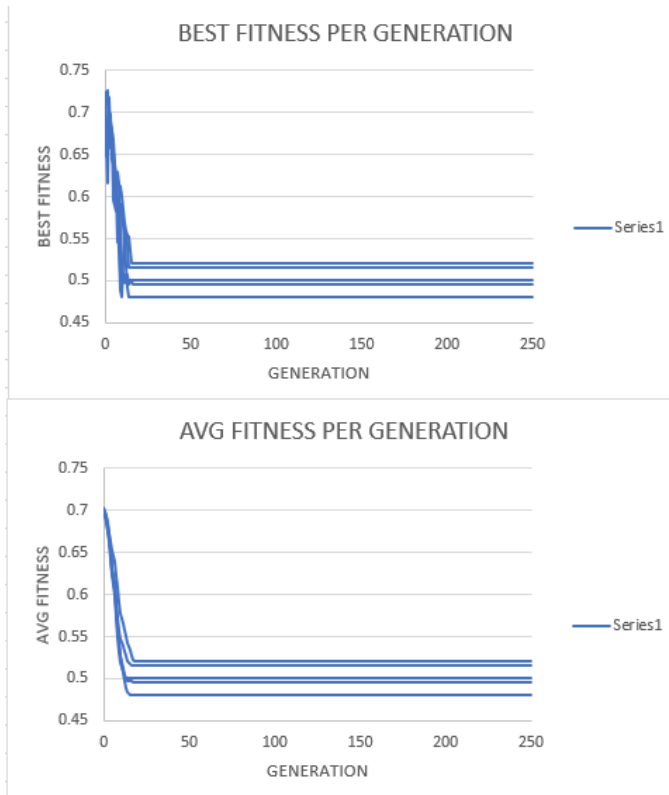


Fig. 16. Two Point Crossover 90 Crossover 0 Mutation 26 Chromosome best fit and avg fit per generation.

Minimum 0.479896931  
 Maximum 0.7270107  
 Sum 638.2080617  
 Count 1255  
 Largest(1) 0.7270107  
 Smallest(1) 0.479896931  
 Confidence Level(95.0%) 0.001790436  
 Figure 16: Best fit and avg fit per generation

#### D. Two Point Crossover 90 Crossover 0 Mutation 40 Chromosome

There is a local maxima reached by about 25 generations. Further generations make no progress.

The final best fitness achieved: 0.612117059

The final key deciphered: jjeubippfyveibscqy-qazuzrvcm-kfyjupazdb

Mean 0.622544525  
 Standard Error 0.00037965  
 Median 0.620657076  
 Mode 0.615215302  
 Standard Deviation 0.013449472  
 Sample Variance 0.000180888  
 Kurtosis 17.99201481  
 Skewness 3.78157748  
 Range 0.110213769  
 Minimum 0.612117059  
 Maximum 0.722330828

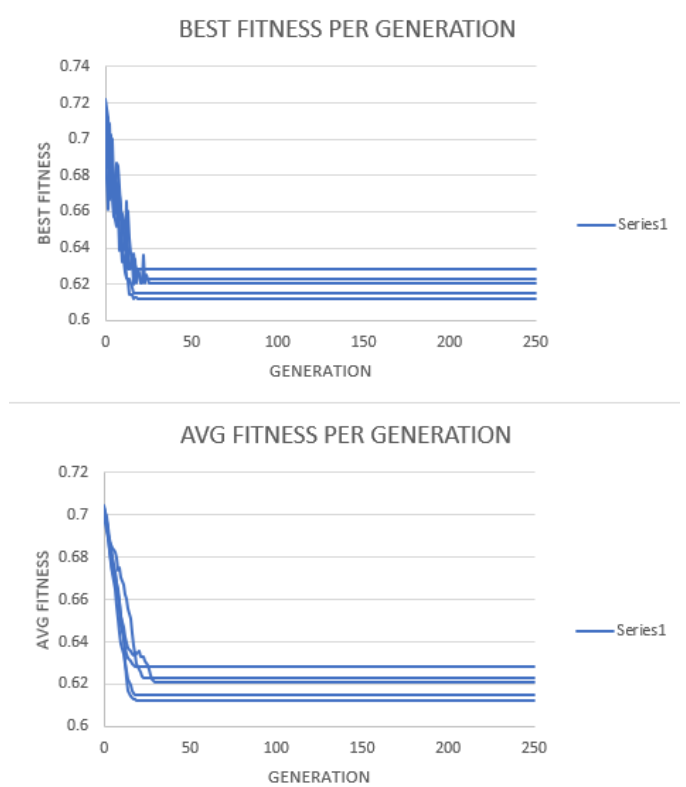


Fig. 17. Two Point Crossover 90 Crossover 0 Mutation 40 Chromosome best fit and avg fit per generation.

Sum 781.2933785  
 Count 1255  
 Largest(1) 0.722330828  
 Smallest(1) 0.612117059  
 Confidence Level(95.0%) 0.000744819  
 Figure 17: Best fit and avg fit per generation

#### E. Two Point Crossover 90 Crossover 10 Mutation 26 Chromosome

There is steady progress, but a local maxima is reached around 200 generations. From there, it slowly progresses.

The final best fitness achieved: 0.146542702

The final key deciphered: uhcshsasutersecuretasmmorh

Mean 0.285036139  
 Standard Error 0.003176685  
 Median 0.24633921  
 Mode 0.173527183  
 Standard Deviation 0.112537191  
 Sample Variance 0.012664619  
 Kurtosis 2.808251728  
 Skewness 1.642446138  
 Range 0.56895181  
 Minimum 0.146542702  
 Maximum 0.715494512  
 Sum 357.720355  
 Count 1255  
 Largest(1) 0.715494512

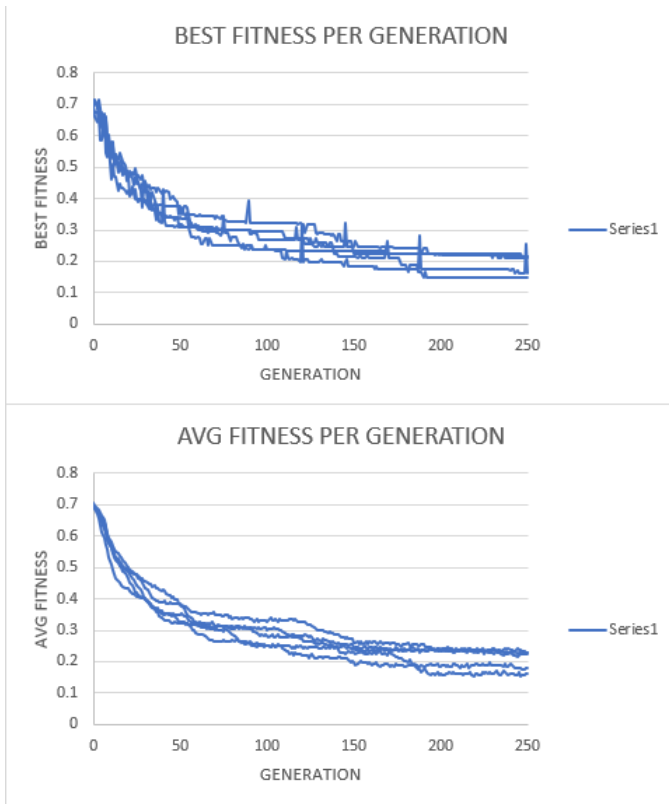


Fig. 18. Two Point Crossover 90 Crossover 10 Mutation 26 Chromosome best fit and avg fit per generation.

Smallest(1) 0.146542702

Confidence Level(95.0%) 0.006232204

Figure 18: Best fit and avg fit per generation

#### F. Two Point Crossover 90 Crossover 10 Mutation 40 Chromosome

There is steady progress, but a local maxima is reached around 50 generations. From there, it slowly progresses.

The final best fitness achieved: 0.535605981

The final key deciphered: niaqbpbiufmpqbppluetiethph-lobgifxtlppa

Mean 0.567165058

Standard Error 0.000964177

Median 0.554731754

Mode 0.544410403

Standard Deviation 0.0341569

Sample Variance 0.001166694

Kurtosis 3.700770736

Skewness 1.965028767

Range 0.182134185

Minimum 0.535605981

Maximum 0.717740166

Sum 711.7921472

Count 1255

Largest(1) 0.717740166

Smallest(1) 0.535605981

Confidence Level(95.0%) 0.001891577

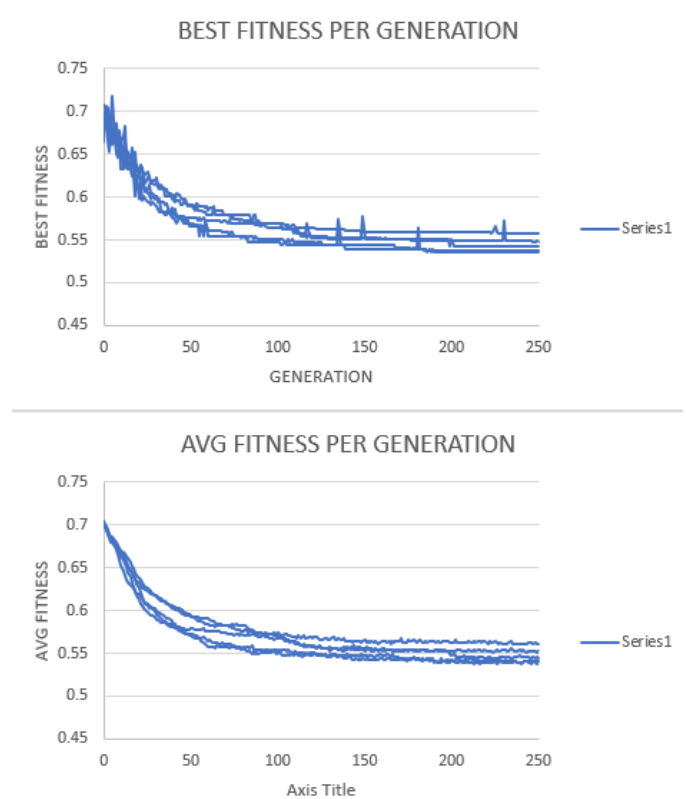


Fig. 19. Two Point Crossover 90 Crossover 10 Mutation 40 Chromosome best fit and avg fit per generation.

Figure 19: Best fit and avg fit per generation

#### G. Two Point Crossover 100 Crossover 0 Mutation 26 Chromosome

Local maxima reached at 20 generations, no progress is made after.

The final best fitness achieved: 0.501973756

The final key deciphered: mipsiglsqsegqjuuijpamyvamm

Mean 0.524077997

Standard Error 0.000842842

Median 0.519132802

Mode 0.506134407

Standard Deviation 0.029858501

Sample Variance 0.00089153

Kurtosis 16.13407584

Skewness 3.588471497

Range 0.216198122

Minimum 0.501973756

Maximum 0.718171878

Sum 657.7178865

Count 1255

Largest(1) 0.718171878

Smallest(1) 0.501973756

Confidence Level(95.0%) 0.001653536

Figure 20: Best fit and avg fit per generation

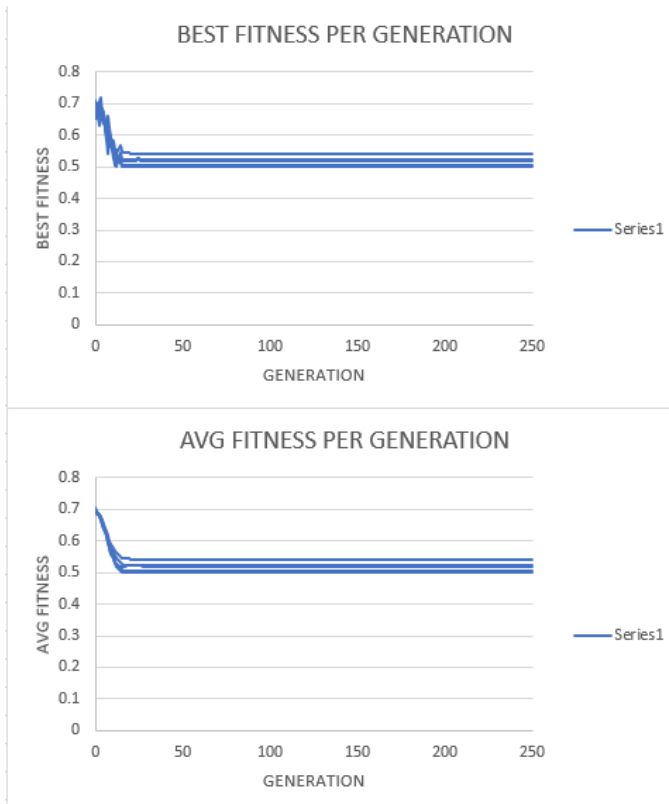


Fig. 20. Two Point Crossover 100 Crossover 0 Mutation 26 Chromosome best fit and avg fit per generation.

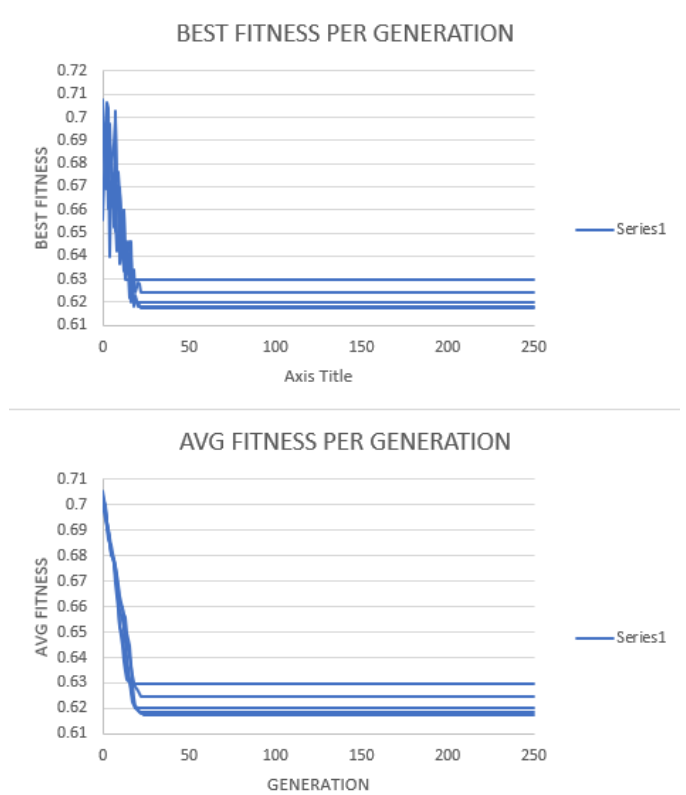


Fig. 21. Two Point Crossover 100 Crossover 0 Mutation 40 Chromosome best fit and avg fit per generation.

#### H. Two Point Crossover 100 Crossover 0 Mutation 40 Chromosome

Local maxima reached at 20 generations, no progress is made after.

The final best fitness achieved: 0.617472834

The final key deciphered: rqrzchla-qgmqgbbxkatzkeppxmzisesataufjp

Mean 0.624811455

Standard Error 0.000357472

Median 0.620070379

Mode 0.629714619

Standard Deviation 0.012663789

Sample Variance 0.000160372

Kurtosis 17.30175736

Skewness 3.894303585

Range 0.090320894

Minimum 0.617472834

Maximum 0.707793728

Sum 784.1383757

Count 1255

Largest(1) 0.707793728

Smallest(1) 0.617472834

Confidence Level(95.0%) 0.000701309

Figure 21: Best fit and avg fit per generation

#### I. Two Point Crossover 100 Crossover 10 Mutation 26 Chromosome

Local maxima reached at 50 generations, slow progression afterwards.

The final best fitness achieved: 0.171146172

The final key deciphered: shchesasupersacuqepassorp

Mean 0.282928873

Standard Error 0.003102733

Median 0.233113512

Mode 0.210342975

Standard Deviation 0.109917359

Sample Variance 0.012081826

Kurtosis 3.044151518

Skewness 1.842486696

Range 0.569783419

Minimum 0.171146172

Maximum 0.740929591

Sum 355.0757359

Count 1255

Largest(1) 0.740929591

Smallest(1) 0.171146172

Confidence Level(95.0%) 0.00608712

Figure 22: Best fit and avg fit per generation

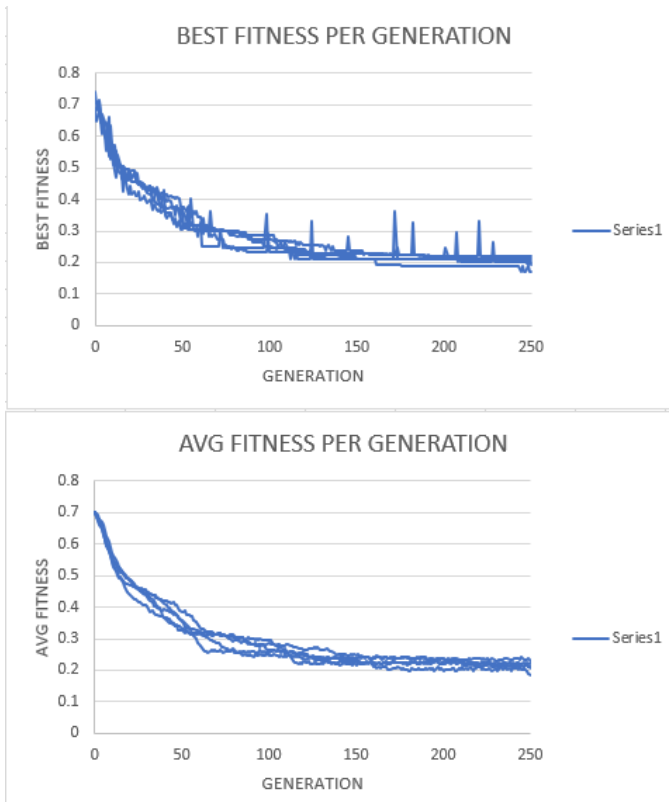


Fig. 22. Two Point Crossover 100 Crossover 10 Mutation 26 Chromosome best fit and avg fit per generation.

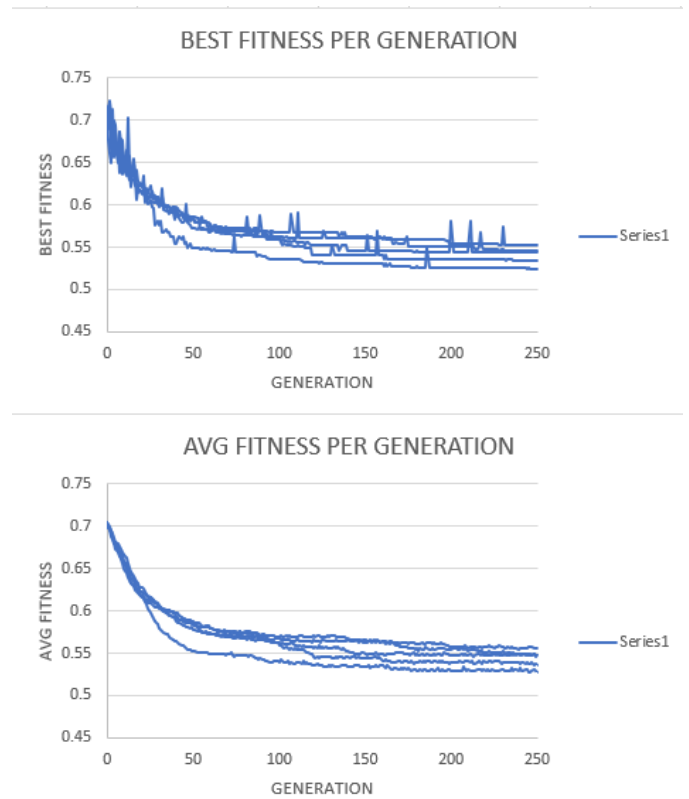


Fig. 23. Two Point Crossover 100 Crossover 10 Mutation 40 Chromosome best fit and avg fit per generation.

#### J. Two Point Crossover 100 Crossover 10 Mutation 40 Chromosome

Local maxima reached at 50 generations, slow progression afterwards.

The final best fitness achieved: 0.524590213

The final key deciphered: ktaeibitiiljftitilibuppostfgecip-plbumcvd

Mean 0.564563596

Standard Error 0.000981003

Median 0.555582861

Mode 0.52612043

Standard Deviation 0.034752985

Sample Variance 0.00120777

Kurtosis 3.850103556

Skewness 1.862111148

Range 0.198591728

Minimum 0.524590213

Maximum 0.723181941

Sum 708.5273132

Count 1255

Largest(1) 0.723181941

Smallest(1) 0.524590213

Confidence Level(95.0%) 0.001924588

Figure 23: Best fit and avg fit per generation

#### IX. DISCUSSIONS AND CONCLUSIONS

The algorithm performed better with random mutation, crossover percentage was relatively ineffective at predicting performance, but still completely necessary to maintain elitism.

Overall, single point crossover with a random flip point performed better than two static point crossover, this is probably because it explored more options, similar to what mutation does to the chromosome, avoiding inbreeding.

It is clear, without some random mutation (for both crossover functions, regardless of crossover percentage), that a local maxima is reached and never expanded upon: the inbreeding becomes too strong.

A similarly scoring best fitness was achieved on multiple parameters, assuming they at least had some mutation and some crossover, the results lead to a belief that with too much crossover and not enough mutation, you will get stuck too early, and with too much randomness, you will never reach a maxima, as the maxima will be lost due to randomness.

Going forward, a longer, much higher generation study should be performed. This can be done without testing for 0 mutation percentage as the maxima will be reached quickly. This would be interesting to see how the algorithm performs at more complex tasks such as the 40 chromosome decryption. To see if the algorithm can make progress much deeper into the potential search space.

## REFERENCES

- [1] Mitchell, Melanie (1996). An Introduction to Genetic Algorithms. Cambridge, MA: MIT Press.
- [2] Khanduri, A. (2019, September 27). Vigenère Cipher. Retrieved from <https://www.geeksforgeeks.org/vigenere-cipher/>.
- [3] Bridge, D. (n.d.). Retrieved from <http://www.cs.ucc.ie/dgb/courses/tai/notes/handout12.pdf>.