Project Report

DoReMi Delivery Optimization System

Group members

Chad Ng

Jin Xin

Xu Dongbin

Sun Hang

Li Xin

1.0 EXECUTIVE SUMMARY

In this report, a delivery routing optimization system is introduced for DoReMi Books Inc. DoReMi Books Inc. is a company started in 1955 and specialize in the supply and sale of classical music scores and music books in the USA with subsidiaries in the major cities of each of the 50 states. In the previous project, our team focused on the automated system for the online music books sale transaction and order handling process as well as introduce improved stock and inventory planning and management capabilities. To further improve the intelligent business process capability, a delivery routing optimization system is necessary because it not only improve the efficiency of delivery process but also directly related to the satisfaction of customer experience.

As a traditional company, DoReMi Books Inc. is encountering changelings from its competitors who have brought on by online sales transactions. In comparison, some of its business processes are still manual and rather traditional. In response to the new challenges, the company decided to conduct a business process improvement exercise to implement intelligent planning and process optimization capabilities.

Our group of five members, brainstormed and decided to build a reasoning system which need conduct search and optimization to get optimal (sub-optimal) solutions. The system we're focusing in this group project is a delivery routing optimization system with hybrid optimization techniques (Clustering, GA and Machine learning) and automate. To be realistic, we use the real cities in one state (Iowa) in United States.

Our primary objective is to apply the search and optimization techniques in DoReMi's delivery routing system. It will significantly improve the delivery efficiency by saving the cost and time, enhance customer satisfaction, and thus competence. Eventually will benefit for the companies in the same industry area.

2.0 PROBLEM DESCRIPTION

With the new online ordering system, DoReMi Books have increased their business very fast and also facing more pressures on order delivery at the same time. All the subsidiaries in the major cities of 50 states of USA are hungry of stock being topped up as soon as possible.

To better meet the demands of each city, DoReMi needs to deliver the products more efficiently, it should be fast but not cost too much. The most cost-efficient solution is using trucks to deliver, one truck is able to travel up to 1000 km per day with high capacity and is able to serve several cities per day. However, there are too many cities in USA and the demands are keep changing, that causes lots of problem when planning the delivery with trucks:

1. How many products should each truck carry?
2. What is the sequence to deliver to each city?
3. How many trucks should be used?
4. How long will each truck take to finish the delivery?

These problems are detailed but reliable, a best solution will increase the delivery efficient easily.

The management and company have enjoyed the benefit from the on-line ordering system, so they decided to improve the process again with a Delivery Optimization System to provide the best delivery process.

2.1 PROJECT OBJECTIVE

We aim to create a DoReMi's Delivery Optimization System to optimize the delivery solution with trucks in each state of America. The system will give the best order delivery solution in the state. With the solutions, the company will have a high efficiency on order delivery and also the cost of delivery will be reduced.

By implementing the proposed new reasoning system, we believe DoReMi Books Inc will begin to deliver order more efficiently over time as the full system gets implemented. DoReMi is set to become much more competitive in the market in near future.

2.2 PROJECT SCOPE

The proposed new Delivery Optimization System will focus on Iowa state of USA on:

(1) a K-means cluster to divide the target state into numbers of area, the same numbers of trucks will be sent for delivery, the capacity of each area will determine the size of truck used.

(2) a Genetic Algorithm to calculate and give the shortest route to link the cities in each clustered area.

(3) The system will give a very direct interface for the user to use, with simple click of button, the system will provide the solutions in Iowa state, and the distance of each truck will travel is shown with the capacity requirement of the route and recommended type (size) of truck.

2.3 PROJECT ASSUMPTIONS

All the delivery will start from the central city Des Moines in Iowa which is also the printing center in the state.

All the delivery will start from Des Moines and return to same place when delivery completed.

All the delivery will finish within one day.

All the cities are available to receive the delivery 24 HRs.

2.4 PROJECT TEAM

The project members titles and roles are listed as followings:

Nr.	Name and Title	Role
1	Chad Ng Choon Beng (Project Lead)	Project Report, Systems Consultation, Timeline, overall conceptualization
2	Xu DongBin (Algorithm Analyst)	Algorithm for Application, App Testing, Use case Prep, Documentation, help with project report
3	Li Xin (Algorithm Analyst)	Algorithm for Application, App Testing, Use case Prep, Documentation, help with project report
4	Sun Hang (Architecture and Design)	Solution Design and Software development
5	Jin Xin (Architecture and Design)	Solution Design and Software development

The team has met up every week to discuss the project status and updates. Tasks were assigned and areas of responsibilities were clearly outlined for each member.

We used Zoom-US video conference call platform to host our meeting once a week and when required. The project team lead recorded meetings recaps for each meeting conference call. The daily communication between team members were done through whatsApps, and our private Slack channel. Resources and sharing were done through Google Share Drive, Slack and Github.

Our team worked together on researching materials, case studies, doing the needed systems analysis and brainstorming and making sure we cover all possible areas of improvement for DoReMi delivery case of this project. The project is successfully executed, and the proposed solution is implemented on time and on target.

3.0 DATA ACQUISITION AND OPTIMIZATION ALGORITHM

3.1 DATA ACQUISTION

In this report, our group focuses on the application of this system in the real situation to solve the real problem. To achieve this objective, it is critical to use a real case in the real world. As a result, Iowa State was chosen and used as an example for system demonstration. All the cities using in the system are from the real map. In current project, the music books delivery optimization system we're building focuses on Iowa state first, after testing and improvement during system refinement cycle, can extend this system to other states or even other countries also. The route distance between each city were got by calculating the distance between each city as shown in Fig. 1.

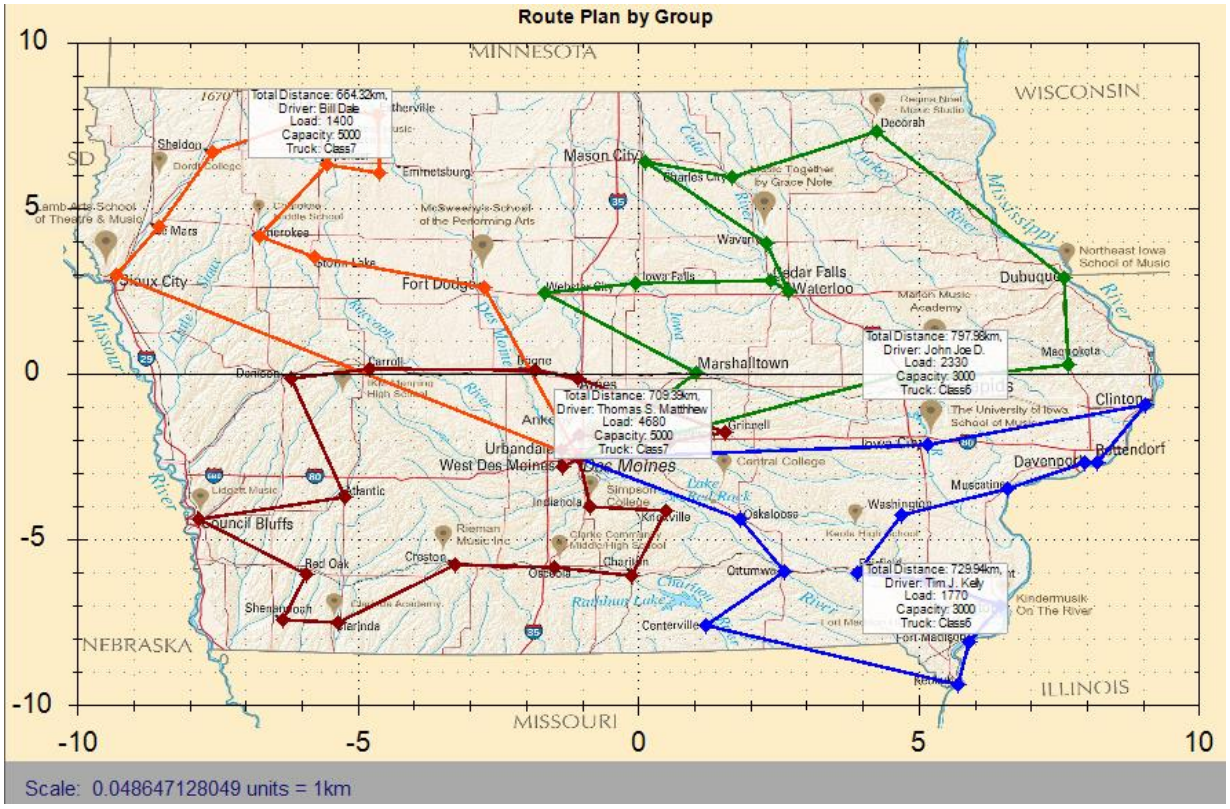

Fig. 1 System demonstration in Iowa state

3.2 OPTIMIZATION ALGORITHM

Nowadays, optimization techniques are widely used in a variety of business areas. For example, Consumers make investment decisions to maximize return, Businesses make pricing decisions to maximize profits, Travelling Salesman Problem (TSP), *et al*. Basically, reasoning systems need conduct search and optimization to get optimal (sub-optimal) solutions. One of the most popular and promising optimization algorithm is genetic algorithm (GA).

In genetic algorithm, the optimization method mimic the genetic process in biological science. The typical steps are (as shown in Fig. 2):

- Generate initial population which comprise a group of random tours
- Choose 2 better parents from population and combine them with crossover to get 2 child tours, the 2 child tours may have high chance to be better than both parents
- To prevent trapped in localized optimized zone, it is essential to have mutation in a certain rate (current set as 3%).
- The offspring are evaluated in terms of fitness score, if demonstrate to be better, then replace the 2 parents
- Repeat above procedures until get an acceptable fitness score

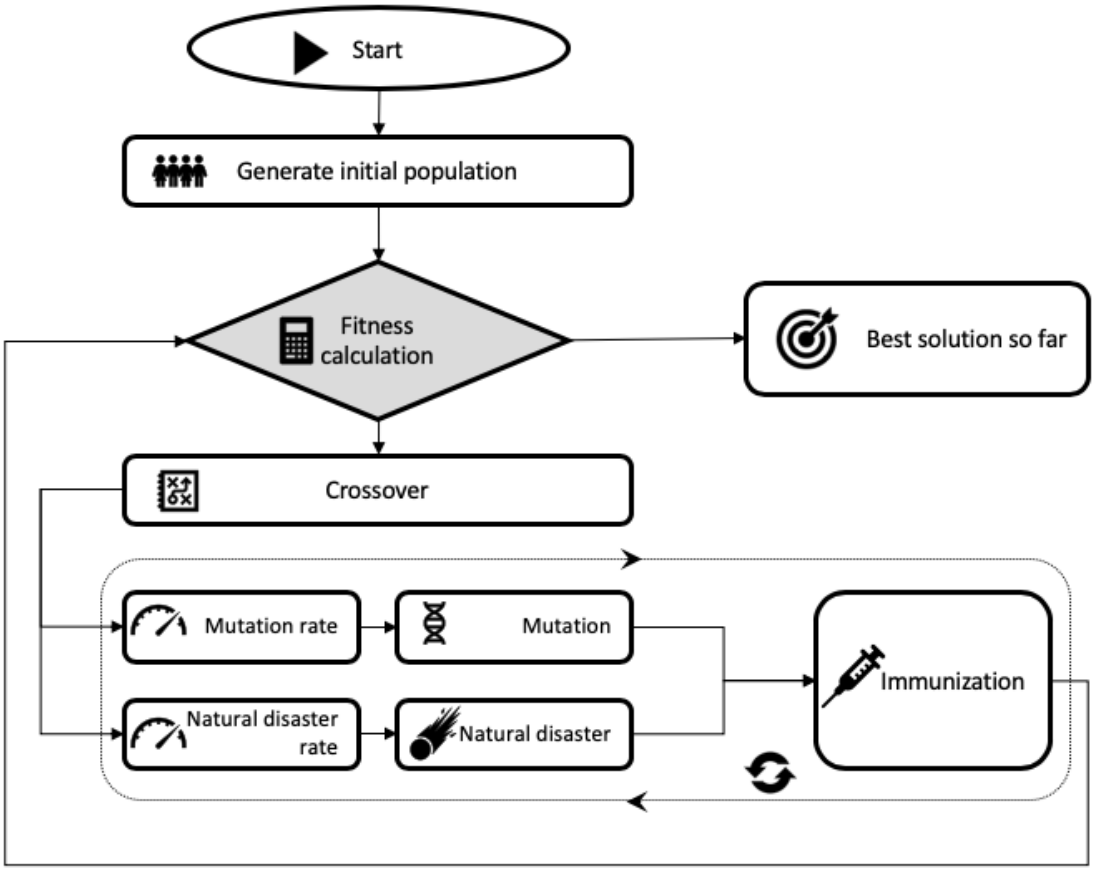

Fig. 2 Flowchart for the improved genetic algorithm

In this project, in order to further improve the capability of genetic algorithm, we used two more algorithm which are natural disaster and immunization. The natural disaster is carried out during the same time as mutation with a certain natural disaster rate (current set as 0.02%), while the immunization is processed after mutation and/or natural disaster. The detailed process for crossover, mutation, natural disaster and immunization are explained in the following part.

- **Chromosome representation**: An array with the length equals to the number of cities were used as the chromosome representation in genetic algorithm calculation.
- **Cross over**: The cross over algorithm in current project is shown in Fig. 3. To create a child tour, firstly select the first city (A) from parent 1, then search for the position of A in parent 2, get the city just after A which is C. Then repeat the above process, search the position of C in parent 1, get the city just after C which is D. Keep repeating this process will make sure child tour can inherit the good gene from both parents, and thus better chance to get higher fitness score than both of parents.

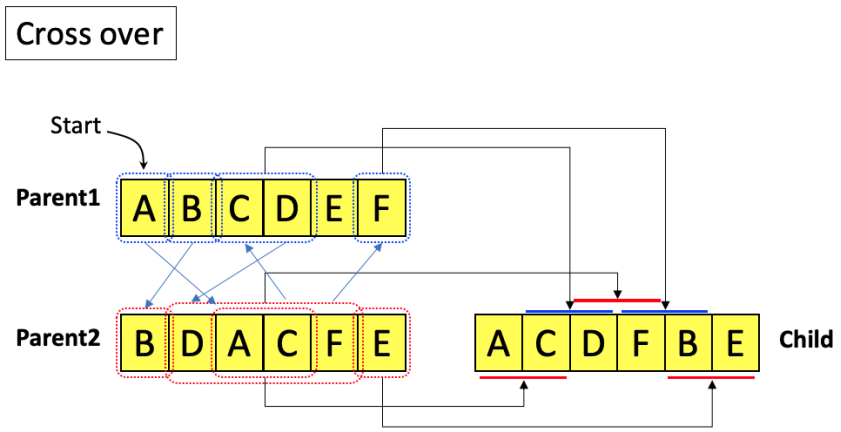

Fig. 3 Crossover algorithm in GA

- **Mutation**: As shown in Fig. 4, mutation typically is an operation to random rearrange the gene in a relatively small region with a certain mutation rate. In current project, mutation algorithm is set to randomly exchange two adjacent cities. The reason to use random position for exchange is because this can make the whole population not be trapped in a local optimization (sub-optimization).

- **Natural disaster**: Just like in the biological science, natural disaster may totally rearrange the gene and give the opportunity for new species to emerge. In this algorithm, we keep the start city position and shuffle all the rest cities randomly. This may lead the optimization to another direction with better fitness score, but the rate is set as low as 0.02% to prevent the shuffle too frequent.

- **Immunization**: Immunization algorithm used here is a criterion check for the tour from the first city to the second city.[1] The basic idea is to filter out the cities which have the distance longer than the average distance as shown in Fig. 5. Since in the initial stage, the connection from the first city is random, this algorithm can only allow the offspring with potential for better fitness score to pass. Therefore, it can improve the quality of the generation and make the optimization more efficient.

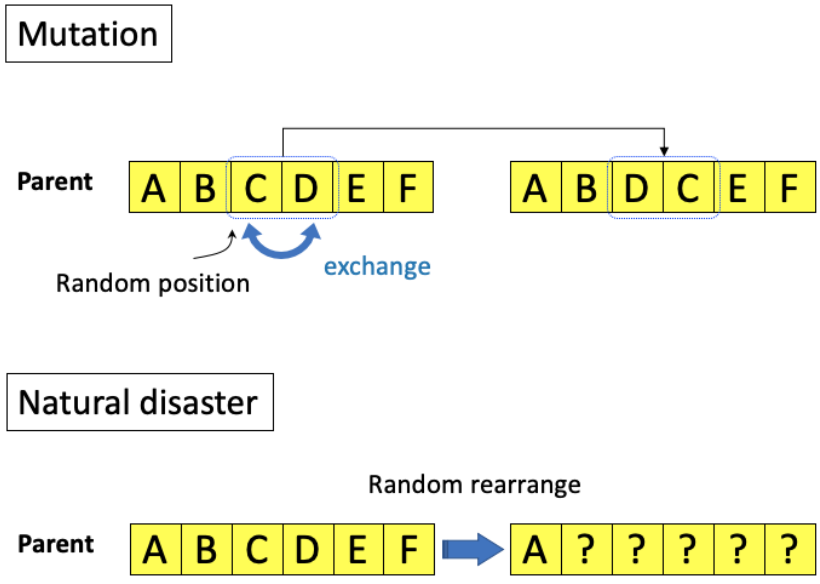

Fig. 4 Mutation and Natural disaster algorithm in GA

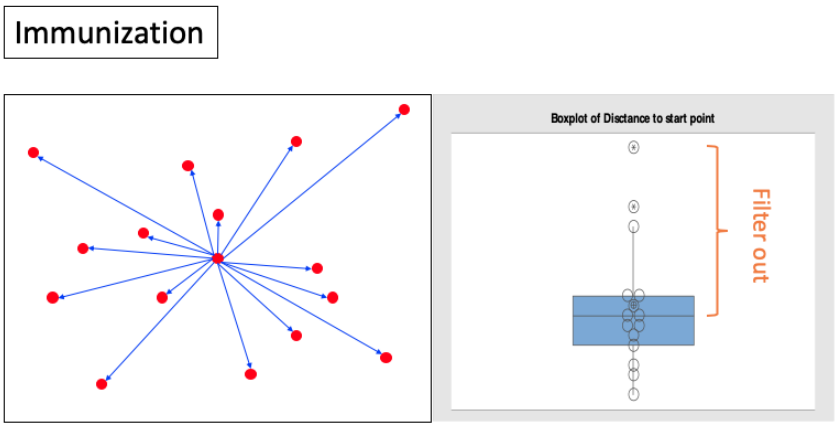

Fig. 5 Immunization algorithm in GA

- **Evaluation**: A preliminary evaluation was done for the proposed algorithm in GA as shown in Fig. 6. It shows at early stage (lower iteration), three algorithms show no obvious difference in fitness optimization. While at higher iteration, the strength of mutation, natural

disaster and immunization began to emerge. It has higher possibility to escape the local optimum and further improve the fitness in the long run.

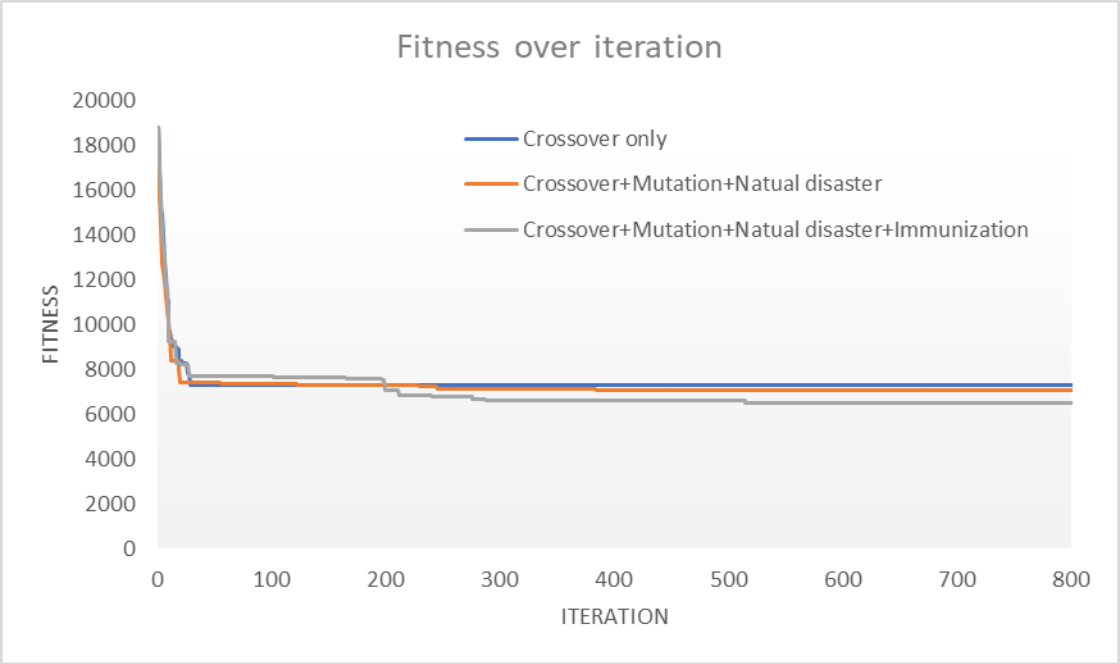

Fig. 6 Preliminary evaluation of mutation, natural disaster and immunization algorithm in GA

3.3 ALGORITHM LIMITATION AND IMPROVEMENT

One of the limitations of genetic algorithm is the optimization may be trapped in a local optimum then it may restrict the algorithm from escape and find the global optimum. In this project, 3 methods (mutation, natural disaster, immunization) were used in the algorithm to avoid this situation and mitigate the drawback. Diversity is critical and essential in generic algorithm since the crossover within homogenous generation will not give new/better solution.

Further improvement may include the feedback from customer and employee as a spiral feedback loop for continuing improvement as shown in Fig.6. More constrains will be considered to add in the optimization to make the delivery more efficient and accurate. However, if there are too many constrains with large number of cities, the optimization process could be too slow and need longer time to converge, as a result, it is critical to compromise between constrains number and system efficiency.

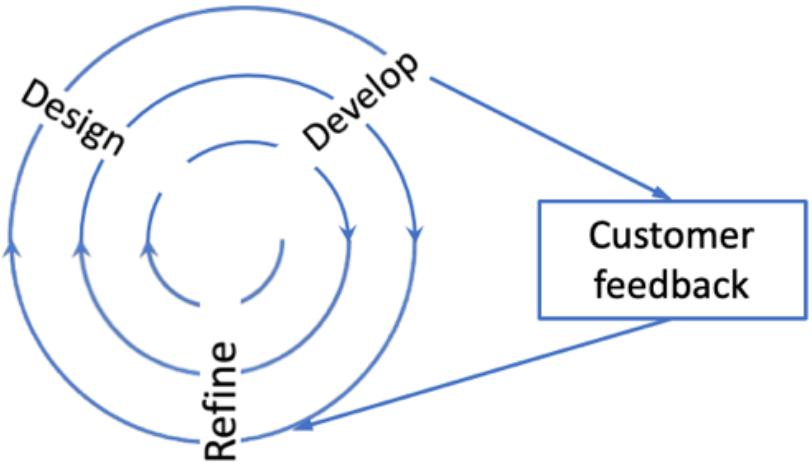

Fig. 6 Spiral feedback loop for continuing improvement

4.0 Solution Outline

The Delivery Optimization System is a Hybrid Reasoning System used Genetic Algorithm (GA) and K-means Cluster to simplify the problem and optimize the solution. K-means cluster will cluster the data as input to Genetic Algorithm and the Genetic Algorithm will calculate the best solution and the max fitness function will feedback to the system to make the system keep running until getting the optimal solution.

4.1 SYSTEM ARCHITECHTURE

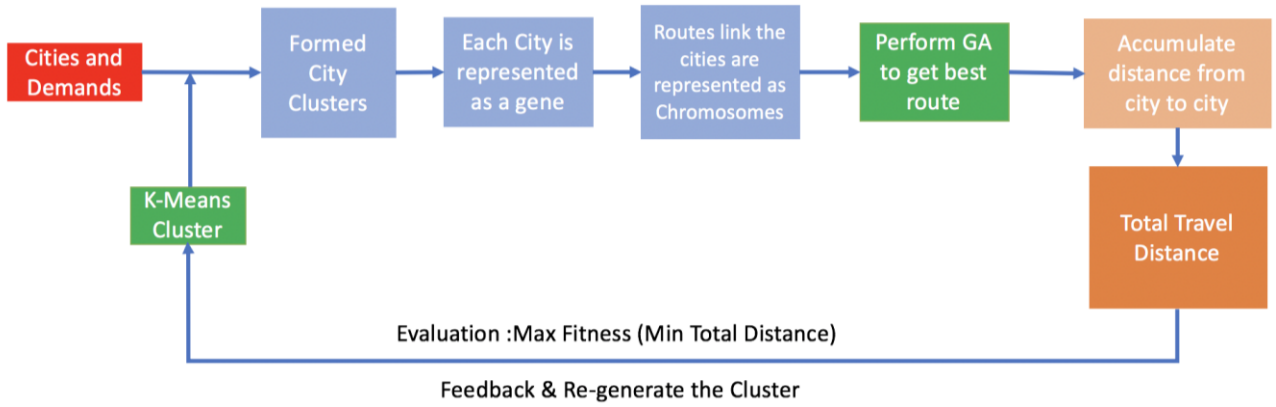

Fig. 7 Block Diagram of DoReMi Book Delivery Optimization System

Upon the orders details, we know the demands of each city in the state Iowa. The data of the cities can be clustered by K-Means clustering method by distance. The K-Means method will cluster the state into 4 areas, each area will be assigned one truck for the delivery, the different capacity trucks will be selected and determined by the total demands of the area.

Then the system will randomly generate two possible routes within each area and these two routes will be used as parents in the genetic algorithm, thus the optimal delivery routes for each area can be generated.

The system will calculate the total travelling distance of all trucks, by going through the max fitness function which requires minimizing the total travel distance, the result will feedback to the system to re-generate the K-Means cluster and keep improving to get the optimal solution.

4.2 SYSTEM'S FEATURES

Ease of Use

The system is easy to be used, with simple click, the delivery route, travel distance, capacity, recommended truck even the driver information will be shown in the page.

Efficient

The system is designed to give the optimal delivery solution to the case, the solution will reduce the delivery time and thus to save lead time of order, save man power and also the petrol costs.

Documented

The system will also save the time to do additional record for each day's delivery information, once follow the recommendation from the system, the company with have those data.

4.3 LIMITATIONS

The system is still under development, there are some unavoidable limitation at current level:

1. Currently only one state data has been added in, the user is not able to select other states in USA.
2. The number of trucks used is fixed at 4 currently, the user is not able to change the number.
3. The system could not solve the problem if different cities have different available and limited delivery slots.

5.0 CONCLUSION & REFERENCES

In this project, our group deployed genetic algorithm and K-means cluster to a hybrid reasoning system which is used to optimize the delivery solution for the customer DoReMi Books Inc. The K-means and genetic algorithm are used and effectively working in the system to simplify the problem, the solution can be easily acknowledged as effective from the fitness outcome.

In development process, our group with five members had very active and fruitful discussions and we learnt a lot from each other since we're from different professions with different skill sets. It is a good opportunity for us not only to apply the reasoning system knowledge and solve real-life problems but also a wonderful time for every group member to experience and learn new things.

5.1 IMPROVEMENTS:

The system will definitely improve the efficiency for DoReMi Books Inc delivery, as this is the Phase one and was done in a limited time, some improvements are expected in the Phase two.

In Phase two, if different cities have different available delivery slot, the system should be able to improve the solution further base on the genetic algorithm to meet all the time window requirement and also the shortest route of delivery.

The demand and delivery time slot of each city will be updated by the on-line ordering system automatedly, it will save more time and reduce the chance of typo.

All the 50 states data should be added in to the system, the user will able to select the state they are planning.

Last but not least, accounts are required to access the system to protect the company critical data and information.

6.0 BIBLIOGRAPHY

1. LC Jiao, and L Wang, "A Novel Genetic Algorithm Based on Immunity," IEEE Transactions on Systems, Man, AND Cybernetics—Part A: Systems and Humans, Vol. 30, No. 5, Sep. 2000