MODIFIED K-NEAREST NEIGHBORHOOD ALGORITHM FOR OPTIMAL SELECTION OF DISTRIBUTION CENTRE IN THE DISASTER RELIEF OPERATION

SITI NABILAH BINTI BASARANG

UNIVERSITI SAINS ISLAM MALAYSIA

MODIFIED K-NEAREST NEIGHBORHOOD ALGORITHM FOR OPTIMAL SELECTION OF DISTRIBUTION CENTRE IN THE DISASTER RELIEF OPERATION

Siti Nabilah Binti Basarang
(Matric No. 3150104)

Thesis submitted in fulfilment for the degree of
MASTER OF SCIENCE

Faculty of Science and Technology
UNIVERSITI SAINS ISLAM MALAYSIA
Nilai

NOVEMBER 2020

BIOGRAPHY

Siti Nabilah Basarang was born on October 31, 1992. She received Bachelor of Computer Science with Honours (Information Security and Assurance) from Universiti Sains Islam Malaysia in 2015. Upon graduation, she has served as Network Engineer in multinational company in Cyberjaya while at the same time pursuing her master's degree by research in the field of Computer Science.

ACKNOWLEDGEMENTS

I would like to First and foremost, I am thankful to Allah S.W.T that through His Blessings and Guidance I am able to accomplish my master's degree thesis entitled "Modified K-Nearest Neighborhood Algorithm For Optimal Selection Of Distribution Centre In The Disaster Relief Operation"

This thesis dedicated to my family, parents and siblings. Thank you for all the love, concern, and strength they provide me in whatever that I choose to do in my life,
I would like to express my deep and sincere gratitude to my supervisor, PM Dr. Waidah Ismail and Dr Nurdiana Azizan, with their enthusiasm, inspiration, expertise, patience and great efforts to explain every single thing clearly and simply have provided a good basis for the present thesis. Throughout my writing phase, they provided encouragement, sound advice, good teaching, and lots of good ideas. I pray to Allah S.W.T they will succeed in every dream that they want to achieve.

Besides, I would like to say big thanks to my course mates that always supporting me. Not forgotten to all lectures and students especially in Faculty of Science and Technology. Thanks for your time, feedback, comment and suggestions.

Finally, I would like to thank again to all of you for your cooperation and may Allah bless all of you for your commendable contributories. I pray to Allah SWT to accept this genuine endeavour as a sincere deed.

ABSTRAK

Operasi bantuan bencana merujuk kepada aktiviti di mana orang membantu mangsa bencana untuk pulih. Bencana boleh dibahagikan kepada bencana alam atau disebabkan perbuatan manusia. Di Malaysia, banjir adalah bencana yang paling terkenal dan ia terus berulang hampir setiap tahun. Kesan kelemahan dalam pemilihan pusat pengedaran untuk operasi bantuan bencana menjadikan sukarelawan sukar menjalankan tugas mereka. Salah satu skop kerja dalam operasi bantuan bencana ialah mengedarkan keperluan mangsa bencana. Oleh itu, lokasi strategik memilih pusat operasi menjadi kebimbangan. Berdasarkan pengalaman sebelum ini, terdapat aduan bahawa tidak semua kawasan bencana dapat dibantu semasa operasi pemulihan bencana. Permasalahan utama dalam memberikan bantuan tidak dapat dilakukan dengan adil. Metod yang digunakan adalah K-Means, K-Means bersama Simulated Annealing (SA) dan K-Means bersama Genetic Algorithm (GA). Sebagai tindak balas kepada masalah ini, kajian telah dilakukan untuk memahami algoritma sedia ada yang digunakan untuk menentukan pusat pengedaran dan membandingkan prestasi K-Nearest Neighbor (KNN) dengan versi yang dipertingkatkan seperti yang dicadangkan dalam kajian ini. Eksperimen yang dimulakan dengan penyelidik menetukan lokasi target dan pusat pengedaran dan melengkapi KNN dengan Algoritma Genetik (GA) dan Simulated Annealing (SA). Perbandingan prestasi telah dibuat dan hasilnya mencadangkan penggunaan GA-KNN dan peratus 21% lebih rendah dibandingkan dengan SA-KNN. Sekaligus kajian ini membantu untuk mencari lokasi pusat pengedaran yang paling optimum dengan pengedaran lokasi target hampir sama setiap pusat pengedaran yang dapat memudahkan usaha pengedaran bantuan untuk masalah nyata dalam bencana kawasan, menjimatkan masa dan kos.

ABSTRACT

Disaster relief operation is refer to an activity where people assist the disaster victim to recover. Inefficiency of distribution centre selection in disaster relief operation makes difficulty for volunteer to perform their humanitarian task. Thus, a strategic location choose of operation centre is being a concern. It has been pointed out that not all disaster area can be covered during disaster recovery operation . The problem of the selection of distribution centre is not done optimally. The methodology by comparison between K-means, K-means with Simulated Annealing (SA), lastly K-means with Genetic Algorithm (GA). In response to the problems, it is needed to understand the existing algorithm used to determine the distribution centre. The K-Nearest Neighbor (KNN) and the use of Genetic Algorithm (GA) and Simulated Annealing (SA) in KNN is proposed to classify and select the distribution centre. The minimization of the fitness value is being the objective in this study. The experiment conducted with demand point and the distribution centre are located by researcher and complement the KNN with the GA and SA. The comparison of the performance has been made and the study found that implementing GA-KNN give the most optimal solution with average 21% of fitness value least compared to SA-KNN. Thus, this study is contributing in finding the most optimal distribution centre location with nearly-equal demand point distribution of each selected location which can facilitates the real-world aid distribution in disaster area, time-wise and cost-wise.

Keywords: distribution centre, optimization, fitness value

ملخص البحث

تشير عملية الإغاثة في حالات الكوارث إلى نشاط يقوم فيه الأشخاص بمساعدة ضحايا الكوارث على التعافي. يمكن أن تكون الكوارث كارثة طبيعية أو بشرية. في ماليزيا ، يعد الفيضان أكثر الكوارث شهرة ويتكرر كل عام تقريبًا. عدم كفاءة اختيار مركز التوزيع في عملية الإغاثة في حالات الكوارث يجعل من الصعب على المتطوع أداء مهمته الإنسانية. يتمثل أحد نطاق العمل في عملية الإغاثة في حالات الكوارث في توزيع الاحتياجات على ضحية الكوارث. وبالتالي ، فإن اختيار موقع استراتيجي لمركز العمليات يعد مصدر قلق. لقد تمت الإشارة إلى أنه لا يمكن تغطية جميع مناطق الكوارث أثناء عملية التعافي من الكوارث. استجابة للمشاكل ، يلزم فهم الخوارزمية الموجودة المستخدمة لتحديد مركز التوزيع. يُقترح أن يكون (GA) والتقليد الصلب (KNN) K-Nearest Neighbor واستخدام الخوارزمية الوراثية لتصنيف وتحديد مركز التوزيع. التقليل من قيمة اللياقة البدنية هو الهدف في هذه الدراسة. وتقع KNN في (SA) تم إجراء .SA و GA مع KNN التجربة التي أجريت مع نقطة الطلب ومركز التوزيع من قبل الباحث واستكمال يوفر الحل الأمثل مع متوسط 21٪ من قيمة اللياقة البدنية GA-KNN مقارنة بين الأداء ووجدت الدراسة أن تطبيق وبالتالي ، تساهم هذه الدراسة في إيجاد أفضل موقع لمركز التوزيع مع توزيع نقطة طلب .SA-KNN مقارنةً بـ متساوية تقريبًا لكل موقع محدد مما يمكن أن يسهل توزيع المساعدات في العالم الحقيقي في مناطق الكوارث ، والوقت المناسب والتكلفة المعقولة.

TABLE OF CONTENT

	Page
BIOGRAPHY	**iii**
ACKNOWLEDGEMENTS	**iv**
ABSTRACT	**vi**
ملخص البحث	**vii**
TABLE OF CONTENT	**viii**
LIST OF TABLES	**xi**
LIST OF FIGURES	**xii**
CHAPTER 1	**1**
INTRODUCTION	**1**
1.1 Background	1
1.2 Motivations	1
1.3 Problem Statement	3
1.4 Research Question	3
1.5 Research Objective	4
1.6 Thesis Structure	4
CHAPTER 2	**6**
LITERATURE REVIEW	**6**
2.1 Introduction	6
2.2 Disaster Management	6

2.3 Disaster relief operation in Malaysia 13
2.3.1 Government Agencies Involved for Disaster Relief. 15
2.4 Phases of disaster response 18
2.5 Post-disaster 20
2.6 Fitness Function 21
2.7 Optimization of disaster relief operation 22
2.7.1 K-Nearest Neighbor (KNN) 22
2.7.2 Simulated Annealing (SA) 25
2.7.3 Genetic Algorithm (GA) 26
2.8 Summary 28

CHAPTER 3 29

METHODOLOGY 29

3.1 Introduction 29
3.2 Methodology Applied 29
3.3 Data Collection 31
3.2.1 Interview with volunteer Representative and Victim 31
3.3 Research Framework 33
3.4 Experiment Setting 35
3.4.1 Problem Formulation and Assumption 35
3.4.2 Mapping of disaster site 36
3.5 Simulated Annealing and Genetic Algorithm based KNN 39
3.6 Fitness function 43

CHAPTER 4 48

FINDING AND DISCUSSION 48

4.1 Introduction 48

4.2 Experimental Result and Analysis 48

4.3 Summary 59

CHAPTER 5 **60**

CONCLUSION AND FUTURE WORK **60**

5.1 Introduction 60

5.2 Achievement 60

5.3 Research Contribution 60

5.4 Limitation 61

5.5 Future Research 61

REFERENCES **62**

LIST OF TABLES

Page

Table 1: Information on Flood Disaster in Malaysia 2016/2017 13

Table 2: Incident management relationship 15

Table 3: Summary of the methodology applied 30

Table 4: List of Coordinate Involved in Experiment 37

Table 5: Summarize the parameter and value 52

Table 6: Result obtained from KNN algorithm 53

Table 7: Result obtained from the SA-KNN and GA-KNN algorithm 57

LIST OF FIGURES

	Page
Figure 1: Disaster Management Level/ Executive Committee	11
Figure 2: National Disaster Management Agency (Nadma)	16
Figure 3: Department of Social Welfare Malaysia (JKM)	17
Figure 4: National Security Council (NSC)	18
Figure 5: Phases of disaster response	19
Figure 6: The Algorithm of KNN	23
Figure 7: Brute Force Algorithm Process	24
Figure 8: The framework of experiment process SA-KNN	33
Figure 9: The framework of experiment process GA-KNN	34
Figure 10: The Structure of Map Gridding	37
Figure 11: The map with 200 population size of DP and 3 DC	39
Figure 12: Pseudocode of Simulated Annealing Algorithm	41
Figure 13: Pseudocode of Genetic Algorithm	43
Figure 14: Flowchart of SA and GA based KNN	45
Figure 15: The map after run KNN algorithm with K set as 60	49
Figure 16: SA-KNN in population of 200 DP	50
Figure 17: Potential DC is marked to be the selection of optimal DC	51
Figure 18: GA-KNN in population of 200 DC	51
Figure 19: The white circle shows the left-out DP that not covered	53
Figure 20: The Result of Fitness value for 200 DP in SA-KNN	54
Figure 21: The Result of Fitness value for 200 DP in GA-KNN	55

Figure 22: Result of SA Experiment with 200 DP 55

Figure 23: Result of GA Experiment with 200 DP 56

Figure 24: Fitness Value of each Algorithm SA and GA-KNN 58