POPULATION SEGMENTATION IN NEIGHBOURHOOD GEOSOCIAL NETWORKING

LOW SHENG LOONG

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS TESIS

JUDUL: POPULATION SEGMENTATION IN NEIGHBOURHOOD GEOSOCIAL NETWORKING

SESI PENGAJIAN: 2016/17

Saya LOW SHENG LOONG
(HURUF BESAR)

Mengaku membenarkan tesis (PSM/Sarjana/DoktorFalsafah) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hak milik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Silatandakan (/)

____	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
____	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
/	TIDAK TERHAD	

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

(TANDATANGAN PENULIS)

Alamat tetap: 287, Jalan Jambu Berasa Jinjang Selatan, 52000, WPKL
Tarikh : 21/8/17

(TANDATANGAN PENYELIA)

Mrs Syarulnaziah Anawar

Tarikh : 21/8/17

CATATAN: * Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana Muda (PSM)
** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa

POPULATION SEGMENTATION IN NEIGHBOURHOOD GEOSOCIAL NETWORKING

LOW SHENG LOONG

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

This report is submitted in partial fulfillment of the requirements for the Bachelor of Computer Science (Computer Networking)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
2016

DECLARATION

I hereby declare that this project report entitled

POPULATION SEGMENATION IN NEIGHBOURHOOD GEOSOCIAL NETWORKING

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT : _____ Date: 21/8/17

(Low Sheng Loong)

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

I hereby that I have read this project report and found this project report is sufficient in terms of the scope and quality for the award of

Bachelor of Computer Science (Computer Networking) with Honors

SUPERVISOR: _____ Date: 21/8/17

(Mrs Syarulnaziah Anawar)

DEDICATION

I want to dedicate my final year project report to my family, friends and for those who have helped me. My supervisor, Mrs Syarulnaziah Anawar, has been guiding and helping me to finish up this project all the long way. I would like to express the deepest gratitude to my beloved parents for their life-long love and support. They have been very encouraging in completion of my project and throughout the semesters of my studies. On top that, I also would like to dedicate this report to my friends and those who have helped me a lot and very supportive throughout the project development.

UTeM
اونيۏرسيتي تيکنيکل مليسيا ملاک
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

DEDICATION

I want to dedicate my final year project report to my family, friends and for those who have helped me. My supervisor, Mrs Syarulnaziah Anawar, has been guiding and helping me to finish up this project all the long way. I would like to express the deepest gratitude to my beloved parents for their life-long love and support. They have been very encouraging in completion of my project and throughout the semesters of my studies. On top that, I also would like to dedicate this report to my friends and those who have helped me a lot and very supportive throughout the project development.

UTeM
اونيۏرسيتي تيکنيکل مليسيا ملاک
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

ACKNOWLEDGEMENT

I would like to show my gratitude to the following respected people in helping me to complete this project. Firstly, I would like to show my gratitude to my project supervisor, Mrs Syarulnaziah Anawar and my project evaluator Dr Mohd Fairuz Iskandar who have given me constructive comments, guidance and support throughout this project. Next, I would like to thank my parents for being supportive in many aspects to finish this project successfully. I would like to thank my course mates for helping me to test and solve the technical and logical issues regarding my system.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

ABSTRACT

Quality of social community will degrade if each resident not social connected. Thus, this project try to tackle all this problem by use suitable algorithm to cluster and segment population of neighbour that implement in geosocial networking mobile application brings high quality of social community and safer housing environment. This algorithm will cluster and segment population of neighbourhood based on Dunbar's number stated by British anthropologist Robin Dunbar suggest cognitive limit of 150 number of people which an individual can maintain stable social relationships. System also will define temporarily and permanent residents by the algorithm. Residents will provide its own information and switch on Global Positioning System (GPS) on cellphone to let server to accurately track residents living place. GPS tracked data will be record in database. All data will be cluster based on modified DBSCAN algorithm and population segmentation based on Dunbar's number for high quality of social community. Thus, the geo-social networking mobile application will connect each resident in its own cluster area. Residents can use this geo-social networking mobile application as a communication platform to organize events, share news feeds, crime watching and swap goods. Therefore, this mobile application will build an neighbour concept, social community, easier platform and brings safer and greener living environment.

اونيۏرسيتي تيکنيکل مليسيا ملاک
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

ABSTRAK

Kualiti sosial kommuniti akan menurun jika setiap penduduk tidak mempunyai sosial penyambungan. Jadi, projek ini cuba untuk mengatasi masalah dengan menggunakan algoritma yang sesuai untuk kluster dan segmen populasi jiran yang melaksanakan dalam aplikasi mudah alih rangkaian geososial demi meningkatkan kualiti sosial kommuniti dan persekitaran perumahan yang selamat. Algoritma ini akan kluster dan segmen populasi kejiranan berasaskan nombor Dunbar dinyatakan oleh British antropologi Robin Dunbar yang cadangkan had kognitif jumlah 150 orang yang mana seseorang dapat mengekalkan hubungan sosial yang stabil. Sistem juga akan menentukan penduduk sementara dan tetap oleh algoritma. Penduduk akan memberikan maklumat tersendiri dan mengguna Global Positioning System (GPS) dalam telefon bimbit untuk membiarkan pelayan tepat mengesan tempat tinggal penduduk. Data yang dikesan oleh GPS akan dirakam dalam pangkalan data. Semua data akan menjadi kluster berdasarkan algoritma DBSCAN yang diubahsuai dan segmen penduduk berdasarkan kepada jumlah Dunbar untuk kualiti sosial yang tinggi. Oleh itu, aplikasi mudah alih rangkaian geo-sosial akan menghubungkan setiap penduduk di kawasan kelompoknya sendiri. Penduduk boleh menggunakan aplikasi mudah alih rangkaian geo-sosial ini sebagai platform komunikasi untuk menyusun acara, berkongsi berita berita, menonton jenayah dan barang swap. Oleh itu, aplikasi mudah alih ini akan membina konsep jiran, komuniti sosial, platform yang lebih mudah dan membawa persekitaran hidup yang lebih selamat dan mesra.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

TABLE OF CONTENTS

CHAPTER	SUBJECT		PAGE
	DECLARATION		**i ii**
	DEDICATION		**iii**
	ACKNOWLEDGEMENTS		**iv**
	ABSTRACT		**v**
	ABSTRAK		**vi**
	TABLE OF CONTENTS		**vii**
	LIST OF TABLES		**xi**
	LIST OF FIGURES		**xii**
CHAPTER 1	**INTRODUCTION**		
	1.1	Introduction	1
	1.2	Research Problem	2
	1.3	Research Question	2
	1.4	Objective	3
	1.5	Scope	3
	1.6	Project Significance	3
	1.7	Conclusion	4
CHAPTER 2	**LITERATURE REVIEW**		
	2.1	Introduction	5
	2.2	Project Introduction	5
	2.3	Taxonomy of Geosocial Networking	6
		2.3.1 Geosocial Networking	6
		2.3.2 Geolocation	7
		2.3.3 Clustering	7

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

	2.3.4	Segmentation	8
	2.3.5	Density-Based	9
	2.3.6	Dunbar's Number	10
	2.3.7	Engagement	11
2.4	Mapping For Geosocial Networking		13
2.5	Proposed Method		13
2.6	Conclusion		14

CHAPTER 3 METHODOLOGY

3.1	Introduction		15
3.2	Methodology		16
	3.2.1	Formulate a question	17
	3.2.2	Design an experiment	17
	3.2.3	Data Gathering	17
	3.2.4	Analyze Data	18
	3.2.5	Publish	18
3.3	Project Milestone		19
3.4	Conclusion		21

CHAPTER 4 AN IMPROVED DBSCAN ALGORITHMS USE DUNBAR NUMBER FOR POPULATION SEGMENTATION

4.1	Introduction		22
4.2	Basic Concept of DBSCAN		23
4.3	Experiment Setup		27
4.4	Design of Improved DBSCAN Algorithm		27
	4.4.1	Improved Accuracy Distance Calculation Between Residents Using Haversine Formula	27
	4.4.2	Defining Threshold Value to Self-Organize Segment Density	28

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

	4.4.3	Improving Social Connectivity By Cluster Re-segmentation Using Dunbar Number	29
4.5	Result		32
4.6	Discussion		33
4.7	Conclusion		34

CHAPTER 5 A NEW METHOD TO DETERMINE RESIDENCY STATUS USING THRESHOLD FREQUENCY

5.1	Introduction		35
5.2	Location Aware		36
	5.2.1	Motivation of Work	36
	5.2.2	Geolocation	36
5.3	Proposed Method		38
	5.3.1	Resident Engagement Incentives	38
	5.3.2	Identify Threshold Frequency	40
5.4	Result		40
5.5	Discussion		41
5.6	Conclusion		42

CHAPTER 6 mySOS: A PROOF OF CONCEPT IMPLEMENTATION

6.1	Introduction		43
6.2	System Architecture		44
6.3	Experiment Setup		46
	6.3.1	Software Specification	46
	6.3.2	Hardware Specification	46
6.4	Application Design and Functionality		47
6.5	Implementation		48
	6.5.1	Emergency Post	49
	6.5.2	Neighbour SOS	54
6.6	Conclusion		64

UNIVERSITI TEKNIKAL MALAYSIA MELAKA
UTeM
اونيۏرسيتي تيکنيکل مليسيا ملاک
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

CHAPTER 7 **CONCLUSION**

7.1	Introduction	65
7.2	Project Summarization	65
7.3	Project Contribution	66
7.4	Project Limitation	67
7.5	Future Works	67
7.6	Conclusion	67
REFERENCES		68

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF TABLES

TABLE	TITLE	PAGE
2.1	Mapping for geosocial networking	13
3.1	PSM 1 Gantt Chart	19
3.2	PSM 2 Gantt Chart	20

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LIST OF FIGURES

DIAGRAM	TITLE	PAGE
2.1	Taxonomy of Geosocial Networking	6
3.1	Methodology of Population Segmentation of Neighbourhood Geosocial Networking	16
4.1	Terms and concept of DBSCAN	24
4.2	DBSCAN algorithm pseudocode dbscan function	25
4.3	DBSCAN algorithm pseudocode ExpandCluster function	26
4.4	Haversine formula for improved DBSCAN (Cite: Virtues of the Haversine. Sky and Telescope)	28
4.5	Determination of cluster group size	30
4.6	Form a rectangle border for arbitrary shape cluster	31
4.7	Inverse Haversine formula	31
4.8	Improved DBSCAN algorithm for low density	32
4.9	DBSCAN algorithm for high density	32
4.10	Improved DBSCAN algorithm for high density	33
5.1	Dataset in database	40
5.2	Result of resident engagement incentive	41
6.1	System architecture and user registration and login process	44
6.2	Application registration interface	45
6.3	Emergency Post	47
6.4	Neighbour SOS button	48
6.5	Temporarily user posting function (Adapted from Dobble Social Network Template)	49

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

6.6	Permanent user posting function (Adapted from Dobble Social Network Template)	50
6.7	Permanent user submit emergency post (Adapted from Dobble Social Network Template)	51
6.8	Emergency post message content (Adapted from Dobble Social Network Template)	52
6.9	Receive notification from emergency post	53
6.10	Neighbour SOS interface button	54
6.11	Neighbour SOS interface	55
6.12	Alert message for temporarily user	56
6.13	SMS message receive by emergency contact person	57
6.14	Application automatic call emergency contact person	58
6.15	Alert message received by permanent user	59
6.16	Neighbour SOS post message	60
6.17	Neighbours received notification	61
6.18	SMS message received by neighbour and emergency contact person	62
6.19	Calling neighbour contact	63

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

CHAPTER 1

INTRODUCTION

UNIVERSITI TEKNIKAL MALAYSIA MELAKA
UTeM

1.1 INTRODUCTION

Nowadays, quality of social community is much important in a local area. Every neighbourhood community should be more communicate and social between each other. Thus, social connection between each resident plays vital part in quality of social community. Quality of social community will degrade if each resident not social connected.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

1.2 RESEARCH PROBLEM

All community, welfare, education activities can organize within an area. Unfortunately, the areas maybe too large to socially bind each neighbour within a community and the community will be so large if based on area. Thus, this had decrease the quality of social community. Next, residents within community need to use traditional ways to determine whether he or she stay within this area. Grouping and determination resident still use traditional ways to identify whether resident staying within the area.

1.3 RESEARCH QUESTION

1. How to identify the density of residential areas?
   - Use improve DBSCAN cluster and segmentation the residential areas with 150 people per cluster.
2. How to determine resident whether he or she stay within the area?
   - State user as temporarily and permanent resident with cumulative point to determine.
3. How to implement this modified algorithm to geosocial neighbourhood?
   - Modified algorithm may implement into social media application with easy interface that ease for resident interaction between each neighbourhood.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

1.4 OBJECTIVE

1. To design an improved DBSCAN algorithms by identify low and high density residential area.
2. To determine resident status temporarily or permanent.
3. To Implement modified algorithm in geosocial neighbourhood.

1.5 SCOPE

In this project, the place of Taman Bukit Melaka, Melaka is chosen as a designated location. This algorithm design to implement in area for grouping neighbourhood into a network social community.

1.6 PROJECT SIGNIFICANCE

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

This algorithm efficiently grouping neighbourhood into a community with Dunbar's number to maintain stable social relationship between each neighbourhood and community. This algorithm will be implement to social network application for community ease to use. Thus, community, welfare, education activities will be increase in the area that enhance racial unity and integration.

1.7 CONCLUSION

This chapter discuss a modified DBSCAN algorithm that suitable use to group neighbourhood community that conjuction with Rukun Tetangga programme. This algorithm will cluster neighbour into a community and segment neighbourhood community based on Dunbar number in that area. This algorithm also will identify temporarily and permanent resident base on cumulative point system. This algorithm will be written in Javascript programming language for better implementation in web and mobile application.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter will discuss the literature review of the project. Literature review is a related study that relies to this project. In this chapter, the articles, books and websites will summaries about important information of the research sources of the project. It also provides a hardy guide and a background of the certain topics. This chapter is the most important step which helps justify the choice of the research or project, establishes the importance of the particular topic, provides background information and discovers how the project is related to the works of others.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2.2 PROJECT INTRODUCTION

Before continue with the project, we need to know what population segmentation and neighbourhood geosocial networking. Data object will represent as user geographical coordinates and cluster into a group. Population segmentation will be applied into the cluster group. Population segmentation is a process of dividing cluster into segments based on variety of categories. Geosocial network is one of the branch of social networking that allow user share information and communicate with other people regarding time and geographical location.

2.3 TAXONOMY OF GEOSOCIAL NETWORKING

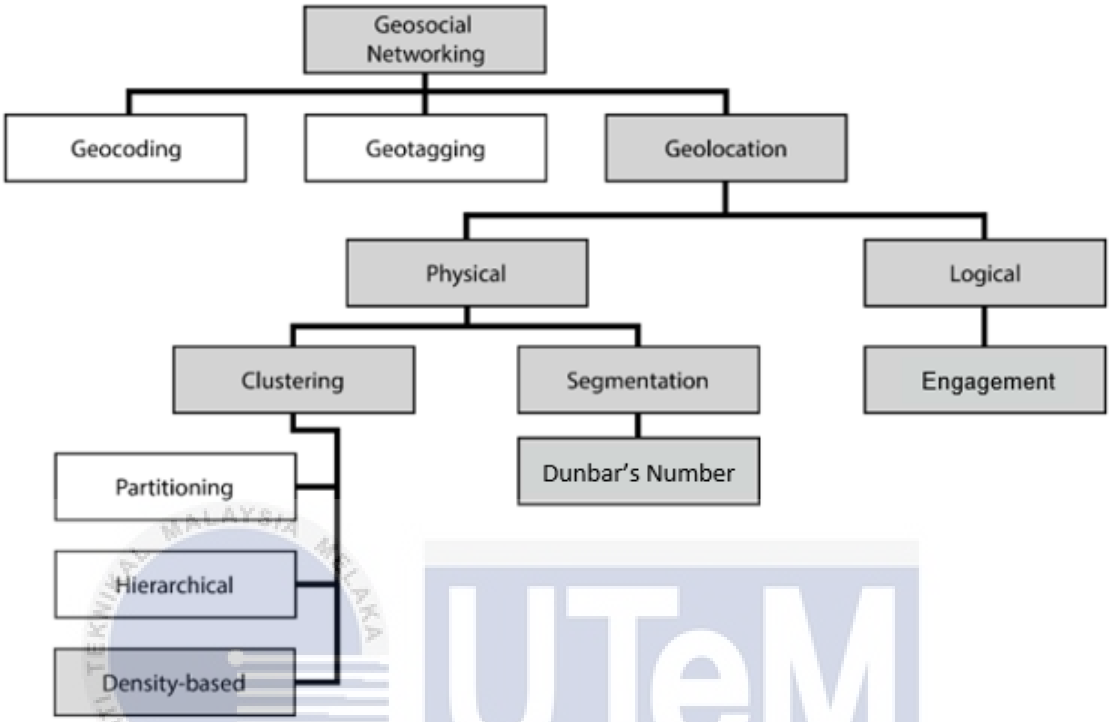

Figure 2.1: Taxonomy of Geosocial Networking

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2.3.1 Geosocial Networking

Geosocial network is one of the branch of social networking that allow user share information and communicate with other people regarding time and geographical location. It became a new research field that start develop for quite recent years. In 2010, a geosocial event detection had created by using collected data from Twitter (Lee and Sumiya 2010). In the same year, a new visualization tool StickViz using k-Neighbors Searches in geosocial networking services developed. (K.-S. Kim et al., 2010). In 2011, Noulas et al. study and understand human mobility pattern in Foursquare geosocial network. Nowadays, famous geosocial networking applications such as Waze, Foursquare and Facebook Places

have millions of active users. In this paper, i try to create a neighbourhood community that will implement into geosocial networking application.

2.3.2 Geolocation

Geolocation is a technique to estimate or identify a person or place geographical location with a set of geographic coordinates. With the assist of GPS system in mobile phone, geolocation will be more accurate within 10 meters or better. Geolocation application such as Foursquare, Brightkite and others encourage user provide details recent visited places, hometown or neighbourhood using GPS system. Geolocation application will put gaming element to let user having interaction between each other like Foursquare, user will earn point if check-in anyplace and badges for progress. (J.Choi et al., 2012). Study shows that people will tends to produce geographic information on geo-games. (S. Matyas et al., 2011)

2.3.3 Clustering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Clustering algorithm have three basic categories that are hierarchical, partitioning and density-based algorithm. (K. Mumtaz & K. Duraiswamy 2010). Large number of data in huge databases can be deals by all these algorithms. Partitioning algorithm construct k clusters given n data object. Each cluster cannot have same common data object and can have many group as many object $k \leq n$. Hierarchical algorithm create hierarchical decomposition presented in a dendrogram, tree that splits set of data object into smaller subset until one data object represents a subset. Hierarchical algorithm can be classified into two types that is divisive and agglomerative. Density-based algorithm design to discover arbitrary shape of cluster have higher density than remainder data object. Low-density region of data object considers outlier or noise.

2.3.4 Segmentation

Segmentation is the process of dividing into segments based on variety of categories. In process of diving and segmenting, those who shared common characteristics such as common interest, common needs, similar lifestyles or even similar demographics profiles will divide into segment.

There are several types of segmentation technique applied in some major sector such as market segmentation, biology segmentation, computing segmentation, demographic segmentation and others.

Demographic segmentation divides a population based on many variables such as age, gender, family size, income, occupation, race and other. There is research has noted that segmentation relying solely on demographic variables may have limited worth when identifying segments that are homogeneous with respect to health-related outcomes (Boslaugh et al., 2005), and would not uncover the different motivations each segment may have for health-related behavior change. Psychographic segmentation helps to reveal these patterns or differences between groups of people who may be similar in age, gender or socioeconomic status. This addition of psychographic variables to behavioural insight adds information about 'why' to the 'who' and 'where' provided by socio-demographic and geodemographic information (French et al., 2011). Psychographic segmentation has been employed in relation to physical activity and dietary choices (Byrd-Bredbenner et al., 2008) using additional information gathered from insight research. The Department of Health Change4Life campaign used segmentation to divide families into six clusters depending on their attitude towards diet and physical activity, levels of obesity and food consumption and demographic make-up to enable interventions to promote healthy weight in families to be more effectively targeted and delivered.

This paper considers the social relationship segmentation of the population who are neighbourhood they within a same cluster group based Dunbar's number.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA