



Faculty of Computer Science and Information Technology

Bipartite-Network-Based Modeling of Habitat Suitability

Liew Chin Ying

**Doctor of Philosophy
(Computational Science)
2016**

Bipartite-Network-Based Modeling of Habitat Suitability

Liew Chin Ying

A thesis submitted

In fulfilment of the requirements for the degree of
Doctor of Philosophy (Computational Science)

Faculty of Computer Science and Information Technology

UNIVERSITI MALAYSIA SARAWAK

2016

ACKNOWLEDGEMENTS

First and foremost I want to thank my Heavenly Father whose abundant grace has strengthened me throughout the process of my PhD pursuit. It is HIM that has enabled me, one step at a time, to come to the completion of my doctorate research.

My deepest gratitude goes to my main supervisor, Assoc. Prof. Dr. Jane Labadin. It has been an honour to be her PhD student. I am indebted to her generous, supportive and excellent supervision and guidance. She is there for me every time I need advices. Our countless sessions of discussion have always been enlightening, motivating, assuring, and fruitful in provoking new thoughts and evaluations to existing ideas. I am thankful for the invaluable experience and training I gained from her through conducting the doctorate research, writing scientific papers, preparing and doing presentations, and attending funded conferences. Thanks for accepting me as who I am and helping me to accomplish my PhD endeavour. Likewise, I would like to express my appreciation to my co-supervisors, Prof. Dr. Wang Yin Chai and Prof. Dr. Andrew Alek Tuen, for their expertise in providing resourceful discussion and valuable suggestions with regards to the respective aspects in my doctorate research that are related to their expertise. Thank you for being patient and kind.

I also wish to acknowledge the members of Sarawak Dolphin Project team – Cindy, Ana and Jenny who are under the leading of Prof. Dr. Andrew Alek Tuen as the project leader – from the Institute of Biodiversity and Environmental Conservation (IBEC) of UNIMAS. My heartiest thanks to Cindy, particularly, for generously sharing her research outcome and her experience and knowledge with Irrawaddy dolphin in Kuching Bay. Thank you for giving your informative explanation diligently.

I am especially grateful for the Computational Science and Mathematics (CSM) post-graduates group, especially those of us who are under the supervision of Assoc. Prof. Dr. Jane Labadin, for the fun time we have throughout the high and low of conducting our respective postgraduate researches. Cynthia, SiewPing, ShianLi and Tay, thank you for your help and friendship and the time we share when attending conferences. Dr. Shakeel Ahmad Kamboh and Dr. Fatai Anifowose, thank you for giving your professional advices promptly despite of your tight research schedules. Dr. Monday Okpoto Eze, thank you for your doctorate work that has inspired the motivation in considering the use of network modeling approach in my PhD research.

My heartfelt thanks go to the Faculty of Computer Science and Information Technology (FCSIT) of UNIMAS for providing a conducive environment, the technical support, and facilities for me in pursuing my PhD degree. I wish to thank mainly the lecturers of CSM for giving constructive suggestions and inputs for my presentations in the department research hour and meeting me in person whenever I need help. I wish to thank the Dean and Deputy Dean for your support, and lecturers for your insightful questions, comments and suggestions for my presentations in the faculty seminar-series. Special thanks to the faculty for entrusting me as the president of the FCSIT Postgraduate Society (PGSoc) for the term of October 2012 to September 2013 of which I am given the privilege to gain valuable experiences I never have before. Thanks to UNIMAS for providing the desirable researching facilities, the resourceful Centre of Academic Information Services and the UNIMAS Graduate School which are well equipped with helpful and friendly staffs.

I also wish to appreciate my colleagues in the Department of Mathematical Sciences of Universiti Teknologi MARA (UiTM) Sarawak. Many of you have taken me in as one of your research project members even when I am a new staff. You have unconsciously cultivated

the passion for research and planted the seed in me to further for a PhD study. I gratefully acknowledge the funding resources and study leave that have made my PhD work possible. I am funded by both the Ministry of Education Malaysia and UiTM for my three years of PhD study through the SLAI-UiTM scholarship with a three-year study leave. I sincerely thank my university, UiTM, and my faculty, Faculty of Computer and Mathematical Sciences, for supporting and approving my PhD study.

Last but not least, I am deeply grateful to my dearest husband for being sacrificial, understanding and supportive faithfully especially during the tough times of my PhD pursuit. I would like to thank my parents for raising me with a love of knowledge and lovingly encouraged me to strive for my PhD dream. I thank my lovely children for constantly being an inspiration to me throughout these years of my postgraduate study. Not forgetting also my dear brothers who have persistently believed in me to complete my PhD study. My family has indeed contributed significantly to the accomplishment of my doctorate degree. Thank you.

DEDICATION

To my Heavenly Abba Father for YOUR great grace and everlasting faithfulness

O LORD, You have searched me and known me.

You know my sitting down and my rising up;

You understand my thought afar off.

You comprehend my path and my lying down,

And are acquainted with all my ways.

For there is not a word on my tongue, But behold, O LORD, You know it altogether.

You have hedged me behind and before, And laid Your hand upon me.

Such knowledge is too wonderful for me;

It is high, I cannot attain it.

(Psalms 139: 1-7a)

To my husband – Soon Phin

To my children – Yee Enn, Yee Yang and Zen Yang

To my parents – Liew Yiu Sheng and Chung Choi Jin

To my brothers – Shuang Chen and Lit Cheng

In memory of my late grand-father – Liew Dao Si

In memory of my late grand-mother – Choo Kee Jiu

ABSTRACT

The approaches used in typical habitat suitability studies are pre-dominantly statistical where huge amount of data is required whereas population dynamics studies that incorporate the aspect of habitat suitability in their deterministic approaches are commonly based strongly on physical laws that entail detailed species-specific demographic values. With the intention to model the habitat suitability of a newly-surveyed species, the Irrawaddy dolphin (ID) in Kuching Bay, this research is characterized by real-world data that is scarce, incomplete and imbalanced, lacking of physical laws to base the modeling, and without demographic values. Solving such problem using the above approaches is extremely challenging. The emergent graph-theoretic bipartite network modeling (BNM) approach is not restricted by the limitations posed by the above approaches and has thus motivated this research in employing it to solve the problems faced. A process-oriented methodology framework adapted and termed as the bipartite-network-based methodology framework (BNB-MF) is devised and followed. The result produced is a bipartite habitat suitability network (BiHSN) model that represents the habitat suitability of the species in terms of the ranking of location nodes of the weighted bipartite network (BiHSN) constructed. This BiHSN consists of thirteen location nodes and thirteen dolphin nodes, each with their respective parameters of which some are quantified through support vector machines (SVM) algorithm, and thirty-eight weighted edges which are quantified through the multiplication rule. The ranking is determined based on the habitat suitability index (HSI) computed for the species for every location node from the implementation of the adapted HITS search algorithm onto the BiHSN. Nodes with higher HSI are ranked higher and thus implies better suitability as the more preferred habitat. Verification analysis conducted reports a root mean square error (RMSE) of 0.0002459 and 0.0007703 for

the location and dolphin nodes respectively, which is much lower than the threshold value of 0.05. Validation analysis performed reports a Spearman's Rank Correlation Coefficient (SRCC) of above 0.900, which is greater than the threshold value set at 0.70. Seven extended analyses are also executed and the outcomes obtained have supported the usefulness of BiHSN Model, and its relevancy to the research scenario. Consequently, this research has shown that the BNM approach following the BNB-MF has been successfully employed, and it is believed to be applicable in solving many other real-world problems.

Pemodelan Berasaskan Rangkaian Bipartit untuk Kesesuaian Habitat

ABSTRAK

Pendekatan utama yang digunakan dalam survei kesesuaian habitat adalah kaedah statistik yang memerlukan saiz data yang besar. Survei kedinamikan populasi yang mengambilkira aspek kesesuaian habitat pula menggunakan pendekatan deterministik yang berlandaskan hukum sains fizikal yang kukuh di samping memerlukan nilai-nilai demografik terperinci sesuatu spesis. Kajiselidik ini bertujuan untuk membangunkan model kesesuaian habitat sejenis spesis yang baharu dikaji iaitu lumba-lumba Irrawaddy (ID) di Teluk Kuching. Kajiselidik ini bercirikan jumlah data kajian dunia sebenar yang jarang, tidak lengkap serta tidak seimbang, kekurangan hukum-hukum sains fizikal untuk digunakan sebagai asas pembinaan model, dan tanpa nilai-nilai demografik. Justeru itu, aplikasi pendekatan-pendekatan di atas dalam menyelesaikan masalah kajian kajiselidik ini amat mencabarkan. Di samping itu, pendekatan pemodelan rangkaian bipartit (BNM) yang berasaskan teori graf tidak terhad oleh batasan yang wujud dalam pendekatan-pendekatan di atas dan sekaligus telah mendorong kajiselidik ini untuk menggunakannya bagi menyelesaikan masalah kajian seperti tersebut di atas. Satu rangkakerja metodologi berasaskan rangkaian bipartit (BNB-MF) dan berorientasikan proses telah dirangka dan digunapakai dalam kajiselidik ini. Hasilnya merupakan sebuah model rangkaian bipartit kesesuaian habitat (BiHSN) yang mewakili kesesuaian habitat spesis ini dalam bentuk susunan kedudukan nod lokasi dalam rangkaian bipartit (BiHSN) berpemberat yang dibentukkan. BiHSN ini merangkumi tiga belas nod lokasi dan tiga belas nod lumba-lumba, masing-masing dengan parameter-parameter mereka yang mana beberapa daripadanya telah dikuantifikasikan melalui pelaksanaan algoritma support vector machines (SVM). BiHSN ini juga mempunyai tiga puluh lapan

penghubung nod, bersama pemberat masing-masing, yang dikuantifikasikan melalui hukum pendaraban. Susunan kedudukan tersebut ditentukan berdasarkan indeks kesesuaian habitat (HSI) yang dinilai melalui pelaksanaan algoritma carian yang telah diubahsuai ke atas BiHSN. Nod dengan HSI yang lebih tinggi akan disusun pada kedudukan yang lebih tinggi. Nod lokasi dengan susunan kedudukan yang lebih tinggi menandakan tahap kesesuaian yang lebih tinggi sebagai habitat yang lebih diutamakan. Analisis verifikasi kajiselidik ini telah melaporkan Ralat Min Punca Kuasa Dua (RMSE) bernilai 0.0002459 dan 0.0007703 bagi nod lokasi dan nod lumba-lumba masing-masing. Nilai RMSE ini adalah jauh lebih rendah daripada nilai ambang 0.05. Analisis validasi kajiselidik ini melaporkan Pekali Korelasi Spearman (SRCC) yang melebihi 0.900. Nilai SRCC tersebut adalah lebih tinggi daripada nilai ambang yang ditentukan pada 0.70. Tujuh analisis lanjutan telah dilaksanakan dalam kajiselidik ini. Hasil analisis-analisis tersebut secara keseluruhannya menyokong kebolegunaan dan praktikaliti Model BiHSN dan relevansinya terhadap senario kajiselidik ini. Sebagai rumusannya, kajiselidik ini telah menunjukkan bahawa pelaksanaan pendekatan BNM mengikut BNB-MF telah berjaya menyelesaikan masalah kajiselidik ini. Adalah dipercayai bahawa pendekatan BNM dapat diaplikasikan dan berupaya menyelesaikan banyak masalah dunia sebenar yang lain.

TABLE OF CONTENTS

Acknowledgements	i
Dedication	iv
Abstract	v
Abstrak	vii
Table of Contents	ix
List of Tables	xiii
List of Figures	xv
List of Equations	xvii
List of Abbreviations	xix
1 INTRODUCTION	1
1.1 Introduction of Study	1
1.2 Background of Study	2
1.3 Definitions	8
1.4 The Motivation and Research Objectives	8
1.5 Scope of Study	9
1.6 Research Methodology	10
1.6.1 Stage 1: Problem Characterization	13
1.6.2 Stage 2: Model Construction	13
1.6.3 Stage 3: Model Analysis and Evaluation	14
1.6.4 Application of BNB-MF	15
1.7 Contributions	16
1.8 Thesis Outline	17
2 LITERATURE REVIEW	19
2.0 Introduction	19
2.1 Habitat Suitability Studies	19
2.1.1 Main Components of Habitat Suitability Studies	21
2.1.2 Common Ecological Theory and Assumptions in Habitat Suitability Studies	25
2.1.3 Approaches of Habitat Suitability Studies	27

2.1.3.1	Statistical Approaches	27
2.1.3.2	Deterministic Approaches	31
2.1.3.3	Other Emergent Scientific Methods	34
2.2	Network Modeling Approaches	37
2.2.1	From Graph Theory to Network Theory	38
2.2.2	Network Modeling Approaches and Ecology Studies	39
2.3	Summary	46
3	BIPARTITE GRAPH FORMALIZATION	47
3.0	Introduction	47
3.1	Current Research Scenario	47
3.1.1	Research Data – Raw Data	48
3.1.2	Research Data – Expert Opinions	53
3.1.3	Quantity of Interest with Variables Concerned	54
3.1.4	Quantity of Interest with Assumptions Concerned	55
3.2	Graph Structure Representation	56
3.2.1	Habitat Suitability Triangle	56
3.2.2	Basic Building Block of the Bipartite Habitat Suitability Network Model	58
3.3	Bipartite Graph Formulation	61
3.3.1	Identification of the BiHSN Nodes	61
3.3.1.1	Location Nodes	61
3.3.1.2	Dolphin Nodes	62
3.3.2	Identification of BiHSN Links between Nodes	63
3.3.3	Complete Bipartite Habitat Suitability Graph for BiHSN	65
3.4	Summary	69
4	BIPARTITE NETWORK FORMULATION	70
4.0	Introduction	70
4.1	Data Pre-Processing	70
4.1.1	Handling of Incomplete Data	71
4.1.2	Handling of Imbalanced Data	74
4.2	Quantification of Node Parameters	77
4.2.1	Quantification of Location Node Parameters	77
4.2.1.1	Quantification of Water Suitability Parameter, W	78
A.	Quantification of W using Support Vector Machines	82
A.1	Implementation of SVM Learning Algorithms in Modeling SVM Water Model	84
A.2	Classification Performance of the Classifier	86
A.3	Attributes Selection Analysis of SVM Water Model	89
A.4	SVM Water Model and Generation of W	91
B.	Quantification of W using Decision Trees	94
C.	Comparison between SVM Water Model and DT Water Model	99

4.2.1.2	Quantification of Fisheries (Food) Availability Parameter, F	100
A.	Quantification of F using Support Vector Machines	103
A.1	Implementation of SVM Learning Algorithm in Modeling SVM Fisheries (Food) Model	104
A.2	Attributes Selection Analysis of SVM Fisheries (Food) Model	105
A.3	SVM Fisheries (Food) Model and Generation of F	106
B.	Quantification of F using Decision Trees (DT)	108
C.	Conclusion on Quantification Technique for W and F	110
4.2.1.3	Quantification of Number of Times a Location is Visited by ID, Fl	111
4.2.1.4	Parameters of Location Node in BiHSN	114
4.2.2	Quantification of Dolphin Node Parameters	116
4.2.2.1	Quantification of Number of Times a Dolphin Visited a Location, Fd	117
4.2.2.2	Quantification of Maximum ID Group Size	118
4.2.3	Values of Node Parameters	120
4.3	Quantification of Link Weight	120
4.4	Summary	125
5	WEB-BASED SEARCH ALGORITHM IMPLEMENTATION, VERIFICATION AND VALIDATION	128
5.0	Introduction	128
5.1	Web-based Search Algorithm	128
5.2	Implementation of HITS Search Algorithm	130
5.2.1	Generation of Hub and Authority Matrix	131
5.2.2	Generation of Hub and Authority Principal Eigenvector	132
5.2.3	Assignment of Nodes' Label and Habitat Suitability Index (HSI)	134
5.2.4	Ranking of Bipartite Nodes in BiHSN	136
5.3	Model Verification	137
5.3.1	Benchmark Verification	138
5.3.2	Analytical Verification	144
5.4	Model Validation	146
5.4.1	Validation through Real-world Data	147
5.4.2	Validation through Past Survey Result	153
5.5	Summary	159
6	EXTENDED EVALUATIONS AND ANALYSES	161
6.0	Introduction	161
6.1	Crowding Analysis	162
6.2	Node Similarity Analysis	165
6.2.1	Similarity Index	166
6.2.2	Pyramidal Similarity Plot	169
6.3	Correlation Analysis	173

6.3.1	Correlation Analysis for Location Node	173
6.3.2	Correlation Analysis for Dolphin Node	175
6.4	Data Size Uncertainty Analysis	178
6.4.1	Data Size Uncertainty Analysis for SVM Water Model	179
6.4.2	Data Size Uncertainty Analysis for SVM Fisheries (Food) Model	182
6.5	Dolphin Node Uncertainty Analysis	185
6.6	Parameter Significance Analysis	188
6.7	Potential Predictive Ability Analysis	194
6.8	Summary	199
7	SUMMARY AND CONCLUSIONS	201
7.0	Introduction	201
7.1	Research Summary	202
7.2	Discussion of Research Result	204
7.3	Main Contributions of Research	206
7.3.1	The General Modeling Research Communities	207
7.3.2	The Body of Knowledge of the Related Research Fields	208
7.3.3	The Specific End-users and Research Communities	211
7.4	Limitation of Research	212
7.5	Recommendations for Future Research	213
7.6	Conclusion	221
	REFERENCES	222
	LIST OF PUBLICATIONS AND PRESENTATIONS	242
	APPENDIX A	244
	APPENDIX B	256

LIST OF TABLES

2.1	Summary of Related Studies on Habitat Suitability Studies	22
2.2	Summary of Related Studies on Ecology Studies Applying Network Modeling Approaches	40
2.3	Summary of Related Studies on Ecology Studies Applying Bipartite Network Modeling Approach	44
3.1	Four Sub-datasets of Research Data	50
3.2	Potential Variables for This Research	54
3.3	Movement of Each Unique Individual ID	65
3.4	Record of Location Node Visited by ID	66
4.1	Results of the Nine SVM ID-Sighting Models	90
4.2	SVM ID-Sighting Models that are Ranked Top Five	91
4.3	Characteristics and Classification Performance of the SVM Water Sub-models	92
4.4	W of 38 Data Points in BiHSN Dataset (SVM Water Model)	94
4.5	Characteristics and Classification Performance of the DT Water Sub-models .	97
4.6	W of 38 Data Points in BiHSN Dataset (DT Water Model)	99
4.7	Comparison between SVM Water Model and DT Water Model	100
4.8	Results of the Four SVM Fisheries-activities Models	106
4.9	Characteristics and Classification Performance of the SVM Fisheries (Food) Sub-models	107
4.10	F of 38 Data Points in BiHSN Dataset (SVM Fisheries (Food) Model)	109
4.11	Characteristics and Classification Performance of the DT Fisheries (Food) Sub-models	110
4.12	Comparison between SVM Fisheries (Food) Model and Possible DT Fisheries (Food) Model	111
4.13	Link Matrix of BiHSN	112
4.14	Parameter Values of W , F , dl , drm , Ti , and pH for 38 Data Points in BiHSN .	115
4.15	Parameter Values of $Fdj : i$	118
4.16	Parameter Values of $Nj : i$	119
4.17	RMSE Values of the Nodes for Four Link Weight Quantification Techniques .	122
4.18	Habitat Suitability Strength (HSS) Matrix of BiHSN	125
5.1	Hub Matrix	132
5.2	Authority Matrix	133
5.3	Principal Eigenvector of the (a) Location Node; and (b) Dolphin Node	135
5.4	The Ranking of (a) Location Nodes; and (b) Dolphin Nodes	136
5.5	RMSE Analysis for Location Nodes	141
5.6	RMSE Analysis for Dolphin Nodes	144

5.7	Habitat suitability Strength (HSS) Matrix of LDF Dataset	151
5.8	The Ranking of (a) Location Nodes; and (b) Dolphin Nodes, for LDF Dataset	152
5.9	Calculation of SRCC for Location Nodes (Real-world Data Validation)	152
5.10	Calculation of SRCC for Dolphin Nodes (Real-world Data Validation)	153
5.11	Calculation of SRCC for Location Nodes (Past Survey Result Validation) . .	158
6.1	Similarity Index Matrix for Location Nodes of the BiHSN Model	168
6.2	Calculation of $LDegR$ and $LGSizeR$ for Location Node	175
6.3	Calculation of $DDegR$ and $DGSizeR$ for Dolphin Node	177
6.4	Summary of Characteristics of the New SVM Water Model	180
6.5	Classification Performance of the New SVM Water Model	181
6.6	SRCC for Comparison of BiHSN Nodes' Ranking (SVM Water Model)	182
6.7	Summary of Characteristics of the New SVM Fisheries (Food) Model	183
6.8	Classification Performance of the New SVM Fisheries (Food) Model	183
6.9	SRCC for the Comparison of BiHSN Nodes' Ranking (SVM Fisheries (Food) Model)	184
6.10	SRCC for (a) ρ_{xDj} and (b) ρ_{xD1toj}	187
6.11	SRCC for 'Leave-one-out' Parameter Analysis	190
6.12	SRCC for 'Leave-multiple-out' Parameter Analysis	190
6.13	SRCC for Extended Parameter Significance Analysis	192
6.14	Ranking of (a) L1–L29; (b) L1–L13 & L30–L40; and (c) L1–L40	198
6.15	SRCC for Potential Predictive Ability Analysis (L1–L13 & L30–L40)	198
A.1	MATLAB Commands for the Interpolation of Incomplete Data	244
A.2	LIBSVM Library Tools and Programs (Chang & Lin, 2011) used	245
A.3	W Generated from 14 SVM Water Sub-models for the BiHSN Dataset	246
A.4	W Generated from Four DT Water Sub-models for the BiHSN Dataset	248
A.5	F Generated from Five SVM Fisheries (Food) Sub-models	249
A.6	Partial Result of Iterations for Hub Principal Eigenvector Generation	250
A.7	Partial Result of Iterations for Authority Principal Eigenvector Generation . .	250
A.8	Calculation of SRCC for Location Nodes (Analytical Verification)	251
A.9	Calculation of SRCC for Dolphin Nodes (Analytical Verification)	251
A.10	Parameter Values of Location Nodes for LDF Dataset (Validation)	252
A.11	Parameter Values of Dolphin Nodes for LDF Dataset (Validation)	252
A.12	Calculation of SRCC for Location Nodes Ranking based on Node Degree and Maximum ID Group Size (Crowding Analysis)	253
A.13	Calculation of SRCC for ρ_{HSL_LDegR} (Location Node Degree Ratio – Correla- tion Analysis)	253
A.14	Calculation of SRCC for $\rho_{HSL_LGSizeR}$ (Location Node Maximum ID Group Size Ratio – Correlation Analysis)	254
A.15	Calculation of SRCC for ρ_{R_DDegR} (Dolphin Node Degree Ratio – Correlation Analysis)	254
A.16	Calculation of SRCC for $\rho_{R_DGSizeR}$ (Dolphin Node Maximum ID Group Size Ratio – Correlation Analysis)	255

LIST OF FIGURES

1.1	Content Coverage of Chapter 1	2
1.2	Variables Characterizing Preferred Habitat of Irrawaddy Dolphin	6
1.3	Bipartite-Network-Based Methodology Framework (BNB-MF)	12
1.4	Application of BNB-MF	16
2.1	Content Coverage of Chapter 2	20
2.2	Three Main Components of Habitat Suitability Studies Structure	25
2.3	Example of Habitat Suitability Map	28
2.4	Summary of Research using PDE in Spatial Interaction and Population Dynamic	33
3.1	Content Coverage of Chapter 3	48
3.2	Scope of Kuching Bay	49
3.3	Re-sight Maps of Twelve Identified Individual ID based on RDF	51
3.4	Habitat Suitability Triangle (HST)	57
3.5	Modified Habitat Suitability Triangle	57
3.6	The Basic Building Block of BiHSN	59
3.7	An Example of BiHSN, H	60
3.8	Actual Location of the BiHSN Location Nodes	62
3.9	Overlaid Location Nodes and Dolphin Nodes onto Figure 3.3	64
3.10	Complete Bipartite Habitat Suitability Graph for ID in Kuching Bay	67
4.1	Content Coverage of Chapter 4	71
4.2	Interpolation of Data for Incomplete depth (de) data	73
4.3	Handling of Imbalanced Data through Under-sampling Method	76
4.4	Parameters of BiHSN Location Node at the (a) Initial Stage; and (b) Final Stage of Process 2.2 in BNB-MF (Figure 1.3)	78
4.5	Indicators of Water Suitability	79
4.6	Generation of Water Suitability Index, W	82
4.7	Implementation Design of Modeling SVM Water Model and Fisheries (Food) Model	85
4.8	Conceptual Representation of Water Suitability Index (W) Generation from SVM Water Model	93
4.9	Implementation Design of Modeling DT Water Model and Fisheries (Food) Model	96
4.10	Conceptual Representation of Water Suitability Index (W) Generation from DT Water Model	98
4.11	Indicators of Fisheries Activities Observance	102
4.12	Generation of Fisheries (Food) Availability Index, F	103

4.13	Conceptual Representation of Fisheries (Food) Availability Index (F) Generation from SVM Fisheries (Food) Model	107
4.14	Quantification of HSS (Link Weight)	123
4.15	Bipartite Habitat Suitability Network (BiHSN)	126
5.1	Content Coverage of Chapter 5	129
5.2	Implementation Design of the Adapted HITS Search Algorithm	131
5.3	Implementation Design of Benchmark Verification	139
5.4	Benchmark Ranking and BiHSN Model Ranking of Location Nodes	140
5.5	Benchmark Ranking and BiHSN Model Ranking of Dolphin Nodes	142
5.6	Implementation Design of Validation using Real-world Data	148
5.7	Bipartite Graph Structure of LDF Dataset used for Validation	149
5.8	Implementation Design of Validation through Past Survey Result	154
5.9	Capture of Literature Review Result (Source: Peter (2012, p. 43))	155
5.10	Literature Review Result overlaid with BiHSN Location Nodes	156
5.11	Assignment of Rank for Literature Review Result and Research Result	157
6.1	Content Coverage of Chapter 6	162
6.2	Crowding Analysis of Location Node with respect to (a) Node Degree; (b) Maximum ID Group Size	163
6.3	Comparing Ranking of Location Nodes based on Node Degree and Maximum ID Group Size with BiHSN Model Ranking	165
6.4	Implementation Design of Pyramidal Similarity Visualization System	170
6.5	BiHSN Pyramidal Similarity Plot	171
6.6	Pyramidal Similarity Plot for L2	172
6.7	Implementation Design of Data Size Uncertainty Analysis	179
6.8	Implementation Design of Potential Predictive Ability Analysis	195
6.9	(a) Actual Position of Location Nodes L1 to L40; (b) Literature Review Result overlaid with Location Nodes L1 to L40	195
6.10	Assignment of Rank for Literature Review Result and Potential Predictive Ability Analysis Result	198
7.1	Content Coverage of Chapter 7	201
7.2	(a) Coexistence Mapping; (b) Complex Network Representation for Location Nodes of BiHSN Model; and (c) Degree Distribution of Complex Network Graph in Figure 7.2(b)	215
B.1	Classification Tree of (a) Sub-model 1; (b) Sub-model 2; (c) Sub-model 3; (d) Sub-model 4 for DT Water Model	256
B.2	Classification Tree of (a) Sub-model 1; (b) Sub-model 2; (c) Sub-model 3 for DT Fisheries (Food) Model	257

LIST OF EQUATIONS

4.1	Overall Accuracy	87
4.2	True Positive Rate (Sensitivity)	87
4.3	False Positive Rate	87
4.4	classregtree for DT Water Model	96
4.5	Specificity	100
4.6	G-mean	100
4.7	classregtree for DT Fisheries (Food) Model	109
4.8	Link Matrix Function	112
4.9	Equation for Fl	113
4.10	Example of Calculation for Fl_1	113
4.11	Equation for Fd	117
4.12	Equation for N	119
4.13	Summation Rule for the Computation of HSS	121
4.14	Multiplication Rule for the Computation of HSS	121
4.15	Sum of the Product Rule for the Computation of HSS	121
4.16	Multiplication of the Sum Rule for the Computation of HSS	122
4.17	Normalization Equation for dl	124
4.18	Example of Calculation for $HSS_{1:3}$	124
5.1	Equation for Hub Matrix Generation	131
5.2	Equation for Authority Matrix Generation	131
5.3	Equation of Principal Eigenvector	133
5.4	Equation for Power Iteration Method	134
5.5	Equation of RMSE (Location Node)	141
5.6	Expression of Calculation for $RMSE_{Loc}$	142
5.7	Equation of RMSE (Dolphin Node)	143
5.8	Expression of Calculation for $RMSE_{Dol}$	143
5.9	Equation of ρ_{HSL_HSS} (Analytical Verification–Location Node)	145
5.10	Expression of Calculation for ρ_{HSL_HSS}	145
5.11	Equation of ρ_{R_HSS} (Analytical Verification–Dolphin Node)	146
5.12	Expression of Calculation for ρ_{R_HSS}	146
5.13	Equation of ρ_{LDFLoc} (Validation using Real-World Data–Location Node)	151
5.14	Equation of ρ_{LDFDol} (Validation using Real-World Data–Dolphin Node)	151
5.15	Expression of Calculation for ρ_{LDFLoc}	153
5.16	Expression of Calculation for ρ_{LDFDol}	153
5.17	Equation of ρ_{LitR} (Validation using Past Survey Result)	157
5.18	Expression of Calculation for ρ_{LitR}	158
6.1	Equation of ρ_{Deg} to Compare BiHSN & Node Degree Ranking	164

6.2	Equation of ρ_{GSize} to Compare BiHSN & Maximum ID Group Size Ranking	164
6.3	Expression of Calculation for ρ_{Deg}	164
6.4	Expression of Calculation for ρ_{GSize}	164
6.5	Equation of Jaccard Index	166
6.6	Example of Calculation of J for (L1,L2) and (L1,L3)	167
6.7	Equation of ρ_{HSL_LDegR} (Correlation Analysis – Location Node)	174
6.8	Equation of $\rho_{HSL_LGSizER}$ (Correlation Analysis – Location Node)	174
6.9	Expression of Calculation for ρ_{HSL_LDegR}	174
6.10	Expression of Calculation for $\rho_{HSL_LGSizER}$	174
6.11	Equation of ρ_{R_DDegR} (Correlation Analysis – Dolphin Node)	176
6.12	Equation of $\rho_{R_DGSizER}$ (Correlation Analysis – Dolphin Node)	176
6.13	Expression of Calculation for ρ_{R_DDegR}	176
6.14	Expression of Calculation for $\rho_{R_DGSizER}$	177
6.15	Equation of ρ_{wLoc} (Data Size Uncertainty Analysis–SVM Water Model)	181
6.16	Equation of ρ_{wDol} (Data Size Uncertainty Analysis–SVM Water Model)	181
6.17	Equation of ρ_{fLoc} (Data Size Uncertainty Analysis–SVM Fisheries (Food) Model)	184
6.18	Equation of ρ_{fDol} (Data Size Uncertainty Analysis–SVM Fisheries (Food) Model)	184
6.19	Equation of ρ_{xDj} for Dolphin Node Uncertainty Analysis	186
6.20	Equation of ρ_{xD1toj} for Dolphin Node Uncertainty Analysis	186
6.21	Equation of $\rho_{ParaLoc}$ for Parameter Significance Analysis	189
6.22	Equation of $\rho_{ParaDol}$ for Parameter Significance Analysis	189
6.23	Equation of ρ_{PreA} for Potential Predictive Ability Analysis	197
6.24	Expression of Calculation for ρ_{PreA}	199

LIST OF ABBREVIATIONS

A	Overall Accuracy
A_{BiHSN}	Overall Accuracy for BiHSN Dataset
A_{ind}	Overall Accuracy for Independent Dataset
ANN	Artificial Neural Networks
BiHSG	Bipartite Habitat Suitability Graph
BiHSN	Bipartite Habitat Suitability Network
BNM	Bipartite Network Modeling
BNB-MF	Bipartite-Network-Based Methodology Framework
CART	Classification and Regression Trees
DT	Decision Trees
f/a-not-observed	Fisheries-activities-not-observed
f/a-observed	Fisheries-activities-observed
G-mean	Geometric or Geometry Mean
HITS	Hypertext Induced Topic Selection
HSI	Habitat Suitability Index
HSI_b	Habitat Suitability Index generated from Benchmark System
HSI_{BiHSN}	Habitat Suitability Index generated from BiHSN Model
HSI_f	HSI obtained from Data Uncertainty Analysis for F Parameter
HSI_w	HSI obtained from Data Uncertainty Analysis for W Parameter
HSS	Habitat Suitability Strength
HST	Habitat Suitability Triangle
IBEC	Institute of Biodiversity and Environmental Conservation
ID	Irrawaddy Dolphin
IUCN	International Union of Conservation of Nature and Natural Resources
LDF	Left-Dorsal-Fin of Irrawaddy Dolphin
LIBSVM	Library for Support Vector Machines

MATLAB	Matrices Laboratory
nSV	Number of Support Vectors
ODE	Ordinary Differential Equations
PDE	Partial Differential Equations
PSU	Practical Salinity Unit
R	Ranking Index for Dolphin Nodes
R_b	Ranking Index generated from Benchmark System
R_{BiHSN}	Ranking Index generated from BiHSN Model
Rk_{LitR}	Ranking of Location Node from Literature Review Result
Rk_{PreA}	Ranking of Location Node from Potential Predictive Ability Analysis
R_f	Ranking Index obtained from Data Uncertainty Analysis for F
R_w	Ranking Index obtained from Data Uncertainty Analysis for W
RBF	Radial Basis Function
RDF	Right-Dorsal-Fin of Irrawaddy dolphin
RMSE	Root Mean Square Error
$RMSE_{Dol}$	RMSE computed for Dolphin Nodes
$RMSE_{Loc}$	RMSE computed for Location Nodes
ROC-AUC	Area Under the Receiver Operating Characteristic Curve
SDP	Sarawak Dolphin Project
SRCC	Spearman's Rank Correlation Coefficient
sus-SVM	Systematic Random Under-sampling SVM
SVM	Support Vector Machines
TNR	True Negative Rate
TPR	True Positive Rate
UCINET	Software for Social Network Analysis
UNIMAS	Universiti Malaysia Sarawak
w-SVM	Weighted SVM
ρ_{Deg}	SRCC for Location Node Degree (Crowding Analysis)
ρ_{fDol}	SRCC for Dolphin Node (Data Uncertainty Analysis for F)

ρ_{fLoc}	SRCC for Location Node (Data Uncertainty Analysis for F)
ρ_{GSize}	SRCC for Location Node Maximum ID Group Size (Crowding Analysis)
ρ_{LDFDol}	SRCC for Dolphin Node (Validation through Real-world Data)
ρ_{LDFLoc}	SRCC for Location Node (Validation through Real-world Data)
ρ_{LiR}	SRCC for Location Node (Validation through Past Survey Result)
ρ_{HSI_HSS}	SRCC for Location Nodes (Analytical Verification)
ρ_{HSI_LDegR}	SRCC for Location Node Degree (Correlation Analysis)
$\rho_{HSI_LGSizeR}$	SRCC for Location Node Maximum ID Group Size (Correlation Analysis)
$\rho_{ParaDol}$	SRCC for Dolphin Node (Parameter Significant Analysis)
$\rho_{ParaLoc}$	SRCC for Location Node (Parameter Significant Analysis)
ρ_{PreA}	SRCC for Location Node (Potential Predictive Ability Analysis)
ρ_{R_DDegR}	SRCC for Dolphin Node Degree (Correlation Analysis)
$\rho_{R_DGSizeR}$	SRCC for Dolphin Node Maximum ID Group Size (Correlation Analysis)
ρ_{R_HSS}	SRCC for Dolphin Nodes (Analytical Verification)
ρ_{wDol}	SRCC for Dolphin Node (Data Uncertainty Analysis for W)
ρ_{wLoc}	SRCC for Location Node (Data Uncertainty Analysis for W)
ρ_{xD1toj}	SRCC for Dolphin Node D1 up to Dolphin Node j (Dolphin Node Uncertainty Analysis)
ρ_{xDj}	SRCC for Exclusion of Dolphin Node j (Dolphin Node Uncertainty Analysis)

CHAPTER ONE

INTRODUCTION

1.1 Introduction of Study

For a newly-surveyed species for example the Irrawaddy dolphin in Kuching Bay, ecological data is scarce and habitat suitability index for the species is not available. In addition, there is no physical law to base on with regards to habitat suitability, and species-specific demographic values are limited, thus, the use of either statistical or deterministic approaches may not be suitable. Therefore, this research attempted to identify and employ an appropriate method to model the habitat suitability of a newly-surveyed species – the Irrawaddy dolphin in Kuching Bay. For such computational modeling problem, the graph-theoretic network modeling approaches which are commented as a heuristic framework and are reported to be applicable in research with little data (Bunn, Urban, & Keitt, 2000) where detail demographic values are normally not required show the potential to be the possible approach. This has motivated us to employ the network modeling approaches in solving our research problems by following a bipartite-network-based methodology framework devised in our research. It has the potential to identify the locations that are more preferred by the species under studied and to rank these locations according to their suitability to function as the habitat through the