



UNIVERSITI PUTRA MALAYSIA

**STRUCTURE OF CO-AUTHORSHIP NETWORK IN UPM JOURNAL
PUBLICATIONS DATABASE FROM 2007-2010**

ZURITA BINTI ISMAIL

IPM 2019 22



**STRUCTURE OF CO-AUTHORSHIP NETWORK IN UPM JOURNAL
PUBLICATIONS DATABASE FROM 2007-2010**

By

ZURITA BINTI ISMAIL

**Thesis Submitted to the School of Graduate Studies, Universiti
Putra Malaysia, in Fulfilment of the Requirements for the
Degree of Master of Science**

June 2018

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in
fulfilment of the requirement for the degree of Master of Science

**STRUCTURE OF CO-AUTHORSHIP NETWORK IN UPM JORNAL
PUBLICATION DATABASE FROM 2007-2010**

By

ZURITA BINTI ISMAIL

June 2018

Chair : Hishamuddin Zainuddin, PhD
Faculty : Institute for Mathematical Research

Co-authorship is one of the most tangible forms of research collaboration. A co-authorship network is a social network in which the authors through participation in one or more joint publications are linked (undirected) to each other. The present work used social network analysis to study co-authorship network of UPM journal publications database for the first four years of Research University (2007-2010) with the aid of Mathematica 11.

The structure of the coauthorship network of publications between 2007 till 2010 was analyzed using the micro-level indicators such as degree centrality, closeness centrality, betweenness centrality and assortativity to observe changes of structure in the network over the early research university years. The study also determines the most connected authors and their influence among coauthors through comparison of their centralities.

Community structure is another typical structure found in social networks. Using modularity approach, the community structures are analysed for each year's co-authorship network and found that then fall naturally onto closely associated areas.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

**STRUKTUR RANGKAIAN KERJASAMA DALAM PANGKALAN DATA
PENERBITAN JURNAL UPM DARI 2007 SEHINGGA 2010**

Oleh

ZURITA BINTI ISMAIL

Jun 2018

Pengerusi : Hishamuddin Zainuddin, PhD
Fakulti : Institut Penyelidikan Matematik

Penulisan bersama adalah salah satu bentuk kolaborasi penyelidikan yang paling ketara. Kolaborasi penulisan bersama adalah rangkaian sosial di mana penulis terlibat dalam satu atau lebih penerbitan bersama (secara tidak langsung) dengan satu sama lain. Penyelidikan ini menggunakan analisis rangkaian sosial untuk mengkaji kolaborasi penulis menerusi pangkalan data penerbitan jurnal UPM bagi Universiti Penyelidikan untuk empat tahun pertama (2007-2010) dengan bantuan perisian Mathematica 11.

Struktur jaringan kerjasama penerbitan antara 2007 hingga 2010 dianalisis dengan menggunakan penunjuk peringkat mikro seperti pemusatan darjah, pemusatan kedekatan, pemusatan keantaraan dan kesisihan untuk melihat perubahan struktur dalam rangkaian kolaborasi sepanjang tahun bermula awal pengiktirafan universiti penyelidikan. Kajian ini juga menentukan penulis yang paling berkolaborasi dan pengaruh mereka di kalangan penulis yang lain melalui perbandingan pemusatan.

Struktur komuniti juga adalah struktur tipikal lain yang terdapat dalam rangkaian sosial. Menggunakan pendekatan modulariti, struktur komuniti bagi kerjasama rangkaian dianalisis untuk setiap tahun dan mendapati ianya kawasan berkaitan antaranya.

ACKNOWLEDGEMENTS

Bismillahirrahmanirrahim. In the Name of Allah, the Most Beneficent and the Most Merciful. Praised be to Allah and the Almighty, Lords of the Worlds for His blessings showered upon me for making the writing of this thesis finally a successful one. However, it would not have been possible without the kind support and help of many individuals. I would like to express my great appreciation for the guidance and assistance received throughout the journey to complete this thesis writing.

First and foremost, I would like to express my very deepest gratitude to my supervisor, Assoc. Prof. Dr. Hishamuddin Zainuddin, for his deepest patience, guidance, times, support and very helpful suggestions throughout this research. His valuable patience, guidance helped me a lot in completing this master thesis. May Allah give him a very good health to continue sharing his knowledge and ideas with other students. My gratitude also to my co-supervisors Prof. Dr. Mohamed Othman for the valuable support and motivation. Many thanks to Dr. Chan Kar Tim for his guidance for programming.

A special thanks to all of the management of Institute for Mathematical Research (INSPERM), UPM and staffs for their support and assistance in my long journey to finish the thesis. I want to thank School of Graduate Studies (SGS), UPM for their kind and generous consideration for me to submit the thesis.

Lastly, my deepest gratitude and appreciation to my lovely husband and my super genius kids for their support and patience. I am really indebted to those who helped me in facing this big challenge in completing this thesis. Thank you very much. May Allah shower all of you with His blessings.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Hishamuddin Zainuddin, PhD

Associate Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Chairman)

Mohamed Othman, PhD

Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

ROBIAH BINTI YUNUS, PhD

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	ii
ACKNOWLEDGEMENTS	iii
APPROVAL	iv
DECLARATION	vi
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER	
1 INTRODUCTION	1
1.1 Introduction to Network	1
1.2 Types of Real World Networks	2
1.3 Networks Models	3
1.3.1 The Random Graph Network	3
1.3.2 Scale Free Network	4
1.4 Problem Statement and Motivations	5
1.5 Objectives	6
1.6 Limits on the Scope Study	6
1.7 Overview	6
2 BACKGROUND AND LITERATURE REVIEW	8
2.1 Co-authorship Network	8
2.2 Co-authorship Network in Publications	9
2.3 Impact of Research and Risk Management	11
3 THEORY AND METHODOLOGY	12
3.1 Defination and Notation	12
3.2 Experimental Design	12
3.3 Data Set Used	13
3.4 Basic Analysis of Social Networks	14
3.4.1 Degree Centrality	14
3.4.2 Closeness Centrality	15
3.4.3 Betweenness Centrality	15
3.4.4 Graph Assortativity	16
3.4.5 Edge Density	16
4 CO-AUTHORSHIP NETWORK ANALYSIS AND FINDINGS FOR JOURNAL PUBLICATIONS OF UPM	17
4.1 Summative Findings	17
4.2 Co-authorship Graph from 2007-2010	18
4.2.1 Co-authorship Graph in 2007	18
4.2.2 Co-authorship Graph in 2008	20

4.2.3	Co-authorship Graph in 2009	21
4.2.4	Co-authorship Graph in 2010	23
4.3	Analysis of Centrality and Graph Assortativity	24
4.3.1	Degree Centrality	24
4.3.2	Closeness Centrality	30
4.3.3	Betweenness Centrality	32
4.3.4	Graph Assortativity	37
5	COMPARISON AUTHORS WITH HIGH CENTRALITIES AND COMMUNITY STRUCTURE FROM 2007-2010	38
5.1	Authors of High Centralities	38
5.1.1	Authors of High Degree Centrality	38
5.1.2	Authors of High Closeness Centrality	38
5.1.3	Authors of High Betweenness Centrality	39
5.2	Community Structure	43
5.3	Summary	47
6	CONCLUSION AND FUTURE WORK	48
	REFERENCES	49
	BIODATA OF STUDENT	53

LIST OF TABLES

Table		Page
1.1	Vertices and edges in network. Some examples of vertices and edges in particular network in M. E. J Newman (2010)	2
4,1	Summary of the Analysis of UPM Research University Publication in 4 years period	17
4.2	Mean, Median and Mode (Closeness Centrality)	30
5.1	Top 20 Authors for Degree Centrality	40
5.2	Top 20 Authors for Closeness Centrality	41
5.3	Top 20 Authors for Betweenness Centrality	42

LIST OF FIGURES

Figure		Page
2.1	An example of a small coauthorship network depicting collaborations among scientists at a private research institution. Nodes in the network represent scientists, and a line between two of them indicates they coauthored a paper during the period of study. This particular network appears to divide into a number of subcommunities, as indicated by the shapes of the nodes, and these subcommunities correspond roughly to topics of research, as discussed by Girvan and Newman (M. Girvan et. al (2002))	10
3.2	Process of Importing and Analyzing dataset	14
4.1	2007 Graph Networks	19
4.2	Giant Component of 2007 Graph Network	19
4.3	2008 Graph Networks	20
4.4	Giant Component of 2008 Graph Network	21
4.5	2009 Graph Networks	22
4.6	Giant Component of 2009 Graph Network	22
4.7	2010 Graph Networks	23
4.8	Giant Component of 2010 Graph Network	24
4.9	Degree Centrality in 2007	25
4.10	Degree Centrality in 2007 (Giant Component)	25
4.11	Degree Centrality in 2008	26
4.12	Degree Centrality in 2008 (Giant Component)	27
4.13	Degree Centrality in 2009	28
4.14	Degree Centrality in 2009 (Giant Component)	28

4.15	Degree Centrality in 2010	29
4.16	Degree Centrality in 2010 (Giant Component)	29
4.17	Closeness Centrality in 2007 (Giant Component)	31
4.18	Closeness Centrality in 2008 (Giant Component)	31
4.19	Closeness Centrality in 2009 (Giant Component)	32
4.20	Closeness Centrality in 2010 (Giant Component)	32
4.21	Betweenness Centrality in 2007 (Giant Component)	33
4.22	Betweenness Centrality in 2008 (Giant Component)	34
4.23	Betweenness Centrality in 2009 (Giant Component)	34
4.24	Betweenness Centrality in 2010 (Giant Component)	35
4.25	Betweenness Centrality in 2007 (Giant Component) and remove 0 frequency	35
4.26	Betweenness Centrality in 2008 (Giant Component) and remove 0 frequency	36
4.27	Betweenness Centrality in 2009 (Giant Component) and remove 0 frequency	36
4.28	Betweenness Centrality in 2010 (Giant Component) and remove 0 frequency	37
5.1	Community Structure in 2007	43
5.2	Community Structure in 2008	44
5.3	Community Structure in 2009	45
5.4	Community Structure in 2010	46

CHAPTER 1

INTRODUCTION

1.1 Introduction to Network

Studies in social networks (Wasserman, S et al, 1994) began a long time ago before complex networks became popular. Social network is a set of people or groups, each of which has connections of some kind to some or all of others. In social network analysis, the people or groups are called 'actors' and the connections 'ties'. Both actors and ties can be defined in different ways depending on the questions of interest. An actor might be a single person, a team, a community or a company. A tie might be a friendship between two people, collaboration or common member between two teams, or a business relationship between companies (M.E.J Newman, 2010).

Social network analysis started in 1930's by psychiatrist, Jacob Moreno, a Romanian immigrant, to America who became interested in the dynamics of social interactions within groups of people. This social network analysis has become one of the most important topics in sociology (Wasserman, S et al, 1994). In recent times, the computer revolution has provided scholars with a huge amount of data and computational resources to process and analyze these data. The size of real networks one can potentially handle has also grown considerably, reaching millions or even billions of vertices. The need to deal with such a large number of units has produced a deep change in the way graphs are approached (Albert Reka et al, 2002).

The most important things to know about social networks is that there are many different possible definitions of an edge in such a network and the particular definition one uses will depend on the questions one wants to answer. Edges may represent friendship between individuals, but may also represent professional relationship, communication patterns, disease infections or many other types of connection. Moreover, the techniques one uses to probe different types of social interaction can also be quite different, so that, different kinds of studies in social network respond to the different questions being addressed.

Historically, the study of networks has been mainly the domain of a branch of discrete mathematics known as graph theory. Since its birth in 1736, when the Swiss mathematician Leonhard Euler published the solution to the Königsberg bridge problem (consisting of finding a round trip that traversed each of the bridges of the Prussian city of Königsberg exactly once), graph theory has witnessed many exciting developments and has provided answers to a series of practical questions (Boccaletti, S et. Al, 2006).

1.2 Types of Real World Networks

Different types of real world network have been widely studied in M.E.J Newman (2010) and will be briefly reviewed in this section. The research of networks is largely dependent based on observations of network properties exhibited by each network.

Networks are also called graphs in the mathematical literature where collections of vertices are joined by edges. Vertices and edges are also called nodes and links in computer science, sites and bonds in chemistry and actor and ties in sociology. Table 1.1 gives some examples of vertices and edges in specific networks in M.E.J Newman (2010).

Table 1.1: Vertices and edges in network. Some examples of vertices and edges in particular network in M.E.J Newman (2010).

Network	Vertex	Edge
Internet	Computer or router	Cable or wireless data connection
World Wide Web	Web page	Hyperlink
Citation network	Article, patent, or legal case	Citation
Power grid	Generating station or substation	Transmission line
Friendship network	Person	Friendship
Metabolic network	Metabolite	Metabolic reaction
Neural network	Neuron	Synapse
Food web	Species	Predation

One can loosely categorize four different types of network, namely social networks, information networks, technological networks and biological networks. For instance, information networks are networks that contain information-theoretic aspects carried by computers or web pages in World Wide Web or citation network of publications. For the World Wide Web, the nodes represent the web pages whereas the edges represent the links between two web pages. Such links are provided by the hyperlinks contains in the web pages and they are directional and are described by directed edges. A.-L. Barabasi et. al (1999) have shown that the World Wide Web possessed the self-organization and scaling properties of random networks. Other examples are the network of citation between US patents and peer to peer networks which are virtual computer networks that allow file sharing among computer users.

Social networks as earlier described are the networks whose vertices represent people and edges represent the interaction between these people. Examples of the social networks are kinship among people, friendship among Facebook users and in society, co-authorship in publications and others. For publication networks, M. E. J Newman (2001) studied the scientific collaboration networks within the disciplines of biomedical sciences, physics and computer science. For friendship network in Facebook, Amanda L. Traud et al (2012) has studied the structure of friendship from various connection group of people all over the world.

1.3 Networks Models

At the same time, much attention has also shifted to create network models which can help us to understand the topological structure of an observed network, how these Real World networks came to be, and how the individual components interact with each other.

1.3.1 The Random Graph Network

A random graph as defined by B. Bollobas (1985) is a collection of points, or vertices, with lines, or edges, connecting pairs of them at random. The study of random graphs has a long history. Starting with the influential work of Paul Erdős and Alfréd Rényi in the 1950s and 1960s discussed in P. Erdos et. al (1959), P. Erdos et. al (1960) and P. Erdos et. al (1961), random graph theory has developed into one of the mainstays of modern discrete mathematics, and has produced a prodigious number of results, many of them highly ingenious, describing statistical properties of graphs, such as distributions of component sizes, existence and size of a giant component, and typical vertex-vertex distances. In almost all of these studies the assumption has been made that the presence or absence of an edge between two vertices is independent of the presence or absence of any other edge, so that each edge may be considered to be present with independent probability p . If there are N vertices in a graph, and each is connected to an average of z edges, then it is trivial to show that $p = z/(N-1)$, which for large N is usually approximated by z/N . The number of edges connected to any particular vertex is called the degree k of that vertex,

and has a probability distribution P_k given by $P_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}$

where the second equality becomes exact in the limit of large N . This distribution we recognize as the Poisson distribution i.e the ordinary random graph has a Poisson distribution of vertex degrees.

Random graphs however turned out to have severe shortcomings as models of real-world phenomena. Studies have been performed on real network such as the network of contacts by which a disease is spread in D. D. Heckathorn (19970,

networks of friendships within a variety of communities discussed in C.C Foster et. al (1963), T. J Fararo et. al (1964) and H. R. Bernard et. al (1988), networks of telephone calls by J. Abello et. al (2000) and W. Aiello et. al (2000), airline timetables by L. A. N. Amaral et. al (2000), and the power grid by D. J Watts et. al (1988), as well as networks in physical or biological systems, including neural networks in D. J Watts et. al (1988), the structure and conformation space of polymers by S. Jespersen et. al (2000), metabolic pathways in D. Fell et. al (2000) and H. Jeong et. al (2000), and food webs in W. Aiello et. al (2000) and L. A. N. Amaral et. al (2000).

It is found in Grossman, J. W et. al (1995) and Grossman, J. W et. al (2002) that the distribution of vertex degrees in many of these networks is measurably different from a Poisson distribution. This strongly suggests, as has been emphasized elsewhere in A. L Barabasi et. al (1999), that there are features of such networks which we would miss if we were to approximate them by an ordinary (Poisson) random graph.

Another very widely studied network is the world-wide web, whose web pages are vertices of the graph and their hyperlinks are edges. Empirical studies have shown that this graph has a distribution of vertex degree which is heavily right-skewed and possesses a fat (power-law) tail with an exponent between -2 and -3 . This distribution is also very far from Poisson, and therefore we would expect that a simple random graph would be a very poor approximation of the structural properties of the world-wide web. In addition, the web differs from a random graph in another in the sense of being directed. Links on one web page to another may be not reciprocated by the other web page.

1.3.2 Scale Free Network

In studies of the networks of citations between scientific papers, De Solla Price, D. J. (1965) showed that the number of links to papers- i.e., the number of citations they received, had a heavy-tailed distribution following a Pareto distribution or power law. Thus that the citation network is said to be scale-free. However, the term "scale-free network", is only defined after a few decades of this discovery. In a later paper in Price, D.J. de Solla (1976) also proposed a mechanism to explain the occurrence of power laws in citation networks, which is called "cumulative advantage". Today it is more commonly known under the name preferential attachment.

Recent interest in scale-free networks started in 1999 with the work of A.-L. Barabási et. al (1999) which mapped the topology of a portion of the World Wide Web. They found that some nodes, which they called "hubs", had many more connections than others and that the network as a whole had a power-law distribution. A few other networks, including some social and biological networks, also had heavy-tailed degree distributions, for which they called them as "scale-

free network". L. A. N. Amaral et. al (2000) showed that most of the real-world networks can be classified into two large categories according to the decay of degree distribution.

A.-L. Barabási et. al (1999) proposed a generative mechanism to explain the appearance of power-law distributions, which they called "preferential attachment" and is essentially the same as that proposed by Price, D.J. de Solla (1976).

1.4 Problem Statement and Motivation

The study of scientific collaboration network in Malaysia is still new. It is not known how collaborative tendencies among scientific researchers are structured in Malaysia's research ecosystem. In other countries, the research in this area is utilized of future research strategies in respective organization.

On 11th October 2006, the Malaysian Cabinet has recognized several public universities as research universities. The declaration of these research universities has profound impact in many aspects of research activities of universities. Universiti Putra Malaysia (UPM) is one of the four universities recognized by the Ministry of Higher Education. Recognition of a research university is based on eight (8) selection criteria determined by the Assessment of Research Universities Committee. These criteria have been developed with a focus on aspects of Research (R) and Development (D) which are based on the criteria adopted by several leading international rating agencies. One of these criteria is to increase the quantity and quality of research and that include increasing number of publications in cited and highly impact factor journals. As one of the strategies, UPM has established mentor-mentee system to encourage established researchers in helping junior researchers to publish in their respective fields.

It is does timely to consider local study of scientific collaboration in Malaysia practically and the research university. In this research, we study the collaborative structure in Universiti Putra Malaysia through coauthorship network for journal publications. Authors (nodes) are linked if they are connected via co-authorship (edges) in papers. From the coauthorship graphs, we will study community of co-authors that play the main role for niche areas in UPM's publications and the various network centralities and structures.

For this study, we constructed coauthorship networks using the raw data of UPM journal publications from UPM's Research Management Centre for the period of 2007-2010. This coincides with the early years of during the early years of UPM's research university. All UPM authors are considered as nodes along with as well as collaborators within and outside UPM. Edges are formed if any

authors or nodes have once coauthored a paper in the year considered. The raw data will be curated for this manipulation in a scientific software. Mathematica 11.0 software will be used to visualize and manipulate the data into various forms and calculate relevant quantities. We use the analysis from developed programs to answer a broad variety of questions about collaboration patterns in UPM and identification of key authors who collaborate prolifically and strategically.

1.5 Objective

The research objectives of this study consist of:

1. To construct graph theoretic structure of UPM scientific coauthorship networks and study their network features.
2. To extract network-theoretic quantities from UPM scientific coauthorship network.
3. To compare the UPM scientific co-authorship network for each year in the period of 2007-2010 in order to extract possible network tendencies.

1.6 Limits on the Scope of Study

It is important to highlight that this study is limited to the network construction of UPM co-authorship network for journal publications the period of 2007-2010 which coincides with the early years of research university status. Since there is no such study before, the present work serves as a baseline study for which future work and analysis can be based on. The journal publications considered are those from journal from Scopus cited journal in all sciences. The selected four years duration of the study is assumed to be sufficient for the network structure to stabilize during the early years of RU status.

By studying co-authorship networks, it must be emphasized that the magnitude of co-authorship does not in any refer to quality of research made by authors. The study merely points out the network structure of co-authorship available in UPM, which may assists engineering of scientific collaboration and in no way makes statement on scientific quality.

Study also makes no attempt of developing new network theoretical structures. It only computes well known quantities from network studies based on UPM journal publications data as stated above.

1.7 Overview

The remainder of this thesis will be organized as follows. In chapter 2, the theoretical background for this thesis is stated. Together with a review of some

of well-known co-authorship networks and known co-authorship network studies local and literature review. Network models are also briefly mentions.

In chapter 3, the dataset used in this work is briefly described. The theory and dataset used in this network are briefly described. This includes construction of co-authorship network, network centralities and other network theoretic-quantities. Programs and computation for the network analysis are developed in using Mathematica software.

Network analysis and findings will be discussed in chapter 4 using the method proposed by M. E. J. Newman (2001a) and M. E. J. Newman (2001b). This analysis includes degree centrality, closeness centrality, betweenness centrality and extended analysis in assortativity as mention earlier.

Comparison of network centralities for each year and their community structure will be discussed in chapter 5. In particular, we will highlight authors of high centralities in those periods and possibly draw some conclusions.

Lastly, the conclusion and potential future research will be presented in the final chapter of this thesis.

REFERENCES

- A. Scala, L. A. N. Amaral and M. Barthélemy (2001). Small World Networks and The Conformation Space of a Lattice Polymer Chain. *cond-mat/0004380*.
- A.-L. Barabasi and R. Albert. (1999). Emergence of Scaling in Random Networks. *Science* 286: 509–512.
- A.-L. Barabási (2003). Scale-Free Networks. *Scientific American*, 288: 50-59.
- A.L. Barabási, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, T. Vicsek (2002). Evolution of the Social Network of Scientific Collaborations. *Physica A*, 311: 590-614
- A.-L. Barabási, R. Albert, H. Jeong (1999). Mean-Field Theory for Scale-Free Random Networks. *Physica A*, 272: 173–187.
- Abbasi, A., Chung, K., and Hossain, L. (2012). Egocentric Analysis of Co-Authorship Network Structure, Position and Performance. *Information Processing and Management*, 48:671– 679.
- Albert, Réka; Barabási, Albert-László. (2002). *Statistical mechanics of complex networks*. *Reviews of Modern Physics*. 74 (1): 47–97.
- Amanda L. Traud, Peter J. Mucha, Mason A. Porter (2012). Social Structure of Facebook Networks. *Physica A: Statistical Mechanics and its Applications*, 391: 4165 – 4180.
- B. Bollobas. (1985). *Random Graphs*. Academic Press, New York.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D. U., (2006). Complex Networks: Structure and Dynamics. *Physics Reports*, 424:175-308.
- Borgatti, S.P. (1995). Centrality and AIDS. *Connections*, 18 (1): 112–114.
- Burt RS. (1995). Structural Holes: The Social Structure of Competition. *Harvard University Press; Cambridge, MA*
- C. C. Foster, A. Rapoport, and C. J. Orwant (1963). A Study of A Large Sociogram: Elimination of Free Parameters. *Behavioural Science*, 8:56–65.
- Cheng, M. Y., Hen, K. W., Tan, H. P., and Fok, K. F. (2013). Patterns of Co-Authorship and Research Collaboration in Malaysia. *In Aslib Proceedings: New Information Perspectives*, 65: 659–674. Emerald Insight.

- D. D. Heckathorn. (1997). Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. *Social Problems*, 44:174–199.
- D. Fell and A. Wagner (2000). The Small World of Metabolism. *Nature Biotechnology*, 18: 1121–1122.
- D. J. Watts and S. H. Strogatz (1988). Collective Dynamics of 'Small-World' Networks. *Nature*, 393: 440–442
- De Solla Price, D. J. (1965). Networks of Scientific Papers. Science, 149 (3683): 510–515.*
- Fatt, C. K., Ujum, E. A., and Ratnavelu, K. (2010). The Structure of Collaboration in the Journal of finance. *Scientometrics*, 85:849– 860.
- Freeman, Linton C. (1979). Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1(3): 215–239.
- Grossman, J. W. & Ion, P. D. F (1995). *Congressus Numerantium* 108:129–131.
- Grossman, J. W (2002). *Congressus Numerantium* 158: 202–212.
- H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.- L. Barabasi (2000). The Large-Scale Organization of Metabolic Networks. *Nature*, 407: 651–654.
- H. R. Bernard, P. D. Kilworth, M. J. Evans, C. McCarty, and G. A. Selley (1988). Studying Social Relations Cross Culturally. *Ethnology*, 2: 155–179.
- J. Abello, A. Buchsbaum, and J. Westbrook (2000). A Functional Approach to External Graph Algorithms. *Proceedings of the 6th European Symposium on Algorithms*.
- J. M. Montoya and R. V. Sole (2002). Small World Patterns in Food Webs. *cond-mat/0011195*.
- Kumar, S. and Jan, J. M. (2013). Mapping Research Collaborations in the Business and Management Feld in Malaysia, 1980-2010. *Scientometrics*, 97:491–517.
- L. A. N. Amaral, A. Scala, M. Barth'el'emy, and H. E. Stanley (2000). Classes of Small-World Networks. *Proc. Natl. Acad. Sci. USA*, 97:11149–11152.
- M. E. J. Newman (2001). The Structure of Scientific Collaboration Networks. *Proc. Natl. Acad. Sci. USA*, 98: 404–409.
- M. E. J. Newman (2001a). Scientific Collaboration Networks: II. Shortest Paths, Weighted Networks, and Centrality. *Phys. Rev. E*, 64 (016132).
- M. E. J. Newman (2001b). Scientific Collaboration Networks: I. Network Construction and Fundamental Results. *Phys. Rev. E*, 64 (016131).

- M. E. J. Newman (2002). Assortative mixing in networks. *Phys. Rev. Lett.* 89: 208701.
- M. E. J. Newman (2004). Coauthorship Networks and Patterns of Scientific Collaboration *Proc. Natl. Acad. Sci. USA*, 101: 5200–5205.
- M. E. J. Newman (2000). Models of the small world, *J. Stat. Phys.* 101, 819–841.
- M. E. J. Newman (2010). *Networks: An Introduction*. Oxford University Press.
- M. Girvan and M. E. J. Newman (2002). Community Structure in Social and Biological Networks. *PNAS*, 99 (12): 7821-7826.
- M. E. J. Newman (2001). *Proc. Natl. Acad. Sci. USA*, 98:404–409.
- P. Erdos and A. Renyi (1959). On Random Graphs. *Publicationes Mathematicae*, 6: 290–297.
- P. Erdos and A. Renyi (1960). On the Evolution of Random Graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5: 17–61.
- P. Erdos and A. Renyi (1961). On The Strength of Connectedness of a Random Graph. *Acta Mathematica Scientia Hungary*, 12: 261–267.
- Price, D.J. de Solla (1976). A General Theory of Bibliometric and Other Cumulative Advantage Processes. *Journal of the American Society for Information Science*. 27(5): 292–306.
- R. J. Williams and N. D. Martinez (2000). Simple Rules Yield Complex Food Webs. *Nature*, 404: 180–183.
- S. Jespersen, I. M. Sokolov, and A. Blumen (2000). Small-World Rouse Networks as Models of Cross-Linked Polymer. *J. Chem. Phys.* 113: 7652–7655.
- T. J. Fararo and M. Sunshine (1964). *A Study of a Biased Friendship Network*, Syracuse University Press, Syracuse, NY.
- Tan, H., Ujum, E., and Ratnavelu, K. (2014). A Scientometrics and Social Network Analysis of Malaysian Research in Physics. *Frontiers in Physics*, 1588:241–245.
- W. Aiello, F. Chung, and L. Lu (2000). A Random Graph Model for Massive Graph. in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*.

Wasserman, S., & Faust, K (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press, 8.

Watts, Duncan J.; Strogatz, Steven H. (1998). Collective Dynamics of 'Small-World' Networks. *Nature*. 393 (6684): 440–442.



BIODATA OF STUDENT

Zurita Ismail was born on 16 June 1980 in Kluang, Johor, Malaysia. From 1987 to 1997 she attended primary and secondary schools in Johor and Melaka. She continued her study at Johor Matriculation in 1998 for one year. Then she pursued her degree in 1999 at Universiti Putra Malaysia with Bachelor of Science (Honour) in Mathematics. She graduated in 2002. She starts working at Institute for Mathematical Research (INSPEM), UPM in 2003 as a Science Officer and change her job as a Research Officer in 2008 until now. She is now working on Complex Networks as her main research and will continue her interest in this research for her expertise.





**UNIVERSITI PUTRA MALAYSIA
STATUS CONFIRMATION FOR THESIS / PROJECT REPORT
AND COPYRIGHT
ACADEMIC SESSION: 2017/2018**

TITLE OF THESIS / PROJECT REPORT:
**STRUCTURE OF CO-AUTHORSHIP NETWORK IN UPM JOURNAL
PUBLICATIONS DATABASE FROM 2007-2010**

NAME OF STUDENT: ZURITA BINTI ISMAIL

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as:

*Please tick (✓)

CONFIDENTIAL

(Contain confidential information under Official Secret Act 1972).

RESTRICTED

(Contains restricted information as specified by the organization/institution where research was done).

/

OPEN ACCESS

I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for:

PATENT

Embargo from _____ until
(date) (date)