EFFICIENT HYBRID REDUCTION FOR BINARY BASED INFORMATION SYSTEM IN SOFT SET THEORY

AHMAD NAZARI MOHD ROSE

A thesis submitted in
fulfillment of the requirement for the award of the
Degree of Doctor Philosophy of Information Technology

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia

MAY, 2016

This thesis is lovingly dedicated to my parents. For their endless love and prayers. And also a heartfelt dedication to my wife and children for the unwavered support, understanding, sacrifices and encouragement.

ACKNOWLEDGEMENT

I would like to express the deepest appreciation to my supervisor, Prof. Dr Mustafa Mat Deris, who has been enthusiastically guiding and encouraging me in this PhD research. His untiring perseverance and relentless energy in seeing and pushing me through in this journey is unmatched and inspiring. And only to the Almighty ALLAH, I seek and ask that Prof. Dr Mustafa to be granted a healthy and a blessed life for his patience in me as his student. Without his persistent help, this thesis would not have been possible. I would like to thank Dr. Tutut Herawan who has introduced me the soft set theory. I would never forget the heartfelt motivation that they have given me. A big thank you to Universiti Tun Hussien Onn Malaysia (UTHM) for giving me the opportunity to pursue my unfinished dream. And a giant thank you also to the Universiti Sultan Zainal Abidin (UniSZA) for the trust in me . I am indeed very thankful to my colleagues for their support and encouragement.

A special thanks to my mother, mother in law and children for all of the sacrifices and prayers. The prayer for me was what sustained me thus far. I would also like to thank all of my friends who supported me in writing, and incented me to strive towards my goal.

Finally, I feel very grateful and blessed to have a wife, Mrs Najmah for providing the resolute support and accompanying me late in the night burning the midnight oil, motivating and aspiring to see me ending this wonderful journey successfully. May ALLAH grant us all, endless bliss and exquisite felicity.

ABSTRACT

In soft set literatures, issues regarding reduction techniques with regards to dataset in soft set have been discussed and analyzed. The existing reduction techniques discussed were the techniques based on rough set guidelines and parameter reduction. All of the proposed techniques have successfully reduced the datasets but the factors of consistency and accuracy are still outstanding. Based on the research done on data transformation in soft set theory, the three newly introduced reduction methods will be integrated into a technique known as Hybrid Reduction in Soft Set (HRSS). HRSS consists of two(2) types of parameter reduction and a newly proposed object reduction. The proposed technique has been implemented and the results were compared to the existing techniques, and HRSS was found to be 100% consistent, accurate and able to reduce the data substantially. With SRR (Soft Set Rough Reduction) and Parameter Reduction (PR) being ineffective with respect to consistency and accuracy, further analysis on the data size achieved by HRSS and Normal Parameter Reduction (NPR) were then considered. HRSS has also demonstrated efficiency when searching for decisional values. Lastly, HRSS has also been found to be the least complexed in terms of the algorithm used. With the results obtained, it is safe to conclude that, decision-making that are based on selected datasets that have undergone the HRSS processing is competent.

ABSTRAK

Dalam kajian set lembut, isu-isu mengenai teknik pengurangan berkaitan dengan dataset di set lembut telah dibahas dan dianalisa. Teknik-teknik pengurangan sedia ada yang dibincangkan adalah teknik berdasarkan garis panduan yang ditetapkan teori set kasar dan pengurangan parameter. Semua teknik yang dicadangkan telah berjaya mengurangkan set data namun faktor-faktor seperti konsisten dan ketepatan masih lagi menjadi isu yang perlu diselesaikan. Kajian kami dalam hal bersangkutpaut dengan transformasi data ini telah membuka peluang kepada kami untuk memperkenalkan tiga teknik pengurangan yang akan disepadukan ke dalam teknik yang dikenali sebagai Pengurangan Hybrid dalam Set lembut (HRSS). HRSS terdiri dari dua jenis pengurangan secara atribut dan teknik baru pengurangan secara objek. Teknik yang dicadangkan itu telah dilaksanakan dan keputusan dibandingkan dengan teknik yang sedia ada, dan HRSS didapati 100% konsisten, tepat dan dapat mengurangkan data dengan baik. Apabila kajian dari kami mendapati SRR (Pengurangan secara Set kasar dalam Set Lembut) dan Pengurangan Parameter (PR) sebagai tidak berkesan dari segi konsisten dan ketepatan, analisis lanjut kepada saiz data yang dicapai oleh HRSS dan Parameter Normal Pengurangan (NPR) terus dipertimbangkan. HRSS juga telah memperlihatkan keupayaan yang cekap dalam mengenalpasti objek optimal. Dari keputusan yang diperolehi, berdasarkan operasi dan ujian kepada set set data terpilih, kami dengan yakin yang HRSS adalah teknik pengurangan data yang terbaik dalam set lembut.

TABLE OF CONTENTS

DECLARATION		**ii**
DEDICATION		**iii**
ACKNOWLEDGEMENT		**iv**
ABSTRACT		**v**
ABSTRAK		**vi**
TABLE OF CONTENTS		**vii**
LIST OF TABLES		**x**
LIST OF FIGURES		**xiii**
LIST OF ALGORITHMS		**xiv**
LIST OF SYMBOLS AND ABBREVIATIONS		**xv**
LIST OF PUBLICATIONS		
CHAPTER 1		**1**
1.1	Background	1
1.2	Problem Statement	7
1.3	Objectives and Scope	9
1.4	Research Significance	10
1.5	Thesis Organization	11
CHAPTER 2		**12**
2.1	Introduction	12
2.2	Information System	15

2.3	Indiscernibility Relation		19
2.4	Reduction		20
2.5	Approximations		22
2.6	Reduction Techniques in Soft Set		24
	2.6.1	Analysis of Attribute Reduction of Soft Set by Maji *et al.* (2003) [14]	27
	2.6.2	Analysis of Parameterization Reduction of Soft Set by Chen *et al.* (2003,2005) [29], [63]	29
	2.6.3	Analysis of parameterization reduction of soft set by Kong *et al.* (2008) [17]	31
2.7	Discussion		32
2.8	Summary		34
CHAPTER 3			**35**
3.1	Soft Decision Making Based on Hybrid Reduction		35
	3.1.1	Supported Set Parameter Reduction	36
	3.1.2	Significance Based Parameter Reduction	45
	3.1.3	Supported Set Object Reduction	47
	3.1.4	The Algorithm of HRSS	53
3.2	Summary		55
CHAPTER 4			**57**
4.1	Proposed Technique		57
4.2	Analysis of Reduction Techniques in Soft Set Theory		58
	4.2.1	Samsung Mobile Phone Dataset	58
	4.2.2	Analysis of Hybrid Reduction of Soft Set (HRSS)	59
	4.2.3	Samsung Result Comparison Analysis	72
	4.2.4	Comparison Analysis From Other Datasets	72
4.3	Discussion		97
4.4	Summary		107
CHAPTER 5			**109**
5.1	Research Summary		109
5.2	Research Contribution		111

5.3 Limitation 112

5.4 Conclusion 113

VITAE **124**

THE LONGEST YARD **125**

LIST OF TABLES

Table 2.1: An information system [83] 16
Table 2.2: Tabular representation of angina diagnosis 18
Table 2.3: Tabular representation of a soft set from the Example 2.1 24
Table 2.4: Tabular representation of a soft set from Example 2.2 27
Table 2.5: A binary table of representation of $R \subset E$ 28
Table 2.6: Perfomance analysis of SRR 29
Table 2.7: A binary table of representation of $S \subset E$ 30
Table 2.8: Performance analysis of PR 31
Table 2.9: A binary table of representation of $R \subset E$ 32
Table 2.10: Performance analysis of NPR 32
Table 3.1: The reduct set of Z 43
Table 4.1: The datasets used in the experiments 58
Table 4.2: Tabular representation of a soft set from Samsung 63
Table 4.3: Objects and attribute for reduction of the soft set from Samsung 70
Table 4.4: Comparison of the identified core sets of Samsung's datasets 71
Table 4.5: Compression percentage on Samsung dataset 72
Table 4.6: Comparison of the identified core sets of Nokia's datasets 75
Table 4.7: Comparison of identified core sets for Car dataset 75
Table 4.8: Comparison of the identified core sets for Animal dataset 76
Table 4.9: Comparison of the identified core sets for heart's dataset 76
Table 4.10: Comparison of the identified core sets for Vote dataset 77
Table 4.11: Comparison of the identified core sets for Student's dataset 77
Table 4.12: Comparison reduced table 78
Table 4.13: Comparison of reduced dataset against Car's dataset 79

Table 4.14: Comparison of reduced dataset against Animal's dataset79
Table 4.15: Comparison of Heart reduced table 80
Table 4.16: Comparison of Votes reduced table 80
Table 4.17: Comparison of Students reduced table 81
Table 4.18: Optimal objects obtained with Samsung dataset 81
Table 4.19: Comparison of Optimal Values Obtained with Nokia dataset 82
Table 4.20: Optimal values comparison with Car's dataset 82
Table 4.21: Comparison of optimal objects with Animal's dataset 83
Table 4.22: Comparison of optimal objects with Heart dataset 83
Table 4.23: Comparison of optimal objects with Votes dataset 84
Table 4.24: Comparison of optimal objects with Students dataset 84
Table 4.25: Comparison of suboptimal values with Samsung's dataset 85
Table 4.26: Comparison of Suboptimal Objects 85
Table 4.27: Comparison of suboptimal objects for Animal's dataset 86
Table 4.28: Comparison of suboptimal objects for Heart's dataset 86
Table 4.29: Comparison of suboptimal objects for Votes dataset 87
Table 4.30: Comparison of suboptimal objects for Student dataset 87
Table 4.32: The decision computational complexity comparison 91
Table 4.33: Consistency achievement for Samsung's dataset 92
Table 4.34: Consistency achievement for Nokia dataset 92
Table 4.35: Consistency achievement for Car dataset 93
Table 4.36: Consistency gauge for Animal dataset 93
Table 4.37: Comparison of consistency in Heart's dataset. 93
Table 4.38: Variation on Samsung's dataset 94
Table 4.39: Accuracy on Samsung's dataset 94
Table 4.40: Variation on Nokia's dataset 95
Table 4.41: Accuracy Variation on Car's dataset 95
Table 4.42: Accuracy Variation on Animal dataset. 96
Table 4.43: Accuracy Variation on Heart's dataset. 96
Table 4.44: Variation on Vote dataset 96
Table 4.45: Variation on Vote dataset 97
Table 4.46: The decision computational complexity comparison 97

Table 4.47: The Samsung comparison results 98
Table 4.48: The Nokia comparison results 99
Table 4.49: The Car comparison results 99
Table 4.50: The Animal comparison results 100
Table 4.51: The Heart comparison results 101
Table 4.52: The Vote comparison results 102
Table 4.53: The Student comparison results 102
Table 4.54: Comparison of accomplishments by HRSS and NPR 106

LIST OF FIGURES

Figure 2.2: The mapping of parameters	23
Figure 2.3: The soft set	23
Figure 2.4: The Soft Set	26
Figure 3.1: The achieved co-occurrence from Example 2.2	37
Figure 3.2: The support of each parameter	38
Figure 3.3: The achieved co-occurrence	43
Figure 3.4: The support of each parameter	44
Figure 4.1: The mapping of parameters	61
Figure 4.2: The soft set representation	62
Figure 4.3: Co-occurrence of reduced set	64
Figure 4.4: The supported set values	65
Figure 4.5: The co-occurrence of the soft set	67
Figure 4.6: The supported set of the reduced set	69

LIST OF ALGORITHMS

3.1	Supported set parameter reduction	42
3.2	Significance based parameter reduction	46
3.3	Supported set object reduction algorithm	51
3.4	The algorithm of hybrid reduction supported set	54

LIST OF SYMBOLS AND ABBREVIATIONS

(F,E)	-	A pair (F,E) is called soft set over universe (U)
HRSS	-	Hybrid Reduction of Soft Set
SBPR	-	Significance Based Parameter Reduction
SRR	-	Soft Set Rough Reduction
PR	-	Parameter Reduction
NPR	-	Normal Parameter Reduction
$R \subset E$	-	Set R as the subset of E
$S=(U,A,V,f)$	-	An information system
$S=(U,A,V_{\{1,0\}},f)$	-	A Binary information system
$IND(B)$	-	An indiscernibility relation
U/B	-	The equivalence class in the partition
$B*$	-	Subset of B
$\|U\|$	-	The cardinality of universe
M_E	-	The maximum value of f_E
$E \setminus a_x$	-	Set of parameters without a_x
$U/(B-\{a\})=U/B$	-	An equivalence class without parameter a
$leastsig(e_m)$	-	Least significant parameter
$mostsig(e_m)$	-	Most significant paramter
$ultmaxsupp(u_{i_j})$	-	Ultimately maximal supported
$\text{supp}(u_{i_k})$	-	Supported set value(k) for an object u_i
$coo(u_i)$	-	The co-occurrence of an object u_i
$\underbrace{\{u_x,u_y,u_z\}_a}_{\max}$	-	A maximally supported set with the value of a

$\underbrace{\{u_x, u_y, u_z\}_0}_{\min}$	-	A minimally supported set with the value of 0
UCI	-	University California Irvine Machine Learning Repository

LIST OF PUBLICATIONS

Journals:

(i) A. N. M. Rose, M. I. Awang, H. H., and M. M. D., "Comparison of Techniques in Solving Incomplete Datasets in Softset," *Int. J. Database Theory Appl.*, vol. 4, no. 3, pp. 1–12, 2011

Proceedings:

i) A. N. M. Rose, M. I. Awang, H. Hassan, A. H. Zakaria, T. Herawan, and M. M. Deris, "Hybrid reduction in soft set decision making," in *Advanced Intelligent Computing*, Springer, 2012, pp. 108–115.

ii) A. N. M. Rose, H. Hassan, M. I. Awang, N. A. Mahiddin, H. M. H. M. Amin, and M. M. Deris, "Solving Incomplete Datasets in Soft Set Using Supported Sets and Aggregate Values," *Procedia CS*, vol. 5, pp. 354–361, 2010.

iii) A. N. M. Rose, T. Herawan, M. M. Deris, and M. M. D., "A Framework of Decision Making Based on Maximal Supported Sets," *Lect. Notes Comput. Sci. Adv. Neural Networks-ISNN 2010, Springer Berlin Heidelber*, pp. 472–482, 2010.

CHAPTER 1

INTRODUCTION

This chapter introduces the context of the research that is to be presented in this thesis. It starts off with an introduction to the general area of uncertain data, soft computing, soft set and discusses the fundamental motivations behind data reduction in soft set theory. Then, it presents the objectives and the scope of the research. The chapter ends with a discussion on the organization of the rest of this thesis.

1.1 Background

In the era where time is money, lengthy process in decision-making is unwelcome. The state of decision-making requires a speedy and accurate decision and is regarded as vital in the process of decision-making in any organization. As the adage says "An inch of time is gold, but an inch of time cannot be purchased by an inch of gold", it explains how time is valuable and is irreplaceable, thus attention must be paid to practices which are not properly managed and time-wasting. It also highlights how bad time management might cause an organization to lose profit and its leading edge. The process of making an accurate and right decision is closely linked to the availability of the information and how the information can be mined into knowledge. The task of making a speedy and accurate decision is crucial to an organization. It is crucial that within the limited time that if there is an increase in poor decision-making, it may lead to the decline in the profit and cutting edge of the organization [1]. The availability of the information and how the information can be mined into knowledge is closely linked to the process of making an accurate and right decision [2]. Therefore, if the right information can be made available through some feasible and competent techniques, then the process of decision-making will not be hindered by the performance of the process.

Huge dataset exists in lots of areas and data reduction is deemed important in areas of image processing, engineering, medical and so forth. The data reduction technique is in dire need because large memory is required when processing large datasets. If an efficient data reduction technique is available, then it will help free lots of computer resources and thus contributing to increase in the performance of the computer.

In image processing, data reduction plays a vital role in ensuring the amount of information that is just sufficient for further processing. Some techniques that has been in used for data reduction in image processing are such as clustering [3] [4], data reduction scheme for triangulated surfaces [5], adaptive quantization [6] and dimensionality reduction [7] [5]. In the engineering field, data reduction has been achieved through the techniques such as fourier transform analysis [8], Laplacian Eigenmaps [9], adaptive reduction [10], histogram reduction method [11] and so forth. While in the field of medical, symptoms observed can also be reduced through the process of different combinations of reduction [12], wavelet based statistical based on the generalized Gaussian distributed (GGD) modeling [13], application of data-reduction algorithms on ECG data [14] and automatic data reduction for the identification of novel biomarkers of renal toxicity [15]. None of the various techniques that were used as part of data reduction in the field of engineering, medical and image processing process mentioned about object reduction as one of the alternatives. Therefore it is also very appropriate and timely, to discuss the approach of object reduction as part data reduction process in this thesis.

Data reduction must be properly taken care of in ensuring that data available are adequate for the process of decision-making. Hand *et al.* [16] has in-fact stressed for the need of sophisticated examination methods to highlight features which would be readily apparent in small data sets. Another important issue of concern is the quality of data, that is ensuring the data in reduced form is still considered as quality data [17]. It is also a very important element in data reduction such that data integrity is still been maintained even though it is now represented in a reduced form. Decision-making is a very critical chore that needs to be complemented with complete but not extraneous information. Some extraneous data that occurred may be due to redundancy.

In a normal database, a proven method of eliminating redundancies is through normalization [17]–[21]. But in this thesis, since the scope of discussion with regards to data representation in soft set theory is only focused to binary form, normalization as practiced in database management will not be discussed.

While the discussion on the redundancy in datasets in soft set theory has led to researchers to the usage of data reduction techniques in eliminating redundancies or non vital data. The soft set theory has an impact in the field of data reduction. The data reduction helps to improve decision making with less association of data. With less association of data, decision-making will be more competent. This is very crucial in situations where quick decision-making is required urgently and it would require the attention in alienating non-vital data from the dataset. An original Boolean-valued database usually contain large amount of data. To some extent, one does not actually need the whole amount of data for decision-making. If huge amounts of information are involved during the process of decision-making, then factors such as processing times and inefficiencies will be the main cause of concern. It is a great concern because any redundant information will have an effect on the processing time due to extraneous data. When longer time is needed to process the data, it will affect the competency of the process of decision-making. In some processes, time is crucial. Therefore, opting for lesser amount of data to be processed will be a better alternative.

As part of data transformation in achieving better performance, data needs to be reduced from its original gigantic size. But the process of data reduction must be handled with proper precautions to ensure the quality of data of the information is maintained.

An issue to addressed during data transformation, is whether all the data are required? Is it necessary, to process all the data, or can the process of decision making be made with less data. Data reduction In order to perform reduction, one may encounter this problem: for a particular property, whether all the attributes in the set of attributes are always necessary to conserve this property [17]. Data reduction must be properly taken care of in ensuring that data that are made available are adequate for the process of decision-making.

An important issue that is of concern is the quality of data in the reduced form is still considered as quality data [17]. Hand & Kamber [18] has defined the quality of data in terms of consistency, accuracy, completeness, timeliness,

believability, and interpretability, which are assessed based on the intended use. Kong *et al.* [22] has defined its main objective of normal parameter reduction is to provide consistency in selecting vital valued objects. And it is very important that, any inconsistencies in the dataset can be identified, as inconsistencies contribute to inaccurate data [23]. Therefore it is the prerogative of any reduction algorithm, that it helps in upholding data consistency.

The advantage that will be gained by efficient data reduction techniques is towards less energy consumption used in the storage, efficient data transmission and storage access [24]. For a highly complex computational algorithm, the resources involved will be a lot, it may sometimes hinder processing efficiency of the datasets. Aspects that need to be considered when evaluating the quality of data is the consistency and the accuracy of object derived when reduction has been completed.

What makes a situation become more complex is that the state of data that is available is uncertain. The techniques that are currently available might contribute to either lengthy process or inaccurate decision. Dennis Lindley [25] had insisted in his book that "uncertainty is everywhere". Interestingly, most of the problems in the world such as in the field of engineering, medical science, social science and pure science, involves uncertain data.

As mentioned, a lot of real life aspects are also surrounded by lots of uncertainty in nature, this is especially true with regards to the field of social sciences, economics, sports, medical sciences, engineering etc. The collected data in these areas may soon be very huge and it would require an intelligent approach in handling them. As pointed out by Aggarwal [26], data which contains errors are referred to as uncertain data sets as those in which the level of uncertainty can be quantified in some way and these imprecise data may come from the following scenarios:

(i) The scientific measurement techniques are also not precise which may also be the reasons for the availability of imprecise data. During the measurement, the data collected may just be an approximation rather than a precise value presentation of the measurement intended to. Even if a value is used, it will also depend on the unit that has been used. Suppose for example, a measuring device is measuring a pen with units of one hundredth of a centimetre, and its length to be 4.30 cm. But, a more sophisticated and expensive device might give a value of 4.30312 cm. And again if measured in millimetre, a different

value would also be obtained. Therefore, there is no limit to how precisely you can measure the pen.

(ii) Many new hardware technologies such as sensors generate data which are imprecise. Sensors are used to provide reading based on contamination, wind reading, data from rainfalls, sea level, motion and so on for observation purposes. The sensors collect and communicate their readings to the sensor databases, the system will then analyse the data received and will highlight the follow-up measures if any, based on the data received and analysed. Ma *et. al* [27] has stated that due to continuous changes and possible errors in these values, the data values recorded in sensor databases may differ from the actual status, thus resulting in imprecise data collected. The quality of apparatus used may also be the hindrance in keeping track with changes in reading continuously. Sometimes the diminishing of power used in the apparatus may also cause problems to the apparatus continuously keeping track of data changes. These highlighted scenarios are all the contributing factors as to why sensored data are imprecise in nature.

Aggarwal [26] had also highlighted that the field of uncertain data management presented a number of challenges in terms of collecting, modeling, representing, querying, indexing and mining the data. The solutions of such problems involved the use of mathematical principles based on uncertainty and imprecision.

Zadeh [28] has stated that some scientist have characterized uncertain data and imprecise data as soft data. Zadeh [28] did also emphasized that the existing mathematics at that time were purely based on techniques from probability theory and statistics when coping with applications dealing with parameter estimation, hypothesis testing and system identification when encountered in various field, such as in medicine, sociology, anthropology and system analysis, and the techniques used are not that effective when dealing with this kind soft data.

When dealing with this soft data, another term known as soft computing has been coined. Zadeh [29] had defined soft computing as actually a technique of computing based on its tolerance to imprecision, uncertainty and approximation in achieving tractability, and better rapport with reality.

Currently, there are a lot of theories and discussions pertaining to the uncertainties. Among them, are the fuzzy set theory [30], rough set theory [31], [32]; and the theory of interval mathematics [33], [34]. However, Molodtsov [35] pointed

out some of the shortcomings in all theses theories. As for the topic of the interval mathematics, taking into account the error of calculations by constructing an interval estimate for the exact solutions of a problem, but it is not sufficiently adaptable for problems with different uncertainties since the interval estimation becoming wider. Among these theories, the fuzzy set theory by Zadeh [30] has been considered as the most appropriate theory to deal with uncertainty. Nevertheless, every union and intersection operations in fuzzy set have three rules of membership function so-called standard, algebraic and boundary that contribute to deficiencies of fuzzy set in setting the membership function in particular cases. According to Molodtsov [35], the reason for all of these difficulties is due to the inadequacy of the parameterized tool of the spoken theories. Consequently, he had introduced the soft set theory as a new mathematical tool, which he claimed to having enough parameterization tools to deal with uncertainties [35]. The soft set theory used parameterization sets as its main solution for problem solving, which makes it very convenient and easy to apply in practice as demonstrated by [35] in various applications. Presently, great advancement of study in soft set theory has been achieved on the theoretical and applications side by the researchers.

Nowadays, problems related to uncertain data can be solved using mathematical principles, and one of them is the Soft Set Theory [12-13]. Soft Set has also now become part of soft computing family due to its ability to deal with uncertain and imprecise data. Soft sets are also called binary, basic, elementary neighborhood systems [35].

Over the past 10 years, there have been a great number of researches on the soft set theory. It has become a topic of interest to researchers in providing solution based on soft set theory in areas of decision making, data reduction, incomplete data sets, clustering, forecasting, fuzzy soft set, and association rules mining. The reason for the interest from researchers as pointed out by Molodtsov [35] was that the main advantages of soft set theory is that it is free from the inadequacy of the parameterization tools, unlike in the theories of fuzzy set, probability and interval mathematics [35].

In recent years, great progresses have been achieved in the soft theories. The researches cover issues on the fundamental soft set theory, soft set theory in abstract algebra and soft set theory for data analysis. There has also been extensive research in decision making in [37], [38], [22], [39]–[51], incomplete dataset in [31]–[35],

applications with soft set [57]–[63], data mining [64]–[66], merging with other mathematical fundamentals [29-39] and so on, has been researched and developed rapidly in meeting the numerous demands in real-world situations. The soft set theory also has an impact in the field of data reduction.

There were also researches that have been done including the fundamental research on soft set theory [36], [67], [68], soft set theory in abstract algebra [69], research on relation to rough sets [70]–[74] and fuzzy soft sets [29], [41], [43], [48], [75]–[82].

One of the frequently talked about topic with regards to soft set was on data reduction as been highlighted in [22], [50], [51], [64], [83], [84]. The main objective of data reduction is to improve decision making with less association of data. An original Boolean-valued database does usually contain large amount of data. To some extent, one does not actually need the whole amount of data for decision-making. For the purpose of data transformation, data needs to be reduced from its original size. In order to perform reduction, one may encounter this problem: for a particular property, whether all the parameters in the set of parameters are always necessary to conserve this property [85].

1.2 Problem Statement

Data reduction techniques can be applied to obtain a reduced representation of the data set that is much smaller in volume, but must still preserve the integrity of the original data and at the same time has the advantage of being more efficient when been processed by the same technique [18]. But the issues of consistency and accuracy with regards to data reduction are still questionable, when some researchers have opted to just maintaining the optimal values and neglecting other ranked values. Even the reduction approach which was based on the rough reduction, where a decision will be selected based on the maximal weighted was still questionable with regards to consistency.

If consistency in terms of objects derived could not be achieved, then decision-making will thus be worthless due to the issues of irregularity. Consistency must be achieved at any cost, at least at the level of optimal and suboptimal values. [37], [22], [50], [86] had discussed the issue of consistency by focusing on the

REFERENCES

[1] M. I. Hwang, "Decision making under time pressure: a model for information systems research," *Inf. Manag.*, vol. 27, no. 4, pp. 197–203, 1994.

[2] A. Neely and Y. Jarrar, "Extracting value from data-the performance planning value chain," *Bus. Process Manag. J.*, vol. 10, no. 5, pp. 506–509, 2004.

[3] S. Eschrich, J. Ke, L. O. Hall, and D. B. Goldgof, "Fast accurate fuzzy clustering through data reduction," *Fuzzy Syst. IEEE Trans.*, vol. 11, no. 2, pp. 262–270, 2003.

[4] E. Bingham and H. Mannila, "Random projection in dimensionality reduction: applications to image and text data," in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, 2001, pp. 245–250.

[5] B. Hamann, "A data reduction scheme for triangulated surfaces," *Comput. Aided Geom. Des.*, vol. 11, no. 2, pp. 197–214, 1994.

[6] W. Jass, "Process for adaptive quantization for the purpose of data reduction in the transmission of digital images." Google Patents, 1995.

[7] M. I. Sezan, "A peak detection algorithm and its application to histogram-based image data reduction," *Comput. vision, Graph. image Process.*, vol. 49, no. 1, pp. 36–51, 1990.

[8] C. F. Hess and J. D. Trolinger, "Panicle Field Holography Data Reduction By Fourier Transform Analysis," *Opt. Eng.*, vol. 24, no. 3, p. 243470, 1985.

[9] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, no. 6, pp. 1373–1396, 2003.

[10] E. H. Lim and D. Suter, "Conditional random field for 3D point clouds with adaptive data reduction," in *Cyberworlds, 2007. CW'07. International Conference on*, 2007, pp. 404–408.

[11] G. J. Kutcher, C. Burman, L. Brewster, M. Goitein, and R. Mohan, "Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 21, no. 1, pp. 137–146, 1991.

[12] J. Zvárová and M. Studen\'y, "Information theoretical approach to constitution and reduction of medical data," *Int. J. Med. Inform.*, vol. 45, no. 1, pp. 65–74, 1997.

[13] S. Gupta, L. Kaur, R. C. Chauhan, and S. C. Saxena, "A wavelet based

statistical approach for speckle reduction in medical ultrasound images," in *TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region*, 2003, vol. 2, pp. 534–537.

[14] J. P. Abenstein and W. J. Tompkins, "A new data-reduction algorithm for real-time ECG analysis," *Biomed. Eng. IEEE Trans.*, no. 1, pp. 43–48, 1982.

[15] E. Holmes, J. K. Nicholson, A. W. Nicholls, J. C. Lindon, S. C. Connor, S. Polley, and J. Connelly, "The identification of novel biomarkers of renal toxicity using automatic data reduction techniques and PCA of proton NMR spectra of urine," *Chemom. Intell. Lab. Syst.*, vol. 44, no. 1, pp. 245–255, 1998.

[16] D. J. Hand, H. Mannila, and P. Smyth, *Principles of data mining*. MIT press, 2001.

[17] N. T. Pang, S. Michael, and K. Vipin, " Introduction to Data Mining," *Addison Wesley*, 2009.

[18] J. Han and M. Kamber, "Data Mining : Concepts and Techniques," 2006.

[19] P. Fitsilis, V. Gerogiannis, A. Kameas, and G. Pavlides, "Producing relational database schemata from an object oriented design," in *EUROMICRO 94. System Architecture and Integration. Proceedings of the 20th EUROMICRO Conference.*, 1994, pp. 251–257.

[20] A. Alsaadi, "Checking Data Integrity via the UML Class Diagram," in *Software Engineering Advances, International Conference on*, 2006, p. 37.

[21] G. K. Gupta, *Introduction to data mining with case studies*. PHI Learning Pvt. Ltd., 2011.

[22] Z. Kong, L. Gao, L. Wang, and S. Li, "The Normal Parameter Reduction of Soft Sets and Its Algorithm," *Comput. Math. with Appl.*, vol. 56, pp. 3029–3037, 2008.

[23] J. Han, M. Kamber, and J. Pei, *Data mining: concepts and techniques: concepts and techniques*. Elsevier, 2011.

[24] Tim Dobert (University of Hamburg), "Energy-Efficiency and Performance Trade-offs of Data Reduction Techniques," *Univeristy of hamburg*, 2014. .

[25] D. V Lindley, *Understanding uncertainty*. John Wiley & Sons, 2006.

[26] C. C. Aggarwal, "Managing And Mining Uncertain Data," *Kluwer Acad. Publ.*, 2009.

[27] Z. M. Ma and L. Yan, "Querying Imprecise Data in Sensor Databases," *Mob. Data Manag. 2008. MDM '08. 9th Int. Conf.* , pp. 205–210, 2008.

[28] L. A. Zadeh, "Possibility Theory and Soft Data Analysis," *Math. Front. Soc.*

Policy Sci., pp. 69–129, 1981.

[29] L. A. Zadeh, "Roles of Soft Computing and fuzzy Logic in the Conception, Design and Deployment of Information/Intelligent Systems," *Comput. Intell. Soft Comput. Fuzzy-Neuro Integr. with Appl.*, pp. 1–9, 1998.

[30] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.

[31] Z. Pawlak, "Rough sets: Theoretical aspects of reasoning about data," *Rough Sets Theor. Asp. Reason. about Data*, vol. 9, no. D, 1991.

[32] Z. Pawlak, "Rough sets," *Int. J. Comput. Inf. Sci.*, vol. 11, no. 5, pp. 341–356, 1982.

[33] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy sets Syst.*, vol. 20, no. 1, pp. 87–96, 1986.

[34] M. B. Gorzalczany, "A method of inference in approximate reasoning based on interval-valued fuzzy sets," *Fuzzy sets Syst.*, vol. 21, no. 1, pp. 1–17, 1987.

[35] D. Molodtsov, "Soft Set Theory - First Results," *Comput. Math. with Appl.*, vol. 37, no. 4–5, pp. 19–31, 1999.

[36] R. B. P.K. Maji and A.R. Roy, P. K. Maji, R. Biswas, and A. R. Roy, "Soft Set Theory," *Comput. Math. with Appl.*, vol. 45, pp. 555–562, 2003.

[37] P. K. Maji, A. R. Roy, and R. Biswas, "An Application of Soft Sets in A Decision Making Problem," *Comput. Math. with Appl.*, vol. 44, pp. 1077–1083, 2002.

[38] A. R. Roy, P. K. Maji, A. R. Roy, R. Biswas, Z. Kong, L. Wang, and Z. Wu, "Application of fuzzy soft set in decision making problems based on grey theory," *J. Comput. Appl. Math.*, vol. 44, no. 6, pp. 412–418, 2002.

[39] Z. Kong, L. Wang, and Z. Wu, "Application of fuzzy soft set in decision making problems based on grey theory," *J. Comput. Appl. Math.*, vol. 236, no. 6, pp. 1521–1530, 2011.

[40] Y. Jiang, H. Liu, Y. Tang, and Q. Chen, "Semantic decision making using ontology-based soft sets," *Math. Comput. Model.*, vol. 53, no. 5, pp. 1140–1149, 2011.

[41] Y. Jiang, Y. Tang, and Q. Chen, "An adjustable approach to intuitionistic fuzzy soft sets based decision making," *Appl. Math. Model.*, vol. 35, no. 2, pp. 824–836, 2011.

[42] K. Gong, Z. Xiao, and X. Zhang, "The bijective soft set with its operations," *Comput. Math. with Appl.*, vol. 60, no. 8, pp. 2270–2278, 2010.

[43] K. Gong, P. Wang, and Z. Xiao, "Bijective soft set decision system based parameters reduction under fuzzy environments," *Appl. Math. Model.*, vol. 37,

no. 6, pp. 4474–4485, 2013.

[44] F. Feng, Y. Li, and N. Ça\ugman, "Generalized uni--int decision making schemes based on choice value soft sets," *Eur. J. Oper. Res.*, vol. 220, no. 1, pp. 162–170, 2012.

[45] Z. Zhang and S. Zhang, "A novel approach to multi attribute group decision making based on trapezoidal interval type-2 fuzzy soft sets," *Appl. Math. Model.*, vol. 37, no. 7, pp. 4948–4971, 2013.

[46] F. Feng, Y. B. Jun, and X. Liu, "An Adjustable Approach of Fuzzy Soft Set Based Decision Making," *234*, vol. 1, pp. 10–20, 2010.

[47] A. M. Isa, A. N. M. Rose, and M. M. Deris, "Dominance-Based Soft Set Approach in Decision-Making Analysis," *Adv. Data Min. LNAI ,Springer, Heidelb.* , vol. 1, no. 7120, pp. 299–310, 2011.

[48] Z. Kong, L. Gao, and L. Wang, "Comment on 'A fuzzy soft set theoretic approach to decision making problems,'" *Comput. Math. with Appl.*, vol. 223, no. 2, pp. 540–542, 2009.

[49] A. N. M. Rose, T. Herawan, M. M. Deris, and M. M. D., "A Framework of Decision Making Based on Maximal Supported Sets," *Lect. Notes Comput. Sci. Adv. Neural Networks-ISNN 2010, Springer Berlin Heidelber*, pp. 472–482, 2010.

[50] D. Chen, E. C. C. Tsang, D. S. Yeung, and X. Wang, "The Parameterization Reduction of Soft Sets and its Applications," *Comput. Math. with Appl.* , vol. 49, pp. 757–763, 2005.

[51] X. Ma, H. Qin, N. Sulaiman, T. Herawan, and J. H. Abawajy, "The parameter reduction of the interval-valued fuzzy soft sets and its related algorithms," *Fuzzy Syst. IEEE Trans.*, vol. 22, no. 1, pp. 57–71, 2014.

[52] A. N. M. Rose, H. Hassan, M. I. Awang, N. A. Mahiddin, H. M. H. M. Amin, and M. M. Deris, "Solving Incomplete Datasets in Soft Set Using Supported Sets and Aggregate Values," *Procedia CS*, vol. 5, pp. 354–361, 2010.

[53] H. Qin, X. Ma, T. Herawan, and J. M. Zain, "Data Filling Approach of Soft Sets Under Incomplete Information," *Lect. Notes Comput. Sci. Springer Verlag*, vol. 6592, pp. 302–311, 2011.

[54] T. Deng and X. Wang, "An object-parameter approach to predicting unknown data in incomplete fuzzy soft sets," *Appl. Math. Model.*, vol. 37, no. 6, pp. 4139–4146, 2013.

[55] A. N. M. Rose, M. I. Awang, H. H., and M. M. D., "Comparison of Techniques in Solving Incomplete Datasets in Softset," *Int. J. Database Theory Appl.*, vol. 4, no. 3, pp. 1–12, 2011.

[56] H. M. A. A N M. Rose. Awang M. I., M. D. Mustaffa, "Solving Incomplete Dataset in Soft Set Theory Based on Decimal Approach," *Abstr. has been Accept. Int. Conf. Comput. Inf. Technol.*, 2011.

[57] W. Xu, Z. Xiao, X. Dang, D. Yang, and X. Yang, "Financial ratio selection for business failure prediction using soft set theory," *Knowledge-Based Syst.*, vol. 63, pp. 59–67, 2014.

[58] Z. Xiao, K. Gong, and Y. Zou, "A Combined Forecasting Approach Based On Fuzzy Soft Sets," *J. Comput. Appl. Math.*, vol. 228, no. 1, pp. 326–333, 2009.

[59] S. J. Kalayathankal and G. S. Singh, "A fuzzy soft flood alarm model," *Math. Comput. Simul.*, vol. 80, no. 5, pp. 887–893, 2010.

[60] S. Das and S. Kar, "Group decision making in medical system: An intuitionistic fuzzy soft set approach," *Appl. Soft Comput.*, vol. 24, pp. 196–211, 2014.

[61] T. Herawan and M. M. Deris, "On multi-soft sets construction in information systems," in *Emerging Intelligent Computing Technology and Applications. With Aspects of Artificial Intelligence*, Springer, 2009, pp. 101–110.

[62] D. Pei and D. Miao, "From Soft Sets to Information Systems," *proceeding 2005 IEEE Int. Conf. Granul. Comput. IEEE GrC'05, IEEE Press*, pp. 617–621, 2005.

[63] X. Y. Li, "An Absolute Information Quantity-based Data Making-up Algorithms of Incomplete Information System," *Comput. Eng. Appl.*, vol. 22, pp. 155–197, 2006.

[64] T. Herawan and M. M. Deris, "Soft Set Theoretic Approach for Dimensionality Reduction ," *Database Theory Appl. Commun. Comput. Inf. Sci.*, vol. 64, pp. 171–178, 2009.

[65] T. Herawan and M. M. Deris, "A Soft Set Approach for Association Rules Mining," *Knowl. Based Syst.*, vol. 24, no. 1, pp. 186–195, 2011.

[66] T. Herawan, A. N. M. Rose, and M. M. Deris, "Soft set theoretic approach for discovering attributes dependency in information systems," in *Advances in Neural Networks-ISNN 2010*, Springer, 2010, pp. 596–605.

[67] M. I. Ali, F. Feng, X. Liu, W. K. Min, M. Shabira, and M. Shabir, "On Some New Operations in Soft Set Theory," *Comput. Math. with Appl.*, vol. 57, no. 9, pp. 1547–1553, May 2009.

[68] K. V. Babitha and J. J. Sunil, "Soft set relations and functions," *Comput. Math. with Appl.*, vol. 60, no. 7, pp. 1840–1849, Oct. 2010.

[69] Y. B. Jun and C. H. Park, "Applications of Soft Sets in Ideal Theory of

BCK/BCI-algebras," *Inf. Sci. (Ny).*, vol. 178, no. 11, pp. 2466–2475, 2008.

[70] F. Feng, X. Liu, V. Leoreanu-Fotea, and Y. B. Jun, "Soft sets and soft rough sets," *Inf. Sci. (Ny).*, vol. 181, no. 6, pp. 1125–1137, 2011.

[71] T. Herawan and M. M. Deris, " A Direct Proof of Every Rough Set Is A Soft Set," *Proceeding Third Asia Int. Conf. Model. Simul.*, pp. 119–124, 2009.

[72] M. Shabir, M. I. Ali, and T. Shaheen, "Another approach to soft rough sets," *Knowledge-Based Syst.*, vol. 40, pp. 72–80, 2013.

[73] M. I. Ali, "A note on soft sets, rough soft sets and fuzzy soft sets," *Appl. Soft Comput.*, vol. 11, no. 4, pp. 3329–3332, 2011.

[74] H. Aktas, N. Cagman, H. Aktaş, and N. Çağman, "Soft Sets and Soft Groups," *Inf. Sci. (Ny).*, vol. 177, no. 13, pp. 2726–2735, Jul. 2007.

[75] P. K. Maji, R. Biswas, and A. R. Roy, "Fuzzy Soft Sets," *J. Fuzzy Math.*, vol. 9, no. 3, pp. 589–602, 2001.

[76] Y. Jiang, Y. Tang, Q. Chen, H. Liu, and J. Tang, "Interval-Valued Intuitionistic Fuzzy Soft Sets and Their Properties," *Comput. Math. with Appl.*, vol. 60, no. 3, pp. 906–918, 2010.

[77] X. Yang, T. Y. Lin, J. Yang, Y. Li, and Y. Dongjun, "Combination of Interval-Valued Fuzzy Set and Soft Set," *Comput. Math. with Appl.* , vol. 58, pp. 521–527, 2009.

[78] P. K. Maji, R. Biswas, and A. R. Roy, "Intuitionistic Fuzzy Soft Sets," *J. Fuzzy Math.*, vol. 9, no. 3, pp. 677–692, 2001.

[79] Y. B. Jun, K. J. Lee, and C. H. Park, "Fuzzy soft set theory applied to BCK/BCI-algebras," *Comput. Math. with Appl.*, vol. 59, no. 9, pp. 3180–3192, 2010.

[80] M. Agarwal, K. Biswas, and M. Hanmandlu, "Generalized intuitionistic fuzzy soft sets with applications in decision-making," *Appl. Soft Comput.*, vol. 13, no. 8, pp. 3552–3566, 2013.

[81] P. Majumdar and S. K. Samanta, "Generalised fuzzy soft sets," *Comput. Math. with Appl.*, vol. 59, no. 4, pp. 1425–1432, 2010.

[82] P. K. Maji, A.R.Roy, and R. Biswas., "On Intuitionistic Fuzzy Soft Sets," *J. Fuzzy Math.*, vol. 12, no. 3, pp. 669–683, 2004.

[83] X. Ma, N. Sulaiman, H. Qin, T. Herawan, and J. M. Zain, "A New Efficient Normal Parameter Reduction Algorithm of Soft Sets," *Comput. Math. with Appl.*, vol. 62, pp. 588–598, 2011.

[84] D. Chen, E. C. C. Tsang, D. S. Yeung, and X. Wang, "Some Notes On The Parameterization Reduction of Soft Sets," , *Proceeding Second Int. Conf.*

Mach. Learn. Cybernatics, pp. 1442–1445, 2003.

[85] Y. Zhao, F. Luo, and S. K. M. Wong, "A General Definition of an Attribute Reduct," *Lect. Notes Artif. Intell.*, vol. 4481, pp. 101–108, 2007.

[86] A. N. M. Rose, A. M. I., T. Herawan, and M. D. Mustaffa, "Hybrid Reduction in Soft Set Decision Making," *ICIC 2011*, p. in Press, 2010.

[87] T. Pang-Ning, M. Steinbach, V. Kumar, and others, "Introduction to data mining," *Libr. Congr.*, p. 74, 2006.

[88] X. Xi, E. Keogh, C. Shelton, L. Wei, and C. A. Ratanamahatana, "Fast time series classification using numerosity reduction," in *Proceedings of the 23rd international conference on Machine learning*, 2006, pp. 1033–1040.

[89] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," *Commun. ACM*, vol. 30, no. 6, pp. 520–540, 1987.

[90] J. G. Cleary and I. H. Witten, "Data compression using adaptive coding and partial string matching," *Commun. IEEE Trans.*, vol. 32, no. 4, pp. 396–402, 1984.

[91] A. Moffat, "Implementing the PPM data compression scheme," *Commun. IEEE Trans.*, vol. 38, no. 11, pp. 1917–1921, 1990.

[92] S. Rosaria, Iris, Adae, A. Hart, and M. Berthold, "Seven Techniques for Dimensionality Reduction," 2014.

[93] I. T. Jolliffe, "Introduction," *Princ. Compon. Anal.*, pp. 1–9, 2002.

[94] J. Borenstein and Y. Koren, "The vector field histogram-fast obstacle avoidance for mobile robots," *Robot. Autom. IEEE Trans.*, vol. 7, no. 3, pp. 278–288, 1991.

[95] P. Xanthopoulos, P. M. Pardalos, and T. B. Trafalis, "Linear discriminant analysis," in *Robust Data Mining*, Springer, 2013, pp. 27–33.

[96] Y. Qian, J. Liang, W. Pedrycz, and C. Dang, "Positive approximation: an accelerator for attribute reduction in rough set theory," *Artif. Intell.*, vol. 174, no. 9, pp. 597–618, 2010.

[97] Y. Zhao, F. Luo, S. K. M. Wong, and (2007) 101-108 Y.Y. Yao, "Lecture Notes in Artificial Intelligence," *A Gen. Defin. an Attrib. Reduct*, vol. 4481, pp. 101–108, 2007.

[98] Z. Pawlak, A. Skowron, and A. Skowron 177 (1) (2007) 3–27, " Rudiments of Rough Sets," *Inf. Sci. (Ny).*, vol. 177(1), no. 1, pp. 3–27, 2007.

[99] T. Herawan, M. M. Deris, A. N. M. Rose, and M. M. Deris, "Soft Set Theoretic Approach for Dimensionality Reduction ," *Database Theory Appl. Commun. Comput. Inf. Sci.*, vol. 64, pp. 171–178, 2009.

[100] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping multidimensional data*, Springer, 2006, pp. 25–71.

[101] Y. Kudo and T. Murai, "Indiscernibility relations by interrelationships between attributes in rough set data analysis," in *Granular Computing (GrC), 2012 IEEE International Conference on*, 2012, pp. 220–225.

[102] R. Slowinski and J. Stefanowski, "Handling various types of uncertainty in the rough set approach," in *Rough Sets, Fuzzy Sets and Knowledge Discovery*, Springer, 1994, pp. 366–376.

[103] J. Stefanowski and A. Tsoukias, "Incomplete information tables and rough classification," *Comput. Intell.*, vol. 17, no. 3, pp. 545–566, 2001.

[104] R. Susmaga, "Tree-like parallelization of reduct and construct computation," in *Rough Sets and Current Trends in Computing*, Springer, 2004, pp. 455–464.

[105] Y. Zhao, Y. Yao, and F. Luo, "Data analysis based on discernibility and indiscernibility," *Inf. Sci. (Ny).*, vol. 177, no. 22, pp. 4959–4976, 2007.

[106] M. Lichman, "{UCI} Machine Learning Repository," *University of California, Irvine, School of Information and Computer Sciences*, 2013. [Online]. Available: http://archive.ics.uci.edu/ml. [Accessed: 01-May-2016].

[107] Mobile88.com, "Category Archives : Comparison," 2000. .

[108] P. C. and A. Silva, "Using Data Mining to Predict Secondary School Student Performance," in *Proceedings of 5th FUture BUsiness TEChnology Conference (FUBUTEC 2008)*, pp. 5–12.

VITAE

The author was born in November, 1962 in Kuala Terengganu. He has started his early education at Sultan Ismail School. As for the secondary education, he was then accepted and enrolled at MARA Junior Science School, Pengkalan Chepa in 1975. After MCE, he was sponsored by MARA and sent to United States of America for his tertiary studies in 1980. He studied Computer Science at Ohio University, Athens, USA. He then graduated and came back to Malaysia in 1984. He then joined Shell Sarawak as Analyst Programmer for a year, before returning back for good to his home state, Terengganu. He then applied for a post at KUSZA and was subsequently accepted as a lecturer at KUSZA in 1987. In 1995, he was then sent abroad to United Kingdom to study Data Communication System at Master level at Brunel University. When KUSZA was upgraded to become a public university, UniSZA, he then enrolled for his PhD studies at UTHM on his own iniative. It was during this PhD stint, he has presented 3 papers at international conferences and written 7 publications related to his PhD work.

THE LONGEST YARD

KUSZA has been transformed and upgraded to the status of the university in 2006. And it was initially named as Univeristi Darul Iman or better known as UDM. The changes in the status quo of the lecturers from KUSZA to the lecturers of UDM, has ushered in a new era in the working environment at UDM. Before this, the lecturers of KUSZA, with the scheme of DG41 were just required to teach besides doing administrative chores if requested. Lecturers of KUSZA, were only required to further their studies to at least up to the Master level. Publishing has never been the main criteria in the career.

When KUSZA was upgraded to UDM by 2006, it was only then when me and my good friend, Mr Isa decided to pursue our education in computer science at the PhD level. We started by registering at UMT, with Prof Dr Mustaffa as our supervisor. We were accepted but UMT has assigned us Prof Dr Yazid as our supervisors, because Prof. Dr Mustaffa has decided to move to UTHM to seek greener pasture. We then decided, to further our education at PhD level at UTHM. In 2007, we were accepted as PhD students at UTHM. When we registered as PhD students, the faculty only had two lectures with PhD at that time and the office was in shoplot building blocks. But now, the faculty has a lot highly caliber lecturers and located in a beautiful building. Our journey started officially in 2007, in the pursuit of excellence in knowledge.

The first year was smooth sailing, we were required to attend compulsory courses and there were lots of travelling between Kuala Terengganu and Parit Raja. Later on in the following years, our group grew from two to five(5). We were joined by En. Ahmad Shukri, En. Rabie and En. Zailani, lecturers from UMT, who have registered as PhD students to Prof. Dr Mustaffa. That means, we have more companions when travelling and as the adage says, the more the merrier. The following years until 2016, would just be routine visits to UTHM. It will include progress report visits and subject registration.

I was lucky, as I have never enjoyed driving for long distance, but my good friend En Isa did not enjoyed travelling being a co-pilot, so he was always behind the wheels, as the driver of the trip. En Isa has always been the driver for the numerous

trips between Kuala Terengganu and Parit Raja from 2007 to 2016. We have lost track in counting the number of trips taken.

The trip to Parit Raja, frankly speaking is an arduous journey and its divided into three parts. The start of the journey, from Kuala Terengganu and Gambang. The middle part, from Gambang to Segamat. The final phase, from Segamat to Parit Raja. Whenever I was required to make a visit to UTHM, I will always be dreading of the long journey. The start of the journey, has always started as a somber and dull start, because we knew it is going to be long trip. We have always tried to initiate our discussion on interesting topics, be it on political issues or campus politicking or anything else that is interesting. This will help us pass the time during the first two phases, without laboring on constant checking on the time to arrive at Parit Raja or where we were. Segamat has always been the first pit stop, where the car will be refueled and we replenish out driving strength with some *Roti Canai* and *Teh Tarik*. After that, it will be another one and the half hour of driving, until we finally reached Parit Raja. And it will be the routine checking-in, resting and obligatory prayers. Our next day, will start as early as possible, be it attending presentation, paying fees or meeting our supervisor. Then, when everything is completed, we head back to Kuala Terengganu as soon as possible. Driving home and reaching home, is always a good feeling.

The group of five(5) enjoyed the companionship of each other during the travelling and doing assignments together. This mutual bonding has taken the stress and pressure out of us during the travelling and doing assignments. But soon our group became smaller again when Dr Zailani completed his PhD in 2012, Dr Ahmad Shukri completed his PhD in 2013 and Dr Rabei completed his PhD in 2014. And we were back to two(2). When one by one, the UMT lecturers completed their PhD, the pressure to finish was starting to mount onto us. This was where Mr Isa took his leave for one semester, off to UTHM to focus on the simulation of his research. While I was tied-up with some management jobs, as the Head of Department (2012-2014) and then as the Deputy Dean (2015 – present). When En Isa left for Parit Raja, I started to feel as if I was in the pressure cooker, and that was when I started writing codes to do the simulation of my reduction techniques. It took me almost six(6) months to finish the coding and another 4 months of testing. By early end of january 2015, I have finished the all the experimentations and data collection.

Then I started to polish my thesis with the collected data and finished the writing of thesis at the end of 2015. I was lucky, because I have written some publications and some of the chapters have already been written starting from 2012. Therefore it was just a matter of integrating and polishing the thesis to the expected level. Alhamdulillah, I have finished my first draft by November 2015 and I submitted my thesis to the faculty in December 2015.

Looking back through the years, one thing that I regretted most is the lack of discipline and efficient time management from myself. Reflecting on what I have managed to achieve in 2015, juggling in between office responsibility, a father and husband, and community engagement, was really an uphill task but nevertheless I did it. If only had I done and focused earlier, I would have finished it earlier. This is because I already had all the chapters 50-70% ready by 2012. It was just the final push that was lacking. But nevertheless, Alhamdulillah. The system that has been in place by the faculty (FSKTM) actually helped me a lot. The progress presentation, the thesis preparation has helped me a a lot when I was in the final phase. The truth is, the presentation for viva, was made-up from the materials that I already had prepared for the various progress reports and it was just a matter of polishing it.

As for the numerous trips that have been made to Parit Raja, the truth was trip was not never an enjoyable ones. Starting from Ajil, the palm oil trees were lining on the roadside waving its leaves until we reached Parit Raja. It was common eyesight all the way to Parit Raja or vice versa. The trips had never given us the chance to mingle with community and society during our journey. That was why, if we were to be asked whom we met and made friendship during trip, the answer is never an easy one. The trip will more colorful and beautiful, if we were able to made more friends and be of help to any part of the society that we met during the journey. It was always a trip, like a business trip, looking forward to complete the deal as soon as possible and head home straight away. To the readers of this passage, first and foremost wherever the destination is, it must be a journey, that we dream-off, a journey that we look forward to and a journey that we will enjoy it. Only then, the physical journey will be pleasant and full of beautiful memories. The physical trip must something we are comfortable with and be pleasant with, only then we will cherish the journey.

My PhD journey is an event-oriented journey. Most of the actions from me were triggered by events. If there were no events, I don't know how I can finish the

journey. The events that occurred, helped push the pace. Everything was unplanned. I have never dreamt to pursue my education up to PhD level. I have always believe, PhD is something very big, something that will change the world. When I started to register as a PhD students, it was because the circumstances in the working environment has changed. Prof Dr Mustaffa has be very insistent years before, that I continue my studies in until PhD level. But I have always brushed it aside. If I were to listen to Prof Dr Mustaffa earlier, I might have done my PhD studies in UMT and finish it there. But fate has a different storyline onto me. Fate has decided that UTHM is where I learnt the trades of being a better lecturer. Even when I have registered as a PhD student at UTHM, the motivation and desire to finish, were not there. If not because of the system in UTHM in monitoring its graduate students and my graduating friends from UMT, I might have not be awaken from my sleep. That is why I have to concur to the adage that "Failure to plan, is planning to fail". The keyword here is plan and stick to the plan.

The day when I passed my PhD defense, four people cried most. My mother, mother-in-law, wife and daughter cried with tears of joy. My mother cried because she said if my father were still alive, he would have been very proud of me. My mother in-law cried, because to her this is the epitome of my pursuit in seeking the knowledge and again my very supportive father in-law was also no longer in this world to witness my success. My daughter cried like a baby, as Allah has finally answered her unwavering dua's in seeing his father passing his PhD defense. My mothers and children were the silent supporters and who have said their prayers vigilantly into the night, asking ALLAH to make it easy and smooth sailing for their loved one, me.

In soldiering-on in the quest for a degree in PhD, to have companions that are truly supportive is important. They will be the motivator and the batteries to charge-up for the sometimes ,wilting desire to finish.

My other half, may ALLAH reward her for the true love, support and prayer that she has steadily been giving. She has been very persistent in supporting me morally, just to see me through. For all the preparations, she has provided during the numerous trips, I am very thankful and ALLAH is the BEST to reward her. Thank you very much.

Prof Dr Mustaffa, I will never forget him. He is a true friend that will remain as my one of best friends forever. En Isa, my friend who I normally travel with, will

also be a friend that I am very indebted to. I will never be able to finish this, if I were doing it alone. His presence and companionship is the mutual symbiosis friendship that helps to motivate and provide strength to finish. Dr Tutut, who taught me soft set theory and tutored me on writing publications, is instrumental in showing me the light in the early stage of my PhD studies when I was still in the dark of what tbe researched on.

The longest yard is comprised of 18 semesters or 9 years of perseverance, though sometimes a bit of lackluster. I am thankful and grateful to ALLAH, as I finally have reached the finishing line. There lots of people involved, it will be a very long list to name them all. My inner PhD journey will linger on forever in my memories. It will remain as a beautiful part of my life. But I am signing-off here by saying a heartfelt thank you to my mothers, wife, children, friends and the faculties, FIK and FSKTM for the understanding and being there with me

Thank you all.