CONTENTS

Acknowledge	ments	d
Abstract in Th	nai	e
Abstract in En	nglish	g
List of Tables		k
List of Figure	s กลยนด์	1
Statement of	Originality in English	n
Statement of	Originality in Thai	0
Chapter 1 Intr	roduction	1
1.1	Background	1
	1.1.1 (k, e)-anonymous model	3
	1.1.2 An incremental privacy breach	5
1.2	Purposes of the Study	6
1.3	Educational Advantages	6
1.4	Research Design	7
1.5	Scope	7
1.6	Method AI UNIVERSIT	7
1.7	Summary	8
Chapter 2 Rel	ated Work	9
2.1	Privacy Breach Problems	9
2.2	Principle privacy preservation models	11
2.3 A	Incremental privacy preservation problems and the solutions	21
2.4	Summary	28
Chapter 3 Bas	sic Definition, Problem statement, and Solution	29
3.1	Basic Definition	29
3.2	Problem Definition.	36
3.3	Existing Issues	36
	3.3.1 Efficiency issue	36

	3.3.2 Effectiveness issue	39
3.4	Summary	40
Chapter 4 Observations and Proposed Algorithm		41
4.1 C	Observation of Incremental Privacy Breach Scenarios	41
4.2 Proposed Algorithm		50
4.3 S	ummary	55
Chapter 5 Experiment		56
5.1	Experiment Preparation	56
5.2	Effects of the value of k	57
5.3	Effects of the value of e	59
5.4	Effects of the Number of Partitions	60
5.5	Effects of $ \Delta D $	62
5.6	Summary	63
Chapter 6 Conclusion		64
6.1	Research Contribution	65
6.2	Future work	65
Reference		67
Curriculum vitae		72
	MAI HAINERS'	
	UNIVE	

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม Copyright[©] by Chiang Mai University All rights reserved

LIST OF TABLES

Table 1.1 The released dataset	
Table 1.2 The public dataset	2
Table 2.1 Linkable dataset	12
Table 2.2 Published dataset	12
Table 2.3 2-Anonymity dataset	13
Table 2.4 3-diverse dataset after generalization	14
Table 2.5 <i>l-diversity</i> skewed sensitive attribute values	15
Table 2.6 (0.5, 3)-diversity dataset after generalization	15
Table 2.7 An example dataset	15
Table 2.8 An example (k, e)-anonymous dataset after permutation	17
Table 2.9 An Example dataset to demonstrate minimum summation of	
error algorithm	20
Table 2.10An example dataset demonstrates a border unchanged impact	
of adding new tuple in an in-range case	22
Table 2.11An example dataset demonstrates a borders changed impact	
of adding new tuple in in-range case	23
Table 2.12An example dataset demonstrates a previous marge impact	
of adding new tuple in border case	23
Table 2.13An example dataset demonstrates a next marge impact	
of adding new tuple in border case	24
Table 5.1 The characteristics of proposed algorithm and	
naïve re-applying algorithm	57

LIST OF FIGURES

Figure 1.1 An Example Dataset for (k, e) -anonymous Model		
(k = 3, and e = 2000)	4	
Figure 1.2 An Incremental Privacy Breach Example	5	
Figure 2.1 Minimum Summation of Error Algorithm	18	
Figure 2.2 Illustration of the minimum summation of error algorithm		
rigure 2.2 mustration of the minimum summation of error argorithm	1)	
Figure 3.1 The illustration for the definition of dataset and the definition of		
(k, e)-Anonymous partition dataset where $k = 3$ and $e = 2,000$.	30	
Figure 3.2 Two versions of partitioned dataset of dataset D_0	31	
Figure 3.3 The illustration of the $(3, 2,000)$ -Anonymous model	32	
Figure 3.4 The shuffled dataset at time 0 and 1 that are satisfied		
(3, 2000)-Anonymous which minimal summation error	33	
Figure 3.5 Difference privacy breach from $p'_a[D_0]$ to $p'_c[D_1]$, p_{ac}	34	
Figure 3.6 Difference privacy breach from $p'_c [D_1]$ to $p'_a[D_0]$, p_{ca}	35	
Figure 3.7 Intersection privacy breach between $p'_a[D_0]$ and $p'_c[D_1]$, p^{\cap}_{ac}	35	
Figure 3.8 Brute force of incremental process	37	
Figure 3.9 The Naïve Re-Applying Algorithm	38	
Figure 3.10 3 versions of the solution dataset from naïve re-applying algorithm		
by considering one version of previously released dataset	39	
Figure 3.11 Results of an intersection and a difference between a partition		
in D_0 and a partition in D_2	40	
Figure 4.1 Overlapping between p_a and p_b	42	
Figure 4.2 An example of two released datasets D_0 and D_1	43	
Figure 4.3 An example situation in which the top part of $p_b[D_1][S]$ is covered		
by some part of $p_a[D_0][S]$	44	
Figure 4.4 An example situation in which the bottom part of $p_b[D_1][S]$ is		
covered by some part of $p_a[D_0][S]$	44	

Figure 4.5 An example situation in which the $p_b[D_1][S]$ shrinks from $p_a[D_0][S]$	44
Figure 4.6 Fully covered situation	46
Figure 4.7 An example situation when the $p_b[D_1][S]$ is fully covered by $p_a[D_0][S]$	47
Figure 4.8 An example situation in which the $p_c[D_n][S]$ is separated into two parts	; 49
Figure 4.9 The examples of 2 situations on data scenarios when the $p_c[D_n][S]$ is	
separated into two parts	49
Figure 4.10 The Proposed Algorithm	51
Figure 4.11 Illustration of proposed algorithm	52
Figure 4.12 An example of the two datasets, to illustrate the algorithm	54
Figure 5.1 Execution time of the proposed algorithm and the naïve re-applying	
when the k value is varied	58
Figure 5.2 Proposed algorithm's execution time and the percentage of discard	
when the value of k is varied	58
Figure 5.3 Execution time of the proposed algorithm and the naive re-applying	
algorithm when the <i>e</i> value is varied	59
Figure 5.4 Proposed algorithm's execution time and percentage of discard when	
the <i>e</i> value is varied	60
Figure 5.5 Execution time of the proposed algorithm and the naive re-applying	
algorithm when the number of partitions is varied	61
Figure 5.6 Proposed algorithm's execution time and the percentage of discard	
when the number of partitions is varied	61
Figure 5.7 Execution time of the proposed algorithm and the naive re-applying	
algorithm when the size of the incremental part (%) value is varied	62
Figure 5.8 Proposed algorithm's execution time and the percentage of discard	
when the size of the incremental part (%) value is varied	62
Figure 6.1 A Deleting Privacy Breach Example	66

STATEMENT OF ORIGNALITY

A proposed algorithm for preventing an incremental privacy breach on (k, e)-Anonymous with all versions of dataset, this algorithm needs to consider only one previous version of the dataset. At the same time, the privacy of all of the released datasets as well as the optimal solution can be always guaranteed.



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่ Copyright[©] by Chiang Mai University All rights reserved

ข้อความแห่งการริเริ่ม

ขั้นตอนวิธีสำหรับปกป้องการละเมิดความเป็นส่วนตัวของข้อมูลในต้นแบบ (k, e)-Anonymous โดยข้อมูลอยู่ในสถานเพิ่มขึ้นตลอดเวลา ขั้นตอนวิธีที่ นำเสนอนี้พิจารณาเพียงข้อมูลฉบับก่อนหน้าเพียงหนึ่งฉบับเท่านั้นและยัง สามารถให้ผลลัพธ์ที่เหมาะสมที่สุดในทุกช่วงเวลา



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่ Copyright[©] by Chiang Mai University All rights reserved