

RIGHTS ASSESSMENT FOR DISCRETE DIGITAL DATA

A Thesis

Submitted to the Faculty

of

Purdue University

by

Radu Sion

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2004

To Amy Brown.

## ACKNOWLEDGMENTS

I want to start by thanking both my advisers. *Mike* was always available, a good friend, with amazing insights and advice (often on personal topics) for a green bean as myself. Be it 2 am or 2 pm I knew he was going to be there with a good word or a sharp research remark. Through his ways he made me experience and understand respect and restraint. I owe much of the enjoyment I get out of research writing to *Sunil* and his many great eye-opening perspectives, including on social dimensions of academia and the much debated peer-review process. It was invaluable to experience his entire tenure process parallel my graduate studies and succeed. I was extremely lucky to have the academic parents I had. I never had to worry about funding, going to (often expensive) conferences and organizing my own schedule. Thank you. Growing up now, “away from home”, I will miss you.

I want to thank *Amy Brown*. For being. *Bogdan Carbunar*, my long-term apartment mate and good friend. *Melissa Dyehouse* for allowing me to teach her skiing and putting up with me for weeks through Europe. *Daniel Aliaga* for the racquetball, mountain biking and car races. *Nicoleta Neagu*, for being there. *Ion Constantinescu*, for shelter, skiing and personality. *Mirela Mustata* for driving through rain, snow and mud for me, in times when I needed a friend. *Jens Palsberg* for the great comedy club experiences and friendship. *Irina Athanasiu* for Pizza Hut. *Murat Kantargioglu*, for Tomatina. *Ladislau Boloni* for latex (the digital kind). *Ton Kalker* for the insights into issues of high dimensionality of usability spaces and friendship. *Dr. Gorman*, clearly the most important person in our department, for great conversations during empty campus holiday weeks, for advice, help and insights into the intricacies of academic bureaucracy. *Renate Mallus* for dancing and smiling. The friends from the Vienna Coffee-shop including *Olga*, *Shared*, *Jacques*, *Mercan* and *Umut*. For games, fun and a great atmosphere.

I would also like to thank my teachers at Purdue. *Doug Comer* for his PhD topic generator and his great TCP/IP class. *Susanne Hambrusch* for the great advice on the association between flower-power and academics. The members of my PhD committee (*Jens Palsberg*, *Elisa Bertino* and *Samuel Wagstaff*) for useful advice and insights. *Debbie Frantz* and the entire staff at CERIAS for great chats and help with speedy travel reimbursements. *Dan Marinescu* for guidance and help in my first year at Purdue. *Eugene Spafford* (*Spaf*) for a great research environment at CERIAS.

I thank my parents for coming up with the idea of having me, my father (1935-2003) for teaching me the power of work and persistence. There are creations out there that become relevant after they cease to exist, often leaving a bitter taste of things that could've been on a backdrop of things that actually were. My mother, for too many things, including reading the Dialogues of Socrates to a three and a half year old version of me as bed-time stories.

I want to thank all of the above and those many others that I ought to but probably forgot due to being stressed and growing older. My graduate student years were the most amazing and fun years in my life so far and you all were a big part of this experience. Thank you.

## TABLE OF CONTENTS

	Page
LIST OF FIGURES . . . . .	x
ABSTRACT . . . . .	xvii
1 Introduction . . . . .	1
1.1 Deployment Scenario . . . . .	2
1.1.1 Rights Protection through Assessment . . . . .	3
1.1.2 Information Hiding vs. Newspaper Digests . . . . .	5
1.2 Watermarking vs. Watermarking . . . . .	7
1.3 Summary of Contributions . . . . .	8
1.3.1 Model . . . . .	9
1.3.2 Numeric Relational Data . . . . .	10
1.3.3 Categorical Data . . . . .	13
1.3.4 Sensor Streams . . . . .	15
1.3.5 Abstract Structures . . . . .	18
1.3.6 Limits . . . . .	21
2 Model of Watermarking (Part One) . . . . .	24
2.1 Model and Definitions. . . . .	25
2.2 Consumer Driven Watermarking . . . . .	30
2.3 Steganography and Watermarking . . . . .	33
2.4 Notations and Primitives . . . . .	34
2.5 Conclusions . . . . .	35
3 Relational Data with Numeric Types . . . . .	36
3.1 Introduction . . . . .	36
3.2 Challenges . . . . .	38
3.2.1 Available Bandwidth . . . . .	40

	Page
3.2.2 Model of the Adversary . . . . .	42
3.3 Simplified Problem: Numeric Collections . . . . .	43
3.3.1 Solution Summary . . . . .	44
3.3.2 Selecting Subsets . . . . .	45
3.3.3 Amplifying Watermark Power . . . . .	46
3.3.4 Resilience Analysis . . . . .	53
3.4 The Relational Database . . . . .	56
3.4.1 Algorithm . . . . .	57
3.4.2 Embedding Optimizations . . . . .	58
3.4.3 On-the-Fly Update-Ability . . . . .	61
3.5 Discussion . . . . .	63
3.5.1 Detection Maps . . . . .	63
3.5.2 Subset Markers . . . . .	64
3.5.3 Primary Key Dependence . . . . .	65
3.6 Experimental Results . . . . .	65
3.6.1 Implementation: wmdb.* . . . . .	66
3.6.2 Experiments . . . . .	66
3.6.3 Scenario: The Wal-Mart Sales Database . . . . .	76
3.7 Related Work . . . . .	78
3.8 Conclusions. Future Research. . . . .	81
4 Relational Data with Categorical Types . . . . .	83
4.1 Introduction . . . . .	83
4.2 Model . . . . .	84
4.2.1 Notation . . . . .	84
4.2.2 The Adversary . . . . .	85
4.3 Categorical Data . . . . .	86
4.3.1 Challenges . . . . .	87
4.3.2 Bandwidth Channels . . . . .	88

	Page
4.3.3 Algorithms . . . . .	89
4.3.4 Multiple Attribute Embeddings . . . . .	94
4.4 Discussion . . . . .	97
4.4.1 Correlation Attacks . . . . .	97
4.4.2 On-the-fly Quality Assessment . . . . .	103
4.4.3 Vertical Partitioning Revisited . . . . .	105
4.4.4 False Positives and Vulnerability to Attacks . . . . .	107
4.4.5 Bijective Attribute Re-mapping . . . . .	109
4.4.6 Data Addition . . . . .	110
4.4.7 Minimizing Alteration Distance . . . . .	112
4.4.8 Blindness, Incremental Updates and Streams . . . . .	113
4.4.9 Multi-Layer Self-Reinforcing Marks . . . . .	114
4.5 Experiments . . . . .	116
4.6 Conclusions . . . . .	119
5 Discrete Streaming Data. Sensor Streams. . . . .	120
5.1 Introduction . . . . .	121
5.2 Challenges . . . . .	123
5.2.1 The Adversary . . . . .	123
5.2.2 Model . . . . .	125
5.2.3 Related Work . . . . .	128
5.3 An Initial Solution . . . . .	130
5.3.1 Overview . . . . .	130
5.3.2 Embedding . . . . .	131
5.3.3 Detection . . . . .	133
5.4 Improvements . . . . .	136
5.4.1 Defeating Correlation Detection . . . . .	136
5.4.2 Repeating Labels . . . . .	139
5.4.3 Reconstructing Labels . . . . .	139

	Page
5.4.4	Hysteresis . . . . . 140
5.4.5	Defeating Bias Detection . . . . . 141
5.4.6	On-the-Fly Quality Assessment . . . . . 144
5.4.7	Finite Window . . . . . 145
5.4.8	Offline Detection . . . . . 146
5.4.9	Labeling Made Safer . . . . . 146
5.4.10	Summarization Revisited . . . . . 147
5.5	Analysis . . . . . 148
5.6	Experimental Results . . . . . 153
5.6.1	Random Alterations . . . . . 155
5.6.2	Sampling and Summarization . . . . . 157
5.6.3	Segmentation. Combinations . . . . . 158
5.6.4	Overhead and Impact on Data Quality . . . . . 159
5.7	Conclusions . . . . . 161
6	Semi-structured Aggregates . . . . . 163
6.1	Introduction . . . . . 163
6.2	Challenges . . . . . 164
6.2.1	The Adversary . . . . . 166
6.3	A Solution . . . . . 167
6.3.1	Tolerant Canonical Labeling . . . . . 167
6.3.2	Tolerant Content Summaries . . . . . 174
6.3.3	Algorithm . . . . . 175
6.3.4	Discussion . . . . . 178
6.4	Implementation and Experiments . . . . . 179
6.4.1	The wmx.* Package . . . . . 179
6.4.2	Experiments . . . . . 180
6.5	Conclusions . . . . . 182
7	Model of Watermarking (Part Two) . . . . . 183

	Page
7.1 First Principle of Watermarking . . . . .	183
7.2 Challenge of Watermarking . . . . .	184
7.3 Limits . . . . .	185
7.3.1 Introduction . . . . .	185
7.3.2 A Sample Watermarking Algorithm . . . . .	187
7.3.3 Analysis . . . . .	189
7.3.4 Discussion . . . . .	193
7.4 Discussion . . . . .	195
7.4.1 Oracle Attacks . . . . .	195
7.4.2 Persuasiveness and Watermark Length. Distance Metrics . . .	197
7.4.3 Note on Collusions . . . . .	199
7.5 Conclusions . . . . .	200
8 The Future . . . . .	203
LIST OF REFERENCES . . . . .	204
A Appendix . . . . .	209
VITA . . . . .	210

## LIST OF FIGURES

Figure	Page
1.1 Introduction: (a) <i>Digital Watermarking</i> conceals an indelible “rights witness” (“rights signature”, watermark) within the digital Work to be protected. (b) In court, a detection process is deployed to prove the existence of this “witness” beyond reasonable doubt (confidence level) and thus assess ownership. . . . .	2
1.2 Introduction: Rights Assessment is useful when valuable content is to be sold/outsourced to potentially un-trusted parties, even if rightfully licensed. . . . .	3
1.3 Introduction: A scenario where resilient information hiding for fingerprinting might reveal which secret agent leaked secret documents to Lex Luthor. . . . .	5
1.4 Introduction: Information Hiding classification according to Petitcolas et al [4] . . . . .	6
1.5 Introduction: Relational Data with Numeric Types – (a) The <b>wmdb.*</b> package. (b) Random attack (non-zero average) on a normally distributed data set. (c) Impact of classification preservation on the available watermarking bandwidth. . . . .	12
1.6 Introduction: Relational Data with Categorical Types – (a) More available bandwidth (decreasing $e$ ) results in a higher attack resilience. (b) The watermark degrades almost linearly with increasing data loss. . . . .	15
1.7 Introduction: Discrete Streaming Data – (a) Watermark survival to epsilon-attacks. (b) Watermark survival to combined sampling and summarization. . . . .	18
1.8 Introduction: Semi-structured Aggregates – Averaged watermark loss over 10 runs of an 8 bit watermark embedded into an arbitrary 32 node graph with 64 edges. Surgery attacks are applied randomly (node removals 60%, link addition 20%, link removal 20%). The labeling scheme was trained for 3 surgeries. . . . .	20

Figure	Page
1.9 Introduction: Model of Watermarking – (a) No matter how sophisticated the watermarking method, there exists a random attack with a success probability of 33% and above (although we might not know what the attack is). It can be seen that a more court convincing $\epsilon_w$ value yields an even higher upper bound on attack success probability (2D cut through (b)). (b) The 3D evolution of the probability of a successful attack. . . . .	22
2.1 Model of Watermarking: (a) A 2-dimensional view of a usability space. A point uniquely identifies a Work in $\mathbb{D}$ (e.g., coordinates in this space are DCT coefficients considered for watermark embedding). Watermarking results in a point $O'$ , a “watermarked” version of $O$ and is naturally represented as a transform in the usability space. (b) Usability vicinities of a certain Work $O \in \mathbb{D}$ for a given marking algorithm. $U_{data}$ is defined by the actual data type of the usability metrics. $U_{max}$ is the maximal allowable usability vicinity with respect to the associated usability domain(s) (e.g., Human Visual System). The results ( $U_{alg}$ ) of a valid marking algorithm with respect to a given Work and all other possible inputs should be contained within the maximal allowable usability vicinity ( $U_{max}$ ) of the Work. $U_{wm}$ is determined by $\epsilon_w$ . . . . .	27
2.2 Model of Watermarking: In actuality, (symmetric) watermarking is based on the use of a common secret (key) $k$ shared between the encoding and detection (e.g., in court) phases. . . . .	29
2.3 Model of Watermarking: In consumer-driven watermarking a set of data constraints are continuously evaluated in the encoding process to ensure quality of the result. . . . .	32
3.1 Relational Data: Rights assessment is important when valuable data is outsourced to a third party. . . . .	37
3.2 Relational Data with Numeric Types: Primitive Mark Power Amplification. Subset selection after sorting on keyed hash of the most significant bits (MSB) of the normalized data items. This enables recovery after various attacks, including re-shuffling/sorting and linear changes. The secrecy of the subsets to which the weak(er) encoding is applied provides a resilience amplification effect. . . . .	47
3.3 Relational Data with Numeric Types: Single Bit Encoding Algorithm (illustrative overview). . . . .	50

Figure	Page
3.4 Relational Data with Numeric Types: Distribution of item set $S_i$ . Encoding of the watermark bit relies on altering the size of the “positive violators” set, $v_c(S_i)$ . . . . .	51
3.5 Relational Data with Numeric Types: (a) Different error correcting (wmdb. sys. RedundancyCoder) plugins can be added/removed at runtime in order to provide an increased level of resilience for the original watermark to be embedded. (b) Example of majority voting over three recovered watermark copies for a 6 bit sized original watermark. . . . .	53
3.6 Relational Data with Numeric Types: Watermark Embedding Algorithm (version using subset markers and detection maps shown). . . .	59
3.7 Relational Data with Numeric Types: Watermark Detection Algorithm (version using subset markers and detection maps shown). . . .	60
3.8 Relational Data with Numeric Types: The <b>wmdb.*</b> package. Application runtime snapshot. . . . .	67
3.9 Relational Data with Numeric Types: The <b>wmdb.*</b> package. Overview.	68
3.10 Relational Data with Numeric Types: Resilience to data surgeries (a) uniform distribution, (b) normal distribution, (c) single subset (1-bit) encoding . . . . .	70
3.11 Relational Data with Numeric Types: Epsilon-attack (zero-average) on normally distributed data. . . . .	72
3.12 Relational Data with Numeric Types: (a) Epsilon-attack (non-zero average) on a normally distributed data set. (b) Impact of guaranteeing a Maximum Allowable Absolute Change on the available watermarking bandwidth. . . . .	74
3.13 Relational Data with Numeric Types: Impact of a classification preservation on the available watermarking bandwidth. . . . .	75
4.1 Relational Data with Categorical Types: (a) Embedding Algorithm (b) Alternative using embedding map (bit size adjustments omitted) .	90
4.2 Relational Data with Categorical Types: Overview of multi-bit watermark encoding. . . . .	91
4.3 Relational Data with Categorical Types: (a) Decoding Algorithm (b) Alternative using embedding map . . . . .	94
4.4 Relational Data with Categorical Types: Defeating vertical partitioning.	95

Figure	Page
4.5 Relational Data with Categorical Types: Handling multiple marks interference. . . . .	98
4.6 Relational Data with Categorical Types: Defeating correlation attacks.	100
4.7 Relational Data with Categorical Types: Defeating correlation attacks revisited (multiple embeddings). . . . .	102
4.8 Relational Data with Categorical Types: Data quality is continuously evaluated. A backtrack log aids undo operations in cases where the watermark embedding would violate quality constraints (see also Chapter 3. . . . .	104
4.9 Relational Data with Categorical Types: Handling extreme multi-set partitioning. . . . .	106
4.10 Relational Data with Categorical Types: Handling attribute remapping.	111
4.11 Relational Data with Categorical Types: Handling an informed Mallory.	115
4.12 Relational Data with Categorical Types: (a) The watermark degrades gracefully with increasing attack size ( $e = 65$ ). (b) More available bandwidth (decreasing $e$ ) results in a higher attack resilience. . . . .	117
4.13 Relational Data with Categorical Types: (a) The watermark alteration surface with varying $c$ (watermark modifications) and attack size. Note the lower-left to upper-right tilt. (b) The watermark degrades almost linearly with increasing data loss. . . . .	118
4.14 Relational Data with Categorical Types: (a) Embedding time dependency as a function of $e$ and $N$ . (b) Detection time requirements are similar to embedding and linear in the size of the data. . . . .	118
5.1 Discrete Streaming Data: Sensor Streams Watermarking Scenario. . .	122
5.2 Discrete Streaming Data: Stream Processing is necessarily bound in both time (stream rate) and space (window). . . . .	126
5.3 Discrete Streaming Data: (a) A sample stream. If all the extremes are considered to be major, then the resulting label bits for $K$ are shown (for $\varrho = 2$ ) (b) $\delta$ -Radius characteristic subset of extreme $\eta$ . . . . .	127
5.4 Discrete Streaming Data: Initial Embedding Algorithm . . . . .	132
5.5 Discrete Streaming Data: Initial Detection Algorithm . . . . .	135
5.6 Discrete Streaming Data: Average exhaustive search iterations required in computing the closest point that satisfies the characteristic subset bit encoding convention (logarithmic scale). . . . .	143

Figure	Page
5.7 Discrete Streaming Data: Overview of proof of concept implementation.	154
5.8 Discrete Streaming Data: Label alteration for increasingly aggressive uniform altering epsilon attacks. (a) Different label bit sizes shown. A smaller label size seems to survive better. (b) Different altered data percentages shown. . . . .	156
5.9 Discrete Streaming Data: Watermark survival to epsilon-attacks. (a) Naturally, increasing $\tau$ and $\epsilon$ values result in a decreasing watermark bias. (b) Same shown for $\epsilon = 10\%$ . (real data) . . . . .	156
5.10 Discrete Streaming Data: (a) Label resilience under sampling conditions. A higher label bit-size naturally yields an increased fragility to sampling. (b) Label alteration for summarization of increasing degree.	157
5.11 Discrete Streaming Data: (a) Watermark survival to summarization. An increasing summarization degree results in a decreasing detected watermark bias. (b) Watermark survival to sampling. A bias of 10 ensures a true-positive probability of 99.999%. (real data) . . . . .	158
5.12 Discrete Streaming Data: (a) Watermark survival to segmentation. (b) Watermark survival to combined sampling and summarization. (real data) . . . . .	159
5.13 Discrete Streaming Data: (a) Computation overhead (iterations) in multi-hash encoding increases with increasing guaranteed resilience (e.g., sampling degree) levels (logarithmic scale). (b) Decreasing the number of considered bit-encoding extremes (increasing $\phi$ ) decreases the impact on mean and standard deviation in the watermarked data.	161
6.1 Semi-structured Aggregates: A webpage as a semi-structure. . . . .	165
6.2 Semi-structured Aggregates: Tolerant Canonical Labeling. Composite Labels are a result of successive training sessions. . . . .	169
6.3 Semi-structured Aggregates: A combination of propagated structural and node content information determines a node label. . . . .	170

Figure	Page
6.4 Semi-structured Aggregates: (a) The surface defining the composite label collisions appearing after 4 stages of training (i.e., $i = 4$ ) with a random generated set of surgeries applied to the graph. It is to be noted that lower $\gamma$ values seem to yield a lower number of composite label collisions but in turn results in a lower resistance to structural attacks (i.e., as labeling will not be as resilient to graph surgeries). (b) The zero-collision (for composite labels) surface in the (iterations, $\alpha$ , $\gamma$ ) space corresponding to the same set of surgeries. Its existence proves the ability to label resiliently (to the considered surgeries) without colliding resulting composite labels. Computed using the <b>wmx.*</b> package. (c) The considered graph. . . . .	173
6.5 Semi-structured Aggregates: Labeling Algorithm. . . . .	174
6.6 Semi-structured Aggregates: Watermark Embedding Algorithm . . .	176
6.7 Semi-structured Aggregates: Watermark Detection Algorithm . . . .	177
6.8 Semi-structured Aggregates: Surfaces defining the composite label collisions appearing after 3 stages of training with a random generated set of surgeries. (a) Tree shaped graph. Much of the web content online is tree-shaped. Again, note that lower $\gamma$ values seem to yield a lower number of composite label collisions. (b) Star shaped graph. Note the smoother shape and the lower collision bounds, compared to (a). The same nodes were used, differently interconnected. Computed using the <b>wmx.*</b> package. . . . .	181
6.9 Semi-structured Aggregates: Averaged watermark loss over 10 runs of an 8 bit watermark embedded into an arbitrary 32 node graph with 64 edges. Surgery attacks are applied randomly (node removals 60%, link addition 20%, link removal 20%). The labeling scheme was trained for 3 surgeries. . . . .	181
7.1 Model of Watermarking: Mallory attacks (different variations). . . .	201
7.2 Model of Watermarking: (a) No matter how sophisticated the watermarking method, there exists a random attack with a success probability (e.g. of 35% and above, shown here for 2-dimensional usability spaces, see Section 7.3.4 for a discussion on high dimensional spaces). It can be seen that a lower $\epsilon_w$ value (more convincing in court) yields an even higher upper bound on attack success probability (2D cut through (b)). (b) The 3D evolution of $P_{sa}$ with varying $\epsilon_w$ and $R_a/\Delta u_{max}$ . . . . .	202

7.3	Model of Watermarking: (a) Sparse maximum allowable usability vicinity, $U_{max} = \cup(U_{max_i})$ , (b) A concave $U_{max}$ does not respect optimality. . . . .	202
-----	--	-----