Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. [http://researchspace.auckland.ac.nz/feedback](http://researchspace.auckland.ac.nz/feedback)

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.
Concepts and Techniques in Software Watermarking and Obfuscation

William Feng Zhu

August 2007

Supervisor: Prof. Clark Thomborson

A thesis submitted in partial fulfillment of the requirements of Doctor of Philosophy in Computer Science

The Department of Computer Sciences
The University of Auckland
New Zealand
ABSTRACT

With the rapid development of the internet, copying a digital document is so easy and economically affordable that digital piracy is rampant. As a result, software protection has become a vital issue in current computer industry and a hot research topic.

Software watermarking and obfuscation are techniques to protect software from unauthorized access, modification, and tampering. While software watermarking tries to insert a secret message called software watermark into the software program as evidence of ownership, software obfuscation translates software into a semantically-equivalent one that is hard for attackers to analyze. In this thesis, firstly, we present a survey of software watermarking and obfuscation. Then we formalize two important concepts in software watermarking: extraction and recognition and we use a concrete software watermarking algorithm to illustrate issues in these two concepts. We develop a technique called the homomorphic functions through residue numbers to obfuscate variables and data structures in software programs. Lastly, we explore the complexity issues in software watermarking and obfuscation.
Acknowledgment

It has taken nine years to officially complete my doctoral study. This is a hard experience, but it is also a fruitful period. I want to use this opportunity to thank all people who helped me in this way or that way.

I would like to thank my supervisor, Professor Clark Thomborson, for valuable guidance and financial support. I have learnt a lot from him. I would also like to thank the advisors for my doctoral study at the University of Auckland – Professor Fei-Yue Wang, Professor Christian Calude, and Doctor Michael Dinneen. The Department of Computer Science also financially supported my doctoral study at Auckland. It deserves my many thanks.

I was fortunate to be part of a very active research group here at Auckland. I have worked closely with and learnt a lot from several students, postdoctorals and faculty members in our group.

Especially, I would like to thank my friends. Mr. Fred He and his wife for their strong support for my study. Dr. Mingkuan Liu also helped me a lot. Thanks also to other friends who encouraged me and helped me.

I would also like to express my gratitude towards my daughter and my family. Their support always makes me overcome every difficulty.

I am also indebted to the following people for their help in my academic career: Huacan He, Dehuang Chen, Jiarui Wu, Huaxiao Zhang, Cheng Ge, Yiyu Yao,
Xindong Wu, Daniel Zeng, Yixin Zhong, T. Y. Lin, Qing Liu, Jingtao Yao, Guoyin Wang, Weizhi Wu, Guilong Liu, Min Xiao, Zhongjin Cheng, Rong Su, Xuekong Yang, Jianxin Feng, Guoliang Dai, Yizhong Zhan, Fangcai Liu, Ning Lu, and Jishou Ruan.
To My Dear Daughter

Miss Ruolin Zhu
# Contents

1 Introduction .................................................. 1
  1.1 Software Security and Protection ......................... 1
  1.2 Software Watermarking and Obfuscation ................. 4
  1.3 Goal, Structure and Contribution ....................... 4

I Software Watermarking ........................................ 7

2 Survey of Software Watermarking ............................ 9
  2.1 Overview of Software Watermarking ...................... 10
  2.2 Taxonomy of Software Watermark ......................... 11
    2.2.1 Classification by Purpose .......................... 12
    2.2.2 Classification by Extracting Technique ............. 12
    2.2.3 Robust and Fragile Software Watermark ............. 13
    2.2.4 Visible and Invisible Software Watermark .......... 13
    2.2.5 Blind and Informed Software Watermark ............ 14
    2.2.6 Tamperproofing Software Watermark ................. 14
  2.3 Attacks on Software Watermark ........................... 14
  2.4 Protection of Software Watermark ........................ 15
  2.5 Evaluation Criteria of Software Watermark ............. 16
## CONTENTS

2.6 Research Platforms for Software Watermarking ............................. 16

2.7 Software Watermarking Algorithms ............................................. 17
   2.7.1 Basic Block Reordering Algorithms ...................................... 17
   2.7.2 Register Allocation Algorithms .......................................... 18
   2.7.3 Spread-spectrum Algorithms .............................................. 18
   2.7.4 Opaque Predicate Algorithms ............................................ 19
   2.7.5 The Threading Algorithm ................................................. 22
   2.7.6 The Abstract Interpretation Algorithm .................................. 22
   2.7.7 The Metamorphic Algorithm .............................................. 22
   2.7.8 Dynamic Path Algorithm .................................................. 23
   2.7.9 The Mobile Agent Watermarking ....................................... 23
   2.7.10 Graph-based Algorithms ............................................... 23
   2.7.11 Birthmarks ...................................................................... 25

2.8 Complexity Problems in Software Watermarking ....................... 26

2.9 Conclusion ............................................................................. 29

3 Extraction .............................................................................. 31
   3.1 Overview ............................................................................ 32
   3.2 Embedding .......................................................................... 33
   3.3 Extracting ........................................................................... 42
   3.4 Representative Extracting .................................................... 49
   3.5 A Software Watermark Embedding and Extracting System ........ 54
   3.6 Conclusions ........................................................................ 55

4 Recognition ............................................................................ 57
   4.1 Recognition ......................................................................... 58
      4.1.1 Recognitions and Partial Recognitions .............................. 58
## CONTENTS

4.1.2 Blind Recognitions ............................................. 65
4.2 A Software Watermark Embedding and Recognition System ........ 67
4.3 Conclusions ......................................................... 68

II Software Obfuscation .......................... 71

5 Survey of Software Obfuscation ............. 73
5.1 Why Obfuscate Software? ......................... 74
5.2 Definition of Software Obfuscation ................. 75
5.3 Taxonomy of Software Obfuscation ............... 75
5.4 Criteria of Software Obfuscation ................. 77
5.5 Status of Software Obfuscation ................. 78
5.6 Conclusion ....................................................... 79

6 Homomorphic Functions .......... 81
6.1 Basic Concepts about Residue Numbers ........ 82
6.2 One-dimensional Homomorphic Obfuscations .... 85
   6.2.1 Definition of Homomorphic Obfuscations ........ 85
   6.2.2 Examples of Homomorphic Obfuscations ........ 86
   6.2.3 Representation of Homomorphic Obfuscations .... 87
6.3 K-dimensional Homomorphic Obfuscations ...... 89
   6.3.1 Basic Definitions ............................................. 89
   6.3.2 Representation of Homomorphic Obfuscations .... 90
6.4 Homomorphic Functions and Orders .......... 93
6.5 Division of Integers by a Constant .......... 94
6.6 Division of Integers by Several Constants .... 96
6.7 Conclusion ....................................................... 97
7 Application of Homomorphic Function

7.1 Array Transformations
   7.1.1 Array Index Change
   7.1.2 Array Folding and Flattening

7.2 Application of Homomorphic Function to Arrays
   7.2.1 Index Change
   7.2.2 Index and Dimension Change
   7.2.3 Array Folding
   7.2.4 Array Flattening

7.3 Conclusions

III Conclusions and Future Work

8 Conclusions and Future Work

8.1 Conclusions
   8.1.1 Software Watermarking
   8.1.2 Software Obfuscation
   8.1.3 Complexity and Security

8.2 Future Work
   8.2.1 Combination of Software Watermarking and Obfuscation
   8.2.2 Applications of Rough Set Theory to Software Security

9 References

A Publications

B Academic services