

Information Theoretic Formulation and Analysis of Performance Issues in Biometric Authentication

by

JAY PRAKASH RAVINDRA BHATNAGAR

Department of Electrical Engineering

Submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy
to the



INDIAN INSTITUTE OF TECHNOLOGY, DELHI
HAUZ KHAS, NEW DELHI, 110016- INDIA
NOVEMBER 2009

Certificate

This is to certify that the thesis titled “**INFORMATION THEORETIC FORMULATION AND ANALYSIS OF PERFORMANCE ISSUES IN BIOMETRIC AUTHENTICATION**” being submitted by Jay Prakash Ravindra Bhatnagar for the award of Doctor of Philosophy is a record of bona-fide work carried out by him under our joint guidance and supervision at the Department of Electrical Engineering, Indian institute of Technology, Delhi. The work presented in this thesis has not been submitted elsewhere either in part or full, for the award of any other degree or diploma.

Dr. Brejesh Lall
Assistant Professor
Electrical Engineering
IIT Delhi

Dr. R K. Patney
Professor
Electrical Engineering
IIT Delhi

Acknowledgements

Sri Radha-Krishna Vande Jagad Gurum. I was blessed with many wonderful teachers and I am deeply indebted to them all. My liking for Communication theory, Information Theory and Applied Math is mainly due to **Professor Zhen Zhang** - my Mentor and Teacher at USC. Sincere gratitude is due to teachers from USC, UC Riverside and Georgia Tech. These settings were first rate in providing me with enough courses to handle research. I returned home (India) due to family concerns and the academic environment here offered many different challenges. I think it is nice to have role model (s) in a lifetime and mine are mostly from 'Theory' world. It is my greatest pleasure and honor to sincerely thank **Professor Robert Gallager** arguably the God of Information theory (as quoted by Zhen Zhang). His exemplary work, teachings and personally encouraging me transported my energy levels and focus to a completely different level. In almost all technical problems and thought processes relating many details in biometrics and communication-coding-information theory, I have sincerely and carefully sheltered myself in **Gallager's** and **Verdù's** methods-teachings. However, the mistakes are solely my own limitations in dealing with so pure research as the works of early Information Theorists. The work developed by me in this thesis, took a flight back in time to the 60's and 80's of Information and Communication theory. It is an almost divine experience to have relived the Golden era while developing Information models and theorems for Pattern recognition using Biometrics.

I am deeply thankful to **Prof. R. Patney** and **Dr. B. Lall** for providing critical comments on the presentation and wording of synopsis – thesis (2007-2010). They were my 'steadfast' well wishers and moral support in many difficult times at IIT Delhi. In particular, Professor Patney made me appreciate finer human values and read much of my work with corrections in presentation. I started PhD in IIT Delhi with **Dr. A. Pathak** who introduced me to biometrics and I will like to thankfully acknowledge this association. This was between 2006- 07 when he was at IIT Delhi. I had my first length of PhD and the second length followed like re-birth and resurrection.

Finally, my two charms: **my Mother and my Father** who have been the invaluable boon of Lord, my Support and reason for my return to India. They have been my dazzling Armor and my human ATM (funds, moral support, advice, help....) all through my tribulations in PhD work. **I dedicate this thesis to them.** I thank **my Sister** for her gracious and loving support. Thanks are due to **Sudarsan, Srinivas, Amioy and Nipun** for the help provided in some simulations and figures produced in the thesis. My Sincere thanks are due to **Liby Thomas, Dr. Jiwani, Dr. Bhattacharya, Bighnaraj** and my peers from the student community at IITD for their warmth and affection. I am grateful to all that has ever happened (even the worst) and I hope to continue making some sense in this dangerously complex world.

Jay R. Bhatnagar

Abstract

The objective of this thesis, as reflected in its title, is investigating fundamental theorems and formulations in information transmission due to Shannon, Gallager to performance related aspects of biometric authentication. This is a recent offshoot of pattern recognition systems that perform recognition of humans with the purpose of providing security. In chapter two we develop constrained capacity as a performance index of biometric authentication using the information-capacity theorem. We illustrate applications of constrained capacity in estimating: authentication performance, structure of cohorts or interferers, performance with fusion rules, and relevance of features. The formulation of capacity developed in this thesis basically provides a fundamental limit given in terms of the number of users in a database, for a given database and matcher function, that can be reliably authenticated. In this way, our capacity formulation characterizes the performance of the authentication system. In another performance related index, the error in discriminating users based on their biometric information and subject to intra-class noise/variability is given by the Probability of Random Correspondence (PRC). In chapter three we investigate information theoretic approaches to formulate PRC of biometric by using formulations based on rate-distortion approach and random coding exponents of noisy sources. As a performance measure that mainly depends on a biometrics, PRC characterizes the limit to reliability constrained by the noise inherent in biometric information independent of the authentication system. In chapter four we propose architecture and algorithm to implement rank level fusion by exploiting the redundancy in the intra-template information due to statistical dependence among feature-subsets in a template. Our motivation in doing so based upon information theoretic techniques such as sequential decoding, LDPC iterative decoding and multiple description of correlated sources We demonstrate our approach for rank level fusion by using palmprint biometrics. In chapter five we propose an iterative algorithm for estimating sample size requirements that can give high confidence of readings from the ROC. In real applications of biometric systems, when query templates of a user are much different than the enrolled templates of the user actual ROC is much different than the calibrated ROC. This necessitates that there be enough number of samples per user in the database to maintain a high confidence level of the ROC. The iterative algorithm, using information theoretic measures, proposed in the thesis is an approach for estimating sample size requirements for reliable authentication.

Contents

Acknowledgements	(i)
Abstract	(ii)
List of Figures	(iii)
List of Tables	(v)
List of Abbreviations and Symbols	(vi)
Chapter 1	
Overview of Thesis	
1.1 Introduction to biometric authentication	1
1.2 Functional model of biometric authentication system	3
1.3 Biometric Authentication and Information Theory: A Parallelism	5
1.4 Organization of Thesis	8
1.5 Performance aspects of biometric authentication	11
1.6 Capacity as a performance measure of biometric authentication	13
1.6.1 Recognition capacity of biometric authentication	14
1.6.2 Some useful results from early information theory	15
1.6.3 Constrained capacity of biometric authentication	18
1.7 Measures of random correspondence of biometrics	19
1.7.1 Compound statistical methods for <i>PRC</i> of fingerprint biometrics	19
1.7.2 Random correspondence based on Divergence measures	21
1.7.3 Error exponents for a binary source-channel	22
1.7.4 A <i>PRC</i> bound of biometrics based on error exponents	24
1.7.5 Measures of Random correspondence of biometrics based on coding theorems	25
1.8 Multimodal biometric systems - Fusion principles	26
1.8.1 Sequential decoding	27
1.8.2 Multimodal biometric systems and fusion at rank level	28
1.8.3 Rank level fusion based on sequential decoding	29
1.9 Estimating sample size requirements for stable <i>ROC</i>	30
1.9.1 Formulations based on statistical approaches	31
1.9.2 Estimating sample size requirements based on AEP	32

Chapter 2

Constrained Capacity of Biometric Authentication

2.1 Introduction	33
2.2 Shannon Capacity for noisy authentication channel	35
2.3 Constrained Capacity of Biometric Authentication System: Proposed Framework	36
2.4 Constrained Capacity of Multimodal (score level) Biometric Systems	44
2.5 Experiments and Discussion	45
2.6 Conclusions	55
Appendix 2.1	56

Chapter 3

Measures of Random Correspondence of Biometrics

3.1 <i>PRC</i> based on a Rate-Distortion framework	58
3.2 <i>PRC</i> of biometrics based on error exponents	62
3.2.1 Motivation	62
3.2.2 Formulating <i>PRC</i> of biometrics using Error exponents	64
3.2.3 Simulations illustrating use of error exponent to calculate <i>PRC</i>	68
3.2.4 Conclusions (<i>PRC</i> based on Error exponents)	73
3.3 Constrained capacity as measure of uniqueness of biometrics	74
3.4 <i>PRC</i> of biometrics based on matcher error statistics	77
3.4.1 Motivation	77
3.4.2 Generating error plot	78
3.4.3 Formulating probability distribution of error clusters	81
3.4.4 Formulating probability distribution of error points	83
3.4.5 Formulating <i>PRC</i> of biometrics	85
3.4.6 Conclusions	87
Appendix 3.1	88