System Architecture for Secure Mobile Internet Voting

By

Surendra Thakur

Submitted in fulfilment of the requirements for the degree of

Doctor of Technology

in the

Department of Information Technology in the
Faculty of Accounting and Informatics

at the

Durban University of Technology
Durban, South Africa

Dedication

I dedicate this thesis to my family – my parents, Deonarian and Manpathy who both have passed on but who will never be forgotten, and my wife Maleni, daughter Mikaela, and son Sachin. Without their constant support and encouragement, the writing of this thesis would not have been possible.

Acknowledgements

I am an unstructured busybody, so sitting and writing without an artefact to anchor me is completely outside my comprehension. It is in this context that I thank my supervisors.

Firstly, Professor Olugbara, who persuaded me to study when I had no such intention. And he did this rather astutely, initially getting me to write papers on my electoral work - which I love - and then, before I realised it, roping me in to start my doctoral thesis.

Secondly, Professor Millham, who joined us later and introduced a calmness which helped keep me on track.

In the course of my consultations, the three of us had huge shouting debates. I do hope all my nervous colleagues who were within earshot will now come to see that for what it was - intellectual rigour - and even adopt this form of discourse.

A special word to Dr Adetiba, a post-doctoral research fellow, who embraced my non-trivial topic with passion and provided enormous support as we elevated the model.

Professor Bawa is my remarkable Vice-Chancellor and boss. He gently harangued me at every opportunity to honour the pledge I had made to my late mom to complete my studies.

My wife endured the brunt of the emotional fallout that is the unfortunate byproduct of months of intellectual toil. My heartfelt apologies and grateful thanks, Maleni, for your support. Mikaela and Sachin Thakur, thank you for inspiring dad with your awesomeness.

Finally, I thank my team at Nemisa e-Skills – thanks for the lift and the push.

Plagiarism Declaration

I, Surendra Thakur, Student number 19650230 know and understand that plagiarism is using another person's work and pretending it is one's own, which is wrong.

I declare this thesis is my own work.

I have appropriately referenced the work of other people I have used.

I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his/her own work.

_____ _____

Student Signature (Mr Surendra Thakur) Date

APPROVED FOR FINAL SUBMISSION

_____ _____

Promoter (Professor Oludayo, O. Olugbara) Date

_____ _____

Joint Promoter (Professor Richard Millham) Date

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANN	Artificial Neural Network
BNN	Biological Neural Network
BYOD	Bring Your Own Device
COTS	Commercial Off The Shelf
DoS	Denial of Service
DCT	Discrete Cosine Transformation
DDOS	Distributed Denial of Service
DRE	Direct Recording Electronic
DWT	Discrete Wavelet Transform
eID	Electronic Identity Card
EMB	Electoral Management Board
EVM	Electronic Voting Machine
FOO	A practical secret voting scheme for large-scale elections conceived by Fujioka, Okamoto and Ohta (1992)
GMM	Gaussian Mixture Model
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HTTP	Hypertext Mark-up Language
HTTPS	Hypertext Transfer Protocol Secure
HMM	Hidden Markov Method
HOG	Histogram of Gradients
IEC	The Electoral Commission of South Africa

iDTV	Integrated Digital Television
LPCC	Linear Prediction Cepstral Coefficients
MOC	Match-on-a-Card
MFCC	Mel-Frequency Cepstral Coefficients
MFDWC	Mel-Frequency Discrete Wavelet Coefficient
NFC	Near Field Communication
NIST	National Institute of Standards and Technology
OTP	One Time Password
OSS	Optical Scan System
PIN	Personal Identification Number
PHP	Hypertext Preprocessor
PVID	Pseudo-Voter Identity
REVS	Robust Electronic Voting System
RFID	Radio-Frequency Identification
SIM	Subscriber Identity Module
SMIV	Secure Mobile Internet Voting
SMS	Short Message Service
SHOG	Spectral Histogram of Distribution
VVAT	Voter Verifiable Audit Trail
VVPAT	Voter Verifiable Paper Audit Trail
WAP	Wireless Application Protocol
WML	Wireless Markup Language

List of Publications

Journal

Thakur, S., Adetiba, E., Olugbara, O.O., Millham, R. 2015. Experimentation using short-term spectral features for secure mobile internet voting authentication. *Hindawi Mathematical Problems in Engineering. In Press.*

Book Chapter

Thakur, S. 2015. E-voting: India and the Philippines – A comparative analysis for possible adaptation in Africa. In: Sodhi, IS. Ed. E*merging issues and prospects in African e-Government*, 28-55. Hershey, PA: Information Science Reference. doi:10.4018/978-1-4666-6296-4.ch003.

Proceedings

Thakur, S. and Boateng, R. 2011. E-voting for good governance and a green world, In: *Proceedings of the Africa Digital Week*. July 25-29 2011. Accra, Ghana: African Institute of Development Informatics and Policy, 55-62.

Thakur, S. and Singh, S. 2012. A study of some e-government activities in South Africa. In: *2012 e-Leadership conference on sustainable e-Government and e-business innovations (E-LEADERSHIP).* IEEE, 1-11.

IEC. 2014. Electronic voting, an enabler or disabler to strengthening electoral democracy? In: *Seminar on Electronic Voting and Counting Technologies.* Cape Town, South Africa, 12 March 2013, 1-20.

Thakur, S., Olugbara, O.O., Millham, R., Wesso, H.W., Sharif, M. and Singh, P. 2014. Transforming the voting paradigm – the shift from inline, to online to mobile voting. In: Asia-Pacific Institute of Management. *Internatio nal Summer School on Information and Communication Technology for Democracy*. New Delhi, India, 9-15 March 2014.

Thakur, S., Olugbara, O.O., Millham, R., Wesso, H.W., Sharif, M. and Singh, P. 2015. Transforming the voting paradigm - the shift from inline, through online to mobile voting. In: *IEEE International Conference On Adaptive Science and Technology (ICAST)*. Lagos, Nigeria, 29 October 2014. IEEE. **Best paper award**

Maphephe, J., Balkaran, R. and Thakur, S. 2014. Digital and interactive content production as part of Lesotho strategic development – a brief study on Lesotho working towards national collaboration for updated civil register and voter register in the spirit of improved service delivery. *African journal of computing and ICT, (*7)4: 85-90. IEEE.

Keynotes

Thakur, S. and Murphy, S. 2010. Digital democracy – using electronic voting to empower the nation**.** *Keynote*. In: *Proceedings of The Conference for Software Quality***.** Johannesburg, South Africa, 18 November 2010.

Thakur, S. 2011. Is it time for E-voting in Africa? the Experiences of the Philippines and India. *Keynote. P*ushing buttons for electoral change. In: *6th annual EISA Symposium*. Nairobi, Kenya, 23-24 November 2011.

Thakur, S. 2012. *Keynote.* In: *The first E-Leadership Conference on Sustainable E-Government and E-Business Innovations in Africa*. University of Pretoria, Pretoria, South Africa, 4-5 October 2012. IEEE.

Thakur, S. 2012. Digital democracy: using e-voting to empower nations? Is it On or Off? *Keynote. I*n: *14th Annual Conference on ZAWWW Applications,* Mangosuthu University of Technology, Durban, South Africa, 7-9 Nov 2012.

Thakur, S. 2013. E-voting a X-national experience. *Keynote*. In: *Seminar on Electronic Voting and Counting Technologies.* Cape Town, South Africa, 11-12 March 2013.

Thakur, S. 2015. Digital democracy, and smart certifications. *Keynote*. In: *The Internet of Everything CISCO Southern African Conference*. Durban, 26-30 May 2015.

Public Lectures

Thakur, S. 2014. Digital Democracy in South Africa: Is it on or off? *Centre for the Civil Society,* University of KwaZulu-Natal, 14 March 2014.

Thakur, S. 2014. Pressing buttons for democracy: a contextual evaluation. *Urban Futures Centre*, Seminar Lecture series, Durban University of Technology seminar lectures. 2 October 2014.

Thakur, S. 2014. Design Digital Festival: Democracy 2020. *Faculty of Arts and Design*. Durban University of Technology, Durban, 12-13 September 2014.

Thakur, S. 2013. IEC commission briefing to commissioners, Johannesburg, South Africa.

Radio Interviews

Thakur, S. 2014. Is SA available for e-voting? Interview by John Maytham on Afternoon Drive, *CapeTalk Radio* (radio broadcast), 21 November 2014. (Online). Available: https://soundcloud.com/primediabroadcasting/is-sa-ready-for-evoting

Thakur, S. 2014. E-voting and the Namibian elections. Interview by Matthew Veeran on NewsBreak. L*otusFM Radio* (radio broadcast), 26 November 2014.

Thakur, S. 2014. The Brazilian elections. Interview by Matthew Veeran on NewsBreak, *LotusFM Radio* (radio broadcast), 28 October 2014.

Thakur, S. 2013. The NOTA option in India. Interviewed by Matthew Veeran on NewsBreak, *Lotus FM Radio* (radio broadcast), 28 September 2013.

Newspaper Articles

Mucunguzi, A. 2010. Conversations on technology: E-voting in Africa: Mr. Collin Thakur interview,. *PCTechMagazine*, Uganda, September-October 2010: 26.

Sidimba, L. 2014. Youth want to swipe or click for democracy. *The Sowetan,* 15 November 2014.

Cohen, S. 2014. Voting in SA at the click of a button. *The Witness,* 6 November 2014: 8.

Artefacts

Thakur, S. and Beer, C. 2014. An interactive token based system to seamlessly recognise documents. A*rtefact* (online). Available: http://www.authenticateit.co.za (Accessed 20 June 2015).

Electoral observation

Thakur, S. 2011. The Zambian elections. International Observer. Electoral Institute for the Sustainability of democracy in Africa.

List of Awards

Commissioned Research

Thakur, S. 2012. Electronic voting: a cross border analysis. *The Electoral Commission*. Pretoria, March 2012.

Thakur, S. and Dávila, R. 2013. The path towards effective solutions: a study on voter registration experiences and technology. Brussels. UNDP.

Abstract

This thesis focuses on the development of an enhanced innovative secure mobile Internet voting system architecture that offers desirable security requirements to theoretically mitigate some of the intrinsic administrative and logistical challenges of voting, inter alia lack of mobility support for voters, voter inconvenience, election misconduct, and possible voter coercion often associated with the conventional poll-site voting system. Systems in existence have tended to revolve around the need to provide ubiquitous voting, but lack adequate control mechanism to address, in particular, the important security requirement of controlling possible coercion in ubiquitous voting. The research work reported in this thesis improves upon a well-developed Sensus reference architecture. It does so by leveraging the auto-coupling capability of near field communication, as well as the intrinsic merits of global positioning system, voice biometric authentication, and computational intelligence techniques. The leveraging of the combination of these features provides a theoretical mitigation of some of the security challenges inherent in electoral systems previously alluded to. This leveraging also offers a more pragmatic approach to ensuring high level, secure, mobile Internet voting such as voter authentication. Experiments were performed using spectral features for realising the voice biometric based authentication of the system architecture developed. The spectral features investigated include Mel-frequency Cepstral Coefficients (MFCC), Mel-frequency Discrete Wavelet Coefficients (MFDWC), Linear Predictive Cepstral Coefficients (LPCC), and Spectral Histogram of Oriented Gradients (SHOG). The MFCC, MFDWC and LPCC usually have higher dimensions that oftentimes lead to high computational complexity of the pattern matching algorithms in automatic speaker authentication systems. In this study, higher dimensions of each of the features were reduced per speaker using Histogram of Oriented Gradients (HOG) algorithm, while neural network ensemble was utilised as the pattern-matching algorithm. Out of the four spectral features investigated, the LPCC-HOG gave the best statistical results with an R statistic of 0.9257 and Mean Square Error of 0.0361. These compact LPCC-HOG features are highly promising for implementing the authentication module of the secure mobile Internet voting system architecture reported in this thesis.

Chapter One – Introduction

The world of today has surpassed the *mobile moment*; the number of mobile devices has approximately exceeded the number of people, proliferating the global society with all kinds of mobile devices (Gallant *et al.* 2014; International Data Corporation 2013). Although the mobile moment has not reached the one-device for one-person nirvana, owing to the fact that not all people have mobile devices and some have many, the fact that it has been surpassed provides a unique opportunity to rethink how voting is conducted. Concurrent with this mobile moment, there has been pragmatic evidence of a perceptible downward trend of voter participation in the process of democratic decision-making (Bittiger 2007; Ellis *et al.* 2006; IDEA 2014). This takes the form of low voter turnout rates, particularly amongst the youth and elderly people (Scott *et al.* 2012). Given the high penetration of mobile devices, one potential way to address this challenge is to connect government electoral resources with mobile devices to enable a wider participation of citizens in the electoral process (Hill and Louth 2006; Hill and Alport 2007).

Mobile devices have now become so embedded in human life that people are utilising them to transact all kinds of business activities (Siau and Shen 2003; McGrane 2013; Zambrano and Seward 2011). Citizens of many countries of the world are beginning to believe that one way to enforce openness, transparency, and accountability in their government's electoral processes is to draw on the power of technology to conduct voting anywhere, anytime (Alvarez and Hall 2010). Voting using mobile devices, with the associated benefits of providing mobile convenience and fostering mass participation in the electoral process, is increasingly being demanded by citizens (McGrane 2013; Zogby and Kuhl 2013; Allen 2006). None of these potential benefits can be realised if mobile Internet voting is not secure.

The use of mobile devices for capturing, counting, and managing elections can pose significant security threats and risks that can jeopardize the real essence of free, fair and acceptable elections. There is, in all likelihood, no person who wants her[1] voting choices disclosed without consent because such information may be used to the

[1] This paper respectfully follows the precedent of Shamos (2004) by symbolically using the female form to refer to voters, in recognition to women's superior numeric numbers as well as the need to reintroduce women into democratic governance.

detriment of her rights as a voter. It is therefore imperative to seek effective ways to secure privacy of data and ensure that this privacy is preserved by a voting system or any other electronic information system that hosts voter data. If security is not adequately protected by shielding the electoral infrastructures from vulnerability, the trust of citizens will be lost and the electoral process will be illegal. It should be possible to always safely access, reliably transmit and maintain confidentiality of the electoral data, and keep these safe from threats and risks posed by malicious groups.

Security, in the context of the electoral process, exists when election data is securely protected against threats and risks (Moynihan 2004). The essential requirements for secure electronic voting systems have been identified in the literature to guide the design of secure voting systems (Qadah and Taha 2007; Gritzalis 2002).

Inspired by the need to enable a wider population of citizens to participate in a truly democratic electoral process without hindrances (Thakur *et al.* 2015; Brücher and Baumberger 2003), the researcher rigorously pursued the non-trivial task of designing a Secure Mobile Internet Voting (SMIV) system architecture. This SMIV system architecture is fundamentally based on the reference architecture of Sensus, a security conscious Internet polling system (Crano and Cytron, 1997), whose underlying concept is derived from Fujioka, Okamoto and Ohta (1992) and their practical secret voting scheme for large-scale elections. This protocol is also known as FOO. The security requirements fulfilled by Sensus - such as eligibility, convenience, and mobility - were implemented differently in this research work from the Sensus approach, and the requirement of incoercibility, which Sensus does not consider, was included.

With regards to the security requirement of eligibility, in Sensus, the Voters Identification Number (VIN) and secret token were used to implement eligibility; in this research work, multimodal authentication, including voice biometric, was used instead. Biometrics is the automated use of science and technology to uniquely identify individuals based on physiological or behavioural characteristics. A primary motivation for using voice biometrics is because it effortlessly and repeatedly recognises an individual with a method that does not require technical knowledge or memory on the part of the subject (in this case, the voter) (Jain, Ross and Nandakumar 2011; Whither Biometrics Committee 2010; Pocovnicu 2009).

The security requirement of convenience was implemented in the Sensus architecture using familiar devices and the casting of the vote in one or two sessions. In this research work, convenience was enhanced using the Near Field Communication (NFC) tag attached to the voter ID card, in an addendum to Sensus (Crano and Cytron, 1997) implementation approach. NFC has the advantage of automatically launching the correct voter application, thus mediating phishing and malware while storing pertinent voter information off-device. NFCs automatic data transfer mediates and reduces voter input errors and saves time (Han, Hu and Kotagiri 2012; Coskun, Ok and Ozdenizci 2011; Ok, Coskun and Aydin 2010).

The security requirement of incoercibility was not considered in the Sensus reference architecture; but incoercibility was realised in this research work using the Global Positioning System (GPS) service to geofence and control the maximum number of voters from a fixed area. Geo-fencing (aka geofencing) is a feature in a software program that uses the GPS to define the geographical boundaries within which the voter declared she would cast her ballot. This approach to controlling for incoercibility, as implemented in this research work, is similar to the maximum number of voters allowable in a conventional polling booth so as to minimise fraudulent or coercive practices.

Regarding the security requirement of mobility, in the Sensus approach, a networked workstation was the access mode for voter participation, while this research supports voter mobility within a predefined precinct. This mobility leverages device familiarity and supports various classes of immobile voters while mimicking the traditional well-known electoral ward concept. The application of the SMIV architecture could be beneficial because it leverages the power of mobile and Internet technologies to conduct frequent elections or referendums in the form of deliberative democracy. In the particular context of South Africa, which is a young stable democracy with little evidence of electoral fraud and violence, mobile Internet voting would enhance efficiency, effectiveness, stability, and foster voting convenience at reduced costs. Since the country also has to balance competing priorities in other sectors, such as healthcare, social welfare, energy and infrastructure development, the capital outlay for these competing priorities may well be supported from cost-saving arising from the adoption of cost effective mobile Internet voting. The business of conducting elections in four or five year cycles, depending on the country, is

indisputably expensive in financial terms (IEC 2014; Thakur 2013). The United Nations Development Programme (UNDP) has spent at least USD 1.2 billion in the last 12 years supporting democracy (Thakur and Dávila 2013). This budget was primarily utilised to purchase equipment for voter registration drives, to create a Voter's Roll in order to improve electoral efficiency, and to mitigate electoral fraud such as ballot stuffing. A legacy benefit of this exercise has been the establishment of a core adult population register in recipient countries.

The use of emerging technologies such as NFC for voter data capturing and storing may well bring about significant savings. In general, mobile Internet voting as touted in this research can potentially offer the following benefits (Thakur *et al.* 2015):

a) Costs of printing and transporting of paper ballots across the country can possibly be reduced. The South African electoral process consumed about 460 tonnes of paper in the 2009 general elections (and over 500 tonnes in 2014) (IEC 2014); Sampath (2013), the Indian Chief Electoral Office, informed the The Electoral Commission of South Africa (IEC) that it used 12,000 tonnes of paper to run its last paper-based election in 1996 (which was the last year it used paper for its elections).
b) Mobile device ubiquity and frugality positions mobile as an irresistible medium to pragmatically engage the youth, the digital natives, the rural, and even the elderly in elections, irrespective of their geographical locations. It also engages specific segments of the electorate, such as diplomats, soldiers, healthcare workers, nomads, and the increasingly influential Diasporas who cannot make it to the physical poll-sites (Thakur 2012; IEC 2014).
c) Mediation of voter mobility and support for increasing the flow of voters can possibly be achieved because mobile devices offer a platform for ubiquitous voting at any time and at one's convenience (Alvarez and Hall 2010; 2004).
d) Mobile Internet voting can be specially designed to assist the blind and the partially sighted voters, as well as voters with mobility impairments, to cast their votes by allowing access from their own habitats. In this context, it can also help the elderly, the nomad and the disabled who may be unable or disinclined to travel to a specific location to vote.

e) Multi-lingual instructions with an effortless interactive interface can be offered without increased printing costs. This eliminates language intolerance.
f) Mobile Internet voting facilitates the concept of Bring-Your-Own-Device (BYOD), which serves to reduce implementation costs.
g) Hacking attempts, such as tempest attacks, which involve the electronic monitoring of radiation from voting screens to capture images and therefore monitor the vote (which is one of the reasons the Dutch stopped e-voting) can be mitigated.

Before outlining the research problem of this study, it is imperative to define some of the key terms of mobile device, election and vote; as well as to enunciate the general security requirements or requirements that an election must rigidly adhere to in order to be considered as both free and fair.

1.1 Definitions

1.1.1 Mobile device (aka a mobile)

For the purposes of this thesis, a mobile device (hereinafter referred to as 'a mobile') is formally defined as a communication device with the following desirable characteristics:

a) A relatively small *form factor,* which refers to the size, shape, style and layout of the device,
b) At least one wireless network interface for network access (data communications), such as Wi-Fi or cellular phone (aka cellphones),
c) Local built-in (non-removable) data storage,
d) An operating system,
e) Applications, available through a web browser or acquired and installed from third parties, must be able to execute on this platform,
f) One or more wireless personal area network interfaces, such as Bluetooth or near-field communications,
g) One or more wireless network interfaces for voice communications, such as cellular,
h) Global Positioning System (GPS), which enables location services,

i) Battery powered for mobility (Souppaya and Scarfone 2013; Karpov 2011).

Popular examples of mobiles include cellular or smartphones, personal digital assistants, phablets, tablets, and notebooks. Mobile computing devices include laptops as well (Han, Hu and Kotagiri 2012).

1.1.2 Election

An *election*, pedantically defined, is a process to obtain accurate data representing a set of participants' answers to a posed question (Mursi *et al.* 2013).

1.1.3 Vote

A *vote* is a participant's response to a posed question, and is generally referred to as a predetermined set of answers (Mursi *et al.* 2013).

1.2 Security Requirements of Elections

Probably the most exhaustive list of principles of secure voting as provided in a literature review, derived 26 security requirements for administering elections (Mursi, *et al. 2*013). These security requirements, along with their formal definitions, illustrate the complexity involved in an electronic voting system, and can be restated as follows:

a) **Eligibility:** Only valid voters who meet certain predetermined criteria are qualified to vote in an election. Some scholars refer to this requirement as **invulnerability** (Cranor and Cytron 1997; Mauw, Verschuren and de Vink 2007).

b) **Authentication:** Only voters who have obtained legal authorization should be able to vote in an election.

c) **Uniqueness/non-reusability:** No voter should be able to vote more than once; no voter may be able to change, duplicate or replicate someone else's vote. Some literatures also use **democracy** for this term, while others even combine this term with eligibility (Al-Saidi 2011).

d) **Privacy:** No one should be able to determine how any individual voter voted in any election. Al-Saidi (2011) helpfully adds that no participant other than the voter can determine the value of the vote cast.

e) **Convenience:** Voters should be able to cast their votes with minimal equipment and skills, no physical restrictions, and without a steep learning curve.

f) **Transparency:** Voters should possess a general knowledge and understanding of the whole voting process.

g) **Walk away:** A voter, after voting, should not be directly or indirectly involved in any other post vote process.

h) **Dispute freeness:** Any voting scheme must provide a mechanism to resolve all disputes at any stage of the election.

i) **Practicality:** A voting scheme should not have assumptions and unnecessarily rigid requirements that may be difficult to implement on a large-scale.

j) **Fairness:** No one should learn the outcome of an election before the official announcement of the result.

k) **Incoercibility:** A voting scheme should be resistant to duress and intimidation. In an electoral process, coercion occurs when an entity tries to swindle the manner in which a vote is cast; deceive a voter to abstain or vote for a particular candidate; or even falsely represent a valid voter by obtaining the voter's credentials. Although some literature also uses the term **uncoercibility** (Howlader *et al.* 2011), we defer to Mursi *et al.* (2013) and therefore use the term incoercibility.

l) **Accuracy/completeness:** Voting systems should record the votes correctly and no record should be omitted.

m) **Soundness:** No reasonably sized coalition of voters or authorities may disrupt the smooth conduct of an election.

n) **Verifiability:** Voters should be able to verify that their votes are correctly counted in the final tally.

o) **Integrity:** Votes should not be able to be modified, either in transit or anywhere else in the process, without detection.

p) **Reliability:** The system must be resilient to randomly generated malfunctions or to the system failure.

q) **Robustness:** The voting system should be sound and resilient regardless of partial failure of the system.

r) **Flexibility:** Equipment should allow for a variety of ballot formats so it can be used for several types of elections concurrently.

s) **Auditability:** There should be reliable and authentic election records that can be used as evidence in the case of voting discrepancies.

t) **Certifiability/function check:** Systems should be testable against essential criteria agreed upon by the law.

u) **Cost effectiveness:** Electoral systems and processes should be affordable.

v) **Voter mobility:** There should be no restrictions on the location from which a voter can cast a vote.

w) **Receipt freeness:** A voter should not be provided with a receipt that proves how he/she voted; this information should not be visible to any other entity. The researcher posits that a voter *could* receive a notification of successful voting, but this notification *must* be private and not available to any other entity. Some literature refers to this as **Uncoercibility** (Al-Saidi 2011).

x) **Verifiable participation:** It should be possible to find out whether a particular voter has participated in an election by casting a ballot or not.

y) **Efficiency:** The voting process should not require too many steps to reach the end of a voting process for voters. This implies that the whole election can be held in a timely manner, with, for example, all computations being done in a reasonable amount of time and voters are not required to wait for other voters to complete the process.

z) **Scalability:** The complexity of the protocols used in a voting scheme is a major factor in its practical implementation. An efficient voting scheme has to be scalable with respect to storage, computation, and communication needs so as to accommodate a larger number of voters (Mursi *et al.* 2013).

It is practically impossible to simultaneously satisfy all of these security requirements or requirements in *any* electoral system, whether it is a paper-based manual process, a

machine assisted mechanical voting process, or a computer automated terminal, as some of these criteria are by definition inherently in conflict with one other. In fact, as Mursi *et al.* (2013) indicate, a conflict exists between *authentication* versus *privacy,* as authentication requires checking the credentials of a voter, while at the same time, privacy requires the protection of the confidentiality of the voter. Another instance is *verifiability* versus *receipt freeness*, which requires enabling the voter to verify that his/her vote is correctly counted and correctly cast, without giving a receipt of the actual vote cast. The principles of frugal systems refer to the quality of being frugal, sparing, thrifty, prudent or economical with respect to resources. In particular, Frugal Information Systems strive to meet the following four information drives:

a) The drive to ubiquitously access information unrestricted by time and space (ubiquity),
b) The drive to precisely identify the characteristics and locations of entities (uniqueness),
c) The drive for information consistency (unison),
d) The drive to overcome the friction of incompatibilities of information systems (universality).

(Watson 2013; Watson, Kunene and Islam 2013; Olugbara and Ndlovu 2014)

The researcher argues that the SMIV architecture meets these requirements because the mobile and GPS allows some time, location and space supporting ubiquity; the NFC and voice biometric supports uniqueness; the confirmation code supports end-toe-end unison; while the auto-launching capability of NFC ensures the right systems are launched (universality).

As mentioned, the basic reference architecture that guides the design of this study's SMIV architecture is the Sensus (Crano and Cytron 1997). However, the researcher took a step further to resolve the conflict between ubiquity (mobility) and uniqueness (coercion). The resolution of this conflict is a distinctive contribution of this current research work because to the best knowledge of the researcher, no extant work has addressed it, and if any, it has not been adequately addressed as proposed in this thesis. The research has used Mursi *et al*.'s (2013) security requirements to provide

the evaluation framework that was used to establish a comparison between Sensus and SMIV architectures.

1.3 Research Problem

The greatest challenge facing the electoral process in many parts of the world is how to securely conduct an election to usher in credible leaders to manage the affairs of a nation, in a way that is globally judged to be free, fair, and acceptable to the generality of citizens and external observers. This challenge is even more convoluted when the population of eligible voters is colossal, such as in the case of India, which has about 814 million eligible voters[2]. Conducting a free, fair, and acceptable election can be an arduous task, especially in a country with enormous population, or in countries where there are reported cases of incessant violence, malpractice, corruption and quasi terrorism.

In a concerted effort to effectively manage electoral processes, technological innovations have been sought. The process of electronic voting is the most complex form of electoral technology upgrade as it touches upon the core of the electoral system and has to adequately satisfy many conflicting requirements. Whilst this upgrade process provides an opportunity to solve many old electoral problems such as human errors, coercion, and inaccessibility; it also opens new electoral problems such as placing system knowledge in the hands of a few, and regrettably also raises the issues of authentication, trust, and transparency. Moreover, one faces the challenge of developing trustworthy electronic voting systems, which include the resolution of authentication, since remote authentication can lead to impersonation (Jain, Ross and Prabhakar 2004; Kinnunen and Li 2010; Khelifi *et al.* 2013). There is also the possibility of insecure Internet protocol, service restricted area, incorrect architecture, security breaches, and real-time management of huge volumes of data that must be dealt with (Kohno, Stubblefield, Rubin and Wallach 2004; Pieters 2009; Thakur 2012). However, these issues are beyond the scope of this thesis.

The task of conducting free, fair, and acceptable elections using technology also poses other challenges to governments and researchers. For governments, besides the political resistance, there is the potential ignominy of failure. Whilst researchers have

[2] http://eci.nic.in/eci/eci.html

endeavoured in recent years to create, refine, and evaluate the capability of electronic voting system architecture to adequately solve election quandaries, there is no record of breakthrough yet because the task is obviously a difficult one. Electronic voting system research is still in the infancy stage, and the increasing rapidity of technological advances adds challenges of its own. Consequently, in designing an effective electronic voting architecture, one has to contend with diverse security requirements, as well as the need to adequately meet user satisfaction requirements (Kohno *et al.* 2004; Pardue, Landry and Yasinsac 2010; Abdelkader and Youseff 2012; Mursi *et al.* 2013).

Within this context, the one prevailing research question that is rigorously pursued in the course of this research work can be lucidly enunciated as follows:

How can emerging technological innovations such as mobile technology, global positioning system service, voice biometrics and near field communication, be embedded into the existing reference architecture of an electronic voting system to improve security requirements?

1.4 Research Aim and Objectives

The overarching aim of this research work was to develop a secure mobile Internet voting system architecture (referred to as a SMIV in this study) that can be implemented to enable massive citizen participation in a truly democratic electoral process without hindrances. The researcher hypothesizes that it should be possible to realise the desired system architecture from the set of security requirements in order to demonstrate that building secure mobile Internet voting systems is realisable. The following research objectives are setup to realise the aim of this research work:

a) To identify and enhance a secure mobile Internet voting system reference architecture that could help increase the public trust by leveraging the functional capabilities of mobile devices to improve the system security.
b) To effectively embed the emerging technological innovations of mobile technology, global positioning system service, voice biometric authentication, and near field communication technology into the identified reference architecture of a mobile Internet voting system for security enhancement.

c) To validate by experimentation, the effectiveness of the voice biometric authentication components of an embedded reference architecture of a mobile Internet voting system.

1.5 Contributions

The research work reported in this thesis concerns the design of a Secure Mobile Internet Voting (SMIV) system architecture that leverages the functional capabilities of mobile devices, near field communication technology, global positioning system service, and voice biometric technology. With respect to 26 security requirements, each electoral innovation causes some to be significantly addressed, and others to be partially addressed or even weakened.

The system architecture being proposed in this research work satisfies the following desiderata:

a) Authentication – through a novel use of voter ID number, voice biometrics (short-term speech features), GPS and geofencing, blinded encrypted keys, and confirmation key to certify uniqueness of the vote.
b) Incoercibility – voter GPS location tracking and control with the GPS location tracking of the place of voting to prevent agenda-driven or disinterested hackers.
c) Accuracy – only one of several identical encrypted ballots gets counted.
d) Privacy – blind signature and data encryption, different servers run registrar and tallier. The pollster does not run on a machine that runs either registrar or tallier. The private ballot confirmation key is used to mediate surreptitious or evident coercion. Personal copy of pollster installed on trusted machine by a voter through anonymous channel.
e) Verifiability – publishing of a list of encrypted ballots, decryption keys and decrypted ballots as well as confirmation key. The voters, through the voter's private ballot, can verify that their votes were counted correctly and correct any mistake anonymously.
f) Convenience – familiar devices and user interfaces, casting of vote in a two-fold engagement, because the voter ID card is affixed with a NFC tag for

auto-loading of voting application and seamless error transfer of data from the NFC tag to the voting application.

g) Mobility – Internet enabled computers and mobile devices enable remote voting within a specified designated precinct mimicking traditional voting practices.
h) Eligibility – NFC-enabled identification is used as authenticating token.
i) Flexibility – XML/HTML, WML and Plain SMS communication, which support deployment on a plethora of devices, have different form factors.

The potency of the SMIV system architecture being proposed in this research work is based on two essential technical properties. First and foremost, it supports the control or tracking of voter location to enforce the incoercibility security requirement. Second, it incorporates state-of-the-art technologies to enhance existing practices, ensure seamless system interoperability, and provide a solid foundation to develop new electronic voting systems. The contributions of this research work to the scientific body of knowledge are succinctly summarized in more detail as follows:

a) The investigation of an approach for SMIV system architecture leveraging the functional capabilities of the state-of-the-art technologies such as near field communication, global positioning system services, and pragmatic voice biometrics, is a unique contribution.
b) The examination of several implementation issues of security requirements of the existing architectural frameworks in order to close a foreseeable gap such as the possible voter coercion, is a novel contribution.
c) The experimental evaluation of the voice biometric authentication component of the proposed mobile Internet voting system architecture is a distinctive contribution.

The original research work reported in this thesis piggybacks on a number of existing techniques and principles to achieve a more pragmatic mobile Internet voting system architecture. The proposed application has a much wider spectrum of applications in electronic voting; for example, large-scale elections, surveys, polls, and small-scale elections such as the student representative and club elections (Cranor and Cryton

1997; Abelkader and Youssef 2012; Baiardi *et al.* 2005; Mauw *et al.* 2007; Fujioka, Okamoto and Ohta 1992).

1.6 Synopsis

The thesis chronicle begins with Chapter 1, which introduces the theme of the research work reported. In this chapter the definitions of concepts germane to this research work are stated and then the security requirements of elections are presented. The general research problems and the associated research question are thereafter discussed. The aim and overarching objectives of the research work are then enunciated. The distinctive contributions of this research work to the body of scientific knowledge are discussed. Chapter 1 of this thesis concludes by presenting the synopses of the thesis to enable a reader to gain a clear picture of the entire contents of the research work beforehand.

The entirety of Chapter 2 of this thesis focuses on the general review of voting systems by detailing the genesis of voting and voting systems in general. This is followed by a comprehensive discussion of the trends of computer voting, Internet voting, and mobile Internet voting. This discussion is deemed necessary to lay a solid foundation for the need to seek a more pragmatic approach for secure authentication in a mobile Internet voting system.

Chapter 3 of this thesis concentrates on secure authentication on the mobile Internet voting system. In the Chapter, several issues concerning security services, knowledge based authentication, possession based authentication, biometric based authentication, and location based authentication are reviewed. The paramount issues discussed in Chapter 3 lay a solid foundation and provide an understanding of the tools used to design a secure mobile Internet voting system architecture.

In Chapter 4 this thesis moves from the discussion on secure authentication in mobile Internet voting system, to the actual design of a secure mobile Internet system architecture. In this Chapter various intriguing components of the system architecture are discussed and compared using a set of security requirements with the state of the art voting system architectures.

Chapter 5 of this thesis introduces the theoretical foundation for the building of secure mobile Internet system architectural components. The literature about suitable components for the system design is large and entangled.

One paramount contribution of this study is thus the review, structure, and duplication of existing theoretical and mathematical characters for the design and development of secure mobile Internet voting system architectural components.

Chapter 6 of this thesis reports on an extensive experimental validation of voice biometric authentication for a mobile Internet voting system. The crucial purpose of this Chapter is to determine and validate a suitable voice authentication model for a secure mobile Internet voting system.

Finally, Chapter 7 of this thesis gives a succinct conclusive commentary of the contributions of the study, as reported in this thesis. Possible future areas of study are suggested to help take the study reported in this thesis to a greater height.

1.7 Delimitations

The development of any electoral system is a massive human, time, and cost intensive exercise. Notwithstanding this fact, the postulation of models and experimental testing and evaluation of parts of the model, actively contribute to the body of the knowledge. It is for this reason that only the voice authentication was tested.

Chapter Two – Development of Electronic Voting Systems

This chapter provides an overview of the development of electronic voting systems from inception, to provide a solid foundation that can help foster a better understanding of the context of elections using electronic means. To understand the roles of electronic voting systems as integral components of the electronic government service, it is important to have a broad understanding of the development of various voting systems. The overview process of voting systems is organised into five core themes in this chapter. They are: genesis of voting, trends in computer voting, Internet voting, mobile voting, and mobile Internet voting.

2.1 Genesis of Voting

The history of voting systems reflects enormous progress made by mankind in advancing technology. In 4 BC, voting in Athens was *viva voce,* i.e. uttered publicly and loudly. Some later examples of democratic voting practices encompassed the *showing of hands* (Rhodes 2004; Sinclair 1991). While these voting styles offered transparency, they introduced vote buying and coercion. In some cases this prevented voters from voting by conscience. They were also difficult to administer and scale. This scaling challenge is perhaps the reason Hermotimus was reputed to have exclaimed: "no I did not count, I estimated" (Rhodes 2004).

In 4 BC, the Athenian Greek voters also used another method of voting by inscribing their choice on discarded pieces of pottery called ostraka, which were placed in an urn and tabulated. Ostraka is the origin of the word ostracise (Saltman 2006; Rhodes 2004; Sinclair 1991; Albright 1942). During the Renaissance period, voting practices included the use of *white balls* for acceptance and *black balls* for rejection of candidates. The balls, called *ballotta,* are the origin of the term, *ballots* (Albright 1942). Elections in India may be traced to 920 AD through the documented processes on stone edicts and carvings. The voters wrote the names of candidates on the palm leaf (Panai olai) and dropped them in pots to be counted, candidate by candidate. The candidate with the highest number of votes was elected (Narasimhan 2012; Nagaswamy 2003). The ballotta, the ostraka, and the Panai olai are examples of the economic sustainability of elections through reusability where the *available* resources were used to record votes.

These methods have enabled voting secrecy, improved counting, allowed for auditing through recounting, and over the years have developed into three main categories of voting devices, namely paper, mechanical, and computer voting.

2.1.1 Paper Voting

The conduct of larger-scale elections became unwieldy to administer as populations increased. The increasing availability of paper, pen and ink[3] spawned the use of the paper ballot. Paper-based voting processes evolved from votes recorded by officials with citizen input or *viva voce.* This approach encouraged nefarious practices by the recording official through deceit, vote buying, coercion and fraud. In order to prevent the electoral officials from recording the votes, the literate voters recorded their votes on *any available* piece of paper. The introduction of party tickets, that are pre-printed listing contesting parties with party icons, introduced convenience and surreptitious monitoring of voters against fraudulent practices. The party agents (vote monitors), along with party officials, were responsible for counting the ballots as voters handed their respective ballots to officials, thereby having a real-time accurate tally of voting in progress. The resurgence of the white unmarked ballot paper in 1856, called the Australian secret ballot paper, which is now colloquially referred to as secret ballot, appeared to solve many of the inherent problems of elections (Saltman 2006; Jones 2003; Albright 1942).

The practice of using paper ballots or mark-sense ballots has proved reasonable for emerging democracies. However, paper-based ballot elections have posed some logistical and administrative challenges, such as difficulty to scale and increasing costs of production. Paper is susceptible to ballot miscount coercion and vote buying. Every form of paper ballot that has been devised has been manipulated with considerable ease (Shamos 2004; Albright 1942). The simplest form is *ballot stuffing,* which is a particular type of election fraud whereby a person who is permitted to vote once, actually submits multiple ballots. This risk was ameliorated over time by specially manufactured government issued ballots that were sufficiently difficult to counterfeit; for example, specialised printing techniques such as

[3] This wasn't always as Socrates, in Plato's Phaedrus, bemoaned writing as a disruptive technology that induces people to cease to exercise their memory and become forgetful (Nehamas and Woodruff 1995).

watermarks were employed to prevent counterfeiting. Notwithstanding these measures, paper voting still makes chain voting possible. C*hain voting* occurs when an attacker is able to obtain a blank ballot (by theft, or other devious methods) on which he marks his candidate choices (Jones 2005). The attacker either convinces or coerces a voter to take the pre-marked ballot to a polling station and exchange it for the blank ballot issued and then return the blank ballot to the attacker.

2.1.2 Mechanical Voting

Fast and accurate tabulation of votes became imperative as the number of voters grew enormously. The processing speed became an essential, although not a sufficient feature in elections. Delays in obtaining the election results had instigated voter suspicion, leading to the development of *mechanical lever machines* that were enormous and unwieldy and used in the USA, and then the *punch card system*. The punch card system is where cards are used to mechanically record votes that are punched contingent to the vote selection, either automatically or by hand (McGaley 2008). This voting system requires voters to mark their ballots by punching holes in paper cards, which are then fed into computerised counting machines either at the local precincts or at the centralised tallying facilities. The punch card development brought about improved tallying speed, automatic ballot count, improved transparency, and mitigated voting irregularities such as ballot stuffing, ballot misinterpretation, over-voting and chain voting. It also reintroduced recounts and auditability of votes to support electoral challenges (Saltman 2006; Jones 2003; Albright 1942).

However, the punch card system presented some intrinsic technical and human challenges, related to card handling, punching and card readers. The event of the year 2000, regarding the *Bush versus Gore* election in the state of Florida in USA, challenged the continued use of the punch card machines in a voting system. A flawed registration system added to the quagmire creating what Justice Wells (2013) referred to as the *perfect storm*. The poor design confused voters in this *Bush versus Gore* election, which led to a famous Supreme Court case and the recounting of the election results. A subsequent Supreme Court intervention halted the recount (Alvarez and Hall 2004; 2010) and Bush was declared the 44th president. This decision was made, notwithstanding that Gore received 544,000 more votes *in toto*

than Bush out of 104 million votes cast nationwide (Saltman 2006). This particular episode illustrates that the design of a ballot paper could influence voting results (Saltman 2006; Alvarez and Hall 2004; 2010). These events provided the specific research impetus for capturing and counting of votes using technologies. The Help America Vote Act (Hava) of 2002 introduced considerable funding (USD 3 billion) to encourage e-voting adoption (Alvarez and Hall 2004; Shamos and Yasinsac 2012).

2.2 Computer Voting

Computer automation of the voting process is widely termed the electronic voting (e-voting) system. The proliferation of computers gave rise to the term e-voting system, which is now formally defined. An e-voting system is an integrated device that uses electronic components to perform one or more of the following functions: ballot presentation, vote capture, vote recording, and vote tabulation (Voting Guidelines 2005). The benefits of e-voting include fast and accurate count; the fact that it is possibly the ultimate green voting solution because of minimal outbound and reverse logistics and attendant carbon saving by minimising voter movement; housebound voting as well as elimination of over votes and spoilt votes; device familiarity; convenience; and the elimination of no-go areas (IEC 2014; Alvarez and Hall 2010).

Traditional computers were found unsuitable as electronic voting terminals, mainly because their operating systems record all operations in a transaction log called a log file. This tendency is a kernel built functionality that is non-trivial to modify (Tanenbaum and Woodhull 2006). It is therefore possible to link a voter to a vote by replaying the transaction log and combining this with publicly obtainable voter queue knowledge (Alveraaz and Hall 2004; 2010). This violates the *privacy* requirement, which states that no one should be able to determine how any individual voted. There is also the inherent challenges of computer voting, such as Denial of Service (DoS), more points of failure as chain-of-custody, increased cost, and perception of outsourcing or privatising democracy (Oostveen 2010; Castro 2011a). In addition, massive voter re-education can introduce a trivialisation of democracy by a point-and-click mentality.

The computer systems currently in use as voting terminals can be classified into three main categories: the optical scan system, the direct recording electronic machines, and the voter verifiable paper audit trail.

2.2.1 Optical Scan System (OSS)

The Optical Scan System (OSS) is an electronic voting system that uses an optical scanner to read marked paper ballots and tally the results. This voting system combines the characteristics of the traditional paper ballot with that of automated vote capture and vote counting computer system. The OSS addresses the security problem of auditability, and it provides speed and accurate interpretation of election results. However, it faces the challenges of grid-connectedness, significant cost, and other complexities such as scanner calibration and technological problems, such as paper jam (McGaley 2008). OSS unintentionally introduced an *undervote*, which is when the voter's marks are either illegible or incomplete, as well as an *overvote,* which is when the voter makes too many marks or the paper ballot smudges. The latter is rejected because the ballot is deemed spoilt (IEC 2012). The percentage of undervotes and overvotes, which is known as the error rate, may thus be high with OSS (Saltman 2006; Jones 2003). OSS requires the ballot to be carefully calibrated for correct interpretation. This negative influence of poor calibration of the paper ballot approach was observed in Scotland in 2007, leading in part to the rejection of over 140,000 (2%) of the votes obtained (Sherrif 2007; Carman, Mitchell and Johns 2008).

2.2.2 Direct Recording Electronic (DRE)

The prevalence of computers in the 1970s led to the development of the first Direct Recording Electronic (DRE). The DRE machines, some of which are the electronic implementations of the traditional mechanical lever machines, are computer-based terminals that allow voters to enter their votes by pressing buttons or touching the images on a computer screen. The DRE system introduced speed and accuracy, reduced costs and transportation of paper ballots, and eliminated overvotes. On the other hand, it introduced *computer-related challenges* such as software bugs, denial-of-service (DoS) attacks, security and *off-grid challenges* such as power and access. However, it reintroduced transparency, auditing and recount challenges. Despite this

there has been a steady adoption of these machines by nations in different contexts (Thakur 2012; McGaley 2008; Oostveen 2007; Saltman 2006; Sinclair *et al.* 2000).

2.2.3 Voter Verifiable Paper Audit Trail (VVPAT)

The Voter Verifiable Paper Audit Trail (VVPAT) system allows a voter to see a printout of the symbol of the candidate to whom he/she has voted. With VVPAT, no computer recount is possible in the event of disputes. Mercuri (2002) suggested a compromise that addresses the concerns of the 'recount on demand' and the 'fully automated' camps. She proposed an innovation to DRE in the form of printing a receipt detailing the vote. This accepted hardcopy counterfoil is dropped into a ballot box. Certain authors (Khono *et al.* 2004; Mercuri 2004) conclude that this is the best electronic voting solution as it allows voter verification. This reintroduces auditability, verifiability and recounts (Mercuri 2000; 2002; 2004). The yin yang journey continues as the print option introduces further costs, another electro-mechanical point of failure, and mandatory real-time support for consumable paper and cartridge change. All modern EVMs now have a VVAT. The twist becomes helical as Castro (2007) claims that paper is no solution to electronic voting. For example, some models print the ballot receipt which the voter (should) place by hand in ballot boxes provided. Many voters took these ballots home for reasons such as souvenir collection, curiosity, forgetfulness, or maliciousness. This has the impact of reducing the reliability of the audit comparison between the computer generated value and the physical count (Thakur 2012; Jones 2003; Saltman 2006).

2.3 Internet Voting

The network of computer networks, called the Internet technology is an information system that brought about Internet voting (i-voting). The concept of i-voting is an electoral system that uses an encryption scheme to allow a voter to transmit a secret ballot securely over the Internet networks (Oostveen and Besselaar 2003). The surge in Internet subscription and increased availability of access points such as computers, mobiles and iDTV, has made i-voting increasingly attractive (Khelifi, *et al.* 2013). I-voting is not just a research topic as Krimmer, Triessnig and Volkamer (2007) list 104 worldwide Internet elections, with 40% being binding elections. I-voting has also drawn interesting queries and e-commerce comparisons such as, "if I can bank

online, why can't I vote online?" Schneider and Woodward (2012) as well as Jefferson (2011) have pointed out that one can check bank statements, allowing one to detect any errors or unauthorised transactions. However, ballot secrecy, which supports the privacy requirement, negates this audit, so the voting system has to be trusted. Jefferson (2011) as well as Simons and Jones (2012) have emphasized the contrary to perceptions of e-commerce transactions that are not safe, pointing in fact that such transactions are highly risky. Banks however accept this risk because of the economic business opportunity. In particular, the voting security and privacy requirements are *unique* and in *tension* in a way that has no comparable analogy in the e-commerce world (Jefferson 2011).

Shamos (2004:13) points out that "altering redundant encrypted write-once computer records is impossible even for experts provided that the Internet voting records are written correctly in the first place." There are several security mechanisms in place to detect such occurrences in real-time with end-to-end verification and audit systems (Benolah *et al.* 2007; Claps and Carter 2013). Almost every Internet voting system has been compromised in a laboratory system, or in a simulated real-time environment (Harris 20313; Thakur 2012). The attacks and errors ranged from software to hardware intrusion, tapping, and tempest-style attacks (Pieters 2009). A tempest attack is what Dutch hackers demonstrated allows voting choices to be known by measuring the electromagnetic radiation of voting screen terminals from a distance. There has been no evidence of fraud that resulted in a court case where electronic rigging was positioned as electoral fraud in an actual election.

The first legally binding i-voting occurred in Arizona, USA in 2000 when voter turnout doubled with young people expressing enthusiasm. The same year, the USA allowed military personnel abroad to vote using the Secure Electronic Registration and Voting Experiment (SERVE). This project cost $22m. However, the momentum slowed significantly after this, due to decentralised elections where counties balance priorities and make individual equipment choices (Done 2002; Kersting and Baldersheim 2004; Jefferson *et al.* 2004a; 2004b).

2.3.1 Mechanism of Internet Voting

The time frame of i-voting is *before* the Election Day, when only site voting occurs, usually paper based. Only eligible voters may vote and *only by themselves*. The voter has to be remotely identified by some detail or token such as ID, pins or eID and associated reader. The voters may not share their details or tokens with any other persons. A voter initially uses this to authenticate and download a client application, which may be used to vote. A voter is allowed to *revote* or *recast* her vote, with the last vote always prevailing and being the only one counted. This diminishes the perceived impact of coercion and makes vote selling futile (Stenerud and Bull 2012), unless the voter 'sold' her token. The revote process is achieved by having a double digital envelop – the other having the voter's name and the inner containing the vote. The outer vote is encrypted with the user's digital signature. If a voter revotes, the envelope containing his/her vote is overwritten, with no consideration to the contents, together preserving the privacy and the uniqueness requirement. If a paper ballot is cast, the voter digital envelope is revoked or deleted with the i-vote having no impact on the results. At the close of polls, the ballot server is disconnected, paper and revotes are reconciled. The outside envelope is removed to eliminate vote trace and contents shuffled and a copy sent to a counting server, which tabulates the result. No revote is allowed on Election Day, mediating last minute Denial of Service (DoS) surges in support of the soundness requirement. There has been no prescription put forward as to the maximum number of revotes in implementation, to the author's knowledge.

One of the open debates regarding I-voting is whether it facilitates social cohesion and community networks, or whether it undermines them by replacing the face-to-face interaction with online interaction (Internet Policy Institute 2001). I-voting is potentially vulnerable to challenges such as *software:* spoofing, automated vote buying, viral attacks on voter PCs or servers, and human errors; *hardware* technology failure whether contrived or operational; or *societal*-human errors induced or reality (Larcom and Liu 2013). Award and Leiss (2011) add an attacker disrupting or shutting the network down with either viruses, worms or spoofing. Adida (2008) suggests the possibility of cyber-war attacks through viruses like Stuxnet used to cripple Iran's nuclear plants. The problem is the uncontrolled platform where voting software or computers may be infected (Healy 2014; Simons and Jones 2012;

Kushner 2013). Ross informs of the 'weaponisation' of the code, and states that cyber offence is easier than cyber defence, referring to the fears of cinema theatres over the Sony movie, The Interview. Sony pulled the movie off the cinema circuit because hackers broke into Sony's computer system, stole private information, and then leaked it online (CNN 2014). Adidia (2008), the developer of Helios[4] - the first web-based, open-audit voting system, pragmatically cautions that all the verifiability doesn't change the fact that a client side corruption in the browser can flip the vote even before it's encrypted (Simons and Jones 2012; Adida 2008). Helios provides strong integrity guarantees and uses a series of voter challenges. We used this in a previous model (Thakur *et al.* 2015). This is an explicit mitigation of this studies migration to the choice of the cast-as-intended SMS.

The Washington DC State developed an Internet system to help facilitate the citizens' interactivity. It had a unique public trial, in the form of a mock election where the public was asked to attempt to compromise its security. A team led by Halderman from the University of Michigan compromised the system, forcing its cancellation. The team found and exploited a Linux shell vulnerability that gave them almost total control of the server software, including the ability to change votes and reveal voters' secret ballots (Wolchok *et al.* 2012). A neutral analysis suggests the public test was a good act of transparency. That the flaws were detected shows the need for such piloting and testing. The developers acknowledge the errors, but point out that they could be addressed. Puiggali, an i-voting software expert developer, asserts that: "Internet balloting has not been perfect, but counters, we have to consider the risks of voting channels that already exist" citing practices such as stuffed ballot boxes (AFP 2014: website). Shamos (2004) alludes to a history of paper ballot fraud. Kelleher (2013) concurs arguing while no voting system can ever be completely secure, this does not mean one should therefore not have voting at all.

Given these critical issues, i-voting authentication procedures must be more secure than the traditional approaches if the adoption of the approach is to be promoted for general elections. Currently the voter registers, either online or in-person, for credentials that allow them to vote online by providing them with a user name and a password for extra security on the system. However this approach to ensuring the

[4] http://heliosvoting.org

security of an i-voting system is insufficient and may be compromised through several ways, such as key theft and malware (Award and Leiss 2011; Alvarez and Hall 2004).

2.3.2 Types of Internet Voting

I-voting tends to improve convenience for voters, positively impacting voter turnout, enhancing accessibility, stimulating youth involvement and mobility, and (arguably) reducing costs (Pammett and Goodman 2013). There are four main types of i-voting system, shown in Table 2.1 (Alveraaz and Hall 2010; Pammet and Goodman 2013).

Table 2.1 Types of Internet Voting (Alvarez and Hall 2004 2010)

Type	Description	Merit	Demerit
Polling Place Internet	Voting from an area under the direct control of an authority	Infrastructure under the direct control and supervision of the EMB	Requires infrastructure to be replicated across entire country without exception
Precinct Internet Voting	Voting from a particular voting district, ward or precinct - for example where the voter resides, through the Internet	Geofencing of the voter promotes local social cohesion; mediates denial-of-service attacks; assists in post-election audits	Geofencing requires mobile devices or an Internet enabled device. Authentication and coercion becomes a challenge
Kiosk Internet Voting	Voting from a dedicated terminal not under the direct control of the authority. This may be at an office or shopping mall but also through the Internet[5]	The public nature may take democracy to the people in familiar settings. It is connected to a closed, controlled network	The absence of the authority may encourage some voters to try to hack the system. Authentication and coercion becomes a challenge. The placement of the kiosk may distract and divert the voter to other activities
Remote Internet Voting	Refers to using a computer that is *not under the physical control* of the authority officials; which is used to cast the ballot over an Internet connection	The mobility offers the voters anyplace anytime convenient voting	Exposes the process to the entire Internet, where hackers may attack a system because it is there; authentication and coercion becomes a challenge

(Sources: Alvarez, and Hall 2004; 2008; Pammet and Goodman 2013)

[5] This may be compared to an ATM still owned and operated by a bank but largely unmanned.

There are thus differences amongst the forms of i-voting, with each type offering different challenges for voter authentication (Mossberger, Tolbert and McNeal 2008). This study focuses on the remote Internet voting type, with a special focus on mobile devices from whence mobile Internet voting emerges.

Mobile Internet is an important component that enables ubiquitous voting because it delivers the Internet to voters who do not have access by other means such as mobile data services. Some literature calls this Mobile Phone Voting System (MPVS) (Kogeda and Mpekoa 2013; Ullah and Umar 2013) while others called it ubiquitous voting (Abdelkader and Youssef 2012). The ability to eliminate voting restriction is gratifying for increasing participation rates, increasing voter satisfaction, simplifying the voting process, and saving time and costs of conducting elections.

The discussion thus far allows one to comprehend and classify e-voting systems through the analysis of incremental improvements. This study now investigates mobile Internet voting, which is a special type of remote Internet voting system, belonging to the genre of uncontrolled e-voting.

2.4 Mobile Voting

The proliferation of mobile devices in different forms and their relatively low costs are alluring features for exploring mobile devices for voting. Poushter and Oates (2015) assert that 9 in 10 South African[6] citizens have a cellphone, which explicitly provides access. Simply stated, mobile voting is the act of voting via mobile devices. It enables *mobile democracy*, which may be defined as using mobile interfaces to improve the relationship between the government and its citizens; it connotes a move toward a more inclusive and participatory democracy. Of course it would be an exaggeration to claim that democratic ties between the government and its citizens may be strengthened only with the help of mobile communication devices (Brücher and Baumberger 2003). Mobiles are also referred to as Common-off-the-Shelf (COTS) devices because they can be purchased from a store. The mobile is fortuitously imbued with useful technologies such as computation capabilities, wireless, voice acquisition, GPS mobile tracking, and NFC reading, which this study's architecture opportunistically incorporates for multi-modal voter

[6] They rate cellphone penetration at 89%, with smartphone penetration standing at 34% penetration.

authentication in our enhanced voting architecture. Weilenmann and Larsson (2002) point out that the mobile phone itself is also shared between friends, who borrow and lend each other's phones, or even share a phone with unknown others, with the purpose of making contact. This increases the value proposition for Mobile voting.

The introduction of mobile voting reintroduces traditional benefits associated with Internet voting, such as increased mobility, ubiquitous voting, ease-of-use and convenience (Thakur *et al.* 2015; Weilenmann and Larsson 2002). That said, the ballot design becomes a challenge when there is a high number of candidates because the mobile user interface becomes cumbersome. In addition, the small form factor of a mobile device, which is a key contributing factor to its ubiquity, makes data entry problematic and sometimes makes interactivity difficult. This can introduce erroneous inputs, making the use of multiple passwords extremely difficult (Ekong and Ekong 2010).

Yet Campbell *et al.* (2014) found *otherwise;* they found that in fact smartphone owners committed fewer errors on a mobile voting system than on traditional voting systems, and even enjoyed the queue-less engagement. They used a sample of 88, of which 48 owned smartphones, and found that there were no reliable differences between the smartphone-based system and other voting methods in efficiency and perceived usability. More important, though, the *same* 48 smartphone owners committed fewer errors on the mobile voting system than they did on the traditional voting systems. From this one can infer the power of familiarity.

Brücher and Baumberger (2003) astutely allude to mobiles providing a form of mDemocracy and even show how mDemocracy enhances eDemocracy, which also helps contextualize e-voting and m-voting. They assert that mDemocracy supports and completes eDemocracy by supporting:

1) *Infrastructure* as most citizens are connected. This effort postulates that this position may free or redirect resources towards unconnected areas,
2) *Media capability* which suggests device familiarity also exists
3) *Location and time independence.*

2.5 Trends in Modern Voting

Besides the trends in computer voting from an academic research perspective, there is a growing trend across the world to practically use advanced e-voting systems to conduct general elections. The goal of this reform slightly varies from one country to another. In some countries the drive is to increase electoral turnout rates, while in other countries the purpose is to reduce electoral malpractices. The common goal, however, is to improve the democratic process by making the voting process more cost effective, convenient and secure. In this section of the thesis, different practical implementations of computer voting systems are discussed by contextually comparing South Africa's view of computer voting with that of two other countries. The review of practical trends of voting systems, when combined with academic perspectives, provides a realistic understanding to inform the design of a secure voting system architecture.

2.5.1 Estonia Perspective

Computer based voting using Internet technology was carried out seven times in Estonia between 2005 and 2014, in local elections, parliamentary elections, and European parliament elections. The voting is permitted from the 10^{th} day of commencement to the 4^{th} day before the final poll day when in-person voting prevails. To vote, Estonians put their electronic identity card (eID) into an appropriate reader attached to their computer and enter passwords (Estonia Authority 2014a; Alvarez, Hall and Trechsel 2009; Kalvet 2009).

The voting system starts and displays the registered candidates in random order of their political parties. The voter makes a selection and the system encrypts the voter's choice. The encrypted vote is opened by the EMB using a private key after the election. A private key is a secret key used for decryption (Estonia Authority 2014a). The encrypted vote may be regarded as a vote contained in an anonymous inner envelope. The voter issues a digital signature to confirm selection. By digitally signing, the voter's personal data or outer envelope is added to the vote encrypted. To ensure that voters are expressing their true will, they are allowed to change their electronic votes by voting electronically again during advance polls, or by voting at the polling station during advance polls (Heiberg and Willemson 2014; Estonia Authority 2014a). Before the ascertaining of voting results during the

Election Day, the encrypted votes and the digital signatures, which identify the voter, are separated. Then anonymous computer based votes are opened and counted. The system opens the votes only if they are not connected to personal data. This study reported in this thesis adopts this approach because it mediates coercion, vote selling or buying. The principle of incoercibility is satisfied with this approach because a voter under duress has the opportunity to change his/her mind at a later time, when the pressure is relaxed.

The time interval between the 4th and the Election Day is necessary in order to ensure there is sufficient time for authorities to eliminate duplicate votes or revotes. For example, if a voter cancels her vote by voting at a polling station, it must be guaranteed that only one vote is counted per voter. To that end, all polling stations are informed of the voters on their list of voters after the end of advance polls and before the Election Day. If it is found at the polling district that the voter has voted both electronically and with a paper ballot, the information is sent to the electronic voting committee and the voter's computer vote is cancelled retaining the uniqueness principle. This time interval may be exploited to develop responses to failures, which supports the respective principles of reliability, robustness and scalability. This thesis consequently recommends repeated vote as an important inclusive step in *any* computer based voting model because it is a good incremental step towards a fully automated election, allowing the electorate a choice of computer voting and then the traditional paper voting, retaining transparency. Most Internet voting schemes proposed in the research literature use cryptographic techniques to achieve the end-to-end (E2E) verifiability. Instead Estonia uses a conceptually simpler design at the cost of having to implicitly trust the integrity of voters' computers, server components, and the election staff (Springall *et al.* 2014).

Estonians supported the computer-based voting because they believed the country's paper-based system to be fraught with widespread fraud (Springall *et al.* 2014*;* Heiberg 2013). Heiberg and Willemson (2014) explain the extension to the Estonian computer-voting scheme, which allows voters to check the cast-as-intended and recorded-as-cast properties of their vote on a mobile device. The scheme was used during the 2013 Estonian local municipal elections and the 2014 European Parliament elections where 3.43% and 4.04% of all Internet votes were verified respectively (Estonia Authority 2014b). The Estonian system supports vote auditing by releasing

the randomised seed that is used for encryption. This study accepts this approach, but acknowledges that as the mobile is the primary communication device used in this work, which coincidently is the medium used by the short messaging service (SMS), it is still reasonable because SMS or e-mail is *another* out-of-band communication band channel method for improved security. This provides the need for simplicity to sustain the principle of accuracy and completeness.

2.5.2 The Norwegian Perspective

Norway started Internet Voting pilots in 2011, with the purpose of increasing the availability of voting systems and reducing costs in the long term. The pilots were discontinued in 2014 over privacy and authentication concerns. Norway, like Australia (Zetter 2003), published all documents, including architecture and software source code to build trust and support transparency. The country received widespread acclaim for the use of verifiable end-to-end cryptographic voting protocol, which allowed third parties to perform robust audits of the count. This achieved a high level of voter trust (Chowdhury 2013; Tallinn 2014; Spycher, Volkamer and Koeni 2012). Norway uses eID and a fingerprint reader that the voter must alternatively purchase or go to the poll-site to vote. The voting system requires the voter to download a Java applet, the voter client, onto the user PC, which possibly exposes the machine to malware. The voter verifies herself using MinID a free two-factor authentication mechanism. She is presented with a graphical interface representing the ballot. The applet encrypts and digitally signs the voter's selections and sends it to the central-voting servers (Stenerud and Bull 2012).

Norway allows computer voting for a whole month mimicking early voting practices. Each voter is given their own unique set of random codes for the different candidates in the election. Once a vote was cast, the code is sent via SMS back to the voter, to allow comparison with the SMS code sent to their phone with the printed code on their voting card. This is called *return code,* which is computed before an election day for verification of the cast-as-intended vote. Voters may vote as many times as they desire between the 10^{th} and 6^{th} day before the election, and may even go to a poll-site, overriding all e-votes (Stenerud and Bull 2012; Øyvann 2013). The return code method was first presented with Chaum's SureVote (Chaum 2001). Return code aims to avoid requiring the voter to trust any computational device or digital signature

in order to vote, improving transparency and understanding of the process (Ryan and Teague 2013). The SMS that is sent uses a different path, known as out-of-band, for additional security for the return code. The user may check and compare her vote. This method provides verifiability and may enhance integrity and trust. The coding sheets are sent to voters via ordinary mail or other trusted channels. These codes have random vote codes against each candidate. This in effect is a private vote codebook that assures votes are cast-as-intended (Stenerud and Bull 2012; Øyvann 2013; Ryan and Teague 2013; Heiberg and Willemson 2014; Barrat *et al.* 2012).

This return code is very much an electronic equivalent of a Voter Verifiable Paper Audit Trail (VVPAT). The return code is created using two lists - the voter list and a greater or equal number of return code sets. Each set has a list of all parties and their corresponding 4-digit return code and a unique ID (Stenerud and Bull 2012). A second independent process takes the return code list and folds it. This is randomly ascribed to any voter and the voter data is printed on the outside, but the inside is invisible. The binding now takes place between the unique ID and the voter. This is uploaded unto the server to the component responsible for sending return codes by SMS. This process ensures no person or component can determine the meaning of the return codes of an individual user. This helps the voter at an individual level, while if a sufficient number verify their votes, it improves overall confidence in the total votes and the process is in agreement with the principle of dispute freeness.

The main purpose of the approach described was to address a Norwegian system requirement specification which stated that: “Even though the e-voting client system may be under outsider control, the e-voting system shall be such that it is not feasible for an outsider to manipulate votes without detection” (Stenerud and Bull 2012:23). Stenerud and Bull (2012) point to an election with 40,000 ballots cast and a manipulation rate of just 1%, and assert that the chances of detecting the manipulation are more than 90% if just 230 voters verify. If 2% of the voters verify their ballots, the same manipulation is detected with a probability of more than 99.9%. Positive side effects were that voters found the return code confidence inspiring, the Electoral management board (EMB) found that it built trust, and this approach was welcomed by disabled voters.

Nonetheless, Norway discontinued Internet Voting pilots in 2014 due to concerns over privacy and authentication. Norway has a strong tradition of consensus on all matters regarding electoral policy. The electronic voting system uptake was not very high. Despite perceived successes, there was political controversy over fears that the security of the revote was insufficient and fears that allowing votes outside the poll-sites can diminish the credibility of the vote (Norway Ministry 2014). Despite this limitation, this study embraces the concept of return codes/renamed confirmation code, albeit differently as explained in Chapter Four. This is due to the end-to-end verifiability comfort that is provided to the voter.

2.5.3 South African Perspective

In this section of the thesis, the situational overview of voting system in the South African context is discussed. South Africa as a country is a young democracy with a population of 53 million; the country's democracy started in 1994 following the historic demise of apartheid. It has a registered voting population of 25.4 million, out of a voting age population (VAP) of 31m; implying that 5.6m (18%) have not registered to vote. In the 2014 elections, 18.7m (73.5%) participated. This means that 6.7m registered voters did not vote, suggesting a non-participation rate of 12.3m. This percentage is significant when compared to the VAP of 31 million. Increased participation may therefore be a driver for the introduction of an electronic voting system to improve turnout. A phenomenon of interest is that in each of the last three general elections in the years 2006, 2009 and 2011, young South Africans constituted the two largest blocks of registered individuals and reflected a sustained rise in registration numbers that was observed across most age groups. For example, the 20-29 age group followed this trend, and in 2011 there were 9% more registered South Africans in this age group (5.53m) than in 2006 (5.08 million) (Scott *et al.* 2012; Schulz-Herzenberg 2014). This commitment arguably demonstrates intent to vote.

However, despite this rise in the number of young registered South African voters, less than half of the country's youths voted in 2014. This comparatively low level of young people's involvement in the electoral process is even starker when considering the voter turnout for the 2011 municipal elections as a proportion of the country's population as a whole. During that period, only 28% of the country's total 20 to 29-

year-old population participated in the elections. Correspondingly, only 37% of the 30 to 39-year-old group voted in 2011 (Scott *et al.* 2012; Schulz-Herzenberg 2014).

The Electoral Commission of South Africa (IEC) is charged with administering elections in South Africa and is perceived to be neutral if one looks at the lack of legal challenges. It has a neutral position on e-voting although it is acutely aware of e-voting, having commissioned an investigation (Thakur 2012) as well as hosting an international seminar on the subject (IEC 2014). The IEC used Prosser and Krimmer's (2004) four dimensions of Electronic Voting, namely Technology, Law, Politics and Society, as well as Krimmer, Schuster and CC's (2008) e-voting readiness index to neutrally assess e-voting in the country. The IEC's conclusion was that SA had the capability, infrastructure, and political maturity to implement e-voting although it has several equally important competing priorities, *inter alia* infrastructure, healthcare and education (IEC 2014).

South Africa has experience with multifactor recognition using voice biometric and other methods on a mass-scale having conducted a voice biometric registration of social grant beneficiaries, which we describe later and use in this study. This proof-of-life test largely resulted in 650 000 people being removed from the social grants system, saving the nation's Treasury R2 billion. It also allowed the indigent to receive their grants in an exponentially more convenient and secure form (Agnitio 2014; Top 2014). It would appear that there is a youth appetite for m-participation with Cupido and Van Belle (2012), in a survey of 131 respondents between the ages of 18-35, finding that there was an appetite for voting by mobile.

Contextual research on voting modalities will help inform the IEC, and indeed other similarly positioned electoral authorities. In the sections that follow, trends in Internet voting, mobile voting, and mobile Internet voting, as well as their mechanisms and types, will be discussed.

2.6 Mobile Internet Voting

Mobile Internet voting (MI-voting) can be defined as ubiquitous voting using mobile devices to access voting services on the Internet. Distributing the processing of votes over many web servers that are installed in tamper-resistant manner provides an environment that can improve security. The security process is made possible using

the Smart Card Web Server (SCWS) on a mobile phone Subscriber Identity Module (SIM) (Kyrillidis *et al.* 2012). In general, mobile Internet technology is the result of the convergence of networks of the traditional Internet technology, broadband mobile networks, and the mobile terminals. The mobile Internet can play an important role to communication by taking the advantages of a large user base, the surging sales of smartphones, tablets and 3G data, and exploration of other mobile electronic devices (Juan and Shoulian 2010). This research adds democracy to the communication advantage. Restated in the simplest terms, mobile Internet refers to ubiquitous access to wireless Internet services at anytime using mobile devices. Mobile Internet technology is a rapidly growing, emerging form of networking and a typical representative of convergence of existing networks (Yuan-Yuan, Jie, and Zhen-Ming 2013). The convergence of the traditional Internet, mobile networks, wireless networks, smart devices and mobile intelligent terminals can be attributed to the development of mobile Internet connectivity.

The deployment of MI-voting systems for elections has the capability to boost participation amongst wider audiences and make the voting process more convenient (Cupido *et al.* 2012; Akilli 2012; Hermanns, 2008; Brücher and Baumberger 2003). It is a common practice nowadays to access the Internet and the Web from mobile phones, especially because the usage fees for access to data bundles has dropped considerably, to a point where anybody who can afford a mobile phone can equally afford mobile Internet access. Almost every mobile phone today comes with an integrated web browser that can display HTML web pages and execute JavaScript. In the past mobile web access was purely limited to the domain of high-end smart phones because only those devices had the processing power to render and display HTML content, and data bundles were expensive. The cheaper feature phones were mostly bound to the Wireless Application Protocol (WAP) as part of the Internet technology (Bao *et al.* 2011; Zakas 2013).

The current theories and practices of elections, apportion the main expenses to voter registration, boundary delimitation, voting operation, counting and transmission of results, dispute adjudication, voter education and information, campaigning by political parties and candidates, and oversight by party representatives and observers (López-Pintor and Fischer 2006). Researchers have estimated the operational cost of elections at between $1 and $3 per voter (López-Pintor and Fischer 2006; Clarkson,

Chong and Myers 2008). This estimate was arrived at based on the cost of the New York i-voting model. It was found that the cost to develop a remote voting system is similar to the current costs; this includes a once-off capital equipment cost after which costs decline per election.

MI-voting will satisfy the electorate with any time, any place, queue-less voter convenience, time-saving and device familiarity, and these benefits are more likely to encourage than discourage voter participation. Economic migrants who prefer to take part in hometown elections will also be appeased (Dave *et al.* 2008). MI-voting also strengthens security as hackers may be disinclined to attack individual mobile devices that are geographically distributed. Needless to say, the attention of hackers seeking to launch attacks may be shifted to base stations or to the server side. This mitigates authentication, trust and confidence. However, MI-voting re-introduces the coercion challenge that is often associated with traditional i-voting. The problem of coercion can manifest in uncontrolled MI-voting because it is possible for a single politician to pay voters to vote at a particular location. Mobility generally increases coercion, which may be difficult to detect and control, especially in an uncontrolled voting system. The Norwegian model of allowing repeated votes before Election Day might mitigate coercion (Heiberg and Willemson 2014) as voters have the chance of changing their mind at a later time, but this is not a sufficient solution.

2.6.1 Mobile Internet Voting Research

MI-voting is evolving as a research endeavour and is attracting the attention of researchers. Ekong and Ekong (2010) proposed a mobile voting prototype using Wireless Markup Language (WML) and Hypertext Preprocessor (PHP). They refer to Nigeria as a country having typically 50 parties in an election and suggest that mobility becomes a pertinent challenge when a large number of parties compete in a general election. Sandler, Derr and Wallach (2008) developed VoteBox, a tamper-evident, verifiable voting system, in which all e-voting terminals (DREs) are connected to a LAN, but disconnected from the WAN or Internet. All critical events are broadcast locally so any observer with a computer can watch and discern *local* discrepancies in real-time. This regrettably does not scale to mobile voting, because of the need to broadcast events as they occur. Verifying remote votes as they occur is non-trivial even if each and every one is broadcast. WAN voting also introduces

possibilities of monitoring and interception because of the nature of the Internet communication.

DynaVote (an e-voting protocol implementation) asserts that it is secure and practical over a network and fulfils core requirements such as privacy, eligibility and accuracy and does not require anonymous channels such as Mix-Net. DynaVote subscribes to the researcher's design principles of not using complex algorithms such as homophobic encryption. It uses a Pseudo-Voter Identity (PVID) scheme and relies on blind signature (Cetinkaya and Koc 2009).

Helios is an implementation that allows a system to be randomly audited in real-time. It is also called the Benaloh (2007) challenge. Thakur *et al.* (2015) proposed this in an earlier model. This occurs in the voting cycle, after a voter has made a selection *and* it is encrypted for transmission. The voter may challenge the vote and decrypt it to determine if the cast vote is encrypted. The challenged vote is discarded. The voter must vote again and will be presented with the same challenge or cast option. A coercer or malware installer has no idea if the voter will accept or challenge the vote - reducing the value of the effort (Popoveniuc *et al.* 2010). Schneider and Woodward (2012) suggest that cryptographically based systems can be safely used, and offer truly verifiable democracy citing work on Helios and Prêt à Voter[7]. Prêt à Voter is a scheme that preserves the secrecy of the vote by creating a non-repudiable link between the voter and her cast ballot. This link is not available to a casual voter. This method is backed by mathematical proofs that the votes have been processed, decrypted and tallied correctly, which is beyond the scope of this study. A variation of this method, together with the Norwegian out-of-band transmission, is used for verification instead.

Elleithy and Rimawi (2006) propose the CyberVote system; it uses a cryptographic protocol to ensure voter authentication through vote capture, transmission, counting and auditing. CyberVote suggests many features now standard in i-voting projects, such as resetting the database at the start of the election and closing the system to voting at the end, with complete denial during counting for protection. Some systems

[7] http://www.pretavoter.com

(such as Estonia and Norway, as described in 2.2.1 and 2.2.2) use separate, disparate systems for counting.

Kim and Hong (2007) assertively suggest that a voter be identified using a wireless certificate without additional registration when a user votes using her mobile terminal, such as a cellular phone or a Personal Digital Assistant (PDA). This places more emphasis on the device authentication process and leverages the triple - of device, wireless certificate, and voter.

Dave *et al.* (2008) also agree in part and suggest that a wireless certificate be issued in advance to the mobile-id. They recommend a voting card, which has the index of the candidates. A message is sent to the GSM in the format <voter-id> <candidate-id> <mobile-id> over the GSM. Therefore 033 02 919981360643 is a vote by voter number 33 for candidate 02 by mobile 919981360643. The votes are stored till the end of the election to allow for reconciliation of duplicate votes.

Similarly, Khelifi *et al.* (2013) suggest a secure mobile system, m-vote, that uses three levels of security - username and passwords; national ID; and fingerprint. This, they say, improves local device security and validation which prevents unauthorised access, as well as mediating attacks by external hackers. This multifactor method is selectively leveraged using a token and voice.

EVox is a working prototype with promise (Herschberg, 1997) that was superseded by Robust Electronic Voting System (REVS), which addressed sustained DDoS attackers from malicious colluding servers. REVS was designed for distributed and faulty environments, namely the Internet (Joaquim, Zúquete and Ferreira 2003), and is consequently selected for deeper analysis in Chapter 4. The source code of REVS is available (REVS 2015).

Clarkson, Chong and Myers (2008) describe Civitas, a coercion-resistant, universally and voter verifiable prototype, suitable for security analysis. This prototype is computationally intensive, suggesting it will not scale easily (Mursi *et al.* 2013).

Kumar *et al.* (2011) proposed a GSM mobile voting scheme which relies on mobile service provider authentication infrastructure to enhance security, provide authentication, improve voters' convenience, and support mobility.

Gentles and Sankaranarayanan (2011; 2012) examine India's voter convenience and poll-site e-voting, where 1.4 million machines have to be deployed into consideration. Based on this experience, they consequently, propose a biometric, authenticated, mobile voting system for Jamaica. Their technology proposes that using fingerprint supported biometric control information and encryption, along with Secure Socket Layer (SSL), would make the software involved in the voting process well secured. In addition, they bind (tie) the credentials to a mobile device to make the system even more robust (Gentles and Sankaranarayanan 2011; 2012). This idea is therefore incorporated into the proposed model. In particular, the mobiles' functionality is increased by converting the device into a voting terminal, which could be used by more than one voter, by leveraging the following technology NFC, GPS and Voice technology.

Ahmad, Hu and Han (2009) suggest that conventional methods use symmetric encryption algorithms, or hybrid symmetric and asymmetric algorithms, at the expense of weaker security strength. They propose elliptic curve cryptography (ECC) algorithm to secure votes in a mobile voting scheme.

Ullah *et al.* (2014) help to reduce the number of steps in m-voting by proposing a cost effective, secure, mobile phone voting system using blind signcryption. This method includes digital signature and encryption in one step. They claim this satisfies privacy, anonymity, integrity and untraceability.

Olaniyi *et al.* (2013) developed a multifactor authentication and integrity system incorporating cryptographic hash function methods. During registration a unique voter-ID is generated, along with a unique grid card. These are linked. A SMS pin is also generated and sent to the voter. The voter uses the SMS and the grid card to vote, which is tied to the unique Voter-ID. The vote, x is cast. The casted vote is encrypted, e(x) and sent to the server. The ballot is also hashed, h(x) and sent to the server for post-election processing. At post-election, the encrypted vote (which should be e(x)) is decrypted and hashed. If it is the original vote, it should yield h(x). Any variation implies integrity violation.

Ayo, Daramola and Azeta (2009) described an integrated e-voting system comprising: the electronic voting machine (EVM), Internet voting (i-Voting), and mobile voting (m-Voting). It examined issues of interoperability of the integrated

system as well as the need for security measures. The authors recommended that emphasis be directed at EVM for in-country voters, while certain classes of voters (citizens living abroad or living with certain deformities) are restricted to special cases of remote voting.

Some commercial companies also offer i-voting products; these include Skytl, Adapt-It and Smartmatic. The reader is referred to Thakur (2012) for a comprehensive list of suppliers of e-voting equipment.

Ayo (2009) speculated that increasing the medium of voters' accessibility with mobiles will provide an alternative platform for the non-physically challenged voters, while also supporting disabled voters, which will drive increased participation. A prototype system was developed for experimentation, which examined and reported on the prospects and challenges of voice voting in Nigeria. The results of this study would provide insights into ways of improving participation of voters in general elections in Nigeria and other democracies.

2.6.2 Mobile Voting System Testing

Kohno *et al. (*2004) suggest certification is an important process to maintain trust. It is also needed to pacify adversaries. Some critics demand full code transparency or an *open source* solution. The private developers of the Australian i-voting project, as a tender requirement, placed the full source code online. This improved trust especially among expert stakeholders, who reviewed the code and found two errors (Thakur 2012). They caution that one must understand what certification means. Hall (2008) pragmatically notes that some view software as intellectual property. He points that even if the code is released, only experts would understand it. He therefore recommends that the code be released to experts who assess and certify this. This is called *white box testing*. On the other hand, *black box testing* occurs when the functionality of the voting process is tested. Here the code is usually not released. Both forms of testing are non-trivial and require a careful evaluation against general information technology principles or availability, integrity and confidentiality, using Mursi *et al.'s* (2013) 26 requirements. As biometrics is used, one may test against Jain *et al.'s* (2004) biometric principles as well. Volkamer (2009) points out that it is

difficult for a voter to validate the integrity and authenticity of the software if she has to install it.

It is worth noting here that opponents of e-voting also improve system integrity. Thakur (2012) matter-of-factly pointed out that just as hackers make computing safer as they find flaws, so too does the anti-voting lobby. One must respectfully add that these lobby groups are reasonable, articulate and measured. If one considers their objections one can build systems to mediate these, to the system's benefit. In the USA Black Box Ballot leads the campaign against e-voting. In Europe the lobby against e-voting is coordinated by European Digital Rights member (EDRi) member Chaos Computer Club; the Dutch foundation *Wij vertrouwen stemcomputers niet* (We don't trust voting computers); and the Irish lobby group, Citizens for Trustworthy Evoting (ICTE). These lobby groups exist because of the risk of electronic errors and the potential for abuse. The group in India that leads the anti e-voting lobby is called Citizens for Verifiability, Transparency and Accountability in Elections (VeTA) (Narasimha Rao 2012). A VeTA computer scientist exposed serious vulnerabilities in India's EVMs in 2010 (Prasad *et al.* 2010). Prasad *et al.* (2010) argue that counting is not repeatable, EVMs can be hacked and no auditing is possible of the EVMs in India.

2.6.3 Mobile Internet Voting Architecture

The concept of reference architecture was defined in the Rational Unified Process, as the predefined structural pattern or set of patterns, possibly partially instantiated, designed and proven for use in particular technical contexts together with supporting artefacts that are harvested from previous projects (Reed 2002). A reference architecture does not represent the whole of the system; it allows for contextual improvement and innovations, which may be selectively crafted from the past to improve the future (Cloutier *et al.* 2010). More importantly reference architecture provides a proven template of solution and the lexicon or common vocabulary with which to discuss implementations. A reference architecture helpfully provides the context to craft the secure MI-voting system based on the existing reference architecture, such as FOO, Sensus, and REVS.

This study will examine these three voting architectures, namely: FOO, a practical, secure, large-scale voting mechanism that addresses the privacy and fairness security requirement (Fujioka, Okamoto and Ohta 1992); Sensus, which incrementally improves on FOO by partially addressing or fully addressing more security requirements such as verifiability, accuracy and invulnerability (Cranor and Cryton 1997); and finally REVS, which is a more recent development that mitigates Internet related failure and challenges, as well colluding agents who practice electronic ballot stuffing (Joaquim *et al.* 2003; REVS 2015). It then presents the SMIV proposed architecture, which amongst others introduces a secure remote authentication and voting architecture. The benefits of SMIV have already been presented in the Chapter One, the introduction.

2.7 Conclusion

This chapter traced the evolution of the development of electronic voting systems. It noted the opportunity that each successive iteration has brought to the fore, while simultaneously pointing to the challenges these opportunities introduced. To wit, voting started with in-person voting, and moved to mechanical voting, and onto computer voting, which comprised OSS, DRE and DRE/VVAT. Thereafter, this chapter then presented Internet voting and the types of Internet voting, along with a comparison and a short case study of Estonia and Norway, before presenting the South African perspective. MI-voting, was then presented and a survey of the different approaches used was presented together with a review of research and testing. Finally, MI-voting Architecture was presented.

Chapter Three - Secure Authentication in Mobile Internet Systems

There are several reasons why authentications in MI-voting systems in particular, and in ubiquitous information systems in general, are difficult to accomplish. These reasons have not been well covered in the explanation of the conventional security requirements for voting systems (Mursi *et al.* 2013). The security requirements for mobile Internet systems go beyond the conventional security issues because of the inherent challenges brought about by mobility. The physical outreach nature of ubiquitous computing systems and devices makes the task of preserving the security of users more challenging. The set of users in the ubiquitous space affects the security properties of the space. Since the nature of group interactions between various users cannot easily be prevented within the space, this dynamism has to be taken into cognisance when designing security mechanisms. As a result, the community of ubiquitous computing researchers considers security to be amongst the biggest challenges to the vision of computing environments (Kagal *et al.* 2002; Zakiuddin *et al.* 2003; Kagal, Finin and Joshi 2001).

The notion of mobile Internet computing derives from pervasive or ubiquitous computing, which refers to the proliferation of many computing devices, sensors and embedded microprocessors that provide the latest computing functionalities. These functionalities also provide specialised services to enhance productivity, and to facilitate seamless interactivity with the surrounding environment and available resources. Ubiquitous computing allows for the realisation of added abstractions that did not exist in traditional computing paradigms. It extends computing boundaries beyond the conventional environment of hardware and software resources to include the ecosystem of physical spaces, smart spaces, building infrastructures, and all devices contained within the ecosystem. The ubiquitous computing is aimed at transforming the dull, passive spaces into interactive, dynamic and programmable spaces that are coordinated through a software infrastructure and populated with a large number of mobile users and smart devices. The ubiquitous computing environment is mobile as is its users, and it is able to adapt to environments with scarce computing resources, whilst at the same time being able to evolve and extend more computing resources to become available. In addition, ubiquitous computing is

able to capture the differing contexts and integrate them with users plus devices (Kagal *et al.* 2002; Kagal, *et al.* 2001).

The pervasiveness and cost effectiveness of mobile computing devices make them suitable for exploitation as voting systems, but security issues stare users in the face. The security issues in mobile computing systems raise the need to find novel security mechanisms. To date, security related issues in the mobile computing environments have not been adequately explored in depth, thereby limiting the applicability of mobile computing devices for the conduct of practical elections. Nevertheless, addressing the security issues associated with ubiquitous computing is germane to the real-world deployment of the technology, especially in the domain of elections. Ubiquitous computing environments raise the complex security issues that require novel security mechanisms that are able to adequately deal with the ubiquity in the practical sense. In fact, the self-same attractive features that make ubiquitous computing environments convenient, portable, cost-effective and powerful, also make them vulnerable to fresh security threats. In order for ubiquitous computing systems to deliver the promise to revolutionise the future of technology in computing, and for it to be widely deployed for the conduct of elections, the associated security issues must be mediated (Al-Muhtadi 2005; Kagal *et al.* 2002; Kagal *et al.* 2001).

In the sections that follow, a brief overview of several issues concerning security services, knowledge based authentication, possession based authentication, biometric based authentication, and location based authentication are reviewed to provide a foundation and an understanding of the tools in order to design a secure MI-voting system architecture.

3.1 Security Services

In this section, the security services that are particularly relevant to the context of this research are explained in the context of mobile Internet or ubiquitous computing. Besides the security requirements of a voting system that were discussed in Chapter 1, verification, identification, and authentication security services will be discussed, to provide an understanding of the subsequent discussions.

A v*erification system* is used to prove that the identity claimed by an individual entity in a system is correct. An entity is regarded as a person, a computer program, a

device, a sensor, or even a physical space. The system performs a one-to-one match to prove that the entity is what was claimed to be in order to authorise access to the system. Authorization is the process of verifying whether an entity has the right to access the protected resources. The entity may provide detailed information such as: name, identification (ID) number, token, or password, to reference the enrolment of the entity. An enrolment is the process whereby an entity registers on a system and her features are extracted, usually in a controlled environment, and used to verify the identity of the entity in future (usually uncontrolled) environments (Jain, Ross and Nandakumar 2011). This is discussed next.

An *identification system,* on the other hand, searches all references of entities in a database for a match of the sample presented. *Identification* is the process of linking an entity with an identity previously captured and stored in a database. This process can be initiated by the entity, typically by typing ID, passing parameters or it can be inferred by the system through sensors and detectors. An identification process may be regarded as a one-to-many relationship and although not covered, may be leveraged at voter registration for enrolment to determine voter eligibility (Sabareeswari and Stuwart 2010; Pato and Millet 2010; Jain, Ross and Prabhakar 2004).

The *authentication service* provides the assurance of the claimed or detected identity of an entity in the system. This service verifies whether the identification of an entity or system user is correct. *Data authentication* provides evidence that a piece of data has originated from a particular authenticated system user. *Location authentication* provides an assurance for the claimed location of the system user. A more general form of location authentication is *context authentication* that provides an assurance for the claimed context that a particular system user operates under (Al-Muhtadi 2005). The notion of context is defined in literature as any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications (Furht 2008). The authentication systems can be classified into the following categories:

a) **Knowledge based authentication** is based on information that authorised voters must remember. Examples include passwords and secret questions.

b) **Possession based authentication** is based on the voter extrinsically possessing a token. Examples include tokens like photo-ID, voter-ID and smart tags such as eID, RFID, NFC.
c) **Biometric based authentication** measures voter's physiological or behavioural characteristics.
    - **Physiological biometric** data relate to the intrinsic physical traits of a voter's physique. These include items such as fingerprints, retinal scans and facial geometries.
    - **Behavioural biometric** data include signatures, handwriting analysis, keystroke detection, and voice pattern recognition.
    - **Chemical biometric** data include properties such as body odour and thermographs (Jain, Ross and Nandakumar 2011; Pocovnicu 2009; Thakur 2012; Patil and Shimpi 2013; Chen *et al.* 2010).

Each of these authentication systems will now be discussed in detail, together with examples.

3.2 Knowledge based Authentication

Knowledge-based person authentication relies heavily on the surrogate representations of identity, such as passwords or ID cards, which can be easily forgotten, lost, guessed, stolen, or shared. Individuals may also impersonate others by presenting forged identification documents. Moreover, traditional security systems do not provide strong post-event recognition such as a voter having voted, and an individual may repudiate that she voted. Consequently, it is becoming increasingly apparent that knowledge-based mechanisms are necessary, but not sufficient for reliable identity management (Jain, Ross and Nandakumar 2011; Pocovnicu 2009; Patil and Shimpi 2013). This may be exacerbated in uncontrolled remote voting for instance.

3.2.1 Password or Secret Question Authentication

The password is a commonly used authentication protocol to validate users prior to allowing them access to a system. This security protocol transmits passwords that are not very well encrypted, over a computer network; and is therefore considered

insecure. Passwords can be easily hacked because they can be shared, stolen, lost, or spied (Pato and Millet 2010). Controlling computing access through a password is a proxy for verifying the presence of a person who may be an imposter that gained illicit access. Patil and Shimpi (2013) reiterate the human difficulty in remembering passwords, citing psychological studies that have proven that humans remember pictures more easily than text. They propose a hybrid graphical password system that is resilient to shoulder surfing attacks. Bao *et al.* (2011) measured the time taken to type an 8-character mixed-case alphanumeric password for desktop and mobile phone systems. The participants achieved 17wpm and only 6wpm respectively, using the two systems. They found that mobile device users were aware of the extra efforts, and they avoid business data on their phones because it would have required a corporate-compliant password. This discussion supports the argument that the use of passwords should be replaced with other means of ensuring a high level of system security.

Nonetheless, one captures secret questions, not passwords, for storage as a backup mechanism to identify a voter who is experiencing a verification failure during an election. An electoral authority, being independent, may at their discretion, include passwords in the full model to make it even more robust. One must now examine the use of public and private key for security. This research proposes that the public key could be stored in the NFC tag to ensure free transfer.

3.2.2 Public and Private Key Authentication

There are numerous tested public and private encryption systems currently in use to ensure high level security controls in information systems. However, within the scope of this thesis, the i-voting models could well replace the password with the private key to ensure a better form of security. This helps to ensure secure transmission of voting data and increase stakeholder confidence. That said, Budurushi, Neumann and Volkamer (2012) suggest that ease-of-use is negated, in current i-voting rollouts, by the need to use two secret pins. Chaum (1983) introduced a blind signature where the content of a message (the vote) is disguised or blinded before it is signed. The resulting blind signature may be publicly verified against the original, not blinded message, much like a regular digital signature.

The complication, as Jain *et al.* (2004) mention, is that the *same* individual biometric measurement taken over *different* times is almost *never* identical. This is because of the sensory conditions (such as imaging), environmental conditions (such as ambient conditions) and other factors (such as user emotion). This variance is the reason the threshold technique *was* introduced (Jain *et al.* 2004). The enrolment process introduces errors such as Failure to Enrol (FTE), which is the inability to create a template because the voter may not have the biometric, or may have a failed biometric capture after maximum consecutive attempts (Modi 2011). The absence may be medical, lifestyle, or circumstantial. Harris (2013:website) reflects on "an unspeakable irony, a challenge that the Sierra Leone voter ID biometrics collectors are facing is how to get fingerprints from people whose hands were cut off." The biometric choice must be mindful of such contextual challenges with an appropriate backup system, which is beyond the scope of this study.

3.3 Possession Based Authentication

Possession-based person recognition also relies on the surrogate representations of identity, such as passwords or ID cards, which can be easily forgotten, lost, guessed, stolen, or shared as discussed in the case of Knowledge based authentication.

3.3.1 Photo authentication

The photo remains the most widely used possession based authentication technique. The capture challenges include *equipment* (camera, grid, lighting, backdrop, optic resolution, lamination); *logistics* (outbound, reverse logistics, and between-use storage); and *human* (support, training and deployment). A further challenge is socio-cultural issues with photo-ID. This is typified by *averseness, such as* "You take my picture, you capture my soul"; or *avoidance,* where people avoid photographs because of concerns about how the images may be used (privacy); or *customary,* where they may be required to remove customary adornments such as headscarves. This introduces unintentional (minority) discrimination. Notwithstanding, photo-IDs are used by almost all developing countries, including countries such as India, Brazil and South Africa (Thakur and Dávila 2013).

3.3.2 Radio-Frequency Identification (RFID) authentication

Radio-Frequency Identification (RFID) is the use of a wireless, contactless system that uses electro-magnetic fields to transmit data from a tag attached to an object, for the purposes of automatic identification and tracking. RFID operates, *inter alia*, in the 13.56 MHz frequency band up to distances of 20 centimetres with data transfer rates of up to 424kbit/s (Thakur 2009). The aspirant voter is issued with RFID Smartcards with voter biodata. To repudiate multiple votes, one method automatically invalidates the card once a vote is cast. The voter is given a VVAT style receipt to aid transparency and auditing, but sadly this mechanism violates *receipt-freeness* and may even expedite vote selling (Oren and Wool 2010).

The RFID technology was used to make voting stations less prone to mechanical failures. However, Oren and Wool (2010) developed attacks that were easy to mount, difficult to detect, and compromised both the confidentiality and the integrity of the election. The RFID also introduces possible *privacy* issues and possible vote monitoring as these may be tracked from a distance with an appropriate reader. Oren and Wool (2012) showed that the RFID system might be attacked or erased with appropriate resources, which violates *fairness* (Thakur 2007; Oren and Wool 2010; Oren, Schirman and Wool 2012). RFID requires a high investment, infrastructure, people and technology, with little application in the election scenario. However, RFID has found success in other parts of the Electoral Cycle; for example, in Cambodia it was used to track ballot paper logistics (O'Conner 2010).

3.3.3 Smart Card authentication

Smart card, as a national electronic ID card (eID), is secure, increasingly available, and is a technology used for voter authentication in several e-voting democracies, *inter alia* Estonia, Australia, Belgium, Germany, Netherlands, Norway, Romania and Latvia. Authentication of the eID and holder takes place using a set of two credentials (Castro 2011b; Chandramouli and Lee 2007). The advent of the smart card provided a significant leap to voting technology, as persons were able to vote conveniently within their own comfort zone. The process requires that a voter places the eID into a reader. Smart cards need costly infrastructure and an appropriate reader, which confines its usage to kiosk voting, particularly in the developing parts

of the world (Castro 2011b; Chandramouli and Lee 2007). The eID embeds user details in a tamper proof microchip, which provides local or non-server centric verification. If that data is fingerprint, the comparison may be client side (Chandramouli and Lee 2007; Budurushi, Neumann and Volkamer 2012).

The implementation of a smart card authentication system is a nontrivial task in the context of MI-voting in developing countries, because a voter will be required to invest in a smart card reader. Smart cards are not utilised by citizens in some democracies, such as the USA, because some authors have cited privacy issues as a bottleneck (Healy 2014; Castro 2011). Budurushi, Neumann and Volkamer (2012) evaluated smart card usage in e-elections in Austria, Estonia and Finland and pointed out that where the smart ID did not have multiple parallel uses, user friendliness was compromised. They also note that eIDs have limited functionality, making it difficult to improve security.

3.3.4 Near Field Communication (NFC) Authentication

NFC is a short-range, wireless connectivity standard, based on the RFID technology that uses magnetic field induction to allow for automatic seamless communication between electronic devices in close proximity. This technology enables users to perform intuitive, safe, contactless transactions, access digital content, and connect electronic devices simply by touching or bringing devices into close proximity (Coskun, *et al.* 2011). NFC operates in the unlicensed 13.56 MHz frequency band with data transfer rates of up to 424kbit/s and transmits data across small distances (4 and 10 cm). This is viewed against Bluetooth's 2.4 GHz frequencies and 10-meter range, or RFID's range of a few centimetres to a few kilometres. This design of NFC reduces the likelihood of unwanted *interception* and makes NFC particularly suitable and secure for crowded areas where correlating a signal with its transmitting physical device becomes difficult. These characteristics position NFC as a secure technology to explore in the context of the MI-voting system (Chen *et al.* 2010). NFC readers are already a ubiquitous feature on smartphones. This almost free cost provides a technology push, driving uptake and demand (Chen *et al*. 2010).

NFC enables ubiquitous communication as opposed to ubiquitous computing. This communication falls into three primary categories - service initiated, peer-to-peer, and payment and ticketing. In the service initiation scenario, the user touches an NFC-

enabled device – such as a mobile phone - against a specially located NFC tag, which typically transmits a small amount of information to the device. This could be some lines of text, a web address (URL), or some simple data (Coskun *et al.* 2011). In this study, it will be the voter data, such as Voter Identification Number and the Voter's Public Key, which is explained in Chapter 4. This data transmission will either replace or mediate the use and theft of passwords and fake profiles through impersonation (Thakur *et al.* 2015). The researcher postulates that NFC or eID may extend beyond just MI-voting and prove to be a multimodal, one-time-single sign-on for users on all systems, extending the scope of this work (Sanjith and Deokaran 2013).

The Fast Identity Online Alliance (FIDO) is a consortium of ICT vendors, *inter alia* Google, Microsoft, Blackberry, Samsung Electronics RSA and Lenovo; chipmaker Qualcomm; mobile chip designer ARM Holdings; Financial services' Bank of America, PayPal and Visa). The FIDO has committed to a password-less future, with the release of open specifications which they claim will bring passwords to an end by enabling authentication through biometrics and hardware tokens (Fido 2014; Solomon 2014a; 2014b). The Fido (2014) infers that NFC and Bluetooth extensions will be added in 2015. This emphatically validates the choice of NFC, which the researcher proposed in New Delhi (Thakur *et al.* 2014).

The NFC technology provides mobile *location-based service* opportunities. A simple example would be tapping a tagged poster with a smartphone and getting electronic coupons. The smartphone is the initiator and the poster the passive target. In this context, location-based services provide for manned or unmanned poll-site identification of voters. All the major telecommunication carriers, credit-card companies and banks support NFC technology. They, along with companies like Google, Amazon and PayPal, started trialling NFC systems. With already made NFC readers a ubiquitous feature on smartphones, this provides a technology push driving uptake and demand (Chen *et al.* 2010).

Ok *et al.* (2010) proposed the use of NFC technology in elections where one tag is placed on each candidate poster choice in the voting booth. The voter taps a candidate choice with a smartphone, this launches NFC communication and the candidate choice becomes input data to the voting application launched on the

mobile. The user types a private key to identify the voter and the encrypted vote is transferred over the network. In this poll-site application, the mobile device is owned and supplied by the authority (Ok *et al.* 2010). This is a novel and intuitive poll-site use, which may well be extended to assist in non-political community decision making. The NFC technology has arguably more opportunity in other parts of the electoral cycle than the RFID technology.

Thakur and Beer (2014) implemented an NFC system to authenticate 941 university degree documents through a website verification process. A user places his/her mobile phone over the NFC, which only contains an index. The index is fed as an input parameter and a website is launched, with an index as the search key, returning the degree awarded and the year of the qualification. The returned results either affirm or reject the authenticity of the document. This live implementation of NFC provides comfort for its implementation in the MI-voting scenario.

3.4 Biometric Based Authentication

Biometrics is the automated use of science and technology to uniquely identify individuals based on their physiological or behavioural characteristics. Biometric technology sits at the intersection of biological, behavioural, social, legal, statistical, mathematical, and computer sciences, sensor physics and philosophy (Jain and Ross and Nandakumar 2011; Pocovnicu 2009; Whither Biometrics Committee 2010). A biometric uses a twofold process (enrolment and verification) to authenticate a person. A biometric system must be robust to exclude illegal votes, while the population must not suffer indignity because of the lack of privileges available to those who have not successfully enrolled (Harel 2008).

There are various types of biometric systems and some of them can be combined to realise a multimodal authentication. A system that uses more than one biometric is widely called a multimodal system, biometrics fusion, or multimodal recognition. Multimodal systems offer greater protection against spoof attacks, as an impostor must fake several biometric characteristics. A multimodal system may significantly improve the voter confidence in the system given the effort to invest in self-identity. On the other hand, these systems require more resources for acquisition, computation, storage, comparison and more time, triggering user inconvenience and expense.

Despite this, multi-modal systems are being increasingly deployed in high-security applications (Jain *et al.* 2011; Pavešić and Ribarić 2009).

3.4.1 Physical Biometric Authentication

Physical biometric authentication refers to a physical attribute of humans that can be captured and used for security purposes. The fingerprint and face authentication are extensively deployed in many practical applications. Galton (1892) in an exhaustive study investigated the individuality and permanence of fingerprints, which was duly adopted by Scotland Yard. The fingerprint is unfairly associated with criminality in some contexts (Wong 2006); nonetheless, it is one of the widely used biometrics, which may impact electoral acceptance. The scientific community agrees that fingerprints are immutable. Fingerprint databases are almost universally deployed for population registration, voter registration, and criminal records (Jain *et al.* 2004; 2011; Galton 1892). However, they provide some challenges, such as *Capture equipment* hardware and software dependency, *Human support*, *Environmental conditions* such as heat introducing sweat, *Enrolment and Comparative errors* (Woodward 1996; Thakur and Dávila 2013).

Facial recognition is innate in humans and is one of the most widely used biometrics. It is almost standard practice to incorporate facial photographs in authentication tokens such as passports and driver's licences. The *natural person* photo-ID is also the current most popular voter identification method. Some registration systems, such as those obtainable in South Africa, combine face portraits with fingerprints for authentication (Thakur and Dávila 2013). The automated comparison of two face images to determine equivalence, is a non-trivial task in computer vision, as several capture challenges emerge. These challenges include variations in posture, illumination, age, facial expressions, makeup, facial hair, and the wearing of headscarves. This is complicated by similarities like identical twins, or biological father and son. The former spawns a spoofing attack known as evil-twin. It has been demonstrated that one could take a still-photograph of someone and place it in front of the phone to unlock it. Security vendors are mitigating this by using, amongst others, video or human motion (Crossman 2012). Paul and Anilkumar (2012) propose a multi-modal online voting model using face and fingerprint, and a transmission innovation of merging of the secret key and pin with the image which is

transmitted. The key is extracted on the server and used for verifying the voter with the image. Afghanistan is undertaking a large-scale biometric project combining face, finger and photo (van de Haar, van Greunen and Pottas 2013).

3.4.2 Behavioural biometric authentication

Behavioural biometric technology is still in its infancy, but has prospects of improving system security. There are three major behavioural characteristics, namely gait, keystroke and voice biometrics. Gait authentication involves people being identified through the analysis of their movement. It requires a video system for recording person movement; gait is unique but computationally intensive. Keystroke biometrics are processes that involve the analysis of typing speed and rhythm of an individual. Voice authentication is the process of verifying the claimed identity of a speaker based on the speech signal emitted by the speaker (voiceprint). In order to verify that the individual speaking is, in fact, who he/she claims to be, the captured voice features have to be matched with recorded samples of the same speaker's voiceprint. This voiceprint is generated, by usually asking a voter to repeat a few expressions in a controlled environment such as during the voter registration, which is called enrolment. There are two types of speaker authentication systems, which are Text-Dependent Speaker Authentication (TD-SA) and Text-Independent Speaker Authentication (TI-SA). TD-SA requires the speaker to say exactly the enrolled voice text password. TI-SA is a process of verifying the speaker identity without restriction on the speech content. Compared to TD-SA, TI-SA is more convenient because the user can speak freely to the system without constraint. However, TI-SA requires longer training and testing utterances to achieve good performance (Liu, Huang and Zhang 2006a; Liu *et al.* 2006b; Shankaranand 2014).

3.5 Proposed authentication model

The authentication model proposed in this project is based on possession biometric, location system, and behavioural biometric, to enforce high-level security in a mobile Internet voting system. In the proposed authentication model, NFC token is being selected for first level security, followed by second level GPS security, and voice as the behavioural biometric for the third level security, as shown in Figure 3.1.

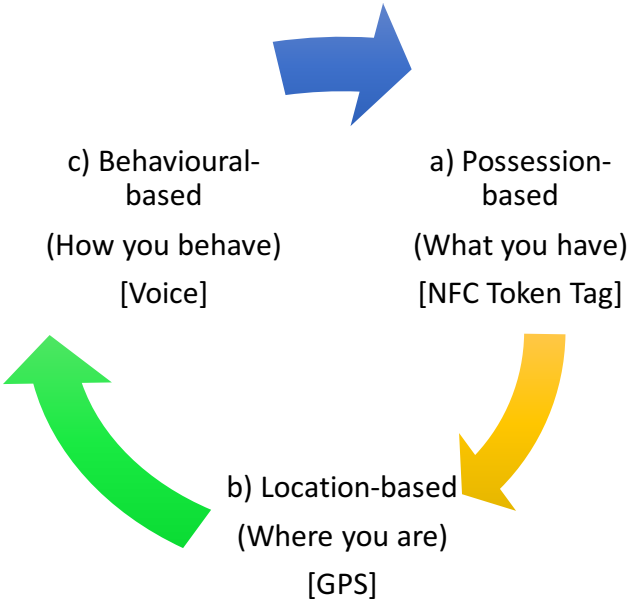

Figure 3.1 Proposed authentication model

There are literature justifications for multimodal security control in real high-security applications (Pavešic and Ribaric 2009). A multimodal system can improve the voter confidence in a system, given the effort invested in self-identification. However, these multimodal security systems require more resources for acquisition, computation and storage efficiency (Jain *et al.* 2011). The combination of multimodal NFC token (possession), GPS (location), and voice biometric (behaviour) provide opportunity to leverage the reduction of resource utilisation.

3.5.1 Possession Based using Near Field Communication (NFC)

The NFC features enable intuitive, safe, contactless transactions, access to digital contents and connection to electronic devices by simple touch or by bringing devices into a close proximity (Campbell *et al.* 2014). Consequently, this study uses possession based NFC as a token that is either stuck or attached to the voter card or ID. The NFC token stores lightweight data of a voter as an important first level security mechanism. NFC is a cost effective technology because the unit price of the current NFC tag is a few US cents and the price is still dropping. The ubiquity of the read capability of a mobile NFC reduces the technology requirements and cost of adoption. This ubiquitous read capability is not the case for Radio Frequency Identification (RFID), because one needs additional dedicated hardware to effectively

use RFID. Moreover, the auto-coupling feature of NFC means that no matter what state a smartphone is in, bringing the smartphone into contact with an NFC will instantly launch the voting system encoded in the tag. This mechanism converts the smartphone into an Electronic Voting Machine (EVM) that requires little or no extra skills by a user to operate. The uniqueness of the electronic storage capability and parsimony features of the NFC tag technology is used to store the voter's ID for authentication. The write-once, publish-only NFC characteristic makes the tag read-only and therefore impervious to overwriting using false data, thus supporting the voter security (Ahson and Ilyas 2011; Ok *et al.* 2010).

Ok, *et al.* (2010) proposed the use of NFC in elections, as shown in Figure 3.2,

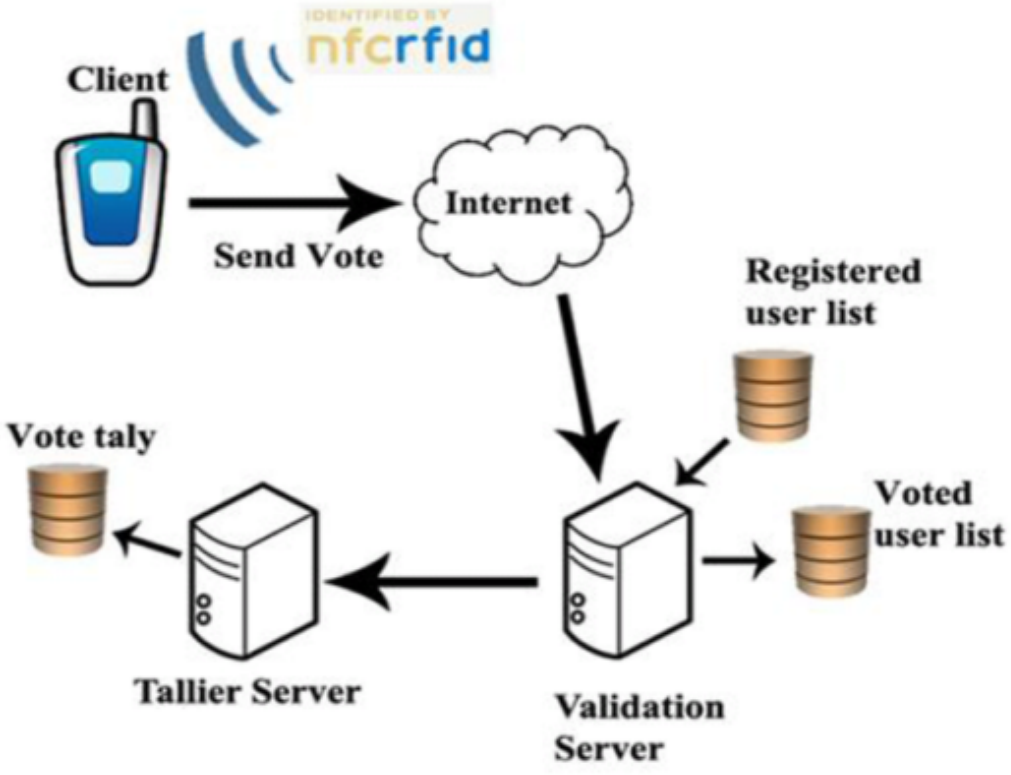

Figure 3.2 NFC Lab: Voting with an NFC token (Ok *et al.* 2010)

where the authorities placed one tag on the picture of each candidate inside the voting booth. The voter taps her choice on the poster and the candidate choice is made. This launches NFC communication and the candidate choice becomes input data to the voting application launched on the mobile. The user types her private key, which identifies her, and the encrypted vote is transferred over the network. This innovative approach for voting eliminates paper and spoilt ballots, but it is a poll-site application

that is limited to a location (Ok *et al.* 2010). This is a novel and intuitive poll-site use, which may well be extended to assist in non-political community decision making.

The use of NFC for voting is novel and curbs surreptitious monitoring, capture and man-in-the-middle attacks. NFC also has opportunity in other parts of the electoral cycle much like RFID.

3.5.2 Location Based Global Positioning System

Even though MI-voting allows voters to vote anywhere and anytime, there is the need to track the voting locations. This control mechanism provides a secure means of controlling possible coercion that may arise if mobile voting is not controlled. The MI-voting system as proposed in this study should satisfy the four characteristics of a frugal system, which are uniqueness, ubiquity, unison and universality. Uniqueness is the drive to precisely identify the characteristics and locations of entities, including a voter and the voting device. Ubiquity is the drive to ubiquitously access information unrestricted by time and space. Unison is the drive for information consistency. Universality is the drive to overcome the friction of incompatibilities of information systems (Watson 2013; Watson, Kunene and Islam 2013; Olugbara and Ndlovu 2014).

The mobile voting system architecture, as proposed in this study, fulfils a unique effort as it is capable of identifying individual voters and their proximity from a cell phone – these are predefined by the system. It therefore meets numerous security requirements. The system architecture identifies individual users, what information to accept or reject and what requests to deny, as well as what pre-defined information to give the users. The system signifies a unison drive, since the system includes the metaphors and operations of the voting system, including election information that is represented by their respective labels, test and images. The metaphors, texts and images are comprehensive enough for voter and user transactions, consistent and easy to comprehend. The system architecture is designed to only accept inputs from authorised users and generates error messages if data elements are received in the wrong format. It also returns confirmatory messages for data successfully received in

the correct format. There is no likelihood of conflicting information, as every system user is associated with a unique identifier.

3.5.3 Behavioural Based Voice Biometrics

The voice biometric provides a secure way to authenticate system users through their natural voice patterns. The use of voice for mobile authentication does not need additional hardware to be installed as voice capture and transmission is an inherent feature of the mobile. The voice biometric is a familiar, intuitive, non-threatening, contactless, frictionless, non-intrusive, culturally deferential method of enrolling and verifying voters by asking them to say something (Jain *et al.* 2004; 2011; Karpov 2011). The key strength of voice authentication systems is the ability to conduct enrolments and verification remotely. The voice biometric provides an advantage that the storage size of a voiceprint is small (Meyers 2004; O' Neil King 2014). It operates in an out-of-band trusted network with both physiological and behavioural, variable, dynamic samples; with no software deployment mediating software-installation-associated compatibility and version control issues. This binds the person, the phone, and the session coherently.

A fraudster may know everything about the target (the voter), but will not be able to *create* a voiceprint, mediating phishing and other Internet attacks. The voice biometric exhibits a low risk of data breach or theft, hacking, social engineering, phishing, brute force attacks, and credential sharing. By comparison, tokens generally exhibit a high risk for data breach and theft, and a medium risk for credential sharing and hacking (O' Neil King 2014; Elliot 2007; Markowitz 2000; Jain *et al.* 2004 2011; Karpov 2011). However, impersonation remains a possibility through playback or replay, audio splicing, voice conversion, voice transformation, or mimicry remains because no authority (or observer) is present to verify the voter and her credentials (Evans, Kinnunen and Yamagishi 2013). These bottlenecks are being addressed with some commercial authentication systems such as Agnitio (Agnitio 2014). Nuance and OneVault declare their products can differentiate between real and recorded voices (Sanjith and Deokaran 2013).

3.6 Conclusion

This chapter presented a literature review and analysis of the different types of authentication systems available, which are considered to be of particular interest to mobile, for possible incorporation into the reference architecture. Multimodal authentication was proposed, as being able to provide the high level of authentication assurance required of voters in elections. The next chapter explores the proposed Secure Mobile Internet Voting architecture further.

Chapter Four - Secure Mobile Internet Voting Architecture

This chapter introduces the researcher's proposed Secure Mobile Internet Voting Architecture (SMIV). The architecture follows the conceptual perspective of the Organisation for the Advancement of Structured Information Standards (OASIS). The SMIV model follows the Electoral Cycle of pre-election phase of voter registration, election phase and post-election phase (EML7.0 2011). These electoral procedures remain basically unchanged despite the evolution of voting protocols. The voting protocols implement four specific sets of tasks, namely registration, collection, validation and tallying. The *registration* process comprises the compilation of the list of eligible voters. The process of *validation* involves checking the credentials of someone who makes an attempt to vote and only allowing the eligible voters to proceed. C*ollection* is a process that involves collecting the voted ballots, while *tallying* is responsible for the counting of the votes. At the tallying phase, before the vote count, the voter's digital signature is removed by the system so that members of the Electoral Authority may open the now anonymous e-votes and count them (Cranor and Cryton 1997; Cranor 1996; Fujioka, Okamoto and Ohta 1992).

In the electoral cycle, no opportunity must be created for fraudulent practices that may breach the sanctity of the electoral process and thereby impair the trust the electorates have in the process (Cranor and Cryton 1997). Consequently, a secure voting system architecture is mandatory to guarantee the sanctity of Internet voting. This chapter revisits seminal models such as FOO, SENSUS and REVS, and combines these architectures with features that the ubiquitous mobile provides. These are discussed in the following sections.

Mursi *et al.* (2013) compiled a standardised set of requirements from literature that e-voting architecture must satisfy in order to realise universal suffrage. These security requirements are used to evaluate voting protocols in this thesis and were presented in Chapter 1. Interestingly, they are not all simultaneously achievable by any conventional or another means of voting. Satisfying the convenience, mobility and flexibility security requirements, in particular, is vital for ensuring a higher participation of voters in elections, although there is a concurrent need for sustaining the security requirements of privacy, eligibility and incoercibility. Although a

number of voting schemes have been proposed in the academic literature over the past 30 years, these schemes were targeted at fulfilling some of the generic requirements for e-voting to minimise electoral frauds. Examples of these voting schemes include absentee balloting, vote by mail balloting, cryptographic protocols, two-agency protocol, one-agency protocol, FOO voting scheme, Sensus, SEAS, and EVOX (Thakur 2014; Mursi *et al.* 2013; Mauw, *et al.* 2007; Baiardi *et al.* 2005; Cranor and Cryton 1997; Fujioka, Okamoto and Ohta 1992).

4.1 Reference Architecture

This section provides guidance for the development and use of SMIV in the form of a description of the widely used reference architectures of FOO, Sensus and Revs for voting systems. Reference architectures have been used in academia and industry to provide information, guidance and direction for focused subject areas. These reference architectures have wide ranging purposes, uses and levels of detail and abstraction. The concept of reference architecture has multiple definitions and meanings, and is relative to the context in which it is used. Reference architecture literature can be found throughout the Department of Defence and industry addressing various subject areas (Cloutier *et al.* 2010). Due to keen interests in the Service Oriented Architecture (SOA), a sufficient amount of existing reference architecture literature is focused on this subject area. Most notable are the efforts by the Organization for the Advancement of Structured Information Standards (OASIS) (EML7.0 2011) and the Open Group Architecture Forum (TOGAF) among others (Bent *et al.* 2008).

4.1.1 Fujioka, Okamoto, Ohta (FOO)

Fujioka, Okamoto and Ohta (1992) proposed a protocol to preserve the privacy and security requirement of voting. This is now referred to as the FOO model and has become a seminal reference in blind signature voting protocols (Joaquim *et al.* 2003). They assert that FOO is a practical and secure secret voting scheme that is suitable for large-scale elections, while concomitantly solving the privacy and fairness problem. FOO comprises as its core basis an administrator, a counter, and the voters. The voters communicate through an anonymous communication channel to mediate vote monitoring which satisfies fairness. It has the desirable stages of *Preparation* where

the voter selects a candidate, completes the ballot, and encrypts the ballot using blind signatures. The *Administration* permits the eligible voter to cast only one vote. In *Voting***,** the administrator signs the ballot and transmits it accordingly. The *Collecting Counter* adds and publishes the vote (Fujioka, Okamoto and Ohta 1992).

The blind signature mechanism of the FOO model enables the voters to get their vote validated from an election authority, while preserving the secrecy of their vote. Blind signatures are the electronic equivalent of signing carbon-paper-lined envelopes, wherein a voter seals a slip of a paper or ballot inside such an envelope and later gets it signed on the outside. When the envelope is opened, the slip will bear the carbon image of the signature. When used in an online voting protocol, a voter encrypts, then blinds the vote, and presents it to a validating authority for validation. After the authority validates the vote, the voter un-blinds the encrypted vote. This yields a validated vote that can no longer be correlated to the original blinded message. The voter then uses an anonymous channel to submit the validated vote to the tallying authorities. It should be noted that there are other encrypted voting protocols such as Mix-nets and Homomorphic encryption protocols. The latter two have a very high mathematical complexity as well as high computational costs (Olusola *et al.* 2012; Benaloh 1987; Chaum 1983; Sako and Kilian 1994).

Protocols within this model are simple, easily manageable, computationally efficient, and naturally mirror "write-in" ballots. The perceived familiarity of "write-in" ballots may simplify training and the *convenience* security requirement. A crucial problem with early protocol schemes was the ability of a malicious server to impersonate absentee voters in the final tally by surreptitiously casting votes on behalf of them, thus violating the enshrined principle of *democracy* (Fujioka, Okamoto and Ohta 1992; Cranor and Cytron 1996; Herschberg 1997). The formal FOO model postulation had a two-phase voting process to achieve fairness. Here voters submitted their encrypted vote and then waited until the *end* of the election period to submit their vote-decryption keys (Fujioka, Okamoto and Ohta 1992). However, Cranor and Cytron (1997) and Herschberg (1997) made a significant adaptation to FOO to permit voters to vote and *walk away*, which is an important Mursi *et al.* (2013) security requirement, which reduces the burden on the voter. That said, these last two adaptations created a risk that a malicious (or not) authority could learn the intermediate results, therefore violating the *fairness* property.

While the FOO protocol satisfies most of the core properties well, it fails to correct the problem that the Validator can surreptitiously cast votes for abstaining voters. These surreptitiously casted votes may be detected by the abstaining voters themselves, or by an auditor who checks the signatures on all the validation requests submitted. However, there is no way to identify the invalid ballots and remove them from the tally. If voters who wish to abstain submit blank ballots, then this problem can be avoided. FOO affirms that the following security requirements are supported: Soundness is supported as invalid votes are detected at the counting stage. Privacy is retained as neither the administrator nor the counter knows the blind signature or the anonymous communication channel used. Unreusability (or non-reusability) is supported as the voter will have to break the blind signature. Eligibility is positioned on the strength of the blind signature and the administrator checking the list of voters. Fairness is supported as the counting of the votes does not affect the voting because the counting takes place after the voting stage. Verifiability is provided for as long as all voters vote as there is then no way for the administrator to dummy vote. This may be achieved due to legislation requiring compulsory voting (Fujioka, Okamoto and Ohta 1992).

4.1.2 Sensus

Cranor and Cytron (1997) designed, implemented and evaluated the Sensus reference architecture, which was an expansion and enhancement on the pioneering work of FOO (1992). Sensus, like FOO, also uses blind signatures to guarantee that only registered voters can vote and can exercise this right once only, while preserving each voter's privacy. Sensus also allows voters to verify independently that their votes were counted. It also provides a mechanism for anonymously challenging the results should their votes be miscounted. While Sensus is deemed suitable for small-scale elections, the authors argue that with minor modifications, it could be scaled to large elections.

The design goals of Sensus were to facilitate the registration, validation, collection and tallying phase of remote Internet voting. This is done in the knowledge that election authorities may collude through ballot stuffing, ballot replacement, ballot destruction, or fraudulent ballot creation; and voters may be impersonated (registered or otherwise) which suggests an opportunity to revisit to model with a view of

appropriate mitigation. The Sensus reference architecture development contributes significantly to the evolution of the proposed SMIV architecture. Table 4.1, which is placed at the end of the chapter because it follows the discussion and shows a detailed comparison, between Sensus, REVS and SMIV within Mursi's *et al.* (2013) security framework. . Figure 4.1 provides the legend of symbols used in the models. Figure 4.2 depicts the Sensus reference architecture featuring the three main components namely validator, pollster and Tallier, and their respective interactions. This is described and discussed in the next section

The *Validator* is responsible for checking voter registration and ensuring each eligible voter casts just one vote. The Validator creates a blinded validation certificate by signing a blinded ballot. The voter then unblinds the validation certificate and submits it to the Tallier with her ballot. The Validator will issue no more than one validation certificate to each registered voter. The Validator uses the registered voter list to obtain each voter's public key and check the signatures on their ballots. The Validator changes the contents of the validation field from 0 to 1 after validating a ballot. This method prevents record keeping of the order in which ballots are validated. This mediates a replay attack (Cranor and Cytron 1997; Cranor 1996).

The *Pollster* presents the ballot to the voter, collects the ballot response from the voter, performs the requisite cryptographic functions on the voter's behalf, obtains validations and receipts, and finally delivers the ballots to the ballot box. This is the only component that the voter must completely trust. Pollsters implement the user interface to capture the ballot. The purpose of encrypting the ballot is to support the security requirement of fairness by preventing vote monitoring (Cranor and Cytron 1997; Cranor 1996).

RVL = Registered Voters List supplied by Registrar and residing on Validator compromising the n-tuple (VIN, ..) for each voter.

BL = Ballot List residing on Tallier,

Rec# = Receipt Number

RL= Received voter List, partially populated by Registrar, with VIN, dk number

B = The vote ballot

GPS = Current coordinate location of the mobile

VB=Validation Bit

VIN = Voter identification number

ek, dk encryption/decryption keys

pk_p, pk_p^{-1} = pollster public private key pair

pk_T = tallier public key

pk_v=validator public key

C=confirmation key

$\{...\}_{pk_i^{-1}}$ message signed party i

$\{...\}_{ek}$ message encrypted by public key ek

GPS_{ea} is the area where a voter commits to vote from

Figure 4.1 Legend for Sensus and SMIV reference architecture

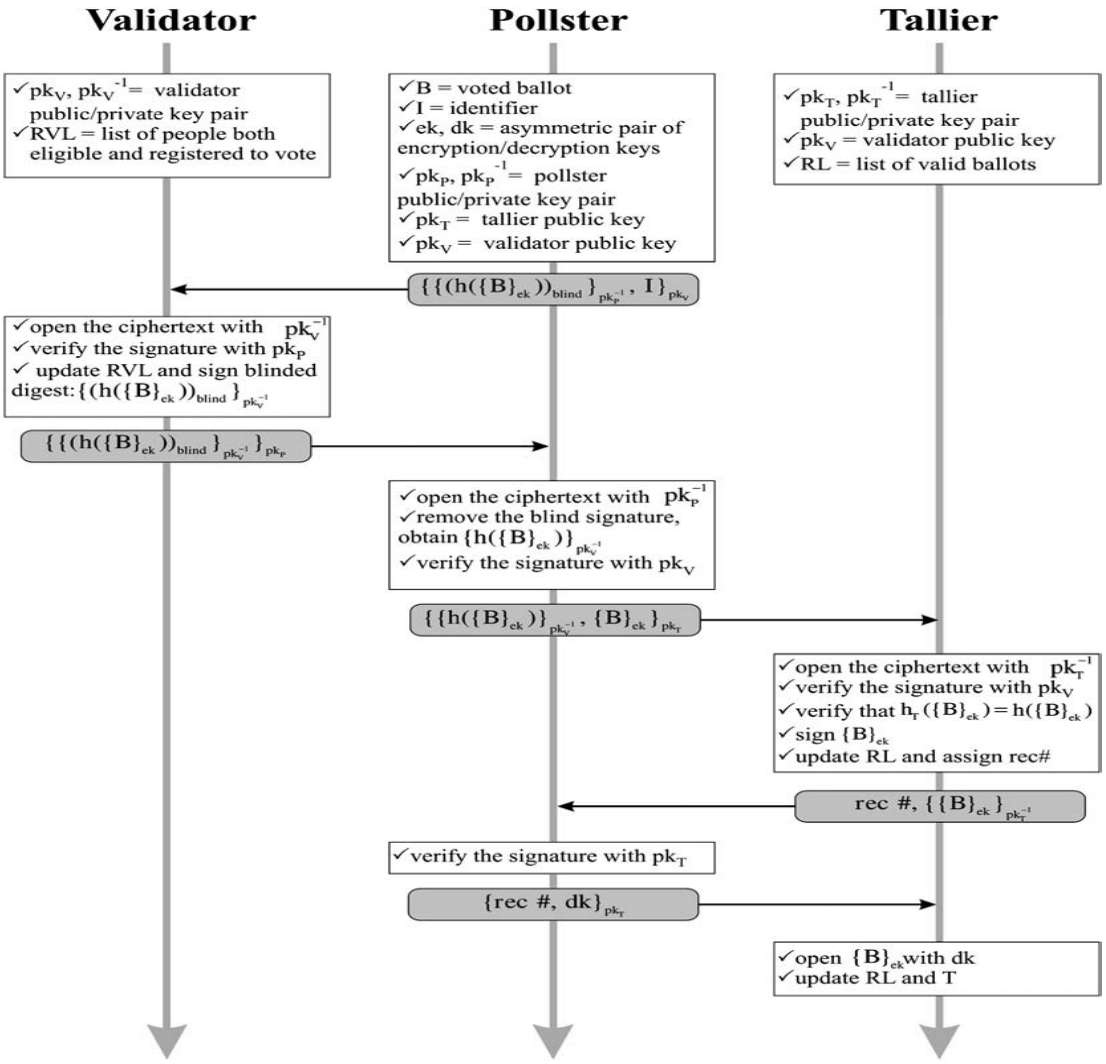

Figure 4.2 Sensus reference architecture (Baiardi 2005)

The *Tallier* is responsible for collecting the vote ballots and tallying the results of the election. Voters first submit encrypted ballots, signed by the Validator to the Tallier. The Tallier checks the authenticity of the validation and verifies that the encrypted ballot is unique among the encrypted ballots received thus far. If the ballot is valid and unique, the Tallier issues a signed receipt to the voter. The voter then submits the ballot decryption key. The Tallier uses the key to decrypt the ballot. After the election, the Tallier publishes a list of encrypted ballots, decryption keys, and decrypted ballots, allowing for independent verification of election results. The Tallier computes a 16-byte digest of each encrypted ballot received and uses it to index the encrypted ballots and receipts. With the FOO protocol, the Tallier responds by placing the encrypted ballot on a list that is published after all the voters vote. Thus, a voter cannot submit his or her decryption key until after the voting phase of

the election is over. As a result, votes cannot be cast in a single session, which has an impact on voter time. In the Sensus protocol, as soon as the Tallier responds by sending a receipt to the voter she may submit the decryption key immediately, completing the entire voting process in one session (verification must still wait until the election is over). Therefore, Sensus protocol alleviates the two-phase requirement of the voter, promoting the walkaway security requirement. However, both Sensus and FOO have a deficiency of allowing the Tallier to vote for absentee voters.

Evaluation of Sensus with respect to security requirements reveals that the system makes the following assumptions:

a) The voter's privacy while casting the vote is not violated through surveillance mechanisms such as the voter allowing someone to look over their shoulder. The latter is referred to as *shoulder surfing*.
b) Voters do not use a multi-user system where other users with root privileges may also detect and determine the vote. Other terms for root user are supervisor or administrator.
c) The voter uses a trusted computer system in which it is not possible for clear text messages to be intercepted, with all communication between voter and election servers occurring over an anonymous channel.
d) The messages from voters will not arrive at the Validator and Tallier in the same predictable order, allowing the Validator and Tallier to collude to link ballots with the voters who cast them, which unintentionally expedites vote purchasing.
e) The encryption algorithms used are sufficiently strong that encrypted messages cannot be decrypted without the proper keys.

The discussion of the security requirements of the Sensus system now follows in this context (Cranor and Cryton 1997; Cranor 1996).

Accuracy: Although it is possible to alter, eliminate, or add votes, this activity is detectable. Therefore, Sensus satisfies all three parts accuracy. Legitimate participating voters may discover if their votes have been altered or eliminated from the final tally by examining the Tallier's published list. These voters can then submit their receipts anonymously along with their ballots and decryption keys to protest the

election results and have them corrected. The only party that can add invalid votes to the final tally is the Tallier. Any party who checks the authenticity of the validation certificates for all ballots can detect these; the final tally may then be corrected.

Invulnerability: Sensus satisfies the invulnerability property completely only if and when all registered voters submit ballots.

Privacy: Sensus partly satisfies the part of the privacy property, because it is not possible for any party to link a ballot to the voter who cast it. However, Sensus does nothing to prevent a voter from proving that he or she voted in a particular way.

Verifiability: On the one hand, Sensus satisfies the verifiability property completely because voters can verify that their votes were counted correctly and correct any mistakes they might find without sacrificing their privacy. On the other hand, it is still not possible for any interested party to verify that *all* votes were counted correctly.

Convenience: Sensus satisfies the convenience property by allowing voters to cast their votes and walk away. In a mock election, experiment participants cast their votes within a few minutes and found the Sensus interface easy to use. The participants recommended adding an Internet browser voting interface to make Sensus more accessible to voters.

Flexibility: Sensus satisfies the flexibility property because the ballot is highly configurable to cater for varying numbers of candidates and referendums.

Mobility: Sensus satisfies the mobility property, as it can be used from any computer connected to the Internet.

Eligibility: Sensus will not accept ballots from those not registered to vote.

Democracy: Sensus supports democracy in part because will it not accept more than one ballot from each registered voter.

Ballot stuffing: Invalid ballots only can be introduced into the final tally by the colluding Validating authority if some voters do not submit ballots.

4.1.3 REVS (Robust Electronic Voting System)

The REVS voting protocol uses blind signatures and distributed authorities to prevent voter impersonation by an authority and provides individual verifiability. REVS addresses fault tolerance and is therefore designed for distributed and sometimes faulty environments, like the Internet, where one may experience device or communication failure. REVS sustains robustness by alleviating the impact of vote interruption by performing an appropriate recovery without weakening the voting protocol. This allows for tolerating a level of real-time server failure with server replication. The redundant servers are prohibited from individually or in collusive unison, corrupting the election outcome. This is achieved through a combination of several participating machines, while computer or communication failure is mediated by maintaining a loosely coupled state. Furthermore, each voter keeps a local state in mobile non-volatile storage, allowing her to stop or resume the electoral process at any time or place. The Servers are replicated for redundancy; thus implying only a subset of servers can be contacted by each voter. Each server keeps a distinct state regarding the participation of each voter in the election, allowing determinism because each voter will get the same answer from each server regardless of the number of times she poses it. Each server alone is not able to act as any voter and cannot provide false replies to voters without being noticed. The collusion of servers in order to interfere with the election (e.g. voting for absentees) is prevented to a certain degree (Joaquim *et al.* 2003).

The REVS protocol uses five types of servers: Commissioner, Ballot Distributor, Administrator, Anonymiser and Counter as shown in Figure 4.3. There is a Voter Module, typically resident on the voter's device, which is used by voters to support all the appropriate interactions with electoral servers (to get the ballot, get it signed by election servers, validate and submit the ballot, etc.). The Voter Module, resident on the voter's device, interacts with the electoral servers during vote casting. The Electoral Servers have the following functions:

Commissioner generates the election's keys, signs ballot questions, and defines the operational setup of the election, such as addresses and public keys of servers, number of required signatures, *etc.*

Ballot Distributer distributes ballots to the voter and distributes the operational configuration as determined by the Commissioner

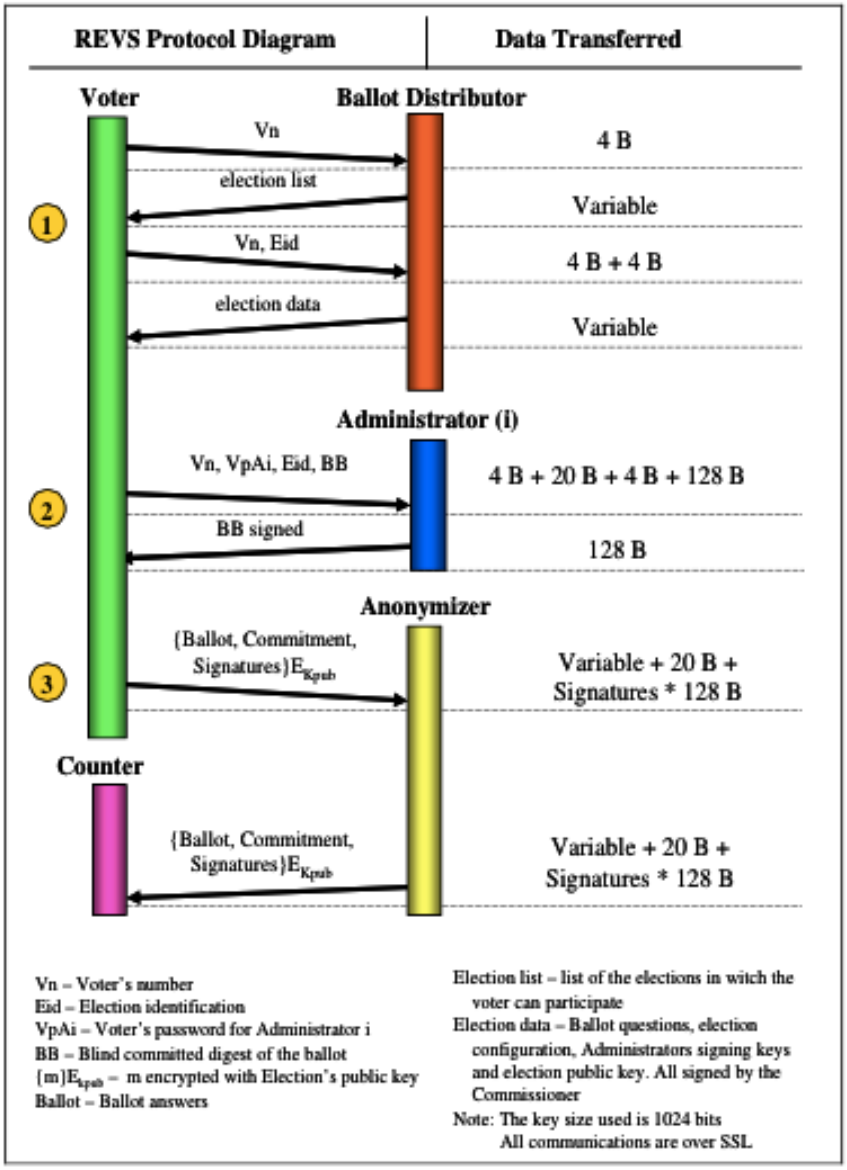

Figure 4.3 REVS sequence diagram (Joaquim *et al.* 2003)

Administrator decides upon the acceptability of a voter ballot. The voter must use a different password to get a signature from each individual Administrator server, and the voter must have a minimum set of signatures (more than half of the signatures from every Administrator server) for ballot acceptability. Because an Administrator only issues signatures for valid voters who have not yet voted, this minimum set makes it impossible for any voter to get two valid votes. Because each Administrator

does not know the passwords of other Administrator servers, they cannot collude to impersonate a voter.

Anonymiser receives the encrypted voter's ballot and provides anonymity to the voter's device by hiding the voter's location, and randomly delaying and shuffling several submitted ballots before submitting them to the Counter; thus preventing the Counter from associating a ballot to a device owner based on time analysis.

Counter verifies the validity of the ballots, and is thus the server that verifies all required signatures are on the ballot. Then the Counter removes the repeated ballots verifying a bit commitment (made by the voter in the ballot signing phase) and performs the tally after the election when the Commissioner releases the election key (private key) to decrypt the submitted ballots.

REVS is scalable, robust and deliberately designed for (sometimes) faulty and distributed environments. This allows the voting process to be terminally interrupted and then resumed in exactly the same state as at the time of interruption. REVS mitigates server collusions; however, REVS reveals the private key after the election, which means that it ultimately does not prevent vote selling. It also does not address malware on the system client-side and it makes the assumption that the system used has special trusted software (Joaquim *et al.* 2003; Zúquete, Costa, Romao 2007). For REVS, the electoral process is all that concerns the realization of an election.

In order to provide a better understanding, it is useful to divide the electoral process into several phases and reiterate some points that have already been made software (Joaquim *et al.* 2003; Zúquete *et al.* 2007):

Registration/Preparation: The list of eligible voters is complied, the ballots prepared and the necessary election arrangements made. This also includes any required pre-election activities.

Validation: The eligible voters' credentials must be verified. Only registered voters are authorized to vote, and only once.

Collection: The collection phase consists of anonymous collection of the voters' ballots.

Verification: The validity of the ballots is verified with only valid ballots used in the tallying phase.

Tallying: The tallying phase comprises counting valid ballots. At the end the tally is published.

Claiming: This is the phase in which all claims should be made and appropriately investigated. At the end of this phase the final results are published.

REVS can be evaluated against some pertinent Mursi's *et al.* (2013) security requirements, as follows:

Accuracy - It is supported as vote cannot be altered, because this will destroy all Administrators' signatures. A voter can verify and report an anomaly if her vote was eliminated from the final tally, by inspecting the list of received votes published by the Tallier. Since the signatures can be verified by anyone and are published with the votes, it is impossible for an invalid vote to be part of the final tally.

Democracy - Each voter can only vote once in each election because only one NFC tag is issued per voter. The Pollster and Validator cooperate to overwrite revotes preserving democracy.

Privacy - Neither authorities nor anyone else can link any ballot to the voter who cast is guaranteed. However, the second part of privacy - that no voter can prove that he voted in a particular way, as in most voting protocols proposed so far, isn't accomplished.

Verifiability - The final tally can be made by anyone, verifying the signatures on the votes and summing all votes. Each voter can verify if its own vote is correct, using the information saved.

4.2 The Secure Mobile Internet Voting (SMIV)

The primary objective of e-voting system design is to ensure that the electoral cycle tasks are carried out electronically and securely. The proposed SMIV architecture is inherently a fundamental organisation of a voting system embodied in its components, the relationships of these components to each other and to the environment, and the principles guiding its design (Cloutier *et al.* 2010).

The proposed SMIV architecture is based on the reference architecture of Sensus (Cranor and Cryton 1997); REVS (Baiardi *et al.* 2005; Mauw *et al.* 2007); and FOO (Fujioka, Okamoto and Ohta 1992). However, security requirements in SMIV, such as incoercibility, eligibility, convenience, accuracy and mobility, are realised differently from the Sensus approach in the study's proposed approach. This difference is motivated by a desire to leverage recent advances in mobile, Internet, GPS, NFC and voice biometric technologies. For example, the voter identification number and secret token were used to implement *eligibility* in Sensus (Cranor and Cryton 1997), while the SMIV's architecture uses VIN, confirmation key, voice biometric, and GPS locations to achieve *eligibility*. The satisfaction of Mursi's *et al.*'s (2013) respective requirements of incoercibility, convenience, mobility, accuracy and transparency in the proposed SMIV is addressed in Section 4.3.

SMIV's use of modern technologies to satisfy the requirements of eligibility, convenience, mobility, incoercibility, accuracy and transparency as proposed in this study, is an important contribution to e-voting research. The sequence diagram of the SMIV architecture is shown in Figure 4.4. The protocol is explained in the section thereafter.

4.2.1 The SMIV Protocol

SMIV uses a multimodal authentication because voice is an emerging technology which cannot, alone, securely identify a person. The underlying theory of voice is explained in Chapter Five.

The following section indicates the protocols used among SMIV components which is illustrated in Figure 4.4. Some assumptions include the following. Figure 4.1 contains the reference for the legends as well as the acronyms for this protocol.

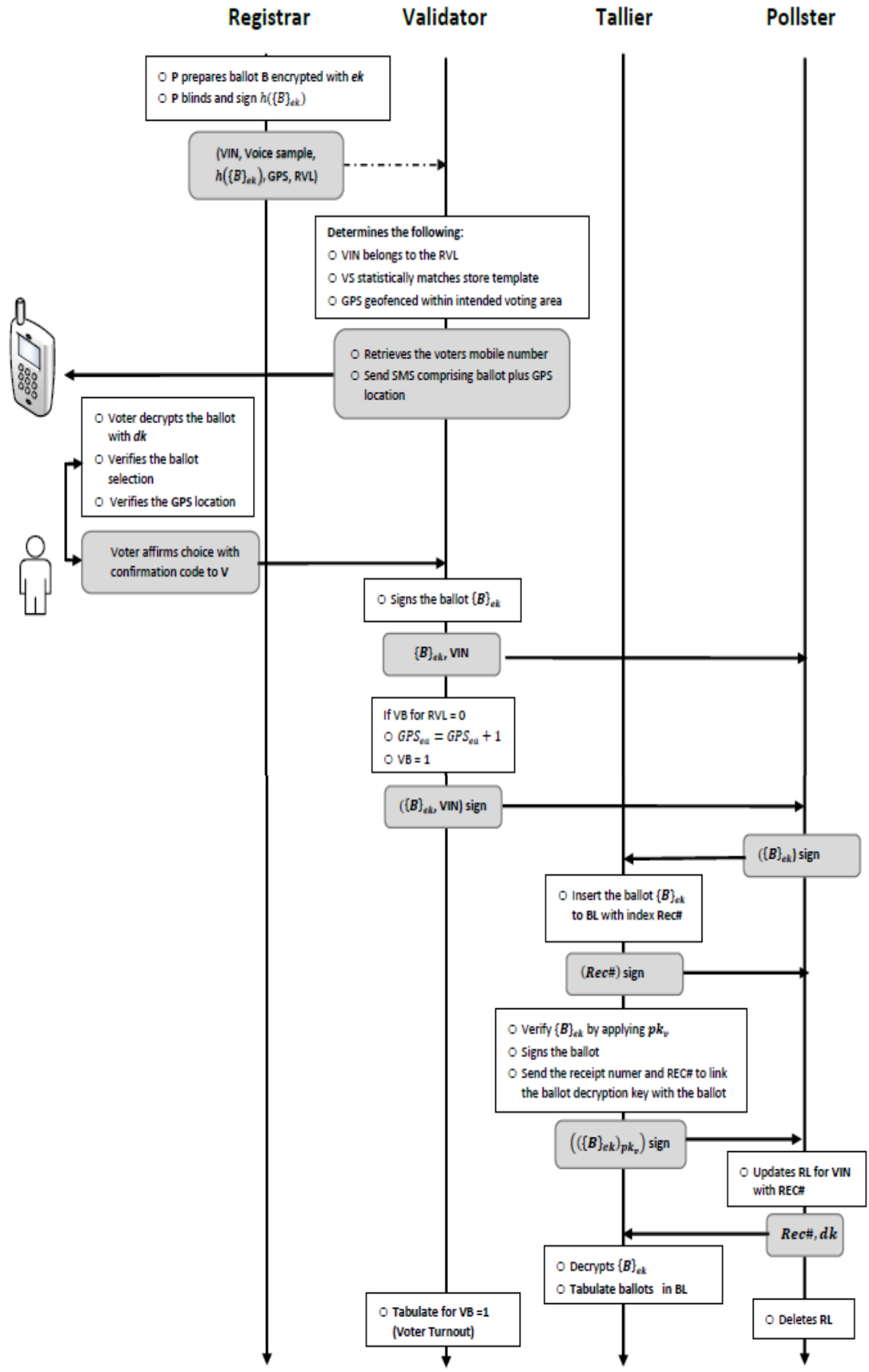

Figure 4.4 The SMIV Reference Architecture

The *Tallier,* **T** holds encrypted ballots in **BL**. *The Validator*, **V** holds the **RVL** supplied by the Registrar *with the Pollster,* **P** holding **RL**. **RL** has one entry per voter. In the event of a revote, the receipt number, **Rec#** for that voter is overwritten. SMIV emulates the electoral district or precinct used in conventional electoral processes by requiring the voter to declare the area from which they will vote. This is addressed using **GPS** and **GPSea** as per Figure 4.1. Whenever one electoral server **X** communicates with **Y**, it signs it with its private key $\{\ldots\}_{pk_x^{-1}}$. **Y** unsigns or verifies the signature with X public key, pk_x. This provides trusted communication and is assumed to occur in all communication between the three entities. This protocol will, subsequently for ease of reading, only include the critical signatures.

1. **P** prepares a vote ballot **B** which is encrypted key with **ek**. **P** then blinds the encrypted message, $h(\{\mathbf{B}\}_{ek})$. This encrypted message is signed by **P**. **P** sends the following to **V**: blinded encrypted message, **VIN**, Voice Sample, and GPS.
2. **V** checks the credentials of the aspirant pollster by determining the following:
   (i) V**IN** belongs to the **RVL**
   (ii) Voice Sample statistically matches the stored template
   (iii) **GPS** is geofenced within the intended voting area
3. **V** retrieves the mobile number from the n-tuple and sends an SMS comprising the ballot and **GPS** location to the voter. The voter decrypts the ballot with **dk** and verifies the ballot selection and the **GPS** location. The voter affirms with the confirmation code to **V** who
   a. Signs the ballot, $\{\mathbf{B}\}_{ek}$ and returns ($\{\mathbf{B}\}_{ek}$, **VIN** to **P**
   b. If the *Validation Bit*, **VB** for the matching entry in **RVL** = 0,
      i. $\mathbf{GPS_{ea}} = \mathbf{GPS_{ea}} + \mathbf{1}$ supporting post-election audits.
      ii. **VB = 1** will support voter turnout computation.
   c. **V** signs and sends ($\{\mathbf{B}\}_{ek}$, VIN) to **P**
4. **P** receives the confirmed tuple:
   i. **P** signs the ballot and sends it to **T**.
   ii. **T** inserts $\{\mathbf{B}\}_{ek}$ into a list of valid ballots, **BL** with index **Rec#**.
   iii. The **Rec#** is signed and sent back to **P**.

T cannot have the **VIN** as this is a direct link to the voter preserving voter privacy. It must however have a link to **dk**, which **P** has. Therefore the **RL** table of **P** needs to have **VIN**, **dk** and **Rec#**

T checks the honesty of **P**, by verifying that the encrypted ballot $\{\mathbf{B}\}_{ek}$ has been validated by **V** by applying pk_v

Finally, **T** signs the encrypted ballot and sends it back to **P** with a receipt number, **REC#** to link the ballot decryption key with the ballot itself.

P updates **RL** for voter **VIN** with **Rec#**

5. At the close of the election
   (i) **P** sends all tuples (**Rec#**, **dk**) to **T**. Using BL, T matches **Rec#** to $\{\mathbf{B}\}_{ek}$ and decrypts.
   (ii) **P** deletes **RL**
   (iii) All decrypted ballots in **BL** are tabulated. Remaining ballots in **T** remain encrypted and are not tabulated.
   (iv) At the end of the election period the **V** tabulates from the **RVL** all participating voters where **VB** = 1. This result is the *Voter Turnout*.

4.2.1.1 The Pollster

The **Pollster (P)** is the intermediary between the voter and the electoral servers. The Pollster is the only component that the voter must completely trust. Upon receipt of the **VIN** from the Registrar, it instantiates the voting process. This happens when the voter taps her NFC-enabled ID on the mobile, launching the voting application. The NFC token supports eligibility.

The Pollster now verifies the voter to mediate impersonation and support incoercibility. The NFC seamlessly transfers the voter VIN and public key to the Pollster while auto-launching the voting application. The Pollster uses the **VIN** to retrieve the corresponding voter n-tuple from the RVL. The Pollster sends the current mobile GPS coordinates with the VIN, to the Validator to ensure that the voter is within her declared electoral area. If the servers indicate the mobile is outside an area the application stops, partially supporting incoercibility. Otherwise, the authentication continues and the remote voter is requested to supply a current voice sample. The sample is used to perform a voice biometric probability match against

the corresponding voter's voice template from the RVL. If this match is statistically acceptable, the authentication process is complete and the voting selection module launched.

In acknowledgement of the time consumed for application loading and latency caused by high traffic, a section of the national anthem may be played to keep the user occupied and inform her that the voting process is proceeding satisfactorily. The anthem instils sanctity to the voting process and also elevates this engagement to more than a mere online interactive session (Bornman 2006; Habib 1997). The SMIV trial will, to the researcher's knowledge, be the first instance of an anthem being played during an online voting environment.

When the Validator accepts a confirmed ballot it signs it and returns it to the Pollster. The Pollster signs and sends this to the Tallier, which in turn sends the Pollster a Receipt number (**Rec#)**. The Pollster updates a list, RL which has the VIN, decryption key and Receipt Number (Rec#). After the election, the Pollster sends all tuples of (Rec#, decryption key) in RL to the Tallier. After sending all of these tuples, the Pollster deletes the RL to remove any voter-ballot links which ensures voting privacy.

As NFC represents a token, it makes the registration process like Sensus (supporting eligibility). The NFC also seamlessly transfers the voter's VIN (supporting convenience). Furthermore, it also auto-launches the voting application (supporting ease-of-use) on the mobile. The Pollster retrieves the current GPS (supporting incoercibility) and obtains voter features, such as voter voice spectral features, through the mobile to verify the voter (supporting democracy and ease-of-use). Pollsters implement the user interface on the voter's device to capture the ballot.

4.2.1.2 The Validator

The **Validator (V)** authenticates the remote user and ensures that one legitimate vote by each registered voter is counted while allowing the revote option. For every voter that exercised the revote option the last confirmed ballot will prevail. The Validation Bit is used to indicate if the voter has cast and confirmed a vote or not. Upon receipt of the encrypted ballot and VIN from the Pollster, the Validator uses the transmitted VIN to uniquely identify the voter in the RVL. The Validator will send an out-of-

bound SMS to the voter with her place of voting concatenated with her encrypted ballot. This SMS return message is also a contribution to previous models. She may check the place or location for correctness and verify the ballot with her private decryption key. This satisfies the end-to-end verifiability security requirement while providing location determinism, and supports incoercibility. If the voter replies, using the out-of-bound SMS, with her correct confirmation key, then the Validator does the following:

a) Changes the contents of the validation field from 0 to 1 to indicate a vote, and to prevent electronic ballot stuffing. This field will be available, post-election, to stakeholders to inform them which citizens participated in the electoral process.
b) Increments the electoral vote count in the corresponding geographical voting area by one.
c) Stores the vote in the corresponding voter's digital envelope overriding the previous one.

If no confirmation key is received the validation field remains zero implying that a vote, while cast, has not been confirmed. These may also be analysed post-election as votes cast but not confirmed.

The Validator will now have, at the end of the election period, a series of registered voters who did not participate, as well as voters who cast and confirmed ballots; it will also have a list of cast and not confirmed ballots. The Validator must publish this list. This public affirmation mediates collusion and ballot stuffing by the Validator addressing the shortcoming identified by Baiardi *et al.* (2005) that one of the entities can cast votes in the place of those abstaining. This is SMIV's contribution to the principal of incoercibility and democracy.

4.2.1.3 The Tallier

The **Tallier (T)** is responsible for collecting and tallying the ballots of the election. After receiving a blinded ballot from the Pollster, the Tallier checks the ballot to determine if it was signed by the Validator as a valid vote. If valid, the Tallier then stores the blinded vote in a digital envelope/slot and returns a receipt number to the Pollster. At the end of the election, it sends a list of receipt numbers to the Pollster.

The Pollster returns this list of receipt, RL numbers corresponding to the last vote cast by each voter. Each receipt number is tupled with the decryption key. Ballots in the Tallier with receipt numbers not corresponding to the receipt number list are discarded as duplicates. From this new list of receipt numbers from the Pollster, the Tallier sends the Pollster a receipt number and receives its matching decryption key in return. Using this decryption key, the ballot is now unblinded and tabulated. After the election, the system publishes a list of encrypted ballots, decryption keys, and decrypted ballots, allowing for independent verification of election results and post-election audits. The unconfirmed ballots are not tabulated and may be used in post-election analysis.

4.2.2 Evaluation of SMIV

The SMIV model is now theoretically evaluated against the following security requirements:

Accuracy: The vote cannot be altered as this will destroy the encryption keys and the signatures, which will be detected. The use of the confirmation key enables the voter to affirm her own vote selection. A voter can also verify and report if her vote was eliminated from the final tally, by inspecting the list of received votes published by the Tallier. The use of GPS supports post election audits and provides for contextual analysis of the voting accuracy.

Privacy: Different servers run validation, registration and tallying. SMIV mediates ballot-voter linkage by maintaining separate lists on each of these servers. All communication between servers is encrypted, using blind signature and data encryption using the RSAREF encryption library through anonymous channels. SMIV's return message does reveal the voter choice although the revote possibility of the revote, which makes determination of a voters ultimate choice problematic. The Ballot anonymity is preserved by deleting any voter-ballot links before tallying or publishing results.

Verifiability: The final tally can be made by anyone, verifying the signatures on the votes and summing all votes. Each voter can verify her own vote is correct, using the information saved. The out-of-band SMS sent to the voter supports end-to-end verifiability of the vote and its location.

Authentication: The remote voting method removes direct, natural person engagement, which mediates, to some extent, fraudulent activity by officials. The threat of fraudulent behaviour now moves to the cloud. Eligible remote voters are authenticated using a multimodal combination of the NFC-enabled voter's ID number, voice biometrics, and GPS location of their voting device. The use of voice as one mode of authentication is a neutral socio-political biometric capture method compared to face and fingerprint (Thakur and Dávila 2013). The use of voice authentication provides a proof-of-life option as well. The SMIV architectural model negates the additional technology requirement of other remote identification models such as e-ID which require at least one costly dedicated smart card reader of US$ 7 per household (Alvarez, Hall and Trechsel 2009); thus elevating and satisfying Mursi *et al.*'s (2013) requirements of cost effectiveness, practicality, scalability and convenience respectively. GPS location mitigates computer generated authentication attacks.

Convenience: The SMIV system should enable voters to vote easily, quickly, with minimal equipment, and with no special knowledge. The familiar mobile allows anytime, GPS-based voting. The NFC tag supports auto-coupling and seamless data transfer. An example of the reduced time requirement is the Estonian experience where traditional voting took 44 minutes while remote i-voting took just 6 minutes (Tsahkna 2013).

Mobility: There should be no geographical restriction with respect to where voters decide to cast their vote. This requirement also implies that the voting system is available and accessible during the voting phase, regardless of where the voter decides to cast his or her vote. On the other hand, one of the strengths of geographical electoral demarcation of areas is that it makes counting and post-election auditing possible. SMIV therefore allows a voter to inform where she will vote from which and accordingly applies a GPS restriction or geofencing.

Flexibility: An e-voting system can be said to be flexible if it allows diversities of ballot question formats, including open-ended questions. SMIV allows for mobile web page formats to represent a wide variety of candidate voting options. The mobile supports WML, HTML, XML and SMS. These formats allow for pragmatic ballot design, although the small form factor of the mobile presents a challenge to certain

classes of ballot, such as when there are a large number of candidates (Ekong and Ekong 2010).

Incoercibility: In an electoral process, coercion occurs when an entity tries to swindle a vote; decoy or influence a voter to abstain or vote for a particular candidate; or impersonate a valid voter by obtaining her credentials. Incoercibility is a vulnerability in remote voting deployments (such as Sensus) which is accentuated by SMIVs use of mobile remote voting. SMIV rebuts computer automated or botnet attacks through the GPS requirement. The use of voice authentication mitigates impersonation while providing a proof-of-life. The use of encryption and decryption keys reduces man-in-the-middle interception and ballot alteration, while the out-of-band confirmation key alerts even an abstaining voter of an illicit attempt to vote. The out-of-band message enables the voter to confirm both the location and choice of location from which a ballot-choice was cast, which permits the voter to affirm the validity of her vote. The revote option reduces the threat of vote buying, selling and subtle coercion. The NFC card supports voting from any appropriate mobile, and will assist a nervous voter to cast an unfettered ballot. The write-once capability reduces impersonation. Abdelkader and Youssef (2012) even suggest verification being sent to multiple devices to combat malware on a device. A coercer or voter buyer is unlikely to make the effort if a voter can simply vote again to override her previous choice. Some authorities use the in-person poll site vote as an ultimate option that can override all Internet votes, including remote votes cast later in time. Further the GPS service will help monitor the number of voters that can vote from a defined geographic location, which is a form of geofencing of voters, supporting the detection of ballot stuffing and post-election audits. Here a voter indicates where she will be during voting and accordingly may *only* vote from that area.

Receipt freeness: Although this is compromised, as only the voter is informed of her vote via the out-of-bound message, the onus is transferred to the voter to preserve her secret vote, perhaps by deleting the message. This has an e-voting analogy of the VVAT or voter verifiable audit trail. Also, the use of the return message in the SMIV architecture reduces collusion between administrator and Tallier as the voter has to be informed of her selection.

Fairness: The Tallier is permitted to only tabulate results at the end of the election. The SMIV architecture uses contactless and very close proximity based NFC, and secure channel and private ballot encryption to mediate monitoring. The use of the revote (multiple votes) from the same user, with the last vote counting, reduces the perceived usefulness of real-time monitoring and supports fairness.

Eligibility: Each voter can only vote once in each election because only one NFC tag is issued per voter. The Validator of SMS uses Validation bit in conjunction with the Tallier to control and remove duplicate ballots.

4.3 Comparison of Sensus, REVS and SMIV

Table 8.1 illustrates the evolution of the proposed SMIV architecture from the Sensus reference architecture and provides a summary of the realisation of the eight different security requirements that are satisfied by SMIV.

Table 4.1 The Sensus, REVS, and proposed SMIV architectures' fulfilment of the e-voting security requirements

Sensus	REVS	The proposed SMIV
Accuracy • Only one ballot gets counted per voter	**Accuracy** • Loosely coupled cooperating servers ensure greater accuracy because more than half of the servers must be simultaneously compromised • The servers cannot collude to affect election outcome	**Accuracy** • Only one of several identically encrypted ballots gets counted per voter • GPS determination prevents agenda-driven or disinterested hackers
Privacy • Blind signature and data encryption using the RSAREF encryption library • Different servers run Validator and Tallier • Pollster does not run on a machine that runs either Validator or Tallier • Installation of personal copy of Pollster on trusted machine by voter • Anonymous channel	**Privacy** • Chose Blind signatures over Homomorphic. The latter is more computationally intensive in a scaled environment • Multiple servers mitigates voter-ballot links • Ballot anonymity persevered by ballot shuffling, and random transmission delays undertaken by Anonymiser • Uses RMI over SSL	**Privacy** • Blind signature and data encryption using the RSAREF encryption library • Different servers run Registrar, Validator, and Tallier • Pollster does not run on a machine that runs either Registrar, Validator or Tallier • Installation of personal copy of Pollster on trusted machine by voter • Anonymous channel • Ballot anonymity preserved by deleting any voter-ballot links before publishing results
Verifiability • Publishing of a list of encrypted ballot, decryption keys, and decrypted ballots • Only voters can verify that their votes were counted correctly and correct any mistake anonymously	**Verifiability** • Publishing of a list of encrypted ballot, decryption keys, and decrypted ballots • Each voter can verify that their own vote is correct • The final tally can be made by anyone • A voter can verify the correctness of their vote and correct this anonymously	**Verifiability** • Publishing of a list of encrypted ballot, decryption keys, and decrypted ballots • Only voters can verify that their votes were counted correctly and report any anomalies • Confirmation key provides end-to-end verifiability to the voter
Authentication • VIN • Token • Public and private keys	**Authentication** • User name, password, and secret PIN	**Authentication** • Voice biometrics (Short term spectral features) • NFC, GPS, VIN • Public and Private keys • No need for special readers for e-ID cards

Convenience[8] • Familiar Workstation and user interfaces • Remote Voting • Password access	**Convenience** • Familiar workstations and user interfaces • Remote voting • Configurable ballots through XML • Voter marks ballot but must get Manager's signatures before submission • Username and password	**Convenience** • Familiar mobile devices and user interfaces • Remote mobile geofenced voting • Casting of vote in one or two sessions • Configurable ballots through XML • Anyplace voter registration • Anyplace precinct voting • Voter ID card affixed with NFC tags for auto-loading of voting application
Mobility • Networked workstation	**Mobility** • Internet-enabled device	**Mobility** • Internet-enabled or mobile device • Geofenced real-time monitoring of maximum voters per precinct • Geofence supports post election audits • Mediates coercion by overt or covert influence • Voice is inherent to mobile, with other biometrics requiring additional acquisition technology
Flexibility • Ballot description language (BDL)	**Flexibility** • XML	**Flexibility** • XML/HTML • WML • Plain SMS
Incoercibility • Not implemented	**Incoercibility** • REVS argues precinct voting is the only way to prevent possible coercion • Multiple servers prevent electronic ballot stuffing	**Incoercibility** • SMIV allows revotes to minimise remote coercion as well as electoral administrators • GPS place of voting • Out-of-band message to voter • Confirmation key • GPS service and geofencing is used to monitor and control the maximum number of voters that may be cast from a geographic area mitigating electronic ballot stuffing

[8] This is Table 4.1 (Continued)

Eligibility[9]	**Eligibility**	**Eligibility**
• Voter's ID number • Secret token • Blinded validation certificate and signed receipt to certify uniqueness of vote • Invulnerability is satisfied if voting is compulsory, that is all voters submit a ballot	• Each voter can only vote once because it is non-trivial to simultaneously deceive multiple usually anonymous servers • Blinded validation certificate and signed receipt to certify • Uniqueness of vote	• Voter's ID number • Voice biometrics (Short term spectral features) • GPS place of voting • Confirmation key • Duplicate votes removed to ensure one vote per voter • One NFC tag per voter • Blinded validation certificate and signed receipt to certify uniqueness of vote

4.4 Conclusion

This chapter presented the FOO, SENSUS, and REVS reference architectures, with a view to informing the rationale behind the design of a proposed evolutionary SMIV model. The SMIV protocol and its communicating components were presented. It was demonstrated that SMIV is an improvement with respect to the security requirements of authentication, eligibility, accuracy, privacy, verifiability, mobility, convenience, democracy, uniqueness, and incoercibility. The chapter ended with a tabular comparison of these requirements with respect to Sensus, REVs and SMIV.

[9] This is Table 4.1 (Continued)

Chapter Five - Theoretical Foundation of Voice Biometric Authentication

The theoretical foundation of the voice biometric authentication system is presented in this chapter. The physiology of voice, features that are extracted from voice, and some algorithms and methods that are usually considered to realise voice biometric authentication systems, are explained and discussed.

5.1 Voice Physiology and Features

The human speech system is made up of three major components, namely, (i) the lungs, (ii) the larynx containing the vocal folds and the glottis, and (iii) the vocal tract with the nasal and mouth cavities. All these components are shown diagrammatically in Figure 5.1.

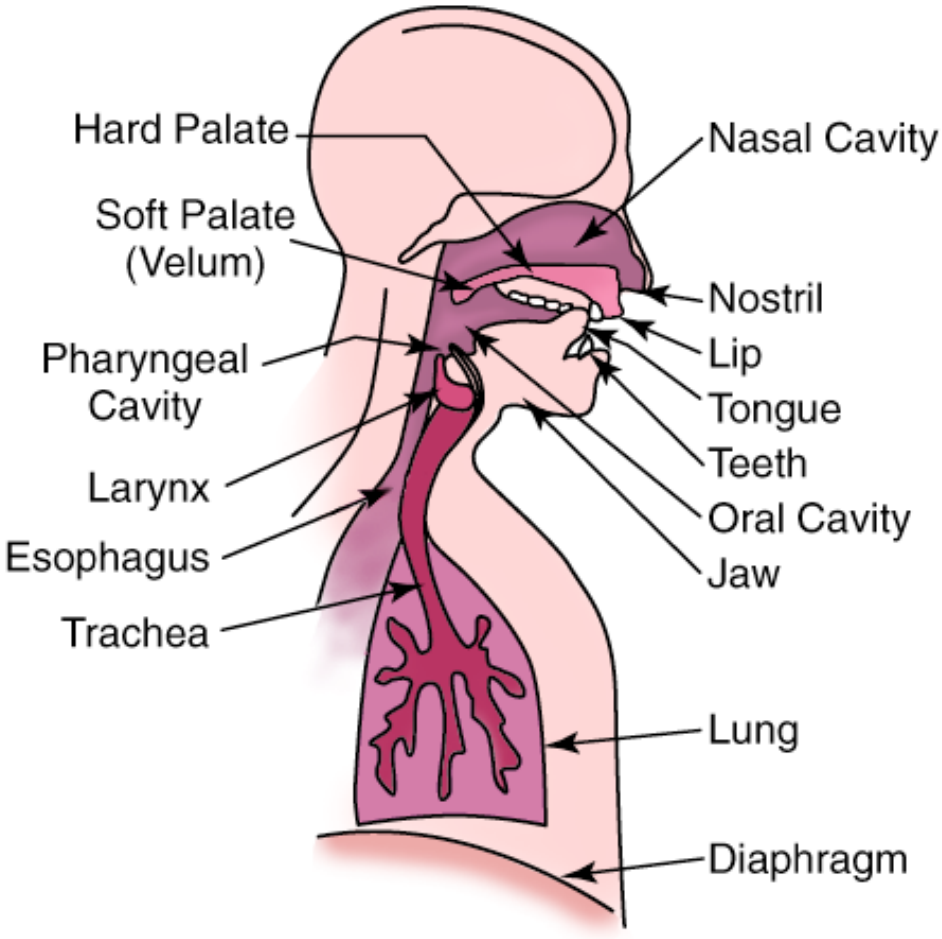

Figure 5.1 Human speech production system (Bouman 2009)

Voice or speech is produced when air is forced out of the lungs, up the trachea and into the vocal tract. The air from the lungs provides the source of power to set up the vibration of the vocal folds - which are two V-shaped, thin sheets of tissue that stretch from the back to the front of the larynx. They contain a space named the glottis and are positioned in various ways to generate speech sound. The physiological characteristics of human speech principally comprise of the vocal tract structures, which possess low within-person variance and strong between-person variance. As a result, physiological characteristics of the human speech system are good candidates for speaker recognition (Kain 2004). However, in a larger context, speaker recognition belongs to the field of voice biometrics, which refers to authenticating persons based on their physiological and behavioural characteristics. The features for voice biometrics can be divided into Prosodic features, High-level features, and Spectral features (Chang 2012; Karpov 2011; Reynolds 2002).

Prosody is a linguistic term for various features of the speaker - such as speaking rhythm, intonation stress, and emotional state of the speaker of the language - that may not be encoded by grammar. Prosodic features span over long periods of speech, such as syllables, words and phrases. Modelling prosodic features for speaker recognition is a challenging task, although recent studies indicate that prosody features improve speaker verification system performance (Chang 2012; Karpov 2011; Volkmann *et al.* 2011). The most important prosodic feature is the fundamental frequency (also called F0), which is defined as the rate of vibration of the vocal folds during voiced speech segments. The advantage of F0 is that it may be extracted even in noisy conditions. The F0 value contains information that is expected to be independent of the speech content. Therefore, combining F0 with spectral features should improve overall recognition accuracy especially in noisy conditions. However, using F0 related features alone have been reported to show poor recognition accuracy (Chang 2012; Karpov 2011).

Human voice characteristics differ not only due to physical properties of the vocal tract, but also due to speaking style and lexicon as well. Listeners can distinguish between familiar people much better, because of the speaker's idiosyncrasies, than between those they have never heard (Kinnunen and Li 2010). These individual speaking styles are referred to as high-level features. High-level features are not yet widely used in modern speaker recognition systems. However, with advancements in

speech technology, it is now possible to utilise efficient phone and word recognisers in the speaker recognition area as well. An overview of recent advances in this area is available in Karpov (2011) and Chang (2012).

The spectral features convey information about the speaker's vocal tract characteristics, such as the location and amplitude of the peaks or *formants* in the spectrum. Formants are the simplest, most discriminative, and most commonly used method in speaker recognition. The speech signal is 'seen' as a quasi-stationary or as a slowly varying signal. Therefore, for analysis, it is assumed to be stationary over relatively short intervals. Thus, spectral analysis involves framing the speech signal into short 20−30ms frames, with each frame overlapping by 10ms. The overlap is called a timeshift and each frame has N points (Chang 2012; Karpov 2011; Kinnunen and Li 2010; Kinnunen, Karpov and Fränti 2006). The formants, also called vocal tract resonances, are the peaks of the spectral envelope. Figure 5.2 shows a spectral envelope and its formants (Chang 2012; Kinnunen and Li 2010).

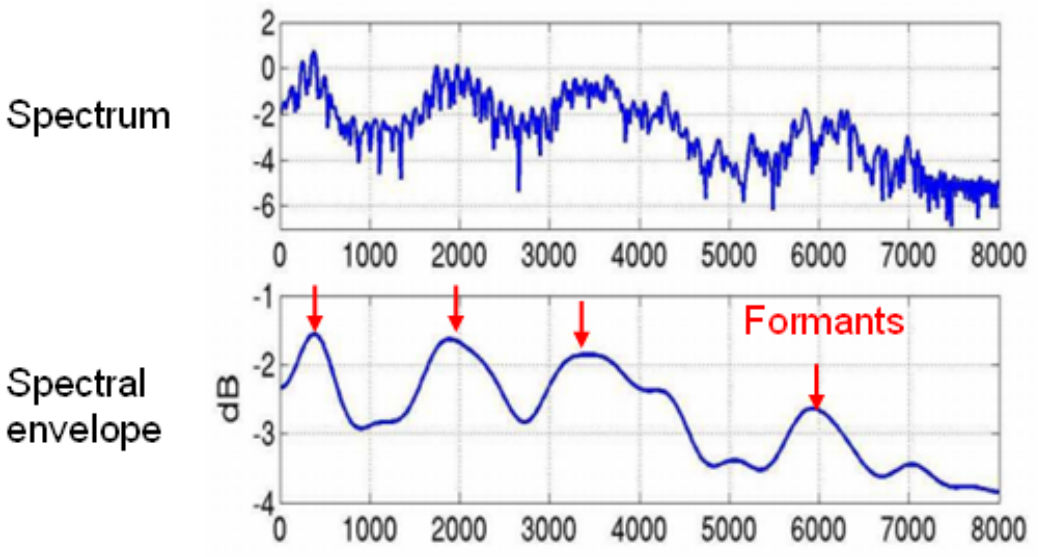

Figure 5.2 The spectrum, spectral envelope and formants (Chang 2012)

The spectrum X[k], equals the spectral envelope H[k], multiplied by the spectral details E[k]

$$|X[k]| = |H[\mathrm{k}]| * |E[\mathrm{k}]| \qquad (5.1)$$

In order to separate the spectral envelope and the spectral details from the spectrum, the log of both sides of the equation are taken as follows:

$$log|X[k]| = log|H[\mathrm{k}]| + log|E[\mathrm{k}]| \qquad (5.2)$$

The log spectrum is now the sum of a smooth signal (the spectral envelope) and a fast varying signal (the spectral details). The spectral envelope is the low frequency components of the log spectrum; i.e. the low frequency cepstrum coefficients. A cepstrum is the result of taking the Inverse Fourier transform (IFT) of the logarithm of the estimated spectrum of a signal as shown in Figure 5.3.

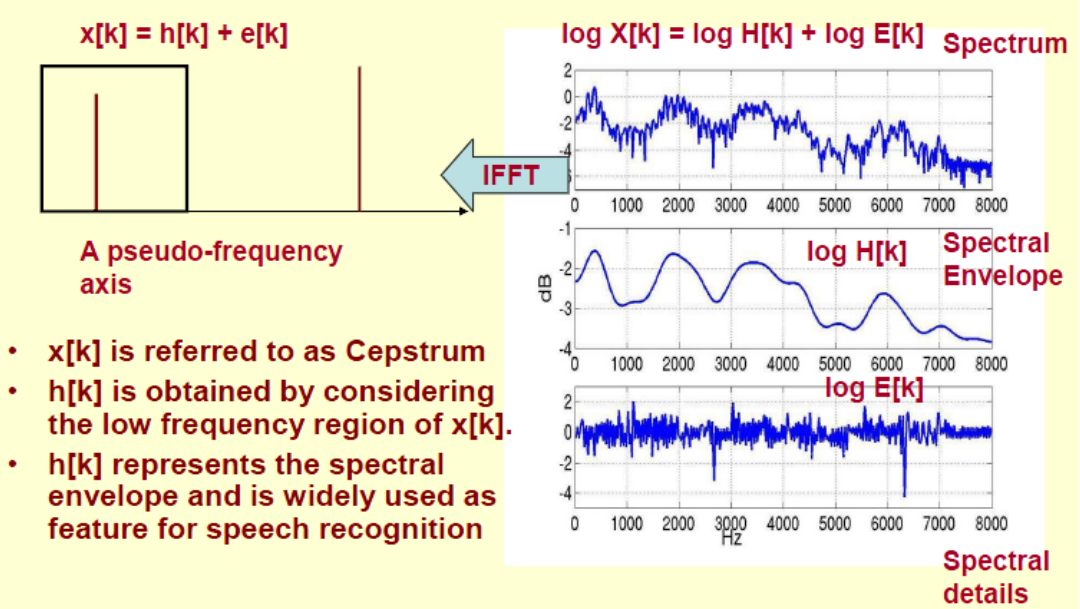

Figure 5.3 Computing the cepstrum coefficients (Chang 2012)

There are many spectral features that convey information about the cepstral of speech signals. These include Mel-Frequency Cepstral Coefficients (MFCC), Linear Predictive Cepstral Coefficients (LPCC) and Mel-Frequency Discrete Wavelet Coefficient (MFDWC) (Deller, Hansen and Proakis 2000; Karpov 2011; Campbell 1997). Each of these features will be thoroughly discussed in 5.2.1 in this thesis.

5.2 Voice Biometric Authentication

The generic block diagram for biometric authentication systems is shown in Figure 5.4. This block diagram also suitably captures the different components of a voice biometric authentication system.

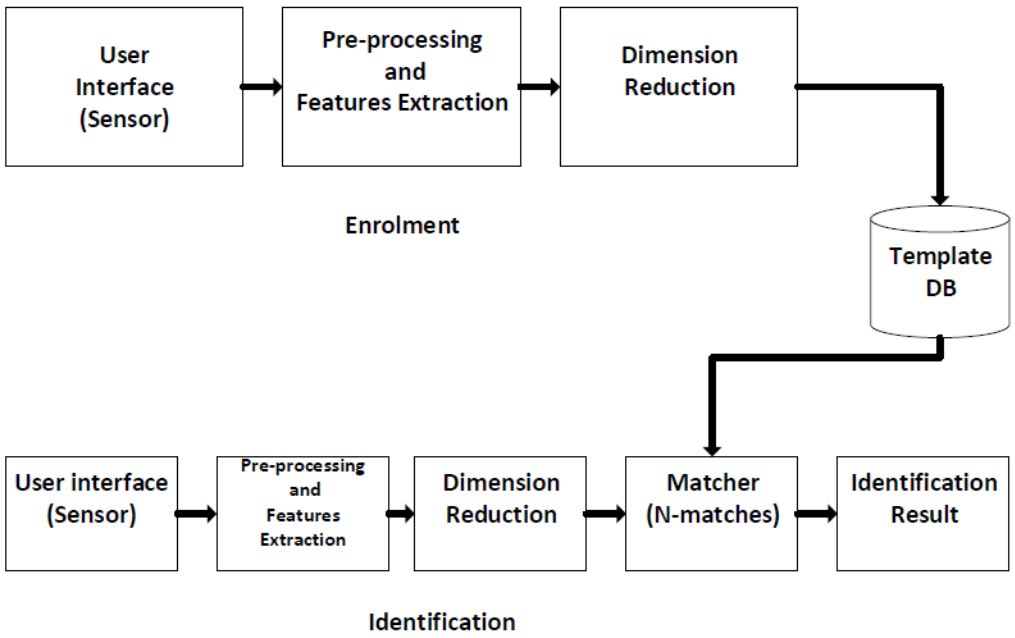

Figure 5.4 Generic block diagram for enrolment and identification of biometric authentication system

As shown in Figure 5.4, the system comprises the speech signal capturing with the interface sensor (microphone), the digitization of the captured analogue speech, pre-processing of the digitised speech, extraction of discriminating speech features, dimension reduction of the extracted features, training of pattern matching model, and identification of speakers through pattern matching with the trained model. These processes are carried out in two major stages, namely, the enrolment and identification stages. In the enrolment stage, the reduced voice features are stored as a voice template in the database or utilised to train the matching model; while in the identification stage, fresh voice features are extracted and matched with the database or trained model in order to ascertain the identity of the speaker.

Voice biometric recognition systems can either be text-dependent or text-independent. Text-dependent voice recognition systems are based on the utterance of a fixed word or phrase while text-independent systems distinguish an individual regardless of the uttered word or phrase. Text-independent voice recognition, even though very challenging to design, offers more protection against scams (Kain 2004). As earlier noted, voice biometrics are also referred to as speaker recognition, which is a different technology than speech recognition in which computer algorithms extract

features of the spoken utterance to determine the word that is spoken (Gaafar *et al.* 2014).

Given the foregoing, a *text-independent* speaker identification approach using short-term spectral features is nominated for fulfilling the authentication requirement of the proposed SMIV architecture in this study. This choice satisfies the biometric *negative recognition* paradigm, which prevents a single individual from using multiple identities (Wayman 2001; Mansfield and Wayman 2002). Consequently, this choice will also satisfactorily enhance the realisation of both the mobility and eligibility security requirements as per our architecture. The concepts and algorithms that are usually adopted for the pre-processing, features extraction, and pattern matching components of voice biometric authentication systems, are discussed in the subsequent subsections.

5.2.1 Pre-processing and Features Extraction

As earlier mentioned, voice biometric systems generally leverage both the physiological and behavioural characteristics of humans to carry out speaker recognition. In speaker recognition, features with high between-speaker variability are selected so that it can be easy to distinguish different speakers. Features that have low within-speaker variability are also strongly considered and this is particularly useful in identity claims. Speaker recognition systems must be robust against noise and distortion. The speech feature must be easy to measure, stable over time, occur naturally, and frequently exhibit little change from one capture environment to another (Chang 2012; Karpov 2011; Kinnunen, Karpov and Fränti 2006). However, the splitting of the continuous speech signal into short frames, which is required for extracting discriminative features, always creates discontinuity or abrupt changes at the frame edges. In order to prevent this, a common pre-processing method is to apply a window function such as a Hamming function. One may also pre-emphasise each frame to boost higher frequency components whose intensity would be otherwise be low due to the downward sloping spectrum of the glottal voice source (Karpov 2011; Deller, Hansen and Proakis 2000). Once the digitised signal is pre-processed, the feature extraction process transforms the signal into feature vectors in which speaker-specific properties are emphasized and statistical redundancies suppressed. The feature specifics properties have hitherto been the dominant features

in speaker recognition systems because of their stability, ease of extraction, and requirement of a small amount of data, text and language independence, and less computational requirements. As pointed out in Section 5.1, the most prominent features in the literature are Mel-Frequency Cepstral Coefficients (MFCC), Mel-Frequency Discrete Wavelet Coefficients (MFDWC), and Linear Prediction Cepstral Coefficients (LPCC) (Rose 2002; Wolf 1972).

Mel-frequency is the measure of the human perception of the frequency content of speech signals on the "Mel-scale". Mel-Frequency Cepstrum (MFC) stands for the power spectrum of the speech, based on a linear cosine transform of a log power spectrum, computed on the non-linear Mel-frequency. The MFCC features are obtained by taking the log of the outputs of a Mel-frequency filter bank, which is subsequently subjected to cepstrum analysis as earlier illustrated in Figure 5.3. The final MFCC feature vectors are obtained by retaining about 12-15 lowest DCT coefficients. Each vector is independent of each other and ordering information is lost. The MFCCs are, therefore, the coefficients that collectively make up the MFC. The frequency bands in the MFC are equally spaced and from research findings in the psychophysical field, it has been established that the Mel scale approximates the auditory system of humans better than a linearly spaced frequency band. Mel-frequency warping of the spectrum gives emphasis on low frequencies that are more important for speech perception by humans. The computational components of the MFCC algorithm are captured in Figure 5.5 (Malode and Sahare 2012; Karpov 2011) and explained as follows.

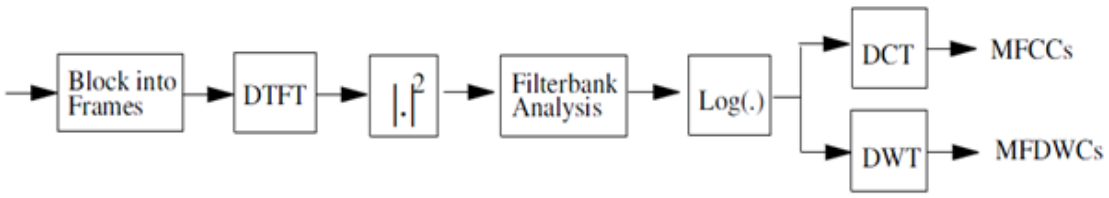

Figure 5.5 Extractions of the MFCCs and MFDWCs (Chang 2012)

Assuming x[n] is the digitised version of the input speech signal with sampling frequency f_s and is divided into P frames each of length N samples with an overlap of N/2 samples; x_p denotes the p^{th} frame of the speech signal x[n] and to compute the

ceptral coefficients of the p^{th} frame, x_p is multiplied with Hamming window. The Hamming window is given as:

$$w(n) = 0.54 - 0.46\cos(\frac{2\pi n}{N-1}), \quad n = 0, 1, \ldots, N-1 \tag{5.3}$$

The windowing function is purposely for smoothening of the signal for the computation of the Discrete Time Fourier Transform (DTFT). The DTFT is used for computing the frequency response of each frame to generate the spectrogram of the speech signal as:

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}, \tag{5.4}$$

The relationship between the Mel-frequency and linear frequency is:

$$mel(f) = \begin{cases} 2595\log_{10}(1+f/700) & if \quad f > 1kHz \\ f & if \quad f < 1kHz \end{cases} \tag{5.5}$$

where mel(*f*) is the Mel-frequency scale and *f* is the linear frequency. The Mel-filter bank filters the magnitude spectrum that is passed to it to give an array output called Mel-spectrum. Each of the values in the Mel-spectrum array corresponds to the result of the filtered magnitude spectrum through the individual Mel-filters. The Mel-spectrum is given as:

$$Y(n) = \sum_{k=0}^{N/2} \left|X[k]\right| * MelWeight[n][k], \qquad 0 < n < M \tag{5.6}$$

where M represents the number of filters. The MFCC features are computed by firstly, taking the log of the Mel-spectrum and then computing the DCT as follows:

$$C_n = \sum_{n=1}^{N} [\log Y(n)]\cos[k(n-\frac{1}{2})\frac{\pi}{N}] \quad \forall \quad k = 1,\ldots,M \tag{5.7}$$

The C_0 is omitted from the DCT computation because it represents the mean value of the input speech that contains little speaker unique information, but rather contains information on the microphone used for recording the speech signal (Deller, Hansen and Proakis 2000; Huang, Acero and Hon 2001). The MFCC feature vector is obtained per speaker by retaining about 12-15 lowest DCT components. Mel-

Frequency Discrete Wavelet Coefficients (MFDWC) are computed in the similar way as the MFCC features. The only difference is that a Discrete Wavelet Transform (DWT) is used to replace the DCT in the last step. Figure 5.5 shows the algorithms for MFCC and MFDWC.

MFDWC features are calculated using similar procedures to the computation of MFCC features. However, the DCT computation in the last step is substituted with the DWT, as shown in Figure 5.5. DWT is acclaimed to allow better localization in both time and frequency domains and based on this, the MFDWC has been shown to give better performance in noisy environments (Bai *et al.* 2012). MFDWCs were used in speaker verification and it was shown that they give better performance than the MFCCs in noisy environments. An explanation for this improvement is that DWT allows good localization both in the time and frequency domain.

Linear Prediction Coding (LPC) is an alternative spectrum estimation method to DTFT. LPCC is also known as all-pole model or the autoregressive (AR) model where the outputs are linearly related to its previous values. It has a good intuitive interpretation both in time domain (adjacent samples are correlated) and in frequency domain (all-pole spectrum corresponding to the resonance structure). Given a signal, s[n] in the discrete time domain, the LPC prediction error is given as:

$$e[n] = s[n] - \sum_{k=1}^{p} a_k s[n-k] \qquad (5.8)$$

where a_k are the coefficients of the predictor.

Assuming s[n] is the speech signal and e[n] is the voice source (or glottal pulses) (Bai *et al.* 2012) as shown in Figure 5.6. Equation (6) is transformed to:

$$E[z] = S[z](1 - \sum_{k=1}^{p} a_k Z^{-k}) \qquad (5.9)$$

The spectral model is therefore given as:

$$H(z) = \frac{1}{1 - \sum_{k=1}^{p} a_k Z^{-k}} \tag{5.10}$$

where a_k are the predictor coefficients that are often computed by minimizing the residual energy using the Levinson-Durbin algorithm (Harrington 1999), and *H(z)* is the spectral model. However, these predictor coefficients are infrequently used as features; rather, they are transformed to the more robust LPCC features using a recursive algorithm proposed by Rabiner and Juang (1993). However, unlike MFCC, the LPCC are not based on perceptual frequency scale, such as Mel-frequency scale (Chang 2012). The three features (Prosodic, High-level, and Spectral features) are experimentally compared in Chapter 6 in order to make the best choice of features for the voice biometric aspect of the SMIV architecture.

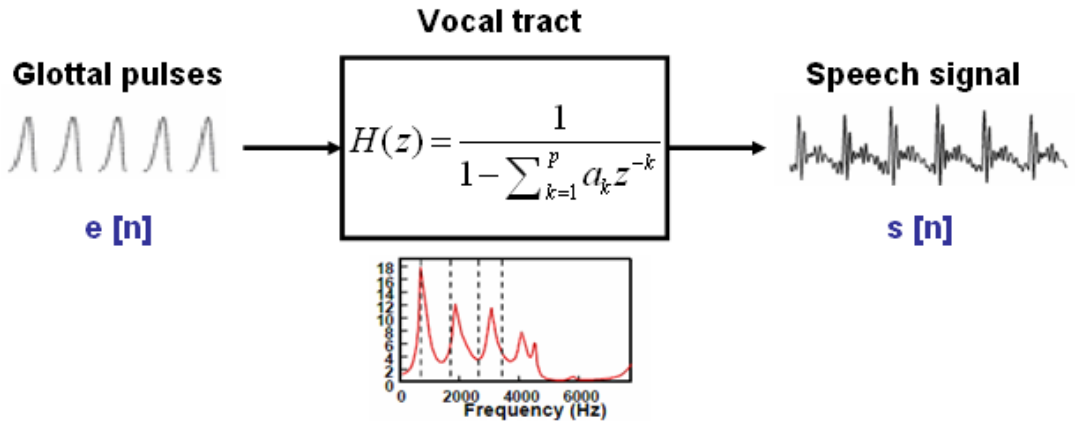

Figure 5.6 Computing the LPCC (Chang 2012)

5.3 Dimension Reduction

The outputs of the feature extraction algorithms described in Section 5.2 (i.e. MFCC, MFDWC and LPCC) are essentially 2-dimensional matrices that can be described analogously as 2-dimensional digital image signals (Kumar 2003). More so, spectral features have been described as the acoustic correlate of the 'colour' of sound in (Kinnunen and Li 2010). Therefore, these matrices can be processed further using digital image processing algorithms to achieve dimension reduction of the speech features. One benefit of this dimension reduction is feature reduction that prevents training samples from growing exponentially as the number of features increase – a

phenomenon that is referred to as the curse of dimensionality. Another benefit of dimension reduction is the decreased computational complexity of the pattern matching models in the speaker recognition systems (Kinnunen and Li 2010; Bellman 1961). Further benefits include mediating issues of poor communication channels, security, and avoiding deliberate compression, which may create loosely-data where some of the original features of the voice are lost (Hecht 2014). The transmission of fewer features over a network is useful as it results in a concomitant reduction of audio file traffic to the server, which further reduces traffic congestion. Examples of dimension reduction methods in signal processing, image processing, and statistics include Principal Component Analysis (PCA), Factor Analysis (FA), Independent Component Analysis (ICA), Projection pursuit, Random projections, and the Histogram of Oriented Gradient (HOG).

The HOG is a recent descriptor developed by Dalal and Triggs (2005) that can effectively capture the local appearance and shape information by encoding the spectral gradient orientation from the output of the short-term features as histograms. The algorithm has been reputed to be successful in recent applications such as speech processing (Kobayashi, Hidaka, and Kurita 2008; Das 2014) This algorithm is adopted for the dimension reduction task in this thesis and further details on the algorithm is presented in the subsequent subsection.

5.3.1 Histogram of Oriented Gradients

Histogram of oriented gradients (HOG) is an image processing technique initially developed to detect edges around objects, but it can also be used to compress data. An image is divided up into overlapping, evenly-sized cells. The changes in pixel values within a cell are calculated from the x and y axis to form a gradient vector. The magnitude and direction of this gradient vector can be calculated from this change in pixel values. Although the pixel values may change in magnitude, in the example of increased brightness or magnitude, the relative difference between pixels remain. Hence, this difference is used to distinguish edges of objects within an image and remains invariant to magnitude. The magnitude of a larger area (block of cells) is calculated and the value is applied to all pixels within that block (normalisation). In voice, the same technique is used. The gradient vectors of the spectral features of voice are calculated per given area or window. Using block normalisation, the

number of values that need to be recorded are greatly reduced (Dalal and Triggs 2005).

The first step in the HOG process involves the calculation of gradient values, which are computed to apply the finite difference approximation, also called the derivative masks, on the inputs. Using the 1-D centred mask, which was demonstrated to be superior to the Sobel or diagonal mask by Dalal and Triggs (2005), the input I is filtered both horizontally and vertically with kernels as:

$$\left.\begin{aligned} D_x &= [-1\ 0\ 1] \\ D_y &= [1\ 0\ -1]^{\mathrm{T}} \end{aligned}\right\} \tag{5.11}$$

where $[.]^{\mathrm{T}}$ is a transpose vector. The x and y derivatives of I are then attained through the convolution operations as:

$$\left.\begin{aligned} I_x &= I * D_x \\ I_y &= I * D_y \end{aligned}\right\} \tag{5.12}$$

The magnitude and orientation of the gradient of I respectively are calculated using the following formulae:

$$|\mathrm{G}| = \sqrt{\mathrm{I}_x^2 + \mathrm{I}_y^2} \tag{5.13}$$

$$\theta = \arctan\left(\frac{I_y}{I_x}\right) \tag{5.14}$$

The next step in the HOG algorithm is orientation binning where the cell histograms are created. Cells in HOG are usually rectangular, although they may be circular in some situations, and histogram channels may either be signed or unsigned. Signed histogram channels are distributed from 0 to 180 degrees while unsigned channels are distributed from 0 to 360 degrees. Using the value from the gradient calculation in the first step, each pixel within the cell provides a weighted vote for an orientation-based histogram channel (Adetiba and Olugbara 2015).

The third step of the HOG calculation is the establishment of descriptor blocks. The cell-orientation histograms are clustered into larger and spatially connected blocks before they can be normalised. The purpose of this clustering is to determine changes

in contrasts. The blocks are either rectangular (R-HOG) or circular (C-HOG) in shape. The R-HOG shape is often a square grid that can be characterised by the number of cells per block, the number of pixels per cell, and the number of cells per histogram. These blocks overlap each other by half of the size of a block.

The final step in HOG calculation is block normalisation. The normalisation vector for a non-normalised vector (v) that contains the histogram in a given block is one of the following norms:

$$\text{L2norm: } f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}} \tag{5.15}$$

$$\text{L1 norm: } f = \frac{v}{\|v\|_1 + e} \tag{5.16}$$

$$\text{L1-sqrt: } f = \sqrt{\frac{v}{\|v\|_1 + e}} \tag{5.17}$$

Where e is a constant whose value will not influence the result. Dalal and Triggs (2005), in their human recognition experiment, determined that the L2-norm and L1-sqrt methods performed similarly while L1-norm performance was the poorest. Thus, the HOG descriptor is the vector, which contains the normalised cell histograms from all of the block regions in the image (Adetiba and Olugbara 2015).

Dalal and Triggs (2005) studied feature sets for human detection, which show that locally normalized HOG descriptors provide excellent performance relative to other existing feature sets including wavelets. Selvan and Rajesh (2012) extended HOG to a spectral band calling it Spectral Histogram of Oriented Gradient (SHOG). The SHOG combines the vital components of both HOG and MFCC to generate highly compact speech features. Experiments were performed using the SHOG with a speech dataset of an ethnic Tamil group in India, and found good improvements in classification of the group into male or female gender. The work of these authors provided a motivation for this study to utilise HOG as dimension reduction algorithm for short-time spectral features (Selvan and Rajesh 2012). The SHOG features were also considered for experimentation in this study and compared with the MFCC-

HOG, MFDWC-HOG and LPCC-HOG dimensionally reduced features. Details of the experiments are reported in Chapter 6.

5.4 Pattern Matching

Pattern matching plays an important role in speaker recognition. The sequences of acoustic vectors that are extracted from individual speaker's speech are patterns, and in the aforegoing sections, the MFCC-HOG, MFDWC-HOG and LPCC-HOG features are examples of speech patterns. Pattern matching involves the comparison of features from input speech with Speaker models so as to recognise a speaker. Assuming the training vectors were used to create a speaker model represented as $\Re$ and test vector extracted from an unknown person speech sample is $\chi = \{x_1, \ldots, x_T\}$, pattern matching task involves the definition of a match score $\mathcal{S}(\chi, \Re) \in \mathbb{R}$ indicating the similarity of χ and $\Re$. The class of the model is what determines if the matching score will be a likelihood (probability), membership (classification) or dissimilarity (distance) value.

Speaker models and pattern matching techniques in the literature are divided into three categories, namely, i) template model, ii) stochastic model, and iii) discriminative model (Campbell 1997). Template model is based on the assumption that feature vectors are inexact replicas of the template, and in this technique, the training and test vectors are compared by measuring the distance between them. Examples of these techniques are Dynamic Time Warping (DTW) and Vector Quantization (VQ) (Soong, Juang and Rabiner 1987; Furui 1981). Stochastic model assumes that speaker voice is a probabilistic source with a fixed probability density function. Training vectors are employed to estimate the parameters of this function and in the testing phase, the conditional probability is evaluated. Eminent examples of stochastic models are the Gaussian Mixture Model (GMM) and Hidden Markov Model (HMM) (Reynolds, 1995; 2009; Naik, Netsch and Doddington 1989). In discriminative model, which is a more recent technique in the speaker recognition literature, the boundaries between speakers are modelled. Prominent examples of this technique are Artificial Neural Network (ANN) and Support Vector Machine (SVM) (Farrell, Mammone, and Assaleh 1994; Campbell, Sturim and Reynolds 2006a). Some of these techniques are discussed in the following subsections.

5.4.1 Vector Quantization

Vector quantization involves the conversion of training data into a codebook. Given a feature space with partitions that represent a different cluster of the training data and each cluster having a centroid, the collection of the centroids is called a codebook. The codebook ultimately represents the Speaker model (Ramachandran *et al.* 2002). Assuming the test template for a speaker recognition system is $\chi = \{x_1, \ldots, x_T\}$ and the reference template is $\Re = \{r_1, \ldots, r_k\}$, using vector quantization theory, the average quantization distortion of χ using $\Re$ as the quantizer is given as:

$$D_Q(\chi, \Re) = \frac{1}{T}\left(\sum_{t=1}^{T} \min_{1 \le k \le K} d(x_t, r_k)\right) \tag{5.18}$$

Where d(.,.) is a distance measure for vectors. Weighted distance measures of the form represented in Equation (19) is often used.

$$d_W^2(x, y) = (x - y)' W (x - y) \tag{5.19}$$

Where $\boldsymbol{W}$ is a weighing matrix used for accentuating discriminative features. Euclidean distance produces a case in which $\boldsymbol{W}$ is an identity matrix. Other distance measures such as Mahalanobis distance are also utilised for Equation (5.19) (Karpov 2011).

5.4.2 Hidden Markov Method (HMM)

HMM can be used to model a random system that changes states according to a transition rule that only depends on the current state. The HMM is a doubly stochastic process with an underlying stochastic process that is not observable (it is hidden), but can only be observed through another stochastic process that produces a set of observable data (Rabiner and Jaung 1986). Let S represent a state alphabet set, and V, the observation alphabet set:

$$S = (s_1, s_2, .., s_n) \tag{5.20}$$

$$V = (v_1, v_2, .., v_M) \tag{5.21}$$

Q is defined as a fixed sequence of length T with the matching observations *O*:

$$Q = q_1, q_2, \ldots, q_T \quad (5.22)$$

$$O = o_1, o_2,..,o_T \quad (5.23)$$

The formal definition of a HMM is (Blumson 2004):

$$\lambda = (A, B, \pi) \quad (5.24)$$

A is a transition array that holds the probability of state *j* following state *i.* *S*tate transition probabilities is autonomous in time and A is defined as:

$$A= [a_{ij}],$$

$$a_{ij} = P(q_t = s_j \mid q_{t-1} = s_i)\ 1<=i,j<=N \quad (5.25)$$

B is the observation array that holds the probability of observation *k* being created from the state *j*, autonomous of *t*:

$$B = [b_i(k)],$$

$$b_i(k) = P\ (x_t = v_k \mid q_t = s_i) \quad (5.26)$$

π is the initial probability array given as:

$$\Pi = [\Pi_i],$$

$$\Pi_i = P(q_1 = s_i) \quad (5.27)$$

To generate a model of a given observation sequence that would be generated by an appropriate HMM, the following steps must be followed.

1. Choose an initial state q1=So according to the initial state distribution Π
2. Set t=1
3. Choose $O_t=V_k$ according to the symbol probability distribution in state S_i
4. Transit to a new state $q_{t+1} = S_j$ according to the state transition probability distribution for state Si, (ex: a_{ij})
5. Set t = t + 1; return to step 3 if t < T; otherwise end procedure

A complete specification of an HMM requires two model parameters (N and M), observation specifications, and three probability measures A, B, and Π. Supervised

learning involves adjusting the model parameters of (A, B, Π) to maximise the probability of a sequence of observed states and explain how these observed states came about (Blumson 2004).

To determine the parameter A that would maximize the log-likelihood of the sequence of observed states, we use the formula:

$$\hat{A}_{ij} = \frac{\sum_{t=1}^{T} 1\{z_{t-1} = s_i \wedge z_t = s_j\}}{\sum_{t=1}^{T} 1\{z_{t-1} = s_i\}} \quad (5.28)$$

The formula provides the maximum likelihood probability of a change from state *i* to state *j* as the number of transitions from *i* to *j* divided by the total number of times that we changed to state *i*. This probability is related to the fraction of time that while in state *i*, the model changed to state *j* (Blumson 2004).

In practical terms, the maximum likelihood parameters based on a speech recognition dataset need to be developed to allow the effective training of the HMM before the maximum likelihood state assignment of the speech signal is determined. The maximum likelihood state transitions of A_{ij} (expected number of transitions from s_i to s_j divided by the expected number of appearances of s_j) B_{jk} is calculated as the expected number of outputs of V_k from S_j divided by expected number of appearances of s_j as follows (Rabiner 1989):

$$\hat{A}_{ij} = \frac{\sum_{t=1}^{T} \alpha_i(t) A_{ij} B_{j\,x_t} \beta_j(t+1)}{\sum_{j=1}^{|S|} \sum_{t=1}^{T} \alpha_i(t) A_{ij} B_{j\,x_t} \beta_j(t+1)} \quad (5.29)$$

$$\hat{B}_{jk} = \frac{\sum_{i=1}^{|S|} \sum_{t=1}^{T} 1\{x_t = v_k\} \alpha_i(t) A_{ij} B_{j\,x_t} \beta_j(t+1)}{\sum_{i=1}^{|S|} \sum_{t=1}^{T} \alpha_i(t) A_{ij} B_{j\,x_t} \beta_j(t+1)} \quad (5.30)$$

These algorithms seek relatively simple maximum likelihood (or maximum *a posteriori*) optimization targets. For complex HMMs, the parameter space may be complex with many spurious local optima that can trap a training algorithm (Eddy 1998).

5.4.3 Support Vector Machines

A support vector machine (SVM) is often used for pattern recognition. Unlike other models that focus on minimising the training error, SVMs work on another requirement, called structural risk minimisation, which minimises an upper bound on the generalisation error (Osuna, Freund and Girosi 1997). SVM uses a specific type of function class classifiers with big "margins" in a feature space prompted by a kernel (Schölkopf and Smola 2002). Training a SVM is similar to solving a linearly constrained quadratic programming problem where the number of variables is equal to the number of data points. A problem emerges when the numbers of data points become large (Osuna, Freund and Girosi 1997). In order to overcome this issue, given many training examples, SVMs break down the problem into a group of smaller tasks. The problem is usually broken down into inactive and active ("working set") parts (Joachim 1999). SVMs use a two class classifier which is developed from sums of a known kernel function $K(\cdot,\cdot)$ to define a hyperplane as follows:

$$f(x) = \sum_{i=1}^{N} a_i y_i K(x, x_i) + b \tag{5.31}$$

Where $y_i \epsilon[1,-1]$ are the target values, $\sum_{i=1}^{N} aiyi = 0$, and $a_i>0$. The vector $x_i \subseteq R^n$ are derived from training and serve as support vectors. The purpose of the hyperplane is to separate selected points into the two predefined classes.

Given a training set

$$S = \{(x_1,y_1),...,(x_l,y_l)\}_{i=l}^{l} \subseteq (X * Y)^{1} \tag{5.32}$$

and a kernel function is given as:

$$K(x_i,x_j) = < \Phi(x_i)\ \Phi\ (x_j)> \tag{5.33}$$

where <.,.> indicates the inner product and Φ plots the inner space X to another higher dimensional feature space F. With Φ appropriately chosen, the selected non-linearly separable sample S may be linearly separated in F. The hyperplane, with the smallest generalisation error, between these two classes is where the margin is the

sum of the distances of the hyperplane from the closest point of the two classes (Osuna, Freund and Girosi 1997).

5.4.4 Artificial Neural Networks

Artificial neural networks (ANNs) is a general pattern recognition mechanism that can learn to discriminate between categories. This implies the usefulness of the neural network model in solving spectral pattern variation problems (Huang *et al.* 1998). ANNs have been used for several years in both speech and speaker recognition systems because of their high accuracy, noise tolerance, and non-linear property (Gaafar *et al.* 2014). A neural network needs to be given only raw data related to the problem; the network sorts through this information, develops ANN learning rules which are used to determine relationships in the data, and from these rules, ANNs develop a model based on the understanding of factors that impact the result. Using this model, ANNs can provide predicted output given an unseen samples of these key factors (Rameshkumar and Samundeswari 2014). ANN generally uses a set of layers of neurons, interconnected via connections links, to process information and pass it on to the next layer. The links have associated weights, each of which is multiplied along with the incoming signal (net input) and the output signal is obtained by applying activation to the net input. The single layer neural net has two input neurons (x_1, x_2), interconnection weights (w_1,w_2) and one output neuron (y) as shown in Figure 5.7 (Rameshkumar and Samundeswari 2014).

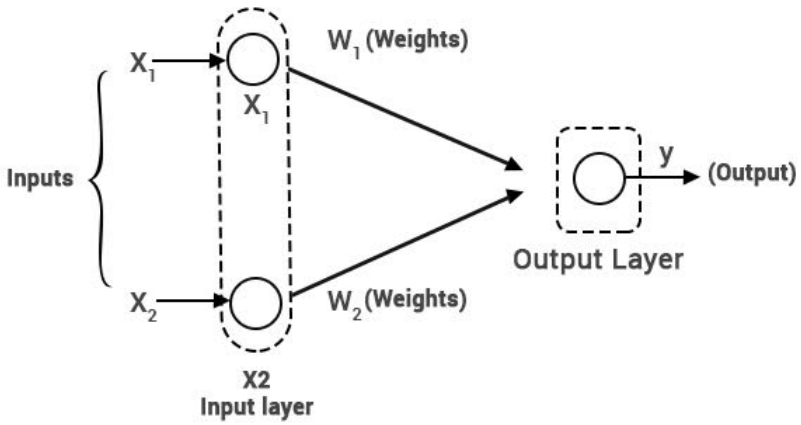

Figure 5.7 A simple artificial neural network (Rameskumar and Samundeswari 2014) (Reproduced)

In a multi-layer model, an artificial neuron has a set of *n* synapses which are linked with their inputs ($x_1,..,x_n$) with each input associated with a weight (w_i). A signal at input *i* is multiplied by the weight w_i; these weighted inputs are summed, and a linear combination of the weighted inputs is obtained. A bias (w_0) that is not linked to any input is added to this linear combination. A weighted sum Z is obtained as follows:

$$Z = w_0 + w_1x_1 + \ldots + w_nx_n \tag{5.34}$$

Successively, a nonlinear activation function *f* is applied to the weighted sum in order to create an output *y*

$$y = f(Z) \tag{5.35}$$

The activation function enables the flexibility and ability of an artificial neuron to approximate functions to be learned. The linear activation function is usually applied in the outer layer with the form:

$$f(z) = z \tag{5.36}$$

Sigmoid-activation functions are S-shaped with the most common being logistic and hyperbolic with their respective representations as follows:

$$f(z) = \frac{1}{1+e^{-az}} \tag{5.37}$$

$$f(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}} \tag{5.38}$$

Multi-layer perceptron (MLP) is a frequently used ANN. It is a non-linear ANN, which consists of layers of neurons. Each MLP consists of at least three layers, which constitute an input layer, one or more hidden layers, and an output layer (Dlashmit 2005). When designing MLPs, the number of hidden layers is critical for its success. Cybenko (1989) demonstrated that a MLP with one hidden layer can approximate any continuous function. Popescu, Balas, Perescu-Popescu and Mastorakis (2009) on the other hand, showed that for large problems, more hidden layers reduce network errors based on acceptance of few local minima.

The selection of the most suitable activation function for the neurons in the various MLP layers is very important for the neural network performance. The linear

activation function is generally used for the input neurons because it transmits the input database directly to the next layer with no transformation. The type of activation function for the output layer neurons is dependent on the function of the problem being solved (Popescu *et al.* 2009). The training dataset for MLP typically consists of patterns (x_p,t_p) where *p* represents the number of patterns and x_p is the N-dimensional input vector (Adetiba and Olugbara 2015). In this thesis, the input vectors would contain the HOG-reduced vector of extracted features of a given voice.

MLP networks are usually trained with a backpropagation algorithm - which is the most popular method of ANN learning (Che, Chiang and Che 2011; Rehman and Nawi 2011). In backpropagation, the network's neuron layers send their signals "forward" and their errors are propagated "backwards". The network receives inputs by neurons in the input layer, while the output (of the network) is by the neurons on the output layer. The backpropagation algorithm uses supervised learning, which implies that the algorithm is provided with examples of inputs and associated desired outputs that the network computes and then the error is calculated. The idea of the backpropagation algorithm is to reduce this error, until the ANN learns and understands the training data. The training commences with random weights, with the goal of adjustment of the weights so that the error will be minimal. The overall goal of the algorithm is to attempt to find weights such that, for every input vector in the training set, the neural network yields an output vector closely matching the prescribed target vector (Rumelhart and McClelland 1986). Although there are multiple variants of backpropagation algorithm including Polak-Riebre, Fletcher-Revees, one-secant, Levenberg Marquardt, quasi-Newton, and scale conjugate gradient (SCG), the scaled conjugate gradient backpropagation (SCG-BP) algorithm was selected in this thesis because it was designed to avoid the more time consuming line search and to utilise this algorithm's well known speed of convergence (Adetiba and Olugbara 2015).

In order to enhance the neural network's capabilities, ensemble learning is utilised. Ensemble learning improves the ANN's performance by giving better accuracy than a single ANN (Adetiba and Ibikunle 2011; Adetiba and Olugbara 2015). In machine learning, the idea of ensemble learning is to engage multiple learners and combine their predictions. Ensemble learning models are known to enhance the performance of single models by giving better accuracy than the individual members of the

ensemble. A highly effective method, among several others, that can be used for constructing ensembles is the manipulation of the training samples to generate multiple models (Parmant, Munro and Doyle 1996). In this method, the learning algorithm is run in several iterations with a different subset of the training samples at each iteration. This method is known to work efficiently with unstable learning algorithms, such as decision tree and neural network. Examples of different algorithms used for manipulating the training data sets are Bagging, Cross-validated committees, and Adaboost (Galar *et al.* 2012). Bagging was developed in 1996 and it means bootstrap aggregation. It is reputed as the first effective method of ensemble learning and one of the simplest (Breiman 1996). Boosting changes, the dispersal of the training set depending on the performance of prior classifiers while bagging does not (Bauer and Kohavi 1999). The method creates multiple versions of a training set by sampling with replacement and each of the resampled data sets is used to train a different model. The output of the model is often combined by averaging or voting, depending on the nature of the problem. Bagging ensemble is adopted for this study to leverage on its benefits such as good performance based on non-correlation of classifiers (Breiman 1996).

Although speech has a complex mixture of several effects, with results in spectral pattern variation, ANN uses its general pattern recognition mechanism that can learn to develop a function to distinguish between categories. In other words, neural networks can be used to resolve spectral pattern variation problems (Huang *et al.* 1998). Once these spectral pattern variation problems are resolved, the outputs must be used to match correctly with the appropriate voter template in the database. This match must be able to refute imposters and accidental recognitions (Rumelhart, Hinton and Williams 1986).

5.5 Conclusion

This chapter has thoroughly discussed the theoretical foundation of voice biometrics as it applies to a voter authentication model. Some of the theoretical concepts discussed here (MFCC, MFDWC, LPCC, HOG, ANN and bagging ensemble) are utilised for the experimental models in the subsequent chapter. As part of the voice biometric model, an innovation in the form of the use of HOG for feature reduction is introduced.

Chapter Six - Experimental Model and Results

This chapter undertakes a voice biometric experiment that will mirror the process a voter will go through. It will in particular conduct and report on four experiments carried out to determine the most appropriate spectral features for implementing the voice biometric authentication aspect of the voting scheme. The spectral features we examined are MFCC, MFDWC, LPCC and SHOG. We utilised the HOG algorithm which is reputed to be a good descriptor of spectral shape and appearance to both capture the discriminative content and dimensionally reduce the image spectrograms of the MFCC, MFDWC and LPCC spectral features (Selvan and Rajesh 2012). The sample of eight speakers was selected to be multi-lingual, from diverse parts of the world. The waveform and spectrograms of each voice was presented to show variations. The neural network ensemble was generally utilised as the pattern matching algorithm in the four experimental models that were set up to investigate the spectral features in this study.

6.1 Data Acquisition

Speech signals of eight different Speakers were recorded at 11025 Hz sampling rate as .wav files using a Logitech microphone. The recording was done in MATLAB R2012a at 16 bits per sample for 10 seconds duration. The utterances made sequentially by the Speakers for each instance of recording are as follows:

- "Hello Hello Hello …"
- "1 2 3 …"
- "A, B, C…"
- "Yes Yes Yes…"
- "No No No…"

The five different utterances were repeated three times by each Speaker to generate fifteen utterances per Speaker for the training data set. One of these five utterances was in languages other than English, such as Zulu (South Africa), Yoruba (Nigeria), Halychyna (Carpathian region of Europe), Ewe (Ghana), and Hindi (India). This is to introduce diversity into the data set and emulate the reality in a typical multi-ethnic

population that may want to implement the voice biometric authentication in the proposed architecture for e-voting. Furthermore, any two of the five phrases were uttered, and recorded for each Speaker to generate two text-independent test samples per Speaker. This culminates in a total size of 136 experimental datasets for both training and testing. The distribution of the experimental datasets is shown in Table 6.1. A Graphical User Interface (GUI) shown in Figure 6.1 was designed and implemented in MATLAB R2012a for recording the utterances. The waveforms of the *"Hello Hello Hello ..."* utterance for each Speaker are shown in Figure 6.2 and Figure 6.3 while the spectrograms of the utterance are shown in Figure 6.4 and Figure 6.5. Both the waveforms and spectrograms (in Figure 6.2, Figure 6.3, Figure 6.4 and Figure 6.5) clearly illustrate the variations in the patterns of the speech signal from one speaker to the other. We created our own speech database for this study rather than using existing databases such as TIMIT, NTIMIT, IISC and YOHO (Wildermoth and Paliwal 2000; Lei, Yang and Wu 2006) because of the need to introduce diverse languages into the data set. This added a unique flavour of investigating the proclaimed text and language independence of spectral features (Kinnunen and Li 2010). A choice of a compact data set was made for our investigation in this study, based on the established fact in the literature that only a small amount of data is necessary for spectral features (Kinnunen and Li 2010).

Table 6.1 Size of the dataset used in the experiments

Number of Speakers	Numbers of speech recorded per speaker		Total
8	Training	15	120
	Testing	2	16
	Total	**17**	**136**

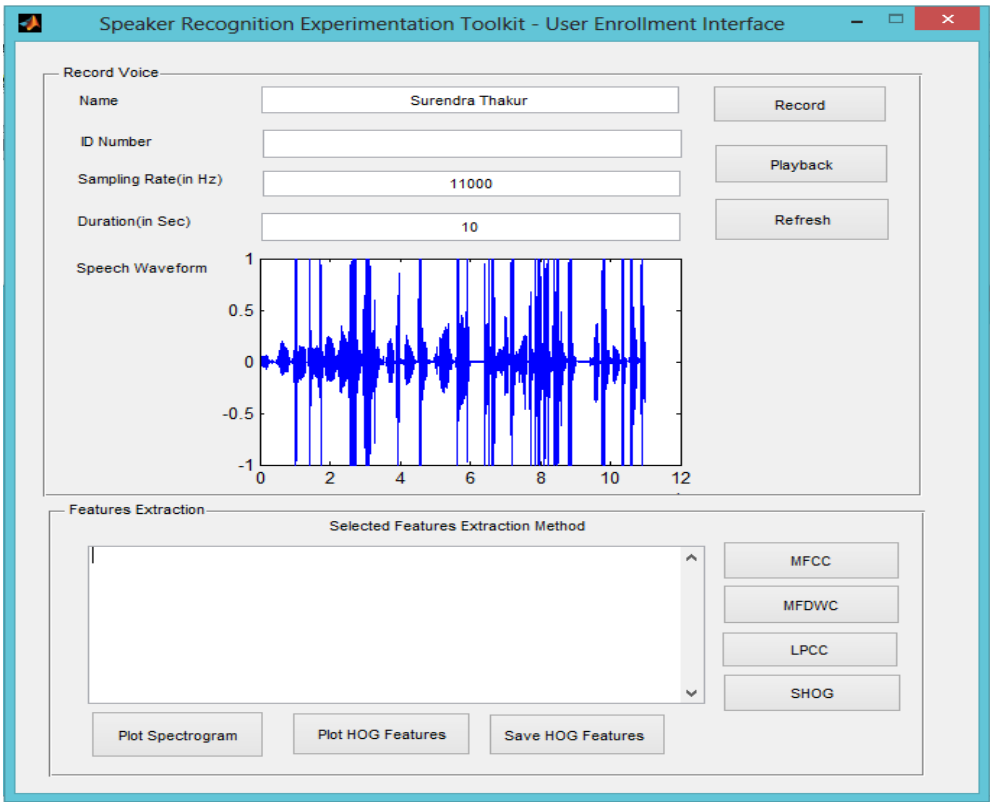

Figure 6.1 Speaker recognition experimentation toolkit (SRET) user enrolment interface, designed by the researcher, using MATLAB R2012a

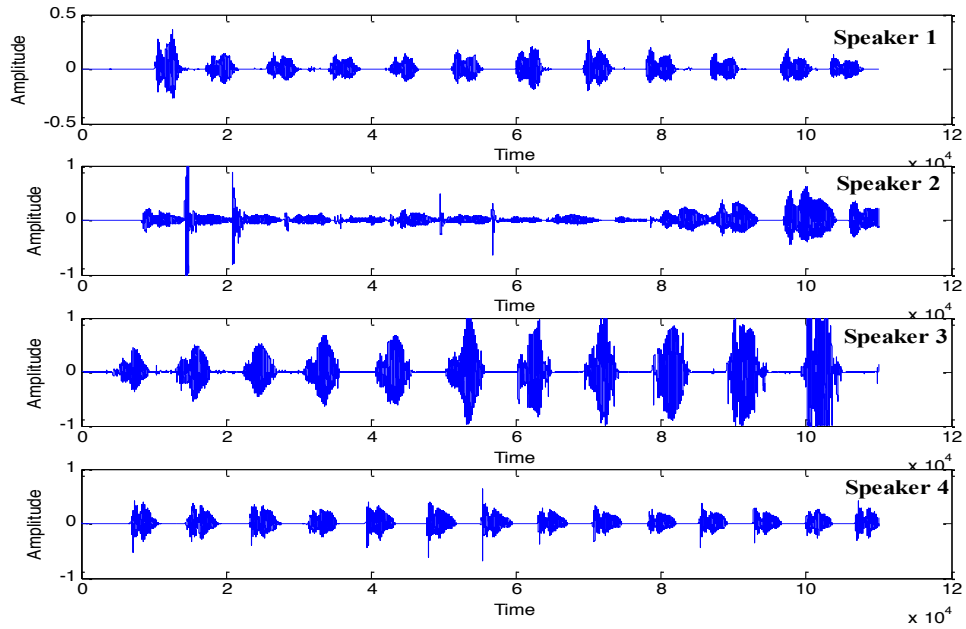

Figure 6.2 Waveforms for the utterance "*Hello Hello Hello …*" by Speakers 1-4[10]

[10] The MATLAB function for plotting the spectrogram accommodates 4 plots per run.

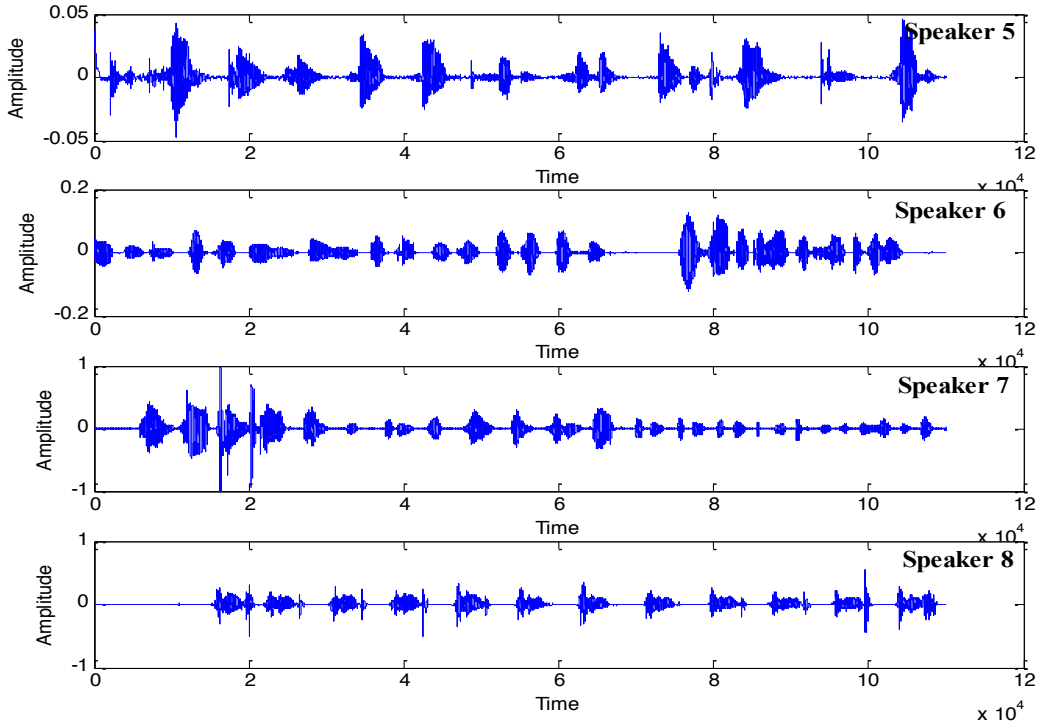

Figure 6.3 Waveforms for the utterance "*Hello Hello Hello …*" by Speakers 5-8

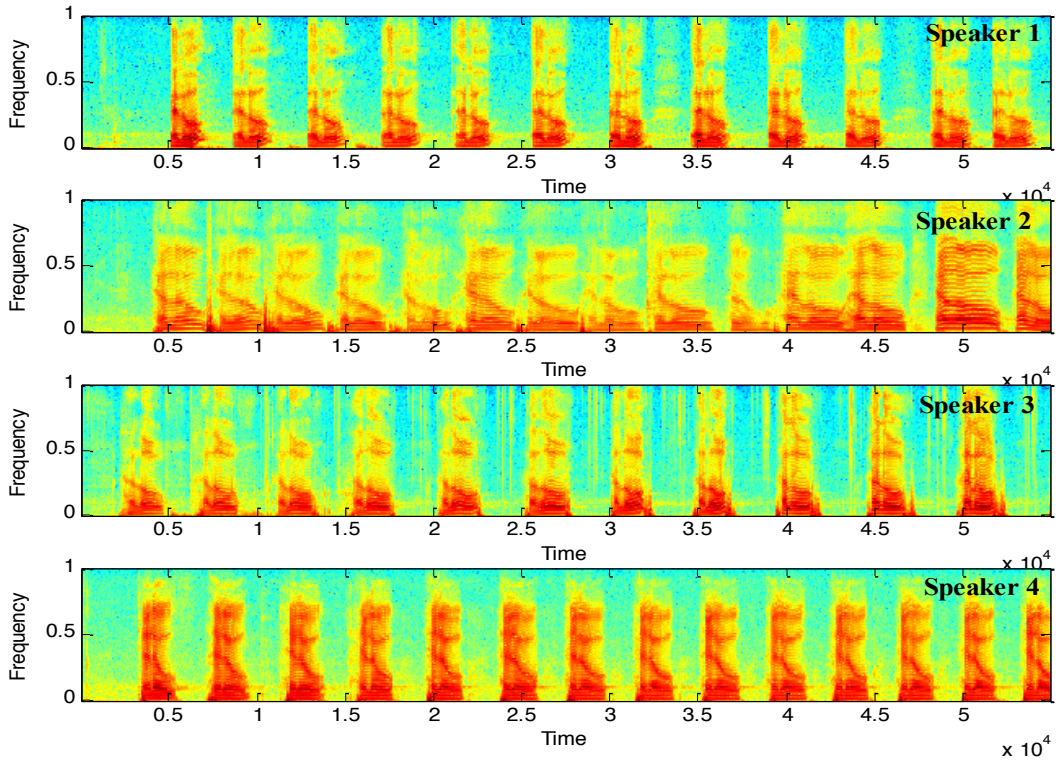

Figure 6.4 Spectrogram for the utterance "*Hello Hello Hello …*" by Speakers 1-4

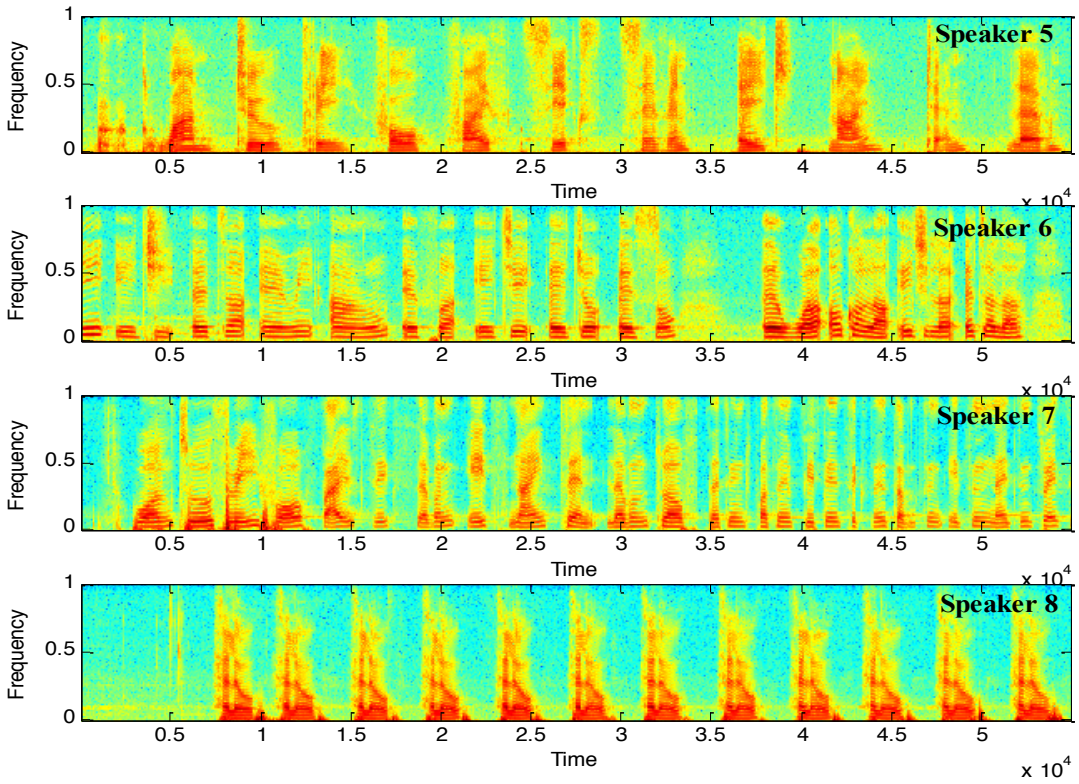

Figure 6.5 Spectrogram for the utterance "*Hello Hello Hello* …" by Speakers 5-8

6.2 Experimental Models

Four different experimental models were designed in this study so as to determine the appropriate combination of algorithms to realise an optimal speaker recognition aspect of the M-Voting system. All the experiments were performed on a computer system with Intel Core i5-3210M CPU operating at 2.50GHz speed. The computer system also has 6.00GB RAM, 500GB Hard Disk and it runs 64bits Windows 8 operating system. As emphasised in Chapter 5, the most prominent features in the literature include MFCC, MFDWC and LPCC (Rose 2002; Wolf 1972). These three features are considered in turns in the experimental models designed in this study to determine the appropriate configuration for the speaker recognition sub-module of the M-Voting System. The following sub-sections contain detailed description of the different models vis-à-vis the report of the results of various experiments that were performed in this study.

6.2.1 Experiment 1

The architecture of the model for experiment 1 is as shown in Figure 6.6. As illustrated in the figure, the first block involves the capturing and digitization of the analogue speech signal using a microphone and the Personal Computer (PC) in the MATLAB R2012a environment. Sample waveforms and spectrograms that were generated from this first block are as shown in Figures 6.2 to 6.5.

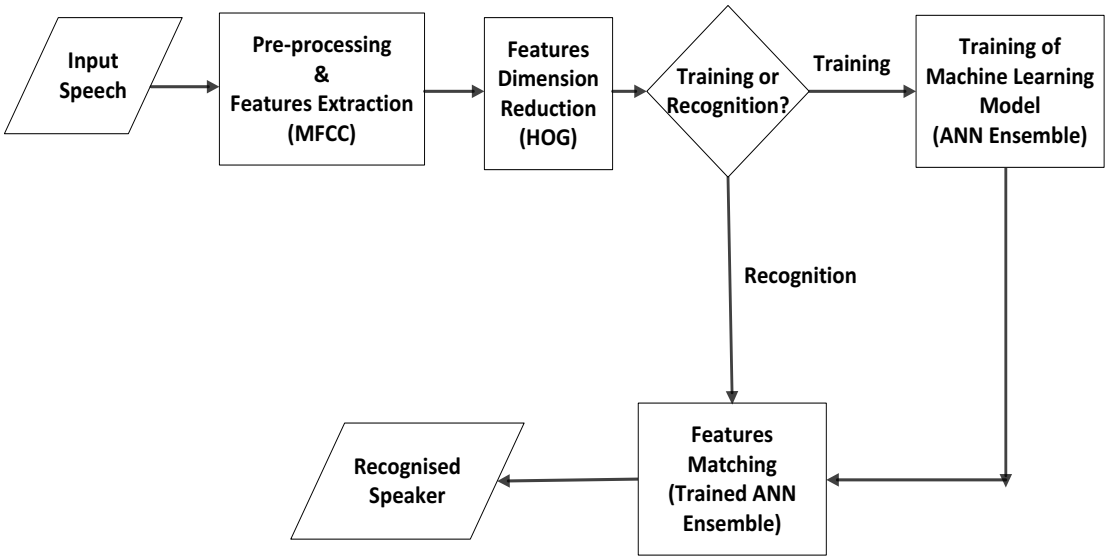

Figure 6.6 Architecture of the model for experiment 1

The pre-processing and features extraction block shown in Figure 6.6 was implemented with the MFCC algorithm. As explained in Chapter 5, Mel-frequency is the measure of the human perception of the frequency content of speech signals on the "Mel-scale". The computational components of the MFCC algorithm are captured in Figure 5.5 and the theory of the computational components of the algorithm has been discussed extensively in Section 5.2.

The MFCC computational components shown in Figure 5.5 were implemented in this study using MATLAB R2012a. The digitised speech signal for each of the seventeen utterances from the different Speakers served as inputs into the MFCC code and each utterance generated a 12x1374 MFCC matrix as outputs.

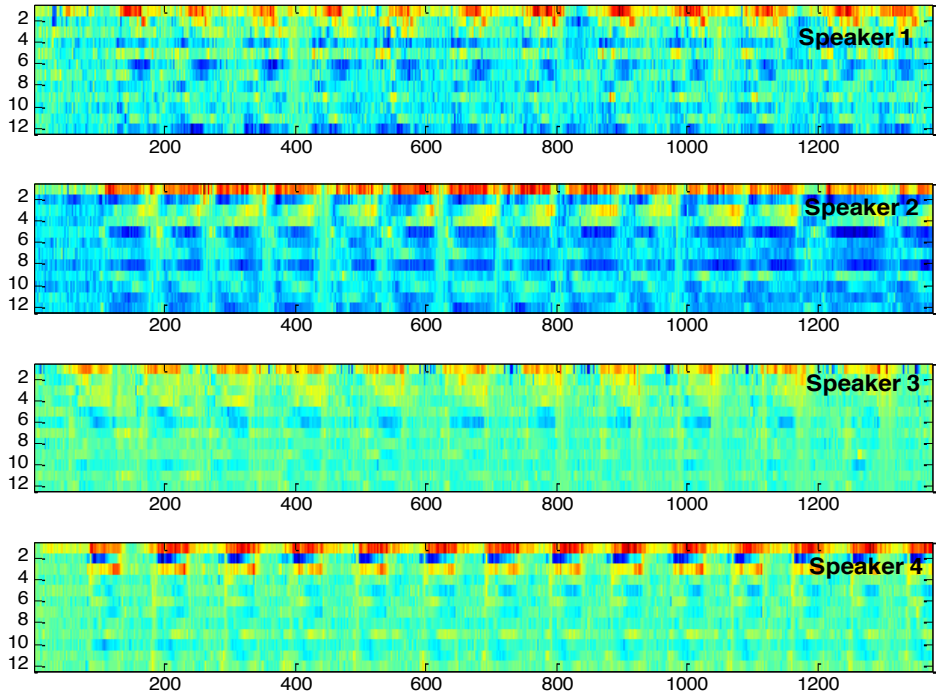

Figure 6.7 The MFCC images for the utterance "*Hello Hello Hello* …" for Speakers 1 – 4

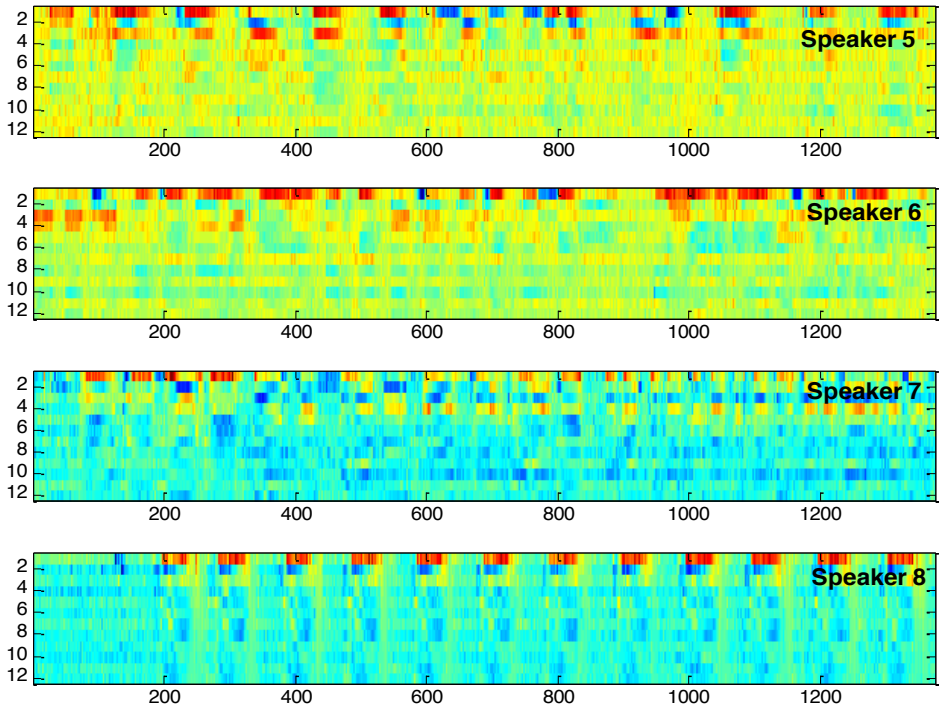

Figure 6.8 The MFCC images for the utterance "*Hello Hello Hello* …" for Speakers 5 – 8

In this study, the HOG block algorithm in the block diagram, was implemented in MATLAB R2012a to reduce the 12x1374 MFCC matrix for each utterance to a feature vector of 81 elements (Selvan and Rajesh 2012). This is an important procedure for reducing the complexity and computational time of the subsequent ensemble-learning network in the model. The time and frequency domain plots of the HOG features for the utterance "*Hello Hello Hello* …" for the 8 Speakers in this study are shown in Figure 6.9 (time domain plot) and Figure 6.10 (frequency domain plot). Both figures illustrate that the dimensionally reduced HOG features for all the Speakers have similar patterns because they represent the utterance of the same set of words; however, despite the similarity in the patterns, the pattern for each of the Speakers is unique both in the time and the frequency domains. This is an illustration of the capability of the HOG algorithm to both reduce the dimension of the extracted MFCC features and still retain the discriminatory features for each of the Speakers in the dataset.

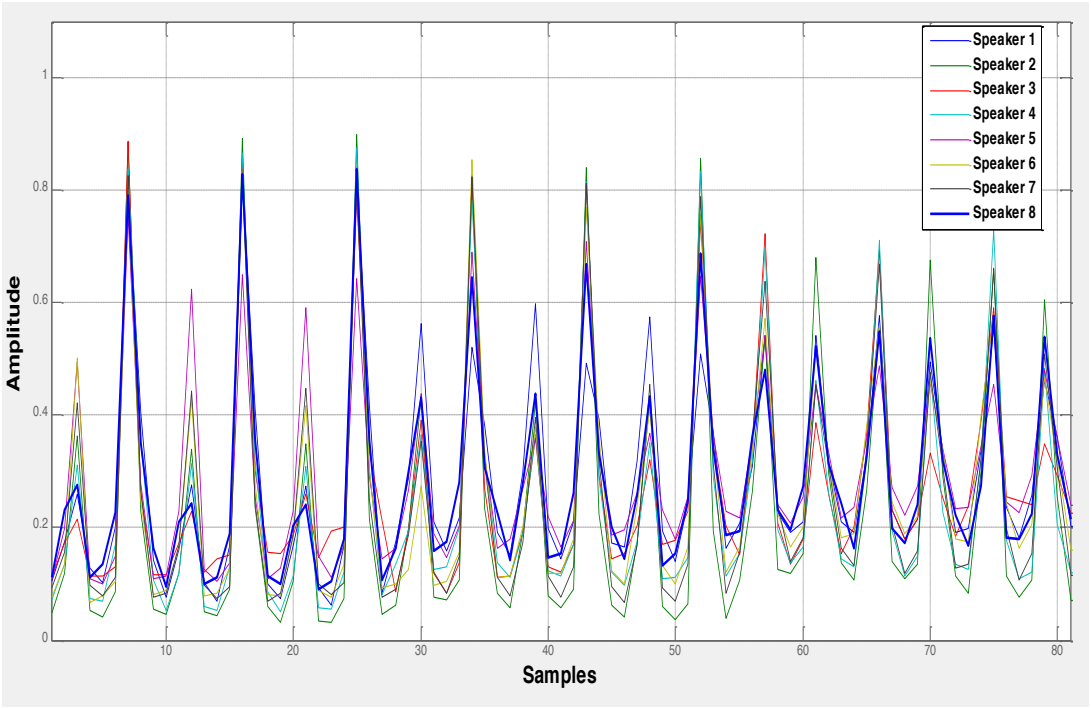

Figure 6.9 Time domain plot of the MFCC HOG features for *"Hello Hello Hello …"* for the 8 Speakers

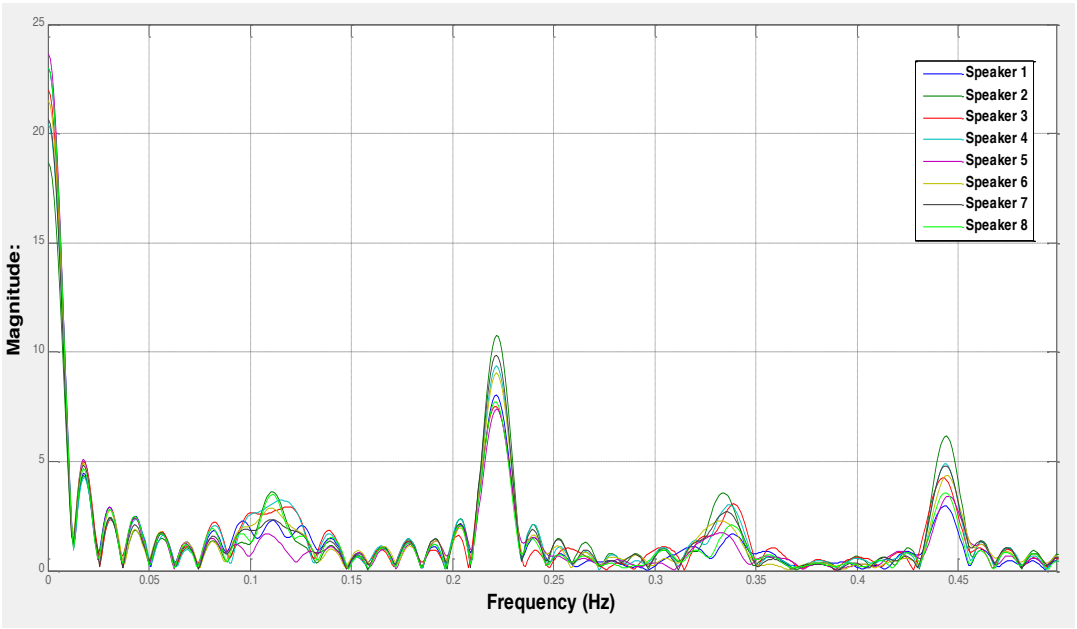

Figure 6.10 Frequency domain plot of the MFCC HOG features for the utterance *"Hello Hello Hello …"* of the 8 Speakers

The next computational block in the model for the current experiment is the design and training of the machine-learning platform, which automates the speaker recognition task in this study. The machine learning method selected for the current experiment is Artificial Neural Network (ANN) ensemble. The theory of Back propagation neural network and ANN ensemble has been discussed extensively in Section 5.4.3 of Chapter 5.

In order to create an ensemble of ANNs, the base ANN has to be properly configured so as to achieve high performance with low network errors. For this study, the configuration of the base ANN are 500 training epochs with the dataset partitioned to 70% training, 15% testing and 15% validation. As mentioned in Chapter 5, the authors in Popescu *et al.* (2009) posit that more hidden layers with a high number of neurons generally lead to small local minimums. On these bases, we selected 2 hidden layers and 80 neurons in each hidden layer for the base ANN. The activation functions selected for each layer of a network are also important in configuring the base ANN for ensemble learning. For the input layer, the linear activation function was selected since this layer is only required to transmit the input data to the subsequent layer without any transformation. The power of MLP-ANN, which is the topology, adopted for the base ANN in this study, comes from non-linear activation in the hidden layer. The most commonly used functions are the logistic and

hyperbolic tangent functions because of their non-linearity and differentiability (Popescu *et al.* 2009). The hyperbolic tangent was however selected for the hidden layer neuron of the base ANN. This is because the function is symmetrical to the origin and decreases the speed of convergence during training. Hyperbolic tangent is also selected for the output layer neuron since the function is adjudged appropriate for binary output patterns (Cybenko 1989; Popescu *et al.* 2009).

There are 81 number of neurons in the input layer for the base ANN in this study in conformity with the number of elements in the HOG feature vector shown in Figure 6.9. We also have **3** neurons in the output layer since there are 8 unique speakers in the dataset and permutations of 3 binary patterns are sufficient for the unique identification of the 8 speakers. The architecture of the fully configured base ANN used in this study is shown in Figure 6.11, while the target outputs binary patterns of the output neurons are shown in Table 6.2.

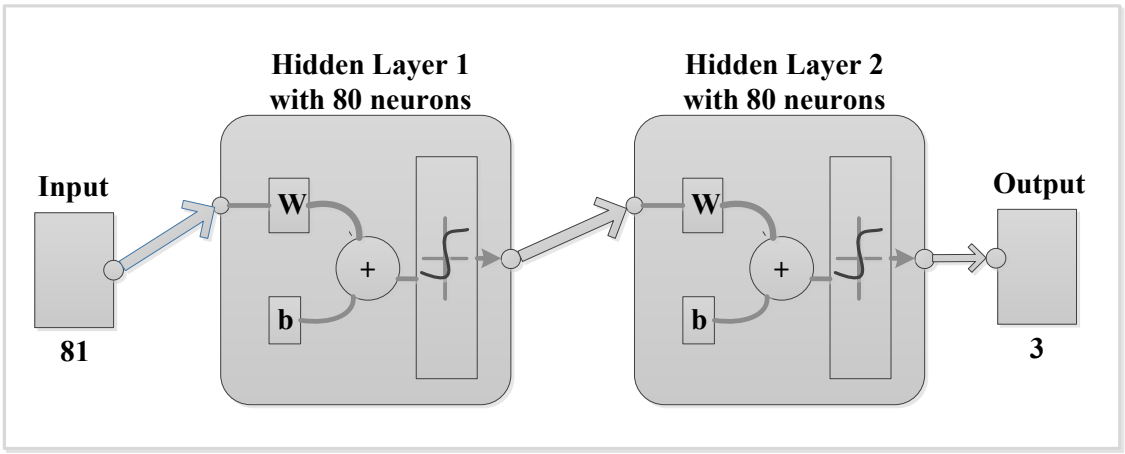

Figure 6.11 Architecture of the configured base ANN

Table 6.2 Target outputs from the ANN for each Speaker

Speaker	Target Output
Speaker 1	0 0 0
Speaker 2	0 0 1
Speaker 3	0 1 0
Speaker 4	0 1 1
Speaker 5	1 0 0
Speaker 6	1 0 1
Speaker 7	1 1 0
Speaker 8	1 1 1

The configuration of the ANN ensemble is another critical aspect in the design of ensemble learning systems. A study by Cherkauer (1996) trained an ensemble of 32 neural networks to identify volcanoes on Venus. In this study, 50, 100 and 200 base models were tested using bagging ensemble and plurality voting for combining their predictions. Our result gave better prediction accuracy with moderate complexity using 100 base models in the ensemble. It is however already established in the literature that having a high number of models is advantageous in problem domains with a small data set, which is the case in this study (Parmanto *et al.* 1996). The bagging ensemble algorithm adopted for this study was implemented in MATLAB R2012a using appropriate functions in the Statistical and Neural Network Toolboxes. The performances of the base ANNs were evaluated based on statistical measurements such as Mean Squared Error (MSE), regression (R value), and coefficient of determination (R^2 value). The coefficient of determination is a measure of the accuracy of prediction of the trained base ANNs and the higher R^2 values indicate better prediction (Kanungo, Sharma and Pain 2014). The R, R^2 and MSEs of the 100 base models in the ANN ensemble for the current experimental model are plotted and shown in Figure 6.12.

From Figure 6.12 (Experiment 1), the average MSE for the ANN ensemble in the first experiment is 0.0430 and the coefficient of determination (R^2 value) is 0.8464. This ANN ensemble was used to predict two test samples from each of the Speakers. The test samples are the acquired speech signals that were not utilised for the training phase of the neural networks. The results that were obtained are shown in Table 6.3.

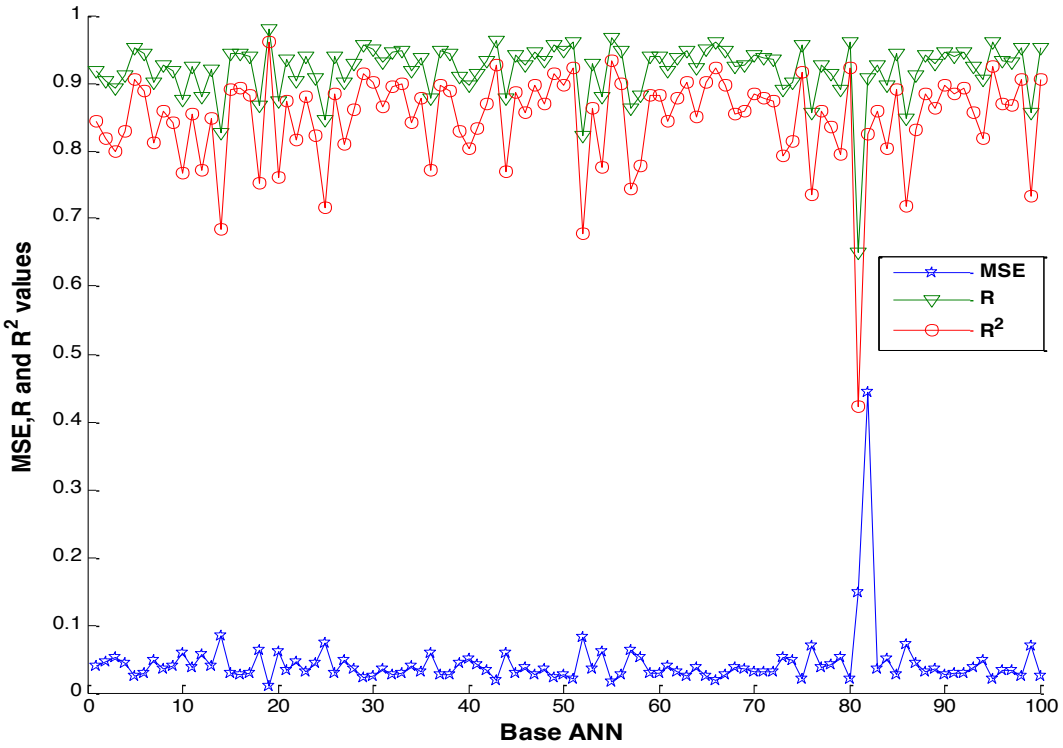

Figure 6.12 MSE and R and R^2 values of the 100 Base ANNs for Experiment 1

As Table 6.3 shows, the model in this Experiment 1 correctly predicted eight test samples out of a total of sixteen test samples. Out of these eight samples, the two test samples for Speaker 1, Speaker 2 and Speaker 6 were correctly predicted. One of the samples was correctly predicted for Speaker 3 and Speaker 8 while none of the samples were correctly predicted for Speaker 4, Speaker 5, and Speaker 7. The test result in the current experiment is not acceptable for the development of a robust Speaker recognition system. This necessitated the setting of another model, which is reported in the next subsection.

Table 6.3 Testing the result of Experiment 1

Speaker ID	Testing Sample	Predicted Output	Target output	Remark
Speaker 1	1	0 0 0	0 0 0	Correct
	2	0 0 0	0 0 0	Correct
Speaker 2	1	0 0 1	0 0 1	Correct
	2	0 0 1	0 0 1	Correct
Speaker 3	1	1 0 0	0 1 0	Incorrect
	2	0 1 0	0 1 0	Correct

Speaker ID	Testing Sample	Predicted Output	Target output	Remark
Speaker 4	1	1 1 0	0 1 1	Incorrect
	2	0 0 1	0 1 1	Incorrect
Speaker 5	1	1 0 1	1 0 0	Incorrect
	2	1 0 1	1 0 0	Incorrect
Speaker 6	1	1 0 1	1 0 1	Correct
	2	1 0 1	1 0 1	Correct
Speaker 7	1	1 0 1	1 1 0	Incorrect
	2	0 0 1	1 1 0	Incorrect
Speaker 8	1	0 1 1	1 1 1	Incorrect
	2	1 1 1	1 1 1	Correct
Total Test Samples = 16		Total Correct Predictions = 8		

6.2.2 Experiment 2

The architecture of the model for Experiment 2 is as shown in Figure 6.13. As illustrated in the figure, the only block that is different from the model for Experiment 1 is the pre-processing and features extraction block. In the current experimental model, MFDWC was selected in place of MFCC as the features. MFDWC is computed in a similar way to the computation of MFCC and the only difference is that a Discrete Wavelet Transform (DWT) is used to replace the DCT block in Figure 6.6. MFDWC has been discussed in Chapter 5.

The MFDWC algorithm was implemented in MATLAB R2012a and the image plots of the output matrices for the utterances "*Hello, Hello, Hello*..." by the 8 Speakers are shown in Figure 6.14 and Figure 6.15. It is clear from these figures that the MFDWC pattern for each speaker is unique.

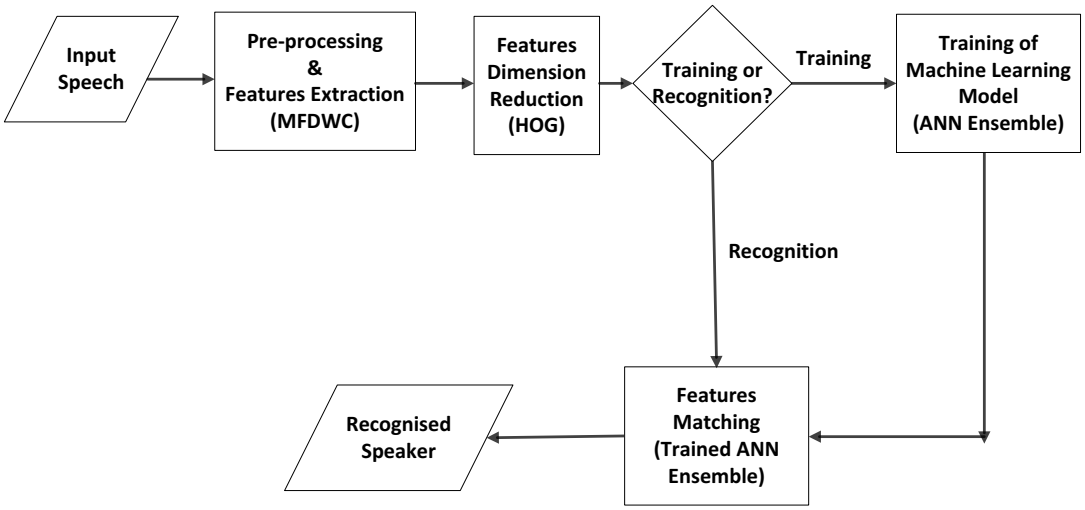

Figure 6.13 Architecture of the model for Experiment 2

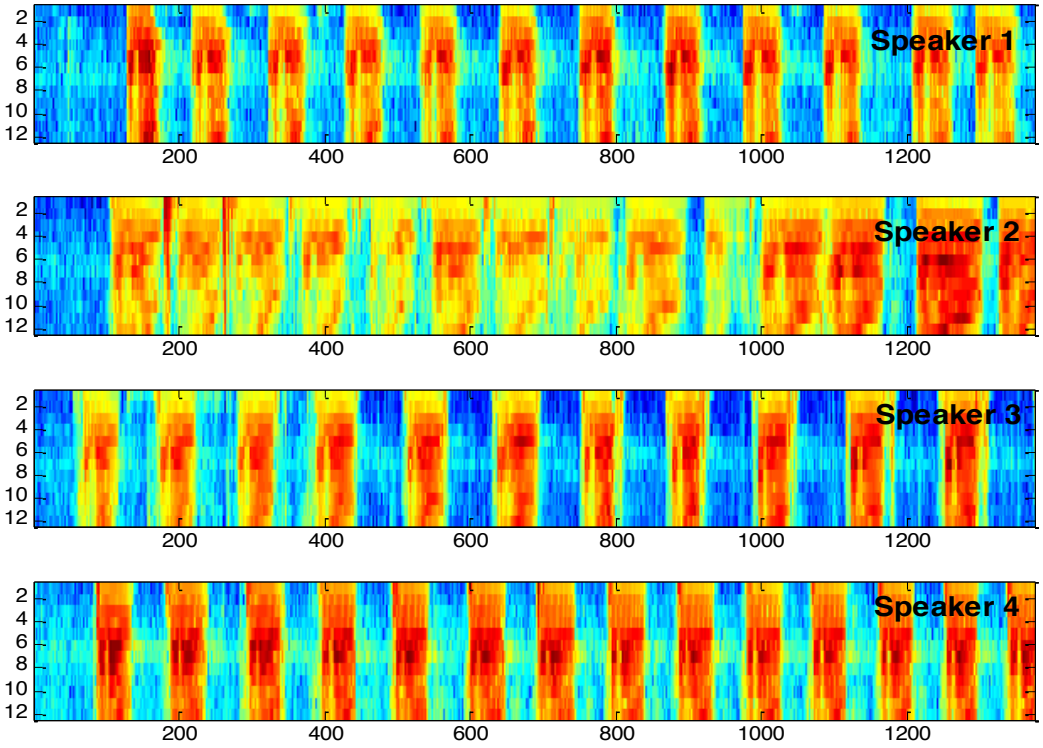

Figure 6.14 MFDWC images for the utterance "*Hello Hello Hello* …" for Speakers 1 – 4

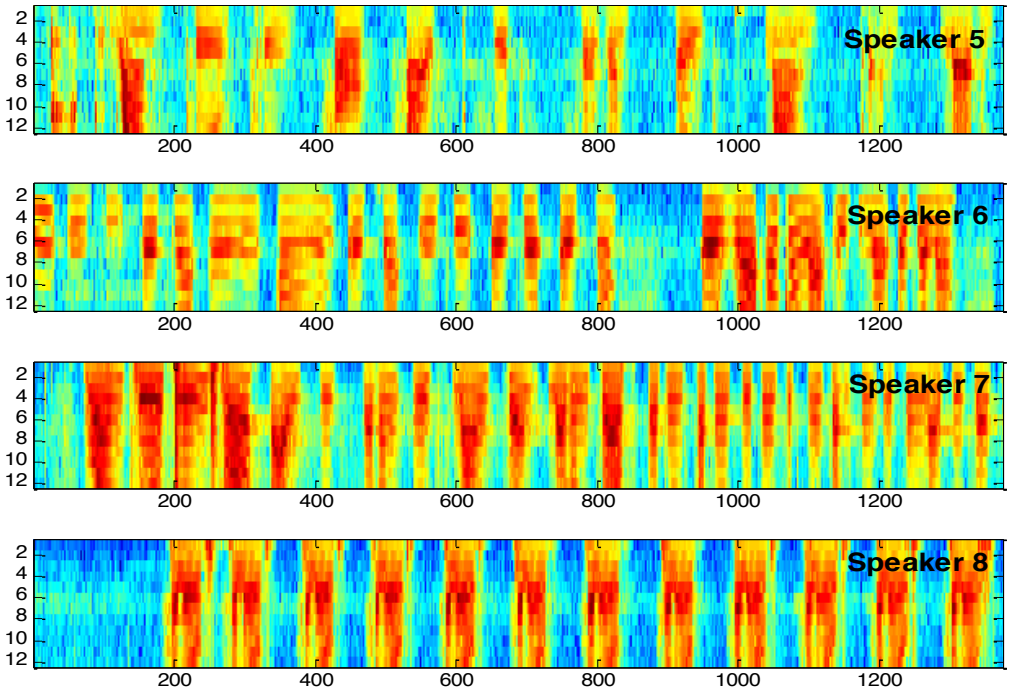

Figure 6.15 MFDWC images for the utterance "*Hello Hello Hello* …" for Speakers 5 – 8

Similar to the procedure in Experiment 1, the HOG algorithm was further utilised to reduce the dimensions of the 12x1374 MFDWC feature matrices in order to obtain 81 element feature vectors for each of the Speakers in this study.

The time and frequency domain plots of the MFDWC-HOG features for the utterance "*Hello Hello Hello* …" for the 8 Speakers are shown in Figure 6.16 and Figure 6.17.

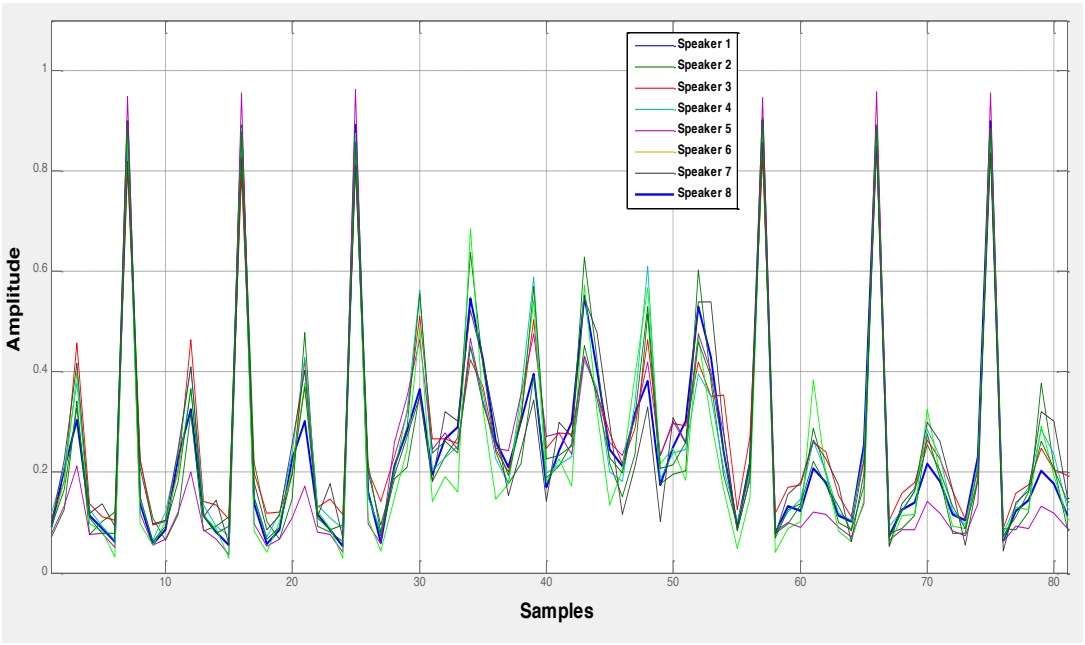

Figure 6.16 Time domain plot of the MFDWC HOG features for the utterance *"Hello Hello Hello …"* of the 8 Speakers

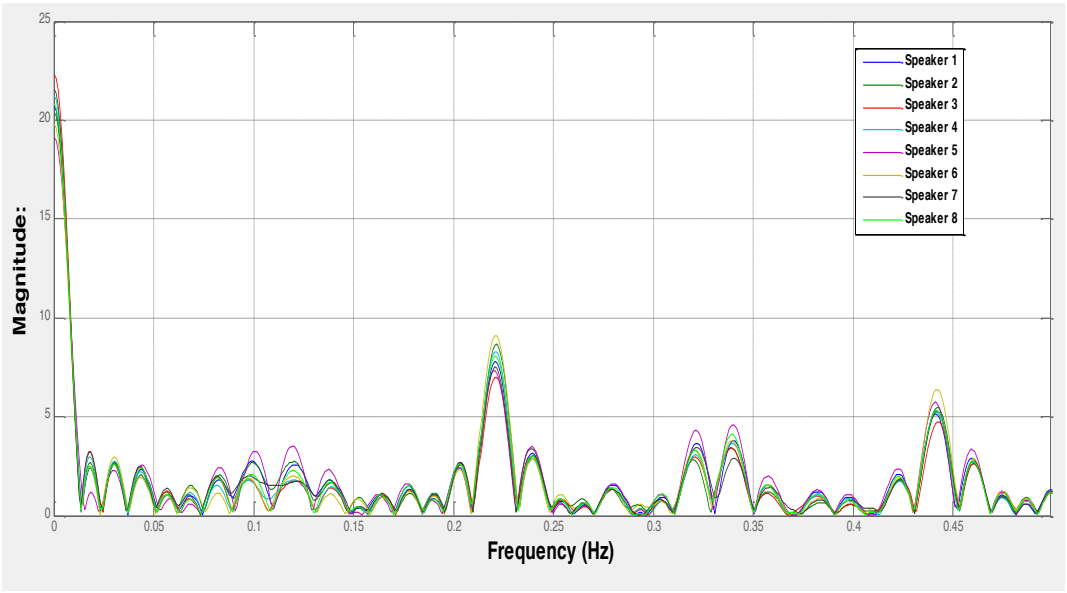

Figure 6.17 Frequency domain plot of the MFDWC-HOG features for utterance *"Hello Hello Hello …"* of the 8 Speakers

As shown in these figures, although the shapes of the MFDWC-HOG features for the different speakers are similar, the sizes of the shapes are unique and this provides a basis for using machine learning approach to uniquely identify each of the speakers. Furthermore, the next block in the model for Experiment 2 (as shown in) is the training of ANN ensemble with the MFDWC-HOG features.

The configuration of the ANN ensemble in Experiment 1 is also used in the current experiment and the result obtained for the 100 base ANNs in this Experiment 2 are shown in Figure 6.18 (Experiment 2(a), 2(b), 2(c) and 2(d)).

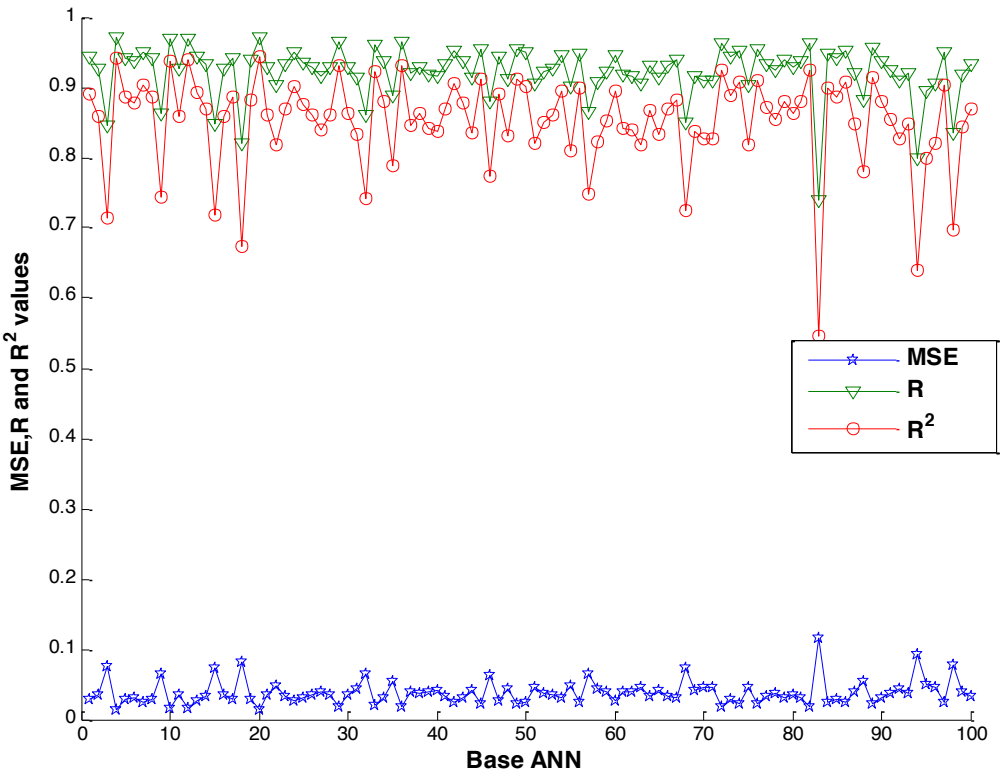

Figure 6.18 MSE, R and R^2 values of the 100 Base ANNs for Experiment 2

The average MSE in Figure 6.18 for the ANN ensemble in the second experiment is 0.0378 and R value is 0.9227 while the coefficient of determination (R^2 value) is 0.8513. The ANN ensemble trained with MFDWC-HOG features in the current experiment gave a slight improvement in the statistical performance parameters over Experiment 1. The ensemble was tested with two utterances, which were not in the training data sets for each of the Speakers, and the results obtained from the test are shown in Table 6.4.

Table 6.4 Testing result of Experiment 2

Speaker ID	Testing Sample	Predicted Output	Target output	Remark
Speaker 1	1	0 0 0	0 0 0	Correct
	2	0 0 0	0 0 0	Correct

Speaker ID	Testing Sample	Predicted Output	Target output	Remark
Speaker 2	1	1 0 1	0 0 1	Incorrect
	2	0 0 1	0 0 1	Correct
Speaker 3	1	0 1 0	0 1 0	Correct
	2	0 1 0	0 1 0	Correct
Speaker 4	1	0 1 0	0 1 1	Incorrect
	2	1 1 0	0 1 1	Incorrect
Speaker 5	1	1 1 1	1 0 0	Incorrect
	2	1 0 1	1 0 0	Incorrect
Speaker 6	1	1 0 1	1 0 1	Correct
	2	1 1 1	1 0 1	Incorrect
Speaker 7	1	1 1 0	1 1 0	Correct
	2	1 1 1	1 1 0	Incorrect
Speaker 8	1	1 1 1	1 1 1	Correct
	2	1 1 1	1 1 1	Correct
Total Test Samples = 16		Total Correct Predictions = 9		

As shown in the Table (6.4), nine samples were correctly predicted out of the 16 samples. The two samples for Speaker 1, Speaker 3, and Speaker 8 were correctly predicted. One sample was also correctly predicted for Speaker 2, Speaker 6, and Speaker 7, while none of the samples for Speaker 4 and Speaker 5 were correctly predicted. It is observed that both Speakers 4 and 5 were also not predicted correctly by the Experiment 1 model. However, as a slight improvement over Experiment 1, the model in the current experiment recognised one of the samples for Speaker 7. This slight improvement is not sufficient to adopt the model in the current experiment and this necessitated the setting up of another model, this is reported in the next subsection.

6.2.3 Experiment 3

The architecture of the model for Experiment 3 is shown in Figure 6.19. The LPCC features extraction algorithm (Rose 2002; Wolf 1972) was selected for the pre-processing and features extraction block in the architecture, and this distinguishes it from the previous architectures in Experiments 1 and 2. A discussion of the theoretical background of LPCC has been done in Chapter 5.

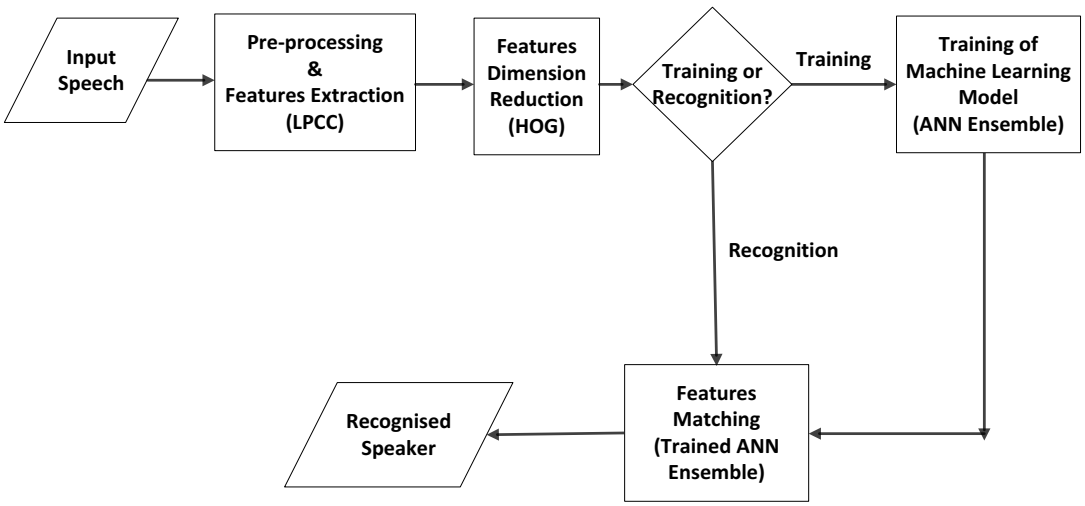

Figure 6.19 Architecture of the model for Experiment 3

The LPCC features extraction algorithm was implemented in MATLAB R2012a and, similar to what was done in Experiments 1 and 2, the image plots of the LPCC feature matrices for the utterances "*Hello, Hello, Hello*..." by the 8 Speakers are shown in Figure 6.20 and Figure 6.21. The patterns of the outputs of LPCC feature matrices shown in these figures are different from the MFCC and MFDWC patterns shown in Figure 6.7 and Figure 6.8 respectively. This is a confirmation of the methodological differences among the different speech features. The patterns of the LPCC feature matrices for each Speaker are also unique and this is a reflection of the discriminatory power of the LPCC features.

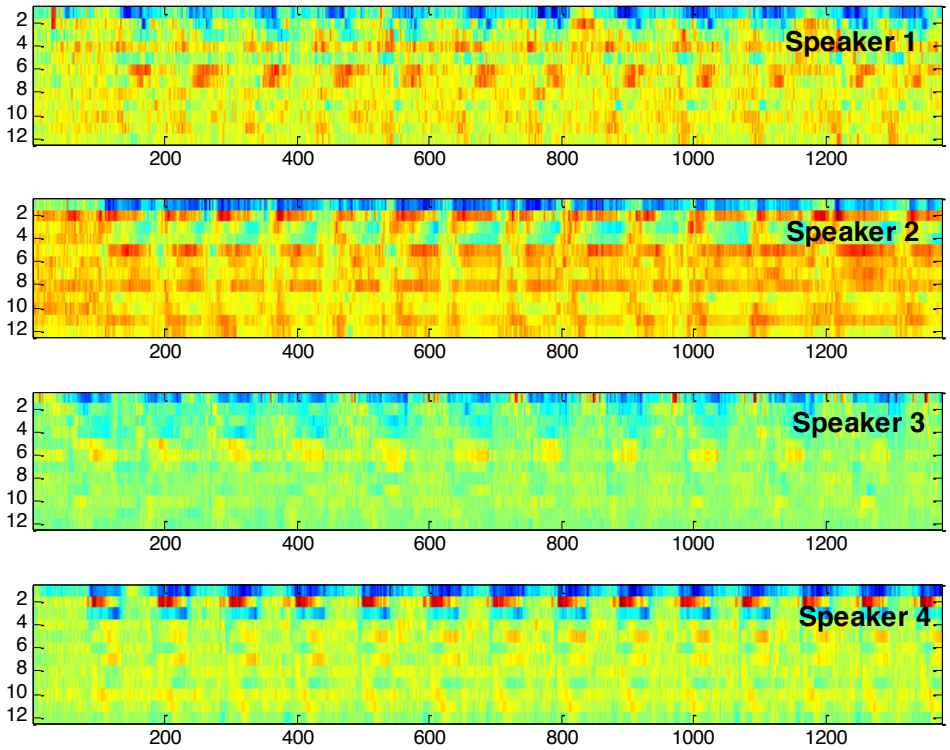

Figure 6.20 The LPCC images for the utterance "*Hello Hello Hello* …" for Speakers 1 – 4

As shown in the architecture for the current experimental model in Figure 6.19, the next procedure is the dimension reduction of the 12x1374 LPCC feature matrices using HOG algorithm, as done in Experiments 1 and 2. The time and frequency domain plots of the 81 elements LPCC-HOG feature vectors obtained for each of the 8 Speakers in this Experiment 3 are shown in Figure 6.22 and Figure 6.23. These features, which are unique for each speaker as shown in both the time and frequency domain plots, are utilised to train the ANN ensemble of the same configuration as was used in Experiment 1 and 2. The values obtained for the MSE, R and R^2 for the 100 base ANNs in the current experiment are shown in Figure 6.24.

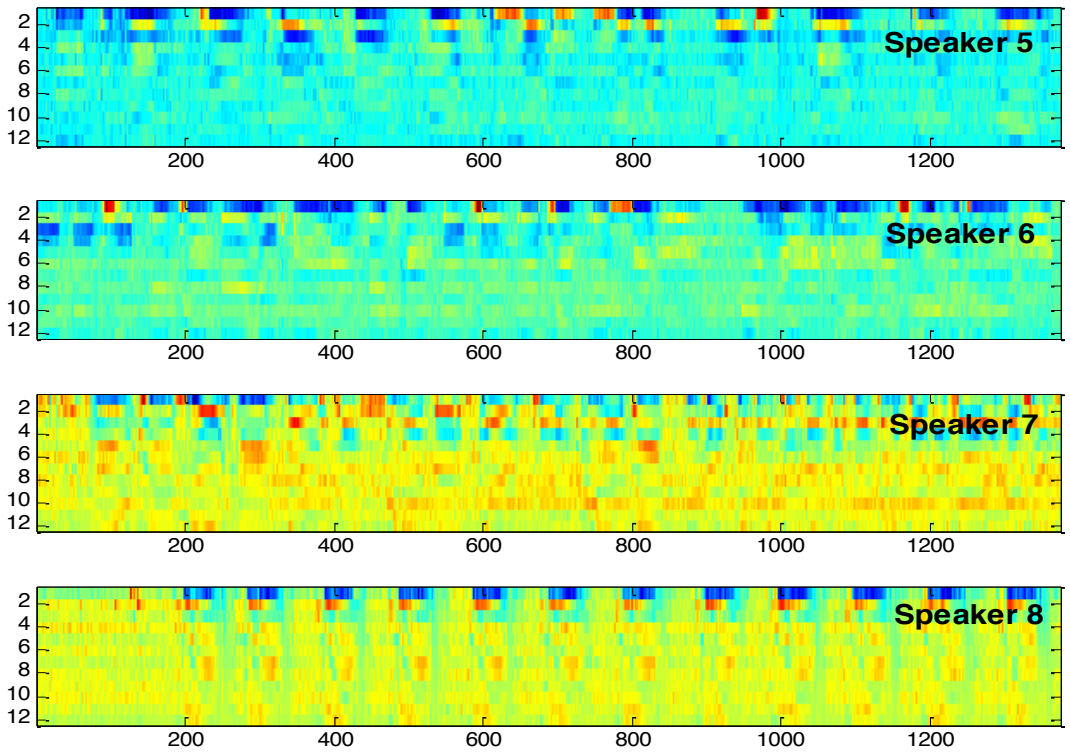

Figure 6.21 The LPCC images for the utterance "*Hello Hello Hello …*" for Speakers 5 – 8

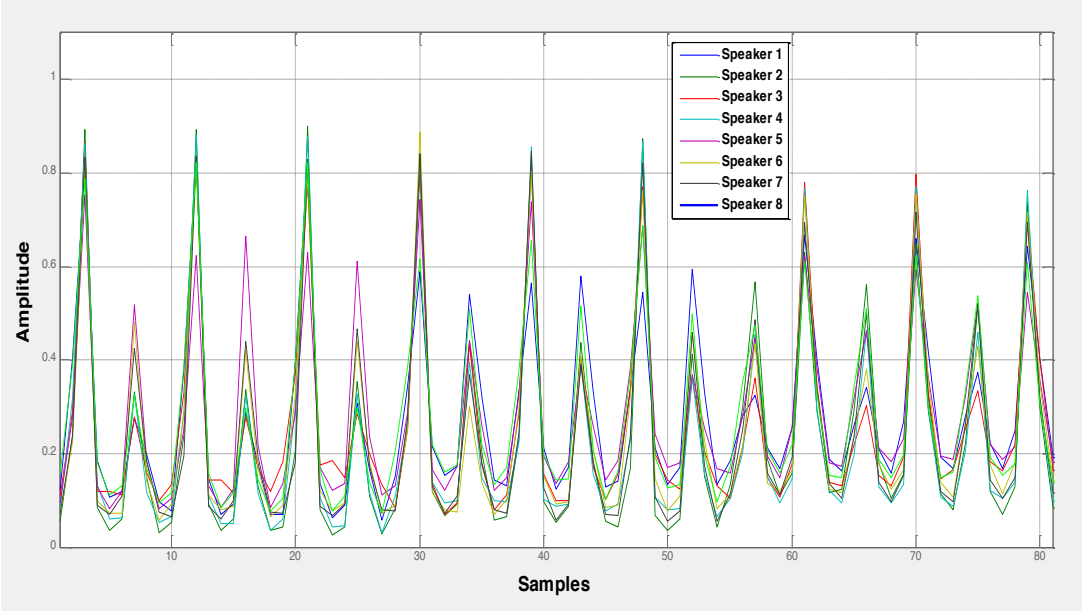

Figure 6.22 Time domain plot of the LPCC-HOG features for the utterance *"Hello Hello Hello …"* of the 8 Speakers

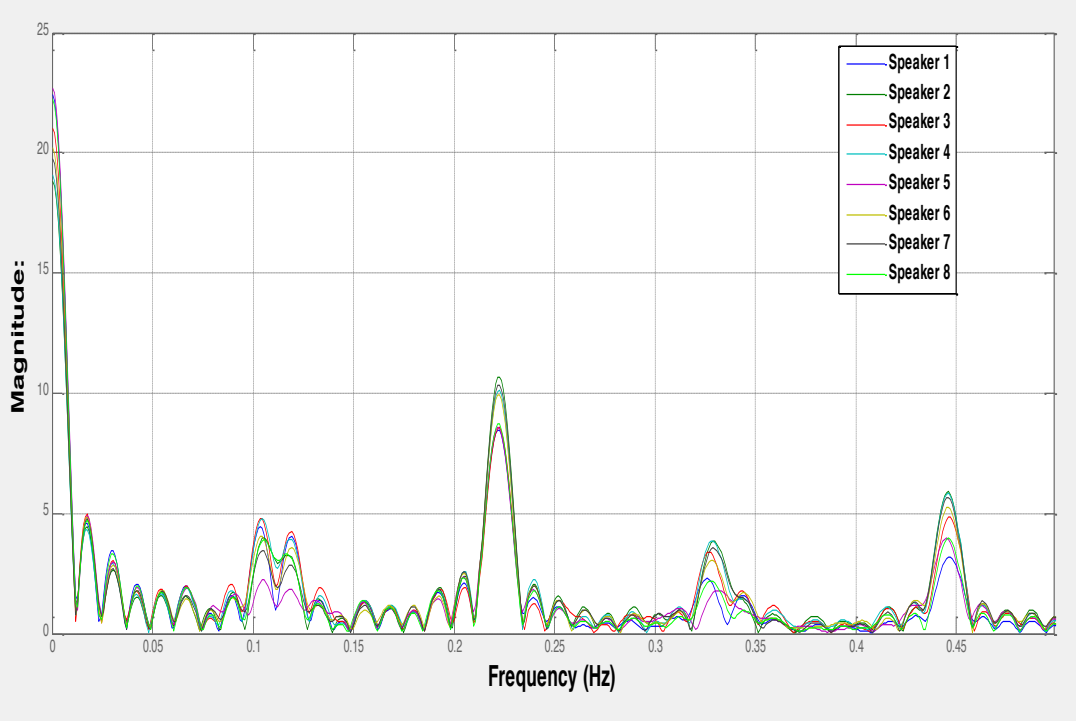

Figure 6.23 Frequency domain plots of the LPCC-HOG features for the utterance *"Hello Hello Hello …"* of the 8 Speakers

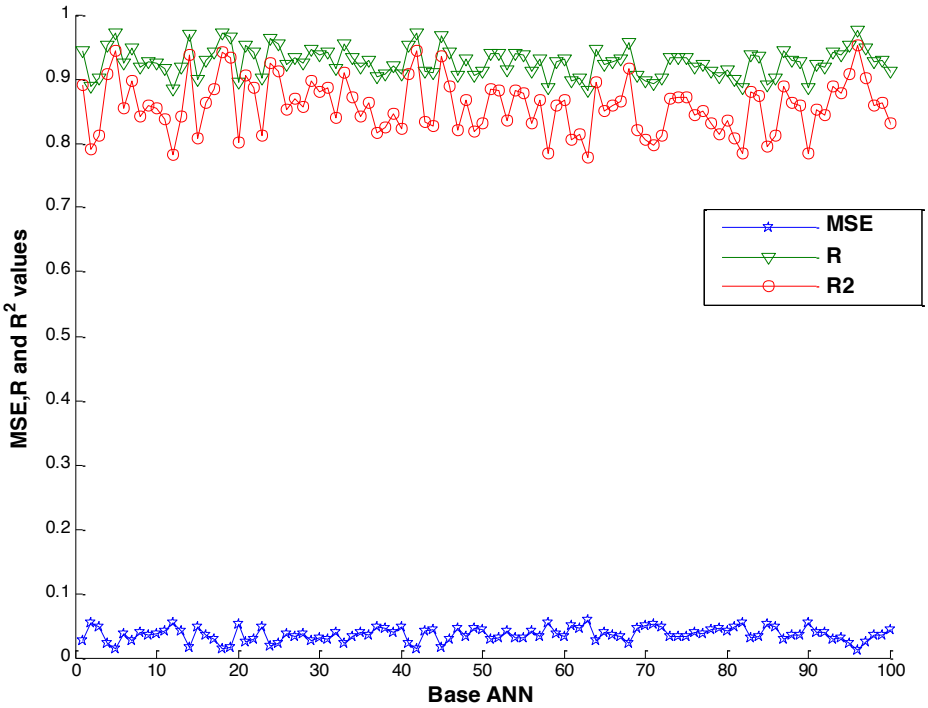

Figure 6.24 MSE, R and R^2 values of the 100 Base ANNs for experiment 3

Figure 6.24 shows an average MSE of 0.0361, R-value of 0.9257 and coefficient of determination (R^2 value) of 0.8575 for the ANN ensemble trained with LPCC-HOG features in Experiment 3. The statistical measures of performance obtained in the current experiment are better than those obtained in the two previous experiments. This is an illustration of a stronger discriminatory capability of LPCC features over both MFCC and MFDWC features. To further validate the current result, the ANN ensemble in this experiment was tested using two test samples for each of the speakers and the outputs of the test are shown in Table 6.5.

Table 6.5 Testing result of Experiment 3

Speaker ID	Testing Sample	Predicted Output	Target output	Remark
Speaker 1	1	0 0 0	0 0 0	Correct
	2	0 0 0	0 0 0	Correct
Speaker 2	1	0 0 1	0 0 1	Correct
	2	1 1 1	0 0 1	Incorrect
Speaker 3	1	0 1 0	0 1 0	Correct
	2	0 1 0	0 1 0	Correct
Speaker 4	1	0 1 1	0 1 1	Correct
	2	1 0 0	0 1 1	Incorrect
Speaker 5	1	1 0 0	1 0 0	Correct
	2	1 0 0	1 0 0	Correct
Speaker 6	1	1 0 1	1 0 1	Correct
	2	1 1 1	1 0 1	Incorrect
Speaker 7	1	1 0 0	1 1 0	Incorrect
	2	1 1 0	1 1 0	Correct
Speaker 8	1	1 1 1	1 1 1	Correct
	2	1 1 1	1 1 1	Correct
Total Test Samples = 16		**Total Correct Predictions = 12**		

As shown in Table 6.5, 12 out of the 16 test samples were correctly predicted by the ANN ensemble, which was trained with the LPCC-HOG features.

The two test samples were correctly predicted for Speaker 1, Speaker 3, Speaker 5, and Speaker 8. In addition, one out of the two test samples was correctly predicted for Speaker 2, Speaker 4, Speaker 6, and Speaker 7. Unlike the results earlier obtained in both Experiments 1 and 2, one or two of the test samples for all the 8 Speakers in the current experiment, were correctly predicted by the ensemble. This is a further validation of the stronger efficacy and discriminatory capability of the LPCC-HOG features over both MFCC-HOG and MFDWC-HOG features. The result obtained with the architecture in this Experiment 3, shown in Figure 6.19 is promising for developing the speaker recognition sub-module of the SMIV model proposed in this study.

We next extend an experiment that successfully classified Tamil-speaking speakers into Male and female to include speaker recognition capability and compare this to Experiments 1 to 3.

6.2.4 Experiment 4

The fourth experimental model reported in this sub-section is based on the Spectral Histogram of Oriented Gradient (SHOG) features that were first reported in the speech processing research community by Selvan and Rajesh (2012) as efficient features for classification of the Tamil language's male/female speakers. Selvan and Rajesh (2012) utilised the HOG algorithm to generate spectral features, rather than for dimension reduction of the short-time speech features (i.e. MFCC, MFDWC and LPCC) as used in the three earlier experiments in this current study. However, the departure sought from the Selvan and Rajesh (2012) study in Experiment 4 of this study is to examine the efficacy of SHOG features for speaker recognition purposes rather than for speech-based classification of persons into male or female gender. The architecture of the model for the current experiment of this study is shown in Figure 6.25.

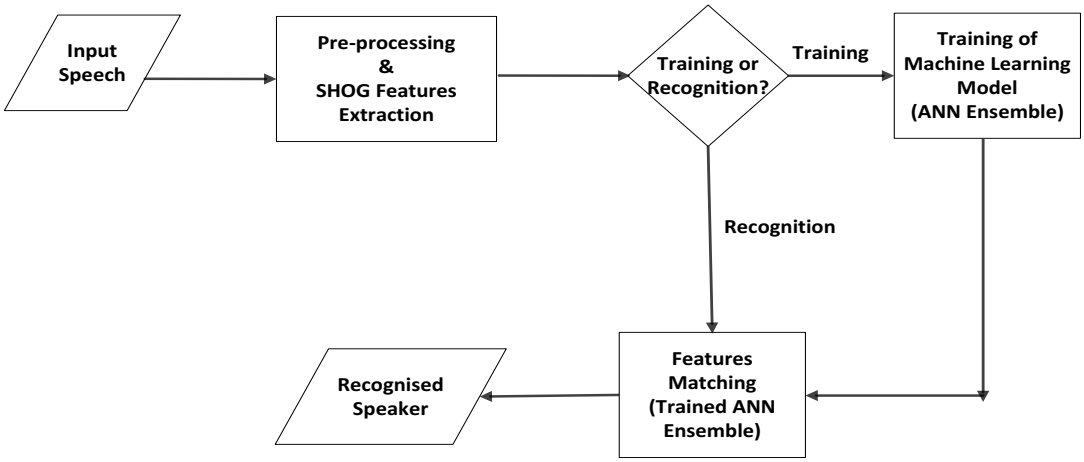

Figure 6.25 Architecture of the model for Experiment 4

As shown in the architecture in Figure 6.25, the features extraction and dimension reduction blocks in the architectures of the previous experimental models have been fused into a SHOG features extraction block. The computational components of the SHOG features extraction block are shown in Figure 6.26.

The spectrograms for the "*Hello Hello Hello ...*" utterances by the 8 speakers in this study have been shown earlier in Figure 6.4 and Figure 6.5. The computational components shown in the SHOG block diagram in Figure 6.26 were implemented in this study using appropriate functions in Image and Signal Processing Toolboxes of MATLAB R2012a.

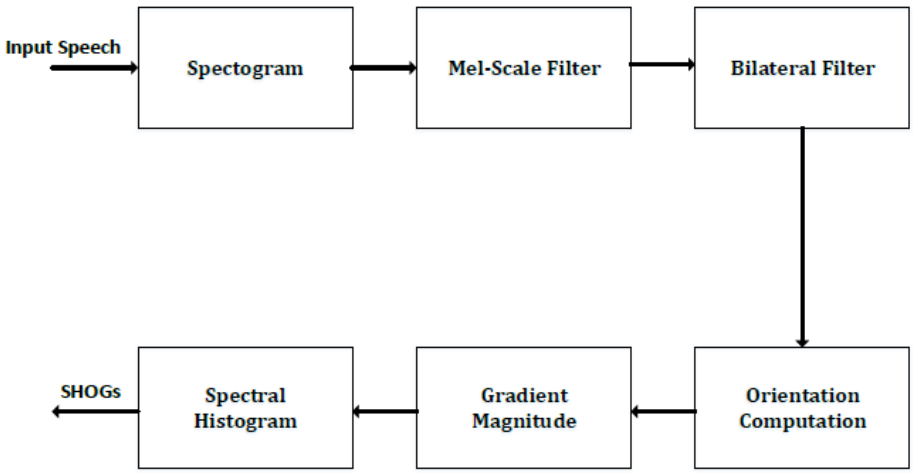

Figure 6.26 Computational components of SHOG (Selvan and Rajesh 2012)

The time and frequency domain plots of the 81 element SHOG features are obtained as outputs from Figure 6.27, Figure 6.28 and Figure 6.29. The SHOG features are unique for each speaker as shown in the time and frequency domain plots. These features are transmitted to the next phase of the current experimental model to train the ANN ensemble with the same configuration as was used in Experiment 1, 2 and 3. The results that were obtained for each of the 100 base ANNs in this fourth experiment are shown in Figure 6.29.

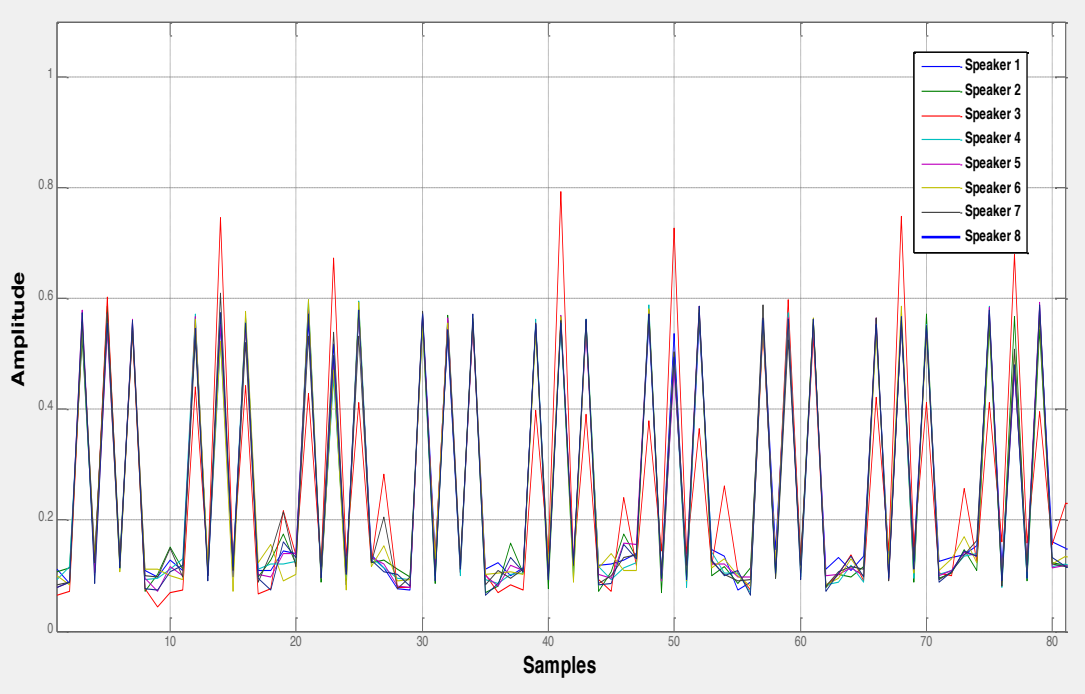

Figure 6.27 Time domain plot of the SHOG features for the utterance *"Hello Hello Hello …"* of the 8 Speakers

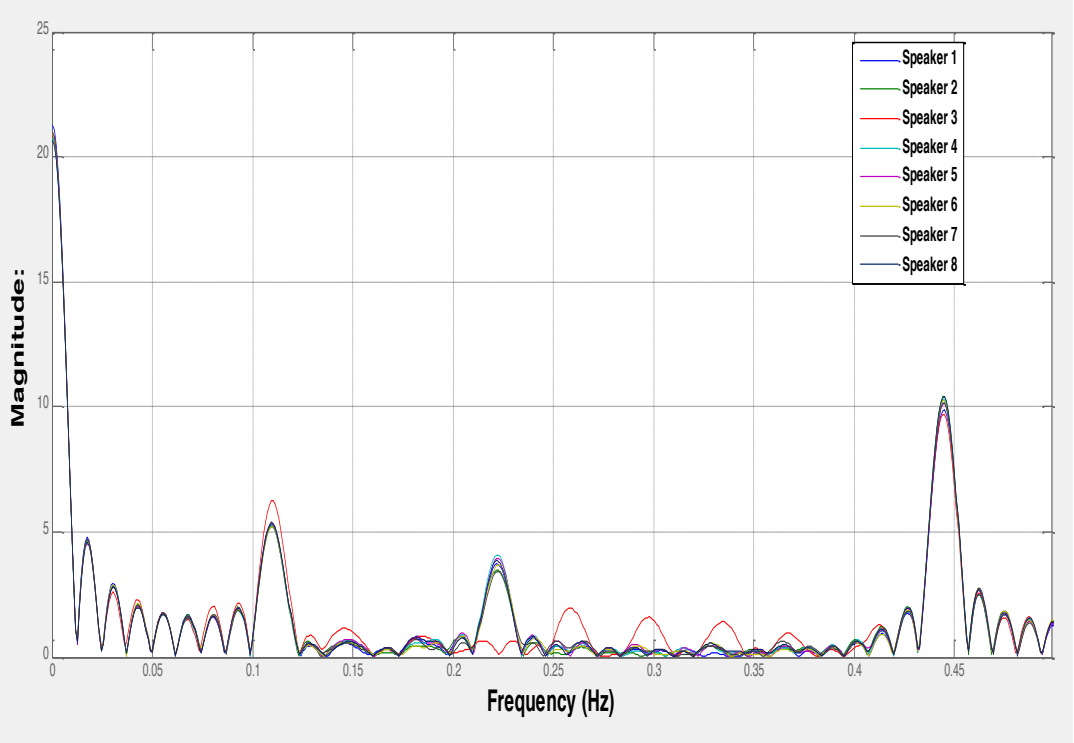

Figure 6.28 Time domain plot of the SHOG features for the utterance *"Hello Hello Hello …"* of the 8 Speakers

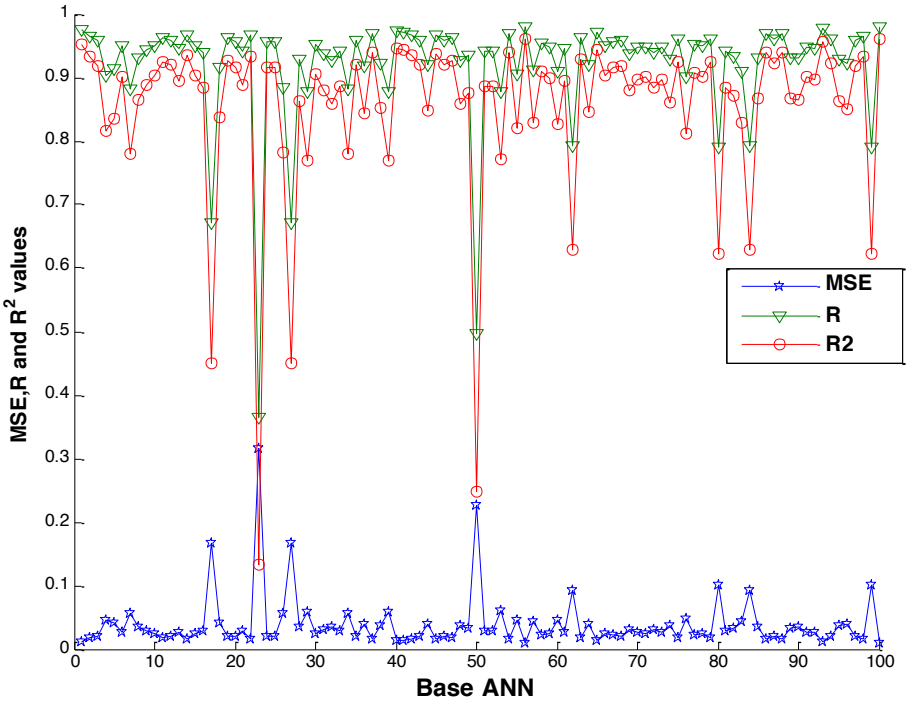

Figure 6.29 MSE, R and R^2 values of the 100 Base for experiment 4

Figure 6.29 shows the result of training the ANN ensemble with SHOG features in the current experiment. As shown in Table 6.6, the average MSE is 0.0386, R-value is 0.9208 and the coefficient of determination (R^2) is 0.8557. These results indicate that using SHOG features gave better performance than MFCC-HOG and MFDWC-HOG features in Experiments 1 and 2 respectively. However, the SHOG features performance is not as good as the result obtained in Experiment 3 where LPCC-HOG features were utilised to train the ANN ensemble. To further test the performance of SHOG features and ANN ensemble for speaker recognition, the test samples utilised in the previous experiments are also used to test the model in the current experiment. The result that was obtained is shown in Table 6.6.

Table 6.6 Testing result of experiment 4

Speaker ID	Testing Sample	Predicted Output	Target output	Remark
Speaker 1	1	1 0 1	0 0 0	Incorrect
	2	0 0 0	0 0 0	Correct
Speaker 2	1	0 0 1	0 0 1	Correct
	2	0 0 1	0 0 1	Correct
Speaker 3	1	0 1 0	0 1 0	Correct
	2	0 1 0	0 1 0	Correct
Speaker 4	1	0 1 0	0 1 1	Incorrect
	2	0 1 1	0 1 1	Correct
Speaker 5	1	0 1 0	1 0 0	Incorrect
	2	0 1 0	1 0 0	Incorrect
Speaker 6	1	1 0 1	1 0 1	Correct
	2	1 0 1	1 0 1	Correct
Speaker 7	1	0 0 0	1 1 0	Incorrect
	2	0 0 0	1 1 0	Incorrect
Speaker 8	1	1 1 1	1 1 1	Correct
	2	1 1 1	1 1 1	Correct
Total Test Samples = 16		**Total Correct Predictions = 10**		

Table 6.6 shows that 10 out of the 16 test samples were correctly predicted in this current experiment. The two test samples for Speaker 2, Speaker 3, Speaker 6, and Speaker 8 were correctly predicted. Furthermore, one out of the two test samples for

Speaker 1 and Speaker 4 were correctly predicted while none of the samples for Speaker 5 and Speaker 7 was correctly predicted. In agreement with the statistical evaluation results shown in Figure 6.7, the number of samples correctly predicted using the SHOG features are more than the ones for MFCC-HOG and MFDWC-HOG features. However, the number of correct predictions with the LPCC-HOG features is more than the ones for SHOG features as illustrated in Table 6.5 and Table 6.6. The experimental results are summarised in Table 6.7.

As shown in Table 6.7, the LPCC-HOG features with ANN ensemble machine learning model gave the best performance out of the four different models that were investigated in this study. On this basis, the LPCC-HOG features and ANN ensemble are nominated for the voters' authentication module of the SMIV system architecture. The result we obtained in this study is in concordance with the position of Kinnunen and Li (2010) who recommended LPCC as one of the best spectral features for practical applications. However, an important contribution of this work to the speech processing literature is the use of the HOG algorithm for dimension reduction of spectral features. This contribution is significant because it serves as a consolidation of the earlier efforts by Selvan and Rajesh (2012) who developed the SHOG for classification of Speakers into different genders.

Table 6.7 Summary of the experimental results

Experimental Model	Extracted Features	Average MSE	Average R	Average R^2	Number of Correct Predictions (Total Samples = 16)
1	MFCC-HOG	0.0430	0.9190	0.8464	8
2	MFDWC-HOG	0.0378	0.9219	0.8513	9
3	LPCC-HOG	0.0361	0.9257	0.8575	12
4	SHOG	0.0386	0.9208	0.8557	10

6.3 Conclusion

A very important achievement of this current study is the discovery of LPCC-HOG as viable and compact spectral features for implementing the authentication module of the SMIV architecture. These features are also very promising for other applications that require a voice biometric based users' authentication module. However, future work will further improve on the current result by: adding other speech signals in more languages, recording the speech signals over mobile phone lines, and experimenting with other features like the Line Spectral Frequencies (LSF) and Perceptual Linear Prediction (PLP). There should also be experiments with other pattern matching models like Hidden Markov Model (HMM), Support Vector Machine (SVM) and Deep Neural Network (DNN). This may help to further enhance the robustness of the authentication module of the proposed SMIV architecture.

Chapter Seven - Results, Conclusions and Future Work

Society must ensure that democracy does not become a census of those who vote. This thesis consequently argues that the SMIV model, through the use of mobile technology, is actually a *digital provide*[11] (Heeks 2010) as opposed to a digital divide. The SMIV model allows multiple users per device, allowing the broader, sometimes unintentionally excluded sections of the community, to also exercise their democratic option in an election process.

This chapter revisits the proposed SMIV model vis-à-vis the research aims and objectives.

7.1 Summary

A review of voting literature was undertaken to discover the best voting practices and e-voting architectures. A further contextual literature review examined authentication and global positioning systems congruent to mobiles. As authentication of the remote voter is both critical and non-trivial, this thesis focused on secure authentication in architectures. The SMIV secure architecture for remote voting was evolved, based on these literature reviews and the reference architectures of FOO and Sensus, deriving an architecture that uses NFC tag (one per voter), GPS system, and voice biometric. The NFC stores baseline information about the voter mobile, while also auto-coupling the voting application. The GPS mimics the traditional precinct ward by geofencing a voter to their preselected voting area. In particular, SMIV uses this multifactor method to moderate the following threats: *impersonation*, by using a strong multimodal authentication model; and *incoercibility,* by GPS coordinate matching and the revote option. Since voice is a novel electoral authentication method, an in-depth study of the voice process and common recognition algorithms was undertaken. This evaluation then selected four different spectral features, namely MFCC, MFDWC, LPCC and SHOG, which were used along with HOG for feature reduction, and ANN ensemble for its pattern matching algorithm for improved accuracy.

[11] As ICTs have spread into poor communities, a few shards of evidence .. emerged; (of) a 'digital provide' that sees those who do not own and those who cannot access ICTs also benefiting (Heeks, 2008:9).

A small albeit linguistically diverse group of eight speakers was used for evaluation. SMIV then tried to verify each speaker using each algorithm. The result of this comparison indicates that LPCC-HOG yielded the best statistical result of an R statistic of 0.9257 and a mean square error of 0.0361. This result is highly promising for authentication in the SMIV architecture. The SMIV architecture was evaluated against the Mursi *et al.* (2013) framework of voting security requirements, which provides an objective analysis of other e-voting architectures, and highlighted the SMIV architecture's contributions to e-voting.

The SMIV architecture therefore suggests that a multimodal remote authentication scheme that incorporates NFC and GPS, in addition to LPCC-HOG biometrics, will provide the necessary improved statistical confidence to accurately identify a voter.

7.2 Analysis of the SMIV Architecture with respect to research aims and objectives

The study's research question was, "How can emerging technological innovations such as mobile technology, global positioning system service, voice biometric and near field communication, be embedded into the existing reference architecture of an electronic voting system to improve security requirements?"

Since the research objectives realise this research question, these objectives can now be restated and the manner in which they have been addressed can be asserted in turn.

7.2.1 Research objective (a)

This section addresses the study's first **research objective 1.4(a)**, which was: "To identify and enhance a secure mobile Internet voting system reference architecture that could help increase the public trust by leveraging the functional capabilities of mobile devices to improve the system security."

In response, this thesis evaluated other e-voting architectures, notably FOO, Sensus, and REVS in terms of their fulfilment of e-voting security requirements. In Sensus, after voters are determined to be eligible, they are issued with a secret token and a blinded/encryption certificate to authenticate them and their vote (Crano and Cytron 1997). In the REVS architecture, voters are authenticated using username and password. SMIV's authentication is handled differently from both the Sensus and

REVS reference architectures; SMIV authenticates voters by leveraging and using a combination of mobile capabilities, namely, GPS, voice biometric and NFC. These capabilities are discussed further in 7.2.2.

7.2.2 Research objective (b)

This section addresses **research objective 1.4(b)**, which was: "To effectively embed the emerging technological innovations of mobile technology, global positioning system service, voice biometric authentication, and near field communication technology into the identified reference architecture of a mobile Internet voting system for security enhancement."

The introduction of common-off-the-shelf technologies, such as mobile devices, appeals to a broader audience of voters, reduces the need for dedicated polling stations, and eliminates the access barriers posed by a polling station.

The thesis argued that the use of GPS mediates incoercibility by ensuring the traditional precinct ward is mimicked through geofencing a voter to their preselected voting area. This further enhances and retains familiarity while mitigating computer generated attacks. Consequently, introducing GPS addresses the challenge that remote voting - in any context - permits subtle, wilful, or malicious coercion, since, in such an event, an afflicted voter is able to move to a safer location to vote.

The use of the NFC tag, which is attached to the voter ID, enables error-free transmission of voter information from the tag to the mobile. The auto-loading of the voting application enhances the ease-of-use security requirement by loading the legitimate voting application, in the process mitigating the risk of inadvertent or malicious loading of incorrect voting applications.

Voice capability is an inherent feature of mobiles, which is leveraged opportunistically in the SMIV architecture. The introduction of voice biometrics as part of the voter authentication process allows a culturally sensitive yet technologically feasible platform to confirm voter identity.

7.2.3 Research objective (c)

The section addresses the study's third **research objective 1.4(c)** which was: "To validate, by experimentation, the effectiveness of the voice biometric authentication components of an embedded reference architecture of a mobile Internet voting system."

It was argued that the development of an electoral system, such as SMIV, is an exercise that is massively human, time, and cost intensive. Notwithstanding this fact, the postulation of models, experimentations and evaluation of parts of the model carried out in this thesis contribute to the body of the knowledge. In this context, this study therefore evaluated and compared several algorithms. The sets of spectral features nominated and compared include Mel-frequency Cepstral Coefficients (MFCC), Mel-frequency Discrete Wavelet Coefficients (MFDWC), Linear Predictive Cepstral Coefficients (LPCC), and Spectral Histogram of Oriented Gradients (SHOG). The MFCC, MFDWC and LPCC are usually high dimensional voice spectral features; which if used directly, oftentimes leads to high computational complexity of the pattern matching algorithms in voice biometrics. To alleviate this complexity while retaining integrity, the higher dimensions of each of the features were reduced to an 81-element feature vector per speaker using Histogram of Oriented Gradients (HOG) algorithm, while neural network ensemble was utilised as the pattern-matching algorithm. Out of the four sets of algorithms investigated, it was found that the LPCC-HOG gave the best statistical results, with an R statistic of 0.9257 and Mean Square Error of 0.0361. Consequently, the LPCC-HOG algorithm will be implemented for voice biometric authentication in the SMIV architecture.

7.3 Future Work

The implementation and field-testing of the SMIV architecture evolved in thesis is highly recommended. The field-testing will explore the implementation model further in regards to the applicable security requirements of Mursi *et al.* (2013), which will provide contextual data. In particular, an efficient voting scheme has to be scalable with respect to storage, computation and communication needs in order to accommodate large numbers of voters in a normative election.

In addition, laboratory experimentation is required for biometric testing against other commercially available mobile devices (Mansfield *et al.* 2001). Further work in refining voice recognition algorithms needs to be undertaken, so as to increase the voice biometric verification accuracy. This work must include a larger voter sample.

The constantly increasing NFC storage capacity will in future enable the offline storage of suitably encrypted biometric data, which may well allow Match-on-a-Card (MOC) testing. As MOC is a local test, it will reduce biometric traffic. MOC is also a scope for further work.

Although GPS geofencing and GPS tracking of mobile devices has been undertaken by some researchers (Moloo 2011), further work in this area is required to satisfy the stringent authentication requirements of an election, and to ensure that such tracking can be performed under a wide variety of environmental conditions.

The additional functionality of the mobile, *inter alia* audio, images and adjustable text size capability of mobile devices, may be utilised to accommodate voters with particular disabilities such as the partially sighted and illiterate. The configurable nature of the mobile supports multiple ballot formats, extending the scope and type of questions poised to voters. There is also scope for future work in this context.

7.4 Conclusion

Through the leveraging of ubiquitous, pervasively available voting modalities such mobile devices, the Internet, GPS, NFC and voice biometrics, this thesis has evolved a new architecture for e-voting. This thesis argues in conclusion that, far from being disruptive, the paradigm shift brought about by the SMIV architecture from inline (paper-based and mechanical), through online, to remote mobile Internet voting, will enhance the opportunities for e-citizen participation because mobile internet voting operates "in the same way as people do everything else in their lives" (Allen 2006:20). That said, this thesis acknowledges that much work remains to further develop and deploy such electoral system.

Reference List

Adida, B. 2008. Helios: web-based open-audit voting. In: *USENIX Security Symposium* (17), 335-348.

Abdelkader, R. and Youssef, M. 2012. Uvote: a ubiquitous e-voting system. In: *Third FTRA International Conference on Mobile, Ubiquitous, and Intelligent Computing (MUSIC), 2012.* IEEE, 72-77.

Adetiba, E. and Ibikunle, F.A. 2011. Ensembling of EGFR mutations' based artificial neural networks for improved diagnosis of non-small cell lung cancer. I*nternational Journal of Computer Applications*, 20(7): 39-47.

Adetiba, E. and Olugbara, O.O. 2015. Lung cancer prediction using neural network ensemble with histogram of oriented gradient genomic features. *The Scientific World Journal,* Vol 2015, Article ID 786013, 1-17.

AFP. 2014. Online voting not reday for worldwide roll-out:study *The Express Tribune.* (online). Available: http://tribune.com.pk/story/772642/online-voting-not-ready-for-worldwide-roll-out-study/ (Accesssed 25 June 2015).

Agnitio. 2014. South Africa social security administration–*Proof of life.* Customer case study (online). Available: http://www.agnitio-corp.com/sites/default/files/SASSA_case_study.pdf (Accessed 18 January 2015).

Ahmad, T., Hu, J. and Han, S. 2009. An efficient mobile voting system security scheme based on elliptic curve cryptography. In: *Third International Conference on Network and System Security, 2009. NSS'09*, IEEE: 474-479.

Ahson, S.A. and Ilyas, M. Ed. 2011. N*ear field communications handbook*. CRC Press.

Akilli, H.S. 2012. Mobile voting as an alternative for the disabled voters. In: *Electronic Voting*. LNI (205), 301-313.

Al-Muhtadi, J.F. 2005. *An Intelligent authentication infrastructure for ubiquitous computing environments*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Al-Saidi, R.A. 2011. *A secure electronic voting scheme based on Evox-MA and REVS eVoting blind signature protocols: Doctoral dissertation*. Middle East University.

Albright, S.D. 1942. T*he American ballot*. Washington D.C: The American Council on Public Affairs.

Allen, R. 2006. *Implementing electronic voting in the UK* (online). UK Government Report. Available: www.communities.gov.uk/corporate (Accessed 23 January 2015).

Alvarez, R.M. and Hall, T.E. 2004. *Point, click and vote - the future of Internet voting*. Washington D.C: Brookings Institution Press.

Alvarez, R.M., Hall, T.E. and Trechsel, A.H. 2009. Internet voting in comparative perspective: the case of Estonia. *PS: Political Science and Politics,* 42(03): 497-505.

Alvarez, R.M. and Hall, T.E. 2010. *Electronic elections: the perils and promises of digital democracy.* Princeton: Princeton University Press.

Anane, R., Freeland, R. and Theodoropoulos, G. 2007. E-voting requirements and implementation. In: *The 9th IEEE International Conference on E-Commerce Technology and the 4th IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, 2007. CEC/EEE 2007.* IEEE: 382-392.

Award, M. and Leiss, E. 2011. Internet voting in the USA: analysis and commentary. In: Prosser, A.Ed. T*ransforming Government: People, Process and Policy*, 5(1): 45-55.

Ayo, C.K. 2009. A framework for voice-enabled m-Voting system: Nigeria a case Study. In: *ECIW2009-8th European Conference on Information Warfare and Security: ECIW2009*. Academic Conferences Limited, 96.

Ayo, C.K., Daramola, J.O. and Azeta, A. A. 2009. Developing a secure integrated e-voting system. In: *Handbook of research on e-services in the public sector: E-Government strategies and advancements*, 278-287.

Bai, J., Xue, P., Zhang, X. and Yang, L. 2012. Anti-noise Speech Recognition System Based on Improved MFCC Features and Wavelet Kernel SVM. *Advances in information Sciences and Service Sciences (AISS)*, (4) (23): 599-607.

Baiardi, F., Falleni, A., Granchi, R., Martinelli, F., Petrocchi, M. and Vaccarelli, A. 2005. SEAS, a secure e-voting protocol: design and implementation. *Computers and Security*, *24*(8): 642-652.

Bao, P., Pierce, J., Whittaker, S. and Zhai, S. 2011. Smart phone use by non-mobile business users. In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, ACM, 445-454.

Bakker, M. 2012. Slides and privileged presentation and discussion to author. *Voting machines in the Netherlands: from general acceptance to general doubt in two years* (online). The Kiesraad Commission, The Hague, 14 April 2012. Available: http://prezi.com/psko-w9mznsa/voting-machines-in-the-netherlands-kenia/?auth_key=5ab560619bb7502af4ba4ca477e8b5410f9e521b

Barrat, J., iEsteve, J.B., Goldsmith, B., and Turner, J. (2012). International experience with e-voting. *International Foundation for Electoral Systems*.

Bauer, E., and Kohavi, R. 1999. An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine learning*, 36(1-2): 105-139.

Bellman, R.E. 1961. *Adaptive control processes: a guided tour*. Princeton: Princeton University Press.

Benaloh, J. 2007. Ballot casting assurance via voter-initiated poll station auditing. In: *Proceedings of the USENIX Workshop on Accurate Electronic Voting Technology*. USENIX Association, 14-14.

Bent, H.V.D., Sante, T.V., Kerssens, D. and Kemmeren, J. 2008. *TOGAF, the open group architecture framework: a management guide. V*an Haren Publishing - 73.

Bittiger, J. 2007. *Voter turnout in Western Europe since 1945: a regional report. S*tockholm: IDEA.

Bornman, E. 2006. National symbols and nation-building in the post-apartheid South Africa. *International Journal of Intercultural Relations*, 30(3): 383-399.

Bouman, C.A. 2009. L*ab 9a-speech processing (part 1)* (image of human voice) Technical report, Connexions, Rice University, Texas.

Breiman, L. 1996. Bagging predictors. *Machine Learning,* 24(2): 123-140.

Brücher, H., and Baumberger, P. 2003. Using mobile technology to support eDemocracy. In: *Proceedings of the 36th Annual Hawaii International Conference on System Sciences, 2003. I*EEE, 1-8.

Budurushi, J., Neumann, S., Volkamer, M. 2012. The scope of e-voting in Switzerland. 2012. Smart cards in Electronic Voting, Lessons learned from applications in legally binding elections and approaches proposed in scientific papers. In: *Electronic Voting*, 257-270.

Burmest, M. and Magkos, E. 2003. Towards secure and practical e-elections in the new era. In: *Secure electronic voting*, Springer US, 63-76.

Campbell, J. 1997. Speaker recognition, a tutorial. In: *Proceedings of the Biometric Consortium. IEEE* , 85(9): 1437-1462.

Campbell, W., Sturim, D. and Reynolds, D. 2006a. Support vector machines using GMM supervectors for speaker verification. I*EEE Signal Processing Letters 13*: 308–311**.**

Campbell, B. A., Tossell, C.C., Byrne, M. D. and Kortum, P. 2014. Toward more usable electronic voting testing the usability of a smartphone voting system. In: *Human Factors, The Journal of the Human Factors and Ergonomics Society*.

Carman, C., Mitchell, J. and Johns, R. 2008. The unfortunate natural experiment in ballot design, the Scottish parliamentary elections of 2007. In: *Electoral Studies*, 27(3): 442-459.

Castro, D. 2007. Stop the presses, how paper trails fail to secure e-voting (online). *Information Technology and Innovation Foundation.* Available: www.itif.org (Accessed 20 June 2015).

Castro, D. 2011. Seven principles for secure e-voting. Letters to the editor. *Communications of the ACM*, 52(2): 8.

Castro, D. 2011b. Explaining international leadership, electronic identification systems (online). *Information Technology and Innovation Foundation,* September. Available: *www.itif.org* (Accessed 20 June 2015).

Cetinkaya, O. and Koc, M.L. 2009. Practical aspects of DynaVote e-voting protocol. E*lectronic Journal of e-Government*, 7(4): 327-338.

Chandramouli, R. and Lee, P. 2007. Infrastructure standards for smart ID card deployment. *Security and Privacy,* 5(2): 92-96. doi:, 10.1109/MSP.2007.34.

Chang, W.W. 2012. Time frequency analysis and wavelet transform tutorial. Time-frequency analysis for voiceprint (speaker) recognition. Available: https://www.scribd.com/doc/198845429/Voiceprint-Speaker (Accessed 2 July 2015).

Chaum, D. 1983. Blind signatures for untraceable payments. *Advances in Cryptology Proceedings of Crypto, 82*(3): 199–203.

Chaum, D. 2001. Surevote, technical overview. In: *Proceedings of the workshop on trustworthy elections* (WOTE'01). California. Available: http://www.iavoss.org/mirror/wote01/pdfs/surevote.pdf (Accessed 2 July 2015).

Che, Z.G., Chiang, T. A. and Che, Z.H. 2011. Feed-forward neural networks training: A comparison between genetic algorithm and back-propagation learning algorithm. *Int. J. Innov. Comp. Inf. Control*, *7*(10): 5839-5851.

Chen, W., Hancke, G.P., Mayes, K.E., Lien, Y. and Chiu, J.H. 2010. NFC mobile transactions and authentication based on GSM network. In:*Second International Workshop on*, *Near Field Communication (NFC). IEEE,* 83-89.

Cherkauer, K.J. 1996. Human expert-level performance on a scientific image analysis task by a system using combined artificial neural networks. In: *Working notes of the AAAI workshop on integrating multiple learned models*, 15-21.

Choo, K.K.R., Boyd, C. and Hitchcock, Y. 2005. E*rrors in computational complexity proofs for protocols*, Berlin Heidelberg: Springer, 624-643.

Chowdhury, M.J. 2013. Comparison of e-voting Schemes, Estonian and Norwegian solution. In: *International Journal of Applied Information Systems,* 6(2): 60-66.

Claps, M. and Carter, P. 2013. Technology spotlight, delivering end-to-end election modernization roadmaps. In: *IDC Government Insights,* September 2013.

Clarkson, M.R., Chong, S., and Myers, A.C. 2008. Civitas, toward a secure voting system. In: *IEEE Symposium on Security and Privacy, 2008. SP 2008.*, IEEE, 354-368.

Cloutier, R., Muller, G., Verma, D., Nilchiani, R., Hole, E. and Bone, M. 2010. The concept of reference architectures. S*ystems Engineering*, 13(1): 14-27.

CNN: The Weaponisation of code. Interview of Alex Ross. The best of Quest (broadcast). 21 December 2014. *CNN.com.*

Coskun, V., Ok, K. and Ozdenizci, B. 2011. *Near field communication (NFC), from theory to practice*. Istanbul: John Wiley and Sons.

Cranor, L.F. 1996. Electronic voting: computerized polls may save money, protect privacy. C*rossroads*, *2*(4): 12-16.

Cranor, L.F. and Cytron, R.K. 1997. Sensus: A security-conscious electronic polling system for the Internet. In: *Proceedings of the Thirtieth Hawaii International Conference on System Sciences, 1997, IEEE,* (3): 561-570.

Crossman, P. 2012. The case for voice biometrics. *The American Banke*r. October 31. Available: http://www.americanbanker.com/issues/177_210/the-case-for-voice-biometrics-1053976-1.html (Accessed 2 July 2015).

Cupido, K., and Van Belle, J.P. 2012. Increased public participation in local government through the use of mobile phones: what do young South Africans think? In: *Proceedings of the 12th European Conference on e-Government, Barcelona, Spain*, 159-168.

Cybenko, G. 1989. Approximation by superposition's of a sigmoidal function. M*ath. Control Signals Systems*, 2(4): 303-314.

Dalal, N. and Triggs, B. 2005. Histograms of oriented gradients for human detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005. I*EEE (1), 886-893.

Das, D. 2014. Activity recognition using histogram of oriented gradient pattern history. *International Journal of Computer Science, Engineering and Information Technology*, 4(4): 23-31.

Dave, M.R., Singh, J.K., Tiwari, M. and Khare, A. 2008 Implementation of intelligent polling system using GSM mobile. *International Journal of Computer Technology and Electronics Engineering (IJCTEE),* 1(2): 109-113.

Deller, J.R., Proakis, J.G. and Hansen, J.H. 2000. D*iscrete-time processing of speech signals*, 2nd ed. New York, NY, USA: IEEE, 516-553.

Done, R.S. 2002. *Internet voting: bringing elections to the desktop* Research Report PricewaterhouseCoopers, Endowment for the Business of Government.

Eddy, S.R. 1998. Profile hidden Markov models. B*ioinformatics*, 14(9): 755-763.

Ekong, U.O and Ekong, V.E. 2010. M-voting: a panacea for enhanced e-participation. A*sian Journal of Information Technology*, 9(2): 111-116. doi: 10.3923/ajit.2010.111.116.

Election Assistance Commission. 2011. A survey of Internet voting (online). Available: http://www.eac.gov/assets/1/Documents/SIV-FINAL.pdf. (Accessed 20 June 2015).

Elleithy, K. and Rimawi, I. 2006. Design, analysis and implementation of a cyber ote system. In: *Advances in Computer, Information, and Systems Sciences, and Engineering*, Netherlands: Springer, 219-226.

Elliott, S.J. 2007. Case study: phone-based voice biometrics for remote authentication. Paper presented at the *RSA Conference*, San Francisco, CA, 6 February 2007 (online). Available: http://www.slideshare.net/bspalabs/2007-case-study-phonebased-voice-biometrics-for-remote-authentication (Accessed 23 February 2015).

Ellis, A. Gratschew, M. Pammett, J. and Thiessen, E. 2006. Engaging the electorate: initiatives to promote voter turnout from around the world: including Voter Turnout data from national elections worldwide, 1945-2006. Stockholm: *International IDEA*, 2006.

Estonia Authority. 2014a. Internet Voting in Estonia, *Estonian National Electoral Committee* (online). Available: http://www.vvk.ee/voting-methods-in-estonia/engindex/ (Accessed 20 June 2015).

Estonia Authority. 2014b. Estonian national electoral committee, *Reports and Statistics about Internet Voting* (online). Available: http://www.vvk.ee/voting-methods-in-estonia/engindex/statistics (Accessed 20 June 2015).

Evans, N., Kinnunen, T. and Yamagishi, J. 2013. Spoofing and countermeasures for automatic speaker verification. 14th *Interspeech* Conference in Lyon, 25-29 August *2013*. In: *INTERSPEECH*, 925-929.

Everett, S.P. 2007. The usability of electronic voting machines and how votes can be changed without detection. PhD Thesis. Rice University.

Evgeny, K. 2011. Efficient speaker recognition for mobile devices. PhD Dissertation, University of Eastern Finland Dissertations in Forestry and Natural Sciences.

Farrell, K., Mammone, R., and Assaleh, K. 1994. Speaker recognition using neural networks and conventional classifiers. *IEEE Trans. on Speech and Audio Processing* 2(1): 194–205.

Fido. 2014. Specifications Overview (online). Available: https://fidoalliance.org/specifications (Accessed 20 June 2015).

Fujioka, A., Okamoto, T. and Ohta, K. 1992. A practical secret voting scheme for large scale election, In: *Proceedings of Auscrypt'92*, LNCS (718), 244–60.

Furui, S. 1981. Cepstral analysis technique for automatic speaker verification. *IEEE Transactions on Acoustics, Speech and Signal Processing,* 29(2), 254–27.

Furht, B. 2008. E*ncyclopedia of multimedia*. New York: Springer Science and Business Media.

Galton, F. 1892. *Finger prints*. London: Macmillan and Company.

Gentles, D. and Sankaranarayanan, S. 2011. Biometric secured mobile voting. In: *Second Asian Himalayas International Conference on Internet (AH-ICI), 2011*. IEEE, 1-6.

Gentles, D. and Sankaranarayanan, S. 2012. Application of biometrics in mobile voting, *International Journal of Communication Networks and Information Security* (*IJCNIS*), IJCNIS, 4(7): 57-68.

Gaafar, T.S., Abo Bakr, H.M and Abdalla, M. I. 2014. An improved method for speech/speaker recognition. In: *International Conference on Informatics, Electronics and Vision (ICIEV), 2014*, 1-5.

Galar, M., Fernandez, A., Barrenechea, E., Bustince, H. and Herrera, F. 2012. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE *Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews,* 42(4), IEEE, 463-484.

Gallant, L.M., Boone, G. and LaRoche, C.S. 2014. Mobile usability: state of the art and implications. *Interdisciplinary Mobile Media and Communications: Social, Political, and Economic Implications*, 344-354.

Gritzalis, D.A. 2002. Principles and requirements for a secure e-voting system. C*omputers and Security*, 21(6): 539-556.

Habib, A. 1997. South Africa - the rainbow nation and prospects for consolidating democracy. A*frican Journal of Political Science,* 2(2): 15–37.

Han, F., Hu, J. and Kotagiri, R. 2012. Biometric authentication for mobile computing applications. In: Li, H. ed. Advanced topics in biometrics, *World Scientific Publishing*, 461-482.

Hall, J.L. 2008. Policy mechanisms for increasing transparency in electronic voting. PhD Thesis. University of California at Berkeley.

Haynes, P. 2014. Online voting: rewards and risks. *Atlantic Council*, 2-5.

Harel, A. 2008. Biometrics, identification and practical ethics. In: Mordini, E. and Green, M. Ed. *Identity, Security and Democracy: The Wider Social and Ethical Implications of Automated Systems for Human Identification*. Amsterdam, NLD: IOS Press, 69-83.

Harrington, J. and Cassidy, S. 1999. *Techniques in speech acoustics*. Dordrecht: Kluwer Academic Publishers.

Harris, B 2013. Biometrics in elections and everyday life. B*lack box voting. E-voting forum* (online). Available: http://www.ronpaulforums.com/showthread.php?432518-Stay-Tuned-for-the-new-BLACK-BOX-VOTING (Accessed 23 June 2015).

Healy, J. 2014. Rewards of online voting: Estonia. *Atlantic Council*, 2.

Hecht, J. 2014. S*pectrum*. IEEE, 51(10): 36-41.

Heeks, R. 2010. Do information and communication technologies (ICTs) contribute to development?. *Journal of International Development*, *22*(5), 625-640.

Heiberg, S. 2013. New technologies for democratic elections. In: *Business Process Management Workshops*. Heidelberg Berlin: Springer, 630-635.

Heiberg, S. and Willemson, J. 2014. Verifiable Internet voting in Estonia. In: *6th International Conference on Electronic Voting: Verifying the Vote (EVOTE), 2014*. IEEE, 1-8.

Herschberg, M. A. 1997. Secure electronic voting over the world wide web. M.Eng Thesis. MIT.

Hermanns, H. 2008. Mobile democracy: mobile phones as democratic tools. *Politics*, 28(2): 74–82.

Hill, L. and Alport, K. 2007. Reconnecting Australia's politically excluded, electronic pathways to electoral inclusion. In*ternational Journal of Electronic Government Research (IJEGR)*, *3*(4), 1-19.

Hill, L. and Louth, J. 2006. Mobilising the youth and the future of British democracy. *International Foundation for Electoral Systems (IFES)* (online). Available:http://www.newcastle.edu.au/Resources/Schools/Newcastle%20Business%20School/APSA/PanelYouthPol/Hill-Lisa-and-Louth-Jonathon.pdf (Accessed 23 January 2015).

Howlader, J., Nair, V., Basu, S. and Mal, A.K. (2011). Uncoercibility in e-voting and e-auctioning mechanisms using deniable encryption. *IJNSA,* 3(2): 97-109.

Huang, X., Acero, A. and Hon, H.W. 2001. *Spoken language processing: a guide to theory, algorithm, and system development*. New Jersey: Prentice-Hall.

Huang, C.C., Wang, J.F., Wu, C.H. and Lee, J.Y. 1998. Speech recognition using dynamic programming of Bayesian neural networks. In: *Central Auditory Processing and Neural Modelling. US:* Springer, 71-76.

IEC. 2014. Electronic voting, an enabler or disabler to strengthening electoral democracy? In: *Seminar on Electronic Voting and Counting Technologies.* Cape Town, South Africa, March 12, 2013, 1-20.

IDEA. 2014. Voter turnout. *Institute for Democracy and Electoral Assistance (IDEA)* (online). Available: www.idea.int/vt (Accessed 23 February 2015).

International Data Corporation. 2013. *Smartphones shipments by manufacturer.* Available: http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats/a#smartphone-shipments (Accessed 23 February 2015).

Internet Policy Institute. 2001. R*eport of the National workshop on Internet Voting. issues and research agenda. N*ational Science Foundation and University of Maryland (online). Available: http://www.verifiedvoting.org/wp-content/uploads/2012/09/NSFInternetVotingReport.pdf (Accessed 2 January 2015).

Jain, A.K., Ross, A. and Prabhakar, S. 2004. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology,* , 14(1), IEEE, 4-20.

Jain, A.K. and Ross, A. 2008. Introduction to biometrics. In: Jain, A.K., Flynn, P Ross, A. *Handbook of Biometrics*. New York: Springer, 1–22. ISBN 978-0-387-71040-2.

Jain, A.K., Ross, A. and Nandakumar, K. 2011. Security Of biometric systems. Chapter In: Jain *et al*. ed. I*ntroduction to Biometrics*. New York: Springer, 259-306.

Jefferson, D. 2011. If I can shop and bank online, why can't I vote online? Voter Verified Report (online). Available: https://www.verifiedvoting.org (Accessed 25 July 2013).

Jefferson, D., Rubin, A.D., Simons, B. and Wagner, D. 2004a. Analyzing Internet voting security. C*ommunications of the ACM*, 47(10): 59-64.

Jefferson, D., Rubin, A.D., Simons, B. and Wagner, D. 2004b. A security analysis of the secure electronic registration and voting experiment (SERVE). 21 January 2004. Available: http://usacm.acm.org/images/documents/serve_report_full_paper.pdf (Accessed 1 July 2015). ACM, 1-34.

Joachims, T. 1999. Making large scale SVM learning practical. In: Schölkopf, B., and Burges, C.J. ed. *Advances in kernel methods: support vector learning*. Cambrodge: MIT press.

Joaquim, R., Zúquete, A. and Ferreira, P. 2003. REVS–a robust electronic voting system. *IADIS International Journal of WWW/Internet*, 1(2): 47-63.

Jones, D.W. 2003. A brief illustrated history of voting. University of Iowa Department of Computer Science (online). Available: http://homepage.cs.uiowa.edu/~jones/voting/pictures/ (Accessed 1 Jan 2015).

Jones, D.W. 2005. Chain voting. In: *Workshop on Developing an Analysis of Threats to Voting Systems,* National Institute of Standards and Technology.

Juan, S., and Shoulian, T. 2010. Operator's mobile Internet strategy in the process of converged network. In: IEEE *International Conference on Management and Service Science (MASS), 2010,* IEEE, 1-4.

Kagal, L., Finin, T. and Joshi, A. 2001. Trust-based security in pervasive computing environments. C*omputer*, 34(12): 154-157.

Kagal, L., Undercoffer, J., Perich, F., Joshi, A. and Finin, T. 2002. Vigil: enforcing security in ubiquitous environments. In: *Grace Hopper Celebration of Women in Computing.* Vancouver: Canada.

Kalvet, T. 2009. Management of technology: the case of e-voting in Estonia. *International Conference on Computer Technology and Development. ICCTD'09*. IEEE, 2009, 512-515.

Kanungo, D.P., Sharma, S., and Pain, A. 2014. Artificial neural network (ANN) and regression tree (CART) applications for the indirect estimation of unsaturated soil shear strength parameters. F*rontiers of Earth Science*, 8(3): 439-456.

Karpov, E. 2011. Efficient speaker recognition for mobile devices. PhD dissertation. University of Eastern Finland Dissertations in Forestry and Natural Sciences.

Kelleher, W.J. 2013. How NIST Has Misled Congress and the American People about Internet Voting Insecurity; or, Internet Voting in the USA: History and Prospects. *Internet Voting in the USA: History and Prospects (March 6, 2013)*.

Kersting, N and Baldersheim, H. 2004. *Electronic Voting and Democracy: A Comparative Analysis.* Basingstoke: Palgrave Macmillan.

Khelifi, A., Grisi, Y., Soufi, D., Mohanad, D. and Shastry, P. V. S. 2013. M-vote: a reliable and highly secure mobile voting system. In: *Palestinian International Conference on Information and Communication Technology (PICICT), 2013*. IEEE, 90-98.

Kim, K. and Hong, D. 2007. Electronic voting system using mobile terminal. *World Academy of Science, Engineering and Technology*, (32)1105-1109.

Kinnunen, T. and Li, H. 2010. An overview of text-independent speaker recognition: from features to supervectors. *Speech communication*, 52(1): 12-40.

Kinnunen, T. Karpov, E. and Fränti, P. 2006. Real-time speaker identification and verification, *IEEE Transactions on Audio, Speech and Language Processing*, 14(1), January 2006. IEEE*,* 277-288.

Kobayashi, T., Hidaka, A. and Kurita, T. 2008. Selection of histograms of oriented gradients features for pedestrian detection. In: *Neural Information Processing. H*eidelberg: Springer Berlin, 598-607.

Kohno, T., Stubblefield, A., Rubin, A. D. and Wallach, D. S. 2004. Analysis of an electronic voting system. In: *2004 IEEE Symposium on Security and Privacy, 2004. Proceedings.*, IEEE, 27-40.

Kogeda, O.P., and Mpekoa, N. 2013. Model for a mobile phone voting system for South Africa. In: *Proceedings of 15th Annual Conference on World Wide Web Applications (ZAWWW 2013).*

Krimmer, R., Triessnig, S. and Volkamer, M. 2007. The development of remote e-voting around the world: A review of roads and directions. In: *E-voting and identity*. Heidelberg: Springer Berlin, 1-15.

Krimmer, R., Schuster, R., and CC, E.V. 2008. The e-Voting readiness index. In: *Electronic Voting*, 127-136.

Kumar, M. 2003. Digital image processing. In: *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology, Proceedings of the Training Workshop, 7-11 July, 2003*. Dehra Dun, India, 81-102.

Kumar, D. Sahay, R. Hegde, G.R. and Jena, D. 2011. A novel simple secure Internet voting Protocol. In: *Proceedings of the 2011 International Conference on Communication, Computing and Security (ICCCS '11),* Odisha, India, 12-14 February 2011. ACM, 586-589.

Kushner, D. 2013. The real story of Stuxnet. S*pectrum, IEEE*, 50(3), 48-53.

Kyrillidis, L., Cobourne, S., Mayes, K., Dong, S. and Markantonakis, K. 2012. Distributed e-voting using the smart card web server. In: *7th International Conference on Risk and Security of Internet and Systems (CRiSIS), 2012*, IEEE, 1-8.

Larcom, J. A. and Liu, H. 2013. Modelling and characterization of GPS spoofing. In: *IEEE International Conference on Technologies for Homeland Security (HST), 2013*. IEEE, 729-734.

Lei, Z., Yang, Y. and Wu, Z. 2006 Ensemble of support vector machine for text-independent speaker recognition. *International Journal of Computer Science, Network and Security*, 6(5): 163-167.

Liu, M., Huang, T.S. and Zhang, Z. 2006a. Robust local scoring function for text-independent speaker verification. In: *18th International Conference on Pattern Recognition, ICPR 2006,* (2). IEEE, 1146-1149.

Liu, M., Ning, H., Huang, T.S. and Zhang, Z. 2006b. A novel framework of text-independent speaker verification based on utterance transform and iterative cohort modeling. *Urbana*, 51: 61801.

López-Pintor, R. and Fischer , J. 2006. G*etting to the CORE: a global survey on the cost of registration and elections.* UNPD.

Malode, A. A, and Sahare, S. 2012. An improved speaker recognition by using VQ and HMM. I*ET Third International Conference on Sustainable Energy and Intelligent System (SEISCON 2012)*, VCTW, Tiruchengode, Tamil Nadu, India on 27-29 December.

Mansfield, T., Kelly, G., Chandler, D. and Kane, J. 2001. Biometric product testing final report. C*omputing, National Physical Laboratory.* UK: Crown Copyright.

Mansfield, A.J., and Wayman, J.L. 2002. *Best practices in testing and reporting performance of biometric devices*, Teddington, Middlesex, UK: Centre for Mathematics and Scientific Computing, National Physical Laboratory, 1-36.

Markowitz, J. A. 2000. Voice biometrics. C*ommunications of the ACM*, 43(9): 66-73.

Mauw, S., Verschuren, J. and de Vink, E. P. 2007. Data anonymity in the FOO voting scheme. *Electronic Notes in Theoretical Computer Science*, 168: 5-28.

McGaley, M.A. 2008. Electronic voting: a safety critical system. PhD dissertation, Department of Computer Science, National University of Ireland.

McGrane, K. 2013. The rise of the mobile-only user. *Harvard Business Review* (online). Available: https://hbr.org/2013/05/the-rise-of-the-mobile-only-us/ . (Accessed 28 May 2015).

Mendez, F. and Serdült, U. 2014. From initial idea to piecemeal implementation: Switzerland's first decade of Internet voting reviewed. In: Zissis, D. and Lekkas, D. ed. D*esign, Development, and Use of Secure Electronic Voting Systems*, Hershey, PA: Information Science Reference, 115-127. doi:10.4018/978-1-4666-5820-2.ch006.

Mercuri, R. 2000. Electronic vote tabulation checks and balances. PhD Thesis. Department of Computer and Information Systems. University of Pennsylvania.

Mercuri, R. 2002. A better ballot box? *Spectrum, IEEE*, 39(10): 46-50.

Meyers, L. 2004. An exploration of voice biometrics. SANS institute (online). Available:http://www.sans.org/reading-room/whitepapers/authentication/exploration-voice-biometrics-1436 (Accessed 20 June 2015).

Modi, S.K. 2011. Biometrics in identity management: Concepts to applications. London: *Artech House*.

Mossberger, K., Tolbert, C.J., and McNeal, R.S. 2008. Digital citizenship. *The internet, society, and participation*, 1.

Moynihan, D.P. 2004. Building secure elections: e-Voting, security, and systems theory. *Public administration review*, 64(5): 515-528.

Mucunguzi, A. 2010. Conversations on technology: e-voting in Africa: Mr. Collin Thakur Interview,. *PCTechMagazine Uganda*, September October 2010, 26.

Mulliner, C. 2010. Privacy leaks in mobile phone Internet access. In: *14th International Conference on*, *Intelligence in Next Generation Networks (ICIN), 2010* IEEE, 1-6.

Mursi, M.F., Assassa, G.M., Abdelhafez, A. and Abo, K.M. 2013. On the development of electronic voting: a survey. *International Journal of Computer Applications*, 61(16): 1-11.

Naik, J., Netsch, L. and Doddington, G. 1989. Speaker verification over long distance telephone lines. In: *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 1989),* Glasgow, 524–527.

Nagaswamy, R. 2003. The Historic Village in Tamil Nadu. Chennai: Tamil Arts Academy.

Narasimha Rao, G.V.L. 2010. Democracy at risk: can we trust our electronic voting machines? New Delhi: *Veta.*

Narasimhan, T. E. 2012. Temple of democracy. *Business Standard.* 7 July 2012.

Nehamus, A. and Woodruff, P. 1995. Plato:Phaedrus (Translators) In: Phaedrus Plato. ed. Indianapolis: Hackett Publishing Company.

EML7.0. 2011. OASIS Election Markup Language (EML) specification version 7.0 committee specification. (online). Available: http://docs.oasis-open.org/election/eml/v7.0/eml-v7.0.html (Accessed 10 June 2015).

O' Connor, M.C. 2010. Costa Rica counts on RFID to monitor ballots. *RFID Journal* (online). Available: http://www.rfidjournal.com/articles/view?7984/2 (Accessed 25 January 2015).

O' Neil King, R. Speech and voice recognition. White paper. *VoiceTrust* (online). Available: http://www.biometricupdate.com/tag/voicetrust (Accessed 14 February 2015).

Ok K., Coskun V., and Aydin M.N. 2010. Usability of mobile voting with NFC technology, *Proceedings of the International Conference on Software Engineering (IASTED),* Innsbruck, AUSTRIA, 16-18 February 2010, 151-158.

Norway Ministry. 2014. *Internet voting pilot to be discontinued*. Government.no. (online). Available: https://www.regjeringen.no/en/aktuelt/Internet-voting-pilot-to-be-discontinued/id764300/ (Accessed 7 March 2015).

Olaniyi, O.M., Arulogun, O.T., Omidiora, E.O. and Adeoye O. 2013. Design of secure electronic voting system using multifactor authentication and cryptographic Hash Functions. *International Journal of Computer and Information Technology (IJCIT)*, 2(6): 1122-1121.

Olusola, O.O., Olusayo, O.E., Olatunde, O.S. and Adesina, G.R. 2012. A Review of the Underlying Concepts of Electronic Voting. In: *Information and Knowledge Management* (2)(1): 8-20.

Øyvann, S. 2013. Vote early, vote often: inside Norway's pioneering open source e-voting trials (online). *Communications of the ACM*. Sep 16, 2013, Available: http://cacm.acm.org/news/167797-vote-early-vote-often-inside-norways-pioneering-open-source-e-voting-trials/fulltext (Accessed 15 June 2015).

Oostveen, A.M., and van den Besselaar, P. 2003. E-voting and media effects, an exploratory study. In: *Conference on New Media, Technology and Everyday Life in Europe,* London, 23-26 April 2003.

Oostveen, A.M. 2007. Context matters. A social informatics perspective on the design and implications of large-scale e-Government systems. PhD Thesis. University Of Amsterdam.

Oostveen, A.M. 2010. Outsourcing democracy: losing control of e‐Voting in the Netherlands. *Policy and Internet*, 2(4): 201-220.

Oren, Y. and Wool, A. 2010. RFID-based electronic voting: what could possibly go wrong? In: *2010 IEEE International Conference on RFID*. IEEE, 118-125.

Oren, Y., Schirman, D. and Wool, A. 2012. RFID jamming and attacks on Israeli e-voting. In: *Smart Objects, Systems and Technologies (SmartSysTech), Proceedings of 2012 European Conference on*. VDE, 1-7.

Osuna, E., Freund, R. and Girosi, F. 1997. Training support vector machines: an application to face detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997. Proceedings.* IEEE*,* 130-136.

Pammett, J.H. and Goodman, N. 2013. Consultation and evaluation practices in the implementation of Internet Voting in Canada and Europe. Ottawa: *Elections Canada.*

Parmanto, B., Munro, P.W. and Doyle, H.R. 1996. Reducing variance of committee prediction with resampling techniques. *Connection Science*, 8(3) & (4): 405-426.

Pardue, H., Landry, J. and Yasinsac, A. 2010. A risk assessment model for voting systems using threat trees and Monte Carlo simulation. In: *2009 First International Workshop on Requirements Engineering for e-Voting Systems (RE-VOTE*). IEEE*,* 55-60.

Patil, K.I. and Shimpi, J. 2013. A graphical password using token, biometric, knowledge based authentication system for mobile devices. *International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN*, 2278-3075.

Pato, J.N. and Millet, L.I. 2010. ed. B*iometric recognition: challenge and opportunities.* The National Academies Press.

Paul, L. and Anilkumar, M.N. 2012. Authentication for online voting using steganography and biometrics, *IJARCET,* 1 December 2012, (10): 26-32.

Pavešić, N and Ribarić, S. 2009. Biometric recognition: an overview in identity, security and democracy E. In: Green, M. ed. I*OS Press*, 2009, 43-55.

Pieters, W. 2009. Combacting electoral traces: the Dutch tempest discussion and beyond. In: Ryan, P.Y.A. and Schoenmakers. ed. V*ote:ID 2009*, LNCS. Heidelberg: *Springer-Verlag Berlin,* 172-190.

Pocovnicu, A. 2009. Biometric security for cell phones. *Informatica Economica* 13(1): 57-63.

Poushter, J. and Oates, R. 2015. CellPhones in Africa: communication lifeline. Washingston DC: *Pew Research Centre.*

Popescu, M.C., Balas, V.E., Perescu-Popescu, L. and Mastorakis, N. 2009 Multi-layer perceptron and neural networks, *WSEAS Transactions on Circuits and Systems*, 8(7): 579-588.

Popoveniuc, S., Kelsey, J., Regenscheid, A. and Vora, P. 2010. Performance requirements for end-to-end verifiable elections. In: L *Proceedings of the 2010 international conference on Electronic voting technology/workshop on trustworthy elections*. USENIX Association, 1-16.

Prasad, H.K., Wolchok, S., Wustrow, E., Halderman, J.A., Kankipati, A., Sakhamuri, S.K., ... & Gonggrijp, R. 2010. Security analysis of India's electronic voting machines. In *Proceedings of the 17th ACM conference on Computer and communications security.* ACM, 1-14.

Prosser, A. and Krimmer, R. 2004. The dimensions of electronic voting technology, law, politics and society. *Electronic Voting in Europe Technology, Law, Politics and Society*, 21-28.

Rabiner, L. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2): 257-286.

Rabiner, L. and Juang, B.H. 1986. An introduction to hidden Markov models. *ASSP Magazine*, 3(1), IEEE, 4-16.

Rabiner, L. and Juang, B.H. 1993. *Fundamentals of speech recognition*. Englewood Cliffs, New Jersey: Prentice-Hall.

Ramachandran, R.P., Farrell, K.R., Ramachandran, R. and Mammone, R.J. 2002. Speaker recognition general classifier approaches and data fusion methods. *Pattern Recognition*, 35(12): 2801-2821.

Rameshkumar, G.P. and Samundeswari S. 2014. Neural network, artificial neural network (ANN) and biological neural network (BNN). S*oft Computing*, 3(3): 1159-1161.

Rehman, M.Z, and Nawi, N.M. 2011. Improving the accuracy of gradient descent back propagation algorithm (GDAM) on classification problem. *International Journal on New Computer Architectures and Their Applications*, 1(4): 838-847.

Reed, P. 2002. Reference architecture: the best of best practices. *The Rational Edge:* IBM.

REVS. 2015. R*obust Electronic Voting System* (online). Available: http://www.gsd.inesc-id.pt/~revs/ (Accessed 15 June 2015).

Reynolds, D. 1995. Speaker identification and verification using Gaussian mixture speaker models. S*peech Communication,* (17): 91–108.

Reynolds, D. 2002. An overview of automatic speaker recognition. In: *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (ICASSP), 4072-4075.

Reynolds, D. 2009. Gaussian mixture models. *Encyclopaedia of Biometrics*, 659-663.

Rhodes, R.J 2004. *Athenian democracy*. London: Oxford University Press.

Rose, P. 2002. Forensic speaker identification, *Taylor and Francis forensic science series*. New York: Taylor and Francis.

Rumelhart, D. Hinton, G.E, and Williams R.J. 1986. Learning representations by back-propagating errors. N*ature*, (323): 533-536.

Rumelhart, D. and J. McClelland. 1986. *Parallel distributed processing*. Cambridge, Mass: MIT Press.

Ryan, P.Y. and Teague, V. 2013. Pretty good democracy. In: *Security Protocols XVII*. Heidelberg: Springer Berlin, 111-130.

Saltman, R.G. 2006. *The history and politics of voting technology. Inquest of integrity and public confidence*. New York: Palgrave Macmillan.

Sampath, S.S. 2013. E-voting: the Indian experience lecture at Electoral Commission of South Africa. S*eminar on counting technologies*, Cape Town, 11-12 March 2013.

Sandler, D. Derr, K. and Wallach, D.S. 2008. VoteBox: a tamper-evident, verifiable electronic voting system. In: *Proceedings of the 17th conference on Security symposium* (SS'08). Berkeley, CA, USA: USENIX Association, 349-364.

Sanjith, R and Deokaran, Y. 2013. Interview OneVault Voice Biometric Company, Johannesburg, 5 November 2013.

Schneider, S. and Woodward, A. 2012. E-voting: trust but verify. *Scientific American* (online). Available: http://www.scientificamerican.com/author/steve-schneider-and-alan-woodward/ (Accessed 21 June 2015).

Schölkopf, B., and Smola, A.J. 2002. L*earning with kernels: support vector machines, regularization, optimization, and beyond*. Cambridge: MIT Press.

Schulz-Herzenberg, C. 2014. Voter participation in the South African elections of 2014. *Institute for Security Studies.* Policy Brief 61.

Scott, D., Vawda, M., Swartz, S. and Bhana, A 2012. Punching below their weight: young South Africans' recent voting patterns. *Human Science Research Council,* 19(3): 19-21.

Sabareeswari, T.C. and Stuwart, S. 2010. Identification of a person using multimodal biometric system. *International Journal of Computer Applications*, 3(9): 12-16.

Sako, K. and Kilian, J. 1994. Secure Voting Using Partially Compatible Homomorphisms. *Proceedings of Crypto'94*, LNCS 839, Springer-Verlag, 411-424.

Selvan, A.M. and Rajesh, R. 2012. Spectral histogram of oriented gradients (SHOGs) for Tamil language male/female speaker classification. *International Journal of Speech Technology,* 15(2): 259-264.

Shamos, M.I. 2004. Paper vs. electronic voting records-an assessment. In: *Proceedings of the 14th ACM Conference on Computers, Freedom and Privacy,* 1-23.

Shamos, M.I, and Yasinsac, A. 2012. Realities of e-voting security. S*ecurity and Privacy,* IEEE, 10(5): 16-17.

Sherrif, L. 2007. Scottish poll probe: e-counting gets 'hold off until safe' verdict. Not to blame, but not a fabulous idea (online). *The Register. 26 October 2007*. Available: http://www.theregister.co.uk/2007/10/26/scottish_elections (Accessed 15 June 2015).

Siau, K., and Shen, Z. 2003. Building customer trust in mobile commerce. *Communications of the ACM*, *46*(4): 91-94.

Sinclair, R.K. 1991. *Democracy and participation in Athens*. Cambridge: Cambridge University Press.

Sinclair, R.C. Mark, M.M., Moore, S.E., Lavis, C. A. and Soldat, A.S. 2000. An electoral butterfly effect. *Nature,* (408): 665-666.

Simons, B. and Jones, D.W. 2012. Internet voting in the US. *Communications of the ACM*, 55(10): 68-77. doi: 10.1145/2347736.2347754.

Snelick, R., Uludag, U., Mink, A., Indovina, M. and Jain, A. 2005. Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 27(3): 450-455.

Solomon, H. 2014. Background: why the use of voice is increasing (online). *IT World Canada,* 11 March 2014. Available: http://www.itworldcanada.com/article/background-why-the-use-of-voice-biometrics-is-increasing/90217#ixzz3LzfMvIjB (Accessed 20 June 2015).

Solomon, H. 2014b. Alliance of tech leaders vow to kill passwords (online). *IT World Canada,* 9 December 2014. Available: http://www.itworldcanada.com/article/alliance-of-tech-leaders-vows-to-kill-passwords/100474?sub=391742&utm_source=391742&utm_medium=top5&utm_campaign=TD (Accessed 20 June 2015).

Soong, F.K., Rosenberg, A.E., Juang, B.H. and Rabiner, L.R. 1987. Report: a vector quantization approach to speaker recognition. A*T&T technical journal*, 66(2): 14-26.

Souppaya, M. and Scarfone, K. 2013. Guidelines for managing the security of mobile devices in the enterprise. *NIST Special Publication*, (800): 124.

Springall, D., Finkenauer, T., Durumeric, Z., Kitcat, J., Hursti, H., MacAlpine, M. and Halderman, J. A. 2014. Security analysis of the Estonian Internet voting system. In: *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, ACM, 70-715.

Spycher, O., Volkamer, M. and Koenig, R. 2012. Transparency and technical measures to establish trust in Norwegian Internet voting. In: *E-voting and Identity*, Heidelberg: Springer Berlin, 19-35.

Stenerud, I.S.G. and Bull, C. 2012. When reality comes knocking Norwegian experiences with verifiable electronic voting. *Electronic Voting,* (205): 21-33.

Tanenbaum, A. and Woodhull, A. 2006. *Operating systems: design and implementation* (3rd ed.) Amsterdam: Prentice-Hall.

Thakur, S. 2009. *An investigation into the nature and extent of the adoption of RFID in the KwaZulu-Natal Province of South Africa*. Masters Dissertation.

Thakur, S., Adetiba, E., Olugbara, O.O., Millham, R. 2015. Experimentation using short-term spectral features for secure mobile internet voting authentication. *Hindawi Mathematical Problems in Engineering. In Press.*

Thakur, S. and Boateng, R. 2011. E-Voting for good governance and a green world, Conference Abstract. In: *Proceedings of the Africa Digital Week*, Accra, Ghana: African Institute of Development Informatics and Policy, July 26-29 2011, 55-81.

Thakur, S. 2012. *Electronic voting: the cross national experience*. Commissioned Research. The Electoral Commission of South Africa (IEC).

Thakur, S. 2013. E-voting: an enabler or disabler to strengthening electoral democracy. *Seminar on electronic voting and counting*, Cape Town, 12-13 March 2014.

Thakur, S. 2012. Digital democracy: using e-voting to empower nations? Is it On or Off? *Keynote. 14th Annual Conference on ZAWWW Applications,* Mangosuthu University of Technology, Durban, South Africa, Nov 7-9 2012.

Thakur, S. and Dávila, R. 2013. *The path towards effective solutions: a study on voter registration experiences and technology*. UNDP.

Thakur, S. 2015. E-voting: India and the Philippines – a comparative analysis for possible adaptation in Africa. In: Sodhi, I. ed. *Emerging Issues and Prospects in African E-Government*. Hershey, PA: Information Science Reference, 28-55. doi:10.4018/978-1-4666-6296-4.ch003.

Thakur, S. and Beer, C. 2014. An interactive token based system to seamlessly recognise documents. *Artefact* (online). Available: http://www.authenticateit.co.za. (Accessed 20 June 2015).

Thakur, S., Olugbara, O.O., Millham, R., Wesso, H.W., Sharif, M. and Singh, P. 2014. Asia-Pacific Institute of Management. Transforming the voting paradigm – the shift from inline, to online to mobile voting. *International Summer School on Information and Communication Technology for Democracy*. New Delhi, India, March 9-15, 2014.

Thakur, S., Olugbara, O.O., Millham, R., Wesso, H.W., Sharif, M. and Singh, P. 2015. Transforming the voting paradigm - the shift from inline, through online to mobile voting. *IEEE International Conference On Adaptive Science and Technology (ICAST)*, Lagos, Nigeria, 29 October 2014, IEEE.

Thakur, S. and Singh, S. 2012. A study of some e-government activities in South Africa. In: *e-Leadership Conference on Sustainable e-Government and e-Business Innovations (E-LEADERSHIP*), University of Pretoria, , October 2012, IEEE, 1-11.

Thakur, S. 2015. E-voting: India and the Philippines – a comparative analysis for possible adaptation in Africa. In Sodhi, IS. ed. E*merging Issues and Prospects in African E-Government*, Hershey, PA: Information Science Reference, 28-55. . doi:10.4018/978-1-4666-6296-4.ch003.

Top, R. 2014. South African agency uses voice authentication to reduce fraud in welfare payment distribution. O*pus Research*. 26 August 2014.

Olugbara, O.O. and Ndlovu, B.N. 2014. Constructing frugal sales system for small Enterprises. *The African Journal of Information Systems*, 6(4), Article 1: 119-139.

Ullah, M., Din, N., Umar, A.I., Amin, N.U. and Amin, S. 2014. An efficient mobile phone voting system based on blind signcryption. In: *4th International Conference on Computer and Emerging Technologies.* Shah Abdul Latif University. Khairpur Mirs, Sindh, Pakistan. 20-12 March 2014.

Ullah, M., and Umar, A.I. 2013. An efficient and secure mobile phone voting system. In: *2013 Eighth International Conference on Digital Information Management (ICDIM),* IEEE, 332-336.

Volkamer, M. 2009. Evaluation of electronic voting: requirements and evaluation procedures to support responsible election authorities (30). Springer Science and Business Media.

van de Haar, H., van Greunen, D. and Pottas, D. 2013. The characteristics of a biometric. In: *Information Security for South Africa, 2013*, IEEE, 1-8.

Voting Guidelines. 2005. Voluntary Voting System Guidelines, *US Election Assistance Commission,* 1(1).

Watson, R.T. 2013. Africa's contributions to information systems. *The African Journal of Information Systems*, 5(4): 125-130.

Watson, R.T., Kunene, K.N. and Islam, M.S. 2013. View from practice frugal information systems (IS). *Information Technology for Development*, Malacca, 19(2): 158–161.

Wayman, J.L 2001. Fundamentals of biometric authentication technologies, *International Journal of Image Graphics*, 1(1): 93-113.

Wells, C. 2013. *Inside Bush v. Gore*. Gainesville: University Press of Florida.

Weilenmann, A. and Larsson, C. 2002. Local use and sharing of mobile phones. In: *Wireless world,* London: Springer, 92-107.

Whither Biometrics Committee. 2010. *Biometric recognition: challenges and opportunities*. National Academies Press, 2010.

Wildermoth, B. and Paliwal, K.K. 2000. Use of voicing and pitch information for speaker recognition. In: *Proceedings of 8th Australian International Conf. Speech Science and Technology,* Canberra, 324-328.

Wolchok, S., Wustrow, E., Isabel, D. and Halderman, J. A. 2012. Attacking the Washington, DC Internet voting system. In: *Financial Cryptography and Data Security*. Heidelberg: Springer Berlin, 114-128.

Wolf, J.J. 1972. Efficient acoustic parameters for speaker recognition, *Journal of Acoustical Society of America,* 51(6-2): 2044–2056.

Wong, A. 2006. Biometrics market: where are we now? *Biometric Technology Today*, 14(9): 7-9.

Woodward Jr, J.D. 1996. Biometrics: identifying law and policy concerns. In: *Biometrics*, US: Springer, 385-405.

Yuan-Yuan, Q., Jie, Y. and Zhen-Ming, L. 2013. Structural analysis of complex networks from the mobile Internet. In: *Proc. Nat. Doctoral Acad. Forum Inf. Commun. Technol*., 1-7 August 2013.

Zakas, N. C. (2013). The evolution of web development for mobile devices. *Queue*, 11(2): 30.

Zakiuddin, I., Creese, S., Roscoe, B., and M. Goldsmith. 2003. Authentication in Pervasive Computing, Position Paper. Presented at *PAMPAS '02 - Workshop on Requirements for Mobile Privacy and Security.* Royal Holloway, University of London.

Zambrano, R. and Seward, R.K. 2011. *Mobile technologies and empowerment*. New York: UNDP Publication.

Zetter, K. 2003. Media Aussies do it right: e-Voting. Email: Wired news. 11 March 2003, *Wired Magazine* (online). Available at http://archive.wired.com/techbiz/media/news/2003/11/61045?currentPage=all (Accessed 20 June 2015).

Zogby, J. and Kuhl, J.S. 2013. *First globals: understanding, managing and unleashing, the potential of our millennial generation.* E-book: John Zogby and Joan Snyder Kuhl Publishers.