Volume 8 • Issue 2 • April 2014

Editor-in-Chief
Dr. Chen-Chi Shing

INTERNATIONAL JOURNAL OF
COMPUTER SCIENCE AND SECURITY (IJCSS)

ISSN : 1985-1553

Publication Frequency: 6 Issues / Year

CSC PUBLISHERS
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND SECURITY (IJCSS)

VOLUME 8, ISSUE 2, 2014

EDITED BY
DR. NABEEL TAHIR

ISSN (Online): 1985-1553

International Journal of Computer Science and Security is published both in traditional paper form and in Internet. This journal is published at the website http://www.cscjournals.org, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJCSS Journal is a part of CSC Publishers
Computer Science Journals
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND SECURITY (IJCSS)

Book: Volume 8, Issue 2, April 2014
Publishing Date: 30-04-2014
ISSN (Online): 1985 -1553

IJCSS Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2014

EDITORIAL PREFACE

This is *Second* Issue of Volume *Eight* of the International Journal of Computer Science and Security (IJCSS). IJCSS is an International refereed journal for publication of current research in computer science and computer security technologies. IJCSS publishes research papers dealing primarily with the technological aspects of computer science in general and computer security in particular. Publications of IJCSS are beneficial for researchers, academics, scholars, advanced students, practitioners, and those seeking an update on current experience, state of the art research theories and future prospects in relation to computer science in general but specific to computer security studies. Some important topics cover by IJCSS are databases, electronic commerce, multimedia, bioinformatics, signal processing, image processing, access control, computer security, cryptography, communications and data security, etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 8, 2014, IJCSS appears with more focused issues. Besides normal publications, IJCSS intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

This journal publishes new dissertations and state of the art research to target its readership that not only includes researchers, industrialists and scientist but also advanced students and practitioners. The aim of IJCSS is to publish research which is not only technically proficient, but contains innovation or information for our international readers. In order to position IJCSS as one of the top International journal in computer science and security, a group of highly valuable and senior International scholars are serving its Editorial Board who ensures that each issue must publish qualitative research articles from International research communities relevant to Computer science and security fields.

IJCSS editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCSS. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJCSS provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
International Journal of Computer Science and Security (IJCSS)

EDITORIAL BOARD

Assistant Professor Vishal Bharti
Maharishi Dayanand University
India

Dr. Parvinder Singh
University of Sc. & Tech
India

Assistant Professor Vishal Bharti
Maharishi Dayanand University,
India

TABLE OF CONTENTS

Volume 8, Issue 2, April 2014

Pages

A Comparative Study: Gammachirp Wavelets and Auditory Filter Using Prosodic Features of Speech Recognition In Noisy Environment

Hajer Rahali *Hajer.Rahali@enit.rnu.tn*
National Engineering School of Tunis (ENIT)
Laboratory of Systems and Signal Processing (LSTS)
BP 37, Le Belvédère, 1002 Tunis, Tunisie

Zied Hajaiej *Zied.hajaiej@enit.rnu.tn*
National Engineering School of Tunis (ENIT)
Laboratory of Systems and Signal Processing (LSTS)
BP 37, Le Belvédère, 1002 Tunis, Tunisie

Noureddine Ellouze *N.ellouze@enit.rnu.tn*
National Engineering School of Tunis (ENIT)
Laboratory of Systems and Signal Processing (LSTS)
BP 37, Le Belvédère, 1002 Tunis, Tunisie

Abstract

Modern automatic speech recognition (ASR) systems typically use a bank of linear filters as the first step in performing frequency analysis of speech. On the other hand, the cochlea, which is responsible for frequency analysis in the human auditory system, is known to have a compressive non-linear frequency response which depends on input stimulus level. It will be shown in this paper that it presents a new method on the use of the gammachirp auditory filter based on a continuous wavelet analysis. The essential characteristic of this model is that it proposes an analysis by wavelet packet transformation on the frequency bands that come closer the critical bands of the ear that differs from the existing model based on an analysis by a short term Fourier transformation (STFT). The prosodic features such as pitch, formant frequency, jitter and shimmer are extracted from the fundamental frequency contour and added to baseline spectral features, specifically, Mel Frequency Cepstral Coefficients (MFCC) for human speech, Gammachirp Filterbank Cepstral Coefficient (GFCC) and Gammachirp Wavelet Frequency Cepstral Coefficient (GWFCC). The results show that the gammachirp wavelet gives results that are comparable to ones obtained by MFCC and GFCC. Experimental results show the best performance of this architecture. This paper implements the GW and examines its application to a specific example of speech. Implications for noise robust speech analysis are also discussed within AURORA databases.

Keywords: Gammachirp Filter, Wavelet Packet, MFCC, Impulsive Noise.

1. INTRODUCTION

In order to understand the auditory human system, it is necessary to approach some theoretical notions of our auditory organ, in particular the behavior of the internal ear according to the frequency and according to the resonant level. The sounds arrive to the pavilion of the ear, where they are directed towards drives in its auditory external. To the extremity of this channel, they exercise a pressure on the membrane of the eardrum, which starts vibrating to the same frequency those them. The ossicles of the middle ear, interdependent of the eardrum by the hammer, also enter in vibration, assuring the transmission of the sound wave thus until the cochlea. The resonant vibration arrives to the cochlea by the oval window, separation membrane between the stirrup, last ossicle of the middle ear, and the perilymphe of the vestibular rail. The endolymphe of the cochlear channel vibrates then on its turn and drag the basilar membrane. The stenocils, agitated by the liquidize movements, transforms the acoustic vibration in potential of action (nervous messages); these last are transmitted to the brain through the intermediary of the cochlear nerve [1]. These mechanisms of displacement on any point of the basilar membrane, can begins viewing like a signal of exit of a pass strip filter whose frequency answer has its pick of resonance to a frequency that is characteristic of its position on the basilar membrane [2]. To simulate the behavior of these filters, several models have been proposed. Thus, one tries to succeed to an analysis of the speech signals more faithful to the natural process in the progress of a signal since its source until the sound arrived to the brain. By put these models, one mentions the model gammachirp that has been proposed by Irino & Patterson. While being based on the impulsion answer of this filter type, it come the idea to implement as family of wavelet of which the function of the wavelet mother is the one of this one. In this paper, a design for modeling auditory is based on wavelet packet decomposition. The wavelet transform is an analysis method that offers more flexibility in adapting time and frequency resolution to the input signal. MFCC are used extensively in ASR. MFCC features are derived from the FFT magnitude spectrum by applying a filterbank which has filters evenly spaced on a warped frequency scale. The logarithm of the energy in each filter is calculated and accumulated before a Discrete Cosine Transform (DCT) is applied to produce the MFCC feature vector. It will be shown in this paper that the performance of MFCC, based on the gammachirp filter and referred to as GFCC, are also compared to GWFCC which integrate the gammachirp wavelet. In the current paper, prosodic information is first added to a spectral system in order to improve their performance, finding and selecting appropriated characteristics related to the human speech prosody, and combining them with the spectral features. Such prosodic characteristics include parameters related to the fundamental frequency in order to capture the into-nation contour, and other parameters such as the jitter and shimmer. The implementation of gammachirp wavelet shows consistent and significant performance gains in various noise types and levels. For this we will develop a system for automatic recognition of isolated words with impulsive noise based on HMM\GMM. We propose a study of the performance of parameterization techniques MFCC, GFCC and GWFCC including the prosodic features proposed in the presence of different impulsive noises. Then, a comparison of the performance of different used features was performed in order to show that it is the most robust in noisy environment. The sounds are added to the word with different signal-to-noise SNR (20dB, 15dB and 10dB). Note that the robustness is shown in terms of correct recognition rate (CRR) accuracy. The evaluation is done on the AURORA database.

This paper is organized as follow; in the next section we briefly introduce the prosodic features and auditory filterbank. Section 3 introduces the gammachirp filter as wavelet. The processing steps of our gammachirp wavelet parameterization are described in section 4. Section 5 demonstrates simulations tested with new method. Finally, conclusions are given in section 6.

2. THEORETICAL FRAMEWORK

In this phase of feature extraction we will represent both of spectral and prosodic features which are combined for the aim of creating a robust front-end for our speech recognition system.

2.1 Prosodic features

A) Pitch (Fundamental Frequency)

The vibration of the vocal folds is measured using pitch and is nearly periodic. Pitch frequency F0 is a very important parameter using to describe the characteristic of voice excitation source. The average rate of the vibration of vocal folds measured in the frequency domain is defined as pitch. The rate of vibration is inversely proportional to the shape and size of the vocal folds. The size of vocal folds is different from speaker to speaker and hence the pitch also contains uniqueness information of a speaker. In general, the size of the vocal folds in men is larger than that in women and accordingly pitch of men is lower than that of women. The average pitch for a male speaker is about 50-300 Hz and for a female speaker it is about 100-500 Hz [3].

B) Jitter

Fundamental frequency is determined physiologically by the number of cycles that the vocal folds do in a second. Jitter refers to the variability of F0, and it is affected mainly because of the lack of control of vocal fold vibration .On the other hand, vocal intensity is related to sub glottis pressure of the air column, which, in turn, depends on other factors such as amplitude of vibration and tension of vocal folds [3]. For our analysis, the following jitter measurements as defined in PRAAT. Mathematically, jitter is the cycle to cycle variation of the pitch period, i.e., the average of the absolute distance between consecutive periods. It is measured in µ sec. It is defined as:

$$\text{Jitter} = \frac{1}{N-1}\sum_{i=1}^{N-1}|T_i - T_{i+1}|. \quad (1)$$

Where Ti is the extracted F0 period length and N is the number of extracted F0 pitch periods. Absolute jitter values, for instance, are found larger in males as compared to females.

C) Shimmer

It is the variability of the peak-to-peak amplitude in decibels. It is the ratio of amplitudes of consecutive periods. It is expressed as:

$$\text{Shimmer (dB)} = \frac{1}{N-1}\sum_{i=1}^{N-1}|20\log(\frac{A_{i+1}}{A_i})|. \quad (2)$$

Where Ai is the peak-to-peak amplitude in the period and N is the number of extracted fundamental frequency periods. Local shimmer (dB) values are found larger in female as compared to males.

2.2 Gammachirp Auditory Filter

The gammachirp filter is a good approximation to the frequency selective behavior of the cochlea [4]. It is an auditory filter which introduces an asymmetry and level dependent characteristics of the cochlear filters and it can be considered as a generalization and improvement of the gammatone filter. The gammachirp filter is defined in temporal domain by the real part of the complex function:

$$g_c(t) = at^{n-1}e^{-2\pi Bt}\ e^{j2\pi f_r t+jc\ln t+jc\varphi}. \quad (3)$$

With

$$B = b.\mathrm{ERB}(f_r) = b.\ (24.7\ +\ 0.108\ (f_r)). \quad (4)$$

With: t > 0

N : a whole positive defining the order of the corresponding filter.

f_r : the modulation frequency of the gamma function.

φ : the original phase.

a : an amplitude normalization parameter.

c : a parameter for the chirp rate.

b : a parameter defining the envelope of the gamma distribution.

ERB(f_r) : Equivalent Rectangulaire Bandwith.

When c=0, the chirp term, c ln (t), vanishes and this equation represents the complex impulse response of the gammatone that has the envelope of a gamma distribution function and its carrier is a sinusoid at frequency f_r. Accordingly, the gammachirp is an extension of the gammatone with a frequency modulation term.

A) Energy

The energy of the impulse response g_c(t) is obtained with the following expression:

$$E_{n,B} = ||g_c||^2 = < g_c, g_c > = a^2\frac{\sigma(2n-1)}{(4\pi B)^{2n-1}}. \quad (5)$$

With σ(n) is the n-th order gamma distribution function. Thus, for energy normalization is obtained with the following expression:

$$A_{E_{n,B}} = \sqrt{\frac{4\pi B^{(2n-1)}}{\sigma(2n-1)}}. \quad (6)$$

B) Frequency response

The Fourier transform of the gammachirp in “(3)” is derived as follows [5].

$$|G_c\ (f)| = \frac{a|\sigma(n+jc)|}{\sigma(n)} * \frac{\sigma(n)}{\left|2\pi\sqrt{((b\mathrm{ERB}(f_r))^2+(f-f_r)^2}\right|^n}e^{c\theta}. \quad (7)$$

$$|G_c\ (f)| = a_\sigma\ |G_T| * e^{c\theta(f)}. \qquad (8)$$

$$\theta\ (f) = \arctan(\frac{f-f_r}{bERB(f_r)}). \qquad (9)$$

$|G_c(f)|$ is the fourier magnitude spectrum of the gammatone filter, $e^{c\theta(f)}$ is an asymmetric function since is anti-symmetric function centered at the asymptotic frequency. The spectral properties of the gammachirp will depend on the $e^{c\theta(f)}$ factor; this factor has therefore been called the asymmetry factor. The degree of asymmetry depends on “c”. If “c” is negative, the transfer function, considered as a low pass filter, where c is positive it behave as a high-pass filter and if “c” zero, the transfer function, behave as a gammatone filter. In addition, this parameter is connected to the signal power by the expression [6]:

$$c = 3.38 + 0.107\ Ps. \qquad (10)$$

C) Basic structure

Figure 1 shows a block diagram of the gammachirp filterbank. It is a cascade of three filterbanks: a gammatone filterbank, a lowpass-AC filterbank, and a highpass-AC filterbank [7]. The gammachirp filterbank consists of a gammatone filterbank and an asymmetric compensation filterbank controlled by a parameter controller with sound level estimation.

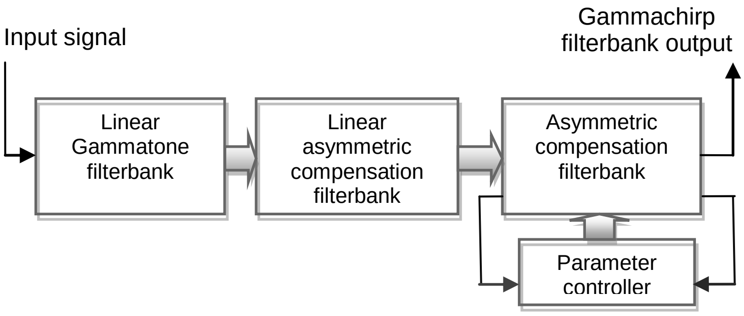
Input signal
Gammachirp filterbank output
Linear Gammatone filterbank
Linear asymmetric compensation filterbank
Asymmetric compensation filterbank
Parameter controller

FIGURE 1: Structure of the Gammachirp Filterbank.

This decomposition, which was shown by Irino in [8], is beneficial because it allows the gammachirp to be expressed as the cascade of a gammatone filter with an asymmetric compensation filter. Figure 2 shows the framework for this cascade approach.

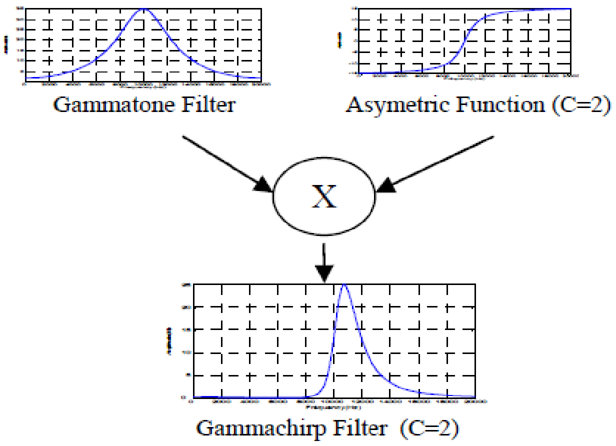

FIGURE 2: Decomposition of the Gammachirp Filter.

3. THE GAMMACHIRP FILTER AS A WAVELET

In this work, a new approach for modeling auditory based on gammachirp filters for application areas including speech recognition. The psychoacoustic model is based on the functioning of human ear. This model analyzes the input signal on several consecutive stages and determines for every pad the spectrum of the signal. The gammachirp filter underwent a good success in psychoacoustic research. Indeed, it fulfils some important requirements and complexities of the cochlear filter [5].

3.1 Wavelet Transform Analysis

The wavelet transform (WT) can be viewed as transforming the signal from the time domain to the wavelet domain. This new domain contains more complicated basis functions called wavelets, mother wavelets or analyzing wavelets. A wavelet prototype function at a scale s and a spatial displacement u is defined as: [9]

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\,\psi\left(\frac{t-u}{s}\right) \quad (u \in \mathrm{IR},\ s \in \mathrm{IR}_+^*). \qquad (11)$$

The WT is an excellent tool for mapping the changing properties of non-stationary signals. The WT is also an ideal tool for determining whether or not a signal is stationary in a global sense. When a signal is judged non-stationary, the WT can be used to identify stationary sections of the data stream. Specifically, a Wavelet Transform function f (t) Є L2(**R**) (defines space of square integrals functions) can be represented as:

$$W_{u,s}(f) = \int_{-\infty}^{+\infty} f(t)\ \psi_{u,s}{}^{*}(t)\,dt. \qquad (12)$$

The factor of scale includes an aspect transfer at a time in the time brought by the term u, but also an aspect dilation at a time in time and in amplitude brought by the terms s and $\sqrt{s}$.

3.2 The Gammachirp Filter As a Wavelet

The Gammachirp function which is a window modulated in amplitude by the frequency f_r and modulated in phase by the parameter c can thus be seen as wavelet roughly analytical [10] [11]. This wavelet has the following properties: it is with non compact support, it is not symmetric, it is non orthogonal and it does not present a scale function. The gammachirp function can be considered like wavelet function and constitute a basis of wavelets thus on the what be project all input signal, it is necessary that it verifies some conditions that are necessary to achieve this transformation. Indeed it must verify these two conditions:

- The wavelet function must be a finished energy “(5)”:

$$||\,g_c||^2 = a^2\,\frac{(2n-1)!}{(4\pi B)^{2n-1}}\ . \qquad (13)$$

$||\,g_c||^2$=1 if a = $\sqrt{\left(\frac{(4\pi B)^{2n-1}}{(2n-1)!}\right)}$ which define the filter of normalized energy.

- The wavelet function must verify the admissibility condition:

$$C_{g_c} = \int_{0}^{+\infty} \frac{|G_c\,(f)|^2}{f}\,df < +\infty. \qquad (14)$$

If the condition “(14)” is satisfied by the function G_c, then it must satisfy two other conditions:

- The mean function g is zero: $G_c(0) = \int_{-\infty}^{+\infty} g_c(t)dt = 0$
- The function G_c (f) is continuously differentiable

To implement the gammachirp function g_c as wavelet mother, one constructs a basis of wavelets then girls $g_{p,q}$ and this as dilating by factor ‘p’ and while relocating it of a parameter ‘q’.

$$g_{p,q}(t) = \frac{1}{\sqrt{p}}\ g_c\left(\frac{t-q}{p}\right). \qquad (15)$$

Studies have been achieved on the gammachirp function [10], show that the gammachirp function that is an amplitude modulated window by the frequency f_r and modulated in phase by the c parameter, can be considered like roughly analytic wavelet. It is of finished energy and it verifies the condition of admissibility. For this family of wavelet, the frequencies of modulation are $f_m = f_r\ .\ s_0{}^{-m}$ and the bandwidths are $B_m = B\ .\ s_0{}^{-m}$, s_0 is the dilation parameter and m∊Z.

The results show that the value 1000 Hz are the one most compatible as central frequency of the Gammachirp function. Otherwise our work will be based on the choice of a Gammachirp wavelet centered at the frequency 1000 Hz. For this frequency range, the gammachirp filter can be considered as an approximately analytical wavelet. The choice of the gammachirp filter is based on two reasons. First reason is that the gammachirp filter has a well defined impulse response, and it is excellent for an asymmetric, level-dependent auditory filterbank in time domain models of auditory processing. Second reason is that this filter was derived by Irino as a theoretically optimal auditory filter that can achieve minimum uncertainty in a joint time-scale representation.

4. IMPLEMENTATION

With the gammachirp filter designed as described above, a frequency-time representation of the original signal, which is often referred to as a Cochleagram, can be obtained from the outputs of the filterbank. It is then straightforward to compute MFCC, GFCC and GWFCC features from the Cochleagram. The remaining of this section presents the details of our implementation. In this study, our objective is to introduce new speech features that are more robust in noisy environments. We propose a robust speech feature which is based on the gammachirp filterbank and gammachirp wavelet.

Figure 3 shows the block diagrams of the extraction of MFCC and GFCC features. Figure 4 shows the block diagrams of the extraction of GWFCC features.

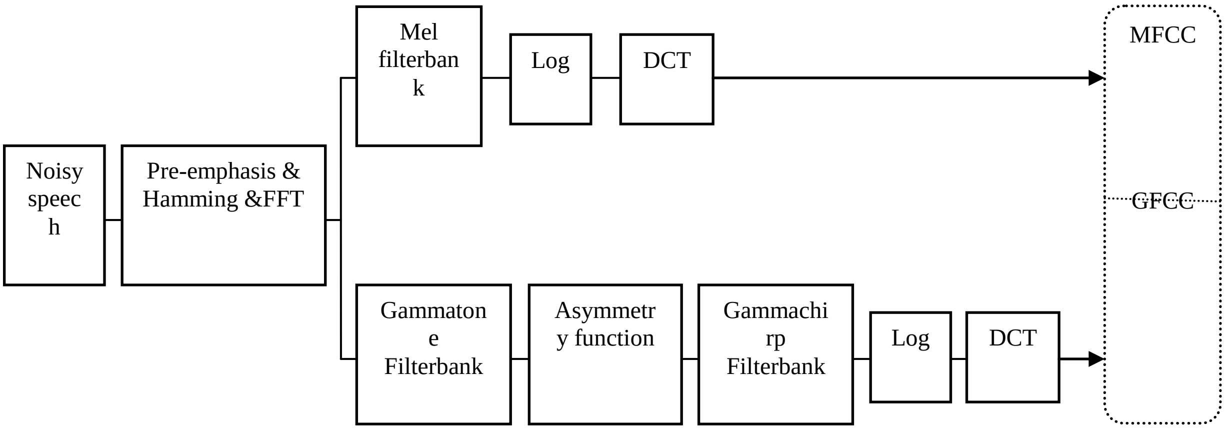

FIGURE 3: Block diagrams of the extraction of MFCC and GFCC features.

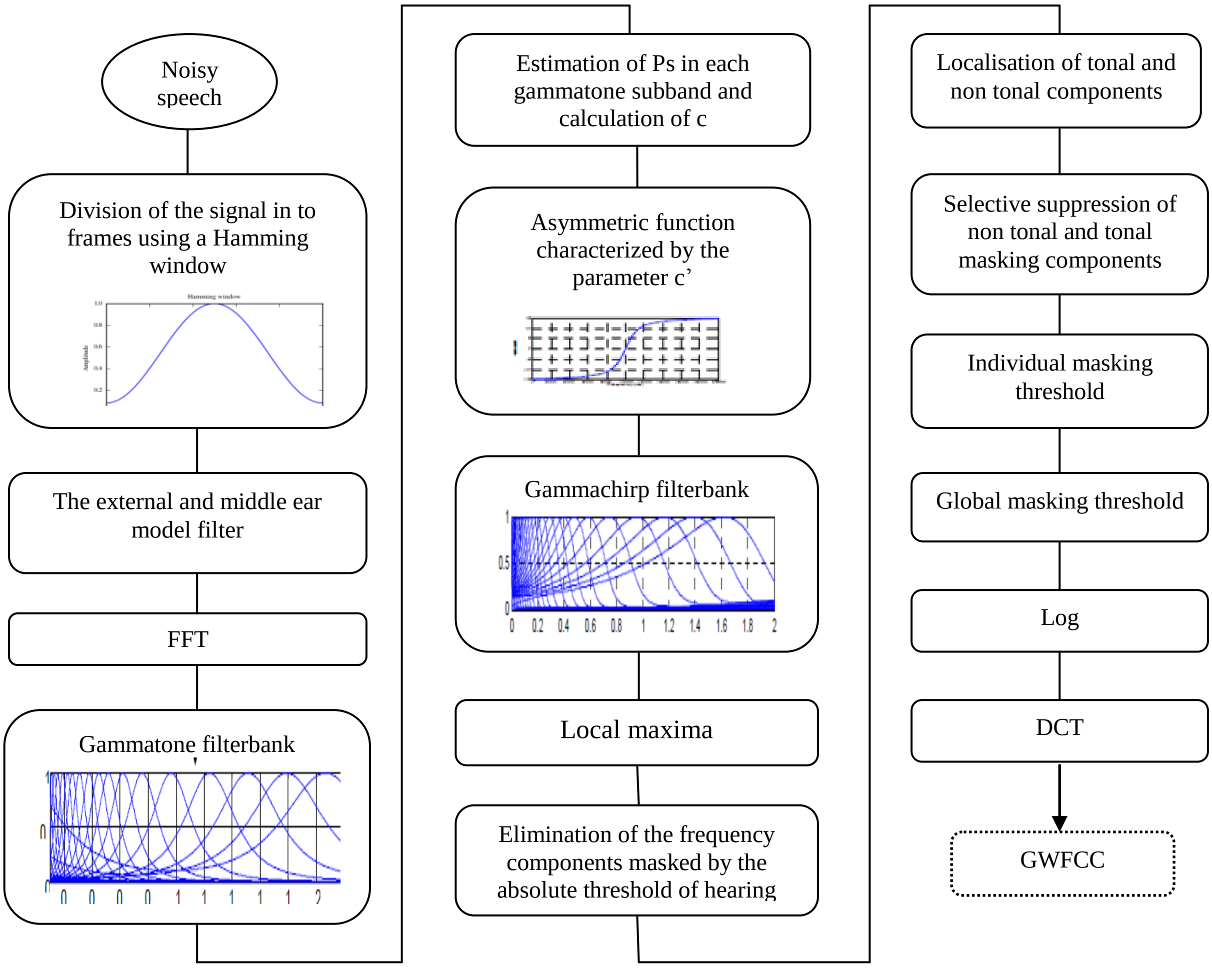
Noisy speech
Division of the signal in to frames using a Hamming window
The external and middle ear model filter
FFT
Gammatone filterbank
Estimation of Ps in each gammatone subband and calculation of c
Asymmetric function characterized by the parameter c'
Gammachirp filterbank
Local maxima
Elimination of the frequency components masked by the absolute threshold of hearing
Localisation of tonal and non tonal components
Selective suppression of non tonal and tonal masking components
Individual masking threshold
Global masking threshold
Log
DCT
GWFCC

FIGURE 4: Block diagrams of the extraction of GWFCC features.

4.1 MFCC and GFCC

Generally, both methods are based on two similar processing blocks: firstly, basic short-time Fourier analysis which is the same for both methods, secondly, cepstral coefficients computation. As illustrated in figure 3, it can be seen that one of the main dissimilarity between MFCC and GFCC is the set of filters used in the extraction. In fact, triangular filter bank equally spaced in the Mel scale frequency axis is used to extract MFCC features, while in GFCC, the gammachirp filterbank are used. The Mel Cepstral features are calculated by taking the cosine transform (DCT) of the real logarithm of the short-term energy spectrum expressed on a mel-frequency scale. After pre-emphasizing the speech using a first order high pass filter and windowing the speech segments using a Hamming window of 20 ms length with 10 ms overlap, the FFT is taken of these segments. The magnitude of the Fourier Transform is then passed into a filterbank comprising of 25 triangular filters. The GFCC are extracted from the speech signal according to the following steps; use the gammachirp filterbank defined in “(2)” with 32 filters and the bandwidth multiplying factor F = 1.5 to bandpass the speech signal. After, estimate the logarithm of the short-time average of the energy operator for each one of the bandpass signals, and estimates the cepstrum coefficients using the DCT. These steps are the main differences between MFCC and GFCC features extraction. The standard MFCC uses filters with frequency response that is triangular in shape (50% filter frequency response overlap). But, the proposed auditory use filters that are smoother and broader than the triangular filterbank (the bandwidth of the filter is controlled by the ERB curve and the bandwidth multiplication factor F). The main differences between the proposed filterbank and the typical one used for MFCC estimation are the type of filters used and their corresponding bandwidth. In this paper, we experiment with two parameters to create a family of gammachirp filterbanks: firstly, the number of filters in the filterbank, secondly, the bandwidth of the filters ERB (f). The bandwidth of the filter is obtained by multiplying the filter bandwidth curve ERB by the parameter F. Experimental results provided in the next section show that both parameters are important for robust speech recognition. The range of parameters we have experimented is 20 – 40 for the number of filters and 1,0 – 2,0 for the bandwidth multiplying factor F. An example of the gammachirp filterbank employing 32 filters and with F = 1.5 is shown in figure 5.

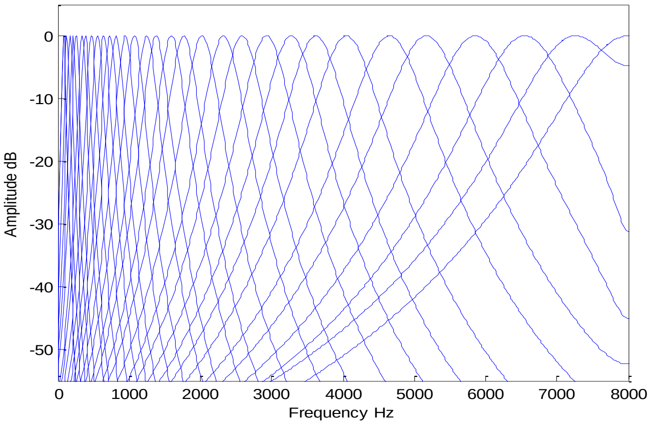

FIGURE 5: A Gammachirp Filterbank with 32 Filters.

4.2 Gammachirp Wavelet Frequency Cepstral Coefficient (GWFCC)

The operating of the new psychoacoustic model is as follows: We segmented the input signal using a Hamming window. The segmented signal is filtered using the non linear external and middle ear model. The output signal of the outer and middle ear model filter is applied to a gammatone filterbank characterized by 32 centers frequencies proposed by the wavelet transform repartition. On each sub-band we calculate the sound pressure level Ps (dB) in order to have the corresponding sub-band chirp term C. Those 32 values of chirp term “c” corresponding to 32 sub-bands of the gammatone filterbank lead to the corresponding gammachirp filterbank. On each sub-band of the dynamic gammachirp filterbank we determine tonal and non tonal components [9]. This step begins with the determination of the local maxima, followed by extracting the tonal components (sinusoidal) and non tonal components (noise) in every bandwidth of a critical band. The selective suppression of tonal and non tonal components of masking is a procedure used to reduce the number of maskers taken into account for the calculation of the global masking threshold. Individual masking threshold takes account of the masking threshold for each remaining component. Lastly, global masking threshold is calculated by the sum of tonal and non tonal components which are deduced from the spectrum to determine finally the signal to mask ratio [12]. After, estimate the logarithm and the cepstrum coefficients using the DCT. In the experiments presented here, a 12 dimensional GWFCC vector is used as the base feature, to which signal log energy is appended, after which velocity and acceleration coefficients (referred to as delta and delta-delta coefficients in the speech community) are calculated for each of the 13 original features, yielding an overall 39 element feature vector for each frame. The complete feature extraction procedure is as shown in figure 6. We note that the addition of delta-cepstral features to the static 13 dimensional GWFCC features strongly improves speech recognition accuracy, and a further (smaller) improvement is provided by the addition of double delta-cepstral features.

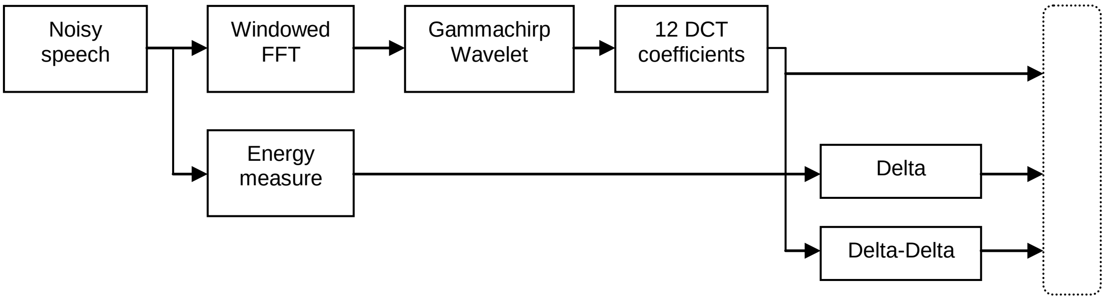

FIGURE 6: Feature extraction with temporal details.

In the next section, we investigate the robustness and compare the performance of the proposed GWFCC features to that of MFCC and GFCC with the different prosodic parameters by artificially introducing different levels of impulsive noise to the speech signal and then computing their correct recognition rate.

5. EXPERIMENTS AND RESULTS

In this section, we investigate the robustness of GWFCC in noise by artificially injecting various types of impulsive noise to the speech signal. We then present speech recognition experiments in noisy recording conditions. The results are obtained using the AURORA databases.

5.1 AURORA Task

AURORA is a noisy speech database, designed to evaluate the performance of speech recognition systems in noisy conditions. The AURORA task has been defined by the European Telecommunications Standards Institute (ETSI) as a cellular industry initiative to standardize a robust feature extraction technique for a distributed speech recognition framework. The initial ETSI task uses the TI-DIGITS database down sampled from the original sampling rate of 20 kHz to 8 kHz and normalized to the same amplitude level [13]. Three different noises (Explosion, door slams and glass breaks) have been artificially added to different portions of the database at signal-to-noise (SNR) ratios ranging from clean, 20dB to 10dB in decreasing steps of 5dB. The training set consists of 8440 different utterances split equally into 20 subsets of 422 utterances each. Each split has one of the three noises added at one of the four SNRs (Clean, 20dB, 15dB and 10dB). The test set consists of 4000 test files divided into four sets of 1000 files each. Each set is corrupted with one of the three noises resulting in a total of (3 x 1000 x 4) 12,000 test utterances. In spite of some drawbacks of the current AURORA task such as the matched test and training conditions, or the absence of natural level variations and variable linear distortions, the AURORA task is of interest since it can demonstrate the potential benefits of using noise robust feature extraction techniques towards improving the recognition performance on a task which (though with matched training and test conditions) has substantial variability due to different types of additive noise at several SNRs.

5.2 Experimental Setup

The analysis of speech signals is operated by using a gammachirp filterbank, in this work we use 32 gammachirp in each filterbank (of 4th order, n = 4), the filterbank is applied on the frequency band of [0 fs/2] Hz (where fs is the sampling frequency), after a pre-emphasis step and a segmentation of the speech signal into frames, and each frame is multiplied by a Hamming windows of 20ms. Generally, gammachirp filterbank and gammachirp wavelet are based on two similar processing blocks: firstly, the speech frame is filtered by the correspondent 4th order gammatone filter, and in the second step we estimate the speech power and calculate the asymmetry parameter c. To evaluate the suggested techniques, we carried out a comparative study with different baseline parameterization technique of MFCC implemented in HTK. The AURORA database is used for comparing the performances of the proposed feature extractor to the MFCC and GFCC features, in the context of speech recognition. For the performance evaluation of our feature extractors, we have used the three noise of the AURORA corpus at four different SNRs (Clean, 20dB, 15dB, 10dB). The features extracted from clean and noisy database have been converted to HTK format using “VoiceBox” toolbox [14] for Matlab. In our experiment, there were 21 HMM models (isolated words) trained using the selected feature GWFCC, GFCC and MFCC. Each model had 5 by 5 states left to right. The features corresponding to each state occupation in an HMM are modeled by a mixture of 12 Gaussians. In the training process, parameters of HMM are estimated during a supervised process using a maximum likelihood approach with Baum-Welch re-estimation. The HTK toolkit was used for training and testing. In all the experiments, 12 vectors with log energy, plus delta and delta-delta coefficients, are used as the baseline feature vector. Jitter and shimmer are added to the baseline feature set both individually and in combination. Table I, II, III and VI shows the overall results. In our experiment, we tested the performance of gammachirp wavelet with additive impulsive noise and prosodic parameter, through recognition of word.

5.3 Results and Discussion

The performance of the suggested parameterization methods GWFCC and GFCC is tested on the AURORA databases using HTK. We use the percentage of word accuracy as a performance evaluation measure for comparing the recognition performances of the feature extractors considered in this paper. %: The percentage rate obtained. Tables I, II and III present the average word accuracy (in %), averaged over all noise scenarios. One Performance measures, the correct recognition rate (CORR) is adopted for comparison. They are defined as:

$$\% \text{ CRR} = \text{no. of correct labels/no. of total labels} * 100\%. \qquad (16)$$

Features	Explosions				Door slams				Glass breaks			
SNR	*Clean*	*20 dB*	*15 dB*	*5 dB*	*Clean*	*20 dB*	*15 dB*	*5 dB*	*Clean*	*20 dB*	*15 dB*	*5 dB*
MFCC (Baseline)	85.45	82.25	78.29	78.44	84.34	80.56	78.98	77.76	87.76	85.43	77.76	76.10
MFCC+Jitter	88.05	84.85	80.89	80.04	86.94	83.16	81.58	80.36	90.36	88.03	80.06	78.70
MFCC+Shimmer	88.45	85.25	81.29	81.44	87.34	83.56	81.98	80.76	90.76	88.43	80.76	79.10
MFCC+Jitter+Shimmer	89.55	86.35	82.39	82.54	88.44	84.66	82.76	81.86	91.86	89.53	81.86	80.20

TABLE 1: Word accuracy (%) of MFCC.

Features	Explosions				Door slams				Glass breaks			
SNR	*Clean*	*20 dB*	*15 dB*	*5 dB*	*Clean*	*20 dB*	*15 dB*	*5 dB*	*Clean∞*	*20 dB*	*15 dB*	*5 dB*
GFCC (Baseline)	89.85	85.23	80.27	80.34	88.24	85.56	81.98	81.76	88.76	87.53	86.76	86.32
GFCC+Jitter	92.45	87.85	82.89	82.94	90.84	88.16	84.58	84.36	91.36	90.13	89.36	88.90
GFCC+Shimmer	92.85	88.23	83.27	83.34	91.24	88.56	84.98	84.76	91.76	90.53	89.76	89.32
GFCC+Jitter+Shimmer	93.95	89.33	84.37	84.44	92.34	89.66	86.06	85.86	92.86	91.63	90.86	90.42

TABLE 2: Word accuracy (%) of GFCC.

Features	Explosions				Door slams				Glass breaks			
SNR	*Clean∞dB*	*20 dB*	*15 dB*	*5 dB*	*Clean∞ dB*	*20 dB*	*15 dB*	*5 dB*	*Clean∞ dB*	*20 dB*	*15 dB*	*5 dB*
GWFCC (Baseline)	92.43	90.17	88.20	85.74	90.35	89.96	88.98	87.70	92.76	91.43	90.86	90.54
GWFCC+Jitter	95.05	92.77	90.80	88.34	92.95	92.56	91.58	90.30	95.36	94.03	93.46	93.14
GWFCC+Shimmer	95.43	93.17	91.20	88.74	93.35	92.96	91.98	91.70	95.76	94.43	93.86	93.54
GWFCC+Jitter+Shimmer	96.53	94.27	92.30	89.84	94.45	94.06	93.08	92.80	96.86	95.53	94.96	94.64

TABLE 3: Word accuracy (%) of GWFCC.

The recognition accuracy for GWFCC, ΔGWFCC and ΔΔGWFCC are obtained and presented in the table VI by different noise. The results are considered for 39 features (GWFCC+ΔGWFCC+ΔΔGWFCC).

Features	Explosions				Door slams				Glass breaks			
SNR	*Clean∞dB*	*20 dB*	*15 dB*	*5 dB*	*Clean∞ dB*	*20 dB*	*15 dB*	*5 dB*	*Clean∞ dB*	*20 dB*	*15 dB*	*5 dB*
GWFCC (13)	83	82.34	80.98	75.74	82.54	80.67	80.54	79.70	85.76	85.47	84.06	83.54
GWFCC+ΔGWFCC (26)	84.23	83	82.54	80.76	90.95	89.56	88.98	85.30	91.56	90.83	90.42	89.86
GWFCC+ΔGWFCC+ ΔΔGWFCC (39)	93.64	91.17	91.20	90.09	94.21	92.87	91.98	90.10	97.76	95.43	92.06	90.87

TABLE 4: Recognition rate (%) of GWFCC, ΔGWFCC and ΔΔGWFCC.

Table I, II and III presents the performance of three voice features in presence of various levels of additive noise. We note that the GWFCC features that are extracted using the gammachirp wavelet exhibit the best CRR. Also, it is observable that the performance of the three features decreases when the SNR decreases too, that is, when the speech signal becoming more noisy. Similarly, the performance of GFCC shows a decrease, but it is a relatively small decrease, whereas the GWFCC features have the overall highest recognition rate throughout all SNR levels. These results assert well the major interest of the gammachirp wavelet and of the auditory filterbank analysis. In additive noise conditions the proposed method provides comparable results to that of the MFCC and GFCC. In convolutive noise conditions, the proposed method provides consistently better word accuracy than all other methods. Jitter and shimmer are added to the baseline feature set both individually and in combination. The absolute accuracy increase is 2.6% and 3.0% after appending jitter and shimmer individually, while there is 4.1% increase when used together. As we can see in the tables, the identification rate increases with speech quality, for higher SNR we have higher identification rate, the gammachirp wavelet based parameters are slightly more efficiencies than standard GFCC for noisy speech (94.27% vs 89.33% for 20 dB of SNR with jitter and shimmer) but the results change the noise of another. We can see the comparison between the two methods parameterization, these GWFCC give better results in generalization and the better performance. The improvement is benefited from using a gammachirp wavelet instead of the auditory filterbank. From the above table VI, it can be seen that the recognition rates are above 90%, this is recognition rates are due to the consideration of using 39 GWFCC features.

From all the experiments, it was concluded that GWFCC has shown best recognition performance compared to other feature extraction techniques because it incorporates gammachirp wavelet features extraction method.

6. CONCLUSION

This paper reviewed the background and theory of the gammachirp auditory filter proposed by Irino and Patterson. The motivation for studying this auditory filter is to improve the signal processing strategies employed by automatic speech recognition systems. In this paper, we concentrated on the implementation of automatic speech recognition in noisy environments. This system uses gammachirp wavelet cepstral features extracted from an audio signal after analysis by gammachirp filterbank. The proposed features (GWFCC) have been shown to be more robust than MFCC and GFCC in noise environments for different SNR values.

Several works have demonstrated that the use of prosodic information helps to improve recognition systems based solely on spectral parameters. Jitter and shimmer features have been evaluated as important features for analysis for speech recognition. Adding jitter and shimmer to baseline spectral and energy features in an HMM-based classification model resulted in increased word accuracy across all experimental conditions. The results gotten after application of this features show that this methods gives acceptable and sometimes better results by comparison at those gotten by other methods of parameterization such MFCC and GFCC.

7. REFERENCES

[1] Miller A., Nicely P. E. (1955). Analyse de confusions perceptives entre consonnes anglaises. J. Acous. Soc. Am, 27, 2, (trad Française,Mouton, 1974 in Melher & Noizet, textes pour une psycholinguistique).

[2] Greenwood, D.D. A cochlear frequency-position function for several species – 29 years later. J.Acous. Soc. Am, Vol. 87, No. 6, Juin 1990.

[3] R.E. Slyh, W.T. Nelson, E.G. Hansen. Analysis of m rate, shimmer, jitter, and F0 contour features across stress and speaking style in the SUSAS database. vol. 4. in Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing, pp. 2091-4, Mar. 1999.

[4] J. O. Smith III, J.S. Abel. Bark and ERB Bilinear Transforms. IEEE Tran. On speech and Audio Processing, Vol. 7, No. 6, November 1999.

[5] T. Irino, R. D. Patterson. A time-domain, Level-dependent auditory filter: The gammachirp. J. Acoust.Soc. Am. 101(1): 412-419, January, 1997.

[6] T. Irino, R. D. Patterson. Temporal asymmetry in the auditory system. J. Acoust. Soc. Am. 99(4): 2316-2331, April, 1997.

[7] T. Irino, M. Unoki. An Analysis Auditory Filterbank Based on an IIR Implementation of the Gammachirp. J. Acoust. SocJapan. 20(6): 397-406, November, 1999.

[8] Irino, T., Patterson R. D. A compressive gammachirp auditory filter for both physiological and psychophysical data. J. Acoust. Soc. Am. Vol. 109, N° 5, Pt. 1, May 2001. pp. 2008-2022.

[9] S. Mallat. A Theory for multiresolution signal decomposition: Wavelet representation. IEEE Trans. Pattern Analysis and Machine Intelligence. Vol. 11. No. 7 pp 674-693 July 1989.

[10] Alex Park. Using the gammachirp filter for auditory analysis of speech. May 14, 2003. 18.327: Wavelets and Filter banks.

[11] Stephan Mallat. Une exploitation des signaux en ondelettes. Les éditions de l'école polytechnique.

[12] H.G. Musmann. Genesis of the MP3 audio coding standard. IEEE Trans. on Consumer Electronics, Vol. 52, pp. 1043 – 1049, Aug. 2006.

[13] H. G. Hirsch, D. Pearce. The AURORA Experiment Framework for the Performance Evaluations of Speech Recognition Systems under Noisy Condition. ISCA ITRW ASR2000 Automatic Speech Recognition: Challenges for the Next Millennium, France, 2000.

[14] M. Brookes. VOICEBOX: Speech Processing Toolbox for MATLAB. Software, available [Mar, 2011] from, www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html.

[15] E. Ambikairajah, J. Epps, L. Lin. Wideband speech and audio coding using gammatone filter banks. Proc. ICASSP'01, Salt Lake City, USA, May 2001, vol.2, pp.773-776.

[16] M. N. Viera, F.R. McInnes, M.A. Jack. Robust F0 and Jitter estimation in the Pathological voices. Proceedings of ICSLP96, Philadelphia, pp.745–748, 1996.

[17] Salhi.L. Design and implementation of the cochlear filter model based on a wavelet transform as part of speech signals analysis. Research Journal of Applied Sciences2 (4): 512-521, 2007□Medwell-Journal 2007.

[18] P. Rajmic, J. Vlach. Real-time Audio Processing Via Segmented wavelet Transform. 10th International Conference on Digital Audio Effect , Bordeaux, France, Sept. 2007.

[19] P.R. Deshmukh. Multi-wavelet Decomposition for Audio Compression. IE (I) Journal –ET, Vol 87, July 2006.

[20] WEBER F., MANGANARO L., PESKIN B. SHRIBERG E. Using prosodic and lexical information for speaker identification. Proc. ICASSP, Orlando, FL, May 2002.

Banking and Modern Payments System Security Analysis

Adam Ali.Zare Hudaib *adamhudaib@gmail.com*
Licensed Penetration Tester EC-Council
Certified Ethical Hacker EC-Council
Certified Security Analyst EC-Council
Wireshark Certified Network Analyst (Wireshark University)
CEH , ECSA , LPT , WCNA
Sweden

Abstract

Cyber-criminals have benefited from on-line banking (OB), regardless of the extensive research on financial cyber-security. To better be prepared for what the future might bring, we try to predict how hacking tools might evolve. We briefly survey the state-of-the-art tools developed by black-hat hackers and conclude that they could be automated dramatically. To demonstrate the feasibility of our predictions and prove that many two-factor authentication schemes can be bypassed, we have analyzed banking and modern payments system security.

In this research we will review different payment protocols and security methods that are being used to run banking systems. We will survey some of the popular systems that are being used today, with a deeper focus on the Chips, cards, NFC, authentication etc. In addition, we will also discuss the weaknesses in the systems that can compromise the customer's trust.

Keywords: Banking Security, Authentication, Chip and PIN, ATM .

1. INTRODUCTION

Cryptology, the science of code and cipher systems, is used by governments, banks and other organisations to keep information secure. It is a complex subject, and its national security overtones may invest it with a certain amount of glamour, but we should never forget that information security is at heart an engineering problem. The hardware and software products which are designed to solve it should in principle be judged in the same way as any other products: by their cost and effectiveness.

However, the practice of cryptology differs from, say, that of aeronautical engineering in a rather striking way: there is almost no public feedback about how cryptographic systems fail.

Most of the development of online financial services has been reactive, doing the minimum amount of work to try and frustrate the attacks which are observed. It has also been quite piecemeal and uncoordinated. Almost all of the defenses have a simple attacker model which only considers those attacks which their prospective target has experienced in the wild. Some of these systems manage to achieve their (fairly limited) goals, but many of them are only partially effective at best [1].

In reaction to the defensive schemes developed by the targets of attacks, many criminals have started to become more sophisticated. This is still lost in the noise of the remarkably successful but simple attacks, which explains why very few people are working on more robust systems. Nevertheless, these new attacks prove that the criminals can adapt to break the defenses which are currently being rolled out.

This thesis is a discussion of the attack and defence landscape surrounding online banking and how these high profile targets and their users can best be protected.

2. BANKING SECURITY

When a bank's system is connected to the internet or intranet, an attack could originate anytime, anywhere. Some essential level of security must be established before business on the internet can be reliably conducted. An attack might be in the form of unauthorized access, destruction, corruption or alteration of data or any type of malicious procedure to cause network failure, reboot or hang. Modern security techniques have made cracking very tedious but not impossible. Furthermore, if the system is not configured properly or the updated patches are not installed then hackers may crack the system using security hole. A wide array of information regarding security hole and their fixes is freely available on the web.

2.1 Banking Security Architecture

In Internet banking as with traditional banking methods, security is a primary concern. The latest methods in Internet banking system security are used to increase and monitor the integrity and security of the systems.

The security of the average Internet banking application is addressed at three levels. The first concern is the security of client information as it is sent from the customer's PC, mobile phones, corporate clients etc. to the Web server. The second area concerns the security of the environment in which the Internet banking server and client information database reside. Finally, security measures are in place to prevent unauthorized users from attempting to log into the online banking section of the Web systems [2].

Data security between the client browser and Web server usually is handled through a security protocol called Secure Sockets Layer (SSL). SSL provides data encryption [3], server authentication, and message integrity for a Internet connection. In addition, SSL provides a security "handshake" that is used to initiate the connection. This handshake results in the client and server agreeing on the level of security they will use and fulfills any authentication requirements for the connection.

Also online banking application supports data encryption. Requests for online banking information are passed on from the Web server to the Internet banking server. The Internet banking application is designed using a three-tiered architecture. The three-tiered architecture provides a double firewall, completely isolating the Web server from the client information SQL database.

The World Wide Web interface receives SSL input and sends requests through a firewall over a dedicated private network to the Internet banking server. The World Wide Web interface is the only process capable of communicating through the firewall to the Internet banking server. Therefore, only authenticated requests communicate with the Internet banking server.

The client information database is housed on a database server, which implements security algorithm in addition to the firewall technology. The client database is usually stored on a RAID-5 drive array, which provides uninterruptible data access, even in the event of a hard drive failure [4].

A security analyzer constantly monitors login attempts and recognizes failures that could indicate a possible unauthorized attempt to log into an account. When such trends are observed, steps will be taken automatically to prevent that account from being used.

Implementation of the SSL security protocol on the Web server and client browser ensures authenticated data has been received from the client. The three-tiered approach of the Internet banking application creates a double firewall which performs information requests over dedicated networks designed to handle specific functions. Placing all business logic and event logging within the Internet banking server creates a controlled environment which allows quick incorporation of Internet security technologies as they evolve. Finally, the security analyzer monitors login attempts in order to prevent unauthorized logins.

Example of banking security architecture is shown on figure 1.

The Open Payment Framework is built entirely on a Service Oriented Architecture (SOA) delivering common, reusable services consisting of a comprehensive data model, choreographed payment business processes and configurable services including parsing, validation, cost based routing, warehousing security, auditing and many more [5].

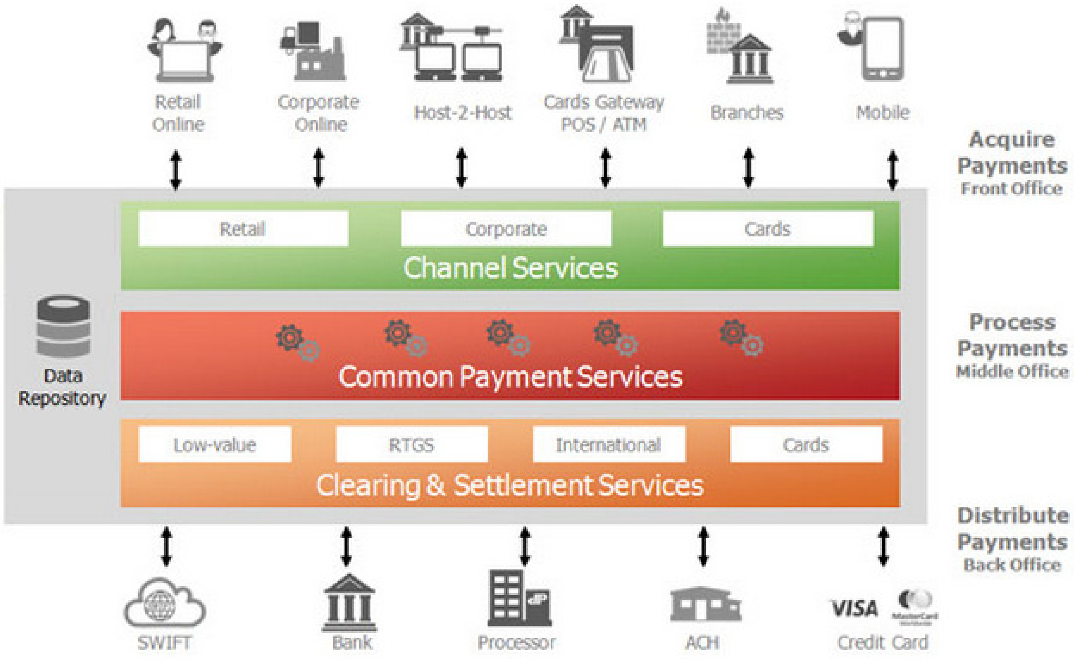

FIGURE 1: Banking Security Architecture.

2.2 Banking Security Attacks and Defense

Notwithstanding an increased number of attacks, the percentage of financial malware detected each month is dropping. The reasons for this are detailed below:

- Malware authors constantly change their programs in order to evade detection by antivirus solutions. However, if the changes made are minor, AV vendors will still be able to detect new malware samples using signatures created for previous variants.
- Banking attacks are usually a multi-step process: social engineering, phishing, and the use of Trojan-Downloaders which then download the financial malware. It's easier for the criminals to modify the Trojan-Downloader programs (which are usually smaller in size, and generally less complex) than the financial malware itself.

Banks have responded to the increased number of attacks by investing more time, money and effort into developing mechanisms for detecting fraud and illegal activity. One safeguard is for an alert to be triggered if a large amount of money is transferred to a 'suspicious' region of the world [6].

In order to sidestep this, cyber criminals have taken to using 'money mules'. Mules are often recruited via seemingly legitimate job offers – for instance, the cyber criminals might advertise for a 'financial manager'. Such services are used because they guarantee anonymity, reducing the likelihood that the cyber criminal will be caught. The remaining funds are the mule's 'commission' – naturally money which has been earned illegally via phishing or financial malware.

When looking at the question of phishing, it's important to have a clear definition of it. This article defines phishing as spoofed messages which allegedly come from a (financial) organization and which are designed to trick the user into giving up confidential information. This is strictly a matter of social engineering, and once malware is involved, the attack can no longer be considered phishing.

Given that phishing continues to be widespread, it is obviously a successful method of attack. Phishing attacks work on all major operating systems. However, there's one major downside from the cyber criminal's point of view: the user has the choice whether or not to click on a link contained in an email, and is then able to choose whether or not to enter his/ her credentials.

This element of choice is inherent in all social engineering approaches. A technical approach involving the use of malware removes this element of choice, making those users who didn't fall for a phishing scam are still a viable target.

Financial malware comes in all shapes and sizes, and will often be tailored to target a single organization. There's no requirement for the cyber criminals to spend time creating unnecessarily complex malware [7]. There are several methods which malware authors can use to get around banking security and harvest user information. For instance, if a bank uses single-factor authentication with a static username and passwords, it's a simple matter of capturing keystrokes. Alternatively, some banks have created dynamic keypads so that the user needs to click a 'random' pattern in order to enter his password. Malware authors use two different methods to circumvent this type of security - they can either create screen dumps when the user visits a specific site or simply gather the information being sent to the site by grabbing the form. In both cases, the stolen data is processed later.

The use of Transaction Authorisation Numbers (TAN) for signing transactions makes gaining access to accounts somewhat more complex. The TAN may come from a physical list issued to the account holder by the financial organisation or it may be sent via SMS. In either case, the cyber criminal does not have access to the TAN [8]. In most cases, malware used will capture the information entered by the user in a way similar to that described above. Once the user enters the TAN, the malware will intercept this information and either display a fake error message, or send an incorrect TAN to the financial site. This may result in the user entering another TAN. An organization may require two TANS to complete a transaction – this depends on the organization and the security systems it has decided to implement. If only one TAN is required to make a transaction, the attack describe above could allow a cyber criminal to make two transactions.

Another method used by cyber criminals is to redirect traffic. Additionally, although the traffic is redirected, it may not be processed in real time, which gives the victim the chance to contact his/ her bank to stop the transaction.

More sophisticated malware will use a MitM attack; this not only enables cyber criminals to attack more banks, but also ensures a higher return, as data is processed in real time. A MitM attack uses a malicious server to intercept all traffic between the client and the server i.e. the customer and the financial organization. Although everything will seem normal to the user, when s/he is asked to authorize a transaction, s/he is actually authorizing a transaction created by the cyber criminal. Malware which uses a MiTM attack typically either hides browser notifications about false web site certificates or, more commonly, shows a fake notification. However, depending on the approach used by the malware, it may do neither of these things, simply because it isn't necessary. A lot of the more sophisticated financial malware which uses MitM attacks also makes use of HTML injection [9].

However, there's a clear trend: the increased usage of two-factor authentication by financial organizations has resulted in an increase in malware capable of defeating this type of authentication. This means that the eventual adoption of two-factor authentication will not have any significant long-term effect. It will simply raise the benchmark for financial malware.

Nonetheless, there is a fundamental problem with two-factor authentication, namely that though the session may be secure, whatever happens during that session goes unchecked. In order to increase security, some additional form of communication, such as the use of a cryptographic token or SMS messages (already implemented by some financial institutions) is required. SMS messages could set limits on the lifetime of the TAN, the account numbers being accessed and the maximum permissible transaction amount.

2.3 Internet Banking Authentication and Attacks

The most recent internet banking security threats are listed below:

- Phishing
- Spyware and Adware
- Viruses
- Trojans
- Keyloggers

The attack tree has one root node, representing the final target of the attacker, which is the compromise of the user's bank account. An intruder may use one of the leaf nodes as a means for reaching the target. To categorize Internet banking attacks, each component of the process should be examined: the user terminal/user (UT/U), the communication channel (CC) and the Internet banking server (IBS). The following types of attacks are identified [10]:

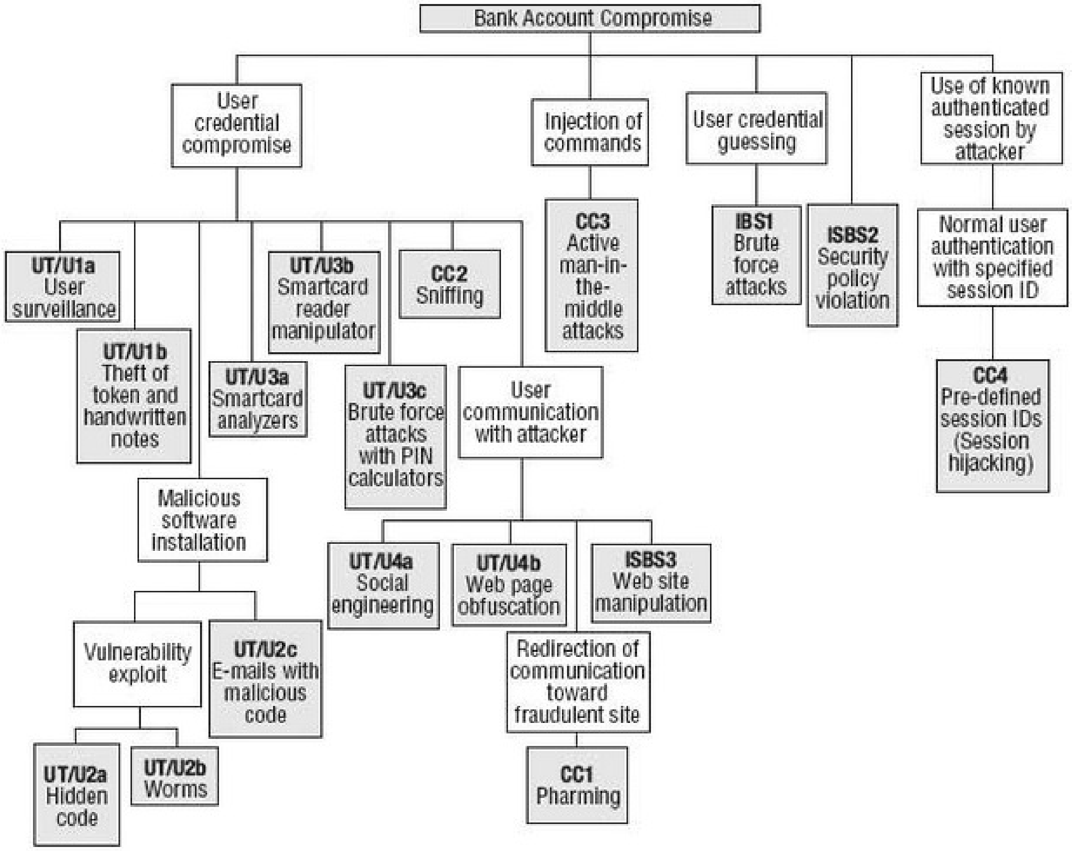
Bank Account Compromise
User credential compromise
Injection of commands
User credential guessing
Use of known authenticated session by attacker
UT/U1a User surveillance
UT/U1b Theft of token and handwritten notes
UT/U3a Smartcard analyzers
UT/U3b Smartcard reader manipulator
CC2 Sniffing
UT/U3c Brute force attacks with PIN calculators
User communication with attacker
CC3 Active man-in-the-middle attacks
IBS1 Brute force attacks
ISBS2 Security policy violation
Normal user authentication with specified session ID
CC4 Pre-defined session IDs (Session hijacking)
Malicious software installation
Vulnerability exploit
UT/U2c E-mails with malicious code
UT/U2a Hidden code
UT/U2b Worms
UT/U4a Social engineering
UT/U4b Web page obfuscation
ISBS3 Web site manipulation
Redirection of communication toward fraudulent site
CC1 Pharming

FIGURE 2: Attacks and Threats Models.

Phishing is a scam where fraudsters ‘fish’ for your personal details by using hoax emails claiming to be from financial institutions. This method continues to be favored by online thieves.

Hoax emails claiming to be from banks are often generated overseas, and are sent in bulk asking recipient to provide sensitive information such as their username, password, Customer Registration Number or Debit Cards / Credit Cards numbers and PINs by providing a link leading to a fake website, enabling thieves to gather the details for later fraudulent use.

You can minimize your chances of being a victim of Phishing scams by:

- Typing actual web-site address into your Internet browser to log on to Internet Banking.
- Treating all emails requesting personal log on information such as username, password or PIN with extreme caution.
- Authentic BankMuscat emails will not request personal details or log on information.
- Immediately deleting emails of unknown origins, no matter how innocent or provocative the subject headings sound.
- Changing your Internet Banking password on a regular basis.
- If you receive an email requesting you to register or enter sensitive details, do not respond and click on any hyperlink. Immediately forward the email to bank [11].

Spyware and Adware

Spyware is a type of software that secretively collects user information while on the Internet. Adware is a type of spyware used by marketers to track Internet user’s habits and interests for the purpose of customizing future advertising material [12]. The information is then used to customize future advertisements directed to the user, or can be sold to a third party for the same purpose.

You can minimize your chances of unintentionally downloading spyware onto your computer, devices.

Viruses

A computer virus is software that affixes itself to another program like a spreadsheet or word document. While active, the virus attempts to reproduce and attach itself to other programs. This can tie up resources such as disk space and memory, causing problems on any home computer. An email virus is the latest type of computer virus that is transported through email messages and usually replicates by automatically distributing itself out to all contacts on the victims email address book.

You can increase your chances of ensuring your computer is free from viruses by:

- Installing anti-virus software, and keeping it updated with the latest virus definitions.
- Downloading and installing security patches for your operating system as soon as they become available.
- Not accepting attachments from emails of unknown sources.
- Installing software from trusted sources only.

Trojans

A Trojan is a destructive program that poses as a harmless application. Unlike viruses, Trojans do no replicate themselves and do not need a host program to attach to. Some Trojans will claim to rid the computer of viruses or other harmful applications, but instead introduce viruses and leave it vulnerable to attacks by hackers and intruders. You can minimize your chances of unintentionally downloading Trojans [130].

Keyloggers

If fraudster installs a software called "keylogger" on the computer or the device on which the customer is accessing Online Banking, the software copies to a file, every keystroke typed on that pc. This sensitive information gets captured that the fraudster can later use for fraudulent purposes and illegitimate access to your account.

There are ways to prevent this from happening.

– You should not use computers to access accounts which are not trusted (like don't use cybercafe, or other people's computers for accessing Online Banking).
– Keep antivirus software updated every day to protect your system and ensure that your system is virus free.

IBS attacks
These types of attacks are offline attacks against the servers that host the Internet banking application. Examples include:

– IBS1: Brute-force attacks in certain password-based mechanisms are reported to be feasible by sending random usernames and passwords. The attacked mechanisms implement a scheme based on guessable usernames and four-digit passwords. The attack mechanism is based on distributed zombie personal computers, hosting automated programs for username- or password-based calculation. This attack may be combined with username filtering methods for determining the identity of the user. These methods filter the different responses of the server, in the case of valid or invalid usernames.

– IBS2: Bank security policy violation—Violating the bank's security policy in combination with weak access control and logging mechanisms, an employee may cause an internal security incident and expose a customer's account.

IBS3: Web site manipulation—Exploiting the vulnerabilities of the Internet banking web server may permit the alteration of its contents, such as the links to the Internet banking login page. This may redirect the user to a fraudulent web site where his/her credentials may be captured [14].

3. PAYMENTS SECURITY

3.1 Banking Security Architecture

Authentication attack can be resisted by cryptographically binding the one-time code to the data of the transaction being attempted – transaction authentication. A robust way to do this is to provide the customer with an electronic signature device with a trustworthy display on which she could verify the transaction data, a trusted path to authorise a digital signature, and a tamper-resistant store for the signing key. Such devices were foreseen by the EU Electronic Signature Directive which provided for signatures thus created to be admissible as evidence in legal proceedings. However such devices typically cost $100 or more.

The Chip Authentication Programme (CAP) [15]is a lower-cost implementation of this general approach. Individual countries have adopted different variants of CAP based on the original specification. Usually it uses the deployed “Chip & PIN” smart card infrastructure. Participating banks have sent out handheld smart card readers with keypads and displays which, with a customer’s card and PIN, generate one-time passwords. Even though Chip & PIN is based on the public EMV standard, the CAP standard is secret and so not subject to scrutiny, despite being a critical security component the public must rely on for banking transactions.

CAP operates in three modes – identify, respond, and sign. These differ in the information a user is asked to enter before a response code is generated. For all three modes a PIN is required first. Thereafter, identify just returns a onetime code; for respond a numerical challenge is required; and for sign an account number and a value are needed. The numerical response code is a compressed version of a MAC computed by the card under its key; it is calculated over the information entered by the customer, a transaction counter, and a flag showing whether the PIN matches the one stored on the card [16].

The implementation of the CAP system is heavily based on the EMV smart card protocol being introduced throughout Europe for credit and debit card point-of-sale transactions. In the UK, EMV is known under the “Chip & PIN” brand. Using EMV as the basis for CAP reduced development and deployment costs; using the existing debit card base meant that the CAP devices themselves did not need to be personalized.

The reader requests a list of all the data records stored by a card. These form a hierarchy, with each node being prefixed by a one or two byte tag. In a standard EMV transaction, these would include account number, public key certificates, signatures, and so on. With CAP, only three entries are of interest – the card data object lists (CDOL1 and CDOL2), identified by tags 0x8C and 0x8D respectively, and the CAP bit filter2, identified by the tag 0x9F56. Tag 0x9F55 is also present on cards, with value 0xA0, but its purpose is unclear.

PIN verification. Once the reader has successfully read all available records, it prompts the customer for a 4-digit PIN. This is sent to the card as the payload to the EMV standard VERIFY command. If three consecutive PIN verifications fail, the card will lock itself until taken to an ATM and reset with the correct PIN. While the EMV standard allows for a transaction to continue if the PIN verification fails or is omitted, the CAP reader requires that the card accept the PIN before continuing [17].

Cryptogram generation. Next, the reader requests an application cryptogram from the card, using the GENERATE AC command. The reader first requests an Authorization Request Cryptogram (ARQC), indicating that it wishes to perform an online EMV transaction. The card then responds with an ARQC, indicating that the card is willing to do so. If this was an EMV transaction, the reader would send the ARQC to the bank for verification, but it cannot do so because it is offline. So the reader then requests an Application Authentication Cryptogram (AAC), indicating that it wishes to cancel the transaction.

A similar transaction flow might be seen during a point-of-sale transaction if a shop is only willing to accept online transactions but fails to connect to the bank (e.g. if the phone line is engaged). This protocol may have been designed so that CAP maintains maximum compatibility with EMV smart card applications. While EMV supports offline transactions by requesting a Transaction Certificate (TC) instead of an ARQC, some card risk-management algorithms may lock up if there are too many consecutive offline attempts. Cancelling the transaction should reset the smart card’s risk-management parameters [18].

Reader response formatting. The response to a GENERATE AC call includes a 16- bit application transaction counter (ATC), a Cryptogram Identification Data (CID) type code, Issuer Application Data (IAD) which includes the result of the PIN verification, and an Application Cryptogram (AC) which is a MAC over all this data. The MAC method used to calculate the cryptogram, and the structure of the IAD, are not specified by the EMV standard, as they are proprietary to the card issuer [19].

The basic principle behind CAP – a trusted user interface and secure cryptographic microprocessor – is sound. However the system has been optimized literally to death. Re-using ATM cards for point of sale and CAP saved money but created a vulnerability to relay attack, and increased the risk of violent mugging and murder. Omitting a server-provided nonce removed assurance that responses are freshly generated. Overloading fields introduce a social engineering vulnerability, as it makes the system model too complex for the average user to be expected to visualize.

3.2 Chip and Skim Cloning EMV Cards with Pre Replay Attack

EMV is now the leading scheme worldwide for debit and credit card payments, as well as for cash withdrawals at ATMs, with more than 1.34 billion cards in use worldwide. US banks were late adopters, but are now in starting to issue EMV cards to their customers. EMV cards contain a smart card chip, and are more difficult to clone than the magnetic-strip cards that preceded them.

EMV was rolled out in Europe over the last ten years, with the UK being one of the early adopters. After it was deployed, the banks started to be more aggressive towards customers who complained of fraud, and a cycle established itself. Victims would be denied compensation; they would Google for technical information on card fraud, and find one or other of the academic groups with research papers on the subject; the researchers would look into their case history; and quite often a new vulnerability would be discovered [20].

We wondered whether, if the “unpredictable number" generated by an ATM is in fact predictable, this might create the opportunity for an attack in which a criminal with temporary access to a card can compute the authorization codes needed to draw cash from that ATM at some time in the future for which the value of the UN can be predicted. We term this scenario the “pre-play" attack. We discovered that several ATMs generate poor random numbers, and that attacks are indeed possible.

EMV did not cut fraud as its proponents predicted. While using counterfeit and stolen cards did become more difficult, criminals adapted in two ways. First, they moved to “card-not-present" transactions (Internet, mail-order, and phone-based payments) which remained beyond the scope of EMV. Second, they started making magnetic-strip clones of EMV cards. There had always been some ATM “skimming" where crooks put devices on ATM throats to capture card data and record PINs; and now that PINs were demanded everywhere and not just at ATMs, the opportunities for skimming increased hugely. The simultaneous deployment of EMV with magnetic strip meant that fallback and backwards-compatibility features in EMV could be exploited; for several years, all ATMs would still accept mag-strip cards, and even once this started to be phased out in the UK for locally-issued cards, it was still possible to use mag-strip clones of UK cards in ATMs in the USA. This is why, soon after the completion of the UK EMV roll-out in 2005, counterfeit fraud went up. Instead of entering PINs only at ATMs, customers were now entering their PIN in POS terminals, which are much easier to tamper with [21].

Total fraud levels were brought down following 2008 through improvements to back-end fraud detection mechanisms which reject suspicious transactions; by more aggressive tactics towards customers who dispute transactions; and by reducing the number of UK ATMs that accept “fallback" magnetic-strip transactions on EMV-issued cards [22]. Fallback fraud is now hard enough to push the criminal community to more sophisticated smart-card-based attacks.

Prior research showed that it was possible to use a stolen EMV card in a POS device without knowing the PIN. Given a suitable man-in-the-middle device, a crook can trick the terminal into believing that the right PIN was entered, while the card thought it was authorizing a chip-and-signature transaction; criminals have now gone on trial in France for exploiting this “no pin" vulnerability.

The specifications and conformance testing procedures simply require that four consecutive transactions performed by the terminal should have unique unpredictable numbers. Thus a rational implementer who does not have the time to think through the consequences will probably prefer to use a counter rather than a cryptographic random number generator (RNG); the latter would have a higher probability of failing conformance testing (because of the birthday paradox) [23].

Even if the UN generation algorithms are patched, a number of powerful attack variants may make pre-play attacks viable for years to come.

Malware. There are already numerous cases of malware-infected ATMs operating in Eastern Europe and depending on the internal architecture of the ATM it may be easy for such malware to sabotage the choice of UN. In fact one bank suggested to us that the ATM that kicked of this whole research project may have been infected with malware.

Supply chain attacks. Such attacks have already been seen against POS terminals in the wild, and used to harvest magnetic strip data. So it is feasible that a criminal (or even a state-level adversary) might sabotage the RNG deliberately, either to act predictably all the time, or to enter a predictable mode when triggered via a covert channel. A suitably sabotaged RNG would probably only be detected via reverse engineering or observation of real world attacks.

Collusive merchant. A merchant might maliciously modify their EMV stack to be vulnerable, or inject replayed card data into the authorization/settlement system. He could take a cut from crooks who come to use cloned cards at their store, or just pre-play transactions directly. In the UK, there was a string of card cloning attacks on petrol stations where a gang bribed store managers to look the other way when PIN pads were tampered with and monitoring devices inserted into network connections; exactly what you need to deploy a pre-play attack. Terminal cut-out. A variant is the terminal cut-out or bypass is where the transaction stream between the merchant terminal and the acquirer is hacked to misreport the unpredictable number when triggered by a particular signal (e.g. a particular account number or a known ARQC). This transaction data stream is not normally considered sensitive within the threat model and can be altered at will by merchant software. The attackers' card performing the replay can then use any UN for which it has an ARQC, and the true random UN made up by the terminal will never see the light of day. This is hard to block: there is no provision in currently deployed EMV cards for the terminal to confirm that its choice of UN was correctly included in the cryptographic MAC. The terminal cut-out could be implemented in malware (and there's evidence of bank botnets looking for POS devices), or in a merchant's back-end system (we have evidence of merchants already tampering with transaction data to represent transactions as PIN-verified when they were not, so as to shift liability) [24].

UN modification in the network. A man-in-the-middle device between a POS device and the acquiring bank, perhaps at a network switch, would also be a good way to deploy such an attack. This could be an attractive way to attack merchants that process high- value transactions, such as jewelers or investment firms, who might guard their premises and take care of their POS equipment yet still fall to a targeted attack. A pre-play attack would be much harder to detect than old-fashioned attacks that just convert deny authorization messages into approve messages.

3.3 Chip Secrets

There are chip attack methods:

Non-invasive attacks observe or manipulate with the chip without any physical harm to it; low-cost: require relatively simple equipment and basic knowledge; time consuming and not always successful. AES is attacked by side-channel attacks such as SPA, DPA, CPA, EMA, DEMA (takes 1 second/1 day); poor signal-to-noise ratio of about –15dB due to low-power operation and multiple sources of noise (clocks, pumps, acquisition).

Invasive attacks almost unlimited capabilities in extracting information and understanding chip functionality; expensive, requires a very sophisticated equipment and knowledge; less time consuming and straightforward for many devices. AES is attacked by partial reverse engineering followed by microprobing (takes 1 day).

Semi-invasive attacks fill the gap between non-invasive and invasive types: direct access to the chip's surface is required but without any physical harm to it; moderate cost: some equipment can be easily built; higher success rate compared to non-invasive attacks; some are easily repeatable and relatively quick to set up. AES is attacked by optical fault injection attack (1 hour) and optical emission analysis (1 week/1 hour).

Ways to improve security:

- turn some ROM areas into reprogrammable Flash areas;
- reprogram low-level features;

- access shadow areas;
- access hidden JTAG registers;
- find the JTAG registers responsible for controlling read sense;
- amplifiers, such that VREF can be adjusted [25].

Bumping attacks are dangerous and can compromise the security in chips – evaluation and protection is necessary. Backside approach helps in modern chips, it is simple to do and does not require expensive optics and precise positioning. Bumping attacks can be used for partial reverse engineering to understand internal data paths and chip structure. The hardware security protection in Actel ProASIC3 FPGAs is under serious threat due to unforeseen problems in the corporate security strategy of the management team. Access path to shadow hardware features brings capability of making ProASIC3 chips more robust and serve security critical applications for the next few years. Embedded memory is more secure than encrypted external memory storage, and encrypted bitstream is even less secure.

3.4 Modern Payments Security: EMV, NFC etc

The total number of purchases on all major worldwide card issuers (American Express, Diners Club, JCB, MasterCard, UnionPay and Visa) increased to a total of 135.33 billion, up 12.1 percent from 2010 on an additional 14.56 billion transactions, the Nilson Report, 2011 report said.

Some statistics:
As of early 2011, 1.2 billion EMV cards were deployed across the globe along with 18.7 million EMV terminals (via IBID). Over a billion smartphones sold by 2012. By 2014, 44% of smartphones will be NFC-compatible (via). Payment card users in Russia: Spring 2011 to Spring 2012: from 49% to 56% (via GfK Rus).

There are Notable IPS (International Payment Systems):

- Visa;
- MasterCard (MC);
- Japan Credit Bureau (JCB);
- Diners Club (DC);
- American Express (AMEX);
- China Union Pay (CUP).

Usually they have security methods: plastic (holograms, watermarks) and cryptography (DES, 3DES, mode: EDE, 2 keys: ABA, cardholder authentication, card authentication, encryption) [26]. Processing cycle begins with cardholder. He receives a card and sign it manually, opens PIN envelope, reads it and burn it. Then issuer (personalization, embossing, encoding, authorization processing, presentment processing). The card is just a static read-only piece of plastic. The acquirer manages terminals and provides services to merchants. Acquirer’s host software provides authorization and presentment processing. The terminal reads card and talks to acquirer’s host.

Transaction phases are shown below:

- Authorization
  Terminal reads card. If cardholder enters PIN, terminal calculates a PIN Block inside PED and PIN Block is encrypted under corresponding TPK. Auth message is sent to Acquirer’s host. Acquirer processes it and sends to IPS. IPS processes it and sends to Issuer. Issuer approves or rejects it and sends the answer back.

- Clearing
  Consists of: terminal reconciliation; acquirer demands satisfaction from the issuer and sends the clearing presentments through the IPS; IPS processes them and sends them to

the Issuer; issuer may not respond, money transfer is automatically performed at the next stage.

– Settlement;
All parties settle their financial positions through the IPS (consolidated funds transfer).

– Dispute resolution.
Terminals usually talk to acquirer’s host in their special protocols: ATM, POS, SSD. But some are built over ISO8583.

NFC system uses devices: tags, smart cards, readers, mobile devices. Ans secures them by NFC Ready and NFC Secure; secure element; authentication; encryption.

The secure element (SE) is a secure microprocessor (a smart card chip) that includes a cryptographic processor to facilitate transaction authentication and security, and provide secure memory for storing payment applications (e.g., American Express, Discover, MasterCard, Visa and other payment applications). SEs can also support other types of secure transactions, such as transit payment and ticketing, building access, or secure identification.

3.5 Verified by Visa and MasterCard Secure Code

Banks worldwide are starting to authenticate online card transactions using the `3-D Secure' protocol, which is branded as Verified by Visa and MasterCard Secure Code. This has been partly driven by the sharp increase in online fraud that followed the deployment of EMV smart cards for cardholder-present payments in Europe and else- where. 3-D Secure has so far escaped academic scrutiny; yet it might be a textbook example of how not to design an authentication protocol. It ignores good design principles and has significant vulnerabilities, some of which are already being exploited.

The primary purpose of 3DS is to allow a merchant to establish whether a customer controls a particular card number. It is essentially a single-sign on system, operated by Visa and MasterCard, and it differs in two main ways from existing schemes such as OpenID or InfoCard. First, its use is encouraged by contractual terms on liability: merchants who adopt 3DS have reduced liability for disputed transactions. Previous single sign-on schemes lacked liability agreements, which hampered their take-up. Few organizations are willing to trust a third-party service provider to authenticate users when they have no recourse in the event of error or attack. (In any case, security economics teaches that you're unlikely to get a secure system if Alice guards it while Bob pays the cost of failure.) Second, in other respects 3DS does not adopt the lessons learned from single-sign on, and breaks many established security rules [27].

Before 3DS can be used to authenticate transactions, cardholders must register a password with their bank. A reasonably secure method would be to send a password to the customer's registered address, but to save money the typical bank merely solicits a password online the first time the customer shops online with a 3DS enabled card known as activation during shopping (ADS). To confirm that the customer is the authorized cardholder, the ADS form may ask for some weak authenticators (e.g. date of birth), although not all banks do even this. From the customer's perspective, an online shopping website is asking for personal details.

The 3DS specification only covers the communication between the merchant, issuer, acquirer and payment scheme, not how customer verification is performed. This is left to the issuer, and some have made extremely unwise choices. For instance, one bank asks for the cardholder's ATM PIN. It's bad enough that EMV has trained cardholders to enter ATM PINs at terminals in shops; training them to enter PINs at random e-commerce sites is just grossly negligent.

Another issuer-specific choice is how to reset the password when a customer forgets it; here again corners are cut. Some banks respond to one or two failed password attempts by prompting an online password reset using essentially the same mechanisms as ADS. In a number of cases, the bank requires only the cardholder's date of birth, which is easily available from public records; with one (UK-government-owned) bank, two wrong password attempts simply lead to an invitation to set a new password [28].

A third variable factor is whether the 3DS implementation asks for a whole password or for some subset of its letters. The idea behind asking for a subset is that a single-round keyboard logging attack does not compromise the whole password. However, this compels users to select relatively simple passwords, and probably to write them down.

3.6 Credit Card Duplication and Crime Prevention Using Biometrics

Until the introduction of Chip and PIN, all face-to-face credit or debit card transactions used a magnetic stripe or mechanical imprint to read and record account data, and a signature for verification. Under this system, the customer hands their card to the clerk at the point of sale, who either "swipes" the card through a magnetic reader or makes an imprint from the raised text of the card. In the former case, the account details are verified and a slip for the customer to sign is printed [29]. In the case of a mechanical imprint, the transaction details are filled in and the customer signs the imprinted slip. In either case, the clerk verifies that the signature matches that on the back of the card to authenticate the transaction. This system has proved to be ineffective, because it has a number of security flaws, including the ability to steal a card in the post, or to learn to forge the signature on the card. More recently, technology has become available on the black market for both reading and writing the magnetic stripes, allowing cards to be easily cloned and used without the owner's knowledge. Fingerprints are one of many techniques used to identify individuals and verify their identify. Matching algorithms used to compare previously stored templates of fingerprints against candidate fingerprints for authentication purposes. Pattern based algorithms compare the basic fingerprint patterns (arch, whole, and loop) between a previously stored template and a candidate fingerprint. The candidate fingerprint image is graphically compared with the template to determine the degree to which they match. The major disadvantage here is that Finger print authentication cannot be successful if the user has a band aid on his finger. Another disadvantage is fingerprint remains the same even if the person is unconscious or dead. This leads to unauthorized use of a person’s fingerprint without his consent. To overcome the limitations of the existing authentication systems of the usage of credit cards, there was proposed a new system of authentication in which authentication is done through two phases. The first phase is verifying the identity of the user using iris recognition and the second phase is the authentication using palm vein technology [30].

Initially the user will be asked to insert his card. The database is checked to verify if such an account exists. If exists, the user will be authenticated using iris recognition. If the user is authenticated in this phase, he will then be asked to stretch out his palm for the vein pattern authentication. This is compared with the stored pattern and if it matches the user is, authenticated.

Users today mostly use textual passwords that follow an encryption algorithm. Mostly textual passwords, nowadays, are kept very simple say a word from the dictionary or their pet names, girlfriends etc.

A typical iris recognition system in involves three main modules:

– Image acquisition is to capture a sequence of iris images from the subject using a specifically designed sensor.

– Preprocessing Stage includes determining the boundary of the iris within the eye image, and extracts the iris portion from the image to facilitate its processing. It includes various stages such as: iris segmentation, iris normalization, image enhancement.

– Feature extraction and encoding is the most key component of an iris recognition system and determines the system’s performance to a large extent. Iris recognition produces the correct result by extracting features of the input images and matching these features with known patterns in the feature database.

Users today mostly use textual passwords that follow an encryption algorithm. Mostly textual passwords, nowadays, are kept very simple say a word from the dictionary or their pet names, girlfriends etc. Years back Klein performed such tests and he could crack 10-15 passwords per day. Now with the technology change, fast processors and many tools on the Internet this has become a Child's Play. Therefore, we use Biometrics in our authentication, which is more customizable and very interesting way of authentication. The vein matching, also called vascular technology is a technique of biometric identification through the analysis of the patterns of blood vessels visible from the surface of the skin. An individual first rests his wrist, on some devices, such that the palm is held centimeters above the device's scanner, which flashes a near- infrared ray on the palm [31].

Unlike the skin, through which near-infrared light passes, deoxygenated hemoglobin in the blood flowing through the veins absorbs near-infrared rays, illuminating the hemoglobin, causing it to be visible to the scanner. Arteries and capillaries, whose blood contains oxygenated hemoglobin, which does not absorb near- infrared light, are invisible to the sensor. The still image captured by the camera, which photographs in the near- infrared range, appears as a black network, reflecting the palm's vein pattern against the lighter background of the palm.

3.7 Security Vulnerabilities of Chip and PIN

Chip and PIN is the brand name adopted by the banking industries in the United Kingdom and Ireland for the rollout of the EMV smart card payment system for credit, debit and ATM cards. The word "chip" refers to a computer chip embedded in the smartcard; the word PIN refers to a personal identification number that must be supplied by the customer. "Chip and PIN" is also used in a generic sense to mean any EMV smart card technology which relies on an embedded chip and a PIN.

The Chip and PIN implementation was criticized as designed to reduce the liability of banks in cases of claimed card fraud by requiring the customer to prove that they had acted "with reasonable care" to protect their PIN and card, rather than on the bank having to prove that the signature matched. Before Chip and PIN, if a customer's signature was forged, the banks were legally liable and had to reimburse the customer. Until 1 November 2009 there was no such law protecting consumers from fraudulent use of their Chip and PIN transactions, only the voluntary Banking Code. While this code stated that the burden of proof is on the bank to prove negligence or fraud rather than the cardholder having to prove innocence, there were many reports that banks refused to reimburse victims of fraudulent card use, claiming that their systems could not fail under the circumstances reported, despite several documented successful large-scale attacks [32].

Chip and PIN cards are not foolproof; several vulnerabilities have been found and demonstrated, and there have been large-scale instances of fraudulent exploitation. In many cases banks have been reluctant to accept that their systems could be at fault and have refused to refund victims of what is arguably fraud, although legislation introduced in November 2009 has improved victims' rights and put the onus on the banks to prove negligence or fraud by the cardholder. Vulnerabilities and fraud are discussed in depth in the main article.

3.8 Synthetic PIN For Authentication and Authorization

There is a new technology that is used for authenticating users and authorising transactions - the Synthetic PIN. Also, the Synthetic PIN solution may be used as an addition for existing security mechanism that service provider has. This solution consists of four components that are shown on the figure 3: the user’s computer and his phone, and a service provider (for example, a bank) and the SyntPIN server. In addition, the components are connected via networks or sound, drawn as grey cloud shapes in the case of the Internet and the telecom network, and drawn as sound waves in the case of sound played over the computer's loudspeaker. A network communications channel is shown as a line, with arrowheads showing directions of communication [33].

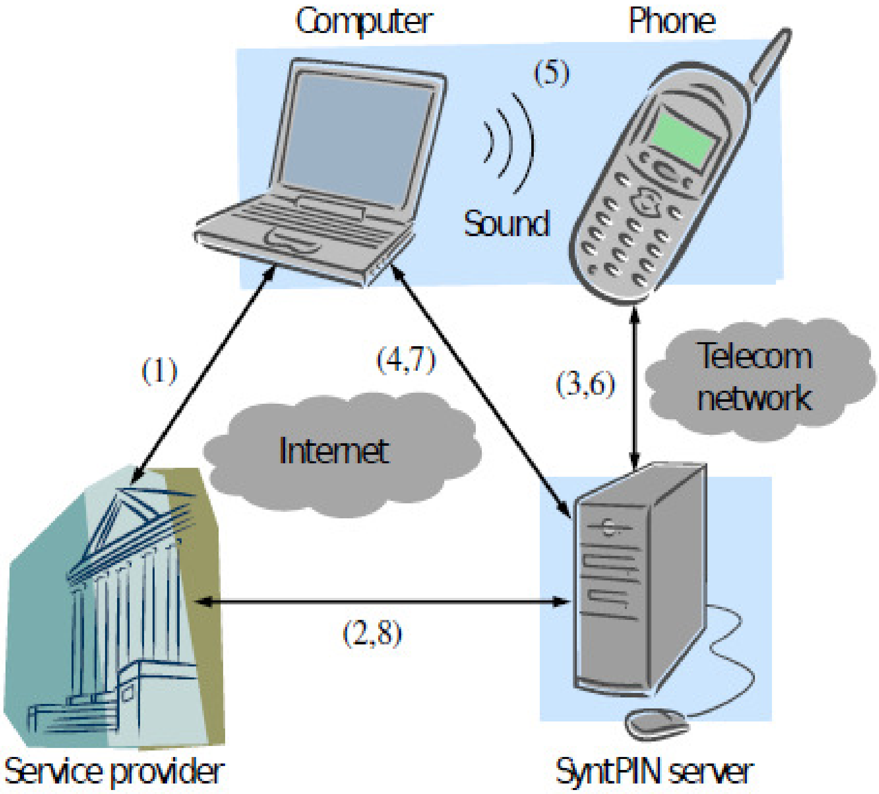

FIGURE 3: Synthetic PIN solution [33].

To perform an authentication or authorisation task the Synthetic PIN solution proceeds as follows. The user is already logged into the service provider's web site, see the channel marked (1) in the diagram. The service provider requests the task by sending a message to the SyntPIN server (2), including the phone number of the relevant user. The SyntPIN server calls the user's phone and waits for the user to answer. Assuming the user answers the call, the server plays a voice message to the user (3). In the case of authentication, SyntPIN server instructs the user to hold the phone up to the loudspeaker of the computer. Then the server sends a unique sound fingerprint4 to the user's computer (4), and the computer starts playing this sound through its loudspeaker (5). This sound is picked up by the phone and transmitted back to the SyntPIN server in the call (6). Once the sound has been received and verified by the SyntPIN server, authentication has succeeded. If the user does not answer the call or hangs up before the sound is transmitted, then the authentication has failed. The SyntPIN server informs the user about authentication success or failure on the user's computer screen (7). Authorisation of a transaction proceeds similarly, the only difference being that the SyntPIN server's call to the user also informs the user about the details of the transaction, using a synthetic voice. Then, the user is instructed to hold the phone up to the loudspeaker of the computer in case he wants to authorise the transaction, or to hang up the call not to authorise it. Last, the SyntPIN server informs the service provider about the result of the authentication or authorisation task, allowing the service provider to take appropriate action (8) [33].

The Synthetic PIN solution has some advantages in compare with the similar security technologies:

1) The PIN code isn't sent to the user. That's why PIN codes couldn't be stolen from unsecure phone, computer or other user's device.
2) SyntPIN server places a call to the user's phone. The user can identify who is calling. And it's hard to divert this call to another number without the knowledge of the user or the SyntPIN server because it requires compromising the telecom network or accessing low level functionality on the phone in order to configure call forwarding. Note that making the user's phone answer a call without the knowledge of the user is hard since to achieve this an attacker must compromise the voice call part of the phone and furthermore the attacker must divert sound from the computer loudspeaker into the answered call. Additional security is ensured if the user's phone is a landline. Mobile phones may be stolen, while landline phones can be protected using traditional physical security measures [33].

3) Also the Synthetic PIN offers to track the position of a user's phone or computer. It can help to thwart attacks: these two units should be in close proximity, otherwise there could be a man-in-the-middle attack in progress. One may also require the user to authorise a transaction by tapping a pre-determined code on the phone's touchscreen or its keyboard. In addition, attacks based on setting up hostile call forwarding may be detected through call forwarding detection mechanisms, depending upon telecom network peering agreements [33].

3.9 The Smart Card Detective

Smart Card Detective (SCD) is a hand-held device, that can protect smartcard users from several attacks, but can also showcase vulnerabilities in the Chip and PIN system. This device contains an ATMEL AVR AT90USB1287 microcontroller that mediates the communication between a smartcard and a terminal, buttons, LEDs and an LCD [36]. The cost of the device has been around $100 (including PCB manufacturing), and in large quantities the expected price is below $20. Using the SCD I developed the Filter Amount application, which was the main goal of the project. This application eavesdrops on a transaction and blocks a payment authorization request until the user verifies the correctness of the transaction. The user is able to check the transaction amount on the LCD and then decide if the transaction should continue or not. Additionally there is a Modify PIN application which replaces the PIN entered on a terminal by a PIN stored in the SCD memory. The main utility of this application is that users do not have to disclose the real PIN and thus can avoid situations where the PIN is seen by criminals looking over the shoulder. Steven Murdoch et al. have recently discovered an important vulnerability in the Chip and PIN system where a PIN transaction can succeed without entering the correct PIN although the receipt will read PIN VERIFIED. It was implemented in SCD. SCD has been successfully tested on a terminal emulator, CAP readers and live terminals [37].

The commercial interest of such device is uncertain. Although such a device can be very useful, carrying yet another gadget every time you go shopping is at least inconvenient. Also the current version of the SCD requires a wired connection between the device itself and the card interface that is inserted into the terminal. However, there are some practical uses of such a device: a user attorney for making high-amount transactions such as buying a car, a research platform for EMV, testing equipment for payment system developers to verify the correct functionality of cards and terminals.

3.10 Chip and PIN Are Broken

The central flaw in the protocol is that the PIN verification step is never explicitly authenticated. Whilst the authenticated data sent to the bank contains two fields which incorporate information about the result of the cardholder verification – the Terminal Verification Results (TVR) and the Issuer Application Data (IAD), they do not together provide an unambiguous encoding of the events which took place during the protocol run. The TVR mainly enumerates various possible failure conditions for the authentication, and in the event of success does not indicate which particular method was used [38].

Therefore a man-in-the-middle device, which can intercept and modify the communications between card and terminal, can trick the terminal into believing that PIN verification succeeded by responding with 0x9000 to Verify, without actually sending the PIN to the card. A dummy PIN must be entered, but the attack allows any PIN to be accepted. The card will then believe that the terminal did not support PIN verification, and has either skipped cardholder verification or used a signature instead. Because the dummy PIN never gets to the card, the PIN retry counter is not altered.

Neither the card nor terminal will spot this subterfuge because the cardholder verification byte of the TVR is only set if PIN verification has been attempted and failed. The terminal believes that PIN verification succeeded (and so generates a zero byte), and the card believes it was not attempted (so will accept the zero byte). The IAD does often indicate whether PIN verification was attempted. However, it is in an issuer-specific proprietary format, and not specified in EMV. Therefore the terminal, which knows the cardholder verification method chosen, cannot decode it. The issuer, which can decode the IAD, does not know which cardholder verification method was used, and so cannot use it to prevent the attack. Because of the ambiguity in the TVR encoding, neither party can identify the inconsistency between the cardholder verification methods they each believe were used. The issuer will thus believe that the terminal was incapable of soliciting a PIN – an entirely plausible yet inaccurate conclusion.

The failure we identify here might be patched in various ways which we will discuss later. But at heart there is a protocol design error in EMV: it compartmentalizes the issuer specific MAC protocol too distinctly from the negotiation of the cardholder verification method. Both of the parties who rely on transaction authentication – the merchant and the issuing bank – need to have a full and trustworthy view of the method used to verify the cardholder; and because the relevant data cannot be collected neatly by either party, the framework itself is flawed [39].

A major contributing factor to the fact that these protocol flaws remained undiscovered is the size and complexity of the specification, and its poor structure.

Core protocol failures are difficult to fix. None of the security improvements already planned by banks will help: moving from SDA to DDA will not have any effect, as these are both methods for card authentication, which occurs before the cardholder verification stage. Neither will a further proposed enhancement – CDA (combined data authentication) – in which the transaction authorization stage additionally has a digital signature under a private key held by the card. This is because the attack we present does not interfere with either the input or output of transaction authentication, so replacing a transaction MAC with a digital signature will not help. One possible work-around is for the terminal to parse the IAD, which does include the result of PIN verification. This will only be effective for online transactions, and offline transactions where CDA is used, otherwise the man-in-the-middle device could tamper with the IAD as it is returned by the card. It would also be difficult to implement because the IAD was intended only for the issuer, and there are several different formats, without any reliable method to establish which one is used by a particular card. However a solution along these lines would require the acquiring banks and the terminal vendors to act together, which for the incentive reasons discussed above would be both slow and difficult.

3.11 Why Cryptosystems Fails In ATM

Nowadays, however, it is clear that ATM security involves a number of goals, including controlling internal fraud, preventing external fraud, and arbitrating disputes fairly, even when the customer's home bank and the ATM raising the debit are in different countries. This was just not understood in the 1970's; and the need for fair arbitration in particular seems to have been completely ignored.

The second error was probably due to fairly straightforward human factors. Many organisations have no computer security team at all, and those that do have a hard time finding it a home within the administrative structure. The internal audit department, for example, will resist being given any line management tasks, while the programming staff dislike anyone whose role seems to be making their job more difficult.

Corporate politics can have an even worse effect, as we saw above: even where technical staff are aware of a security problem, they often keep quiet for fear of causing a powerful colleague to lose face. Finally, we come to the `consultants': most banks buy their consultancy services from a small number of well known firms, and value an `air of certainty and quality' over technical credentials. Many of these firms pretend to expertise which they do not possess, and cryptology is a field in which it is virtually impossible for an outsider to tell an expert from a charlatan. The author has seen a report on the security of a national ATM network switch, where the inspector (from an eminent firm of chartered accountants) completely failed to understand what encryption was, and under the

heading of communications security remarked that the junction box was well enough locked up to keep vagrants out [40].

4. E–COMMERCE AND MOBILE BANKING

4.1 Banking Security Architecture

Banking fraud cannot be eliminated without a dedicated, trusted security device. Common forms of e-banking fraud is not sufficient to protect against the criminals. avenues of attack are implementable by today’s fraudsters. There is a more robust scheme for authentication and authorization of online transactions by using a trusted device to create a very small trusted computing base, enabling secure communication with a bank without relying on the security of any of the intervening computers. This includes the computer which the customer is using to access the e-banking web site. The device forms a trusted path from the bank to the customer. Most solutions at best provide a trusted path to the user’s computer (many do not even do this), however, general purpose computers are not themselves trustworthy agents of the user’s intentions. This has been seen through the many exploits and Trojans, some of which specifically target Internet banking, to which general purpose computers are subject [41].

The proposed device negates the problems with a compromised computer by providing a trusted, authenticated path to the user over which all transactions are authorized. Because it is guaranteed that each transaction will have the correct details shown to the user the principle on which all of the attacks on online banking are based is removed. Thus, even the most powerful attack, the Trojan, is prevented.

In addition, because the device is the minimum necessary to provide the desired functionality it is possible to audit it for security vulnerabilities. It is also possible to build it with some amount of tamper resistance and hence protect it against attackers with much larger resources than is normally the case. This means that real assurances can be made that the authorization seen by the bank is the same as the one shown to the user, the only way to stop the whole class of attacks [42].

As always no solution is a panacea. There are a number of drawbacks to the proposed system. Firstly interoperability. In the past systems like this have failed because of interoperability issues. The proposal tries to mitigate a number of these, helped by the recent standardization of I/O connections and the emergence of portable languages such as Java. Suggestions for alternative methods of communication have also been made which further ameliorate those problems.

Secondly, there are still some avenues of attack left open. Obviously if an attacker can threaten the customer directly, or deceive them sufficiently, the customer may deliberately authorize a transaction to the attacker. There is only so far that technical solutions can go to prevent such abuses and such attacks are outside the scope of this work. This device also does nothing to keep the transaction log secret. The primary interface for transactions is still the computer, with just transactions being confirmed through the device. Protecting against reading of the transaction log requires all interaction to be done through the trusted path. This would significantly increase the cost and reduce the ability to audit the device.

4.2 Payment by Mobile

The primary elements of mobile payments technology include NFC, SE, and TSM.

The use of Near Field Communications (NFC) for mobile payments is governed by the ISO 18092 standard and has the following attributes [43]:

- Is limited to a 424 kilobits per second data transfer rate.
- Supports communication ranges up to approximately 0.2 meters.
- Offers no native encryption.

Under the typical scenario, NFC communications are established automatically when two compatible devices are brought within range of each other; however, the NFC technology in mobile computing and other devices used for mobile wallet transactions is typically tuned for a much shorter range, on the order of a few millimeters.

Since NFC offers no native encryption, mobile payments using NFC must be coupled with a Secure Element (SE) which is a cryptographic module in the mobile device. The exact implementation of a SE in the mobile device has still not been standardized and there are 3 competing options: 1) build it into a chip on the mobile device; 2) implement it into the existing SIM chip; 3)implement through micro SD cards. ISIS and MasterCard are leveraging the SIM approach while Google wallet is using phone that have built in modules. A major challenge for the adoption of mobile banking technology and services is the perception of insecurity. In the survey conducted by the Federal Reserve, 48% of respondents cited their main reason for not using mobile banking was “I’m concerned about the security of mobile banking”. In the same study, respondents were asked to rate the security of mobile banking for protecting their personal information and 32% rated it as somewhat unsafe and very unsafe, while 34% were not sure of the security. These statistics represent a significant barrier to the use of mobile banking products and services [44].

The security risks associated with mobile devices are very similar to any other computing device with a few key exceptions:

- Mobile devices have a smaller form factor and therefore are more susceptible to loss or theft.
- Mobile devices are more personal and there will be a tendency for users to use devices in a more personal and confidential way.
- Security controls and tools available have not matured to accommodate the constraints of limited processing power and limited battery life.

The key risks to the mobile device include:

- Malware.
- Malicious applications.
- Privacy violations relative to application collection and distribution of data.
- Wireless carrier infrastructure.
- Payments infrastructure/ecosystem.
- SMS vulnerabilities.
- Hardware and Operating System vulnerabilities.
- Complex supply chain and new entrants into the mobile ecosystem.
- Lack of maturity of Fraud tools and controls.

The mobile banking and payments ecosystem is complex and dynamic. It is not clear who will emerge as the winner(s) in the growing space from a financial services, application provider or technology perspective. Security and the perception of security will clearly play a role in who ends up dominating. Traditional financial service companies (banks, processors, and card associations) clearly have an advantage from controlling the existing banking and payments infrastructure. The extent to which they can strategically extend their products and services in a way that maintains the customer’s trust in their services be key to their success. A foundational element of that trust is the security of the products and services. The wireless carriers are challenged by entering a segment with little financial service experience. Wireless carriers are challenged by being perceived as simply a wireless bandwidth pipe and have struggled with this since the advent of wireless data. Application providers (Google, Apple) within this space clearly hold an edge relative to innovation and speed to market, however, lack of focus on security and privacy will inhibit progress [45]. Additionally, both wireless carriers and application providers are at a clear disadvantage in terms of understanding the regulatory environment faced by current financial service providers.

4.3 Protecting E−Commerce Bank and Credit Card Systems

The protection of electronic commerce systems pulls together a lot of the topics. Failures come from misconfigured access control, implementation blunders, theft of network services, inappropriate use of cryptology—you name it.

Consequently, a lot of work was done in the 1990s on beefing up intrusion detection. There are a number of generic systems that do anomaly detection, using techniques such as neural networks, but it’s unclear how effective they are. When fraud is down one year, it’s hailed as a success for the latest fraud-spotting system; when the figures go up a few years later, the vendors let the matter pass quietly [46].

Credit card numbers are indeed available on the Net, but usually because someone hacked the computer of a merchant who disobeyed the standard bank prohibition against retaining customer credit card numbers after being paid.

Likewise, fraudulent Web-based transactions do occur, but mainly because of poor implementation of the system whereby cardholder addresses are checked during authorization. The real problem facing dot-coms is disputes. It is easy to repudiate a transaction.

The critical importance for online businesses is that, if more than a small percentage of your transactions are challenged by customers, your margins will be eroded; and in extreme cases your bank may withdraw your card acquisition service.

The existing cryptographic protection mechanisms used by the bank card industry— the PINs used at ATMs and some point-of-sale terminals, and the CVVs, which make card forgery more difficult—are largely ineffective online, so new mechanisms were developed. The most widely used is the Secure Sockets Layer protocol (SSL) [47], an encryption system bundled with most Web browsers.

Most of the problems facing online businesses are no different from those facing other organizations, and the network security risks are not much different from those facing traditional businesses. The real increased risks to an e-business have to do with ways in which traditional risk management mechanisms don’t scale properly from a world of local physical transactions to one of worldwide, dematerialized ones. Credit card transaction repudiation is the main example at present. There are also significant risks to rapidly growing companies that have hired a lot of new staff but that don’t have the traditional internal controls in place.

4.4 Authentication Solutions For Ecommerce and E−Banking

Bank growth and profitability is linked to eBanking. Customers prefer online banking because it is more flexible than high street or phone banking, and it offers banks the opportunity for growth and cost savings. However, eBanking depends on secure authentication and user trust.

X Info Tech is a one-stop shop for complete eBanking security solutions, including hardware, software, consulting and design, training, maintenance and support as well as device customization and fulfillment. With global reach and unique technology [48].

When it comes to remote banking authentication, you need a system than can grow with you. X Info Tech lets you deploy a low-cost, simple system today and still provide an upgrade path for the future.

System supports a wide variety of Two-Factor Authentication solutions, including:

- One Time Password (OTP).
- Double Authentication.
- Challenge-response.
- Sign-What-You-See.
- Secure Domain Separation.
- Dynamic Signatures.
- Electronic Signatures.

The system is completely flexible, allowing you to mix and match users with different devices and authentication schemes. This approach simplifies your backend IT while maximizing flexibility.

For example, System lets you get started with Printed Card or Scratched off Card or simple One Time Password (OTP) Token and, as risks and markets change, seamlessly upgrade to more advanced devices. You can even offer other service providers a multi-issuer authentication service using your authentication system.

The result is a system that lets banks balance the demands of cost, usability and security over time. It is low-risk, scalable, secure, flexible and, above all,future-proof [49].

X Info Tech, as a Two-Factor Authentication, offers protection from all existing kinds of fraud attacks.

Authentication solution includes generation of an OTP – One Time Password. The OTP can be generated on a smart card (presented by a secure device), token, mobile phone or sent by text message.

Benefits using the Token based approach:

- Cost effective devise.
- Provides strong two-factor authentication together with online password.
- Low logistic costs.
- Portability: Token is small and portable - convenient to bring with you at all times.
- A single press on the button generates a One Time Password.
- User-friendly functionality.
- Quick roll-out.
- Smooth personalization, personalize a whole batch in factory or a single device at the bank office.

The Mobile Solution is a set of different technologies allowing authentication to be performed through already existing infrastructures. As part of the secure devices family they emphasize different capabilities with respect to security, usability and the look & feel experience. The set of media utilized offer different solutions in terms of service activation - all easy and cost effective, ranging from self-activation to Over The Air activation (OTA) [51].. The Mobile Solution enables PIN protected One Time Passwords (OTP), Signatures, Challenge/Response functionality and other services in strong Two-Factor Authentication schemes [56].

5. CONCLUSIONS

Assessing the security of Internet banking applications requires specialized knowledge on vulnerabilities, attacks and countermeasures, to gain an understanding of the threats, how they are realized and how to address them. The case study in this article demonstrated that the use of the attack tree should facilitate the work of auditors, security consultants or security officers who wish to conduct a security assessment of an Internet banking authentication mechanism.

We presented our analysis of banking and modern payments system security. We found serious logic flaws in leading online, mobile, e-commerce etc. banking applications. We discussed the weaknesses in the systems that can compromise the customer's trust. Although, we showed and analyzed ways of defense from security threats.

Most of the problems facing online businesses are no different from those facing other organizations, and the network security risks are not much different from those facing traditional businesses. The real increased risks to an e-banking have to do with ways in which traditional risk management mechanisms don’t scale properly from a world of local physical transactions to one of worldwide, dematerialized ones. Credit card transaction repudiation is the main example at present. There are also significant risks to rapidly growing companies that have hired a lot of new staff but that don’t have the traditional internal controls in place.

We believe that our study takes some steps in the banking security problem. We analyzed payments security, found problems, analyzed existing security solutions and proposed new ways to solve payments security. They are more effective and up-to-date. In future work we are considering the security challenges that come with new banking payment systems. Fundamentally, we believe that the variety and changes of banking systems demands new security approaches and research efforts on ensuring the security quality of the systems it produces.

6. REFERENCES

[1] Anderson, R.J., Needham, R.M. “Robustness principles for public key protocols”, CRYPTO 1995. LNCS, vol. 963, pp. 236–247 [1995].

[2] “APACS: Online banking usage amongst over 55s up fourfold in five years“. Internet: http://www.apacs.org.uk/media_centre/press/08_24_07.html [Aug, 2007].

[3] “APACS announces latest fraud figures”. Internet: http://www.apacs.org.uk/APACSannounceslatestfraudfigures.htm [Sep, 2008].

[4] “RedTeam: iTAN online-banking security system”. CAN-2005-2779. Internet: http://www.redteam-pentesting.de/advisories/rt-sa-2005-014.txt [Aug, 2005].

[5] “EMVCo, LLC: EMV 4.1”. Internet: http://www.emvco.com/ [Aug, 2004].

[6] Taylor, M. “Police think French pair tortured for pin details”. The Guardian. Internet: http://www.guardian.co.uk/uk/2008/jul/05/knifecrime.ukcrime [Jun, 2008].

[7] Jenkins, R. “Brainless thugs get life term”. The Times. Internet: http://www.timesonline.co.uk/tol/news/uk/crime/article3850647.ece [May, 2008].

[8] Wong, R.M., Berson, T.A., Feiertag, R.J. “Polonium: an identity authentication system”. IEEE Symposium on Security and Privacy, p. 101 [1985].

[9] Drimer, S., Murdoch, S.J. “Keep your enemies close: Distance bounding against smartcard relay attacks”. In: USENIX Security Symposium [Aug, 2007].

[10] Finn, C. “MTN not budging on fraud issue”. IOL technology. Internet: http://www.ioltechnology.co.za/article.page.php?iSectionId=2885&iArticl%eId=4402087 [May, 2008].

[11] Lomas, N. “Government gateway 2.0 looks to fatter future”. Internet: http://www.silicon.com/publicsector/0,3800010403,39168629,00.htm [Oct, 2007].

[12] “Make Card Readers Optional”. Internet: http://www.stopthecardreaders.org/ [2008].

[13] Samuel, H. “Chip and pin scam ‘has netted millions from British shoppers”. Telegraph. Internet: http://www.telegraph.co.uk/news/newstopics/politics/lawandorder/3173346%/Chip-and-pin-scam-has-netted-millionsfrom-British-shoppers.html [Oct, 2008].

[14] “Cronto: Products datasheet”. Internet: http://www.cronto.com/download/Cronto_Products_Datasheet.pdf [2010].

[15] Davida, G., Frankel, Y., Tsiounis, Y., Yung, M. “Anonymity control in E-cash systems”. FC 1997. LNCS, vol. 1318, pp. 1–16 [1997].

[16] Kerckhoffs, A. “La cryptographie militaire”. Journal des sciences militaires 9, 5–38 [1983].

[17] Bohm, N., Brown, I., Gladman, B. “Electronic commerce: Who carries the risk of fraud?” The Journal of Information, Law and Technology (3). Internet: http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2000_3/bohm/ [Oct, 2000].

[18] “Banking Code Standards Board”. The banking code. Internet: http://www.bankingcode.org.uk/ [March, 2008].

[19] Drimer, S., Murdoch, S.J., Anderson, R. “Thinking inside the box: system-level failures of tamper proofing”. IEEE Symposiumon Security and Privacy, Oakland, pp. 281–295 [May, 2008].

[20] Burrows, M., Abadi, M., and Needham, R. “A logic of authentication. ACM Transactions on Computer Systems 8”, pp.18-36 [1996].

[21] Choudary, O. “The smart card detective: a hand-held EMV interceptor. Master's thesis”, University of Cambridge. Internet: http://www.cl.cam.ac.uk/~osc22/scd/ [June 2010].

[22] “CreditCall”. EMV.LIB Integration Guide. Internet: http://www.level2kernel.com/emvlibfidocumentation.html [June, 2010].

[23] de Ruiter, J., and Poll, E. “Formal analysis of the EMV protocol suite”. Theory of Security and Applications (TOSCA 2011), vol. 6693 of LNCS, Springer, pp. 113-129 [March, 2011].

[24] Drimer, S., and Murdoch, S. J. “Keep your enemies close: Distance bounding against smartcard relay attacks”. USENIX Security Symposium [August, 2007].

[25] Drimer, S., Murdoch, S. J., and Anderson, R. “Thinking inside the box: system-level failures of tamper proofing”. IEEE Symposium on Security and Privacy (Oakland), pp. 281-295 [May, 2008].

[26] “EMVCo. Terminal level 2, test cases”. Type Approval [Nov, 2011].

[27] “EMVCo, LLC. EMV 4.2”. Internet: http://www.emvco.com/ [June, 2004].

[28] Jack, B. “Jackpotting automated teller machines redux”. Presentation at Black Hat USA. Internet: http://blackhat.com/html/bh-us-10/bh-us-10-archives.html [June, 2010].

[29] Kelman, A. “Job v Halifax PLC (not reported) case number 7BQ00307”. Digital Evidence and Electronic Signature Law Review , vol. 6 [2009].

[30] Markettos, A. T., and Moore, S. W. “Frequency injection attack on ringoscillator-based true random number generators”. Workshop on Cryptographic Hardware and Embedded Systems, pp. 317-331 [2009].

[31] Moon, D., Flatley, J., Hoare, J., Green, B., and Murphy, R. “Acquisitive crime and plastic card fraud: Findings from the 2008/09 British crime survey”. Statistical bulletin, Home Ofice, April 2010. Internet: http://webarchive.nationalarchives.gov.uk/20110218135832/http://rds.homeoffice.gov.uk/rds/pdfs10/hosb0810.pdf [April, 2010].

[32] Murdoch, S. J. “Reliability of chip & PIN evidence in banking disputes”. Digital Evidence and Electronic Signature Law Review, vol. 6, Pario Communications, pp. 98-115 [Nov, 2010].

[33] Synthetic PIN for Authentication and Authorisation”. Internet:

http://protectoria.com/Secure-Authentication [June, 2014].

[34] Needham, R. M., and Schroeder, M. D. “Using encryption for authentication in large networks of computers”. Commun. ACM 21, pp. 993-999 [Dec. 1978].

[35] “3-D Secure system overview”. Internet: https://partnernetwork.visa.com/vpn/global/retrieve_document.do?documentRetrievalId=119 [2011].

[36] “RBS Secure Terms of Use”. Internet: https://www.rbssecure.co.uk/rbs/tdsecure/terms_of_use.jsp [Dec, 2009].

[37] “APACS. 2008 fraud figures announced by APACS”. Internet: http://www.ukpayments.org.uk/media_centre/press_releases/-/page/685/ [March, 2009].

[38] Nicholas Bohm, Ian Brown, and Brian Gladman. “Electronic commerce: Who carries the risk of fraud?” The Journal of Information, Law and Technology, (3) [Oct, 2000].

[39] “Cronto”. Internet: http://www.cronto.com/download/Cronto_Products_Datasheet.pdf [2012].

[40] Saar Drimer, Steven J. Murdoch, and Ross Anderson. “Optimized to fail: Card readers for online banking”. Financial Cryptography, LNCS 5628. Springer [2009].

[41] “Internet Retailer. Verified by Visa security program used as bait in phishing scams”. Internet: http://www.internetretailer.com/dailyNews.asp?id=13764 [Jan, 2005].

[42] Jon Varco. “Verified by Visa update”. Internet: http://www.barclaycardbusiness.co.uk/information_zone/customer_forum/pdf/1315_jon_varco_visa.pdf. [2012].

[43] Yuhang Ding, Dayan Zhuang and Kejun Wang. “A Study of Hand Vein Recognition Method”, The IEEE International Conference on Mechatronics & Automation Niagara Falls, Canada [July, 2005].

[44] Shi Zhao, Yiding Wang and Yunhong Wang. “Extracting Hand Vein Patterns from Low-Quality Images: A New Biometric Technique Using Low-Cost Devices”, Fourth International Conference on Image and Graphics [2007].

[45] H. Proença, and A. Alexandre. "Towards noncooperative iris recognition: A classification approach using multiple signatures". IEEE Trans. vol. 29, pp. 607-612 [2007].

[46] Masaki Watanabe, Toshio Endoh,Morito Shiohara, and Shigeru Sasaki. ”Palm vein authentication technology and its applications”, The Biometric Consortium Conference, USA, pp.1-2 [September 19-21, 2005].

[47] Mohamed Shahin, Ahmed Badawi, and Mohamed Kamel. ”Biometric Authentication Using Fast Correlation of Near Infrared Hand Vein Patterns”, International Journal of Biological and Medical Sciences, vol 2,No.1, pp. 141-148 [winter, 2007].

[48] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn. “Image understanding for iris biometrics: A survey”, Computer Vision and Image Understanding, vol. 110, no. 2, pp. 281–307 [2008].

[49] J. Daugman. “Probing the Uniqueness and Randomness of Iris Codes:Results from 200 Billion Iris Pair Comparisons”, Proceedings of the IEEE, vol. 94, no. 11 [2006].

[50] E. M. Newton, P. J. Phillips. “Meta-Analysis of Third-Party Evaluations of Iris Recognition”, IEEE Transactions on Systems, Man, and Cybernetics, vol. 39, no. 1, pp. 4–11 [2009].

[51] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. “Robust Face Recognition via Sparse Representation”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 2, pp. 210–227 [2009].

[52] Dan Kaminsky. “DNS rebinding and more packet tricks”. 24th Chaos Communication Congress. Internet: http://events.ccc.de/congress/2007/Fahrplan/track/Hacking/2393.en.html [Dec, 2007].

[53] Jochen Topf. “HTML form protocol attack”. BugTraq posting. Internet: http://www.remote.org/jochen/sec/hfpa/hfpa.pdf [Aug, 2001].

[54] “Obscure. Extended HTML form attack”. Technical report, EyeonSecurity. Internet: http://www.hackerz.ir/e-books/Extended%20HTML%20Form%20Attack.pdf [2002].

[55] Adrian Pastor. “BT home flub: Pwnin the BT home hub - exploiting IGDs remotely via UPnP”. GNUCitizen. Internet: http://www.gnucitizen.org/blog/bt-home-flub-pwnin-the-bt-home-hub-5/ [Jan, 2008].

[56] Adrian Pastor and Petko D. Petkov. “Hacking the interwebs”. GNUCitizen. Internet: http://www.gnucitizen.org/blog/hacking-the-interwebs/ [Jan, 2008].

DNS Advanced Attacks and Analysis

Adam Ali.Zare Hudaib *adamhudaib@gmail.com*
Licensed Penetration Tester
Certified Ethical Hacker
Network Security Defence
Research &Troubleshooting
CEH , ECSA , LPT , WCNA
Poland

Esra'a Ali Zare Hudaib *israa_hudieb@eng.hu.edu.jo*
Computer & Engineering Department
The Hashemite University
Amman . Jordan

Abstract

Nowadays DNS is used to load balance, failover, and geographically redirect connections. DNS has become so pervasive it is hard to identify a modern TCP/IP connection that does not use DNS in some way. Unfortunately, due to the reliability built into the fundamental RFC-based design of DNS, most IT professionals don't spend much time worrying about it. If DNS is maliciously attacked — altering the addresses it gives out or taken offline the damage will be enormous. Whether conducted for political motives, financial gain, or just the notoriety of the attacker, the damage from a DNS attack can be devastating for the target.

In this research we will review different DNS advanced attacks and analyze them. We will survey some of the most DNS vulnerabilities and ways of DNS attacks protection.

Keywords: DNS, DoS, Cache Poisoning, DNSSEC, DNS Hijacking.

1. INTRODUCTION

Denial of Service (DoS) attacks can be classified into two major categories. In the first one, the adversary featly crafts packets trying to exploit vulnerabilities in the implemented software(service or protocol) at the target side. This class of attacks includes outbreaks like the ping of death[1]. In the second one, the aggressor attempts to overwhelm critical system's resources, i.e. memory, CPU, network bandwidth by creating numerous of well-formed but bogus requests. This type of attack is also well known as flooding. DoS attacks are a threat to almost every service in the Internet and DNS is no exception. These attacks against or related to DNS servers are also classified into two types. One is to directly flood DNS servers by sending a large number of DNS requests or other useless traffic.

Since the DNS servers cannot easily distinguish the legitimate requests from the attack traffic, they would simply accept both of them and send the responses [2]. The effective and deployable defense against this attack is to over-provision the network capacity and numbers of servers [3].The other attack strategy is to exploit DNS servers to amplify attack traffic. The attacker craftsa DNS request that gets a response significantly larger than the request itself, e.g., a 50-byterequest for a 500-byte response. The amplified response is replied to a spoofed third-party victim machine. Under this attack, both the amplifying DNS server's upstream bandwidth and the third-party machine's downstream bandwidth could be exhausted. Due to traffic amplification, an attacker can exhaust the bandwidth of its victims even if his bandwidth is 10 times smaller [4].An effective defense against spoofing-based DoS attacks on DNS servers requires source address spoof detection. Assuming a DNS server can distinguish between spoofed requests from real ones, it can selectively drop those spoofed ones with little collateral damage.

In this paper, we analyze different types of the DNS amplification attacks and ways of protection.

2. DNS ADVANCED ATTACKS AND ANALYSIS

2.1 DNS and The Most Common Security Issues

The Domain Name System (DNS) is a hierarchical, distributed database that contains mappings between names and other information, such as IP addresses.

DNS allows users to locate resources on the network by converting friendly, human-readable names like www.microsoft.com to IP addresses that computers can connect to. An often-used analogy to explain the Domain Name System is that it serves as the "phone book" for the Internet by translating human-friendly computer hostnames into IP addresses. For example, www.xyz.com translates to the addresses 20.52.88.12 (IPv4).

By default DNS works on port number 53 on TCP (Transmission Control Protocol) and UDP (user Datagram Protocol).

DNS is a crucial part of a network and hence securing DNS essentially become quite important. If a DNS is compromised, an attacker can easily prevent normal operations going in the network, can route computers to whatever spoofed IP address or resources he wants, steal information which and a lot of such malicious activity [6].

There are the essential functions of a DNS:

1. DNS is responsible for locating services like DC, etc. for authenticating the services on the network There are the essential functions of a DNS:
2. It is responsible for locating resources like Web Servers, Mail servers, etc. on the network.
3. And obviously, translating Computer names to IP and vice versa.

There are several most common security issues for DNS.

1. Unauthorized Authoritative DNS Record Changes – Changes to authoritative DNS records which point end users to computer systems outside of your control can have the most damage to your business's brand. This type of attack is typically done to either send users to a site which provides a negative marketing message, or to a location mirroring your site where account credentials can be harvested. This attack is particularly devastating because users are typically unaware anything untoward has happened [7].

2. Denial of Service Attacks – Denial of Service (DoS) or Distributed Denial of Service Attacks (DDoS) are done to make your DNS service unavailable and thus create the impression your business is offline or closed down (website, portals, VPNs, FTP, VoIP, email, etc.). This type of attack is one of the easiest to perform and can be one of the hardest to defend against. One of the least recognized impacts to a business that suffers a DNS outage from a DDoS attack is the negative effect it has on your search engine rankings.

3. Recursive DNS Spoofing/Cache Poisoning – Outside of a business's control, the Recursive DNS server an end user utilizes is typically set by the user's network administrator. Recursive DNS servers communicate the Authoritative DNS records a business sets to an end user's device. Unfortunately, many Recursive DNS servers are not well maintained or protected and can be easily compromised to give out false responses. This has the same down stream effect of an Unauthorized Authoritative DNS record change [8].

The main vulnerabilities:

1. Denial of Service.
   a. Harm and block DNS traffic.
   DNS is an effective DOS attack vector for a few reasons: DNS usually uses the UDP as its transport; most of autonomous systems allow source-spoofed packets to enter their network; there is a lot of Open DNS Resolvers on the Internet.

   The attack overloads the system by using: DNS reflectors, amplification, botnet; DDOS, recursive malformed requests, impersonation

2. Data Modification.
   a. Query/Request Redirection.
   b. DNS cache poisoning.
   c. DNS ID hacking.

   Query/Request redirection uses man-in-the-middle position, breaks of the chain of trust. DNS spoofing forges a fake answer. DNS ID hacking succeeds in impersonating a DNS server. DNS cache poisoning sends user to malicious site.

3. Zone Enumeration.
   Not really considered as an attack. Most considered as a threat as it allows attackers to gather information Precedes an attempt at an attack.

4. Tunnels.
   Uses DNS TCP transport mechanism. DNS TCP is used for: failover transport: switch from UDP to TCP; secondary zone transfer; DNSSEC and IPv6 traffic; EDNS is often badly supported by customer network.

   Attacks use TCP channel to tunnel other protocol and run malicious software [9].

2.2 Types of DNS Attacks and How To Deal with Them

DNS servers work by translating IP addresses into domain names. When DNS is compromised, several things can happen. However, compromised DNS servers are often used by attackers one of two ways. The first thing an attacker can do is redirect all incoming traffic to a server of their choosing. This enables them to launch additional attacks, or collect traffic logs that contain sensitive information.

The second thing an attacker can do is capture all in-bound email. More importantly, this second option also allows the attacker to send email on their behalf, using the victim organization's domain and cashing-in on their positive reputation. Making things worse, attackers could also opt for a third option, which is doing both of those things.

There are three common types of DNS attacks.

The first type of DNS attack is called a cache poisoning attack. This can happen after an attacker is successful in injecting malicious DNS data into the recursive DNS servers that are operated by many ISPs. These types of DNS servers are the closest to users from a network topology perspective, von Wallenstein wrote, so the damage is localized to specific users connecting to those servers.

If DNSSEC is impractical or impossible, another workaround is to restrict recursion on the name servers that need to be protected. Recursion identifies whether a server will only hand out information it has stored in cache, or if it is willing to go out on the Internet and talk to other servers to find the best answer.

"Many cache poisoning attacks leverage the recursive feature in order to poison the system. So by limiting recursion to only your internal systems, you limit your exposure. While this setting will not resolve all possible cache poisoning attack vectors, it will help you mitigate a good portion of them," Chris Brenton, Dyn Inc.'s Director of Security [10].

The second type of DNS attack happens when attackers take over one or more authoritative DNS servers for a domain. In 2009, Twitter suffered a separate attack by the Iranian Cyber Army. The group altered DNS records and redirected traffic to propaganda hosted on servers they controlled. The ability to alter DNS settings came after the Iranian Cyber Army compromised a Twitter staffer's email account, and then used that account to authorize DNS changes. During that incident Dyn Inc. was the registrar contacted in order to process the change request. Defense against these types of attacks often include strong passwords, and IP-based ACLs (acceptable client lists). Further, a solid training program that deals with social engineering will also be effective. Unfortunately, all the time and resources in the world can be placed into securing a webserver, but if an attacker can attack the authoritative server and point the DNS records at a different IP address, to the rest of the world its still going to look like you've been owned. In fact it's worse because that one attack will also permit them to redirect your email or any other service you are offering. So hosting your authoritative server with a trusted authority is the simplest way to resolve this problem.

The third type of DNS attack is also the most problematic to undo. It happens when an attacker compromised the registration of the domain itself, and then uses that access to alter the DNS servers assigned to it.

"At this time, those authoritative nameservers answered all queries for the affected domains. What makes this attack so dangerous is what’s called the TTL (time to live). Changes of this nature are globally cached on recursive DNS servers for typically 86,400 seconds, or a full day. Unless operators are able to purge caches, it can take an entire day (sometimes longer) for the effects to be reversed," von Wallenstein wrote. The main advice for authoritative DNS is to host authoritative servers within the organization, allowing for complete control [11].

2.3 Amplification Attacks

The amplification attacks are some of the largest, as measured by the number of Gigabits per second (Gbps). That size of an attack is enough to cripple even a large web host. Even from a cost perspective, the attack doesn't end up adding to our bandwidth bill because of the way in which we're charged for wholesale bandwidth.

DNS Amplification Attacks are a way for an attacker to magnify the amount of bandwidth they can target at a potential victim. Imagine you are an attacker and you control a botnet capable of sending out 100Mbps of traffic. While that may be sufficient to knock some sites offline, it is a relatively trivial amount of traffic in the world of DDoS. In order to increase your attack's volume, you could try and add more compromised machines to your botnet. That is becoming increasingly difficult. Alternatively, you could find a way to amplify your 100Mbps into something much bigger [12].

The original amplification attack was known as a SMURF attack. A SMURF attack involves an attacker sending ICMP requests (i.e., ping requests) to the network's broadcast address (i.e., X.X.X.255) of a router configured to relay ICMP to all devices behind the router. The attacker spoofs the source of the ICMP request to be the IP address of the intended victim. Since ICMP does not include a handshake, the destination has no way of verifying if the source IP is legitimate. The router receives the request and passes it on to all the devices that sit behind it. All those devices then respond back to the ping. The attacker is able to amplify the attack by a multiple of how ever many devices are behind the router (i.e., if you have 5 devices behind the router then the attacker is able to amplify the attack 5x, see the figure 1below).

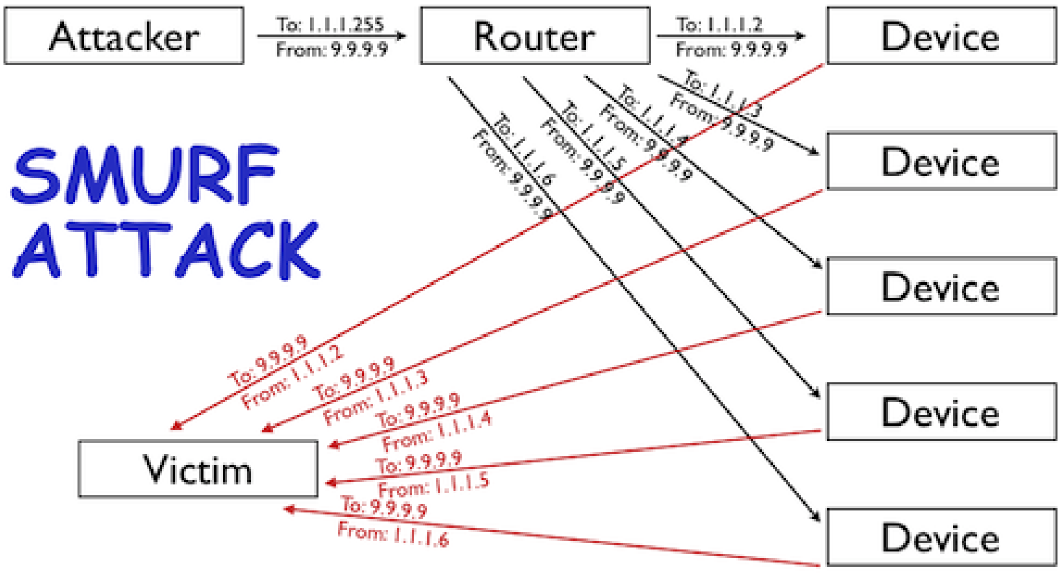

FIGURE 1: SMURF Attack.

SMURF attacks are largely a thing of the past. For the most part, network operators have configured their routers to not relay ICMP requests sent to a network's broadcast address. However, even as that amplification attack vector has closed, others remain wide open [13].

There are two criteria for a good amplification attack vector: 1) query can be set with a spoofed source address (e.g., via a protocol like ICMP or UDP that does not require a handshake); and 2) the response to the query is significantly larger than the query itself. DNS is a core, ubiquitous Internet platform that meets these criteria and therefore has become the largest source of amplification attacks.

DNS queries are typically transmitted over UDP, meaning that, like ICMP queries used in a SMURF attack, they are fire and forget. As a result, their source attribute can be spoofed and the receiver has no way of determining its veracity before responding. DNS also is capable of generating a much larger response than query.

The key term that I used a couple times so far is "open DNS resolver." The best practice, if you're running a recursive DNS resolver is to ensure that it only responds to queries from authorized clients. In other words, if you're running a recursive DNS server for your company and your company's IP space is 5.5.5.0/24 (i.e., 5.5.5.0 - 5.5.5.255) then it should only respond to queries from that range. If a query arrives from 9.9.9.9 then it should not respond.

The problem is, many people running DNS resolvers leave them open and willing to respond to any IP address that queries them. This is a known problem that is at least 10 years old. What has happened recently is a number of distinct botnets appear to have enumerated the Internet's IP space in order to discover open resolvers. Once discovered, they can be used to launch significant DNS Amplification Attacks.

Nowadays there's been an increase in big DDoS attacks. It's in large part because the network operators listed above have continued to allow open resolvers to run on their networks and the attackers have begun abusing them.

2.4 DNS Hijacking

DNS hijacking or DNS redirection is the practice of subverting the resolution of Domain Name System (DNS) queries. This can be achieved by malware that overrides a computer's TCP/IP

configuration to point at a rogue DNS server under the control of an attacker, or through modifying the behavior of a trusted DNS server so that it does not comply with internet standards. These modifications may be made for malicious purposes such as phishing, or for self-serving purposes by Internet service providers (ISPs) to direct users' web traffic to the ISP's own web servers where advertisements can be served, statistics collected, or other purposes of the ISP; and by DNS service providers to block access to selected domains as a form of censorship [14].
A number of consumer ISPs such as Cablevision's Optimum Online, Comcast, Time Warner, Cox Communications, RCN, Rogers, Charter Communications, Verizon, Virgin Media, Frontier Communications, Bell Sympatico, UPC, T-Online, Optus, Mediacom, ONO, TalkTalk and Bigpond (Telstra) use DNS hijacking for their own purposes, such as displaying advertisements or collecting statistics. This practice violates the RFC standard for DNS (NXDOMAIN) responses and can potentially open users to cross-site scripting attacks.

The concern with DNS hijacking involves this hijacking of the NXDOMAIN response. Internet and intranet applications rely on the NXDOMAIN response to describe the condition where the DNS has no entry for the specified host. If one were to query the invalid domain name (fakeexample.com), one should get an NXDOMAIN response - informing the application that the name is invalid and taking the appropriate action (for example, displaying an error or not attempting to connect to the server). However, if the domain name is queried on one of these non-compliant ISPs, one would always receive a fake IP address belonging to the ISP. In a web browser, this behavior can be annoying or offensive as connections to this IP address display the ISP redirect page of the provider, sometimes with advertising, instead of a proper error message. However, other applications that rely on the NXDOMAIN error will instead attempt to initiate connections to this spoofed IP address, potentially exposing sensitive information.

Examples of functionality that breaks when an ISP hijacks DNS:

- Roaming laptops that are members of a Windows Server domain will falsely be led to believe that they are back on a corporate network because resources such as domain controllers, email servers and other infrastructure will appear to be available. Applications will therefore attempt to initiate connections to these corporate servers, but fail, resulting in degraded performance, unnecessary traffic on the internet connection and timeouts.

- Many small office and home networks do not have their own DNS server, relying instead on broadcast name resolution. Many versions of Microsoft Windows default to prioritizing DNS name resolution above NetBIOS name resolution broadcasts; therefore, when an ISP DNS server returns a (technically valid) IP address for the name of the desired computer on the LAN, the connecting computer uses this incorrect IP address and inevitably fails to connect to the desired computer on the LAN. Workarounds include using the correct IP address instead of the computer name, or the DhcpNodeType registry value to change name resolution service ordering.

- Browsers such as Firefox no longer have their 'Browse By Name' functionality (Where keywords typed in the address bar take you to the closest matching site.).

- The local DNS client built into modern operating systems will cache results of DNS searches for performance reasons. If a client switches between a home network and a VPN, false entries may remain cached, thereby creating a service outage on the VPN connection.DNSBL anti-spam solutions rely on DNS; false DNS results therefore interfere with their operation [15].

- Confidential user data might be leaked by applications that are tricked by the ISP into believing that the servers they wish to connect to are available.

- User choice over which search engine to consult in the event of a URL being mistyped in a browser is removed as the ISP determines what search results are displayed to the user; functionality of applications like the Google Toolbar do not work correctly.

- Computers configured to use a split tunnel with a VPN connection will stop working because intranet names that should not be resolved outside the tunnel over the public Internet will start resolving to fictitious addresses, instead of resolving correctly over the VPN tunnel on a private DNS server when an NXDOMAIN response is received from the Internet. For example, a mail client attempting to resolve the DNS A record for an internal mail server may receive a false DNS response that directed it to a paid-results web server, with messages queued for delivery for days while retransmission was attempted in vain.

- It breaks Web Proxy Autodiscovery Protocol (WPAD) by leading web browsers to believe incorrectly that the ISP has a proxy server configured.

- It breaks monitoring software. For example, if we periodically contact a server to determine its health, a monitor will never see a failure unless the monitor tries to verify the server's cryptographic key.

In some cases, the ISPs provide subscriber-configurable settings to disable hijacking of NXDOMAIN responses. Correctly implemented, such a setting reverts DNS to standard behavior. Other ISPs, however, instead use a web browser cookie to store the preference. In this case, the underlying behavior is not resolved: DNS queries continue to be redirected, while the ISP redirect page is replaced with a counterfeit dns error page. Applications other than web-browsers cannot be opted out of the scheme using cookies as the opt-out targets only the HTTP protocol, when the scheme is actually implemented in the protocol-neutral DNS system.

2.5 DoS Attacks

Perpetrators of DoS attacks typically target sites or services hosted on high-profile web servers such as banks, credit card payment gateways, and even root nameservers. This technique has now seen extensive use in certain games, used by server owners, or disgruntled competitors on games, such as server owners' popular Minecraft servers. Increasingly, DoS attacks have also been used as a form of resistance. Richard Stallman has stated that DoS is a form of 'Internet Street Protests’. The term is generally used relating to computer networks, but is not limited to this field; for example, it is also used in reference to CPU resource management [16].

One common method of attack involves saturating the target machine with external communications requests, so much so that it cannot respond to legitimate traffic, or responds so slowly as to be rendered essentially unavailable. Such attacks usually lead to a server overload. In general terms, DoS attacks are implemented by either forcing the targeted computer(s) to reset, or consuming its resources so that it can no longer provide its intended service or obstructing the communication media between the intended users and the victim so that they can no longer communicate adequately.

Denial-of-service attacks are considered violations of the Internet Architecture Board's Internet proper use policy, and also violate the acceptable use policies of virtually all Internet service providers. They also commonly constitute violations of the laws of individual nations.

The United States Computer Emergency Readiness Team (US-CERT) defines symptoms of denial-of-service attacks to include:

- Unusually slow network performance (opening files or accessing web sites);
- Unavailability of a particular web site;
- Inability to access any web site;
- Dramatic increase in the number of spam emails received—(this type of DoS attack is considered an e-mail bomb);

- Disconnection of a wireless or wired internet connection;
- Long term denial of access to the web or any internet services [17].

Denial-of-service attacks can also lead to problems in the network 'branches' around the actual computer being attacked. For example, the bandwidth of a router between the Internet and a LAN may be consumed by an attack, compromising not only the intended computer, but also the entire network or other computers on the LAN.

If the attack is conducted on a sufficiently large scale, entire geographical regions of Internet connectivity can be compromised without the attacker's knowledge or intent by incorrectly configured or flimsy network infrastructure equipment.

A denial-of-service attack is characterized by an explicit attempt by attackers to prevent legitimate users of a service from using that service. There are two general forms of DoS attacks: those that crash services and those that flood services [18].

A DoS attack can be perpetrated in a number of ways. Attacks can fundamentally be classified into five families:

1. Consumption of computational resources, such as bandwidth, memory, disk space, or processor time.
2. Disruption of configuration information, such as routing information.
3. Disruption of state information, such as unsolicited resetting of TCP sessions.
4. Disruption of physical network components.
5. Obstructing the communication media between the intended users and the victim so that they can no longer communicate adequately.

A DoS attack may include execution of malware intended to:

- Max out the processor's usage, preventing any work from occurring.
- Trigger errors in the microcode of the machine.
- Trigger errors in the sequencing of instructions, so as to force the computer into an unstable state or lock-up.
- Exploit errors in the operating system, causing resource starvation and/or thrashing, i.e. to use up all available facilities so no real work can be accomplished or it can crash the system itself
- Crash the operating system itself.

In most cases DoS attacks involve forging of IP sender addresses (IP address spoofing) so that the location of the attacking machines cannot easily be identified and to prevent filtering of the packets based on the source address.

A Distributed Denial of Service Attack (DDoS) occurs when multiple systems flood the bandwidth or resources of a targeted system, usually one or more web servers. This is the result of multiple compromised systems (for example a botnet) flooding the targeted system with traffic. When a server is overloaded with connections, new connections can no longer be accepted. The major advantages to an attacker of using a distributed denial-of-service attack are that multiple machines can generate more attack traffic than one machine, multiple attack machines are harder to turn off than one attack machine, and that the behavior of each attack machine can be stealthier, making it harder to track and shut down. These attacker advantages cause challenges for defense mechanisms. For example, merely purchasing more incoming bandwidth than the current volume of the attack might not help, because the attacker might be able to simply add more attack machines.

Malware can carry DDoS attack mechanisms; one of the better-known examples of this was MyDoom. Its DoS mechanism was triggered on a specific date and time. This type of DDoS involved hardcoding the target IP address prior to release of the malware and no further interaction was necessary to launch the attack.

A system may also be compromised with a trojan, allowing the attacker to download a zombie agent, or the trojan may contain one. Attackers can also break into systems using automated tools that exploit flaws in programs that listen for connections from remote hosts. This scenario primarily concerns systems acting as servers on the web. Stacheldraht is a classic example of a DDoS tool. It utilizes a layered structure where the attacker uses a client program to connect to handlers, which are compromised systems that issue commands to the zombie agents, which in turn facilitate the DDoS attack. Agents are compromised via the handlers by the attacker, using automated routines to exploit vulnerabilities in programs that accept remote connections running on the targeted remote hosts. Each handler can control up to a thousand agents. In some cases a machine may become part of a DDoS attack with the owner's consent, for example, in Operation Payback, organized by the group Anonymous [19].

These collections of systems compromisers are known as botnets. DDoS tools like Stacheldraht still use classic DoS attack methods centered on IP spoofing and amplification like smurf attacks and fraggle attacks (these are also known as bandwidth consumption attacks). SYN floods (also known as resource starvation attacks) may also be used. Newer tools can use DNS servers for DoS purposes. Unlike MyDoom's DDoS mechanism, botnets can be turned against any IP address. Script kiddies use them to deny the availability of well known websites to legitimate users. More sophisticated attackers use DDoS tools for the purposes of extortion – even against their business rivals.

Simple attacks such as SYN floods may appear with a wide range of source IP addresses, giving the appearance of a well distributed DoS. These flood attacks do not require completion of the TCP three way handshake and attempt to exhaust the destination SYN queue or the server bandwidth. Because the source IP addresses can be trivially spoofed, an attack could come from a limited set of sources, or may even originate from a single host. Stack enhancements such as syn cookies may be effective mitigation against SYN queue flooding, however complete bandwidth exhaustion may require involvement [20].

If an attacker mounts an attack from a single host it would be classified as a DoS attack. In fact, any attack against availability would be classed as a Denial of Service attack. On the other hand, if an attacker uses many systems to simultaneously launch attacks against a remote host, this would be classified as a DDoS attack.

2.6 Prevention of The Attacks

There are some ways of DNS attacks prevention.

1. Usage of the best practices configurations.
    a. Run software in secure environment.
    b. Identify data flow.
    c. ACLs.
    d. Stealth Architecture.

2. Enabling DNSSEC.

3. Monitoring DNS Traffic.
    a. Short term analysis (peak detection).
    b. Long term analysis (abnormal behavior).

By server secure environment is meant: running up-to-date software version; checking that the operating system is also having all security fixes; efficient IP comes into an appliance format with a single upgrade process that updates: operating system, services, software.

Also you must identify data flow; run caching, resolver, authoritative server. You should separate the functions as possible and disable unwanted features It will help into preventing attacks. A public authoritative server should never be recursive [21].

Access control list is very important too.

ACLs are used to control what information will be published. With data flow identification, you can choose who will be able to:

- Allow query (server and zone level);
- Allow query cache (server level);
- Allow transfer (server and zone level);
- Allow update (zone level);
- Blackhole (server level);
- Negative Cache (zone level).

There are library of SmartArchitecture DNS templates. One of them is DNS Stealth: State of the Art Internet DNS architecture (see the figure 2).

DNSSEC is used to protect against query/request redirection. DNSSEC creates a chain of trust between the client and the authoritative server. Based on key exchange inside specific signed resource records.

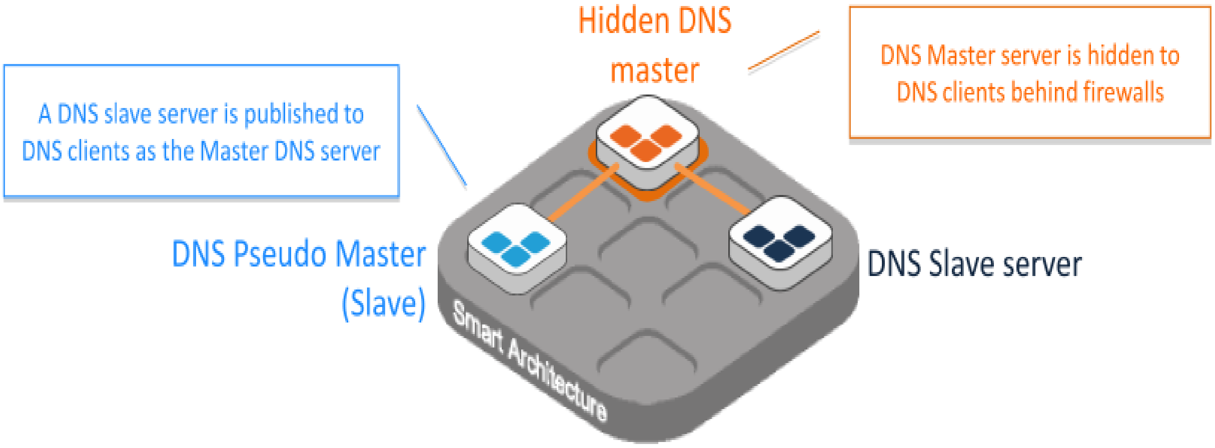
Hidden DNS master
DNS Master server is hidden to DNS clients behind firewalls
A DNS slave server is published to DNS clients as the Master DNS server
DNS Pseudo Master (Slave)
DNS Slave server
Smart Architecture

FIGURE 2: DNS Stealth Architecture.

2.7 Security Best Practices

Registrar Lock Your Domain Names – One of the simplest protections that can be used is lock all of your domain names at your registrar [22].

Outsource DNS Services – In today's world it is typically unrealistic to maintain your own DNS name servers in a way that both protects them from attacks and maintains global performance, and it is naive to use the free DNS services of a domain registrar. Cloud based managed service providers are your best bet for both Authoritative & Recursive DNS. Neustar (UltraDNS), DynDNS, Verisign, Amazon (Route 53), and Community DNS (European focused) are some of the top IP Anycasted Authoritative DNS providers to consider. OpenDNS, Neustar (UltraDNS), DynDNS & Google are the top Recursive DNS providers to consider. The investment in a cloud based DNS provider will protect your business from many of the common attacks, and free you from having to manage the devices yourself.

Utilize Strong Access Controls – As with any critical IT infrastructure, only allow users access to DNS administration for what they need to manage, lock down access to these critical accounts to known IP ranges, utilize strong password controls, and whenever possible use two factor authentication.

Activate DNSSEC On Your Domain Names – DNSSEC counters cache poisoning attacks by verifying the authenticity of responses received from name servers. It effectively prevents responses from being tampered with, because in practice, signatures are almost impossible to forge without access to the private keys.

Continuously Monitor Your Critical Services & DNS Records – Utilize an advanced SIEM like the one available from Savanture to monitor all of critical services and monitor your DNS records for changes from outside your network. UltraTools.com provides a free DNS monitoring service that many top organizations use. Additionally, monitoring the activity level on your services can show when traffic suddenly gets directed away.

Promote The Use of Protected Recursive DNS Servers – Usage of one of the top Recursive DNS providers for network. Many times there is no cost to this, only a configuration change [23].

Protect DNS Service Against DDoS Attacks – Usafe of one of the top Authoritative DNS providers that provides DDoS protection for DNS service. For public facing services that require DDoS protection, lower your DNS Time to Live (TTLs) settings to 300 (5min) so it can redirect traffic quickly if you come under attack and need protection.

3. CONCLUSIONS

We presented our analysis for DNS attacks. We found serious logic flaws in advanced attacks mechanisms. We discussed the weaknesses in the DNS systems and ways of its protection.

Much of the Internet's DNS infrastructure remains open and unprotected—characterized by a lack of dedicated security personnel, poor traffic visibility and unrestricted access to DNS recursors. Yet security threats against DNS infrastructure are serious—and growing.

We believe that our study takes some steps in the security problem space that DNS infrastructure has brought. We believe that our study brings some new chain of trust between the client and the authoritative server in DNS security. In future work we are considering the security challenges that come with other advanced DNS attacks. Fundamentally, we believe that vulnerabilities of DNS demands new research efforts on ensuring the security quality of the systems.

4. REFERENCES

[1] ”Denial of Service Attack via ping”. Internet: http://www.cert.org/advisories/CA-1996-26.html [Dec, 1996].

[2] Sun Changhua, Liu Bin, Shi Lei. “Efficient and low-cost hardware defense against DNS amplification attacks”. IEEE Global Telecommunications Conference, GLOBECOM 2008 [May, 2008].

[3] Li M, Li J, Zhao W. “Simulation Study of Flood Attacking of DDOS”, Icicse: International Conference on Internet Computing in Science and Engineering, Proceedings [June, 2008].

[4] Guo Fanglu, Chen Jiawu, Chiueh Tzi-Cker, Spoof detection for preventing DoS attacks against DNS servers, 26th IEEE International Conference on Distributed Computing Systems, ICDCS [Feb, 2006].

[5] Kambourakis G., Moschos T., Geneiatakis D., Gritzalis S, Detecting DNS Amplification Attacks, Critical Information Infrastructures Security, v(5141), pp. 185 – 196.

[6] Bau J., Mitchell J., A security evaluation of DNSSEC with NSEC3, Citeseer [May, 2010].

[7] Li Wei-min, Chen Lu-ying, Lei Zhen-ming, Alleviating the impact of DNS DDoS attacks , Proceedings of the 2010 2nd International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC 2010), pp. 240-243 [Dec, 2010].

[8] Scalzo F, Recent DNS Reflector Attacks Verisign. Internet: http://www.nanog.org/mtg-0606/pdf/frank-scalzo.pdf [Dec, 2006].

[9] Sen J, A Robust Mechanism for Defending Distributed Denial OF Service Attacks on Web Servers, Arxiv preprint arXiv: 1103.3333 [Jul, 2011].

[10] Dittrich D, Distributed Denial of Service (DDoS) Attacks/tools. Internet: http://staff.washington.edu/dittrich/misc/ddos [Oct, 2012].

[11] The Measurement Factory, Domain name servers: pervasive and critical, yet often overlooked, The Measurement Factory DNS Survey. Internet: http://dns.measurement-factory.com/surveys/sum1.html [Nov, 2005].

[12] Singh A, Singh B, Joseph H, Vulnerability Analysis for DNS and DHCP, Vulnerability Analysis and Defense for the Internet, pp. 111-124 [Dec, 2008].

[13] Beverly R and Bauer S, The spoofer project: inferring the extent of source address filtering on the Internet, USENIX workshop on Steps to Reducing Unwanted Traffic on the Internet, 2005.272 X. Ye et al. /Journal of Computational Information Systems 9, pp. 265–272 [May, 2013].

[14] V. Paxson, “An analysis of using reflectors for distributed denial-of-service attacks, ” SIGCOMM Comput. Commun.Rev., vol. 31, no. 3, pp. 38 – 47 [May, 2011].

[15] K. Rikitake, “A Study of DNS Transport Protocol for Improving the Reliability, ” Ph.D. dissertation, Graduate School of Information Science and Technology, Osaka University [Oct, 2005].

[16] M. de Vivo, G. O. de Vivo, R. Koeneke, and G. Isern, “Internet vulnerabilities related to TCP/IP and T/TCP, ” SIGCOMM Comput. Commun. Rev., vol. 29, no. 1, pp. 81 – 85 [Dec, 1999].

[17] V. Ramasubramanian and E. G. Sirer, “The design and implementation of a next generation name service for the internet, ” SIGCOMM Comput. Commun. Rev., vol. 34, no. 4, pp. 331 – 342 [Feb, 2004].

[18] H. Yang, H. Luo, Y. Yang, S. Lu, and L. Zhang, “HOURS: Achieving DoS Resilience in an Open Service Hierarchy, ” in Proc. IEEE DSN04 [March, 2004].

[19] ICANN SSAC, SSAC Advisory SAC008 DNS Distributed Denial of Service (DDoS) Attacks. Internet: http://www.icann.org/committees/security/dns-ddos-advisory-31mar06.pdf [Feb, 2006].

[20] Huiming Yu, Xiangfeng Dai, Baxliey T, Xiaohong Yuan, Bassett T, A Visualization Analysis Tool for DNS Amplification Attack, Proceedings of the 2010 3rd International Conference on Biomedical Engineering and Informatics (BMEI 2010) [May, 2010].

[21] IPTraf - An IP Network Monitor. Internet: http://iptraf.seul.org/ [Jan, 2014].

[22] S. Murdoch and R. Anderson. “Verified by Visa and MasterCard SecureCode: or, How Not to Design Authentication”. Financial Cryptography and Data Security, pp. 42-45 [Jan, 2010].

[23] “SSL: Intercepted today, decrypted tomorrow”. Netcraft, pp. 10-12 [May, 2013].

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Computer Science and Security (IJCSS)* is a refereed online journal which is a forum for publication of current research in computer science and computer security technologies. It considers any material dealing primarily with the technological aspects of computer science and computer security. The journal is targeted to be read by academics, scholars, advanced students, practitioners, and those seeking an update on current experience and future prospects in relation to all aspects computer science in general but specific to computer security themes. Subjects covered include: access control, computer security, cryptography, communications and data security, databases, electronic commerce, multimedia, bioinformatics, signal processing and image processing etc.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCSS.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 8, 2014, IJCSS is appearing with more focused issues. Besides normal publications, IJCSS intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

IJCSS LIST OF TOPICS
The realm of International Journal of Computer Science and Security (IJCSS) extends, but not limited, to the following:

- Authentication and authorization models
- Computer Engineering
- Computer Networks
- Cryptography
- Databases
- Image processing
- Operating systems
- Programming languages
- Signal processing
- Theory
- Communications and data security
- Bioinformatics
- Computer graphics
- Computer security
- Data mining
- Electronic commerce
- Object Orientation
- Parallel and distributed processing
- Robotics
- Software engineering

CALL FOR PAPERS

Volume: 8 - **Issue: 3**

i. Submission Deadline : April 30, 2014 **ii. Author Notification:** May 31, 2014

iii. Issue Publication: June 2014

CONTACT INFORMATION

Computer Science Journals Sdn BhD

B-5-8 Plaza Mont Kiara, Mont Kiara

50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6204 5627
Fax: 006 03 6204 5628
Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2014
COMPUTER SCIENCE JOURNALS SDN BHD
B-5-8 PLAZA MONT KIARA
MONT KIARA
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6204 5627
FAX: 006 03 6204 5628
EMAIL: cscpress@cscjournals.org