





Dhinaharan Nagamalai  
Jan Zizka (Eds)

# Computer Science & Information Technology

4<sup>th</sup> International Conference on Information Technology, Control, Chaos,  
Modeling and Applications (ITCCMA-2017)  
Dubai, UAE, May 27~28, 2017



**AIRCC Publishing Corporation**

## **Volume Editors**

Dhinaharan Nagamalai,  
Wireilla Net Solutions, Australia  
E-mail: dhinthia@yahoo.com

Jan Zizka,  
Mendel University in Brno, Czech Republic  
E-mail: zizka.jan@gmail.com

ISSN: 2231 - 5403  
ISBN: 978-1-921987-67-0  
DOI : 10.5121/csit.2017.70701 - 10.5121/csit.2017.70706

This work is subject to copyright. All rights are reserved, whether whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the International Copyright Law and permission for use must always be obtained from Academy & Industry Research Collaboration Center. Violations are liable to prosecution under the International Copyright Law.

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

## Preface

The 4<sup>th</sup> International Conference on Information Technology, Control, Chaos, Modeling and Applications (ITCCMA-2017) was held in Dubai, UAE, during May 27~28, 2017. The 3<sup>rd</sup> International Conference on Computer Science, Engineering & Information Technology (CSITY 2017), The 3<sup>rd</sup> International Conference on Artificial Intelligence & Fuzzy Logic Systems (AIFZ 2017), The 3<sup>rd</sup> International Conference on Networks & Communications (NWCOM 2017) and The 3<sup>rd</sup> International Conference on Signal and Image Processing (SIGPRO 2017) was collocated with The 4<sup>th</sup> International Conference on Information Technology, Control, Chaos, Modeling and Applications (ITCCMA-2017). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The ITCCMA-2017, CSITY-2017, AIFZ-2017, NWCOM-2017, SIGPRO-2017 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, ITCCMA-2017, CSITY-2017, AIFZ-2017, NWCOM-2017, SIGPRO-2017 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the ITCCMA-2017, CSITY-2017, AIFZ-2017, NWCOM-2017, SIGPRO-2017.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

Dhinaharan Nagamalai  
Jan Zizka

## Organization

### General Chair

Natarajan Meghanathan,  
Brajesh Kumar Kaushik,

Jackson State University, USA  
Indian Institute of Technology - Roorkee, India

### Program Committee Members

Aguilar C	Universidad de Los Andes, Venezuela
Ahmad Rawashdeh	University of Central Missouri, United States
Ahmad T. Al-Taani	Yarmouk University, Jordan
Ahmed Korichi	University of Ouargla, Algeria
Ahmed Mohamed Khedr	Sharjah University, UAE
Alberto Magrenan	International University of La Rioja (UNITE), Spain
Ali Javadi	Iran University of Science and Technology, Iran
Ali Salem	University of Sfax, Tunisia
Ameera Saleh. Jaradat	Yarmouk University, Jordan
Amel B.H.Adamou-Mitiche	University of Djelfa, Algeria
Ammar Al-Masri	Albalqa Applied University, Jordan
Andre F.Dantas	UnP - Universidade Potiguar, Brasil
Ankit Chaudhary	Truman State University, USA
Atallah Mahmoud AL-Shatnawi	Al al-Byte University, Jordan
Azeddine Chikh	University of Tlemcen, Algeria
Chang-Hyun	Korea Marine Equipment Research Institute, Korea
Dabin Ding	University of Central Missouri, United States
Daniel Gomes	Estácio de Sá, Brasil
Dinyo Omosehinmi	Colossus Technology, Nigeria
Dongpo Xu	Northeast Normal University, China
Duta Cristina-Loredana	University Politehnica of Bucharest, Romania
Efthimios Alepis	University of Piraeus, Greece
Elaheh Pourabbas	National Research Council, Italy
Emad Al-Shawakfa	Yarmouk University, Jordan
Emad Awada	Applied Science University, Jordan
Erman Cakit	Aksaray University, Turkey
Ethirajan Rajan	Pentagram Group of Companies, India
Eyad Al-Zobaydi	Al-Isra University, Jordan
Eyad M. Hassam ALazam	Yarmouk University, Jordan
Fatih Korkmaz	Cankiri Karatekin University, Turkey
Fernando Bobillo	University of Zaragoza, Spain
Fernando Tello Gamarra	Federal University of Santa Maria, Brazil
Ferran Torrent	Universitat de Girona, Girona
G Sriharee	King Mongkut's University of Technology, Thailand
Gammoudi Aymen	University of Tunis, Tunisia
Gonabad Branch	Islamic Azad University, Iran
Hamed Al-Rubaiee	University of Bedfordshire, United Kingdom

Hamid Alasadi	Basra University, Iraq
Hayet Mouss	Batna Univeristy, Algeria
Hongzhi	Harbin Institute of Technology, China
Hossein Jadidoleslamy	MUT University, Iran
Isa Maleki	Islamic Azad University, Iran
Islam Atef	Alexandria University, Egypt
Jalel Akaichi	University of Tunis, Tunisia
Jamal El Abbadi	Mohammadia V University Rabat, Morocco
John Tass	University of Patras, Greece
Jun Zhang	South China University of Technology, China
Khaled Almakadmeh	Hashemite University, Jordan
Khalid M. Oqla Nahar	Yarmouk University, Jordan
Kishore Rajagopalan	Prairie Research Institute, US
Liviu Octavian Maftciu-Scai	West University of Timisoara, Romania
Liyakath Unisa	Prince Sultan University, Saudi Arabia
Liyakathunisa Syed	Prince Sultan University, Saudi Arabia
Mahdi Mazinani	University of Tehran, Iran
Mahdi Salarian	University of Illinois, USA
Mahmoud Shafik	University of Derby, UK
Manoj Vasanth Ram	Lead Hardware Engineer, USA
Masoumeh Javanbakht	Hakim Sabzevari University, Iran
Maysam Toghraee	Yasouj Science and Research Branch, Islamic
Mohamed Tounsi	Prince Sultan University, Saudi Arabia
Mohamedmaher Benismail	King Saud University, Saudi Arabia
Mohammad Ashraf OTTOM	Yarmouk University, Jordan
Mohammad Rawashdeh	University of Central Missouri, United States
Mohammad Zarour	Prince Sultan University, Saudi Arabia
Mohammed Al-Sarem	Taibah University, KSA
Mohammed Ghazi Al-Zamel	Yarmouk University, Jordan
Mostafa Ashry	Alexandria University, Egypt
Mourchid Mohammed Ibn	Tofail University Kenitra, Morocco
Mudassir Khan	King Khalid University, Saudi Arabia
Narges Shafieian	Azad University, Iran
Nicolas H. Younan	Mississippi State University, USA
Noura Taleb	Badji Mokhtar University, Algeria
Ouafa Mah	Ouargla University, Algeria
Paulo Roberto Martins de Andrade	University de Regina, Canada
Quang Hung Do	University of Transport Technology, Vietnam
Rafael Stubs Parpinelli	State University of Santa Catarina, Brazil
Renato Carrijo	Universidade Federal de Uberlandia, Brazil
Siuly Siuly	Victoria University, Australia
Wonjun Lee	The University of Texas at San Antonio, USA
Yong Li	University of Tennessee, USA
Yousfi Abdellah	Mohammed V University, Morocco
Zainab A. Khalaf	Basrah Universit, Iraq
Zhendong Shao	University of Toronto Scarborough, Canada

## **Technically Sponsored by**

**Computer Science & Information Technology Community (CSITC)**



**Networks & Communications Community (NCC)**



**Digital Signal & Image Processing Community (DSIPC)**



## **Organized By**



**Academy & Industry Research Collaboration Center (AIRCC)**

## TABLE OF CONTENTS

### **4<sup>th</sup> International Conference on Information Technology, Control, Chaos, Modeling and Applications (ITCCMA-2017)**

**Development of Tool to Promote Web Accessibility for Deaf ..... 01 - 15**  
*Ahmad A. Alhamed and Maha S. Alqhtani*

### **3<sup>rd</sup> International Conference on Computer Science, Engineering & Information Technology (CSITY 2017)**

**CERT Strategy to Deal with Phishing Attacks ..... 17 - 28**  
*Shahrzad Sedaghat*

### **3<sup>rd</sup> International Conference on Artificial Intelligence & Fuzzy Logic Systems (AIFZ 2017)**

**Efficient Use of Hybrid Adaptive Neuro-Fuzzy Inference System Combined  
with Nonlinear Dimension Reduction Method in Production Processes ..... 29 - 43**  
*Saad Bashir Alvi, Robert Martin and Johannes Gottschling*

### **3<sup>rd</sup> International Conference on Networks & Communications (NWCOM 2017)**

**Classification of Smart Environment Scenarios in Combination with a Human  
Wearable Environment Communication Using Wireless Connectivity..... 45 - 58**  
*Kristof Friess and H.C. Volker Herwig*

**Channel Estimation for the ISDB-T<sub>B</sub> FBMC System Using Neural Networks :  
A Proposal of Application of Back Propagation Training Algorithm..... 59 - 69**  
*Jefferson Jesus Hengles Almeida, P. B. Lopes, Cristiano Akamine and  
Nizam Omar*

### **3<sup>rd</sup> International Conference on Signal and Image Processing (SIGPRO 2017)**

**Multi Resolution Lattice Discrete Fourier Transform (MRL-DFT) ..... 71 - 90**  
*Fred L. Fontaine*

# DEVELOPMENT OF TOOL TO PROMOTE WEB ACCESSIBILITY FOR DEAF

Ahmad A. Alhamed and Maha S. Alqhtani

Department of Information Systems, King Saud University, Riyadh, KSA

## **ABSTRACT**

*This paper presents an intelligent tool to summarize Arabic articles and translate the summary to Arabic sign language based on avatar technology. This tool was designed for improving web accessibility and help deaf people for acquiring more benefits. Of the most important problems in this task, is that the deaf people were facing many difficulties in reading online articles because of their limited reading skills. To solve this problem the proposed tool includes a summarizer that summaries the articles, so deaf people will read the summary instead of reading the whole article. Also the tool use avatar technology to translate the summary to sign language.*

## **KEYWORDS**

*Accessibility, Arabic sign language, Key terms summarizer, Avatar.*

## **1. INTRODUCTION**

The world wide web has evolved very fast in recent years and become necessary for everybody. It is a rich source of information and documents in all fields. Unfortunately, deaf people cannot take full advantage of these advances in technology. Some Web sites are not sufficiently accessible to deaf people, making it difficult or impossible for them to contribute to the Web.

Web accessibility means that the web is accessible to every body even to people with disabilities. In another word, people with disabilities can perceive, understand, and interact with the Web. There are many attempts to improve the web accessibility by providing closed captioning and transcriptions, whenever a sound appears. In addition, the accessibility for deaf can be improved with the utilization of the sign language.

Deaf people face problems in reading electronic documents or articles that are available on certain web pages. They have a hard time to understand the written language because of their limited vocabulary and language skills. The poor language capability for deaf people is because they cannot listen and practice the basic language techniques[45]. According to the previous studies in the field of the deaf, reading level of the deaf is lower than the reading level of hearing people. Furthermore, the level of experience achieved by deaf children in four years is equivalent to the level of one year for hearing children [24].

Most web pages offer written information such as news, articles or other documents. These web pages may lose a number of visitors who are deaf. Because deaf readers have difficulties with many skills that are required for reading such as word identification, vocabulary. Also they have problems with reader-based skills like working memory [16]. With increased advancement of technology text summarization has a number of applications that are used in the analysis of a high volume text documents. With text summarization tools the online documents can be summarized into a shorter form. So, deaf can read only the summary instead of reading the whole document. Automatic text summarization tools can be helpful for deaf people by reducing the time and efforts needed to read and understand the documents. Moreover, many deaf people follow sign language which is a lot easier than the written word. Web accessibility can be improved by integrating text summarization tools with sign language to provide better access for the deaf.

This paper aims to develop a new tool to summarize the online articles in order to produce a summary that contains the main points of the original articles. And interpretation of these summarized articles to sign language by using 3D avatar technology.

The proposed tool will be responsible for two major tasks. The first task is to build Automatic Arabic text summarization tool to summarize certain Arabic articles to create shorter versions of these articles. There are two types of summarization extractive and abstractive. Extractive by selecting important pieces from the original documents to produce shorter form. An Abstractive summarization generates summaries whose material is not presented in the input text by understanding of the main concepts in a document and then use new words that are not present in the original document to express these concepts. Abstraction methods are still difficult to achieve because it requires text analysis, modeling, and language generation techniques [40].

There are different methods to implement extractive summaries. The most important ones are: the linear methods that give a score for each sentence depending on heuristic measures, such as Naive Bayesian classifier, Latent Semantic Analysis (LSA) which is inspired by latent semantic indexing. And a Graph Based methods that represent the sentences in the text as nodes with a directed graph; the two sentences are connected if they share some words may they refer to some similarity [40].

The second task of the proposed tool is to translate the output summary from the previous step to Arabic sign language using avatars technology. Deaf people use sign language as their first language to communicate with others [10]. Sign language relies on facial expressions and body language to convey meaning, rather than transferred to the acoustically sound patterns.

The proposed tool will include a dictionary of the most used words in Arabic with their corresponding sign files. Visualization of signs by using computers have been evolved from pictures to video clips and finally to 3D Avatar. We use 3D Avatar because sign files are very smaller than videos and pictures; thus the amount of storage space required is minimal. Also, the possibility of moving the 3D avatar at different angles making it easy for the deaf to understand the signal [8].

The main objective of this paper is to improve web accessibility for deaf people by developing a tool facilitating communication through the Web for deaf. This tool will summarize Arabic articles and translate the output summary to sign language. Thus will help deaf people to convey the meaning of these articles in less time and efforts.

## **2. RELATED WORK**

This part presents a review of previous studies related to web accessibility for deaf people, automatic Text summarization and translating Arabic Text to ArSL:

### **2.1 Web Accessibility**

An attempt in [27] for making e-learning available to persons with disabilities. It presented a Web accessibility framework for the Arab disabled users and facilitated their learning. The proposed system depends on extracting the website content via reading; then the derived content is presented in ArSL animations for the deaf users or speech for the blind users. In order to translate the extracted material the system use ArSL animated library which consists of each Arabic word and the corresponding ArSL animated GIFs. This system implemented in PHP language and MySQL. Improving web accessibility for the deaf people is relatively possible and inexpensive. The best solution is the use of closed captioning and transcriptions, whenever a sound appears. Web accessibility can also be improved with the utilization of the sign language. Most people who are deaf understand and use the sign language which is easier than the written word. Unfortunately, implementation of sign language is expensive because it requires a sign interpreter. In order to help the deaf user to understand the audio material. Some websites involve captioning within the video or provide a link located near the video or audio material to access the captions and transcriptions.

### **2.2 Automatic Text Summarization:**

R. Dixit and S.Apte [37] proposed a method for text summarization based on fuzzy logic to extract important sentences based on their features as a summary. The fuzzy logic system includes four components: fuzzifier, defuzzifier ,an inference engine, and the fuzzy knowledge base. In the fuzzifier, inputs are translated into linguistic values by using a membership function. Then, the inference engine uses fuzzy IF-THEN rules to derive the linguistic values. In the final step, the output variables from the inference engine are converted to the final values by the defuzzifier using a membership function for representing the final sentence score. Then all the sentences of a text are ranked in a descending order according to their scores. Sentences that have the highest score are extracted as a text summary. G. Al\_Gaphari et al.[16] presented and discussed a graph based centrality algorithm for automatic Arabic text summarization. The graph-based algorithm extracts the most important sentences in a document by computing the similarity between two sentences and assessing the centrality of each sentence in a cluster. The researchers use a real-world dataset for testing and validating the software summarizer performance, and the result was very promising. They concluded that the proposed summarizer may outperforms other summarizers used for text summarization. Sobh et al. [40] Integrates Naive Bayesian and Genetic Programming (GP) classification methods in an optimized way to extract the summary. The process of automatic summarization is becoming very important with an increased number of information sources available on the web. The proposed system use Naive Bayesian classifier and Genetic Programming classifier in parallel. Naive Bayesian Classifier classifies each sentence based on its features such as Sentence length, position in the paragraph, or Sentence similarity. Genetic Programming Classifier is automated learning of computer programs. GP was used in many fields, for example, financial market, image processing, signal processing, and pattern recognition. Authors choose to use the Discipulus1 GP system which is the world's first and fastest Genetic Programming system that writes computer programs automatically in Java, C, and

Intel assembler code. They used the downloadable free version with default settings for cross-over when running the tool for classification. Authors concluded that by integrating both classifiers, the union for integration increases the result summary and intersection for integration decreases the summary size. Finally, the proposed system is optimized and able to produce summaries comparable to human-generated summaries.

Many systems have been investigated in the area of automatic text summarization, and some of these are available online, and others are used as commercial software. In this section, we will list some of these systems:

- SweSum: It is automatic text summarizer for Swedish by Martin Hassel and Hercules Dalianis. It summarizes Swedish news in HTML/text format on the web. This system calculates the frequency of the keywords, in which sentences they appeared, based on statistical and heuristic methods .
- WebSummarizer: It is an application to summarize web pages and documents to produce interactive outlines and Visual Summaries. It also offers summarization of Wikipedia articles and supports many languages such as German ,French and Spanish.
- TextCompactor: This summarizer calculates the frequency of each word in the text. Then, each sentence score is calculated based on the frequency count of the words it contains. The most important sentence is considered to be a sentence with the highest frequency count.

### **2.2.1 Arabic Text Summarization Systems:**

- The Summarizer of Aramedia: This Summarizer used for summarizing Arabic and English texts. It extracts the important sentences based on linguistic analysis of the input text. The user can preview these ideas instead of reading the whole document .
- Sakhr Summarizer: Sakhr is a commercial online Arabic text summarization system. This Summarizer identifies the most relevant sentences within a text based on a set of text-mining tools. The Summarization engine applies the Sakhr Corrector to correct the input Arabic text automatically from the common Arabic errors, and the Keywords Extractor to identify a list of important keywords that to identify the important sentences [40].
- SARA© : Is an automatic Arabic text summarization developed by RDI ( Research & Development International). It extracts the most important sentences based on the morphological analysis of the words and mathematical analysis of the sentences and their relationships. The user of the system can easily control the size of the summary (e.g., 20% or 50%, and so on).And the system is valid for all types of texts such as news, science, arts, sports, etc.

### **2.3 Text To Sign Language**

Wraya, et. al, [4] presented a system designed to support transactions between a post office clerk and a deaf customer. This interactive translates the clerk's speech into British Sign Language (BSL) and displays the signs on the screen, using a specially-developed avatar. However, the

system is limited to post office operations. The system operates through a series of stages, start when a deaf person enters the post office and requests the service using gestures. Once the clerk know that the customer is deaf he/she start speaks into his/her headset microphone. Then the automatic speech recognizer selects possible messages from an extensive set of pre-entered targets and displays those with the highest probability on a screen. The clerk confirms the message intended. Then the message contents, are translated into a set of sign instructions and displayed on a video screen. The authors recognized that the system potentially offers access to a range of information that they may previously have had difficulty obtaining.. The developed system achieved an accuracy of identification of the signed phrases of 61% for complete phrases and 81% for sign units. Email is a convenient communication tool for most of the people. Unfortunately, deaf people cannot take the full advantages of these techniques. Non-deaf people are easy to use Email to communicate with their family, friends or other people. However, deaf people have difficulties to study the nature language. They have a hard time to write a proper sentence or even a short mail because they have limited vocabulary capabilities and weak language skills. When they read a sentence, they do not understand the whole sentence properly. Authors proposed a method to convert Chinese texts to a speech and sign language simultaneously. With this proposed system, a deaf people can easily read an Email or articles by a translated sign language stream [45]. M. JEMNI and O. ELGHOUL, [26] Developed a Web-based interpreter of Sign Language (SL). This tool facilitates the communication between hearing people who do not know sign language and deaf individuals. Authors discussed three main approaches: • Writing based systems The first transcription of signs languages based on writing or drawing symbols. Many systems of transcription appeared like HamNoSys and SignWriting. These systems used to encode sign language in a linear way.

### 2.3.1 Video-Based Systems

The video-based systems are a video sequence corresponding to a human interpreter. This video sequence has many limitations in performance, the contents are not easily reusable, and therefore a human translation in a language of signs is necessary and remains very costly [26].

### 2.3.2 Avatar based systems

Avatar based systems this new technological context, involve the creation of a virtual character. Avatar is more precisely and more interested in communication systems in sign language by animation of virtual 3D persons. The proposed tool is called WebSign it based on the technology of avatar. The input of the tool is a text in natural language, and the output is a real-time interpretation in sign language. The tool uses a dictionary of word and signs to perform the interpretation. The first version of Web Sign tool is finalized, and the authors attempt to contacts with specialized deaf associations to train users and evaluate the system [26].

M. Mohandes [28] discussed the importance of Sign Language for deaf people as their natural way of communication just like a speech for non-deaf people. Deaf people often communicate with non-deaf people through a sign language interpreter. Sign language interpreter is not always available when needed. Therefore, there is an increasing need to provide an automatic interpreter for translating signs to speech and the reverse using the modern technology. He proposed a system that converts Arabic text to Arabic sign language. The system was developed using Java Server Pages (JSP), implicitly having all the benefits of the Java programming language. The system has a database to store Arabic words with the corresponding signs representation. When the word is typed a letter by letter, the system automatically starts searching for the word if the

word is available in the database then the recorded clip will be shown. If the word does not have a corresponding sign in the dictionary, then finger spelling is done by displaying images that representing each letter of the word. The author concluded that this proposed system can be used to teach deaf or their families the Arabic Sign Language.

### **3. WEB ACCESSIBILITY AND DEAF PEOPLE**

Deafness has different meanings, but those meanings are relevant in a way or another; however, different definitions of deafness have been given; the Wikipedia, defined deafness as following:

*"Hearing loss, deafness, hard of hearing, or hearing impairment is a partial or total inability to hear. It affects the development of language for children and for adults it can cause work-related difficulties".*

Deaf people are people who suffer from a weakness or loss of hearing. Deaf people are facing many difficulties when communicating with their society. The main solution to help deaf people is the use of Sign language. Sing language provides an easy way of communication with each other's and it is the easiest way to learn for any deaf. When they want to interact with normal people, they need a translator who understands both languages.

Deaf people use the internet, like any other non-deaf people. While the Web provides a large amount of information and resources in all fields of knowledge, it also poses difficulties when content is not designed to be accessible. Deaf users cannot take full advantage of the web because of their hearing disability. Because they are deaf they cannot listen to any video; they cannot listen to any audio. For example, if deaf users want to use YouTube they cannot listen to any of video of audio material which causes boringness in the film. Deaf people cannot take full advantage of the web they cannot even listen to any educational video to increase their knowledge. Some users with a minor hearing disability can be treated and helped by using some mechanical instruments to get them hear full sound. But such devices are expensive. Another difficulty is that many deaf people do not have highly developed reading skills. Because they have a limited reading vocabulary. So, the access to the information on web pages will be difficult.

Some web pages are designed to be accessible for all users even for people with disabilities by providing content suitable for all type of users this is called web accessibility. Web accessibility can be defined as that people with disabilities can use the Web, i.e., can perceive, understand, and interact with the Web. When web pages are accurately designed or made, developed and modified then all users or clients have equal access to information and functions.

Most Web sites have accessibility obstacles that make it difficult for many people with disabilities to use the Web sites. Here some types of web accessibility obstacles that deaf people commonly encounter from poorly designed websites and web tools.

- Lack of sign language to supplement the information and text that is difficult to read.
- Web sites that offer Audio content, such as videos with voices, without captions or transcripts.

- Web sites that include captions or transcripts for the Audio or video contents but do not present options to adjust the text size and colors for captions.
- Web sites that offer services, including web applications, which rely on interaction using voice only.

#### **4. ARABIC SIGN LANGUAGE**

Sign Language (SL) is an essential communication tool used among deaf people. Sign language is non-voice communication used by people with special needs impaired (deaf) or audio (dumb). Sign language relies on facial expressions and body language to convey meaning, rather than transferred to the acoustically sound patterns [5]. Sign language primary components consist of Manual Features (MF) and Non-Manual Features (NMFs) .Manual Features are signs that are presented by one or both hands in different shapes, movements , orientations and locations. While NMFs are features that do not require hands, it consist of the several facial expressions, head tilting, mouthing, shoulder raising, , and similar signals that we add to our hand signs to create meaning , feeling and/or represent the morphological and syntactic markers of a sentence [5].

As an independent language, Arabic sign language has its own rules. ArSL is limited to represent nouns, verbs, and adjectives. Its sentences have particular characteristics can be illustrated as follow:

- Names: There is no exact representation of the personal name in sign language. Instead, deaf people depend on a special mark of a person to describe that person such as a big nose, his job ...etc. Also, finger spelling is used to spell out the name of a person or any nouns .
- Verbs: are used in a simple form without specifying the time (past, present, future). Verbs are attached with words that indicate the time of occurrence (tomorrow, yesterday, now...etc) .
- Adjectives: This is one of the most important parts of any sentence. It is used in the Arabic language to simplify the structure of the sentences.
- Plural: are represented as a singular attached with the sign of number that indicates the quantity .For example, the word (كتابان) are expressed in sign language by two signs one for the word (كتاب) and the other for the sign of the number two .
- Prepositions and adverbs: are expressed in the context of articulation by specifying movements, orientations, and locations [5].

#### **5. TOOL DESIGN AND IMPLEMENTATION**

In this section, we describe the design of our tool that summarize Arabic text and translate it to Arabic sign language.

## 5.1 Tool Design

The proposed tool architecture is illustrated in Figure 1. The tool is divided into two main phases, the key terms summarizer and the process of converting the output summary to sign language.

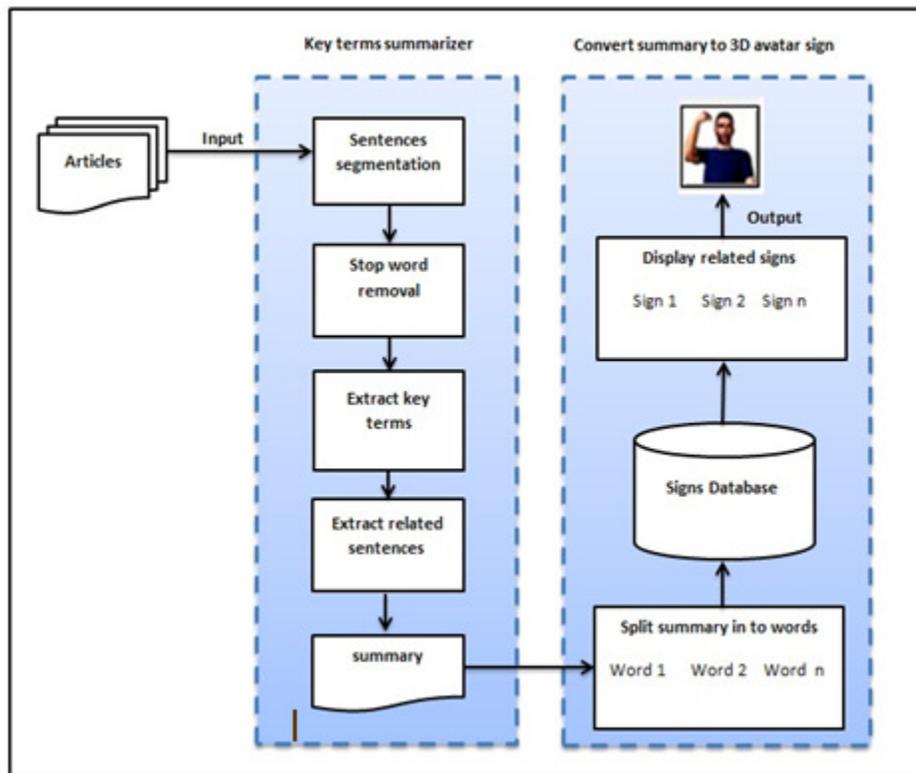


Figure 1: Proposed Tool Architecture

The process goes into several steps:

- Key Terms summarizes read the text and then identify the boundaries of the sentences.
- Remove stop words and extract the Key terms which is the most frequent terms in the article.
- Extract the summary by extracting only the shortest sentence between the sentences belonging to the same key term.
- Convert the extracted summary into sign language by phrasing the summary into words.
- Search the signs data base for the related signs.
- Display the signs in the web page as animation by the signing avatar.

## 5.2 Tool Implementation

The tool was implemented using PHP language to integrate the tool components. In this subsection we will talk about our tool development phases: 1) summarizer building; 2) signs building; 3) Database (DB) building; 3) user interface design.

### 5.2.1 Summarizer building

#### 5.2.1.1 Read TXT File

In this phase we will use PHP programming language to read an Arabic file into our tool, and then we will apply our approach on it.

#### 5.2.1.2 Sentence Segmentation

Segmentation is the process of identifying the boundaries of Segmentation is the process of determining the boundaries of the sentences. In Arabic language punctuation marks are used to divide the sentences which determine the end of each sentence, such as question marks (?), exclamation marks (!), commas (,), semicolons (;), colons (:), and periods (.), are chosen to split the text into sentences.

#### 5.2.1.3 Stop Word Removal

This algorithm will remove the Arabic stop words which consist of the following word categories: Adverbs, Prepositions, Pronouns, Conditional Pronouns, Interrogative Pronouns, Referral Names/ Determiners, Relative Pronouns, Transformers (verbs, letters), Verbal Pronouns

#### 5.2.1.4 Key terms extraction

The algorithm will extract the Key term which is the most frequent term in the text, by scanning the text for the terms and calculate the key terms frequency in the original text by applying this equation:

$$\text{Key term rank} = \Sigma (\text{Key term Frequency})$$

Only one sentence of the sentences that are belonging to one Key term will be extracted .In order to avoid redundant selection of the sentences the length for each sentence will be calculated and give a rank for each sentence, then extract only the shortest sentence between the sentences belonging to the same Key term to be in the output summary. The length of sentence will be calculated using this formula :

$$\text{Length Of Sentence} = \Sigma (\text{Count Of All Terms In Sentence})$$

### 5.2.2 Algorithm

For our tool we will use the following algorithm:

1. Load text file

2. Make preprocessing phases:

2.1 Sentences segmentation

2.2 Stop word removal

3. Extract and rank Key terms

(The most frequent terms in the text)

4. Extract the summary

(The sentences which are referring to one Key terms)

If (two sentences belong to the same Key term)

- Calculate the length of every sentence

(Number of terms in the sentence)

- Extract the sentence with the shortest length

5- Display the summary.

### **5.2.3 Signs building**

For the representation of signs we used a 3D animated virtual agent, the avatar (VGuido) developed by Televirtual Ltd. Also we used HamNoSys to generate the animations from the descriptions of signs. HamNoSys is a system for writing signed languages was developed by Hamburg University. Since computers cannot process the symbols of HamNoSys, SiGML, was used to translate HamNoSys symbols into SiGML which is a modified version of XML. This information is processed in a synthesizer, which, taking into account the description of the geometry of the animated agent. Finally, the description of the signs is represented by the virtual avatar .

Also we use eSign software which is an Editor tool, designed in the project eSign provides an editing environment .In eSign editor document window, the signed text is represented by sequences of glosses. Glosses are the labels for signs. Frist, we select one technical article and after summarizing it using previous step .Then we synthesis the words that are presented in that summary by achieving the following steps:

1- We created a gloss for each word and then from HamNoSys window we enter the notations of manual and non-manual movements as needed (Figure 2).





Figure 3: The Proposed Tool Interface

## 6. CONCLUSIONS

In this paper, a tool to help deaf people to get better access to the web was proposed. The proposed tool is composed of two main phases; summarizing and translating phase. For summarizing phase we select key terms summarizer which produce a summary. This summarizers show promising result. Second phase include the creating of Arabic signs using 3D avatar technology and store these signs in the database. Deaf people have difficulties in reading skills for that reasons our tool provide a simple user interface with clear icons. Moreover, additional improvements can be applied for this tool in the future upon the needs of deaf people.

## ACKNOWLEDGEMENTS

The authors would like to thank all people who contributed in make this accomplishment.

## REFERENCES

- [1] Aarti Patil, Komal Pharande, Dipali Nale, and Roshani Agrawal. "Automatic text summarization." *International Journal of Computer Applications*, (2015):17.
- [2] Abdel-fattah, Mahmoud A. "Arabic sign language: a perspective." *Journal of Deaf Studies and Deaf Education* (2005): 212-221.
- [3] Ahmed Ibrahim, Tarek Elghazaly, Mervat Gheith. "A Novel Arabic Text Summarization Model Based on Rhetorical Structure Theory and Vector Space Model." *International Journal of Computational Linguistics and Natural Language Processing*, (2013):480-485.
- [4] Alison Wray, Stephen Cox, Mike Lincoln, and Judy Tryggvason. "A formulaic approach to translation at the post office: reading the signs." *Language & Communication*, (2004):59-75.
- [5] Ameera Almasoud, and Hend Al-Khalifa. "SemSignWriting: A Proposed Semantic System for Arabic Text-to-SignWriting Translation." *JSEA*, (2012) :604-612.
- [6] Ani Nenkova, Kathleen McKeown. "Automatic summarization." *Foundations and Trends® in Information Retrieval*, (2011):103-233.

- [7] Aqil Azmi, and Suha Al-Thanyyan. "A text summarizer for Arabic." *Computer Speech & Language*, (2012):260-273.
- [8] Basmah faraj, Wafa AL-rajhi, and Yahya elhadj. "Avatar Based Approach for Teaching Arabic Sign Language." *J. of Commun. & Compute. Eng*, (2012) : 43.
- [9] Chang, Jae-Young ,KIM, and Il-Mi. " Research Trends on Graph-Based Text Mining." *International Journal of Software Engineering and Its Applications*, (2014):147-156.
- [10] Daan Hermans ,Harry Knoors ,Ellen Ormel ,and Ludo Verhoeven. "The Relationship Between the Reading and Signing Skills of Deaf Children in Bilingual Education Programs." *Journal of Deaf Studies and Deaf Education*, (2008):518-530.
- [11] David Inouye, Jugal K. Kalita."Comparing twitter summarization algorithms for multiple post summaries. "In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, (2011):298-306 .
- [12] Dipanjan Das, Andr\_e F.T. Martins." A survey on automatic text summarization." *Literature Survey for the Language and Statistics II course at CMU* ,(2007): 192-195.
- [13] Erkan, Radev ." LexPageRank: Prestige in Multi-Document Text Summarization." In *EMNLP*, (2004) :365-371.
- [14] Fatima AL-Khawaldeh, and Venus Samawi ."Lexical cohesion and entailment based segmentation for arabic text summarization (lceas). " *The World of Computer Science and Information Technology Journal (WSCIT)*, (2015):51-60.
- [15] Ganesan, K., Zhai, C., & Han, J. . "Opinosis: a graph-based approach to abstractive summarization of highly redundant opinions." In *Proceedings of the 23rd international conference on computational linguistics*,(2010):340-348.
- [16] Ghaleb Al\_Gaphari, Fadl M. Ba-Alwi, and Aimen Moharram ."Text Summarization using Centrality Concept." *International Journal of Computer Applications*,( 2013):5-12.
- [17] Ghanem, S. and Albidewi, I. (2013)." An Avatar Based Natural Arabic Sign Language Generation System for Deaf People." *Computer Engineering and Intelligent Systems*, (2013):2222-1719.
- [18] Hongyuan Zha."Generic summarization and keyphrase extraction using mutual reinforcement principle and sentence clustering." In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval* ,(2002):113-120.
- [19] Inderjeet Mani ."Summarization Evaluation: An Overview." In *Proceedings of the NTCIR Workshop* ,(2001).
- [20] Judith D. Schlesinger, Dianne P. O'Leary, John M. Conroy." Arabic/English Multi-document Summarization with CLASSY—The Past and the Future." *International Conference on Intelligent Text Processing and Computational Linguistics*, (2008):568-581.
- [21] Martin Hassel. "Evaluation of automatic text summarization." *Licentiate Thesis, Stockholm, Sweden*, (2004):1-75.
- [22] Marty Bray , David Pugalee , Claudia P. Flowers ,and Bob Algozzine. "Accessibility of Middle Schools' Web Sites for Students with Disabilities." *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, (2007):169-178.

- [23] Matjaž Debevc, Primož Kosec, Andreas Holzinger. "Improving multimodal web accessibility for deaf people: sign language interpreter module." *Multimedia Tools and Applications*, (2011):181-199.
- [24] Mehrez Boulares, Mohamed Jemni. "Toward an example-based machine translation from written text to ASL using virtual agent animation." *arXiv preprint (2012):1203.3023*.
- [25] Mine Berker. "Using genetic algorithms with lexical chains for automatic text summarization. 2011. PhD Thesis. Bogaziçi University.
- [26] Mohamed Jemni, Oussama Elghoul. "Towards Web-Based automatic interpretation of written text to Sign Language." In *First International conference on ICT & Accessibility*, (2007).
- [27] Mohammed Abo El-Soud, A.E. Hassan, M. S. Kandil, and Samaa M. Shohieb. "A Proposed Web Based Framework E-Learning and Dictionary System for Deaf Arab Students." *International Journal of Electrical & Computer Sciences IJECS-IJENS*, (2010):24-50.
- [28] Mohandes. "Automatic translation of Arabic text to Arabic sign language." *AIML Journal*, (2006): 15-19.
- [29] N Dheer, C Kumar. "Automatic Text Summarization: A Detailed Study." *International Journal of Science and Research*. (2016): 429-433.
- [30] Nabil Alami, Mohammed Mekkassi, and Nouredine. "Automatic texts summarization: Current state of the art." *Journal of Asian Scientific Research*, (2015):1.
- [31] Nabil Hewahi, Kathrein Abu Kwaik. "Automatic Arabic Text Summarization System (AATSS) Based on Semantic Features Extraction." *International Journal of Technology Diffusion*, (2012):12-27.
- [32] Peter Turney. "Learning algorithms for keyphrase extraction." *Information retrieval*, (2000):303-336.
- [33] Qasem A. Al-Radaideh, and Mohammad Afif. "Arabic text summarization using aggregate similarity". In the *international Arab conference on information Technology*. (2011).
- [34] Rada Mihalcea. "Graph-based ranking algorithms for sentence extraction, applied to text summarization." In *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions*, (2004):20.
- [35] Rafeeq Al-Hashemi. "Text Summarization Extraction System (TSES) Using Extracted Keywords." *Int. Arab J. e-Technol*. (2012): 164-168.
- [36] Rasim Alguliev, and Ramiz Aliguliyev. "Evolutionary algorithm for extractive text summarization." *Intelligent Information Management* (2009):128-138.
- [37] Rucha S. Dixit, Prof. Dr.S.S.Apte. "Improvement of Text Summarization using Fuzzy Logic Based Method." *IOSRJCE*, (2012):05-10.
- [38] S.M. Halawani and Zaitun A.B. "An avatar based translation system from arabic speech to arabic sign language for deaf people." *International Journal of Information Science and Education*, (2012):13-20.
- [39] Shashank gupta, Anushree Jagrawal, and Neha mathur. "Automatixc Text Summarization using Page Rank and Genetic Algorithm." *Journal of Rajasthan Academy of Physical Sciences*, (2014):171-179.

- [40] Sobh, Darwish. and Fayek."Evaluation Approaches for an Arabic Extractive Generic Text Summarization System." In: proceeding of 2nd International Conference on Arabic Language Resource and Tools. (2009):150-155.
- [41] Thomas Hanke. "HamNoSys-representing sign language data in language resources and language processing contexts." In LREC, (2004).
- [42] Waleed AlSanie, Touir, and Mathkour. "Towards an infrastructure for Arabic text summarization using rhetorical structure theory". Master's thesis, King Saud University, Riyadh.(2005):2-70.
- [43] Yihong Gong , Xin Liu." Generic text summarization using relevance measure and latent semantic analysis." In Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval ,(2001) 19-25.
- [44] Yosra Bouzid, Mohamed Jemni. "An Avatar based approach for automatically interpreting a sign language notation." In: Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on. IEEE, (2013) :92-94.
- [45] Yang, H., Liao, W. and Long Lay, Y. "Text to Sign Language Conversion." ,(2001)":241-249.

## **AUTHORS**

Ahmad A. Alhamed ,Ph.D.  
Assistant Professor  
Information Systems Department.  
College of Computer and Information Sciences.  
King Saud Universtiy, Riyadh, Saudi Arabia.  
aalhamed@ksu.edu.sa

Maha Saas Alqhtani,  
Teacher Assistance,  
Information Systems Department.  
College of Computer and Information Sciences.  
King Khalid Universtiy, Abha, Saudi Arabia .

*INTENTIONAL BLANK*

# CERT STRATEGY TO DEAL WITH PHISHING ATTACKS

Shahrzad Sedaghat

Faculty of Engineering Department, Jahrom University, Jahrom, Iran

## **ABSTRACT**

*Every day, internet thieves employ new ways to obtain personal identity people and get access to their personal information. Phishing is a somehow complex method that has recently been considered by internet thieves. First, the present study aims to explain phishing, and why an organization should deal with it and its challenges of providing. In addition, different kinds of this attack and classification of security approaches for organizational and lay users are addressed in this article. Finally, the CERT strategy – which relies on three principles of informing, supporting and helping- is presented to deal with phishing and studying some anti-phishing.*

## **KEYWORDS**

*Phishing, CERT Center, Anti-phishing*

## **1. INTRODUCTION**

Phishing is a kind of internet fraud which is defined as an attempt to steal people's sensitive information such as username, password, and credit card information by foisting itself as a trusted website. This attack is a sample of social engineering technique, in which the key to success is the power to gain individuals' trust. Further, it is obvious that attackers welcome anything that can help them to be appeared as legal and legitimate. These attacks often begin with an email trying to seem legal to the victim and direct him to a fake website or get him to send his information. By sending an email to numerous victims, phishers hope to trap a few of them and their success rate increases statistically [1].

Since a wide range of users is the primary purpose of a phisher, and many of these victims are not informed enough to recognize and deal with such attacks, informing organizations and internet users can decrease the success rate of many of these attacks. On one hand, phishing aims to obtain and misuse personal information, and the disclosure issue of obtaining such information about the people in each country affords a lot of money for the government. However, given the increasing ignorance of internet users regarding internet fraud and the need for academic and scientific specialized centers to support, the CERT specialized center responsible for informing, helping, supporting vulnerabilities and cyber security incidents, was established to meet specialized needs of society, improve technical engineering knowledge regarding the security of

computer systems, transform the results to society, create innovative technologies to meet the contemporary needs and problems, and provide opportunities for a better future in the field of cyber security. As a centralized organization and with a combination of part-time and full-time employees, this center tries to investigate the security issues and problems to be ready for a quick and proper reaction to problems and challenges of dealing with internet attacks such as phishing.

## **2. IMPORTANCE OF DEALING WITH PHISHING BY GOVERNMENTS AND ORGANIZATIONS**

Governments, military centers, and organizations keep a massive amount of confidential information of their employees and clients, some of which has reduced the security to a product level and by using a simple software product, they can secure the information. Protecting confidential information is a commercial and an ethical and legal need in many cases. The existence of a security issue for the formation of phishing attacks can influence an organization in the following ways:

- Reduced income and increased costs
- Loss of reliability and reputation of an organization
- Disclosure of information to competitors and misusing them
- Disruption of current processes and negative side effects on organization and government activities

Phishing activities can reveal personal information about the clients of an organization. Therefore, it is not pleasant for any organization to lose the trust of its clients and investors by involving in such an issue. In addition, the disclosure of personal information of a client due to the negligence of the organization can result in an individual complaint of the respective organization and even prosecution and bankruptcy.

## **3. CHALLENGES AHEAD IN PROVIDING SECURITY AGAINST PHISHING**

- The possibility of creating complex phishing website and illegally selling them in online markets
- Availability of Key-logger software that can secretly record and collect activities related to computer username and password
- Providing the phishers with sensitive information of their accounts by the clients for a false reward for participating in fake surveys,
- Another reason of increasing phishing is a very high return on investment. Phishing has an incredible return on investment.

### 4. WHAT ARE THE STEPS INVOLVED IN SECURING AGAINST PHISHING ATTACK?

The following four trees show the multiple methods of phishing to achieve success. Moreover, some countermeasures have been considered in these trees by which the victim can prevent from attacking progress. The first tree (Figure 1) shows the steps of phishing that are common in all of its formation methods. Open relay mail servers that allow spammers to hide their identity, and throwaway accounts that make it possible to send an email via temporary accounts, and regular email accounts allowing the users to receive an email in Email client software, provide excellent context to start a phishing attack.

Other trees (Figures 2-4) investigate three general methods of a successful phishing, including Trojan, deceit, and spywares [3].

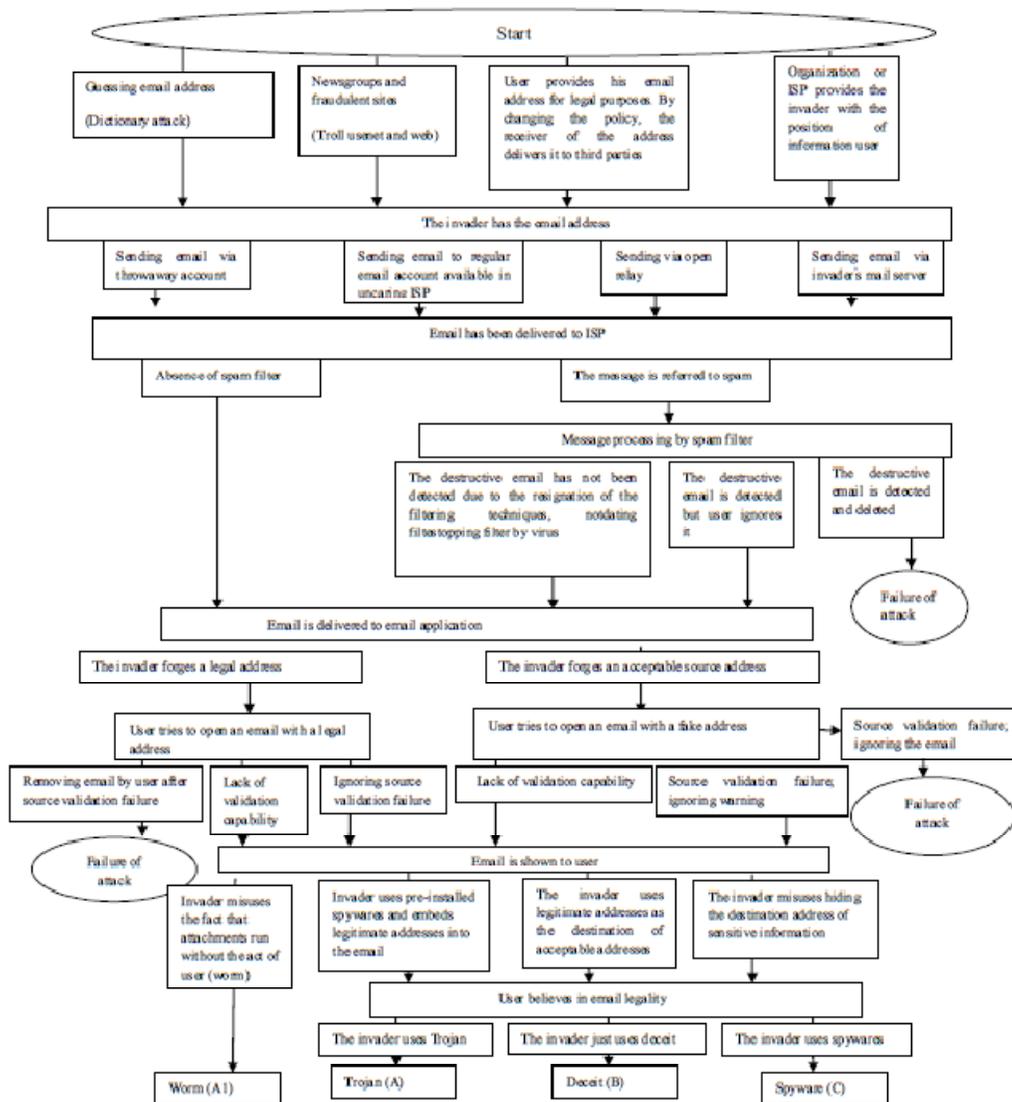


Figure 1- Common sections of phishing attack

Some destructive software may have been attached to the email sent to the victim, that do not behave as he expects and damages his system as soon as he opens them. These destructive codes also spread via pop-up windows. The host-based anti-viruses, anti-spams, and IDSs which register all suspicious movements of the target machine by performing a little service named Agent, and report it to the server via SNMP, have an important role in stopping attacks in this method.

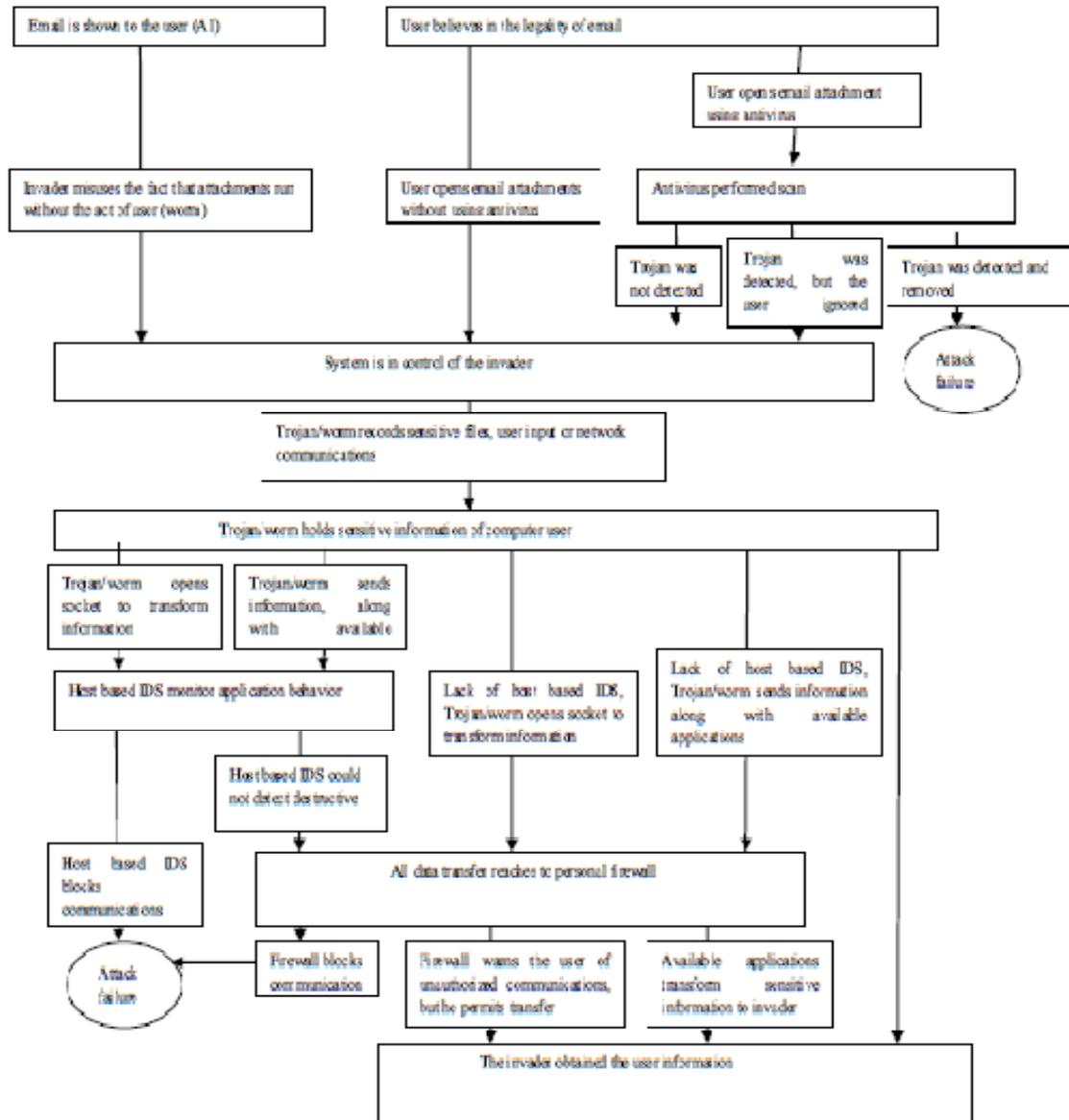


Figure 2- The use of Worm and Trojan in phishing

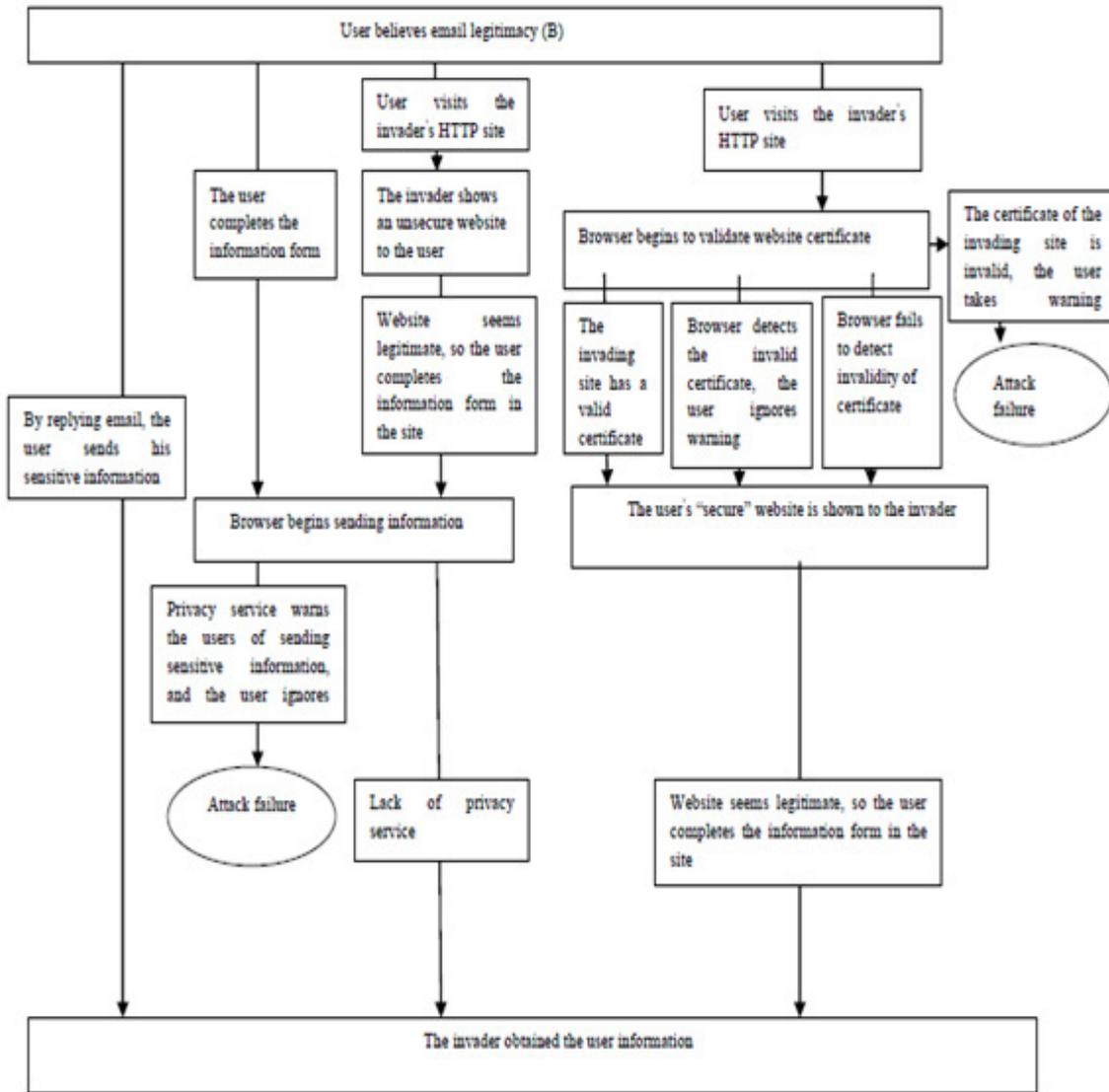


Figure 3- The use of deceit in phishing

Relying on deceit and without misusing any additional software, the attacker continues the attack in this method. If the victim trusts the browsers' warnings of an invalid certificate of the website, and the invading website uses an SSL, some protections are taken place.

In this method, the invader imports a spyware (in the form of a destructive code) to the victim's system. When the victim enters the legal web address intended by the invader, this software activates and directs the victim to the fake website of the invader, or hears the user's communication with a legal website, and provides the invader with the personal information of the victim by any means. Spywares are designed so that they do not perform any demolition or any other action to prevent the victim from suspecting the hearing. Host-based spyware detection specialized software, commercial anti viruses, personal firewalls, and intrusion detection systems can often stop attacks in this type of phishing.

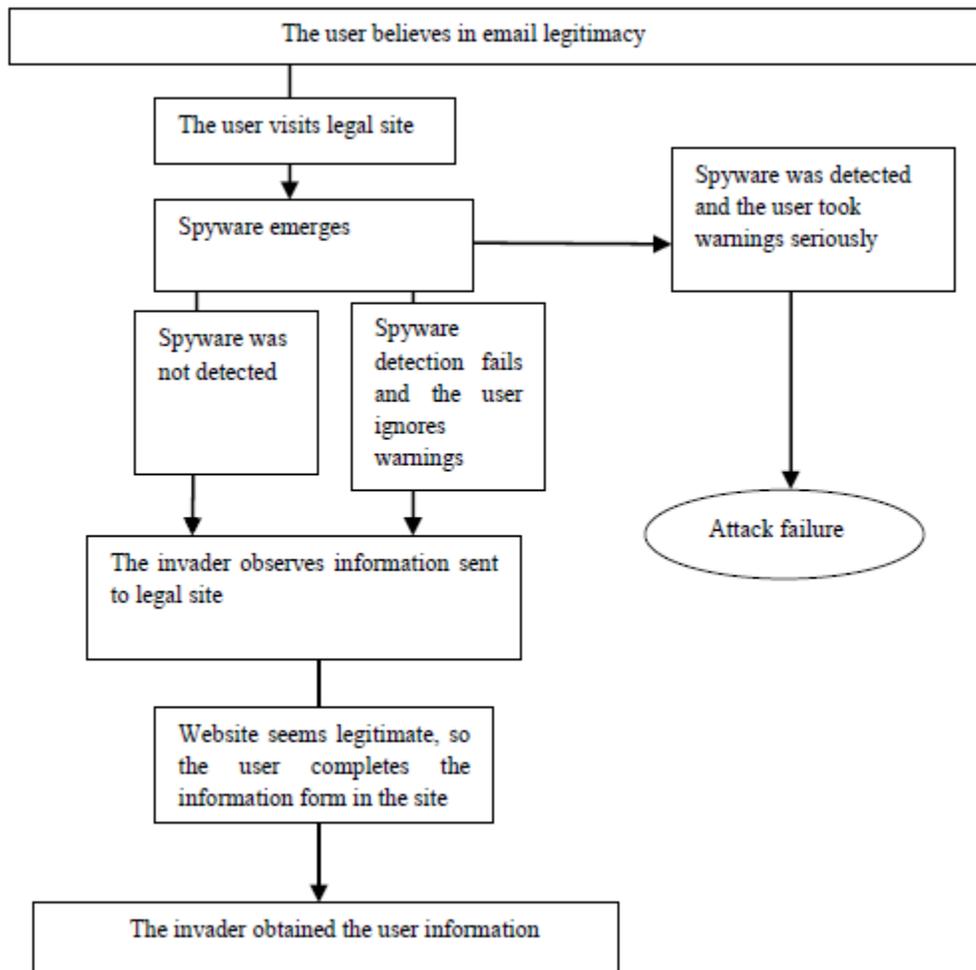


Figure 4- The use of spyware in phishing

## 5. SECURITY APPROACHES TO DEAL WITH PHISHING

Approaches to preventing phishing are generally classified into two groups

### 5-1 Organizational approaches

The primary approach of any organization to deal with any kind of internet attack is to shift the policy and create a relationship between new policies and clients. One of these policies is implementing a method to write the sent emails so that they are not confused with fake emails. In addition, providing users with a solution to validate emails can complement the mentioned policies. Moreover, a victim's information is usually stolen when logging into a website, hence, using a stronger authentication algorithm that does not need entering the sensitive information of the users into organization's website, makes it more difficult for phishers to work. Organizations can provide an additional security layer for their total system. We will explain the key approaches to deal with phishing in organizational environments as follows:

Links that direct the recipient user to a website should not be attached to emails. Although attaching these links provides a simple experience for users, the best secure alternative is culture building the practice of typing and copying the URL into the browser. Organizations should avoid embedding information forms in emails, and never request users for information via such forms. Using such forms makes it difficult for users to recognize the legitimacy of emails. Of course, this approach should be accompanied by educating clients whose emails receiving from the company do not contain any information request form.

Using the Sender Policy Framework (SPF), which is a mechanism to validate the sender of the email in terms of source domain, creates a strong security level for any receiver. This capability is free to use and the steps are such that the server manager should set settings for the organization's mail server in services such as OpenSPF.org. Then, he receives a unique TXT record for his mail server depending on the applied settings and publishes it in the DNS server used by the mail server. Now, the mail server validates all email sender addresses according to the definitions applied to the SPS service and TXT record.

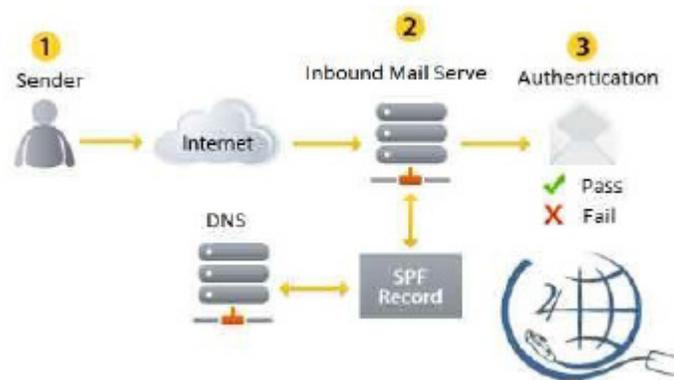


Figure 5- SPF Mechanism

An organization should be able to use the name and photo of the client in writing emails sent for special contracts. Also, it is important to inform clients that every email with a specific content has several contents for validation. In such conditions, the users are able to recognize the limitation of email without installing any additional software or hardware, but the organization is obliged to collect photos of its clients. But, if obtaining the users' photos is impossible or very difficult for an organization, the sequential counting of emails is the best way to create a barrier for phishing. Below is a sample of sequential counting:

- ✓ Date of the current email: Jan. 16, 2017
- ✓ Serial number of the current email: JJH0017
- ✓ Last email we sent to you: JJH0016 on Dec. 10, 2016
- ✓ Serial number of the next email we will send to you: JJH0018

It is worth mentioning that the above approach requires adding serial numbers and an additional security layer to the data base of the company. However, it is still less possible for an organization to prevent phishing, and applying a strong mechanism of authentication can

strengthen this possibility. Users can either employ a token creating a disposable password for each authentication to make phishing almost impossible to happen- since they are not aware of the authentication information and cannot accidentally reveal them, or they can choose another method of authentication. Some of the banks are using a technology similar to that of smart cards of GEMP and MAST, which require a connection to another hardware to be detected. In communicating with any organization, the user should carry a token unique to that organization and pay a separate fee for each token. Therefore, using simulator software of token or mobile phone compatible PDAs is recommended to address these approach issues. The application procedure of PDAs is that they create a Pin to be used on the website upon the request of the user. This approach also requires an organizational training plan for a proper configuration of software by clients. Tokens cannot, by themselves, prevent revealing other identification information such as father name, etc. Internet monitoring of fraudulent and phishing websites by computer security organizations, by means of the development of something like whitelists and with brand and key content confirmation test and publishing it among internet users can decrease the presence of fake websites over the internet. The monitoring services' providers implement an agent-based solution for an ongoing monitoring of web content and active search of all cases such as logo, brand, or key content, for their clients. Anti-virus scanning at a company gateway of HTTP and SMTP related data traffic can reasonably reduce Trojan, Backdoor programs, and Spywares entrance. This is an easy and quick approach for organizations so that, instead of updating all desktop scanners, they can just update the gateway scanner; however, these two scanners should come together as part of the defense layer strategy against destructive codes since some of encrypted traffics must be scanned at the desktop and mobile phone users are not always protected by gateway scanner while away from their office. With the possibility of the existing phishing websites, organizations should block their access to the internal network by filtering the content and configuring rules blocking destructive URLs in the gateway or using anti-spams. Anti-spams delete or isolate spams in one of the methods of refinement marking of email subject although spammers always change their spamming techniques and no solution is absolutely responsive.

## 5-2 Approaches for lay users

Being suspicious and acting conservatively is regarded as the users' main approach to deal with phishing when entering sensitive information over the internet for any reason (for example, typing the address of any website into the browser and not using links) [4]. If a phishing is powerful enough that users cannot visually recognize the legitimacy of a website or email, there are various software that can help them in this regard. But, it should be noted that this kind of software is not error free, and they may act as a kind of internet attack. Therefore, introducing and instructing buying, installing, maintaining and the way these products work, can ensure a right choice by users. Using anti-spam on desktop and gateway do not allow an attack on spam review by the user. Gateway is for the users of the internet (ISP) and email (MSP) provider companies. There are various methods to inject anti-spam filters into the email processing cycle. One of the three-layer methods involve the following steps:

- ✓ Row 1: installing filter on ISP and MSP for all email client and users
- ✓ Row 2: installing filter for all users on the local LAN network
- ✓ Row3: installing filter on the personal desktop by means of installing security applications

The above-mentioned approach requires the cooperation of ISP and MSP. However, according to the statistics, most internet users employ anti-viruses that are mostly run in the background and do not have any effect on their activities. Hence, along anti-viruses, anti-spyware can help to prevent potential spywares. File signatures must be updated regularly so that the above-mentioned approach can prevent attacks. Anti-spyware users should be aware that these kinds of software may also remove spywares that are the prerequisite of legal software performance. There are commercial software packs that can monitor the output traffic of the web- for a series of data definable by the user, which provide the users with desktop privacy services. In such situations, if the output traffic includes identification information of the user, the data are not be exported without user confirmation, even if the destination URL is not clear and it helps to prevent all social engineering attacks. The challenge confronting such an approach is to detect sensitive information and putting them in the Permit List and Deny List for specific websites.

## **6. DEALING WITH PHISHING BASED ON THE CERT MISSION**

Electronic transactions and online banking are mainly designed for consumer convenience and reducing the cost of banking operations. But, a sharp increase in the electronic banking crimes has weakened the success of this plan and customers prefer physical presence to perform a secure transaction. Opponents of online transactions argue that phishing threats %86 of home users and %14 of home investors. At the global level, about 3000 phishing has been reported per month, more than %80 of which have to do with financial institutions [2]. The existence of security holes in the configuration of organizations' servers and not being aware of them, provides the context for phishing attacks. However, these types of attacks are designed so that reducing the sustained damages is almost impossible after their occurrence.

The CERT center is obliged to analyze different kinds of phishing and accordingly, present protective solutions based on the environmental variables of an organization, at the proper time and before attack occurrence. Coping with phishing can be based on three principles, based on CERT mission.

### **6-1. Informing**

The main reason of phishing occurrence is the ignorance of internet users, and most approaches to cope with it require a training plan; hence, informing is regarded as the most important principle to fight against phishing. CERT center should include an educational campaign entitled "mechanisms for secure communications" in its organizational agendas, where it explained different types of phishing first and introduced all anti-phishing software and their method of updating as well as how to use software and hardware tokens to create disposable passwords. It should be noted that people who are in charge of providing electronic activities in banks and organizations are in priority as the target of these training courses, because training should be done in such a way that that security approaches be spread in a network manner by these individuals, and even naïve users also stay safe against such attacks.

In addition, CERT centers across the country should coordinate and create a common database to recognize and list fake websites and share it among the internet users so that professional users help to complement this database in addition to being aware of such addresses.

## 6-2. Support

After identifying the kind of phishing, the CERT center should analyze its reasons for an occurrence such as the unfamiliarity of internet users with strong technologies and algorithm of encryption. Based on the studies conducted in CERT center, nearly %80 of lay users cannot distinguish a phishing email from a safe and correct email. A kind of phishing attack which is regarded so destructive and insensible is that a third party hears them without the knowledge of sender or receiver during information transfer.

By using PGP (Pretty Good Privacy) software to digitally sign your messages, we can make sure that only the intended receiver will receive your message. PGP works based on encrypting the public key and has a pair of keys. To receive a private email from an individual, you should first provide that individual with the public key- which is an easily transferable block of text you have created using this software- and keep the private key for yourself. Know the individual encrypts his message using your public key, signs it with his private key and then send it to yourself. Now, after receiving email and using the sender's public key which you have been already provided with, you can make sure that it is legitimate and no change has been made on its text and decrypt it. It is obvious that if you are the sender of the email, all these processes will happen in reverse. Another capability of PGP software is to encrypt and decrypt the information of clipboard memory by which you can ensure the accuracy of email by copying and pasting links attached to.

S/MIME (Secure/ Multipurpose Internet Mail Extensions) is a universal standard which is used to encrypt emails. Using this service requires a third party license for both public uses over the internet and personal uses in private organizations. When you send email using this service, your email is, in fact, signed digitally and when it is received, the signature is reconciled with sender's information to ensure the identity of the sender. This service also allows for an optional encryption using public keys which are a synthesis of your digital license and password.

## 6-3. Computer help

Conducting a survey and collecting all causes of occurring a phishing attack, the CERT center should try to remove similar security weaknesses to prevent the occurrence of phishing via conventional methods. Because they are completely new, testing anti-phishing software by internet users has a high rate of risk, and some of these kinds of software detect phishing attack after their occurrence and when it is too late. The last version of popular browsers of Mozilla and Internet Explorer include a proper anti-phishing capability. Phishtank (phishtank.com) argues that it can detect phishing websites and offer solutions to fight against them, where users list phishing websites themselves and add them to their databases. The CERT center should test all available anti-phishing software in a comprehensive plan, and introduce those having higher success rate during the test. Some of these kinds of toolbar anti-phishing software are as follows: Windows Live Toolbar is available in IE7, is installed on IE6, and is produced by Microsoft having blocked %51 of phishing websites so far.

- Google Toolbar is a Google product, is available on Firefox, and has blocked %85 of phishing websites by now. This toolbar which is linked to Google database is updated every 2 hours.
- NetCraft Toolbar is produced by the English company of NetCraft which can be installed on Firefox and IE and has blocked %95 of phishing websites so far.

- NC Toolbar is produced by Symantec Company which can be installed on IE and has blocked % 76 of fake websites so far. Preventing Key-logger software is of the advantages of this tool.
- eBay Toolbar is an eBay company product, specialized for using this website, enjoys a proper graphical interface. This tool has managed to block %90 of phishing attacks on eBay websites.

## **7. PROPOSED SOLUTIONS FOR CERT CENTERS TO PREVENT PHISHING ATTACK**

Creating a common internet page for banks that contain links to all electronic payment portals is a proper solution to decrease theft through fishing. Based on this plan, a page entitled “Central Bank” is created with a special standard and logo. By entering this page, users can enter the portal of other banks. Creating just a page and logo of banks increases public confidence and in turn, can decrease phishing related crimes to a considerable extent. Now, when we search the name of a bank several portals and pages related to the respected bank open, the existence of multiple pages allows criminals to create phishing pages. It is emphasized that it is always easier to apply all the mentioned security approaches and protective layers on an internet page than to control the security of multiple bank portals over the internet.

## **8. CONCLUSION**

Phishing is different from the first-grade traditional internet fraud, on a scale of deception that can really be committed to sabotaging. This kind of attack via email and World Wide Web that allows fraudulent to obtain thousands and millions of victims per minute without any cost was discussed in the present study. The important principle in phishing is that the invader must continue to gain the trust of the victim for a successful attack.

Since there is no face-to-face communication between the invader and the victim, the clients have little information to decide to work with a legitimate email or website. The technical and ultimate solution to cope with phishing attacks is to make a significant change in internet infrastructure which is beyond the ability of any of the institutions to implement. In the present study, some approaches were proposed by which CERT centers can somehow cover the inability to change the infrastructure to create a protective layer against these attacks, and it is hoped that they are put into practice.

## **REFERENCES**

- [1] J. Zhang, Y. Pan, Z. Wang and B. Liu, "URL Based Gateway Side Phishing Detection Method," IEEE Trustcom/BigDataSE/ISPA, Tianjin, China, pp. 268-275, 2016.
- [2] M. Volkamer, K. Renaud, B. Reinheimer, A. Kunz, "User experiences of TORPEDO: TOoltip-powered phishing email DetectiOn ", Computers & Security, February 2017.
- [3] G. Tally, R. Thomas, T. V. Vleck, "Anti-Phishing: Best Practices for Institutions and Consumers", McAfee Research Technical Report # 04-004, September 2004.

- [4] Y. Zhang, S. Egelman, L. Cranor, J. H. Carnegie, Mellon University “Phinding Phish: Evaluating Anti-Phishing Tools”, Tech Report: CMU-CyLab-06-018, 2006.

## AUTHORS

She received her B.E. and M.S. degree in computer and Information Technology engineering from Yazd University, Iran in 2008 and 2010 respectively and is currently pursuing her Ph.D. in computer engineering in Sharif University of Technology, Iran. In 2011, she joined the department of computer and information technology engineering, Jahrom University, as an instructor. My research interests are computer network, its security and social networking in general.



# EFFICIENT USE OF HYBRID ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM COMBINED WITH NONLINEAR DIMENSION REDUCTION METHOD IN PRODUCTION PROCESSES

Saad Bashir Alvi, Robert Martin and Johannes Gottschling

Mathematics for Engineer, Duisburg-Essen University, Duisburg

## **ABSTRACT**

*This research study proposes a novel method for automatic fault prediction from foundry data introducing the so-called Meta Prediction Function (MPF). Kernel Principal Component Analysis (KPCA) is used for dimension reduction. Different algorithms are used for building the MPF such as Multiple Linear Regression (MLR), Adaptive Neuro Fuzzy Inference System (ANFIS), Support Vector Machine (SVM) and Neural Network (NN). We used classical machine learning methods such as ANFIS, SVM and NN for comparison with our proposed MPF. Our empirical results show that the MPF consistently outperform the classical methods.*

## **KEYWORDS**

*Fuzzy Inference System, ANFIS, Neural Network, Support Vector Machine, KPCA*

## **1. INTRODUCTION**

Automatic fault prediction is an important topic of research in metal industry [1]. Since the beginning of the first industrial revolution, industries are striving to produce fault-free products in least possible amount of time. Customer expectations are ever-increasing in term of quality and availability of products [2-4]. A fault in a product during the production process can be due to a single cause or a combination of causes and industries are still using trial and error methods to minimize them[5][6]. Faults in production can be reduced by analysis of the production process data [7][8]. Data capturing of the production processes for analysis started manually but due to human-errors, the quality of the data captured was compromised. Now, with the automation of data capturing using sensors, the noisy data is getting reduced. With the advancement of the technology and automation, more and more data is available for analysis to optimise the production processes [9].

In the current globalized industry era with emphasis on automated smart industries, the analysis of measured data plays an important role to improve the quality of the products [10]. Large

number of parameters have an influence on the quality of a product. Deviations of process parameters can have negative effect on the production performance. Measuring and evaluating all the appropriate process parameters with a suitable method ensure a consistently high quality and productivity in the automated production environment. The application of Machine Learning (ML) methods in these processes is motivated mainly by two objectives: the *prediction* of quality properties from the measured data and the identification of *key process indicators*, i.e. the process parameters which have the strongest influence on the outcome. These objectives help to improve the quality of the products and to understand the implicit relations among the parameters in the production processes, which in turn result in reducing faulty products during production.

In this work, we will concentrate on the first objective, namely the *prediction* of quality properties from the measured data by suggesting a novel methodology for automatic fault prediction. The learnt methodology will help to reach target *predictions*, which will ensure stability of the production process regulation and repeatability of the process conditions resulting in quality products with minimal scrap. The potential of the proposed methodology is demonstrated by using three actual production datasets.

In our approach, we use three stage process. In first stage, the input data is split into mutually exclusive sets of training, validation and test data. In the second stage, Support Vector Machine (SVM) [11][12] and Neural Network (NN) [13][14] are trained using the original data whereas Adaptive Neuro Fuzzy Inference System (ANFIS) [15] is trained on the data transformed by Kernel Principal Component Analysis (KPCA) [16][17]. In the last stage, we utilize a novel *fusion* method to combine these different prediction functions and obtain a "Meta Prediction Function" (MPF). Results during these stages are also collected and performance is measured. Figure 1 shows major data processing units of proposed framework.

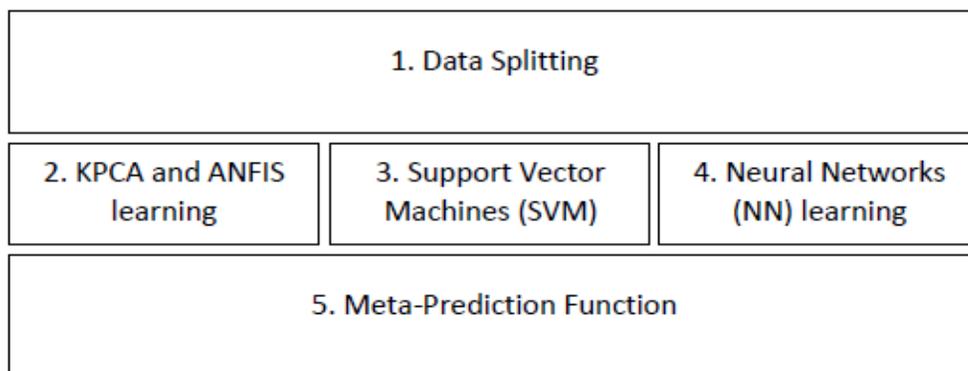


Figure 1 - Major data processing blocks

The performance of the MPF is compared with the performance of the classical machine learning methods such as Neural Networks, SVM and ANFIS.

## 2. RELATED WORK

Lehaet *al* proposed a novel integrated method into a production plan realized in a physics-based realistic simulator. Supervised machine learning techniques namely Model Trees and Neural Networks were integrated. The online learning and on the fly control code modification were allowed by integrating the learning capability into the control process. Averaging was used for

measuring the produced optimization times through learning which outperform times of a production process. [18]

Ashouret *al* proposed a method for automated identification for machined surfaces in manufacturing. Image processing and computer vision technologies were used for automated identification for reduction in inspection time and avoidance of human error due to inconsistency and fatigue. SVM classifier with different kernels were investigated for the categorization of machined surfaces into six machining processes. The gray level histogram was used as discriminating feature. Experimental results suggested that the SVM with the linear kernel outperformed for a dataset consisting of seventy-two workpiece images. [19]

Yuan *et al* proposed a novel method for improving the machining quality of thin walled flexible workpieces. Machining platform was established for thin-walled flexible workpieces. Sparse Bayesian learning based method was used to predict the future deformation. The dual mode predictive controller was developed to reduce the machining vibrations and quality of the workpiece surface was improved. Experiments were performed for thin-walled flexible workpieces and effectiveness of proposed method were demonstrated through machining experiments. [20]

Rostamiet *al* proposed an efficient Equipment Condition Diagnosis (ECD) model for Fault Detection and Classification (FDC). Different machine learning techniques such as SVM, K-Means clustering and Self-Organizing Map (SOM) were used to develop ECD model. Principal Component Analysis (PCA) was used to project the abnormal observations into normal models. Experiments were conducted with industry data. The proposed Model showed effectiveness for classifying the fault fingerprints to give evident guidelines in explaining the detected faults. [21]

Costa *et al* used ANFIS and CART (Classification and Regression Tree) for classifying the defects that occur during the production process for packing of glass. The database project was named as "Newglass". Different operating variables of the furnace and the percentage defects in the end products of the factory model were used in experiments. Finally, ANFIS, CART and glass packing manufacturing results were compared. [22]

### 3. METHODOLOGY

#### 3.1 Support Vector Machine

*Support Vector Machine* (SVM) is a supervised learning method [11][12] which computes a set of hyperplanes in a higher dimensional feature space based on the given data for classification and regression problems.

For each dataset, optimal settings for the SVM method were chosen with experimentation. Input training and validation data was normalized with z-score normalization. Following kernel functions were considered for learning:-

the Polynomial kernel,

$$k(X, Y) = (X^T Y + r)^p, \quad (1)$$

where  $p$  is the power of the polynomial and  $r$  is a shifting parameter;  
the Gaussian kernel,

$$k(X, Y) = e^{-\frac{\|X-Y\|^2}{2\sigma^2}}, \quad (2)$$

where sigma ( $\sigma$ ) is an adjustable parameter;

Sigmoid kernel as shown in equation 3 below,

$$k(X, Y) = \tanh(\gamma X^T Y + r) \quad (3)$$

where rho ( $\gamma$ ) is the scaling parameter of the input data and  $r$  is the shifting parameter controlling threshold of mapping.

The performance of the selected optimal settings was calculated on the test data and is shown in the experiments section. The predictions from SVM were used as input in the MPF described in section 3.5.

### 3.2 Neural Networks

*Neural Networks* (NN) [13][14] are a non-linear supervised learning method based on a network of so-called neurons which are interconnected by weighted links. NN learns by adjusting the weights to optimal values based on the given training and validation data.

In this work, a NN is trained using one of the most popular backpropagation learning algorithm with multilayer perceptron topology. The used feedforward NN consisted of three layers: an input layer, one hidden layer and an output layer. The number of neurons in the input layer is equal to the total number of independent variables and the number of neurons in the output layer is equal to the number of dependent variables, whereas the hidden layer contains eight neurons for all considered datasets. The input layer receives the input from the independent variables and forwards it to all the neurons in the hidden layer. The neurons in the hidden layer apply their activation function to the weighted sum of their inputs and compute an output. The output layer then computes the predicted value for the dependent variable(s).

The back-propagation learning algorithm used here to train the multilayer network consists of two passes. In the forward pass, with randomly selected weights and the input given by the training data, the algorithm produces an output for the dependent variable(s). An error is then calculated based on the difference between predicted and actual output. In the backward pass, this error is propagated backwards through the network from the output layer to the input layer and weights in the network are modified using the *delta rule*. The formula used to calculate the change in weights of hidden and output layer neurons are shown in equation 4 and in equation 5 respectively:-

$$\Delta w_{ij}(p) = \beta \cdot \Delta w_{ij}(p-1) + \alpha \cdot x_i(p) \cdot \delta_j(p) \quad (4)$$

$$\Delta w_{jk}(p) = \beta \cdot \Delta w_{jk}(p-1) + \alpha \cdot y_j(p) \cdot \delta_k(p) \quad (5)$$

where the indices  $i, j, k$  refer to input, hidden and output layers,

$\alpha$  is the learning rate,

$\beta$  is the Momentum with a value between 0 and 1,

$x_i(p), y_j(p)$  are the output of neurons  $i$  in the input layer and  $j$  in the hidden layer at iteration  $p$ ,

$\delta_j(p), \delta_k(p)$  are the error gradients at the neurons  $j$  in hidden layer and  $k$  in the output layer at iteration  $p$ .

During training, we used 1000 epochs. Weights of the NN were optimised using settings of learning rate 0.1 and momentum 0.2. The Unipolar Sigmoid function was used as the activation function.

During the learning phase, training and validation data is normalised. Then NN is trained and the parameter settings which produce best results are preserved and the performance of these settings is calculated on the test data.

The predictions from NN are also used as input in the MPF in section 3.5.

### 3.3 Kernel Principal Component Analysis

*Kernel Principal Component Analysis* (KPCA) is a nonlinear dimension reduction method, introduced by Sholkoptet *al.*[16][17], which maps data from the input space to a lower dimensional feature space while retaining maximum possible variance in the data.

For each dataset, KPCA was applied using the Polynomial kernel and Gaussian kernel, which are defined in equations(1) and (2) respectively. For the Polynomial kernel, settings used were  $p = 1, 2, 3; r = 0$ . For Gaussian kernel, the settings used were  $\sigma = 0.5, 1.0, 1.5$ .

The results are shown for the kernel settings which provided the most accurate prediction results.

### 3.4 Adaptive Neuro-Fuzzy Inference System

*Adaptive Neuro-Fuzzy Inference System* (ANFIS) [15] is a hybrid neuro-fuzzy model which consists of 5 layers. In the first layer, for the inputs, degree of membership of the chosen membership function is computed. In the second layer, firing strength of the rules is calculated using t-norm (product for AND and maximum for OR) operators. In the third layer, computed firing strengths are normalized. Forth layer is the adaptation layer in which rule consequent parameters are computed. Fifth layer is the summation layer which computes sum of all the computed consequents. In learning phase of ANFIS, membership function parameters and rules consequent parameters are optimised.

In our approach, we use the ANFIS, which combines the advantages of fuzzy expert systems with those of classical NN, in combination with KPCA. It has already been successfully employed for data prediction in a variety of fields. KPCA was used to transform the data because for some of the selected parameter settings like input partition method: 'grid partitioning', ANFIS was not immediately applicable to the number of variables as large as 10. In addition, even in the case of small number of variables, it was observed that ANFIS produced comparatively better results for same parameter settings on the transformed data from KPCA in comparison to the original data as input.

Before the training in ANFIS start, a Fuzzy Inference System (FIS) is initialized with information about the number of input variables for the selected dataset, the selected number and type of membership functions and the number of rules. The rules for the FIS were initialized based on the Grid Partitioning (GP) method or Subtractive Clustering (SC) method. In case of GP method,

input membership function type as Gaussian was selected, output membership function type as Linear was selected and in different independent runs for best parameter settings, number of membership functions for input parameters were selected as 2, 3 or 4. For SC method, influence radius was tested with inputs 0.1, 0.2, 0.3, 0.4 and 0.5 for different runs. It was trained using NN backpropagation algorithm and least square methods.

In case of KPCA and ANFIS, the training and validation data was collectively normalised and the corresponding symmetric kernel matrix was computed. The Eigen-values and Eigen-vectors were computed for the computed kernel matrix. Then principal components of the training and validation data were computed using the computed kernel matrix and Eigen-vectors. These principal components along with the corresponding dependent variable values were used to train ANFIS and parameter settings which produced the best result on the validation data were saved. Finally, the performance was calculated on the principal components of the test data. These predictions were also saved for later use in the MPF.

### 3.5 Meta Prediction Function

Our *Meta Prediction Function*(MPF) is inspired by the principle of ensemble methods [23]. The ensemble method provides an optimized prediction by combining the results obtained from the individual learning methods. The main component of our proposed MPF is the combination of KPCA and the accumulator module which is implemented using machine learning methods Multiple Linear Regression (MLR)[24], NN and SVM. The learning algorithm is given below for MPF.

*Step 1:* Apply different ML methods such as SVM, NN and ANFIS as discussed in sections 3.1, 3.2, and 3.4 and save the prediction to be used as input.

*Step 2:* The predictions from the ANFIS, SVM and NN are pre-processed and multicollinearity was removed from them by using KPCA as shown in equation 6

$$[PC1_i, PC2_i, PC3_i] = KPCA(NN_i, SVM_i, ANFIS_i) \quad (6)$$

*Step 3:* The computed principal components obtained from KPCA are used as the independent variables as shown in the equations 7, 8 and 9. These independent variables along with the computed predictions are used to fit the MLR and to find the optimal settings for the NN and SVM.

$$MPF\_value\_ML_{(i)} = \beta_0 + \beta_1 PC1_i + \beta_2 PC2_i + \beta_3 PC3_i + Error \quad (7)$$

$$MPF\_value\_NN_{(i)} = NN\_algo(PC1_i, PC2_i, PC3_i) \quad (8)$$

$$MPF\_value\_SVM_{(i)} = SVM\_algo(PC1_i, PC2_i, PC3_i) \quad (9)$$

Then to compute the performance of the MPF on the test data, independent variables of the test data are also normalized based on the respective mean and standard deviation of training and validation data variables, its kernel matrix is computed and the principal components of the computed kernel matrix are computed using the Eigen-vectors of the training and validation data.

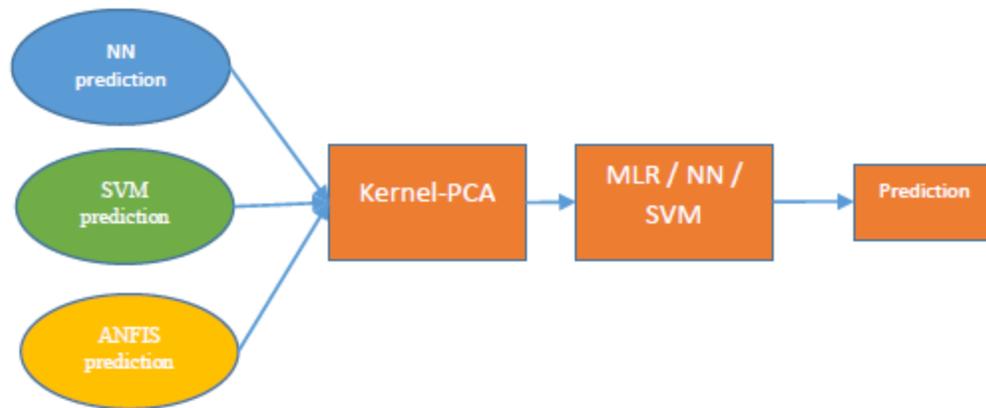


Figure 2 – Proposed Meta Prediction Function (MPF)

To make a comparison with classical machine learning techniques, the test data results of these three MPFs were compared with the best results of the SVM, NN and ANFIS obtained in sections 3.1, 3.2 and 3.4 respectively.

## 4. EXPERIMENTS

In order to evaluate our proposed MPF, we checked its performance on three different datasets related to different fields of actual production data.

The first dataset is related to plastic deformation. It contains 4 variables and 1248 measurements. The dependent variable is the measured flow stress, while the three independent variables are the temperature, the strain and the strain rate during the plastic deformation.

The second dataset is related to the pressure strength of green molding sand. It consists of 5 variables and 1076 measurements. The dependent variable is the green compressive strength, while the 4 independent variables are water content, carbon content, bentonite and clay content.

The third dataset is related to predicting the expected final cost of an ordered product. It consists of 11 variables and 4878 measurements. The dependent variable is the final cost, while the independent variables include delivery weight, material number, length, width, height, molding process, primer, iron ratio and two time related parameters.

To perform the experiments, the input data was divided into three parts. 60% of the input data was used for training and 20% was used for validation of the selected machine learning methods to select their respective optimal parameter values and 20% of the data was used for testing. The results of these methods for test data are shown in section 4.2.

During testing of MPF, the predictions from ANFIS, SVM and NN were transformed using KPCA.

### 4.1 Performance Evaluation Measure

To check the performance of the selected machine learning methods, three different error measures were computed namely *Root Mean Square Error* (RMSE), *Relative Root Mean Square*

Error (RRMSE) and *Symmetric Mean Absolute Percentage Error* (SMAPE), which are defined as

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |O_i - P_i|^2} \quad (10)$$

$$RRMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{|O_i - P_i|^2}{|O_i|^2}} \quad (11)$$

$$SMAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|O_i - P_i|}{(|O_i| + |P_i|)/2} \quad (12)$$

where  $n$  is the total number of patterns.  $O_i$  is the computed output and  $P_i$  is the predicted output from the used machine learning method.

## 4.2 Results and Discussion

### 4.2.1 Results for dataset Flow Stress

First dataset we used to observe our method performance contains actual measurements from metal forming experiments.

For NN, same settings were used as described in the section 3.2 For SVM, Gaussian kernel with sigma ( $\sigma$ ) value of 0.5 produced best results.

Since the number of independent variables is smaller in this instance, after computing the principal components, all three of them were selected, covering 100% of the variance in the data. For ANFIS, Linear kernel along with subtractive clustering (radius: 0.2) produced best results as shown in the Table 1 below.

Table 1: Flow stress - Machine learning tool results

Error type	Learning Methods		
	NN	SVM	ANFIS
RRMSE	0.0266	0.0336	0.0033
SMAPE	5.0818	9.4270	0.9259
RMSE	21.878	26.4546	2.6550

The following Table 2 shows that the results of the MPF as described in Section 3.5.

Table 2: Flow stress - MPF results

Error Type	Kernel: Polynomial		
	NN	SVM	MR
RRMSE	0.0033	0.0033	0.0033
SMAPE	0.9569	0.9225	0.9278
RMSE	2.6397	2.6454	2.6533

For all the considered error measures, the results as shown in Figure 3 clearly show that the combination of KPCA and ANFIS as well as the MPF with MLR outperformed NN as well as SVM.

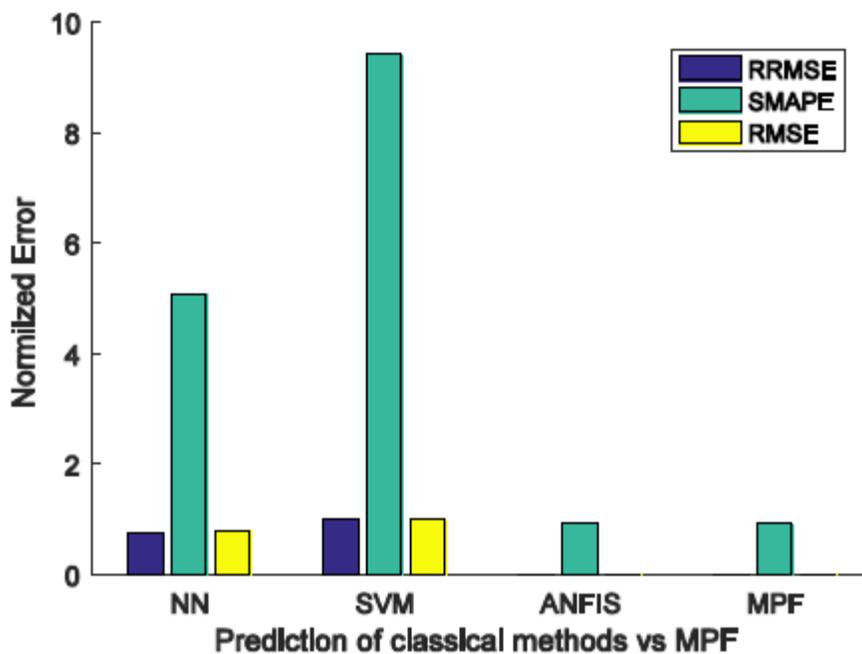


Figure 3: Results of all prediction methods for Flow Stress dataset

#### 4.2.2 Results for dataset Pressure strength of green molding sand

Second dataset we used to compare the introduced methodology results is related to computing the pressure strength of green molding sand.

For NN, same settings were used as described in the section 3.2 For SVM, Gaussian kernel with sigma ( $\sigma$ ) value of 5.0 produced best results.

For this dataset also, all four principal components were selected covering 100% of the variance. For ANFIS, Linear kernel along with subtractive clustering (radius: 0.3) produced best results. Computed results are shown in Table 3 below.

Table 3: Pressure shock resistance - Machine learning tools results

Error type	Learning Methods		
	NN	SVM	ANFIS
RRMSE	0.0425	0.0449	0.0426
SMAPE	4.4149	4.5869	4.3690
RMSE	1.1688	1.2348	1.1708

Again, for appropriate parameters, the results of the MPF as shown in the table 4 outperform all of the separate prediction results on the test data.

Table 4: Pressure shock resistance - MPF results

Error Type	Kernel: Polynomial		
	NN	SVM	MR
RRMSE	0.0481	0.0434	0.0422
SMAPE	4.9242	4.4232	4.3410
RMSE	1.3227	1.1944	1.1615

As we can see in the Figure 4, our proposed MPF outperforms all the individual learning methods.

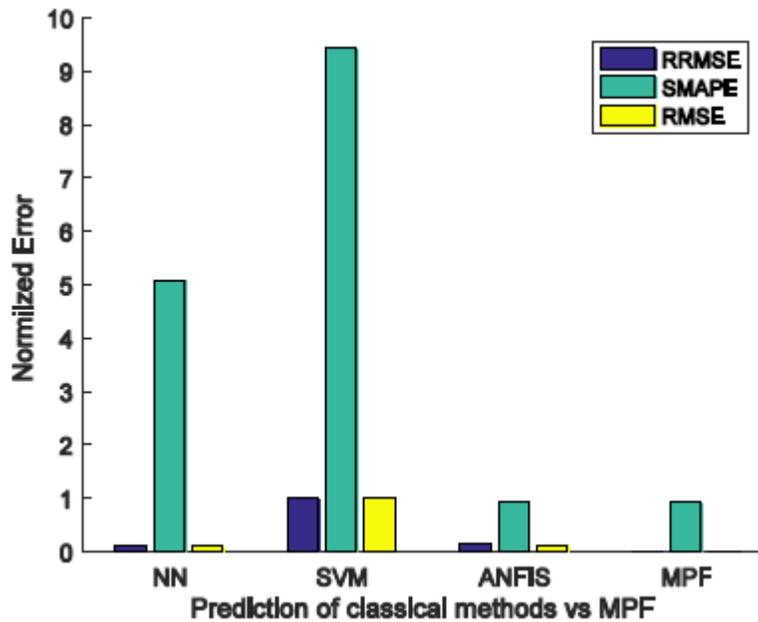


Figure 4: Results of all prediction method for Pressure Strength dataset

### 4.2.3 Results for dataset Final cost

Final cost dataset contains data recorded from actual production environment and is used to predict the expected final costs of the new ordered product.

For NN, same settings were used as described in the section 3.2 For SVM, Gaussian kernel with sigma ( $\sigma$ ) value of 5.0 produced best results.

In this case, seven out of ten principal components were selected covering 95.91% of the variance in the original data. For ANFIS, Linear kernel along with subtractive clustering (radius: 0.2) produced best results. Results are shown in Table 5 below.

Table 5: Final cost - Machine learning tool results

Error type	Learning Methods		
	NN	SVM	ANFIS
RRMSE	0.0143	0.0157	0.0233
SMAPE	9.9220	21.1273	23.9995
RMSE	347.7719	382.5654	565.8913

In this case, the prediction provided by the ANFIS in combination with KPCA was not able to achieve the accuracy of the classical methods. However, the MPF performed best in comparison with all the individual methods as shown in the Table 6 below. The ensemble method with MR produced best predictions.

Table 6: Final costs - MPF results

Error Type	Kernel: Polynomial		
	NN	SVM	MR
RRMSE	0.0139	0.0218	0.0140
SMAPE	10.9912	66.8833	8.5959
RMSE	336.4586	529.2121	339.2884

Figure 5 given below shows the effectiveness of the MPF in comparison with the performance of SVM and NN.

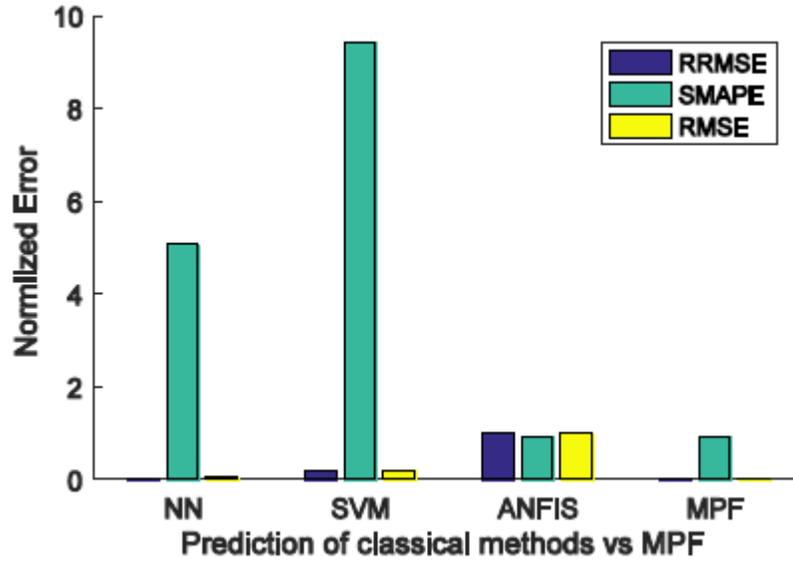


Figure 5: Results of all prediction methods for Final Costs dataset

For the test data, a random sample of prediction results from all the prediction methods used in this work are shown in figure 6. Close observation reveals that the MPF is performing better in comparison with other prediction methods. In Figure 7, which only shows the prediction by MPF, clearly indicates that the MPF is able to predict the test data very accurately.

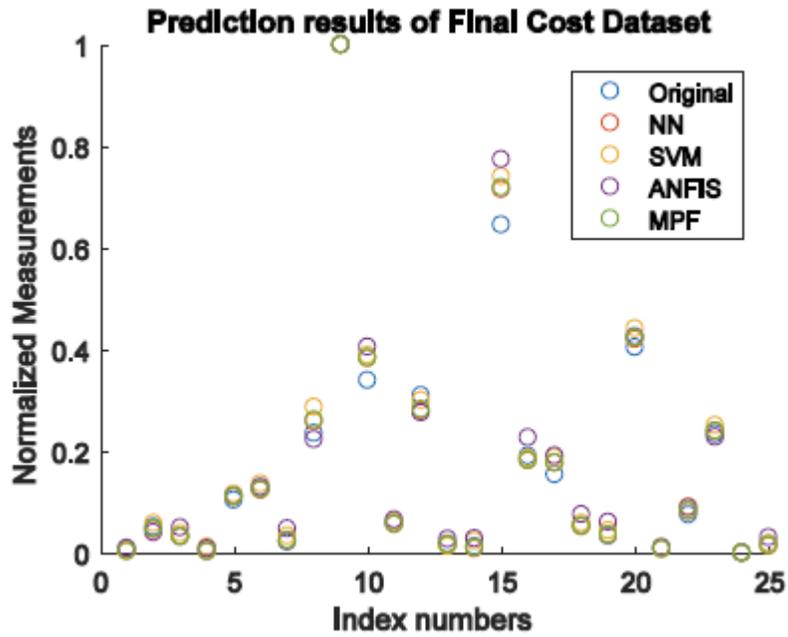


Figure 6: Predictions of sample of measurements for all prediction methods

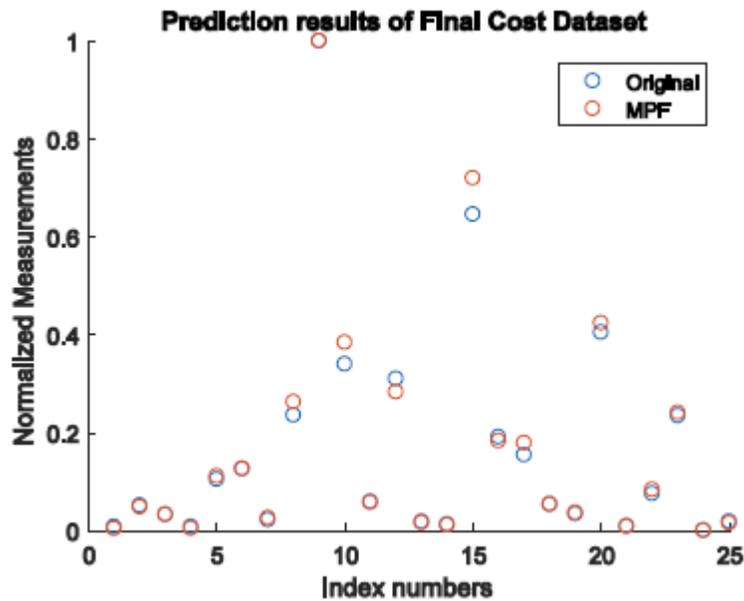


Figure 7: Predictions of sample of measurements for Meta-Prediction Function

## 5. CONCLUSION AND FUTURE WORK

In this paper, we introduced a novel methodology which combines results of different learning methods for quality prediction in production processes. Overall based on RMSE, SMAPE and RRMSE measures, our proposed MPF outperformed the other individual learning methods. Results show that the novel approach of transforming the results of different machine learning methods using KPCA and again applying learning techniques to the obtained data returned very good predictions. Also, it is shown that for the MPF, the combination of a linear kernel for KPCA with MLR produced the best results. For future work, the proposed methodology can be transferred for real time monitoring of production processes to identify parameter deviations which produce faulty parts.

In the future, we will also explore the use of other machine learning methods such as Decision Trees [25] and Bayesian Networks [26] and incorporate them in MPF. Finally, we note that usage of our MPF is not only limited for fault prediction of foundry data, but can also be used for forecasting cost, expenditure or sales prices of production process.

## REFERENCES

- [1] Ding, S. X., Yin, S., Peng, K., Hao, H., & Shen, B. (2013). A novel scheme for key performance indicator prediction and diagnosis with application to an industrial hot strip mill. *IEEE Transactions on Industrial Informatics*, 9(4), 2239-2247.
- [2] Neely, A. (1999). The performance measurement revolution: why now and what next?. *International journal of operations & production management*, 19(2), 205-228.
- [3] Enkel, E., Perez-Freije, J., & Gassmann, O. (2005). Minimizing market risks through customer integration in new product development: learning from bad practice. *Creativity and Innovation Management*, 14(4), 425-437.

- [4] Kan, S. H., Basili, V. R., & Shapiro, L. N. (1994). Software quality: an overview from the perspective of total quality management. *IBM Systems Journal*, 33(1), 4-19.
- [5] Sosna, Marc, Rosa Nelly Trevinyo-Rodríguez, and S. Ramakrishna Velamuri. "Business model innovation through trial-and-error learning: The Naturhouse case." *Long range planning* 43.2 (2010): 383-407.
- [6] Shingo, S., & Dillon, A. P. (1989). *A study of the Toyota production system: From an Industrial Engineering Viewpoint*. CRC Press.
- [7] Dilts, D. M., Boyd, N. P., & Whorms, H. H. (1991). The evolution of control architectures for automated manufacturing systems. *Journal of manufacturing systems*, 10(1), 79-93.
- [8] Kang, S., Kim, E., Shim, J., Cho, S., Chang, W., & Kim, J. (2017). Mining the relationship between production and customer service data for failure analysis of industrial products. *Computers & Industrial Engineering*, 106, 137-146.
- [9] Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27-34.
- [10] Ribeiro, L., & Barata, J. (2011). Re-thinking diagnosis for future automation systems: An analysis of current diagnostic practices and their applicability in emerging IT based production paradigms. *Computers in Industry*, 62(7), 639-659.
- [11] Schölkopf, Bernhard; and Smola, Alexander J.;" *Learning with Kernels*", MIT Press, Cambridge, MA, 2002, ISBN 0-262-19475-9.
- [12] Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural processing letters*, 9(3), 293-300.
- [13] Lawrence, Jeanette, "Introduction to Neural Networks", California Scientific Software Press 1994, ISBN 1-883157-00-5.
- [14] Craven, M. W., & Shavlik, J. W. (1997). Using neural networks for data mining. *Future generation computer systems*, 13(2-3), 211-229.
- [15] Jang, J-SR. "ANFIS: Adaptive-Network-based Fuzzy Inference System." *IEEE transactions on systems, man, and cybernetics* 23.3 (1993): 665-685.
- [16] B. Scholkopf, A. Smola, and K. Muller. "Nonlinear component analysis as a kernel eigenvalue problem". *Neural Computation*, 10(5):1299–1319, 1998.
- [17] Yang, J., Frangi, A. F., Yang, J. Y., Zhang, D., & Jin, Z. (2005). KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition. *IEEE Transactions on pattern analysis and machine intelligence*, 27(2), 230-244.
- [18] Leha, A., Pangercic, D., Rühr, T., & Beetz, M. (2009, September). "Optimization of simulated production process performance using machine learning". In *Emerging Technologies & Factory Automation, 2009. ETFA 2009. IEEE Conference on* (pp. 1-5). IEEE.
- [19] Ashour, M. W., Khalid, F., Halin, A. A., & Abdullah, L. N. (2015, October). "Machining process classification using PCA reduced histogram features and the Support Vector Machine". In *Signal and Image Processing Applications (ICSIPA), 2015 IEEE International Conference on* (pp. 414-418). IEEE.

- [20] Yuan, Y., Zhang, H. T., Wu, Y., Zhu, T., & Ding, H. (2016). Bayesian learning-based model predictive vibration control for thin-walled workpiece machining processes. *IEEE/ASME Transactions on Mechatronics*.
- [21] Rostami, H., Blue, J., & Yugma, C. (2016, December). Equipment Condition Diagnosis and Fault Fingerprint Extraction in Semiconductor Manufacturing. In *Machine Learning and Applications (ICMLA), 2016 15th IEEE International Conference on* (pp. 534-539). IEEE.
- [22] H.R. d. N. Costa and A. La Neve, "A study on the application of regression trees and adaptive neuro-fuzzy inference system in glass manufacturing process for packaging," 2016 Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS), El Paso, TX, USA, 2016, pp. 1-4.
- [23] Dietterich, Thomas G. "Ensemble methods in machine learning." *International workshop on multiple classifier systems*. Springer Berlin Heidelberg, 2000.
- [24] Aiken, L. S., West, S. G., & Pitts, S. C. (2003). Multiple linear regression. *Handbook of psychology*.
- [25] Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3), 660-674.
- [26] Jansen, R., Yu, H., Greenbaum, D., Kluger, Y., Krogan, N. J., Chung, S., ... & Gerstein, M. (2003). A Bayesian networks approach for predicting protein-protein interactions from genomic data. *science*, 302(5644), 449-453.

## AUTHORS

**Saad Bashir Alvi** is currently a Ph.D. student and a research associate in the Institute for Technologies of Metals at the University of Duisburg-Essen. He has previously worked as a software engineer and received his M.Sc. from University of Bonn, Germany in 2013. His research interests include machine learning and Neural Networks with special emphasis on comparative studies in predictive analytics and focus on dimension reduction as well as neuro fuzzy modeling with applications to production processes.



**Robert Martin** received his Ph.D. from the University of Duisburg-Essen University, Germany in 2016 and is currently working at the Institute for Technologies of Metals in Duisburg.



**Prof. Johannes Gottschling** is working at the Institute for Technologies of Metals at the University of Duisburg-Essen and is head of the chair "Mathematics for Engineers".



*INTENTIONAL BLANK*

# CLASSIFICATION OF SMART ENVIRONMENT SCENARIOS IN COMBINATION WITH A HUMAN-WEARABLE-ENVIRONMENT-COMMUNICATION USING WIRELESS CONNECTIVITY

Kristof Friess and Prof. Dr. H.C. Volker Herwig

Department of Applied Computer Science,  
University of Applied Science Erfurt, Erfurt, Germany

## **ABSTRACT**

*The development of computer technology has been rapid. Not so long ago, the first computer was developed which was large and bulky. Now, the latest generation of smartphones has a calculation power, which would have been considered those of supercomputers in 1990. For a smart environment, the person recognition and re-recognition is an important topic. The distribution of new technologies like wearable computing is a new approach to the field of person recognition and re-recognition. This article lays out the idea of identifying and re-identifying wearable computing devices by listening to their wireless communication connectivity like Wi-Fi and Bluetooth and building a classification of interaction scenarios for the combination of human-wearable-environment.*

## **KEYWORDS**

*Wireless Network, Smart Environment, Wearable-Computing, UbiComp, Interaction-Scenarios*

## **1. INTRODUCTION**

With the growing market of worn computer systems like smartphones and smartwatches, in short wearables, the possible interaction between human and computer has changed. In a short time, also the interaction between human, computer and the environment will change. There are unlimited use cases where a human uses the computer as an assistant for filtering data, processing information or storing context information. Now and in the near future, there are a lot of use cases coming up, where not the device itself helps the human to become smarter, but the environment, based on the knowledge about the human, acts smartly. What those scenarios are and how they can be used is the core contribution of this paper. The base technology to implement those scenarios are technologies to recognize humans in the environments.

After an introduction to the state of human recognition, an approach using the wearable computing devices as a tag for smart environments is given. Then, Smart Environments are defined, meaning what are Smart Environments and what is done with a Smart Environment. On the basis of this definition, the interaction between the human and the wearable is enhanced by the interaction with the environment. This extension represents the flow of information between human-wearable and the environment and forms the basis for the classification of Smart Environments, which is developed afterwards.

## **2. HUMAN RECOGNITION**

### **2.1 Established Recognition Techniques**

The application field of human recognition and re-recognition is already there. There are many situations where video surveillance detects movements of humans for security reasons, for examples in factories, banks, airports, or country borders. Also at huge events, video surveillance is used for monitoring individuals and groups to detect panic situations, timely. But all those systems bring no additional benefit, apart from security, to the human, who is being recognized. The detection quality using video analysis and live images suffers, when people in the pictures are overlapping or are very small. When the overlap increases (more people), or people get smaller (because of the distance) in the image, then negative-positive results will be much higher [1] [2]. To re-recognize an object, in this case a human, features of the human (for example clothing) are needed. The required characteristics used for re-recognition can only be found at the human himself [2] [3, 4]. So, if the human changes his clothes, he can no longer be identified as the same human as before.

Another possible way of identifying humans is that every human inside the environment in focus is wearing a tag (e.g. infrared sender), so that each person can be detected. A problem is the sharing of the tag [4]. An alternative for this is the manipulation of the environment with sensors, like pressure sensors inside the ground panel of a floor (e.g. “active Floor” [6]). So, the environment is only knowing there is someone inside the floor, but not who, and also it does not know whether it is a returning human (re-recognition) [6].

The use of a tag, however, appears promising especially when considering the development of the spread of wearables [6] [7] [8]. Wearables, such as a smartphone, usually have wireless communication technology (Wi-Fi, Bluetooth). These technologies can be used to recognize people within an environment by different identification features (such as the network address) [4] [10]. Since most people carry a wearable, the equipment and a smart environment can be realized with simple means.

### **2.2. Recognition using Wearable Computing**

The first wearables were developed more than 1000 years ago. In 1268, the philosopher Roger Bacon wrote about lenses for optical use – the glasses. This is the earliest mention of eyeglasses and the very first step towards wearables [11] [12]. Nearly 400 years later, an abacus ring appeared that was small and portable, and could be used as a calculator [12]. In the following years, more and more wearables were developed, often with mechanical technology such as small pocket watches. Then, in 1966 the first wearable computer was developed. Ed Thorp and Claude Shannon created a portable computer to predict the position of the ball in the casino game roulette

[6] [7]. By now, the market for wearable computing has grown significantly and continues to grow [6] [7].

The question whether a computer is a wearable or not is answered by the pioneers in wearable computing. A definition of wearable computing comes from Thad Starner. He defines the wearable computer as an extension of the human body with passive use, similar to riding a bike. The idea is a very close connection of humans and computers, which allows a new way of perception that would not be possible without a computer. The computer is closely connected to the human and needs to make the same experience as the user. The computer learns from the user's environment and how he acts in certain situations. As the computer learns this, its ability in assisting the user increases [13].

Another definition of wearable computing is presented by Bradley J. Rhodes in 1997 in his article about wearables as remembrance agents [14]. He describes wearables based on the characteristics – portable while operational, hands-free use, sensors (GPS, camera) not only for input, proactive (getting in touch with the user), and many others. According to Bradley, a wearable computer should have many of those characteristics. He tries to show a clear distinction between a Palmtop PC and a wearable computing.

Next to Rhodes, Steve Mann published his definition of wearable computing at the International Conference on Wearable Computing ICWC-98 in 1998 [15]. Early on, he started to experiment with enhancements to his body. For example, he packed an 8-bit microprocessor 6502 in a backpack to control photo graphics equipment in 1981. He defined these devices as a new form of interaction between humans and computers. The computers are small, worn on the body, the system is always running, always ready, and always available. Those are the biggest differences to PDAs, handheld devices, and laptops. This new form of permanent availability and a given consistency of the user interface leads to new synergies between humans and computers [15] [16]. In detail, he sub-classifies wearables into three main characteristics (constancy, augmentation, mediation), which refer to the execution, and into six basic properties (un-monopolizing, unrestrictive, observable, controllable, attentive, communicative) in communication between humans and computers [15].

Looking at the different definitions of wearable computing, a clear concept reveals: Wearables are devices which should be integrated into our daily lives. They are always with us, and monitor us and the environment using sensors. The wearables are intended to assist people and support them secondarily (inform, analyze) while they are performing a primary activity. For that reason, wearables are a perfect basis for a new approach in personal recognition and re-recognition and in setting up some new concepts for smart environment.

The current methods of human recognition and re-recognition allow only limited automation. That can be explained by weaknesses in the algorithms [2]. People detection using video analysis and live images suffers from the challenge that people are overlapping and very small. As already stated before, the re-detection is challenged by the unique identification like clothes.

To recognize people across multiple cameras and over a longer period of time, an additional source of information is required, that is less likely to change, such as clothing. By considering modern technology like smartphones, which is already used by more than 49 million people [8] in Germany, there is a recognition feature that does not change every day.

Due to the high frequency of use, almost 75% of smartphone owners take their device with them [21] [22]. That technology is a good starting point to explore new approaches of recognition. In addition to the smartphone, the wearable computing market continues to grow as well [6] [7]. The devices range includes simple fitness trackers, smartwatches, as well as head-up displays like Google Glasses.

An extensive examination of the currently marketed wearables, done by the author, identified that most units have a wireless communication interface. In addition to GSM / EDGE / 3G and LTE, the most common technology is Wi-Fi, followed by Bluetooth and Bluetooth Low Energy. Other communication interfaces such as ANT + and NFC could also be identified, but currently the relevance and the distribution of ANT + is not as high as of Bluetooth and Wi-Fi. Furthermore, the NFC range is too short for widespread recognition.

Human recognition using Wi-Fi with a non-active connection to a hot spot and an unchanged hardware already has some experimental progress. This is made possible by periodically sent probes of the active Wi-Fi modules [4]. Therefore, wearables can be detected and optionally localized without being actively connected to a Wi-Fi [4]. For example, a smartphone sends a Wi-Fi sample to detect access points close-by every 15 to 60 seconds, depending on the battery state [10]. This type of access point detection is used in multiple studies. In a high traffic street, 7,000 devices could be detected within a time frame of 12 hours [10]. In another study, around 160,000 devices could be detected during three hours [4, 19, 19]. A company in London was able to recognize people using this technology for several weeks. They could identify more than 530,000 smartphones within two weeks [23]. Further experiments have shown that the recognition using Wi-Fi probes is promising as well [24] [25] [23].

In addition to the already existing tests with Wi-Fi, there are other approaches using Bluetooth. Bluetooth was created in 1998 with the merge of several companies under the name Bluetooth Special Interest Group (SIG) [27] [28]. In the following year, version 1.0 was released. Bluetooth is now installed in most phones and is therefore another possibility for mobile positioning and identification of mobile devices. With the release of Bluetooth 4.0 standards in 2010, the new low energy technology was created, which makes it possible to maintain a radio link with low power consumption for a longer time [27] [29] [27]. Using this specification, a company enabled the opportunity for recognizing objects as well as people using small beacon modules. Using the beacons, it is also possible, according to a previously performed calibration, to locate people with a smartphone in premises [31] [32].

It has been shown that several approaches for human recognition and re-recognition using radio wave technologies are available. Almost all of the previously revealed methods use unmodified hardware carried by the people. For this concept, we focus on Wi-Fi and Bluetooth connectivity. Because with the current state of the art knowledge, we know that these two wireless technologies have the possibility to re-identify a user using the MAC or IDs of devices, so this are the best enabling technology for a smart environment. With the scenarios described above, the technology used has to be evaluated in regards to sources of error and the ability for application.

### **3. SMART ENVIRONMENT BASED ON WEARABLE COMPUTING**

As early as the early 1990s, Mark Weiser and his colleagues came to the conclusion that the PCs of the world and the change to the laptop is only an intermediate step in the transformation to the

smart environment [33]. The idea of Weiser is that the computer disappears in the background of the environment and is thus not the main focus of the utilizing human being. On the contrary, man should concentrate on the fellow human beings and his activities [33] [19]. The merging of the environment with computers, the concept that computers can appear in various forms at all times, is summarized under the term ubiquitous computing. Ubiquitous computing (or brief ubicomp) is the idea of a “physical world richly and invisibly interwoven with sensors, actuators, displays, and computational elements, embedded seamlessly in the everyday objects of our lives and connected through a continuous network” [19], from this idea the Smart Environment developed. smart environment, or even Smart Space, is a “region of the real world that is extensively equipped with sensors, actuators and computing components” [31] [32]. So an environment with many sensors, motors and computer components as it has already described in [33]. The described environment opens up new possibilities for interaction and usability. It is a combination of several heterogeneous systems “systems and networks, between systems and systems, and between systems and people” [32].

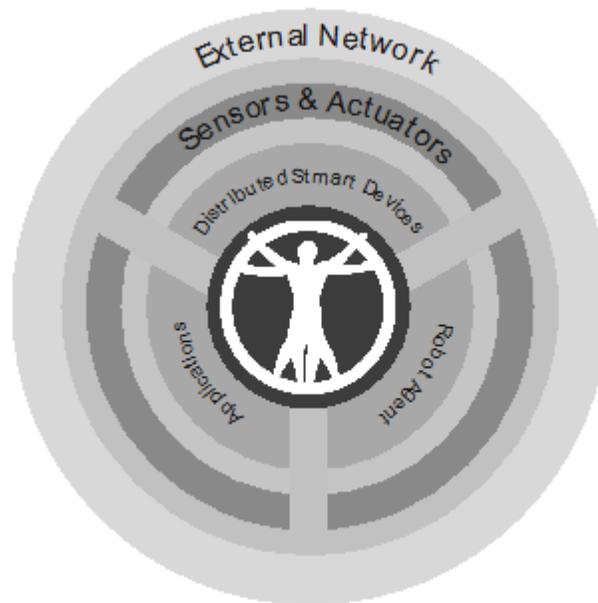


Figure 1. Schematic representation of smart environments [37]

Cook provides a similar description: Smart Environments are a “...small world where all kinds of smart devices are continuously working to make inhabitants' lives more comfortable...” [33]. In this smart or intelligent environment, it is possible to autonomously assess the nature and the individuals and to make decisions from this knowledge. The goal is to improve comfort for the individual and the experience [33]. A schematic view on smart environment is given in Figure 1.

Based on these smart environment descriptions, the following definition is used for this article:

Smart Environments are environments with a significant amount of computer components, motors, and sensors. These components, however, unlike the personal computer, are fused with the environment and thus invisible to the individuals who are in it. The environment is connected via an internal and external communication network and knows how to deal with the data it raises. On the basis of the information gathered, the environment can perform tasks that help individuals and increase comfort.

Examples of Smart Environments are ubiquitous. Smart Smarties, SmartHouses, SmartHealths, SmartLaboratories, SmartLaboratories and SmartRooms are a way of realizing smart environments [33] [32] [34].

Already in 1989, there were first experiments to realize a smart environment. At the Olivetti Research Labs/ University of Cambridge, people were able to be located in rooms and also in the building using an "Active Badge". This was possible because the active mark periodically transmitted an infrared signal to the sensors positioned in the room [39]. However, it could only be ascertained whether a person is in a room, but not where. In order to improve localization, a new active marker was used in 1995. With the help of transmitted ultrasonic waves, it was now possible to determine the position of a person with an accuracy of up to 3cm in the room. The name for the new marker was "Active Bat". This is due to the fact that a bat uses a similar system for position detection [39]. In addition to the identification of persons in rooms using markers, there are also systems that function without a marker. A good example is the Smart Floor. By installing pressure sensors in the floor slabs of a hallway, persons walking across could be recognized [36].

It is clear that a smart environment can only be created, if the environment knows that there is a person, and in the best case, which person is in it. The listed examples are only a small set of actual implementations, but they illustrate the need for manipulation. Manipulation means that the environment must be equipped with sensors in such a way that they can recognize people and / or that the persons must be "manipulated" by carrying a marker. The latter can be supported by the wearing of a wearable.

Because by carrying smartphones, smartwatches and other portable computers with radio technologies, there is an "active tag", a marker for the environment. The fact that most people already carry an active marker every day promotes the development of a smart environment with significantly fewer sensors. This is because the environment can use the wearables' functional technologies within a certain range for the communication or identification of the person and recognize, locate and identify the person. With this information and the information that the environment receives from the other sensors and components, the support and thus the extension of the experiences of the persons within the environment can begin.

When considering the flow of information between humans and computers as defined by Steve Mann, the wearable encompasses the human being and a permanent interaction between man and wearable is possible [15]. This flow of information is now extended and allows the wearable, which is to be understood as a protection for the privacy as well as the extension of the senses [15], a communication with the environment. For this information flow to work, it is necessary that the wearables are equipped with an active wireless technology such as Bluetooth or Wi-Fi, for data exchange.

The interaction between human-computer defined by Steve Mann has been expanded in the Figure 2 by the information flow (in- / output) between the human computer environment. This shows how the communication between the environment and the human being and the environment and the wearable functions. The wearable always has an active functionality and always sends an output that identifies the person. Since this identification must be activated continuously, a connection with the environment in Figure 2 is indicated by a continuous arrow. The environment itself can provide information to people or the computer / wearable, either via radio technologies, sensors, screens or audio. However, this is not always necessary and is

therefore shown in the figure with a dashed arrow. The same is true for the human. He can be a supplier of information, for example, by talking, pressing a switch or being captured by a camera.

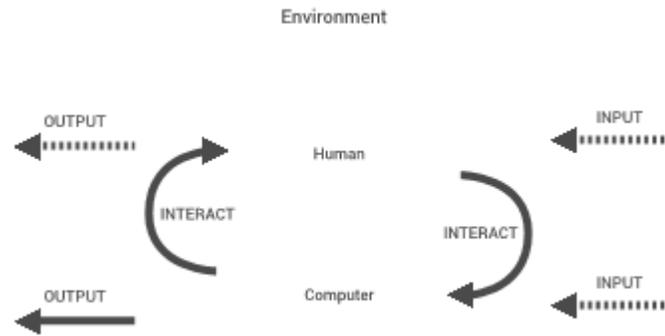


Figure 2. Human-Wearable-Environment information flow

#### 4. CLASSIFICATION OF SMART ENVIRONMENT SCENARIOS

A second factor plays an important role in considering the flow of information between human-wearable and the environment. This is the way communication / interaction with the environment takes place and how data is exchanged. In order to describe this communication and interaction in a more general way, a classification of scenarios takes place. When considering different scenarios for Smart Environments, it could be revealed that these differ in the provision of the information of human-wearable, in the nature of the interaction which originates from the environment and in the kind of the feedback.

It was previously defined that the wearables have a passive communication based on their radio technologies with the environment. The wearable and also the human being get nothing in return from this kind of data transmission, therefore this is called a *Passive Tag*. It is different when the wearable uses the radio connection to transmit information to the environment – *Active Tag*.

**Passive Tag** – The human being and the wearable, too, are not aware that they are perceived by the environment. The wearable is detected by the active radio technology (e.g. Wi-Fi, Bluetooth). Based on this, the human-wearable can be identified. Depending on how the environment processes the information about the human, it can learn the behavior of human beings and support them in the future through interactions.

**Active Tag** – The human being wants the environment to perceive him and provides additional information about the wearable. This information will personalize the people towards the environment. For example, people could share the shopping list with the environment and this responds with an optimized shopping route. Further information can be people's interests, such as politics, music taste and sports. All this information is used to optimize the interaction and personalization of the feedback to the people, for example in the form of tailored advertising on screens, adapted room temperature or music.

With the information that the environment receives from the active or passive tags, it can make decisions for the interaction, how the environment reacts to the human-wearable. In considering how a smart environment interacts with the human-wearable combination, three different interaction possibilities can be classified.

**Direct Interaction** – There will be a directly returned feedback to the human-wearable. This may be accomplished by performing various tasks, such as turning lights on / off, adjusting room temperature or tracking the person with the favorite music when walking through the premises of a building. Ultimately, the environment directly affects and supports the people.

**Indirect Interaction** – The human-wearable does not get any direct feedback from the environment. The environment uses the information for analysis. The information obtained (e.g. the number of persons, movement currents, actions of the persons) are used to learn future actions / decisions (e.g. create a new direct interaction automatic open and close windows or recognizing dangerous situations).

**No Interaction** - The human-wearable gets nothing from the environment, all the information that the environment captures are used for analysis. This is the case, for example, when analyzing a trade fair. It is determined how many people have attended a stand. This information serves purely the statistics and is not returned to the human-wearable in any form.

The smart environment is designed to improve the user experience, to expand the comfort of the people and to support them in the everyday life from the background. Not always, as defined above, direct communication with the human-wearable occurs. However, the surrounding environment collects data. A distinction in which form the environment processes the data and deduces actions from it is decisive. There are four different types of feedback.

**Navigation** – All scenarios supporting the control and navigation of human-wearable within the environment. An example is a person who has just entered a train. This person does not know where there are still free seats, but the train does. Therefore, direct control of this person is done using the wearables or monitors to find a free seat. A similar scenario is a parking garage. The surrounding area would directly navigate to the next free space.

**Content** – All scenarios in which the environment uses the collected information about human-wearable to provide optimized content for it. This is the case, for example, with tailored advertising and optimized messages. The human being gets the information he is interested in.

**Observation** – Identifies all scenarios in which the environment is used as a pure observer. This is the case, for example, when analyzing the number of visitors in buildings or at trade fairs.

**Trigger** – Identifies all scenarios in which the environment triggers an action for human-wearable, for example open a door or turning on/off the light.

In order to understand why this classification has been decided, some scenarios will follow, which illustrate the respective information flows and interactions. The scenarios constructed in this way are fictive and their realization and meaningfulness should not be evaluated. It is important, however, that there are people in it who carry a wearable with active radio technology (such as Bluetooth or Wi-Fi).

### **Smart Train: My Seat**

At the main station of Erfurt is Sebastian S., he has an electronic ticket from Erfurt to Berlin on his smartphone and is waiting for his ICE. While waiting for his train, his smartphone informs him that he is currently in the wrong area. His reserved seat is in the area E. After the train has

entered, Sebastian S. enters the train, on a monitor directly in front of him is indicated that his seat 32F is on the left side. Sebastian S. follows this instruction and quickly finds his seat by a flashing light.

This scenario is an active tag. Because the environment got the information from the electronic ticket and could use this information to carry out a supporting navigation. Sebastian S. was localized and navigated on the basis of his smartphone. The interaction was direct with the human-wearable.

*Classification:* active tag, direct interaction, navigation

### **Smart Train: Detect a free Seat**

At the main station of Dresden is Lucas L., he has an electronic ticket from Dresden to Munich on his smartphone and is waiting for his ICE. While waiting for his train, a monitor informs him that there are still lots of free seats in the Wagon 21. The wagon 21 stops in area B. After the train has arrived, Lucas enters. The monitor in the entrance area informs him that there are still a lot of free seats on the right and recommends to go to the right.

This scenario is a passive tag. Because the environment only gets the information that there is a person with a wearable in it. On the basis of this information and further information about the current characteristics in the incoming train, navigation could be carried out by means of environmental elements. Lucas L. was localized and navigated on the basis of his smartphone. The interaction was direct with the human-wearable.

*Classification:* passive tag, direct interaction, navigation

### **Optimized Advertisement**

Maria M. is at the fair CEBIT. She visits different stands longer, others shorter. Certain booths she does not visit at all. As she moves around the trade fair, the monitors display ads that relate to topics based on the booths Maria M. has already visited.

In this example, in addition to indirect and direct interaction, two different types of feedback are also shown. First, it is assumed that Maria M. does not provide any information about her wearable. That is, it has a passive tag. In order to personalize the advertising for Maria M., the surroundings must learn about the interests of Maria M. by means of the up-and-coming days at the respective stands. This learning is positioned in the indirect interaction with human wearable. If there is sufficient knowledge, the environment can derive action decisions from this – direct interaction. Maria M. also has the possibility to wear an active tag. In doing so, they pass on information about their interests to the environment and the environment can start directly with the interaction. In the nature of the feedback, there is initially the presentation of content and with the recommendation of further fair stands a navigation takes place.

*Classification:* passive / active tag, indirect / direct interaction, content / navigation

**People-Flow**

The passenger flows are measured in a shopping center. The measurements serve to analyze the area in which advertising is best placed: where visitors are most likely to stay and which areas are not visited at all.

This is a scenario where visitors are tracked and analyzed with their passive tag. This is done by observation only. The environment does not derive any action decisions from this information.

*Classification:* passive day, no interaction, observation

**Smart Buildings**

Every morning at around 8 o'clock, Katrin K. arrives at the office. At first, she switches on the light in the entrance area, then in the corridor and in her room.

One day, when Katrin K. arrives, the light turns on automatically in the entrance area, then in the corridor and in her room.

In this scenario, it is obvious that the environment can derive its own action by means of the persons and the actions. The environment is first to be classified as an indirect interaction. After the environment has learned how to support the human-wearable, it is a direct interactor.

*Classification:* passive day, indirect / direct interaction, observation / trigger

These are only a small part of scenarios that can be described with this classification. During the review, further scenarios were conceived and found that this classification is useful for building further possible scenarios and ideas for Smart Environments.

**5. CONCLUSIONS**

Wearable computing is a good system for supporting Smart Environments. Through the focus of wearable computing which learns from humans and supports them in everyday life, the wearable (for example, a smartphone) becomes an assistant. The integrated wireless technologies provide a data exchange interface, which can be used by other wearables or environments for the information exchange.

In the case of a human wearable environment, it is assumed that the wearable is always worn by humans. The interaction is, as Steve Mann defined, encapsulated and continuous. That is, the wearable has the possibility to demand the attention of humans and to interact with them. The human being always has the possibility to interact with his wearable. The environment surrounds people with their wearable. A continuous flow of information from the wearable to the environment takes place via the radio technology. The environment can decide how and whether it interacts with the human-wearable and / or exchanges information.

The possibility that the environment does not necessarily have an interaction / information flow with the human wearable is the basis for the creation of a classification of interaction scenarios between humans and the environment. It was possible to identify a wearable as an active or

passive information provider in the environment. A wearable with active radio connection is a passive tag, the radio connection and an identifier of the device can identify the person within the environment. If the wearable additionally provides information about the person, this is referred to as an active tag. Whether active or passive, the environment can learn from the actions of the person and derive their own actions from it. This learning is referred to as indirect interaction because the environment first processes the data, and it is only when the environment develops actions (switching on light) that direct interaction occurs. However, it may also be that the environment only functions as an information collector, that is, there is no interaction. Depending on which actions the environment chooses to interact with the human-wearable, the feedback differs. If the environment switches on the light when entering the room, then this is a trigger. If the environment responds with, for example, advertising on monitors, this is a substantive feedback. If the environment tries to control the people in it by means of directional indications on the monitors, this is a navigation feedback. Or as briefly described above, the environment analyzes the people within without feedback and only observes them. Thus, different scenarios can be classified with the distinction into active or passive tags, the type of interaction (direct, indirect or none), and the type of feedback (content, navigation, observation, trigger).

## REFERENCES

- [1] E. Corvee, S. Bak and F. Bremond, "People detection and re-identification for multi surveillance cameras," VISAPP - International Conference on Computer Vision Theory and Applications - 2012, 2012.
- [2] P. Dollar, C. Wojek, B. Schiele and P. Perona, "Pedestrian Detection: An Evaluation of the State of the Art," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012.
- [3] F. J. Baek, M. K. Islam and Joong-Hwan, "Person Detection, Re-identification and Tracking Using Spatio-Color-based Model for Non-Overlapping Multi-Camera Surveillance Systems," Smart Computing Review, vol. 2, no. 1, p. 42, 2012.
- [4] M. V. Barbera, A. Epasto, A. Mei, V. C. Perta and J. Stefa, "Signals from the Crowd: Uncovering Social Relationships Through Smartphone Probes," in Proceedings of the 2013 Conference on Internet Measurement Conference, New York, NY, USA, 2013.
- [5] S. Poslad, Ubiquitous Computing: Smart Devices, Environments and Interactions, London: John Wiley & Sons, Ltd., 2009, p. 42.
- [6] M. Addlesee, A. Jones, F. Livesey and F. Samaria, "The ORL Active Floor," IEEE Personal Communications, 1997.
- [7] GfK, "Absatz von Smartwatches in Deutschland in den Jahren 2013 bis 2015," 2015. [Online]. Available: <http://de.statista.com/statistik/daten/studie/459093/umfrage/absatz-von-smartwatches-in-deutschland/>. [Accessed 10 03 2016].
- [8] IDC, "Prognose zum Absatz von Wearables weltweit von 2014 bis 2020 (in Millionen Stück)," 01 01 2016. [Online]. Available: <http://de.statista.com/statistik/daten/studie/417580/umfrage/prognose-zum-absatz-von-wearables/>. [Accessed 29 03 2016].
- [9] Statista, "Anzahl der Smartphone-Nutzer in Deutschland in den Jahren 2009 bis 2016 (in Millionen)," [Online]. Available: <http://de.statista.com/statistik/daten/studie/198959/umfrage/anzahl-der-smartphonenuutzer-in-deutschland-seit-2010/>. [Accessed 26 07 2016].

- [10] A. B. M. Musa and J. Eriksson, "Tracking Unmodified Smartphones Using Wi-fi Monitors," in Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems, New York, NY, USA, 2012.
- [11] B. Rhodes, "A brief history of wearable computing," 2001. [Online]. Available: <https://www.media.mit.edu/wearables/lizzy/timeline.html>. [Accessed 29 03 2016].
- [12] L. d. Medici, "The History Of Wearable Technology - Past, Present And Future," 15 11 2015. [Online]. Available: <https://wtvox.com/featured-news/history-of-wearable-technology-2/>. [Accessed 29 03 2016].
- [13] T. Starner, "The challenges of wearable computing: Part 1," IEEE Micro, vol. 21, no. 4, pp. 44 - 52, Juli-August 2001.
- [14] B. J. Rhodes, "The wearable remembrance agent: a system for augmented memory," Wearable Computers, 1997. Digest of Papers., First International Symposium on, pp. 123-128, 1997.
- [15] S. Mann, "Definition of "Wearable Computer"," 1999. [Online]. Available: <http://wearcam.org/wearcompdef.html>. [Accessed 10 03 2016].
- [16] S. Mann, "WEARABLE COMPUTING as means for PERSONAL EMPOWERMENT," 25 08 1998. [Online]. Available: <http://wearcam.org/icwckeynote.html>. [Accessed 10 03 2016].
- [17] BITKOM, "Das Handy als ständiger Begleiter," 2015. [Online]. Available: [http://www.bitkom.org/de/presse/64046\\_77337.aspx](http://www.bitkom.org/de/presse/64046_77337.aspx). [Accessed 01 03 2015].
- [18] TNS Infratest Mobile Club 2013, "Smartphone-Nutzung und ihre Einsatzorte," 2015. [Online]. Available: [http://www.tns-infratest.com/Presse/pdf/Presse/2013\\_05\\_06\\_TNS\\_Infratest\\_Mobile\\_Club\\_Mobile-Landscape\\_Charts.pdf](http://www.tns-infratest.com/Presse/pdf/Presse/2013_05_06_TNS_Infratest_Mobile_Club_Mobile-Landscape_Charts.pdf). [Accessed 02 03 2015].
- [19] M. Weiser, R. Gold and J. S. Brown, "The origins of ubiquitous computing research at PARC in the late 1980s," IBM SYSTEMS JOURNA, vol. 38, no. 4, 1999.
- [20] K. a. Shubber, "Presence Orb uses Wi-Fi to detect if buses and bars are full," 01 04 2015. [Online]. Available: <http://www.wired.co.uk/news/archive/2014-04/22/presence-orb>. [Accessed 20 09 2015].
- [21] R. Lim, "Tracking Smartphones Using Low-power Sensor Nodes," in Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, New York, NY, USA, 2013.
- [22] J. Freudiger, "How Talkative is Your Mobile Device?: An Experimental Study of Wi-Fi Probe Requests," in Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks, New York, NY, USA, 2015.
- [23] K. Li, C. Yuen and S. Kanhere, "SenseFlow: An Experimental Study of People Tracking," in Proceedings of the 6th ACM Workshop on Real World Wireless Sensor Networks, New York, NY, USA, 2015.
- [24] Bluetooth, "History of the Bluetooth Special Interest Group," 04 01 2015. [Online]. Available: <http://www.bluetooth.com/Pages/History-of-Bluetooth.aspx>. [Accessed 01 03 2015].
- [25] M. Sauter, Communication Systems for the Mobile Information Society, John Wiley & Sons Ltd.

- [26] Bluetooth, "Bluetooth Smart Technology: Powering the Internet of Things," 2015. [Online]. Available: <http://www.bluetooth.com/Pages/Bluetooth-Smart.aspx>. [Accessed 04 03 2015].
- [27] K. T. Davidson, C. Cuffi, Akiba and Robert, Getting Started with Bluetooth Low Energy, Media, O'Reilly, 2014.
- [28] Estimote, "Indoor Location," 2015. [Online]. Available: <http://estimote.com/indoor/>. [Accessed 06 03 2015].
- [29] Apple, "Understanding iBeacon," 2015. [Online]. Available: <https://support.apple.com/en-us/HT202880>. [Accessed 06 03 2015].
- [30] M. Weiser, "The Computer for the 21st Century," SIGMOBILE Mob. Comput. Commun. Rev., vol. 3, no. 3, pp. 3-11, 1999.
- [31] P. A. Nixon, W. Wagealla, C. English and S. Terzis, Security, Privacy and Trust Issues in Smart Environment, 2004.
- [32] P. N. Dobson, G. Lacey and Simon, Managing Interactions in Smart Environments: 1st International Workshop on Managing Interactions in Smart Environments (MANSE'99), Dublin, December 1999, London: Springer-Verlag, 1999, pp. 1,3.
- [33] D. J. C. Das and K. Sajal, Smart Environments: Technologies, Protocols, and Applications, John Wiley & Sons, Inc., 2005, pp. 3,4,9.
- [34] F. Marquardt and A. M. Uhrmacher, "Evaluating AI Planning for Service Composition in Smart Environments," Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia, vol. MUM '08, pp. 48-55, 2008.
- [35] S. Poslad, Ubiquitous Computing: Smart Devices, Environments and Interactions, John Wiley & Sons, Ltd., 2009, p. 42.
- [36] M. Addlesee, A. Jones, F. Livesey and F. Samaria, "The ORL Active Floor," IEEE Personal Communications, 1997.
- [37] FairControl, "Controlling von Live-Kommunikation," 2013. [Online]. Available: [www.faircontrol.de](http://www.faircontrol.de). [Accessed 06 08 2013].
- [38] GMM Gelszus Messe-Marktforschung, "GMM - Experten für mobile Marktforschung vor Ort," 01 01 2013. [Online]. Available: <http://www.gmm-marktforschung.de/>. [Accessed 26 08 2013].
- [39] drinktec, "Lead Tracking," 01 01 2013. [Online]. Available: <http://www.drinktec.com/en/Home/ForExhibitors/exhibitor-services/leadtracking>. [Accessed 10 08 2013].
- [40] Statista, "Anzahl der Smartphone-Nutzer in Deutschland in den Jahren 2009 bis 2015 (in Millionen)," ComSource, 24 09 2015. [Online]. Available: <http://de.statista.com/statistik/daten/studie/198959/umfrage/anzahl-der-smartphonenuutzer-in-deutschland-seit-2010/>. [Accessed 24 09 2015].
- [41] E. Corvee, S. Bak and F. Bremond, "People detection and re-identification for multi surveillance cameras," VISAPP - International Conference on Computer Vision Theory and Applications -2012, Februar 2012.

**AUTHOR**

Kristof Friess

Since 2016 – Doctoral Student, University of Applied Science Erfurt  
2014 – Master of Science, University of Applied Science Erfurt  
2011 – Bachelor of Science, University of Applied Science Erfurt



Volker Herwig

He is professor for business systems at the University for Applied Sciences in Erfurt, Germany. His focus is on mobil business systems and IT strategy. He worked 15 years in the industry in the USA and Germany.



# CHANNEL ESTIMATION FOR THE ISDB-T<sub>B</sub> FBMC SYSTEM USING NEURAL NETWORKS: A PROPOSAL OF APPLICATION OF BACK-PROPAGATION TRAINING ALGORITHM

Jefferson Jesus Hengles Almeida, P. B. Lopes,  
Cristiano Akamine and Nizam Omar

Postgraduate Program in Electrical Engineering and Computing  
Mackenzie Presbyterian University, Sao Paulo, Brazil

## **ABSTRACT**

*Due to the evolution of technology and the diffusion of digital television, many researchers have studied more efficient transmission and reception methods. This fact occurs because of the demand of transmitting videos with better quality using new standards such 8K SUPER Hi-VISION. In this scenario, modulation techniques such as Filter Bank Multi Carrier, associated with advanced coding and synchronization methods, are being applied, aiming to achieve the desired data rate to support ultra-high definition videos. Simultaneously, it is also important to investigate ways of channel estimation that enable a better reception of the transmitted signal. This task is not always trivial, depending of the characteristics of the channel. Thus, the use of artificial intelligence can contribute to estimate the channel frequency response, from the transmitted pilots. A classical algorithm called Back-propagation Training can be applied to find the channel equalizer coefficients, making possible the correct reception of TV signals. Therefore, this work presents a method of channel estimation that uses neural network techniques to obtain the channel response in the Brazilian Digital System Television, called ISDB-T<sub>B</sub>, using Filter Bank Multi Carrier.*

## **KEYWORDS**

*Channel estimation, Artificial intelligence, ISDB-T<sub>B</sub>, FBMC, Neural Networks.*

## **1. INTRODUCTION**

Digital TV standards in current use allow for the transmission of standard or high definition video content. However, consumers are demanding more resolution for more realistic experiences while watching TV. Therefore, researchers all over the world are working in the development of the concepts that will enable the broadcast of ultra-high definition content. These concepts include novel modulation schemes, powerful channel estimation, intelligent receptors, antenna arrays, etc. Filter Bank Multi Carrier (FBMC) is a modulation technique that has been applied as an alternative to the Orthogonal Frequency Division Multiplexing (OFDM) [1]. This trend is due to the fact that FBMC does not use the Cyclic Prefix (CP), increasing significantly the system data rate [2]. Thus, when FBMC is applied to digital television system such as Integrated Services

Digital Broadcasting Terrestrial B (ISDB-T<sub>B</sub>) associated with channel coding and synchronization techniques such as Low Density Parity Check (LDPC) and Bootstrap, 8K transmission may be made viable. Nevertheless, the transmission of more information per second using the same frequency bandwidth leads to an increase in bit error rate if improved channel estimation algorithms, powerful error correcting codes and novel equalizers are not used. However, the channel estimation becomes a little more complex, due to the characteristics of the used filters and degrading components present on the channel [3]. In this article, an intelligent channel estimation algorithm using Neural networks is reported.

The channel estimation is a crucial stage for the perfect reception of digital TV signals because of the interferences that are generated by several sources on the channel. In the specific case of the terrestrial broadcast, these are Additive White Gaussian Noise (AWGN), multipath, and others [4]. Therefore, it is necessary to use different techniques and processes that make possible the removal and minimization of this effects, allowing the reception of the transmitted signal in an appropriated way.

Several techniques can be used to estimate the frequency response of the channel. Among them, we can highlight those that use pilots associated to interpolation methods that adequately minimize the AWGN and multipath effect [5]. However, this process is not always trivial, and can be improved with the use of Artificial Intelligence (AI).

The AI can be understood as a set of algorithms to solve complex problems, being able to resolve, to make decisions, and to develop a method of learning, according to the situation to which it is applied [6]. In this context, Neural Networks (NN) are used to solve problems through the simulation the connection between brain neurons, using specific activation and training algorithms [7].

The herein proposed channel estimation method uses a NN, trained through the Back-propagation algorithm to calculate the channel response and to permit the correct equalization in a scenario with AWGN and multi paths. This technique is applied to an FBMC version of the ISDB-T<sub>B</sub> digital TV standard. The overall system is simulated on GNUR adio environment. The presented results show the feasibility of this new system even when transmission id performed in a channel with severe multipath interference.

## 2. ISDB-T<sub>B</sub>

The Brazilian Digital Television System (SBTVD), ISDB-T<sub>B</sub>, is the terrestrial digital TV standard adopted by 18 countries in the world. It employs a bandwidth of 6, 7, or 8 MHz and can transmit One-Segment (1SEG), Standard Definition Television (SDTV) and High Definition Television (HDTV), according to the combination chosen for the 13 Segments available. In the Figure 1 is possible to see the bandwidth segmentation used in 6 MHz. The 3 modes of operation have different parameters that are detailed in Table I [8].

Table 1. ISDB-T<sub>B</sub> transmission parameters.

Mode	Parameters				
	<i>Carriers</i>	<i>Useful Carriers</i>	<i>Pilots</i>	<i>Symbol period</i>	<i>IFFT</i>
1	1405	1248	157	0,252 ms	2048
2	2809	2496	313	0,504 ms	4096
3	5617	4992	625	1,008 ms	8192

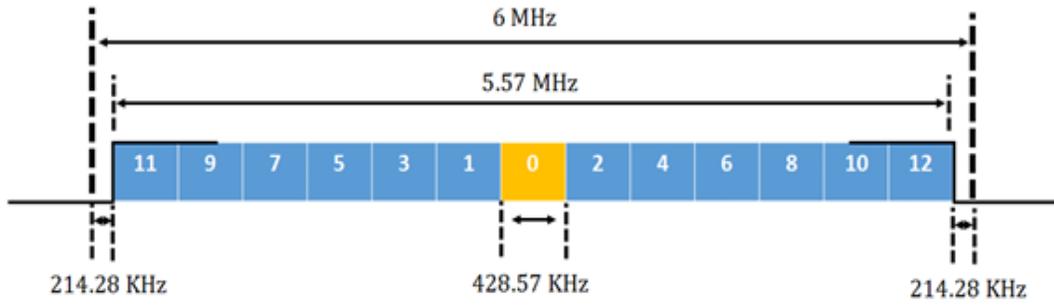


Figure 1. Segmentation of the ISDB-T<sub>B</sub> channel.

### 3. FBMC

The Filter Bank Multi Carrier (FBMC) modulation technique consists of dividing the available band into small equally spaced small segments [9]. For this purpose, filter banks complying with (1) are used.

$$B_k(f) = H\left(f - \frac{k}{M}\right) = \sum_{i=0}^{L-1} h_i e^{-j2\pi i\left(f - \frac{k}{M}\right)} \quad (1)$$

where,  $f$  is the frequency,  $M$  is the number of subcarriers,  $L$  is the number of filter coefficients, and  $k = 0 \dots M - 1$ .

Applying the Z transform, the polyphase decomposition, and making  $W_M = e^{-\frac{2j\pi}{M}}$ , (2) can be derived.

$$B_k(z) = \sum_{p=0}^{M-1} W_M^{-kp} z^{-p} H_p(z^M) \quad (2)$$

By expressing this equation in matrix notation, (3) is obtained.

$$\begin{bmatrix} B_0(z) \\ \vdots \\ B_{M-1}(z) \end{bmatrix} = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & W_M^{-(M-1)^2} \end{bmatrix} \begin{bmatrix} H_0(z^M) \\ \vdots \\ z^{-(M-1)} H_{M-1}(z^M) \end{bmatrix} \quad (3)$$

The implementation of this equation is depicted in the block diagram in the Figure 2.

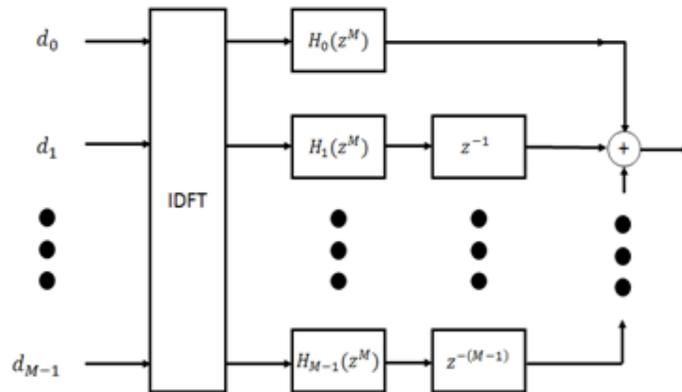


Figure 2. FBMC system model.

The filters are designed according to the zero inter-symbol interference Nyquist criterion to avoid phase and amplitude distortions [10]. FBMC employs Offset Quadrature Modulation (OQAM), so that the orthogonality is obtained between symbols and not between subcarriers [11]. Thus, it is not necessary to use the CP, making it possible to increase the data rate of the system.

#### 4. PILOT BASED CHANNEL ESTIMATION

To estimate the frequency response of the channel, ISDB-T<sub>B</sub> performs the constant transmission of pilots. The position of the pilots depends on a Pseudorandom Binary Sequence (PRBS) sequence that has a generator polynomial equal to  $x^{11} + x^9 + 1$  [8].

After the transfer function ( $H_p$ ) is found using (4), where  $Y_p(k)$  and  $X_p(k)$  are the pilot amplitudes received and transmitted through the  $k$ -th subcarrier, an interpolation method that can be linear, cubic, among others, is used to estimate the responses of the other subcarriers.

$$H_p(k) = \frac{Y_p(k)}{X_p(k)} \quad (4)$$

In the case of linear estimation, we use (5).

$$H(k) = (1 - a) \cdot H_p(k) + a \cdot H_p(k + 1) \quad (5)$$

where  $a$  is a constant determined by the relation between the distance of the position of the subcarrier where it is desired to estimate the response of the channel to the position of the nearest pilot.

In the case of cubic interpolation, (6) is used.

$$H(k) = A(a) \cdot H_p(k) + B(a) \cdot H_p(k + 1) + C(a) \cdot z(k) + D(a) \cdot z(k + 1) \quad (6)$$

where  $A(a)$ ,  $B(a)$ ,  $C(a)$ , and  $D(a)$  are constants related to  $a$  and  $z(m)$  is the second derivative obtained from the pilot information matrix [12].

#### 5. NEURAL NETWORKS

Neural networks can be understood as algorithms that seek to simulate the functioning of the human brain, starting from the construction of small computational entities that act as a human neuron [13]. To do so, we use units called perceptrons (Figure 3) which have as input parameters an input ( $x$ ), a gain ( $w$ ), an activation function ( $f$ ) and an output ( $y$ ).



Figure 3. Perceptron.

The activation function can be linear or not depending on the desired application. The most common are the logarithm and the sigmoid. Perceptrons can be combined to form layers that are interconnected to generate larger and more complex networks. For the network to work correctly, it is necessary to perform the network training, using a set of known inputs and outputs, so that the gains are properly adjusted and the actual inputs generate the desired responses [14]. Among the training techniques, one of the most used is the Back-propagation Training Algorithm.

### 5.1. Back-propagation Training Algorithm

This technique uses a generalization of the Least Mean Square [15]. The activation function is defined as (7).

$$f(\alpha) = \frac{1}{1-e^{-\alpha}} \quad (7)$$

Initially random weights ( $w$ ) are defined for the inputs ( $x$ ). Then from the desired response ( $d$ ) the error is calculated by (8).

$$erro = 0.5 \cdot (d - y)^2 \quad (8)$$

where  $y$  is the output of the activation function.

Then the weights are updated, using (9).

$$w(t + 1) = w(t) - \delta \cdot erro \cdot x \quad (9)$$

where  $t$  is the previous iteration of the algorithm has been set and  $\delta$  is a chosen gain. Finally, the algorithm is repeated until the desired response is obtained at the output of the system and the weights are properly adjusted.

## 6. PREVIOUS DEVELOPED MODEL OF ISDB-T<sub>B</sub> USING FBMC

In [2] and [16], a modified ISDB-T<sub>B</sub> system using FBMC was developed, using GNU Radio Companion (GRC) as simulation environment. The channel estimation algorithm used did not employ any Artificial Intelligence feature. For this reason, the present work expands those articles by using a different approach based on Neural Networks.

### 6.1. GRC

The GRC is a computational tool that allows the development of processing blocks to for simulating communications systems [17]. It is an open source and free software that makes possible the interface between the created model and software radio peripherals [18]. It uses a Graphical User Interface (GUI) that facilitates the software handling [19] [20].

The block codes are created using the C/C++ or Phyton languages and the interconnection among these blocks is described only in Phyton [21].

The processed data sources on GRC are of complex (8 bytes), float (4 bytes), int (2 bytes), or byte (1 byte) types. The used terminology of GRC is presented on Table II [22].

Table 2. GRC terminology.

Name	Definition
Block	Processing Unit with ins or outs
Port	Block in or out
Source	Data generator
Sink	Data consumer
Connection	Data flow from an output to an input
Flow Graph	Set of blocks and connections
Item	Data unit
Stream	Continuous flow of items

Name	Definition
IO Signature	In and out description

Then using GRC and programming language the system can be simulated.

## 6.2. Flow Graph of ISDB-T<sub>B</sub> FBMC

The transmitter implemented on GRC is presented in the diagram shown on Figure 4.

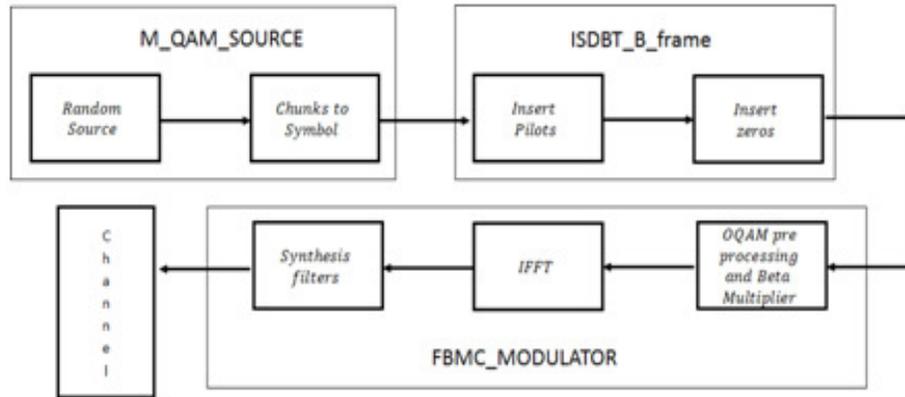


Figure 4. ISDB-T<sub>B</sub> FBMC transmitter.

As it can be seen at the transmission side, an information source generates data that is modulated, formatted according to the standard, processed by OQAM pre-processing, multiplied by Beta, modulated through the IFFT and Synthesis filters and transmitted through the channel.

The receiver diagram is shown in the Figure 5.

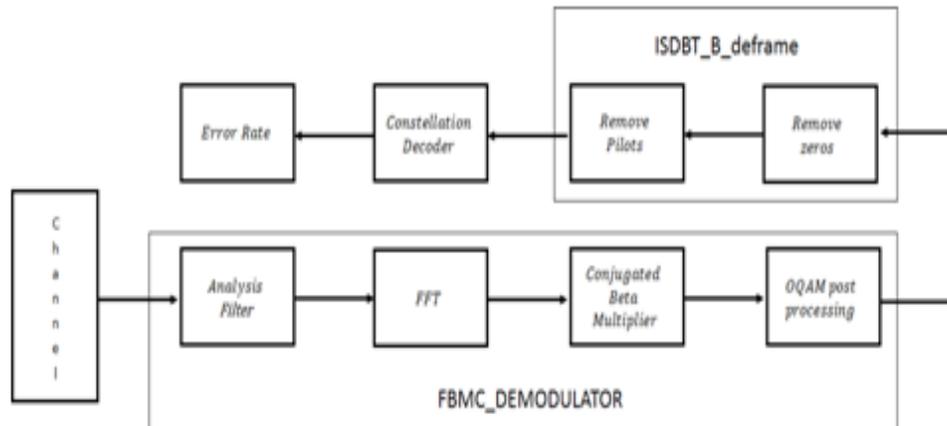


Figure 5. ISDB-T<sub>B</sub> FBMC receiver.

At the reception, the data goes through the analysis filters and FFT, multiplied by conjugated Beta, processed in OQAM post processing. After the zeros and pilots are removed, the data is decoded to calculate the Bit Error Rate (BER).

The GRC environment flow graph is depicted in Figure 6.

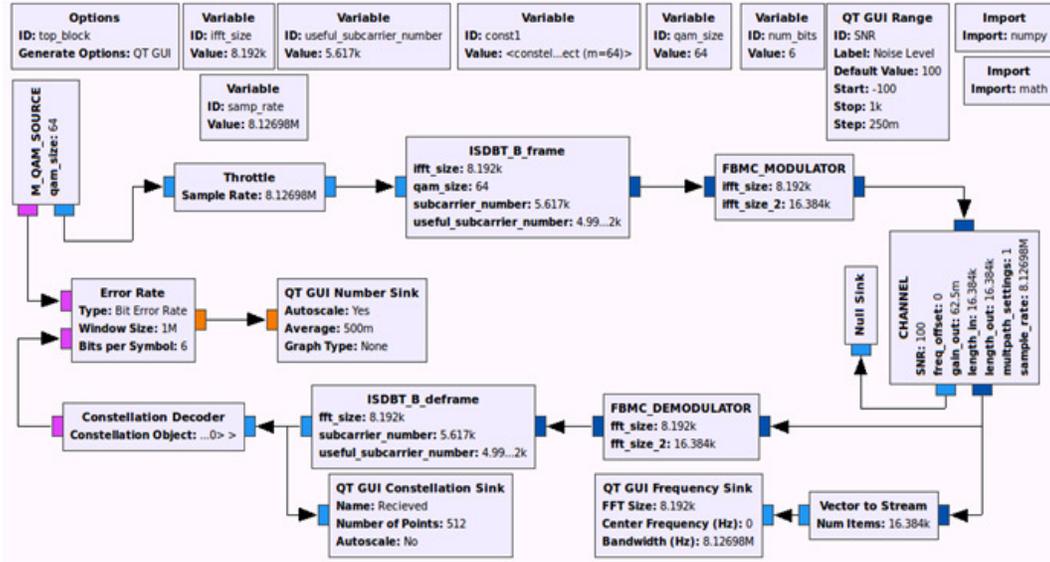


Figure 6. Flow Graph of ISDB-T<sub>B</sub>FBMC.

## 7. PROPOSED IA ESTIMATION METHOD

To accomplish the AI estimation, a simple NN with one perceptron for each real and imaginary part of received symbol is used. When the system initiates, four FBMC symbols are sent as training sequence and the weights of NN are trained using the Back-propagation method. Then regular operation starts and the received data symbols are equalized by the trained system.

The Flow Graph used was shown in Figure 5. Inside the “ISDBT\_B\_deframe” hierarchical block, three different channel estimators were implemented: the linear interpolation and the cubic interpolation, both at the time and frequency, and a neural network estimator trained with the Back-propagation technique (Figure 7).

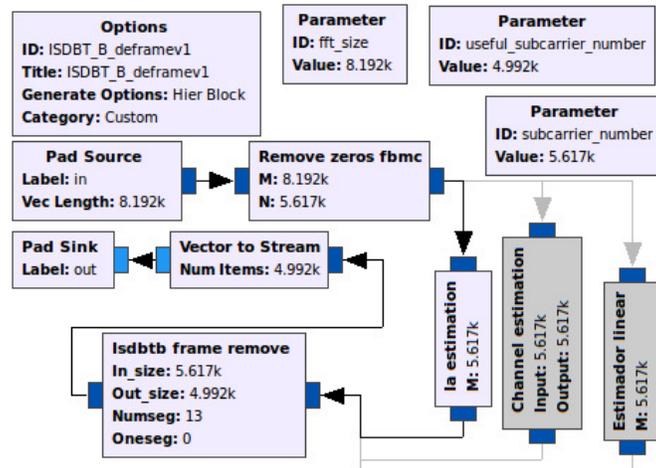


Figure 7. Content of the hierarchical block “ISDBT\_B\_deframe”.

Thus, it was connected each channel estimator at the system and the results could be collected.

## 5. RESULTS

The analysis were made using the ISDB-T<sub>B</sub> FBMC in mode 3 as in Table I. The Bit Error Rate (BER) curves were observed on two scenarios. The first is characterized by AWGN and modulation level equal to 64 (64-QAM) or 6 bits per symbol (Figure 8). The second uses the Brazil A digital TV channel model [21], which applies 6 paths with 0, 0.15, 2.2, 3.05, 5.86, and 5.93 microseconds of delay and 0, 13.8, 16.2, 14.9, 13.6, and 16.4dB of attenuation respectively, using the modulation level 4 (4-QAM) (Figure 9). In this last case, it was necessary to reduce the QAM modulation due to the long time required to perform the real-time simulation.

It can be observed that the use of neural networks brought to the system an increase of robustness at  $10^{-5}$  BER level around to 2.1dB in the case only of AWGN and around to 2dB in the case which there is AWGN and multipath.

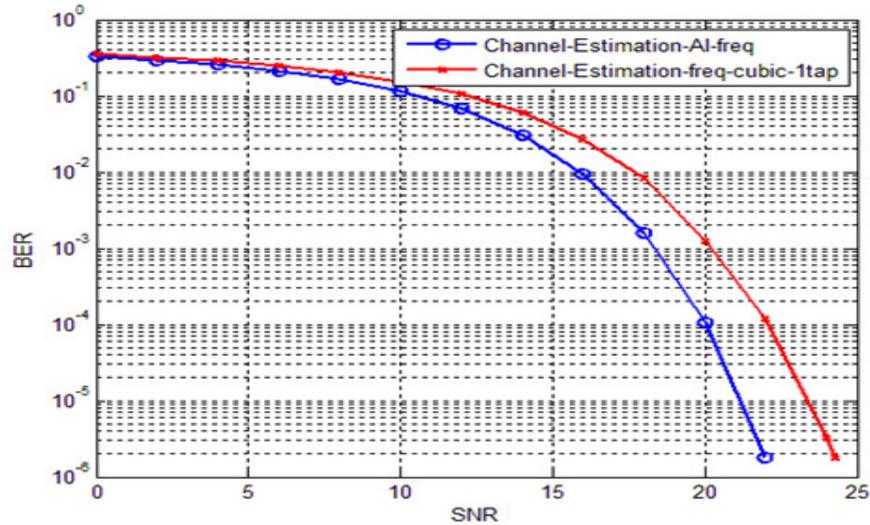


Figure 8. BER curves of ISDB-T<sub>B</sub>FBM Cusing 64-QAM with AWGN inserted.

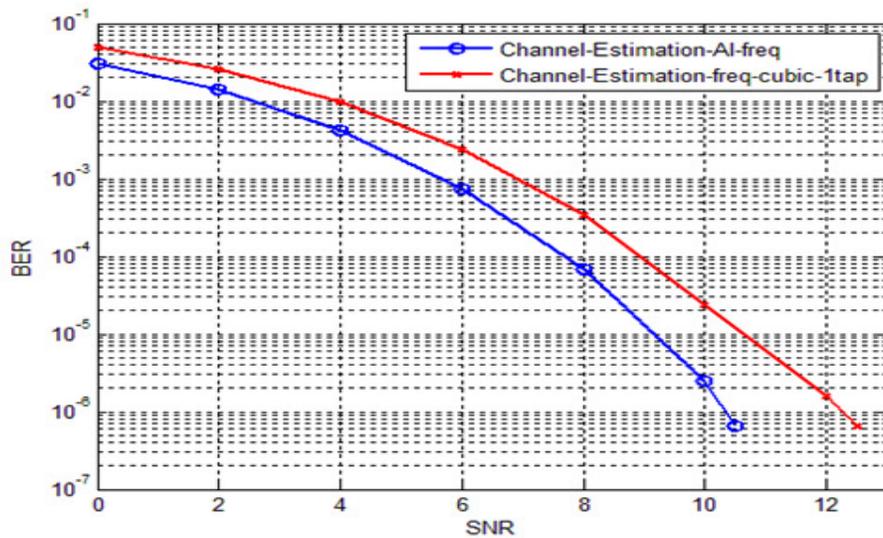


Figure 9. BER curves of ISDB-T<sub>B</sub>FBM Cusing 4-QAM with AWGN and multipath inserted.

## 6. CONCLUSION

Current digital TV standards were established to enable the broadcast of standard or high definition video. Nevertheless, nowadays consumers are demanding even higher definition content. For this reason, researchers are working on new standards that will enable a higher information transmission rate in terms of bits/s/Hz. This is the case of FBMC which was also proposed for forthcoming 5G cellular standard. But this increase in bit rate on the same frequency bandwidth implies in an increase in bit error rate if novel channel estimation and equalization are not created.

The use of artificial intelligence applied to channel estimators for FBMC opens a new field of research. The possibility of employing smart algorithms that can learn even in the presence of nonlinear interference is paramount to the success of more spectrally efficient modulation techniques.

In this paper, it was showed that the application of a simple neural network to the problem of channel estimation in the FBMC modified ISDB-Tb digital TV standard is feasible. The presented technique achieved an increase in 15% in robustness in a channel with several multipath interferences. It was also shown that the Back-propagation training algorithm allows the estimation of the channel frequency response and contributes to minimize the bit error rate. In future, other kind of neural networks will be investigated, such as a recursive network, since it can improve the results and required computing effort.

## ACKNOWLEDGEMENTS

The authors would like to thank the MACKPESQUISA, Coordination for the Improvement of Higher Level Personnel (CAPES) and National Counsel of Technological and Scientific Development (CNPq) for the partial financial subsidies for this research.

## REFERENCES

- [1] Bellanger M, Mattera D, Tanda M. A filter bank multicarrier scheme running at symbol rate for future wireless systems. *Wireless Telecommunications Symposium (WTS), 2015*, 2015; 1–5, doi:10.1109/WTS.2015.7117247.
- [2] Almeida J. J. H., Akamine C, Lopes P. B. A proposal for the next generation of isdb-tb using fbmc in a sdr implementation on gnu radio environment. *2016 8th IEEE Latin-American Conference on Communications (LATINCOM), 2016*; 1–6, doi:10.1109/LATINCOM.2016.7811601
- [3] Bellanger M. FBMC physical layer: a primer. *PHYDYAS*, January 2010; :1–31 URL: [http://www.ict-phydyas.org/teamspace/internal-folder/FBMC-Primer\\_06-2010.pdf](http://www.ict-phydyas.org/teamspace/internal-folder/FBMC-Primer_06-2010.pdf).
- [4] Andreas M. F. *Wireless Communications*. Second edition, John Wiley & Sons Ltd: California, 2011.
- [5] Ishini, A. K., Akamine, C.; 2009. Técnicas de estimação de canal para o sistema isdb-tb. *Revista de Radiodifusão*.
- [6] Russel S, Norvig, P. *Inteligência Artificial*. 2a ed., Rio de Janeiro: Campus, 2004.
- [7] Lippmann R, An introduction to computing with neural nets, in *IEEE ASSP Magazine*, vol. 4, no. 2, pp. 4-22, Apr 1987. doi: 10.1109/MASSP.1987.1165576
- [8] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 15601: *Televisão digital terrestre - Sistema de transmissão*. Rio de Janeiro, 2008.

- [9] Cherubini G, et al. Filter bank modulation techniques for very high speed digital subscriber lines. *IEEE Communications Magazine*, v. 38, n. 5, p. 98-104, May 2000.
- [10] Bellanger M. G. Specification and design of a prototype filter for filter bank based multicarrier transmission. *Acoustics, Speech, and Signal Processing*, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on, vol. 4, 2001; 2417–2420 vol.4, doi: 10.1109/ICASSP.2001.940488.
- [11] Siohan P, Siclet, C, Lacaille N. Analysis and design of OFDM/OQAM systems based on filterbank theory. *IEEE Transactions on Signal Processing*, v. 50, n. 5, p. 1170-1183, May 2002. ISSN 1053-587X.
- [12] Kang S. G, Ha Y. M, Joo E. K. A comparative investigation on channel estimation algorithms for ofdm in mobile communications. *IEEE Transactions on Broadcasting*, v. 49, n. 2, p. 142-149, June 2003.
- [13] Haykin S. *Redes Neurais: Princípios e Prática*. Bookman, 2001, Hamilton, Ontário, Canadá. 2ª Edição.
- [14] Burse K, Yadav R. N, Shivastava S. C. Channel Equalization Using Neural Networks: A Review, in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 3, pp. 352-357, May 2010.
- [15] Lippmann, R. An introduction to computing with neural nets, in *IEEE ASSP Magazine*, vol. 4, no. 2, pp. 4-22, Apr 1987. doi: 10.1109/MASSP.1987.1165576
- [16] Almeida J. H, Akamine C, Lopes P. B. A proposal for the next generation of ISDB-Tb using FBMC in a SDR implementation on GNU radio environment, *IEEE Latin-American Transactions*. 2017, July. Available  
in: <http://www.revistaieeela.pea.usp.br/issues/vol15issue7July2017/Vol15issue7July2017TLA.htm>
- [17] Maheshwarappa MR, Bridges CP. Software defined radios for small satellites. *Adaptive Hardware and Systems (AHS)*, 2014 NASA/ESA Conference on, 2014; 172–179, doi:10.1109/AHS.2014.6880174.
- [18] Stoica RA, Severi S, de Abreu GTF. On prototyping iieee802.11p channel estimators in real-world environments using gnuradio. 2016 IEEE Intelligent Vehicles Symposium (IV), 2016; 10–15, doi:10.1109/IVS.2016.7535356.
- [19] Larroca F, Guridi PF, Sena GG, Gonzalez-Barbone V, Belzarena P. gr-isdbt: An isdb-t 1-segment receiver implementation on gnu radio. *Computing Conference (CLEI)*, 2015 Latin American, 2015; 1–8, doi:10.1109/CLEI.2015.7360050.
- [20] Müller A. DAB Software Receiver Implementation. Technical Report 2008. URL: <http://people.ee.ethz.ch/~andrmuel/files/gnuradio/>.
- [21] Vachhani K, Mallari RA. Experimental study on wide band fm receiver using gnuradio and rtl-sdr. *Advances in Computing, Communications and Informatics (ICACCI)*, 2015 International Conference on, 2015; 1810–1814, doi:10.1109/ICACCI.2015.7275878.
- [22] Gnuradio. <http://gnuradio.org/redmine/projects/gnuradio/wiki> 2016. Accessed in 07/11/2016.
- [23] ITU Radio communication Study Groups: Document 6E/TEMP/131-E, Guidelines And Techniques For The Evaluation Of DTTB Systems, 19 March 2003.

## AUTHORS

**Jefferson Jesus Hengles Almeida** was born in Cotia, on May 1992. Received his B.Sc. degree in Electrical Engineering from Mackenzie Presbyterian University, São Paulo, Brazil, in 2014. Received his M.Sc Degree in Electrical Engineering from Mackenzie Presbyterian University, São Paulo, Brazil, in 2016. He is currently studying his Ph.D degree in Electrical Engineering in Mackenzie Presbyterian University. His current research involves broadcasting areas, digital television transmission systems studies, and software defined radio.



**Paulo Batista Lopes** was received the B.Sc. and M.Sc. in EE from the Universidade Federal do Rio de Janeiro, Brazil, in 1978 and 1981, respectively, and the Ph.D. in EE from Concordia University, Montreal, Canada, in 1985. From 1985 to 1988, he was with Elebra and CMA, two Brazilian companies, working on the design of several Communication equipments. From 1988 to 1999, he was with Texas Instruments as a DSP specialist. In 1999, he moved to Motorola-SPS (later to become Freescale Semiconductor) as a Sales and Application manager. Since 2009, he has been with Universidade Presbiteriana Mackenzie as a professor in the School of Engineering. His research interests are Circuit Theory, Digital Signal Processing, and Analog Circuit Design.



**Cristiano Akamine** received his B.Sc. degree in Electrical Engineering from Mackenzie Presbyterian University, São Paulo, Brazil, in 1999. He received his M.Sc. and Ph.D. degree in Electrical Engineering from the State University of Campinas (UNICAMP), São Paulo, Brazil, in 2004 and 2011 respectively. He is a professor of Embedded Systems, Software Defined Radio and Advanced Communication Systems at Mackenzie Presbyterian University. He has been a researcher in the Digital TV Research Laboratory at Mackenzie Presbyterian University since 1998, where he had the opportunity to work with many digital TV systems. His research interests are in system on chip for broadcast TV and Software Defined Radio.



**Nizam Omar**, Mechanical Engineer ITA 1974. Master in Applied Mathematics, ITA 1979, Ph.D. in Computer Science - PUC RIO 1989. He is Professor at the Mackenzie Presbyterian University in Artificial Intelligence and its applications in Education, Engineering and Economics.



*INTENTIONAL BLANK*

# MULTI RESOLUTION LATTICE DISCRETE FOURIER TRANSFORM (MRL-DFT)

Fred L. Fontaine

Department of Electrical Engineering, Cooper Union, New York City, USA

## **ABSTRACT**

*In many imaging applications, including sensor arrays, MRI and CT, data is often sampled on non-rectangular point sets with non-uniform density. Moreover, in image and video processing, a mix of non-rectangular sampling structures naturally arise. Multirate processing typically utilizes a normalized integer indexing scheme, which masks the true physical dimensions of the points. However, the spatial correlation of such signals often contains important information. This paper presents a theory of signals defined on regular discrete sets called lattices, and presents an associated form of a finite Fourier transform denoted here as multiresolution lattice discrete Fourier transform (MRL-DFT). Multirate processing techniques such as decimation, interpolation and polyphase representations are presented in a context which preserves the true spatial dimensions of the sampling structure. Moreover, the polyphase formulation enables systematic representation and processing for sampling patterns with variable spatial density, and provides a framework for developing generalized FFT and regridding algorithms.*

## **KEYWORDS**

*Image Processing, Video, Medical Imaging, Multirate Signal Processing, Fourier Transform.*

## **1. INTRODUCTION**

A number of imaging applications involve data located on non-rectangular grids. Examples include tomographic schemes such as medical computerized tomography (CT) and certain synthetic aperture radar (SAR) systems, in which data is often acquired on a polar grid [1,2]. In magnetic resonance imaging (MRI), data is acquired in the two- or three-dimensional spatial frequency domain on sometimes quite complex patterns, such as the spiral shown in Figure 1. Different types of video signals have spatio-temporal data on different sampling structures, such as interlaced versus progressive. As illustrated with the spiral pattern, these irregular sampling patterns often have variable spatial density. Many common signal processing applications such as filtering and spectral analysis are difficult to accomplish, and conversion among sampling structures without introducing artifacts can be challenging [1,3]. Data is often converted to a uniform rectangular grid, through an interpolation process called regridding, to facilitate processing, as in [4], but there may be significant artifacts caused by density mismatch.

Some methods have been developed for certain irregular structures as special cases, as in [5,6,7], as well as for uniform but non-rectangular sampling patterns, called lattices, as in [3,8]. A normalized integer indexing scheme is typically used, which does not reflect true underlying geometric factors such as distance and direction. However, for many multidimensional signals, significant information is contained in the spatial correlation properties. For example, in [9], wireless network nodes are located on scattered points in a lattice and information regarding geometric location is vital for proper signal processing. In [10], a lattice based approach is used

to analyse interference signals. In certain digital communication applications, as in [11,12,13], signal sets are defined on lattices, and the geometric properties of the lattices are central to the development of receiver algorithms and performance analysis.

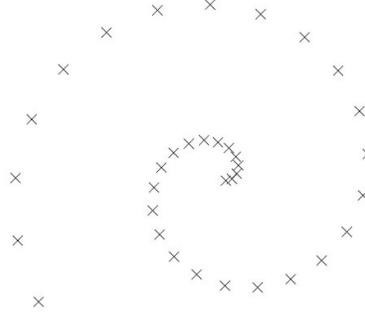


Figure 1: Sample points along a spiral.

Although conversion between different sampling structures is common, usually a single pattern containing various spatial sampling densities is not employed. This paper presents a systematic approach for dealing with multidimensional sampling patterns, including possibly with variable density. The results are formulated using physical coordinate systems for the spatial and frequency domains to facilitate development of algorithms such as filtering and interpolation (regridding) for sampling structure conversion. In particular, we study a finite Fourier transform suitable for this context, denoted here as the multiresolution lattice discrete Fourier transform (MRL-DFT). The MRL-DFT is used to guide the design of filters, sampling structure conversions, and polyphase structures, as well as the study of lattices to facilitate manipulating sampling patterns with variable densities.

We first describe lattices as sampling structures, viewed both in the standard normalized coordinates [3] and in physical coordinate systems, and develop an associated Fourier theory. We then discuss various signal processing operations and properties of lattice structures based on the MRL-DFT, and then provide a detailed example of how the methods developed here can facilitate various signal processing applications.

## 2. LATTICES

In this section, we introduce some notation and summarize basic definitions and results regarding lattices, and describe the topic under study in this paper.

### 2.1. Notation and Definitions

We consider complex valued signals whose domains are certain subsets of  $D$ -dimensional real space,  $\mathcal{R}^D$ . If  $\mathcal{S} \subset \mathcal{R}^D$ ,  $\vec{r} \in \mathcal{R}^D$  and  $A$  is a  $D \times D$  matrix, then  $A\mathcal{S} + \vec{r} = \{A\vec{s} + \vec{r} : \vec{s} \in \mathcal{S}\}$ . If  $\mathcal{S}$  is a finite set,  $|\mathcal{S}|$  denotes the number of elements in  $\mathcal{S}$ .

We can associate vectors with column vectors, and in this form, the dot product (or inner product) of vectors  $\mathcal{R}^D$  can be expressed as  $\langle \vec{a}, \vec{b} \rangle = \vec{a} \cdot \vec{b} = \vec{a}^T \vec{b}$ . In this paper, we will work with a non-radian frequency domain; for example, if the spatial vector  $\vec{r}$  has units of meters, the frequency vector  $\vec{f}$  has units of (meter)<sup>-1</sup>. This suggests the following notation:

$$e(\alpha) = \exp(j2\pi\alpha)$$

Note that  $e(\alpha)=1$  iff  $\alpha \in \mathcal{Z}$ . Then the Fourier transform of a signal  $x(\vec{r})$  can be thought of as a representation of the signal as a superposition of sinewaves of the form  $e(\langle \vec{f}, \vec{r} \rangle)$  for a suitable collection of frequency vectors  $\vec{f}$ .

**Definition 1:** A  $D$ -dimensional lattice  $\mathcal{L} \subset \mathcal{R}^D$  is a set of the form:

$$\mathcal{L} = V\mathcal{Z}^D = \{\vec{r} = V\vec{n} : \vec{n} \in \mathcal{Z}^D\}$$

for some  $D \times D$  nonsingular matrix  $V$ .

In this paper, all matrices we consider (e.g., lattice generators) are invertible  $D \times D$  real. Given a lattice  $\mathcal{L}$ , the associated matrix  $V$ , called a generator of the lattice, is not unique. We denote a lattice generated by  $V$  as  $\mathcal{L}_V$ . A signal  $x(\vec{r})$  is said to have support on a lattice  $\mathcal{L}$  if it is defined (specified) for all  $\vec{r} \in \mathcal{L}$  and is either undefined or has value 0 for any point  $\vec{r} \notin \mathcal{L}$ . A signal is periodic with respect to a lattice  $\mathcal{L}$  if  $x(\vec{r}) = x(\vec{r} + \vec{\xi})$  for all  $\vec{\xi} \in \mathcal{L}$ , and if  $V$  is a generator of  $\mathcal{L}$  then we say  $x$  is periodic with respect to  $V$ .

An integer matrix  $M$  is called unimodular if  $|\det M| = 1$  and this occurs iff  $M^{-1}$  is also an integer matrix. The following lemmas summarize basic results regarding lattices [3].

**Lemma 1:**  $\mathcal{L}_U$  is a sublattice of  $\mathcal{L}_V$ , that is,  $\mathcal{L}_U \subset \mathcal{L}_V$  iff  $U=VM$  for some integer matrix  $M$ , and  $\mathcal{L}_U = \mathcal{L}_V$  iff  $U=VE$  for some unimodular matrix  $E$ .

**Lemma 2:** Points in a lattice form an abelian group under vector addition, and a sublattice is a subgroup. The translated sets of a sublattice,  $\mathcal{L}_U + \vec{r}$  for  $\vec{r} \in \mathcal{L}_V$ , called cosets, form a group under set addition, called the quotient group and denoted  $\mathcal{L}_V/\mathcal{L}_U$ , and then  $|\mathcal{L}_V/\mathcal{L}_U| = |\det M|$ . As a consequence, the value of  $|\det M|$  does not depend on the choice of generators for the lattices.

The density of a lattice, denoted by  $\rho(\mathcal{L})$ , is the number of lattice points per unit volume, and is given by  $\rho(\mathcal{L}) = \frac{1}{|\det V|}$ . Thus, a sublattice is sparser (less dense) than the parent lattice.

Sampling patterns with variable spatial density arise by combining points from different sublattices or their cosets in different regions. In Figure 2, the central region contains points from a dense lattice, and the outer region contains only points in a sublattice. We refer to this as a multiresolution sampling pattern.

For a lattice  $\mathcal{L}_V$ , a corresponding unit cell  $\mathcal{U}_V \subset \mathcal{R}^D$  is a set of points such that, for all  $\vec{r} \in \mathcal{R}^D$ , there is a unique vector  $\vec{r}_0 \in \mathcal{L}_V$  such that  $\vec{r} - \vec{r}_0 \in \mathcal{U}_V$ . Two examples of unit cells are the sets  $V[0,1)^D$ , called the fundamental parallelepiped (FPD) of  $V$ , and  $V[-\frac{1}{2}, \frac{1}{2})^D$ . Regarding the cosets (distinct translations) of  $\mathcal{L}_V$  in  $\mathcal{R}^D$ , a unit cell constitutes a minimal complete collection of coset representatives. Addition of points in the unit cell modulo the lattice corresponds to the quotient group  $\mathcal{R}^D/\mathcal{L}_V$ . The volume of a unit cell is  $|\det V|$ , the reciprocal of the density of the lattice.

If  $\mathcal{L}_U \subset \mathcal{L}_V$ , we denote by  $\mathcal{J}(U, V)$  a complete set of coset representatives. That is, for every coset of  $\mathcal{L}_U$  in  $\mathcal{L}_V$ , there is a unique element  $\vec{r} \in \mathcal{J}(U, V)$  such that the coset can be expressed as  $\mathcal{L}_U + \vec{r}$ . We call such  $\mathcal{J}(U, V)$  an index set.

**Lemma 3:** If  $\mathcal{L}_U \subset \mathcal{L}_V$  with  $U=VM$ , then  $\mathcal{J}(U, V)$  is a valid index set iff it has the form:

$$\mathcal{J}(U, V) = \mathcal{U}_U \cap \mathcal{L}_V$$

for some unit cell  $\mathcal{U}_U$  of  $U$ . Moreover, every index set contains exactly  $|\det M|$  points.

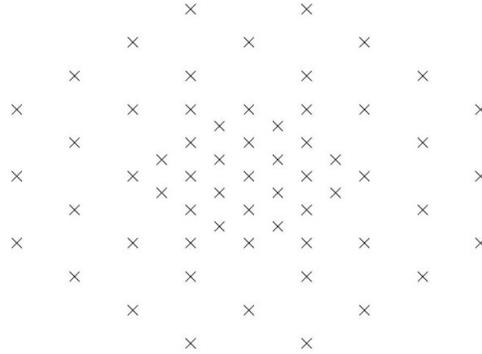


Figure 2: A multiresolution sampling pattern.

The quotient group  $\mathcal{L}_V/\mathcal{L}_U$  can be interpreted as addition of points in an index set  $J(U, V)$  modulo the lattice  $\mathcal{L}_U$ . Figure 3 shows a sublattice (with points marked with 'o') and a coset (with points marked with 'x') that together comprise a denser lattice. In this case, the quotient group contains two elements.

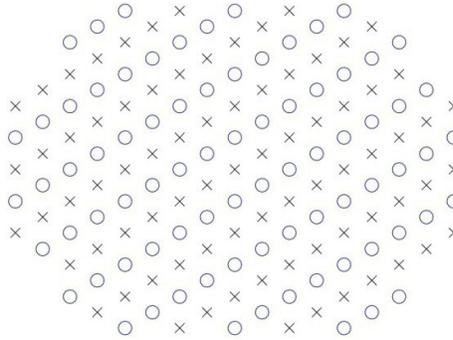


Figure 3: A sublattice ('o') and its coset ('x') together forming a denser lattice.

For a matrix  $A$ , let  $\tilde{A} = A^{-T}$ . Given a lattice  $\mathcal{L}_V$ , we call  $\mathcal{L}_{\tilde{V}}$  the reciprocal lattice, and  $\mathcal{U}_{\tilde{V}}$  the reciprocal unit cell. The density of the reciprocal lattice is the reciprocal of the density of the original lattice. A dense lattice will correspond to small unit cells, a sparse reciprocal lattice and large reciprocal unit cells. Note that if  $U = VM$  then  $\tilde{V} = \tilde{U}M^T$ . Hence, whereas  $\mathcal{L}_U \subset \mathcal{L}_V$ , we get  $\mathcal{L}_{\tilde{V}} \subset \mathcal{L}_{\tilde{U}}$ . Also,  $|\mathcal{L}_V/\mathcal{L}_U| = |\mathcal{L}_{\tilde{U}}/\mathcal{L}_{\tilde{V}}| = |\det M|$ .

## 2.2. Signals on Lattices with Finite Extent

Consider a signal on a lattice  $\mathcal{L}_V$  with a finite number of nonzero values, at points that are confined to within a unit cell  $\mathcal{U}_U$  of a sublattice. In this case, the signal has support within an index set  $J(U, V)$ , and by extending these values periodically with respect to  $\mathcal{L}_U$ , we can now interpret this as a signal over the quotient group  $\mathcal{L}_V/\mathcal{L}_U$ . There exists a finite Fourier transform that can be defined for such signals, and the focus of this paper is its study. We call this Fourier transform the **multiresolution lattice discrete Fourier transform (MRL-DFT)**.

The MRL-DFT results in a frequency domain signal that has support in  $\mathcal{L}_{\tilde{U}}$  and is  $\mathcal{L}_{\tilde{V}}$ -periodic, i.e., a signal over  $\mathcal{L}_{\tilde{U}}/\mathcal{L}_{\tilde{V}}$ . The spectrum is thus identified by its values over an index set  $J(\tilde{V}, \tilde{U})$ ,

the complete set of points in the reciprocal lattice for  $U$  that lie within a reciprocal unit cell for  $V$ . This corresponds to a generalized result of sampling theory: the density of samples in the original (here, spatial) domain determines the extent (bandwidth) in the frequency domain, and the extent in the original domain (spatial size) determines the density of points in the frequency domain.

Associated with this, we develop several multirate signal processing techniques, such as decimation and interpolation, polyphase components, filtering (i.e., convolution) and associated frequency domain concepts. The multiresolution approach is to consider not only a single sparse (low resolution) sublattice of a dense (high resolution) lattice, but also chains of intervening lattices with intermediate densities, that is:

$$\mathcal{L}_U \subset \mathcal{L}_{U_1} \subset \mathcal{L}_{U_2} \subset \dots \subset \mathcal{L}_{U_k} \subset \mathcal{L}_V$$

This study facilitates development of applications such as representing signals with varying sampling densities in different regions of space, formulating fast algorithms for computing convolution and Fourier transforms, and obtaining techniques for regridding, i.e., interpolating data measured at possible irregularly sampled points to multiresolution lattices.

A theory of multidimensional multirateprocessing has been developed, for example in [3]. However, in that context, it is traditional to represent the signals as having support in  $\mathcal{Z}^D$ . Consider a lattice  $\mathcal{L}_V \subset \mathcal{R}^D$  with an identified generator  $V$ . As  $\mathcal{L}_V = V\mathcal{Z}^D$ , if  $x$  is a signal over  $\mathcal{L}_V$  then the value  $x(\vec{r})$  at a point  $\vec{r} \in \mathcal{L}_V$  can be associated with the unique integer index vector  $\vec{n}$  such that  $\vec{r} = V\vec{n}$ . Let us introduce the following notation:

$$x_V[\vec{n}] = x(V\vec{n})$$

We call this form lexicographic indexing. An integer vector  $\vec{n}$  is associated with a point in physical space through the matrix  $V$ . The problem is physical dimension, distance and direction are lost. Consider a discrete signal on a rectangular grid, with a rectangular pixel (or voxel) having respective dimensions  $\Delta x_i, 1 \leq i \leq D$ . Then we can take  $V = \text{diag}\{\Delta x_i\}$ . But these dimensions are hidden in the lexicographic indexing form. As another example, consider changing the lattice generator, say  $V' = VE$  where  $E$  is unimodular. Then:

$$x_{V'}[\vec{n}] = x_V[E\vec{n}]$$

Since  $E\mathcal{Z}^D = \mathcal{Z}^D$ , this change of generator matrix corresponds to a permutation, or rearrangement of lexicographic indices. But the underlying signal, and the features and information contained in the signal, has not actually changed. This artificial indexing can mask spatial correlations which often important to consider.

The act of decimation corresponds to keeping only points within a sublattice, say generated by  $U=VM$ . Decimation is typically expressed in lexicographic form as:

$$x_U[\vec{n}] = x_V[M\vec{n}]$$

This operation is called decimation with respect to the integer matrix  $M$ . This suggests a significant change in signal, and certainly the notion of spatial proximity is lost. In a physical sense, though, the only true change in the signal is that some values are discarded (effectively, set to 0), but the physical positions of the retained points do not change at all. The geometric structures of the underlying lattices are lost as well, because the change from say  $\mathcal{L}_V$  to  $\mathcal{L}_U$  as the underlying lattice of support is masked by forcing indexing via points in  $\mathcal{Z}^D$ . Consider the case of rectangular sampling mentioned above. Decimating in the 1-direction by a factor of 2, basically changing  $\Delta x_1$  to  $2\Delta x_1$  but no change in the other directions, is a nonuniform change of the shape

of a pixel (voxel) that is not evident in lexicographic form. Procedures to design filtering structures, as in [14], and fast algorithms for spectral analysis, as in [15], are usually cast in lexicographic form, and thus are often not tied to the actual geometric structure of the underlying signals that need to be processed in a number of applications.

Here we develop a multirate signal processing approach retaining the physical indices, while relating results back to the traditional lexicographic index form. Using lexicographic notation can help determine techniques for data storage and retrieval of the signals of interest, but the exploration and development of signal processing techniques is facilitated by working with the physical coordinates.

### 3. FOURIER TRANSFORMS

There are generally four forms of the Fourier transform, corresponding to signals that are either continuous or discrete, and aperiodic (infinite in extent) or periodic (finite in support and extended periodically). If the starting point is the continuous aperiodic form, the others can be obtained by sampling in the spatial or frequency domains, or both. Here we obtain formulas for all, preserving the physical nature of the spatial or frequency domain coordinates, with connections to the normalized (lexicographic) forms.

#### 3.1. Multidimensional forms of CTFT, DTFT and Fourier Series

We start with the continuous aperiodic form of the Fourier transform, where the spatial and frequency domains are both  $\mathcal{R}^D$ . This is a multidimensional form of the continuous-time Fourier transform (CTFT). For a signal  $g: \mathcal{R}^D \rightarrow \mathcal{C}$  the Fourier transform is a signal  $G: \mathcal{R}^D \rightarrow \mathcal{C}$  where:

$$G(\vec{f}) = \int_{\mathcal{R}^D} g(\vec{r}) e(-\langle \vec{f}, \vec{r} \rangle) d\vec{r}$$

$$g(\vec{r}) = \int_{\mathcal{R}^D} G(\vec{f}) e(\langle \vec{f}, \vec{r} \rangle) d\vec{f}$$

Our approach to obtaining the other forms of the Fourier transform is to represent a discrete signal as a superposition of impulses, and pass it through the continuous form of the Fourier transform.

Note that  $\delta(\vec{r} - \vec{r}_0) \rightarrow e(-\langle \vec{f}, \vec{r}_0 \rangle)$  and  $e(\langle \vec{f}_0, \vec{r} \rangle) \rightarrow \delta(\vec{f} - \vec{f}_0)$ .

Take a signal over a lattice,  $g: \mathcal{L}_V \rightarrow \mathcal{C}$ , with generator matrix  $V$ . The corresponding frequency domain will be continuous, but periodic with respect to the reciprocal lattice, that is, the spectrum is a signal  $G: \mathcal{R}^D / \mathcal{L}_{\vec{v}} \rightarrow \mathcal{C}$ . To see this, start with the impulsive signal:

$$g_0(\vec{r}) = \sum_{\vec{\xi} \in \mathcal{L}_V} g(\vec{\xi}) \delta(\vec{r} - \vec{\xi})$$

Substitution into the Fourier transform yields a spectrum that is a superposition of spectral functions of the form

$$\phi_{\vec{\xi}}(\vec{f}) = e(-\langle \vec{f}, \vec{\xi} \rangle)$$

where  $\vec{\xi} \in \mathcal{L}_V$ . Note  $\vec{\xi} = V\vec{n}$  for some  $\vec{n} \in \mathcal{Z}^D$  and  $\langle \vec{f}, V\vec{n} \rangle = \langle V^T \vec{f}, \vec{n} \rangle$ . Therefore,  $\phi_{\vec{\xi}}(\vec{f} + V^{-T} \vec{m}) = \phi_{\vec{\xi}}(\vec{f})$  for  $\vec{m} \in \mathcal{Z}^D$  and hence  $\phi_{\vec{\xi}}$  is periodic with respect to the reciprocal lattice  $\mathcal{L}_{\vec{v}}$ , and the spectrum is completely determined by its values in one reciprocal unit

cell  $u_{\tilde{V}}$ . We thus obtain the multidimensional form of the so-called discrete-time Fourier transform (DTFT) for signals over a lattice:

$$G(\vec{f}) = \sum_{\vec{r} \in \mathcal{L}_V} g(\vec{r}) e(-\langle \vec{f}, \vec{r} \rangle)$$

$$g(\vec{r}) = |\det V| \int_{u_{\tilde{V}}} G(\vec{f}) e(\langle \vec{f}, \vec{r} \rangle) d\vec{f}$$

Substituting  $\vec{\xi} = V\vec{n}$ ,  $\vec{n} \in \mathcal{Z}^D$ , and  $\vec{f} = \tilde{V}\vec{u}$ ,  $\vec{u} \in [-\frac{1}{2}, \frac{1}{2}]^D$ , yields the more standard, normalized (lexicographic) form of the multidimensional DTFT:

$$G[\vec{u}] = \sum_{\vec{n} \in \mathcal{Z}^D} g[\vec{n}] e(-\langle \vec{u}, \vec{n} \rangle)$$

$$g[\vec{n}] = \int_{\vec{u} \in [-\frac{1}{2}, \frac{1}{2}]^D} G[\vec{u}] e(\langle \vec{u}, \vec{n} \rangle) d\vec{u}$$

A similar approach can be used to obtain the multidimensional form of the Fourier series. We start with a signal whose spectrum has support on the reciprocal lattice of  $V$  in the frequency domain. By representing the spectrum as a sum of impulses located at the reciprocal lattice points, we find the signal in the spatial domain is continuous but periodic with respect to the original lattice. Thus, continuous lattice-periodic signals  $g: \mathcal{R}^D / \mathcal{L}_V \rightarrow \mathcal{C}$  have discrete spectra with support on the reciprocal lattice,  $G: \mathcal{L}_{\tilde{V}} \rightarrow \mathcal{C}$ , with Fourier transform given by:

$$G(\vec{f}) = \frac{1}{|\det V|} \int_{u_V} g(\vec{r}) e(-\langle \vec{f}, \vec{r} \rangle) d\vec{r}$$

$$g(\vec{r}) = \sum_{\vec{f} \in \mathcal{L}_{\tilde{V}}} G(\vec{f}) e(\langle \vec{f}, \vec{r} \rangle)$$

Expressing both the spatial and frequency domains in normalized (lexicographic) form results in the more conventional multidimensional Fourier series formulas:

$$G[\vec{k}] = \int_{\vec{\zeta} \in [-\frac{1}{2}, \frac{1}{2}]^D} g[\vec{\zeta}] e(-\langle \vec{k}, \vec{\zeta} \rangle) d\vec{\zeta}$$

$$g[\vec{\zeta}] = \sum_{\vec{k} \in \mathcal{Z}^D} G[\vec{k}] e(\langle \vec{k}, \vec{\zeta} \rangle)$$

In normalized form, the spatial signal is viewed in  $\mathcal{R}^D / \mathcal{Z}^D$  and the spectrum is a signal over  $\mathcal{Z}^D$ .

### 3.2. MRL-DFT

We now present the multiresolution-lattice discrete Fourier transform (MRL-DFT), which is a multidimensional form of the DFT that retains the geometric structure of space and frequency.

The MRL-DFT applies to signals with support on one lattice, periodic with respect to a sublattice. To justify the formulation, we start with a special case, which is signals over  $\mathcal{Z}^D$  that are periodic

with respect to an integer matrix,  $N$ . This yields the conventional form of the multidimensional DFT, which is in lexicographic form.

In one dimension, the DFT applies to a time-domain signal where the time is integer valued and periodic with period  $N$ . To apply Fourier analysis, the time domain must be a group, and in fact the group is the cyclic group  $\mathcal{Z}_N = \{0, 1, \dots, N-1\}$  (i.e., modulo- $N$  addition), which is isomorphic to the quotient group  $\mathcal{Z}/N\mathcal{Z}$ , which we recognize as a simple lattice quotient group. The frequency domain is also  $\mathcal{Z}_N$ . The parameter  $N$  is called the radix of the DFT. The simplest multidimensional form assumes a signal over an  $N_1 \times N_2 \times \dots \times N_D$  grid, and the DFT can be applied separately in each dimension, resulting in the formulas:

$$G[\vec{k}] = \sum_{0 \leq n_i \leq N_i - 1} g[\vec{n}] e\left(-\sum \frac{n_i k_i}{N_i}\right)$$

$$g[\vec{n}] = \frac{1}{N_1 N_2 \dots N_D} \sum_{0 \leq k_i \leq N_i - 1} G[\vec{k}] e\left(\sum \frac{n_i k_i}{N_i}\right)$$

Define the matrix  $N = \text{diag}\{N_1, N_2, \dots, N_D\}$ . The scaling factor  $N_1 N_2 \dots N_D = |\det N|$ , and also:

$$\langle \vec{k}, N^{-1}\vec{n} \rangle = k^T N^{-1} n = \sum \frac{n_i k_i}{N_i}$$

Let us define  $\mathcal{J}_N = \{\vec{n}: 0 \leq n_i \leq N_i - 1\}$ . This set is a minimal set of coset representatives of  $\mathcal{L}_N$  in  $\mathcal{Z}^D$ . The sums are both taken over this set. In generalizing this to non-diagonal integer matrices, care must be taken because, in this case,  $N = N^T$ . In order to arrive at the correct form for a general integer matrix, we need a mathematical result. In what follows, for an integer matrix  $M$ , let  $\mathcal{J}_M$  be an index set for  $M$  in the integer lattice, that is,  $\mathcal{J}_M = \mathcal{U}_M \cap \mathcal{Z}^D$  for any unit cell  $\mathcal{U}_M$ . For  $\vec{n}, \vec{k} \in \mathcal{Z}^D$ , and  $D \times D$  invertible integer matrix  $N$ , define:

$$\phi_N(\vec{n}, \vec{k}) = e(\langle \vec{k}, N^{-1}\vec{n} \rangle) = e(\langle \tilde{N}\vec{k}, \vec{n} \rangle)$$

**Lemma 4:** For  $\vec{k} \in \mathcal{Z}^D$ ,  $\phi_N(\cdot, \vec{k})$  is  $\mathcal{L}_N$ -periodic, and for  $\vec{n} \in \mathcal{Z}^D$ ,  $\phi_N(\vec{n}, \cdot)$  is  $\mathcal{L}_{N^T}$ -periodic. Moreover, the following orthogonality conditions hold[1]:

$$\frac{1}{|\det N|} \sum_{\vec{n} \in \mathcal{J}_N} \phi_N(\vec{n}, \vec{k}) = \begin{cases} 1, & \vec{k} \in \mathcal{L}_{N^T} \\ 0, & \vec{k} \notin \mathcal{L}_{N^T} \end{cases}$$

$$\frac{1}{|\det N|} \sum_{\vec{k} \in \mathcal{J}_{N^T}} \phi_N(\vec{n}, \vec{k}) = \begin{cases} 1, & \vec{n} \in \mathcal{L}_N \\ 0, & \vec{n} \notin \mathcal{L}_N \end{cases}$$

Using this Lemma, we can obtain the DFT formulation simply as a standard (finite-dimensional) orthogonal expansion:

**Definition 2:** The discrete Fourier transform (DFT) with respect to an integer matrix  $N$  maps spatial domain signals over  $\mathcal{Z}^D/N\mathcal{Z}^D$  to frequency domain signals over  $\mathcal{Z}^D/N^T\mathcal{Z}^D$  as follows:

$$G[\vec{k}] = \sum_{\vec{n} \in \mathcal{J}_N} g[\vec{n}] e(-\langle \vec{k}, N^{-1}\vec{n} \rangle)$$

$$g[\vec{n}] = \frac{1}{|\det N|} \sum_{\vec{k} \in \mathcal{J}_{N^T}} G[\vec{k}] e(\langle \vec{k}, N^{-1}\vec{n} \rangle)$$

Even if we started with a signal truly on an integer lattice, this formulation gives a normalized (lexicographic) frequency domain. We would expect the spectrum to lie at points on a lattice generated by  $\tilde{N}$ . But for a general integer matrix  $N$ ,  $\tilde{N}$  is not an integer matrix, and instead this form of the DFT indexes the frequency domain at integer vectors in the set  $\mathcal{J}_{N^T}$ .

We propose the MRL-DFT as a formulation that preserves the geometric features of the spatial and frequency domains. The proper form of the equations is provided by the following theorem.

**Theorem 5:** Let  $\mathcal{L}_U \subset \mathcal{L}_V$  be lattices in  $\mathcal{R}^D$ . Then the Fourier transform of a spatial signal over  $\mathcal{L}_V/\mathcal{L}_U$  is a frequency domain signal over  $\mathcal{L}_{\tilde{U}}/\mathcal{L}_{\tilde{V}}$ , and the transform and inverse transform formulas are given by:

$$G(\vec{f}) = \sum_{\vec{r} \in \mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)} g(\vec{r}) e(-\langle \vec{f}, \vec{r} \rangle)$$

$$g(\vec{r}) = \frac{1}{|\mathcal{J}|} \sum_{\vec{f} \in \mathcal{J}(\mathcal{L}_{\tilde{V}}, \mathcal{L}_{\tilde{U}})} G(\vec{f}) e(\langle \vec{f}, \vec{r} \rangle)$$

We call this the **multiresolution lattice discrete Fourier transform (MRL-DFT)**.

**Proof:** Since  $g(\vec{r})$  is  $\mathcal{L}_U$ -periodic, the previous discussion yields that its Fourier transform  $G(\vec{f})$  has support in  $\mathcal{L}_{\tilde{U}}$ . Then the Fourier series formulation yields:

$$G(\vec{\zeta}) = \frac{1}{|\det U|} \int_{\mathcal{U}_U} g_0(\vec{r}) e(-\langle \vec{\zeta}, \vec{r} \rangle) d\vec{r}$$

where  $g_0$  is a signal over continuous-space, but is actually concentrated only at points in the lattice  $\mathcal{L}_V$ . In fact,  $g_0$  is an impulsive function, with impulses located at the lattice points. The integral becomes a sum at points  $\mathcal{U}_U \cap \mathcal{L}_V$ , which is  $\mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)$ , and we obtain the MRL-DFT in the theorem, up to the scaling factor  $|\det U|$ . Now examine the reverse situation: start with  $g(\vec{r})$  having support on the lattice  $\mathcal{L}_V$  with a spectrum  $G(\vec{f})$  that is  $\mathcal{L}_{\tilde{V}}$ -periodic. The generalized inverse DTFT yields an integral over the continuous frequency domain region  $\mathcal{U}_{\tilde{V}}$ , but the integrand is impulsive since it lies at discrete lattice points only, and the integral becomes the inverse MRL-DFT formula, except that the scaling factor  $1/|\mathcal{J}|$  is replaced with  $|\det V|$ . To account for the scaling factors, we recognize that in Fourier transform and inverse transform formulas, what is significant is not the individual scaling factors, but their product. We can rescale the defined spectrum, for example, and since  $|\det \mathcal{J}| = \frac{|\det U|}{|\det V|}$  the proposed MRL-DFT and inverse MRL-DFT formulas are confirmed.

We note that if both the spatial and frequency domains in the MRL-DFT are expressed in lexicographic form, with  $U=VN$  for an integer matrix  $N$ , the MRL-DFT reduces to the standard DFT described previously.

### 3.3. Properties of the MRL-DFT

As with any Fourier transform, we expect certain properties to hold, such as convolution in one domain corresponds to multiplication in the other domain. All properties listed here are special cases of properties of abstract Fourier transforms, and can be derived in that context, for example through group theory. Therefore we do not provide complete proofs for all properties listed here, but highlight some examples to illustrate the application of lattice theory. Here,  $\mathcal{F}\{x\}$  denotes the MRL-DFT of  $x$ , and  $\mathcal{F}^{-1}\{X\}$  denotes the inverse MRL-DFT of  $X$ .

**Definition 3:** Let  $h, g$  be signals over  $\mathcal{L}_V/\mathcal{L}_U$ . Then the convolution  $y = h * g$  is a signal over  $\mathcal{L}_V/\mathcal{L}_U$  and is given by:

$$y(\vec{r}) = \sum_{\vec{s} \in \mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)} h(\vec{s})g(\vec{r} - \vec{s})$$

and the correlation  $\phi_{hg}$  is a signal over  $\mathcal{L}_V/\mathcal{L}_U$  given by:

$$\phi_{hg}(\vec{r}) = \sum_{\vec{s} \in \mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)} h(\vec{r} + \vec{s})g^*(\vec{s})$$

In each case, the terms in the sum are  $\mathcal{L}_U$ -periodic. Convolution and correlation in the frequency domain are defined similarly, i.e., without any scaling factors in the sums.

**Lemma 6:** If  $x(\vec{r}) = g^*(-\vec{r})$  then  $\phi_{hg} = h * x$ .

**Lemma 7:**  $\mathcal{F}\{x^*(\vec{r})\} = X^*(-\vec{f})$  and  $\mathcal{F}\{x^*(-\vec{r})\} = X^*(\vec{f})$ .

**Theorem 8:** Convolution in the spatial domain corresponds to multiplication in the frequency domain, and conversely, specifically:

$$\mathcal{F}\{h * g\} = H \cdot G$$

$$\mathcal{F}\{h \cdot g\} = |J|H * G$$

**Corollary 9:** The MRL-DFT of the correlation function  $\phi_{hg}$  is  $H \cdot G^*$ .

Now matching  $\phi_{hg}(\vec{0})$  to the inverse MRL-DFT of  $H \cdot G^*$  evaluated at  $\vec{0}$  results in the following:

**Corollary 10:** Parseval's theorem for the MRL-DFT:

$$\sum_{\vec{r} \in \mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)} h(\vec{r})g^*(\vec{r}) = \frac{1}{|J|} \sum_{\vec{f} \in \mathcal{J}(\mathcal{L}_V, \mathcal{L}_U)} H(\vec{f})G^*(\vec{f})$$

We omit the algebraic details of these results, except to point out two fundamental concepts that underlie the proof of the convolution theorem. The first is the factorization of the exponential function:

$$e(\langle \vec{f}, \vec{r} \rangle) = e(\langle \vec{f}, \vec{r} - \vec{s} \rangle) \cdot e(\langle \vec{f}, \vec{s} \rangle)$$

The second deals with summing over the index set. The difficulty is that if  $\vec{r}, \vec{s} \in \mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)$ , in general,  $\vec{r} - \vec{s}$  is not in the index set. However, the signals and the exponential functions appearing

in the sums are  $\mathcal{L}_U$ -periodic. If  $\vec{\xi} \in \mathcal{L}_V$ , let  $(\vec{\xi})_U \in \mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)$  be the unique point in the index set such that  $\vec{\xi} - (\vec{\xi})_U \in \mathcal{L}_U$ . We can think of  $(\vec{\xi})_U$  as the value of  $\vec{\xi}$  modulo the sublattice  $\mathcal{L}_U$ . Then for each  $\vec{s}$ , the mapping  $\vec{r} \rightarrow (\vec{r} - \vec{s})_U$  is a one-to-one mapping of  $\mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)$  onto itself.

#### 4. MULTIREOLUTION LATTICES

This section lays the groundwork for signal processing on multiresolution lattices. Specifically, we consider chains of lattices of varying densities, say  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$ . Let  $W=VN$ ,  $U=WQ$ , so that  $U=VM$  with  $M=NQ$ . To avoid the trivial case, where  $\mathcal{L}_W$  equals either  $\mathcal{L}_U$  or  $\mathcal{L}_V$ , we require that neither  $N$  nor  $Q$  be unimodular. Therefore, we can generate intervening sublattices if we can obtain non-trivial factorization of integer matrices. Such factorizations, in fact, will allow us to describe the structure of  $\mathcal{L}_V/\mathcal{L}_U$ .

Another issue we must address is that, in order to actually perform computations such as convolution, correlation or MRL-DFT, we need to be able to obtain valid index sets for specific lattices. It is not obvious how to find a specific set of points that is a valid choice for  $\mathcal{J}(\mathcal{L}_U, \mathcal{L}_V)$ . The next section describes an important mathematical result we will rely on.

##### 4.1. Smith form for integer matrices

**Proposition 11:** Any  $D \times D$  nonsingular integer matrix  $M$  can be factored as  $M = EKF$  with  $E, F$  unimodular, and  $K = \text{diag}\{k_1, k_2, \dots, k_D\}$  with each  $k_i$  a positive integer. This form is not unique, but the  $k_i$ 's are unique up to a permutation. This is called the Smith form of an integer matrix. [16]

Note that  $|\det M| = \prod k_i$ . This allows us to identify all possible combinations of  $k_i$ 's. For example, if  $M$  is  $3 \times 3$  and  $|\det M| = 12$ , the Smith form of  $M$  will result in one of the following (up to a permutation): (12,1,1), (6,2,1), (4,3,1), or (3,2,2).

We now seek a factorization  $M=NQ$  where neither  $N$  nor  $Q$  is unimodular. This hinges on the factorization of  $K$ . Let  $K=AB$  where  $A, B$  are diagonal matrices with positive integers on the diagonals. At the  $i^{\text{th}}$  position, we must have  $k_i = a_i b_i$ . To avoid a trivial factorization, we require at least one  $a_i > 1$  and at least one  $b_i > 1$ . For example if  $K = \text{diag}\{4,3,1\}$  we could take  $A = \text{diag}\{4,1,1\}, B = \text{diag}\{1,3,1\}$  or  $A = \text{diag}\{2,3,1\}, B = \text{diag}\{2,1,1\}$ , or  $A = \text{diag}\{2,1,1\}, B = \text{diag}\{2,3,1\}$ , or  $A = \text{diag}\{1,3,1\}, B = \text{diag}\{4,1,1\}$ . There are no other possibilities! Then with  $N=EA$  and  $Q=BF$ , we have a nontrivial factorization  $M=NQ$ .

Given the Smith form  $M=EKF$ , and  $U=VM$ , we can write  $U'=V'K$  where  $U' = UF^{-1}, V' = VE$ . Since  $F$  is unimodular, so is  $F^{-1}$ . Therefore,  $U', V'$  generate the same lattices as  $U, V$ , and we have found a representation of the lattices such that the generators are related by a diagonal integer matrix. But finding an index set then is straightforward. For example, if  $K = \text{diag}\{4,3,1\}$ , then the entries in the index set are vectors of the form  $V'\vec{n} = VE\vec{n}$  where  $\vec{n} = (n_1, n_2, 0)$  with  $0 \leq n_1 \leq 3, 0 \leq n_2 \leq 2$ . This yields the precise mathematical structure of the quotient group.

**Theorem 12:** If  $U=VM$  and the Smith form of  $M$  is  $EKF$  with  $K = \text{diag}\{k_1, k_2, \dots, k_D\}$ , then the quotient group  $\mathcal{L}_V/\mathcal{L}_U$  is isomorphic to the direct product of cyclic groups as follows:

$$\mathcal{Z}_{k_1} \times \mathcal{Z}_{k_2} \times \dots \times \mathcal{Z}_{k_D}$$

It is well known that every finite abelian group is the direct product of cyclic groups. The significance of this theorem is that we can use the Smith form to find the particular decomposition for a lattice quotient group. The Smith form allows us to generate proper index

sets, and to obtain integer matrix factorizations to identify and generate all intervening lattices between  $\mathcal{L}_U$  and  $\mathcal{L}_V$ .

#### 4.2. The MRL-DFT viewed on lattice chains

In this section, we consider  $U=VM$ , and lattices that lie between those generated by  $U$  and  $V$ , i.e., we study lattice chains such as  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$ . Let  $\mathcal{S}(U, V)$  denote the space of complex valued signals with support on  $\mathcal{L}_V$  that are periodic with respect to  $\mathcal{L}_U$ . The MRL-DFT for such signals will be in  $\mathcal{S}(\vec{V}, \vec{U})$ . We will also use the notation  $\mathcal{S}(V)$  to denote signals with support on  $\mathcal{L}_V$  that are not periodic. If we start with a signal with finite number of nonzero values, but the points of support lie completely within some unit cell  $\mathcal{U}_U$  of  $U$ , then we will consider it extended periodically with respect to  $\mathcal{L}_U$  to apply the techniques developed here. We apply the same principle similarly in the frequency domain.

The first result is a common property of Fourier transforms: translation in either the spatial or frequency domain corresponds to a linear phase shift in the other domain.

**Lemma 13:** Let  $x \in \mathcal{S}(U, V)$ ,  $\vec{r}_0 \in \mathcal{L}_V$ ,  $\vec{f}_0 \in \mathcal{L}_{\vec{U}}$ . Also let  $X$  be the  $\mathcal{L}_V/\mathcal{L}_U$  MRL-DFT of  $x$ . Then the MRL-DFT of  $x(\vec{r} + \vec{r}_0)$  is  $e(\langle \vec{f}, \vec{r}_0 \rangle)X(\vec{f})$  and the MRL-DFT of  $e(\langle \vec{f}_0, \vec{r} \rangle)x(\vec{r})$  is  $X(\vec{f} + \vec{f}_0)$ .

**Proof:** The results follow from a straightforward application of factorizations of the exponential function such as:

$$e(\langle \vec{f}, \vec{r} + \vec{r}_0 \rangle) = e(\langle \vec{f}, \vec{r} \rangle) \cdot e(\langle \vec{f}, \vec{r}_0 \rangle)$$

Take  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$ . Consider a signal  $x \in \mathcal{S}(U, W) \subset \mathcal{S}(U, V)$ , specifically  $x$  is 0 for points in  $\mathcal{L}_V$  that are not in  $\mathcal{L}_W$ . We can compute the MRL-DFT with respect to either  $\mathcal{L}_W$  or  $\mathcal{L}_V$  as the lattice of support. However, since the signal itself is essentially the same, we hope the resulting spectra will match. That is in fact what happens. This confirms our formulation truly preserves the physical nature of the signals under study. To prove it, we need the following.

**Lemma 14:** If  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$ , and we have valid index sets  $\mathcal{J}(U, W), \mathcal{J}(W, V)$ , then  $\mathcal{J}(U, V) = \mathcal{J}(U, W) + \mathcal{J}(W, V)$  is a valid index set. Moreover, if  $\vec{0} \in \mathcal{J}(W, V)$  then the result of this construction is that  $\mathcal{J}(U, W) \subset \mathcal{J}(U, V)$ .

**Theorem 15:** Let  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$ , and  $x \in \mathcal{S}(U, W) \subset \mathcal{S}(U, V)$ . Let  $X^{(W)}$  denote its  $\mathcal{L}_W/\mathcal{L}_U$  MRL-DFT, and  $X^{(V)}$  its  $\mathcal{L}_V/\mathcal{L}_U$  MRL-DFT. Both have support on  $\mathcal{L}_{\vec{U}}$ , and for all  $\vec{f} \in \mathcal{L}_{\vec{U}}$ :

$$X^{(V)}(\vec{f}) = X^{(W)}(\vec{f})$$

This relation has the same form in lexicographic notation, that is

$$X^{(V)}[\vec{k}] = X^{(W)}[\vec{k}] \quad \forall \vec{k} \in \mathcal{Z}^D$$

**Proof:**  $X^{(V)}(\vec{f})$  is computed as a sum over  $\mathcal{J}(U, V)$  for some index set. By Lemma 14, we can chose  $\mathcal{J}(U, W), \mathcal{J}(W, V)$ , with  $\vec{0} \in \mathcal{J}(W, V)$ , so that  $\mathcal{J}(U, V) = \mathcal{J}(U, W) + \mathcal{J}(W, V)$ . With this construction, the only point in  $\mathcal{J}(W, V)$  that is also in  $\mathcal{L}_W$  is  $\vec{0}$ , and every point in  $\mathcal{J}(U, W)$  is in  $\mathcal{L}_W$ . Therefore,  $\mathcal{J}(U, V) \cap \mathcal{L}_W = \mathcal{J}(U, W)$ . Therefore, the signal  $x$  is zero at all other points in  $\mathcal{J}(U, V)$ , so the sum over  $\mathcal{J}(U, V)$  reduces to a sum over  $\mathcal{J}(U, W)$  and the MRL-DFTs match.

Since both spectra have support on the same lattice, their lexicographic forms are both related through the same matrix,  $\tilde{U}$ , so the lexicographic forms match as well.

The next result is that restriction to a sublattice in space or frequency corresponds to periodicity in the other domain.

**Theorem 16:** Let  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$  and  $x \in \mathcal{S}(U, V)$ . Then  $x$  has support on the sublattice  $\mathcal{L}_W$  iff its  $\mathcal{L}_V/\mathcal{L}_U$  MRL-DFT is  $\mathcal{L}_{\tilde{W}}$ -periodic, and  $x$  is  $\mathcal{L}_W$ -periodic iff its  $\mathcal{L}_V/\mathcal{L}_U$  MRL-DFT has support on  $\mathcal{L}_{\tilde{W}}$ .

**Proof:** We prove the first statement. The second follows from the symmetry between the MRL-DFT and its inverse. First assume  $x$  is zero outside  $\mathcal{L}_W$ . Then by Theorem 15,  $X^{(V)}(\vec{f}) = X^{(W)}(\vec{f})$ , but  $X^{(W)}$  must be  $\mathcal{L}_{\tilde{W}}$ -periodic. So this part of the theorem is trivial. Now assume, conversely, that  $X^{(V)}(\vec{f})$  is  $\mathcal{L}_{\tilde{W}}$ -periodic. That means  $X^{(V)} \in \mathcal{S}(\tilde{W}, \tilde{U})$ . Perform an inverse  $\mathcal{L}_W/\mathcal{L}_U$  MRL-DFT to obtain a spatial domain signal  $\tilde{x} \in \mathcal{S}(U, W)$ . Initially,  $\tilde{x}$  is only defined on points in the lattice  $\mathcal{L}_W$ . Now let us extend its domain of definition to  $\mathcal{L}_V$  with the rule that  $\tilde{x}(\vec{r}) = 0$  for  $\vec{r} \notin \mathcal{L}_W$ . Then  $\tilde{X}^{(V)} = \tilde{X}^{(W)}$ , and  $X^{(V)} = \tilde{X}^{(V)}$ . Therefore,  $X^{(V)} = \tilde{X}^{(V)}$ . Since the MRL-DFT is invertible,  $x = \tilde{x}$ .

We can now generalize the prior theorem to describe signals who support is restricted to a coset of a sublattice in the spatial domain, or a coset of a reciprocal sublattice in the frequency domain.

**Corollary 17:** Let  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$ , and  $x \in \mathcal{S}(U, V)$ . Then  $x$  has support only on a coset  $\mathcal{L}_W + \vec{r}_0$  for some  $\vec{r}_0 \in \mathcal{L}_V$  iff its MRL-DFT satisfies:

$$X(\vec{f} + \vec{\zeta}) = e(-\langle \vec{\zeta}, \vec{r}_0 \rangle) X(\vec{f})$$

Similarly, the MRL-DFT has support only on a coset  $\mathcal{L}_{\tilde{W}} + \vec{f}_0$  for some  $\vec{f}_0 \in \mathcal{L}_{\tilde{U}}$  iff it satisfies:

$$x(\vec{r} + \vec{\xi}) = e(\langle \vec{f}_0, \vec{\xi} \rangle) x(\vec{r})$$

**Proof:** For the first part, the frequency domain property can be shown to be equivalent to the condition that  $Y(\vec{f}) = e(\langle \vec{f}, \vec{r}_0 \rangle) X(\vec{f})$  is  $\mathcal{L}_{\tilde{W}}$ -periodic, and the previous theorem can be applied. Similarly for the second statement.

### 4.3. Regridding: upsampling, downsampling and general lattice conversion

Here we discuss changing the sampling structure in the spatial or frequency domain. In classical sampling theory, decimation in time causes aliasing in frequency, and upsampling in time causes imaging distortion in frequency. These results can be mitigated by filtering, which can be implemented via convolution in time or windowing in frequency. Similar results occur if the situations are reversed, which does arise in applications that involve spectral analysis. For example, in MRI, spectral measurements are regridded to convenient point sets via interpolation.

Results are given in the physical domains, in terms of MRL-DFT, and then are reduced to the lexicographic forms. We start with a lattice chain  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$  and assume a fixed choice of index sets, each of which are assumed to contain  $\vec{0}$  where:

$$\mathcal{J}(U, V) = \mathcal{J}(U, W) + \mathcal{J}(W, V)$$

By default, we use generator matrices  $U, W, V$  where  $U=WN$ ,  $W=VQ$ , and hence  $U=VM$  with  $M=NQ$ . Also, we start with a signal  $x \in \mathcal{S}(U, V)$  (periodic case) or  $x \in \mathcal{S}(V)$  (aperiodic case).

**Definition 4:** If  $x$  has support on  $\mathcal{L}_V$ , decimation (downsampling) of  $x$  to  $\mathcal{L}_W$ , denoted  $y = (\mathcal{L}_V \searrow \mathcal{L}_W)x$ , is the restriction of  $x$  to the lattice  $\mathcal{L}_W$ . That is  $y = x$  at points in  $\mathcal{L}_W$ , but we set  $y = 0$  at points in  $\mathcal{L}_V$  not in  $\mathcal{L}_W$ .

**Definition 5:** If  $x$  has support on  $\mathcal{L}_W$ , upsampling of  $x$  to  $\mathcal{L}_V$ , denoted  $y = (\mathcal{L}_W \nearrow \mathcal{L}_V)x$ , expands the domain of  $x$  to  $\mathcal{L}_V$  by specifying  $y = x$  at points in  $\mathcal{L}_W$  and  $y = 0$  at points in  $\mathcal{L}_V$  not in  $\mathcal{L}_W$ .

Note that decimation does significantly change a signal, in that it discards a set of values. Upsampling does not truly change a signal. It is more a technical change that allows us to consider a denser lattice of support, but the additional values are all 0. Based on the previous discussion, this does not actually change the MRL-DFT, and yet we expect an imaging phenomenon. In fact, a signal in  $\mathcal{S}(U, V)$  normally has a spectrum that is completely specified by its support in a reciprocal unit cell  $\mathcal{U}_{\vec{V}}$ . However, because its MRL-DFT matches that for a signal in  $\mathcal{S}(U, W)$ , this spectrum appears in the smaller region  $\mathcal{U}_{\vec{W}}$ , periodically extended via the dense reciprocal lattice  $\mathcal{L}_{\vec{W}}$ . With  $W=VQ$ , it turns out there are  $|\det Q|$  copies of the spectrum that appear in  $\mathcal{U}_{\vec{V}}$ , and these copies are seen because changing to a denser lattice in the spatial domain corresponds to a wider range of support in the frequency domain.

We expect the decimation process to introduce aliasing, which actually changes the spectral values. To obtain this result, note that decimation can be achieved by multiplying the original signal by a 1-0 function. Specifically, for a set  $\mathcal{A}$ , let  $\chi_{\mathcal{A}}$  be a signal that is 1 in  $\mathcal{A}$  and 0 outside of  $\mathcal{A}$ ; this is called the indicator function. Then direct computation of the  $\mathcal{L}_V/\mathcal{L}_U$  MRL-DFT yields that the spectrum of  $\chi_{\mathcal{L}_W}$  is  $|\det N|\chi_{\mathcal{L}_{\vec{W}}}$  where  $U=WN$ . Decimation corresponds to multiplying in the spatial domain by the indicator function of the sparser lattice, and thus the frequency domain equivalent is governed by a convolution with an indicator function. This results in the following formula, which is the aliasing phenomenon:

**Theorem 18:** If  $y = (\mathcal{L}_V \searrow \mathcal{L}_W)x$ , then the MRL-DFT of  $y$  is related to the MRL-DFT of  $x$  via:

$$Y(\vec{f}) = \frac{1}{|\mathcal{J}(W, V)|} \sum_{\vec{\zeta} \in \mathcal{J}(\vec{V}, \vec{W})} X(\vec{f} + \vec{\zeta})$$

The lexicographic definitions of decimation and upsampling, and the corresponding frequency domain formulas, are obtained by recognizing the matrices that relate the various lattices. For example,  $U=VM$  and  $\vec{V} = \vec{U}M^T$ . With  $W=VQ$  and  $y = (\mathcal{L}_V \searrow \mathcal{L}_W)x$ :

$$y_W[\vec{n}] = y(W\vec{n}) = y(VQ\vec{n}) = x(VQ\vec{n}) = x_V[Q\vec{n}]$$

Here we have been careful to define the physical matrix associated with the lexicographic notation. The more traditional form is to omit this, and we obtain the standard formula[3] for decimation of signals on  $\mathcal{Z}^D$  with respect to an integer matrix  $Q$ , denoted as  $y = (\downarrow Q)x$ , via

$$y[\vec{n}] = x[Q\vec{n}]$$

Upsampling to a denser lattice, as defined here, becomes, in lexicographic form, equivalent to upsampling a signal on  $\mathcal{Z}^D$  with respect to an integer matrix  $Q$ , denoted  $y = (\uparrow Q)x$ :

$$y[\vec{n}] = \begin{cases} x[Q^{-1}\vec{n}] & \text{if } \vec{n} \in Q\mathcal{Z}^D \\ 0 & \text{if } \vec{n} \notin Q\mathcal{Z}^D \end{cases}$$

The following gives the lexicographic forms for aliasing and imaging in the frequency domain. Note that  $\tilde{W} = \tilde{U}N^T$ ,  $\tilde{V} = \tilde{W}Q^T$ .

**Corollary 19:** Let  $X[\vec{k}] = X(\tilde{U}\vec{k})$ ,  $Y[\vec{k}] = Y(\tilde{U}\vec{k})$ . If  $y = (\mathcal{L}_V \searrow \mathcal{L}_W)x$  then:

$$Y[\vec{k}] = \frac{1}{|\mathcal{J}(W, V)|} \sum_{\vec{m} \in \mathcal{J}(Q^T)} X[\vec{k} + N^T \vec{m}]$$

and if  $y = (\mathcal{L}_W \nearrow \mathcal{L}_V)x$  then:

$$Y[\vec{k}] = X[N^T \vec{k}]$$

In order to avoid aliasing when decimating, we can first apply a filter  $H$  to bandlimit the signal  $x$  to a reciprocal unit cell  $\mathcal{U}_{\tilde{W}}$ . This gives rise to a decimation filter operation:

$$y = (\mathcal{L}_V \searrow \mathcal{L}_W)Hx$$

where the ideal filter has a spectrum  $H(\vec{f}) = \chi_{\mathcal{L}_{\tilde{W}}}(\vec{f})$ . This can be implemented by windowing (multiplying pointwise) in the frequency domain, or by convolution with the inverse transform  $h$ , which can be computed directly via the inverse MRL-DFT. The same ideal anti-aliasing filter also removes imaging distortion, i.e., retains only one copy of the spectrum when upsampling. Thus, the ideal interpolation filter is given by:

$$y = H(\mathcal{L}_W \nearrow \mathcal{L}_V)x$$

We can also specify ideal filtering operations for general sampling lattice conversion. Suppose  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$  and  $\mathcal{L}_U \subset \mathcal{L}_P \subset \mathcal{L}_V$  but the intervening lattices  $\mathcal{L}_W, \mathcal{L}_P$  are not, in general a sublattice of the other. We want to convert a signal with support in  $\mathcal{L}_W$  to one with support in  $\mathcal{L}_P$ , with ideal filtering to suppress both aliasing and imaging. In general, this may require some information loss (i.e., part of the original signal band may have to be suppressed). The ideal sampling rate conversion process is:

$$y = (\mathcal{L}_V \searrow \mathcal{L}_P)H(\mathcal{L}_W \nearrow \mathcal{L}_V)x$$

where the spectrum of  $H$  is the indicator function for  $\mathcal{U}_{\tilde{W}} \cap \mathcal{U}_{\tilde{P}}$  (actually, points in  $\mathcal{L}_{\tilde{V}}$  that lie in this region).

Depending on the application, multirate filtering is not always designed to suppress aliasing or imaging, and in each of these configurations, the selected  $H$  may be chosen based on different criteria. Additionally, it is recognized that the proposed schemes are not computationally efficient, though they are presented in forms that are directly related to the target application. In an interpolation filter, for example, a signal with many 0 values is fed to the filter, and thus many multiply operations are multiply by 0, which is wasteful. Similarly, with a decimation filter, the filter is applied first, and then many computed values are discarded. The polyphase concept is useful in designing computationally efficient interpolation and decimation filtering algorithms, as well as in developing FFT algorithms. We will define polyphase components in our context below, both in the physical domain and in lexicographic form, and present one fundamental result that allows for efficient polyphase decomposition of filtering operations.

Let  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$ , and  $x \in \mathcal{S}(U, V)$ . Then each polyphase component of  $x$  with respect to  $\mathcal{L}_W$  is comprised of the values of  $x$  on a single coset of  $\mathcal{L}_W$  inside  $\mathcal{L}_V$ . With  $W=VQ$ , note that there are

$|detQ|$  distinct polyphase components. To be specific, each polyphase component is a signal in  $\mathcal{S}(U, W)$ , indexed by points in  $\mathcal{L}_V$ . If  $\vec{k} \in \mathcal{L}_V$ , then the corresponding polyphase component of  $x$  is defined as:

$$x_{\vec{k}}(\vec{r}) = x(\vec{r} + \vec{k}) \text{ for } \vec{r} \in \mathcal{L}_W$$

Selecting the vectors  $\vec{k}$  from any valid index set  $\text{set}\mathcal{J}(W, V)$  generates all the distinct polyphase components.

**Theorem 20:** If  $x, h$  have support in  $\mathcal{L}_V$  (periodic or aperiodic), and  $\mathcal{L}_W \subset \mathcal{L}_V$ , then  $y = h * x$  (i.e., either circular or linear convolution) can be computed in polyphase form as:

$$y_{\vec{m}} = \sum_{\vec{k} \in \mathcal{J}(W, V)} h_{\vec{k}} * x_{\vec{m} - \vec{k}}$$

**Corollary 21:** A decimation filter can be implemented efficiently via polyphase form as:

$$(\mathcal{L}_V \searrow \mathcal{L}_W)(h * x) = \sum_{\vec{k} \in \mathcal{J}(W, V)} h_{\vec{k}} * x_{-\vec{k}}$$

and an interpolation filter can be implemented efficiently via polyphase realization as follows:

$$\left( h * ((\mathcal{L}_W \nearrow \mathcal{L}_V)x) \right)_{\vec{k}} = h_{\vec{k}} * x$$

We note some final remarks on polyphase representation. The roles of space and frequency can be reversed, and we can define polyphase components by considering cosets of sublattices in the frequency domain. This is useful, for example, for regridding in the frequency domain from limited spectral measurements, or for implementing windowing operations (pointwise multiplication) in the spatial domain. Combining polyphase decomposition with prior results regarding the Fourier transform of signals with support on a coset of a lattice (Corollary 17) also leads to the development of FFT algorithms (i.e., generalizations of the standard decimation-in-time and decimation-in-frequency formulations).

## 5. RESULTS

Here we discuss an example to illustrate some important concepts that were developed in this paper. Start with the matrix  $V$  given by:

$$V = \begin{bmatrix} \sqrt{3} & \sqrt{3} \\ -1 & 1 \end{bmatrix}$$

Then  $\mathcal{L}_V$  is generated by two vectors of equal length at an angle of  $120^\circ$ , and is called a hexagonal lattice [5]. We want to form multiresolution lattice chains  $\mathcal{L}_U \subset \mathcal{L}_W \subset \mathcal{L}_V$  and  $\mathcal{L}_U \subset \mathcal{L}_P \subset \mathcal{L}_V$  such that neither  $\mathcal{L}_W$  nor  $\mathcal{L}_P$  is a sublattice of the other. Then our goal is to describe the process of regridding, that is, for converting a signal sampled on  $\mathcal{L}_W$  to one with support on  $\mathcal{L}_P$ . In particular, we want to determine the ideal interpolation filter.

Let us take  $U = VM$  where  $M = \text{diag}\{8, 8\}$ . At first, choosing diagonal  $M$  may seem simplistic, but recall that the Smith form of an integer matrix led to the result that we can always choose the generators of a lattice and sublattice so the matrices are related by a diagonal integer matrix. Now we can generate intervening lattices by (nontrivial) factorizations of  $M$ . Here, we will take  $W = VN$

where  $N = \text{diag}\{2,1\}$ , and  $P=VQ$  where  $Q = \text{diag}\{1,2\}$ . We can confirm that neither  $\mathcal{L}_W$  nor  $\mathcal{L}_P$  is a sublattice of the other by checking that neither  $W^{-1}P$  nor  $P^{-1}W$  is an integer matrix.

Figure 4 shows points in  $\mathcal{L}_W$ , marked 'x', points in  $\mathcal{L}_P$ , marked '+', and points lying in both lattices marked '\*'. Note that these common points themselves form a lattice  $\mathcal{L}'$  which is a sublattice of  $\mathcal{L}_W$  and  $\mathcal{L}_P$ , and which contains  $\mathcal{L}_U$  as a sublattice. The process for regridding has the form  $(\mathcal{L}_V \searrow \mathcal{L}_P) H(\mathcal{L}_W \nearrow \mathcal{L}_V)$ , which means first upsample to a dense lattice, apply an interpolation filter  $H$  which in this case combats aliasing and imaging distortion, and downsample to the target lattice. In the frequency domain  $H$  has support on the reciprocal lattice  $\mathcal{L}_{\bar{U}}$  and is in fact 1 at points inside  $\mathcal{U}_{\bar{W}} \cap \mathcal{U}_{\bar{P}}$ , and 0 outside, for some choice of unit cells. Here, we take the unit cell for a lattice generated by  $A$  to be  $A[-\frac{1}{2}, \frac{1}{2}]^D$ . We can check if a point  $\vec{\zeta}$  is inside such a set by computing  $A^{-1}\vec{\zeta}$ . For this example, Figure 5 shows the frequency points  $\vec{f} \in \mathcal{L}_{\bar{U}}$  and highlights the points where ideal  $H(\vec{f}) = 1$  with filled circles.

We can perform regridding of a signal  $x$  from  $\mathcal{L}_W$  to  $\mathcal{L}_P$  by working in the frequency domain:

1. Compute the  $\mathcal{L}_W/\mathcal{L}_U$  MRL-DFT of  $x$ .
2. Multiply pointwise by  $H(\vec{f})$ , which in this case just means retain the spectral coefficients that lie within  $\mathcal{U}_{\bar{W}} \cap \mathcal{U}_{\bar{P}}$ , and discard (i.e., set to 0) the rest.
3. Compute the  $\mathcal{L}_P/\mathcal{L}_U$  inverse MRL-DFT.

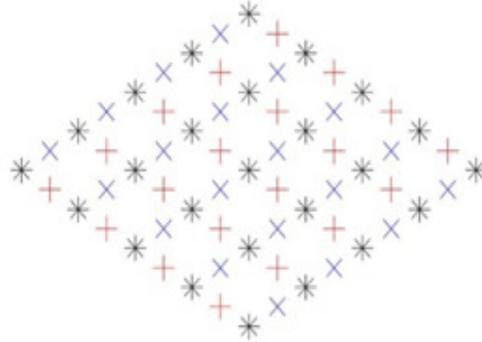


Figure 4: Two lattices, marked 'x' and '+' respectively, with common points marked '\*'. We wish to convert a signal sampled on the 'x' lattice to one with support on the '+' lattice.

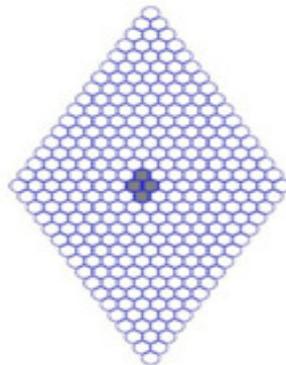


Figure 5: The passband of an ideal interpolation filter, as seen on the reciprocal lattice.

The operation can also be performed in the spatial domain. For this, we need to convolve with the impulse response  $h(\vec{r})$  of  $H$ , which has support on  $\mathcal{L}_V$ . The coefficients  $h$  (which are complex valued, in general) can be found by computing the  $\mathcal{L}_V/\mathcal{L}_U$  inverse MRL-DFT of  $H$ . The result for this case is shown in Figure 6. In order to illustrate the filter coefficients, we show a contour plot of  $|h(\vec{r})|$  if it were computed on the continuum  $\vec{r} \in \mathcal{R}^D$  and then mark with 'x' the points on  $\mathcal{L}_V$  where it should be sampled. The filter coefficients satisfy symmetry conditions  $h(x, y) = h(x, -y)$  and the values at points in the first quadrant are given in Table 1.

Table 1: Computed ideal regridding coefficients.

$\vec{r}$	$\vec{n}$	$h(\vec{r})$
(0,4)	(-2,2)	$2 + j0$
$(\sqrt{3}, 5)$	(-2,3)	$1 + j(-\sqrt{2} + 1)$
(0,2)	(-1,1)	$(\sqrt{2} + 2) + j0$
$(\sqrt{3}, 3)$	(-1,2)	$(\sqrt{2} + 1) - j1$
$(2\sqrt{3}, 4)$	(-1,3)	$1 - j1$
(0,0)	(0,0)	$2 + j0$
$(\sqrt{3}, 1)$	(0,1)	$(\sqrt{2} + 2) - j\sqrt{2}$
$(2\sqrt{3}, 2)$	(0,2)	$2 - j2$
$(3\sqrt{3}, 3)$	(0,3)	$(2 - \sqrt{2}) - j\sqrt{2}$
$(2\sqrt{3}, 0)$	(1,1)	$(\sqrt{2} + 1) - j(\sqrt{2} + 1)$
$(3\sqrt{3}, 1)$	(1,2)	$1 - j(\sqrt{2} + 1)$
$(4\sqrt{3}, 2)$	(1,3)	$0 - j\sqrt{2}$
$(4\sqrt{3}, 0)$	(2,2)	$0 - j2$
$(5\sqrt{3}, 1)$	(2,3)	$(-\sqrt{2} + 1) - j1$
$(6\sqrt{3}, 0)$	(3,3)	$(-\sqrt{2} + 1) + j(-\sqrt{2} + 1)$

Then, conceptually, the spatial domain regridding algorithm is:

1. Upsample from  $\mathcal{L}_W$  to  $\mathcal{L}_V$  by inserting 0's.
2. Interpolate the values by convolving with  $h$ .
3. Decimate down to  $\mathcal{L}_P$ .

This process, in reality, would be realized in polyphase form to achieve computational efficiency, as described in this paper.

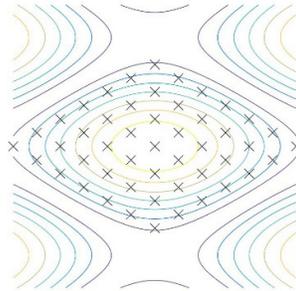


Figure 6: Ideal regridding filter coefficients, located at the lattice points marked 'x', with an indicated contour plot of their amplitude.

Observe that the process of identifying sublattices, and developing multirate algorithms, such as regriding, is relatively straightforward because the physical nature of the spatial and frequency domain coordinates is retained. Directly finding the lexicographic indices of the support of the interpolation filter, for example, would be much more challenging than the approach taken here.

## 6. FUTURE WORK

The MRL-DFT and related results presented in this paper form the foundation for the derivation of a number of important signal processing algorithms. One area to be studied is the formulation of FFT algorithms for multiresolution lattices. The results presented here on spectral properties of signals with support on cosets of sublattices could potentially be used to generate general tree-structured FFT algorithms for multiresolution lattices, in which spectral information of multiresolution signals can be studied directly, without conversion to a universal (high density) resolution. Another topic for further inquiry is formulating fast algorithms for regriding (interpolation), based on generalization of fast convolution operations known for 1-D (and separable multidimensional) cases.

## 7. CONCLUSIONS

We have presented a framework for processing signals on general, variable density sampling grids in the spatial and frequency domains. We formulated the signal processing operations in terms of physically meaningful spatial and frequency coordinates, which retains the significance of spatial correlations in the signal. The underlying mathematical approach was to consider lattice chains, and develop the MRL-DFT as the finite Fourier transform over associated quotient groups. This led to formulation of important relationships, such as the convolution theorem, aliasing and imaging associated with multirate operations, and development of interpolation (regriding) algorithms. We also presented techniques for analysing the lattice structures, for example construction of multiresolution sublattices. Special properties, such as the spectrum associated with signals restricted to sublatticecosets, and the polyphase form, can lead to the further development of special algorithms such as generalized FFT. We related our results to the standard lexicographic form that uses normalized integer indices. In the end, data is stored and processed in rectangular arrays, as computer memory has that structure, but algorithms are developed and understood better in a physical coordinate system. The techniques presented here are useful in a myriad of imaging applications where data is obtained in either the spatial or frequency domains on a non-rectangular grid or possibly an irregular point set with potentially variable density. For example, rather than being forced to artificially regrid data collected at variable densities to a grid with uniform density, the interpolation can be targeted for a multiresolution lattice structure, which can be processed directly. This can help avoid artifacts caused by regriding to densities that are locally too high or too low compared to that of the original data.

## REFERENCES

- [1] Dudgeon, D. & Mersereau, R. (1983). Multidimensional digital signal processing. Saddle Brook, NJ: Prentice-Hall.
- [2] Xing, S. & Li, Y. & Wang, X. (2013). Three-dimensional reconstruction of man-made objects using polarimetric tomographic SAR. IEEE Trans. Geosci. & Rem. Sens., 51(6), 3694-3705. doi: 10.1109/TGRS.2012.2220145.
- [3] Vaidyanathan, P.P. (1993). Multirate systems & filter banks. Saddle Brook, NJ: Prentice-Hall.

- [4] Van de Walle, R. & Lemahieu, I. (2000). Reconstruction of MR images from data acquired on a general nonregular grid by pseudoinverse calculation. *IEEE Trans. Med. Imag.*, 19(12), 1160-1167. doi: S 0278-0062(00)10881-X.
- [5] Conroy, B. (2005). Reconstruction of spiral MRI using FFT on hexagonal lattice structures (Master's Thesis). Retrieved from Cooper Union Archives. (C 7A2 EE 2005 C754).
- [6] Yang, Y. & Liu, Q. (2017). Pseudo-polar Fourier transform-based compressed sensing MRI. *IEEE Trans. Biomed. Eng.*, 64(4), 816-826. doi: 10.1109/TBME.2016.2578930.
- [7] He, X.Y. & Zhou, X.Y. (2012). Fast 3D-ISAR image simulation of targets at arbitrary aspect angles through nonuniform fast Fourier transform (NUFFT). *IEEE Trans. Ant. & Propag.*, 60(5), 2597-2602. doi: 10.1109/TAP.2012.2189717.
- [8] Rundblad, E. & Astola, J. & Egiazarian, K. (2002). Fast algorithms of multidimensional discrete nonseparable K-wave transforms. *IEEE Trans. Sig. Proc.*, 50(6), 1496-1507. doi: S1053-587X(02)04400-8.
- [9] Moustakas, A. & Mertikopoulos, P. & Bambos, N. (2016). Power optimization in random wireless networks. *IEEE Trans. Info. Theory*, 62(9), 5030-5058. doi: 10.1109/TIT.2016.2594183.
- [10] Ezzeldin, Y. & Seddik, K. (2016). Pseudo-lattice treatment for subspace aligned interference signals. *IEEE Trans. Vehic. Tech.*, 63(9), 4728-4734. doi: 10.1109/TVT.2014.2317753.
- [11] Peng, S. & Liu, A. & Wang, H. (2017). Hexagonal multicarrier faster-than-Nyquist sampling. *IEEE Access*, 5(1), 3322-3339. doi: 10.1109/ACCESS.2017.2674666.
- [12] Huang, Y.C. (2017). Lattice index codes from algebraic number fields. *IEEE Trans. Info. Theory*, 63(4), 2098-2112. doi: 10.1109/TIT.2017.2655034.
- [13] Zhang, W. & Wang, W. (2016). Signal shaping and precoding for MIMO systems using lattice codes. *IEEE Trans. Wireless Comm.*, 15(7), 4625-4634. doi: 10.1109/TWC.2016.2543213.
- [14] Muramatsu, S. & Furuya, K. & Yuki, N. (2017). Multidimensional nonseparable oversampled lapped transforms: theory and design. *IEEE Trans. Sig. Proc.*, 65(5), 1251-1264. doi: 10.1109/TSP.2016.2633240.
- [15] Rauh, A. & Arce, G. (2017). Optimized spectrum permutation for the multidimensional sparse FFT. *IEEE Trans. Sig. Proc.*, 65(1), 162-172. doi: 10.1109/TSP.2016.2599483.
- [16] Horn, R. & Johnson, C. (2012). *Matrix Analysis*. Cambridge, UK: Cambridge Press.

## AUTHOR

**Fred L. Fontaine** is the Jesse Sherman Professor and chair of the Department of Electrical Engineering at The Cooper Union in New York City. He has a Ph.D. (Stevens Inst.), M.E. (Cooper Union), and B.E. (Cooper Union) in electrical engineering, and an M.S. (Courant Inst. NYU) in math. He has extensive experience as a research consultant in industry. His areas of interest include multidimensional and multi rate signal processing, radar, communication systems, probability and stochastic processes, machine learning and processor architectures. He is a member of Tau Beta Pi, Eta Kappa Nu and IEEE.



## **AUTHOR INDEX**

*Ahmad A. Alhamed 01*

*Cristiano Akamine 59*

*Fred L. Fontaine 71*

*Jefferson Jesus Hengles Almeida 59*

*Johannes Gottschling 29*

*Kristof Friess 45*

*Lopes P. B 59*

*Maha S. Alqhtani 01*

*Nizam Omar 59*

*Robert Martin 29*

*Saad Bashir Alvi 29*

*Shahrzad Sedaghat 17*

*Volker Herwig H.C 45*