HELLENIC REPUBLIC
National and Kapodistrian University of Athens
Department of Informatics and Telecommunications

ABSTRACTS OF DOCTORAL DISSERTATIONS

Athens 2017
Volume 12

HELLENIC REPUBLIC
National and Kapodistrian University of Athens

Department of Informatics and Telecommunications

ABSTRACTS OF DOCTORAL DISSERTATIONS

The Committee of Research and Development Activities

A. Eleftheriadis

M. Koubarakis

E.S. Manolakos

T. Theoharis

ISSN: 1791-7948

Volume 12

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
Panepistimiopolis, 15784 Athens, Greece

PREFACE

This volume includes extended abstracts of Doctoral Dissertations conducted in the Department of Informatics and Telecommunications, University of Athens, and completed from 1/2016 to 12/2016.

We publish this volume to demonstrate the breadth and quality of the original research performed by our Ph.D. students and faculty and to facilitate the dissemination of their innovative research results. We are happy to present the 12th yearly collection of this kind and expect this initiative to continue in the years to come. The submission of an extended abstract in English is required by all graduating doctoral students in our Department.

We would like to thank all graduates who contributed to this volume and hope that this was a positive experience for them. Finally, we would like to thank PhD candidate Nikos Bogdos for his help and attention to detail in putting together this volume.

The painting in the cover is called *Movement of Vaulted Chambers* (1915) by the artist *Paul Klee* (1879-1940).

The DiT Dept. Committee on Research and Development Activities

A. Eleftheriadis
M. Koubarakis
E.S. Manolakos (publication coordinator)
T. Theoharis

Athens, May 2017

Table of Contents

Bioinformatics methods for the discovery of network signatures towards understanding of underlying molecular mechanisms and investigation of candidate drugs

Marilena Bourdakou*

National and Kapodistrian University of Athens,
Department of Informatics and Telecommunications
kmarlen1988@gmail.com

Abstract. Systemic approaches are essential in the discovery of disease-specific genes, offering a different perspective and new tools on the analysis of several types of molecular relationships, such as gene co-expression or protein-protein interactions. However, due to lack of experimental information, this analysis is not fully applicable. The aim of this study is to reveal the multi-potent contribution of statistical network inference methods in highlighting significant genes and interactions. We have investigated the ability of statistical co-expression networks to highlight and prioritize genes for breast cancer in terms of: (i) classification efficiency, (ii) gene network pattern conservation, (iii) indication of involved molecular mechanisms and (iv) systems level momentum to drug repurposing pipelines. We have found that statistical network inference methods are advantageous in gene prioritization, are capable to contribute to meaningful network signature discovery, give insights regarding the disease-related mechanisms and boost drug discovery pipelines from a systems point of view.

KEYWORDS: network inference methods, gene expression data, co-expression networks, molecular mechanisms, drug repurposing

1 Introduction

Breast cancer is a major public health problem, since it remains the most frequently diagnosed cancer and ranked second as a cause of death in women population. Outbreaks are increasing in most countries, despite current efforts have been made to avoid the disease [1]. This happens because breast cancer is a complex disease with many contributing factors affecting the progress of the disease. Despite the fact that many studies have been conducted, neither the exact etiology of the breast cancer, nor

* Dissertation Advisor: Emmanouil Sagkriotis, Associate Professor.

the mechanisms behind the heterogeneity from patient to patient are known. For this, the diagnosis and the treatment of breast cancer remain a both challenging and fascinating task [2].

With the rapid development of genome-wide gene expression profiling methodologies, many bioinformatics data analysis pipelines have been developed to identify breast cancer related genes and discover gene signatures for prognosis and treatment prediction. However, since breast cancer is a complex disease, it should be determined not only by individual genes, but also by the coordinated effect of numerous genes [3]. The information behind gene interaction networks is of great importance due to the fact that all cellular functions are regulated by gene patterns, where the presence or absence of an interaction may cause the emergence of a disease.

Network analysis and graph theory support the study of interactions among relatively large number of genes in order to conclude to large lists of statistically significant genes [4]. Several bioinformatics tools prioritize genes by combining gene expression data with the protein-protein interaction (PPI) network through a random walk approach to enrich the candidate genes and finally re-rank them. The majority of these methods necessitate prior knowledge to re-rank genes accordingly. However, due to the absence of functional characterizations for a significant number of genes, these approaches are not fully applicable [5]. Genome-wide association studies (GWAS) have recognized DNA variants that are related to common complex diseases but for many of these studies, functional associations between genes and diseases are unknown [6].

In order to overcome this hurdle, several network inference methods have been adopted to construct statistical co-expression networks, based on gene expression data. These network inference approaches identify groups of genes that are highly correlated in expression levels to multiple samples according to a variety of correlation functions and algorithms [7].

In this study, we investigate the ability of statistical co-expression networks to highlight and prioritize significant genes at four different breast cancer molecular subtypes, including Luminal A, Luminal B, HER2 and Triple Negative as well as at four different disease stages (I-IV) in terms of: (i) classification efficiency, (ii) gene subnetwork conservation, (iii) involved molecular mechanisms investigation and (iv) potential boost to drug repurposing pipelines.

Specifically, we have used mRNA gene expression microarray data concerning Breast Invasive Carcinoma, retrieved from The Cancer Genome Atlas – TCGA (http://gdac.broadinstitute.org/runs/stddata__latest/samples_report/BRCA.html), to reconstruct 17 different networks (twelve based on mathematical correlation and six based on the literature) of the top differentially expressed genes. Using a mathematical function that combines gene expression data with custom networks, we prioritized genes based on each network. Furthermore, in order to investigate the quality of each prioritized gene list, we elucidated the impact of each one over sample discrimination, by applying a hold out validation scheme using the TCGA data as training set and a number of Breast cancer datasets from the transcriptional data repository Gene Expression Omnibus GEO (http://www.ncbi.nlm.nih.gov/geo/) as test sets. Using the network inference method that performed the highest classification score, we constructed co-expression networks for all datasets (train and test sets) to

find the most significant gene-gene links that recur in all networks. With the proposed pipeline, we concluded to breast cancer specific network patterns per subtype and stage. Analyzing each pattern we concluded in specific mechanisms per subtype and stage related to cellular community (cell communication, focal adhesion), signaling (in terms of extracellular matrix and cytokine receptor interactions), cell growth and death (cell cycle), immune system (including complement and coagulation cascades and toll like receptor signaling pathway), endocrine system (ppar and adipocytokine signaling pathway), carbohydrate, lipid and amino acid metabolism (glycolysis/gluconeogenesis, fatty acid and glycerolipid metabolism, bile acid biosynthesis, as well as tyrosine, phenylalanine, glycine, serine, threonine metabolism) and xenobiotics biodegradation and metabolism (3 chloroacylic acid and 1,2 methylnaphthalene degradation, metabolism of xenobiotics by cytochrome p450). Interestingly, all the derived network patterns include genes found in breast cancer specific regions of significant somatic copy number alterations (SCNA) [8]. Finally, the genes from the conserved network patterns were used in a drug repurposing pipeline, revealing drugs that have the potential to suppress breast cancer specifically for each molecular subtype and stage of the disease.

2 Methods

2.1 Datasets and preprocessing

Reference Set: TCGA mRNA (microarray) gene expression data for Breast Invasive Carcinoma cases are obtained from Firehose (http://gdac.broadinstitute.org/). From a total 587 samples (526 primary solid tumor samples and 61 primary solid normal samples - 17.814 genes), we have selected a subset of tumor data containing information regarding breast cancer staging, HER2, ER and PR status with their corresponding normal samples. Concerning staging, selection of stages I, II, III and IV was performed based on the clinical records accompanying each sample, while for the case of subtyping, the selection was performed as followed: (i) Luminal A for ER+ and/or PR+ , HER2- , (ii) Luminal B for ER+ and/or PR+ , HER2+ , (iii) HER2 for ER- , PR- , HER2+ and (iv) Triple Negative for ER-, PR-, HER2-. The eight distinct TCGA dataset were statistically analyzed with the LIMMA R package in order to select the Differentially Expressed Genes (DEGs) in breast cancer samples compared with the normal ones. The top 1000 genes of each sub-dataset with p-value < 0.01 and q-value < 0.01, sorted based on their log Fold Change absolute value, were used as the reference sets in our analysis.

Validation Sets: We searched in Gene Expression Omnibus accessed on 19 November 2015 using queries, in order to find microarray datasets for each breast cancer stage and subtype. Finally we concluded in 7 independent datasets from which, one contain clinical feature from both stages and subtypes.

2.2 Network Reconstruction

We have examined 3 major categories of statistical network inference methods: (i) Mutual Information-based methods, (ii) Correlation-based methods and (iii) Tree-

based methods. Also, we utilized Biological information-based network methods and one ensemble scheme using all statistical network inference methods. More specifically, we have used 11 network inference methods to reconstruct gene co-expression networks for each dataset including the top 1000 DEGs from the TCGA dataset. All the selected methods are implemented in R packages. Six mutual information based methods are used (Aracne.a, Aracne.m, CLR, MRNET, MRNETB and C3NET), four correlation based (Lasso, Adaptive Lasso, GeneNET and WGCNA) and one tree based – Genie3. Furthermore, we have used the Cytoscape platform and more specifically the GeneMania plug-in to reconstruct a gene network using biological information. The GeneMANIA algorithm inside the homonymous plugin obtains information from a combination of potentially heterogeneous sources. This plug-in uses a large data set unifying functional networks comprising approximately 800 networks for 6 organisms including Homo sapiens. Using the Homo sapiens network we constructed a sub – network for the top 1000 DEGs from the TCGA dataset merging 5 Network types (Co-expression, co-localization, physical interaction, genetic interaction and pathway). We also used the manually curated human signaling network [9] based on the literature since 2005 (Version 6). The signaling network contains more than 6,000 proteins and 63,000 relations from different data sources including BioCarta, CST Signaling pathways, Pathway Interaction database (PID), iHOP, and many review papers on cell signaling. The signaling network comprised of three different relations (activation, inhibition and physical interactions). This network was used not only as a whole network (all relations), but was further divided into three sub-networks based on the different relation types.

Finally, we have created a union unique gene list based on the different top 100 re-ranked gene lists from the eleven statistical network inference methods. Based on the highest frequency of the appearance, the minimum mean rank and the minimum coefficient of variation across all statistical network inference methods we selected the top 100 genes.

2.3 Gene re-ranking using underlying networks

In order to investigate the influence of the reconstructed 17 gene networks (12 statistically and 5 biologically inferred) on gene prioritization, we applied a method that allows for a custom network selection combining the log fold change absolute values with the selected underlying network in order to re-rank the initial DEGs. The basic idea of the method is the reconciliation of the gene expression values taking into account an underlying gene network. This approach is available as part of the Biorithm software in the Network Reconciliation package [10].

2.4 Scoring the ranked gene lists

Each method is scored according to the maximum achieved mean classification accuracy across datasets, modified by two multiplicative weights: w_n that is related to the number of genes required for the maximum accuracy and w_{cv} that is related to the coefficient of variation (CV) of the classification accuracy along the first 100 genes.

Finally, we calculated the average score of each method across the stages and the subtypes.

3 Results

3.1 Evaluation of gene re-ranking through a classification scheme

The top 1000 re-ranked gene lists for each subtype and stage, along with the initially ranked list, gave us a total number of 18 ranked gene lists. In order to evaluate each list, we elucidated the impact of the top 100 genes from each list over sample discrimination, by applying a hold out validation scheme. More precisely, we employed a Support Vector Machine (SVM) – based classification scheme using the e1071 R package through sequential gene selection of the first 100 genes, using as Train set the expression values of each top 100 gene list from the reference set (TCGA) and as Test sets the expression values of the same top 100 genes from a number of independent GEO datasets (discovery sets) available for each subtype and stage. We followed the same procedure for each top 100 gene lists and we calculated the mean classification accuracy from the discovery datasets in a sequential gene selection manner. Figures 1 and 2 show the box plots of the mean classification accuracies of the top 100 sequential genes for each network approach using the Page Rank reconciling method for each stage and subtype. We observe that, in most cases the median classification performances of the top 100 gene lists from network inference methods are either better or equivalent compared to the median performance of the initial gene list.

Each ranking method is scored according to the maximum achieved mean classification accuracy across datasets, evaluated by a score (see Methods). The maximum average score for breast cancer stages and subtypes was achieved by Genenet network inference method and MRNETB, respectively. For this reason we adopted them for the rest of our analysis. It is worth mentioning that the selected statistical network inference methods achieved a higher or equivalent score compared to the initial ranking in most cases.

3.2 Deriving a common Network Pattern

We applied the Genenet and MRNETB network inference methods to reconstruct gene co-expression networks for each of the available dataset for each stage and subtype. In order to highlight any common gene network pattern, we found the common edges across all datasets. We performed a dynamic filtering to keep only the highly weighted gene - gene links. Finally, we came up with 205 genes-nodes and 216 edges for Stage I, 561 genes-nodes and 896 edges for Stage II, 289 nodes and 380 edges for Stage III and 132 genes-nodes and 169 edges for Stage IV. As far as subtypes are concerned, we came up with 196 genes-nodes and 872 edges for Triple Negative, 201 genes-nodes and 272 edges for Luminal A, 155 genes-nodes and 305 edges for Luminal B and 544 genes-nodes and 573 edges for HER2.

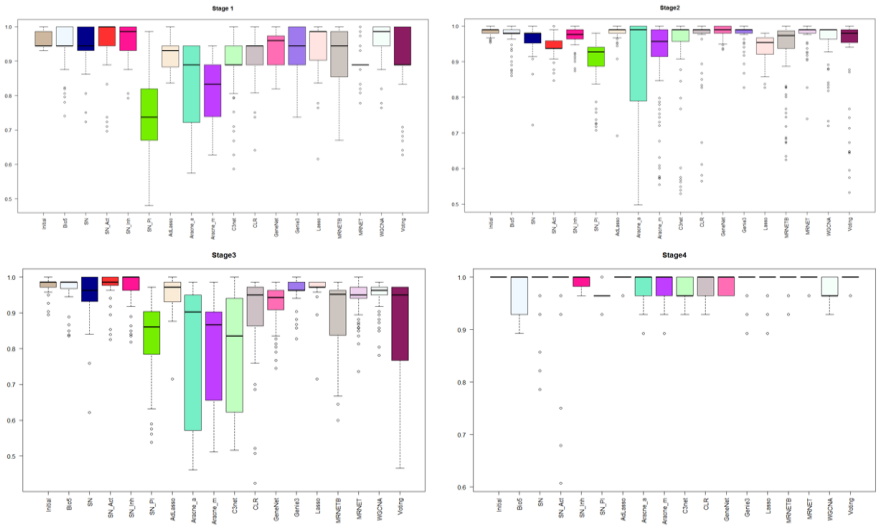


Figure 1: Box plots of the mean accuracy rates of the top 100 sequential genes from all ranked and re-ranked gene lists in combination with PageRank reconciling method, using hold out validation with train set the TCGA expression values and test set the expression values from GEO independent datasets for breast cancer stages.

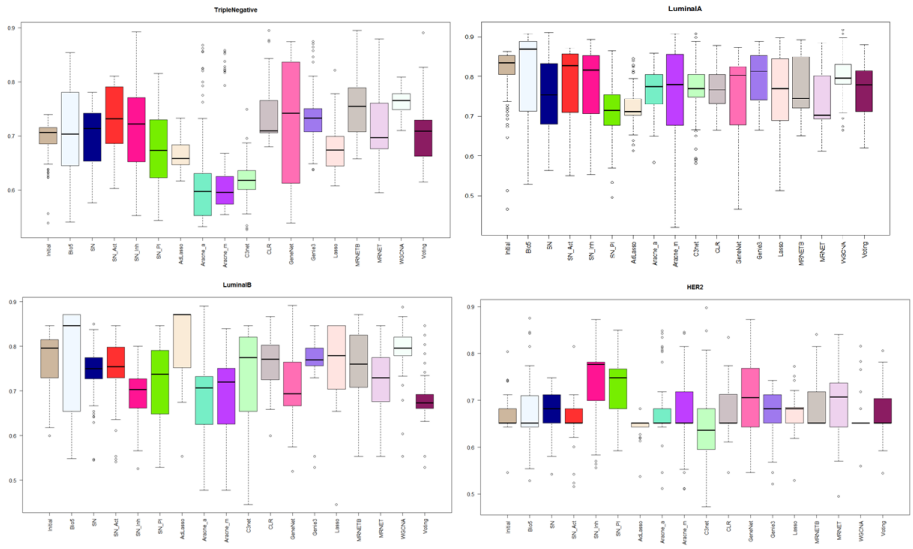


Figure 2: Box plots of the mean accuracy rates of the top 100 sequential genes from all ranked and re-ranked gene lists in combination with PageRank reconciling method, using hold out validation with train set the TCGA expression values and test set the expression values from GEO independent datasets for breast cancer subtypes.

3.3 Network inference, underlying mechanisms

We used the Enrichr web-based software application in order to find the underlying significant biological pathways derived from genes of each network pattern. Common and exclusive mechanisms of each stage and subtype were further investigated.

Following pathway analysis of our findings for the case of Staging, we have found four exclusive stage-related pathways including phenylalanine metabolism for Stage II, peroxisome proliferator-activated (PPAR) signaling pathway and glycolysis and gluconeogenesis for Stage III and toll like receptor signaling pathway for Stage IV.

For the case of Luminal A, Luminal B, HER2 and TN subtypes, we have found seven exclusive subtype-related pathways, including glycine serine and threonine metabolism pathway for Luminal B, glycerolipid metabolism, fatty acid metabolism, complement and coagulation cascades and bladder cancer for HER2 and small cell lung cancer and metabolism of xenobiotics by cytochrome p450 for TN.

3.4 Network inference and drug repurposing

The network patterns were further processed in order to investigate their contribution regarding the discovery of potential drugs for breast cancer subtypes and stages. Actually, genes that constitute the common network patterns from each subtype and stage were divided into up and down regulated, based on their Fold Change from the initial statistical analysis of the TCGA reference sets. The up and down regulated genes formed disease signatures that were queried in a well-established drug repurposing pipeline. Namely, LINCS-L1000 (http://www.lincscloud.org/) is the advanced version of cMap [11] with significantly increased number of drug treatments, cell types and gene signatures based on L1000 high throughput technology. We used the LINCS-L1000 detailed report and we collected the top 20 drugs for each gene list with the most negative enrichment scores. The negative score suggests that the drugs are considered to be inhibitors. We then derived a list of 80 drugs regarding the stages (20 drugs per stage) and 80 drugs regarding the subtypes (20 drugs per subtype). DrugBank database, as well as ChemSpider tool was used to find their chemical structures. The resulted drug lists (names and structures) were further evaluated via ChemBioServer [12], a web application for searching, filtering and comparing drug structures. More specifically, we compared each top 20 drug list from LINCS with 25 known FDA-approved Breast cancer therapeutic drugs. Hierarchical clustering using tanimoto similarity (Soergel distance) was applied to each of the top 20 drug list from LINCS and the 25 known FDA-approved Breast cancer therapeutic drugs. In synopsis, the unique drugs for the breast cancer stages were 63 and for the breast cancer subtypes 58, as we have located common drugs across them.

To further examine the resulted drugs, we constructed a super network that combines each of the top 20 drugs extracted from our analysis with the 25 FDA approved breast cancer drugs, with their target genes and finally with the respective common network pattern. We used the DrugBank database in order to find the target

genes of all drugs from LINCS and the 25 FDA approved Breast Cancer drugs. GeneMANIA plug-in was applied to identify which genes from each pattern were physically interacting with the target genes. Our goal was to understand the correlations between drugs, drug targets and conserved co-expressed genes from a network-based view, in order to outline small paths that are of great importance in breast cancer stages and subtypes. Each network consists of four sub-networks, two drug – drug similarity networks, a drug – target network and a drug target – common pattern genes co-expression network, as shown in Figures 3-4. In Figures 3-4, the yellow cycles represent each top 20 drug list from LINCS and the green cycles the 25 FDA Breast cancer Drugs. Edges between the two cycles represent their structural similarity. As much thicker is the edge, the greater the similarity between the drugs. Only edges with similarity greater than 0.5 are presented. Grey cycles (Figures 3-4) depict the target genes. As we described above, we found the corresponding target genes of the total drugs by means of the DrugBank database. Drug- target associations are represented with red dots. Purple ellipses typify top 100 genes from each common network pattern. Blue edges represent physical interactions between target genes and genes from each common network pattern.

As shown in Figure 3, one drug out of 25 FDA approved Breast cancer drugs, Gemcitabine, was proposed as repurposed drug by the LINCS for breast cancer stage I. Furthermore, Gemcitabine is quite similar (tanimoto similarity greater than 80%) with Clofarabine and Kinetin-riboside (repurposed drugs from LINCS). Clofarabine is also an anti-cancer, antineoplastic chemotherapy drug and is classified as an antimetabolite. Moreover, Vinblastine – Breast Cancer drug was found to be greater than 60% structurally similar with Sepantronium bromide (repurposed drug from LINCS), which is a small-molecule proapoptotic agent with potential antineoplastic activity. Vinblastine has three target genes TUBA1A, TUBB and JUN. The latter was found to physically interact with three genes (ATF3, FOS and EGR1) of the breast cancer stage I network pattern. As shown in Figure 4, Idarubicin (repurposed drug from LINCS) was also found to be 85% structurally similar with Doxorubicin and Epirubicin and they are all topoisomerase 2 inhibitors (TOP2A). Super Networks were constructed and analyzed for each stage and breast cancer subtype.

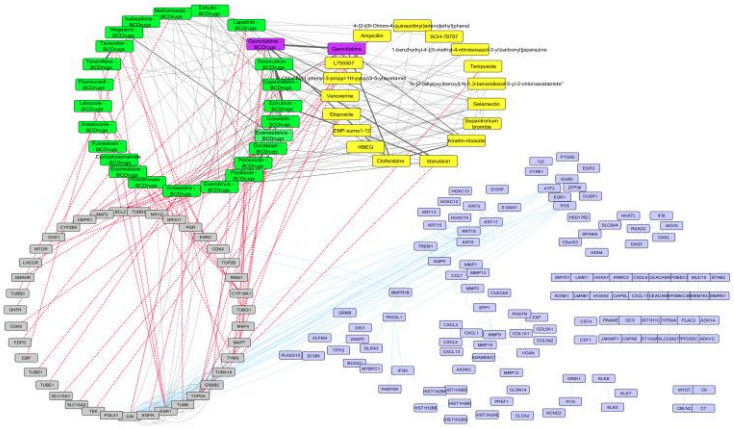


Figure 3. Super Network for breast cancer Stage I- consists of 4 sub-networks: 1) two drug – drug networks: with yellow cycle are represented the 21 drugs from LINCS and with green cycle the 20 therapeutic breast cancer drugs 2) drug – target network: grey round rectangles represent the target genes of all drugs (red dots edges) and 3) target - pattern genes network: physical interactions (blue edges) between target genes and genes from the network pattern (purple ellipses). One out of the 25 FDA approved Breast cancer drugs (Gemcitabine), was found in the top 20 drug list from LINCS from breast cancer stage I (dark magenta).

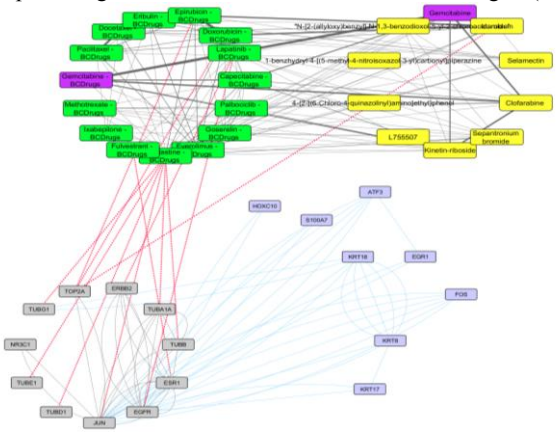


Figure 4. Highlighted target genes that physically interact with genes from the breast cancer stage I common network pattern and their corresponding repurposed drugs from LINCS, along with their structurally similar Breast cancer drugs.

4 Discussion

In the present work, we used eleven network inference methods and one ensemble scheme to reconstruct gene co-expression networks in order to examine their contribution in identifying significant genes and gene-gene links related to different breast cancer stages and subtypes. During this assessment, we demonstrated that, in

most cases of breast cancer stages and subtypes, the statistical co-expression networks produce either similar or more enriched lists with significant genes (in terms of maximum classification accuracy achieved) for each breast cancer stage and subtype than the conventional statistical approach or the networks based solely on the biological information extracted from the literature. Actually, the dominance of statistical networks is profound in the analysis of breast cancer subtypes, whereas in the case of stage analysis, the simple statistical method (Initial) and the signaling network based on inhibition (SN_I) give slightly better (almost equivalent) scoring than statistical networks.

Furthermore, our analysis concluded to eight network patterns, four for the stages (I, II, III and IV) and four for the subtypes (Triple Negative, Luminal A, Luminal B and HER2). Additionally, we further analyzed the gene patterns, in order to investigate potential mechanisms and drugs for breast carcinomas staging and subtypes. As described in the previous section, we have found four exclusive stage-related pathways. Peroxisome proliferator-activated (PPAR) signaling pathway has been implicated in the pathology of numerous diseases including obesity, diabetes, atherosclerosis, and cancer. More specifically, PPAR signaling pathway has been reported as a possible important predictor of breast cancer response to neoadjuvant chemotherapy [13]. Five dehydrogenase (ADH) isoenzymes and aldehyde dehydrogenases (ALDH) genes from the breast cancer Stage III network pattern were involved in the glycolysis and gluconeogenesis pathway. It has been reported that patients with advanced breast cancer had changes in the activity of activity of ADH isoenzymes and ALDH [14]. Furthermore, from the breast cancer Stage IV pattern we have found an exclusive pathway - toll like receptor signaling pathway for which it is well known that supports tumor cell growth in vitro and in vivo [15]. For the case of breast cancer subtypes, we have found seven exclusive subtype-related pathways. Hyperactivation Glycine serine and threonine metabolism pathway drives to oncogenesis and recent developments support that this pathway may provide novel opportunities for drug development and biomarker identification of human cancers [16]. Moreover, from the Triple Negative pattern we found the metabolism of xenobiotics by cytochrome p450 pathway. Cytochromes P450 (CYPs) play a pivotal role in cancer formation and cancer treatment as they participate in the inactivation and activation of anticancer drugs [17].

Most of the specific mechanisms per subtype and stage are related to cellular community, signaling, cell growth and death, immune and endocrine systems, carbohydrate, lipid and amino acid metabolism as well as xenobiotics biodegradation and metabolism. Furthermore, all the derived network patterns include genes found in breast cancer specific regions of significant somatic copy number alterations (SCNA) [8]. These results are fully aligned to the up-to-date recognized cancer hallmarks related to cell growth, metabolism, immune system, inflammation and genome duplication [18].

The resulted network patterns were also analyzed by means of LINCS drug reposition pipeline. Two out of 25 therapeutic FDA approved breast cancer drugs (Gemcitabine and Palbociclib) were also found as repurposed drugs from LINCS. In Stage I, two repurposed drugs Clofarabine and Kinetin-riboside were found to be

structurally similar to Gemcitabine. Clofarabine seems to have potential efficacy in epigenetic therapy of solid tumours, especially at early stages of carcinogenesis [19]. For Stage II, Cladribine (repurposed drug) was found to be structurally similar with Triciribine (repurposed drug) and Gemcitabine and Capecitabine Breast cancer drugs. In clinical trial (June, 2015) triciribine phosphate, when given together with paclitaxel, doxorubicin hydrochloride, and cyclophosphamide, works in treating patients with stage IIB-IV breast cancer (https://clinicaltrials.gov).

Moreover, in Stage III Ruxolitinib and Pyrvinium-pamoate repurposed drugs from LINCS have been found as structurally similar with Letrozole and Vinblastine Breast cancer drugs respectively. An ongoing clinical trial (October, 2015) compares the overall survival of women with advanced (Stage III) or metastatic (Stage IV) HER2-negative breast cancer who receive treatment with Capecitabine in combination with Ruxolitinib versus those who receive treatment with Capecitabine alone (https://clinicaltrials.gov). Irinotecan has been examined in a clinical trial in Phase II in order to find its objective response rate in patients with metastatic breast cancer (Stage IV) (https://clinicaltrials.gov).

In case of repurposed drugs for breast cancer subtypes, we have found that Etoposide and Teniposide (repurposed drugs) as structurally similar with two Breast cancer drugs Epirubicin and Doxorubicin in Triple Negative subtype. These four drugs are topoisomerase ii inhibitors (TOP2A) and Etoposide has been found as effective drug in Chinese women with heavily pretreated metastatic breast cancer [20].

In Luminal A, the target genes of Vorinostat physically interact with two genes (RUNX1T1 and SMYD1) from the Luminal A pattern. It has been reported that Vorinostat in combination with Tamoxifen may treat patients with hormone therapy-resistant breast cancer [21]. In Luminal B, F10 and EGFR genes from Luminal B pattern are also target genes of Menadione (repurposed drug from LINCS) and Lapatinib Breast cancer drug. Menadione has been examined on its antiproliferative action on breast cancer cells50. Finally in HER2 subtype, Palbociclib is also a Breast cancer drug that was found from the drug repurposing analysis of HER2 pattern. It has quite similar structure with WZ-4002 repurposed drug which is a novel, mutant inhibitor of EGFR.

Finally, the action of the remaining mechanisms and drugs found from LINCS may be further investigated since they have been derived from significantly relevant genes related to breast cancer stages and subtypes.

References

1. Howell, A. *et al*. Risk determination and prevention of breast cancer. Breast Cancer Res 16, 446, doi:10.1186/s13058-014-0446-2 (2014).
2. Hutchinson, L. Breast cancer: challenges, controversies, breakthroughs. Nat Rev Clin Oncol 7, 669-670, doi:10.1038/nrclinonc.2010.192 (2010).
3. Zhang, J. et al. Weighted frequent gene co-expression network mining to identify genes involved in genome stability. PLoS Comput Biol 8, e1002656, doi:10.1371/journal.pcbi.1002656 (2012).

4 Cheng, F. et al. Prediction of drug-target interactions and drug repositioning via network-based inference. PLoS Comput Biol 8, e1002503, doi:10.1371/journal.pcbi.1002503 (2012).
5 Chen, J., Aronow, B. J. & Jegga, A. G. Disease candidate gene identification and prioritization using protein interaction networks. BMC Bioinformatics 10, 73, doi:10.1186/1471-2105-10-73 (2009).
6 Nayak, R. R., Kearns, M., Spielman, R. S. & Cheung, V. G. Coexpression network based on natural variation in human gene expression reveals gene interactions and functions. Genome Res 19, 1953-1962, doi:10.1101/gr.097600.109 (2009).
7 Emmert-Streib, F., Glazko, G. V., Altay, G. & de Matos Simoes, R. Statistical inference and reverse engineering of gene regulatory networks from observational expression data. Front Genet 3, 8, doi:10.3389/fgene.2012.00008 (2012).
8 Zack, T. I. et al. Pan-cancer patterns of somatic copy number alteration. Nat Genet 45, 1134-1140, doi:10.1038/ng.2760 (2013).
9 Cui, Q. et al. A map of human cancer signaling. Molecular systems biology 3, 152, doi:10.1038/msb4100200 (2007).
10 Poirel, C. L. et al. Reconciling differential gene expression data with molecular interaction networks. Bioinformatics 29, 622-629, doi:10.1093/bioinformatics/btt007 (2013).
11 Lamb, J. et al. The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. Science 313, 1929-1935, doi:10.1126/science.1132939 (2006).
12 Athanasiadis, E., Cournia, Z. & Spyrou, G. ChemBioServer: a web-based pipeline for filtering, clustering and visualization of chemical compounds used in drug discovery. Bioinformatics 28, 3002-3003, doi:10.1093/bioinformatics/bts551 (2012).
13 Chen, Y. Z. et al. PPAR signaling pathway may be an important predictor of breast cancer response to neoadjuvant chemotherapy. Cancer chemotherapy and pharmacology 70, 637-644, doi:10.1007/s00280-012-1949-0 (2012).
14 Jelski, W., Chrostek, L., Markiewicz, W. & Szmitkowski, M. Activity of alcohol dehydrogenase (ADH) isoenzymes and aldehyde dehydrogenase (ALDH) in the sera of patients with breast cancer. Journal of clinical laboratory analysis 20, 105-108, doi:10.1002/jcla.20109 (2006).
15 Ahmed, A., Redmond, H. P. & Wang, J. H. Links between Toll-like receptor 4 and breast cancer. Oncoimmunology 2, e22945, doi:10.4161/onci.22945 (2013).
16 Amelio, I., Cutruzzola, F., Antonov, A., Agostini, M. & Melino, G. Serine and glycine metabolism in cancer. Trends in biochemical sciences 39, 191-198, doi:10.1016/j.tibs.2014.02.004 (2014).
17 Rodriguez-Antona, C. & Ingelman-Sundberg, M. Cytochrome P450 pharmacogenetics and cancer. Oncogene 25, 1679-1691, doi:10.1038/sj.onc.1209377 (2006).
18 Wang, E. et al. Predictive genomics: a cancer hallmark network framework for predicting tumor clinical phenotypes using genome sequencing data. Seminars in cancer biology 30, 4-12, doi:10.1016/j.semcancer.2014.04.002 (2015).
19 Lubecka-Pietruszewska, K. et al. Clofarabine, a novel adenosine analogue, reactivates DNA methylation-silenced tumour suppressor genes and inhibits cell growth in breast cancer cells. Eur J Pharmacol 723, 276-287, doi:10.1016/j.ejphar.2013.11.021 (2014).
20 Yuan, P. et al. Oral etoposide monotherapy is effective for metastatic breast cancer with heavy prior therapy. Chin Med J (Engl) 125, 775-779 (2012).
21 Munster, P. N. et al. A phase II study of the histone deacetylase inhibitor vorinostat combined with tamoxifen for the treatment of patients with hormone therapy-resistant breast cancer. British journal of cancer 104, 1828-1835, doi:10.1038/bjc.2011.156 (2011).

Spectrum sharing and management techniques in mobile networks

Konstantinos Chatzikokolakis*

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
kchatzi@di.uoa.gr

Abstract. Radio spectrum has loomed out to be a scarce resource that needs to be carefully considered when designing 5G communication systems. Spectrum sharing is considered unavoidable for 5G systems and this thesis provides a solution for adaptive spectrum sharing under multiple authorization regimes based on a novel architecture framework that enables network elements to proceed in decisions for spectrum acquisition. The decision making process for spectrum acquisition proposed is a novel Adaptive Spectrum Sharing technique that uses Fuzzy Logic controllers to determine the most suitable spectrum sharing option and reinforcement learning to tune the fuzzy logic rules, aiming to find an optimal policy that Mobile Network Operators (MNOs) should follow in order to offer the desirable Quality of Service to their users, while preserving resources (either economical, or radio) when possible. The final contribution of this thesis is a mechanism that ensures fair access to spectrum among the users in scenarios in which conveying spectrum license is not prerequisite.

Keywords: spectrum sharing, spectrum management, fuzzy logic, reinforcement learning, fair resource usage, genetic algorithms.

1 Dissertation Summary

The mobile communications have experienced an exploding growth of the connected devices over the past few years. Quantitative results reported indicate that this phenomenon is not expected to change in the near future and many efforts will be spent in research, standardization and regulation for facilitating the service requirements of 5G networks. The latest reports show that more than 11.5 billion mobile devices will be connected by 2019.

3GPP, motivated by the increased mobile data traffic volume has encouraged the research community to move towards three directions namely: a) spectral efficiency improvement, b) higher network cell density and c) exploitation of underutilized radio spectrum resources. The first solution includes Coordinated Multiple Point (CoMP) transmission using sophisticated MIMO techniques and interference management mechanisms. The second area deals with the addition of extra layer cells in the net-

* Dissertation Advisor: Nancy Alonistioti, Assistant Professor

work with base stations that cover smaller areas compared to macro and micro Base Stations (BSs). These solutions include femto cells and the use of relay nodes. The third aspect, which is the main focus of this thesis, deals with the extension of spectrum opportunities for mobile broadband access. Nowadays, spectrum resources are allocated to Mobile Network Operators (MNOs) from the National Regulatory Authority (NRA) and through network planning are used in different geographical areas. Re-allocating spectrum resources that are not fully utilized to congested areas is not a trivial procedure that may cause undesirable effects to the network or in other circumstances (e.g. re-farming spectrum initially given to other communication systems) may take months or even years to complete. However, as the capacity needs for mobile broadband access increase it is expected that dynamic, adaptive and fast solutions that deal with spectrum scarcity will arise in the near future. Such solutions will perform flexible spectrum management and enable spectrum sharing among multiple communication systems, since the complementary solutions given by 3GPP (i.e., MIMO antennas and CoMP) will not be sufficient to cover the capacity needs. Towards, flexible spectrum management 5G communication systems should be carefully designed to overcome spectrum scarcity and Mobile Network Operators (MNOs) will need to revisit business models that were not of their prior interest (e.g., Cognitive Radio) or consider adopting new business models that emerge (e.g. Licensed Shared Access) so as to cover the extended capacity needs. Up to now, MNOs have been reluctant investing for extra network technologies that would offer spectrum flexibility and preferred following traditional exclusive access scheme for their dedicated spectrum resources, which led to reduced spectrum utilization. MNOs hesitation towards spectrum sharing has been also reinforced by the fact that spectrum sharing techniques proposed in the literature focus on single authorization regimes, limiting thus the flexibility and the potentials of spectrum sharing among multiple communication systems. Thereafter, new flexible mechanisms that will handle spectrum efficiently and will exploit the benefits of various authorization regimes are required.

Towards this direction, this thesis aims at providing a solution for spectrum sharing under multiple authorization regimes based on a novel architecture framework that enables the network components to proceed in decisions for spectrum acquisition and exchange information that will lead to the realization of the proposed concept. The contributions of this dissertation move towards four directions, namely a novel functional architecture in conjunction with information model and data model for enabling spectrum sharing under multiple authorization regimes, a fuzzy logic based spectrum controller giving the opportunity to mobile networks to choose effectively the most proper sharing scheme taking into account network conditions and spectrum market demands, its corresponding learning mechanism based on reinforcements that enables this scheme to adapt the decision making process over time and finally, a complementary scheme for fair resource usage applied in general authorization regime in which there is no need for spectrum license.

Regarding the first contribution of this thesis, we have proposed a novel architecture framework, including functional elements incorporated in (either existing or new) network entities, which drive the decision making process of spectrum acquisition and lead to the realization of the proposed spectrum sharing concept. The incorporation of

Spectrum Controller, a logical entity that is responsible for requesting additional spectrum resources to the operator's network and the Spectrum Manager, a logical (and not necessarily centralized) unit that is responsible to gather information on available spectrum and grant access based on the received spectrum requests have been proposed in the context of this thesis. The whole process is regulated by National Regulation Authorities (NRAs) and the framework may be applied upon multiple spectrum sharing scenarios such as the Licensed Shared Access (LSA), the Co-primary spectrum sharing and other light-licensing sharing schemes [1]. A possible instantiation through a Software-Define Network (SDN) has also been introduced. In our proposal we have assumed a fully SDN capable network for configuring both core and access network elements. MobileFlow forwarding engine (MFFE) and MobileFlow controller (MFC) are considered to be the key enablers of the deployment in the configuration of the core network and the Evolved SoftRAN (E-SoftRAN) is the key enabler when configuring the network elements in the access network [2].

The second contribution of this thesis is related to the decision making process that will enable spectrum sharing under multiple authorization regimes. Based on the proposed architecture Spectrum Controller is responsible to perform the decision making process for spectrum acquisition. In this thesis, we propose a novel spectrum sharing technique that uses Fuzzy Logic controllers to determine the most suitable spectrum sharing option. Fuzzy Logic Controllers (also called Fuzzy Inference Systems) consist of three parts, namely the fuzzifier, the inference system, and the defuzzifier. The first part is responsible to map (fuzzify), the input values to the extent that these values belong to a specific state (e.g., low, medium, high using the input membership functions). The input is a numerical value limited to the universe of discourse of the input variable (it could be a real value, integer, natural, etc.) and the output is a fuzzy degree of membership (always in between the [0,1] interval). The second part (inference system) is responsible to apply the fuzzy operators, apply the implication method and aggregate all inputs. More specifically it uses "if … then…" rules to identify the relation of the inputs to the outputs; each rule results to a certain degree for every output. Then, the output degrees for all the rules of the inference phase are being aggregated by using the output membership functions. Finally, the defuzzifier will perform the defuzzification procedure aggregating the outcomes of all the rules and producing a single crisp value. This value captures the decision of the decision maker. In our contribution, Fuzzy Logic Controllers' decisions take into account network conditions, spectral efficiency and the rules preserved in each Fuzzy Logic Controller, which are defined based on the special features of each spectrum sharing option. Several spectrum sharing options exist based on various authorization regimes, which may be divided into two categories, vertical and horizontal spectrum sharing depending on the predefined priority that each communication system has. In vertical sharing concept there is a license-holder, also known as primary user or incumbent, that could grant usage rights to licensees (as in [3]) or the other players (i.e. besides the license-holder) could use the spectrum in opportunistic way [4]. In [5], a rule-regulated distributed and collaborative spectrum sharing approach is proposed. The solution aims at improved system fairness and spectrum utilization and reduced signaling overhead

but lacks flexibility. However, solutions that enable spectrum sharing on unlicensed basis fail to give QoS guarantees to the users.

In horizontal sharing the communication systems that use the same spectrum have equal rights of usage. Inter-operator spectrum sharing is a typical paradigm of horizontal sharing that has emerged over the past years [6][7][8]. A partially distributed implementation method using game theory and learning algorithms proposed in [6], focusing on sharing in multiple licensed bands and aiming to reduce network latency and call dropping rate. In [7], a game theoretical framework that enables Dynamic Spectrum Access through a utility function that takes into account network measurements is proposed. In [8], authors proposed a coordination protocol to enhance utilization between mobile operators using auctions. The spectrum sharing protocol is based on one-shot games between operators without using operator-specific information exchange. Game-theoretic approaches though, induce significant computational complexity to the network, rely on predictive behavior from MNOs and occasionally assume the knowledge of information that is not possible to be obtained.

All these solutions focus on a single sharing scheme limiting thus the potentials for spectrum sharing. In addition, the game-theoretic approaches either assume cooperation between MNOs or rely on the good-willingness of an MNO, which though is impractical for real systems. On the other hand, the mechanism proposed in this thesis is a flexible solution for optimizing the spectrum acquisition process by exploiting multiple sharing schemes (i.e. co-primary and LSA schemes)[9][10]. Using Fuzzy Logic to design spectrum sharing algorithm under various authorization regimes is a novel approach, and, to our knowledge, none similar solutions that enable flexible spectrum sharing exist in the literature.

The third contribution of this thesis is related to the fact that permanent manual configuration should be avoided in such system. Thus, we have developed a reinforcement learning technique that allows dynamic adaptation of the decision making process of the fuzzy logic system over time. Reinforcement Learning (RL) is based on learning process that maps situations (also known as states) to actions so as to maximize a numerical value named long term reward [11]. In RL the learner is not instructed to take specific actions, as in most forms of machine learning, but instead is free to explore the environment (i.e., moving among states) by taking the actions that yield the most reward. In some RL cases, actions may affect not only the immediate reward, but also the next state as well and, through that, all subsequent rewards. «Trial and error» search and «delayed reward» are the two characteristics that distinguish reinforcement learning from other machine learning schemes [12].

There are three fundamental classes of methods for solving a reinforcement learning problem, namely Dynamic Programming, Monte Carlo methods, and, Temporal Difference methods. Dynamic Programming solutions are well developed mathematically but require a complete and accurate model of the environment, which is not available in many application scenarios. Monte Carlo methods do not require a model and are very simple, but are not suitable for step-by-step incremental computation. Temporal Difference methods on the other hand, do not require an accurate model of the environment and are suitable for step-by-step incremental computations, but are more complex and depend on the dimension of the search space. In general, Temporal

Difference is simpler and possible to work both in online and offline fashion [13], making it thus the most attractive method in our case. The most representative algorithm of Temporal Difference method is Q-Learning that works by estimating the values of state-action pairs. The value Q(s, a) is defined to be the expected discounted sum of future payoffs obtained by taking action <a> from state <s> and following an optimal policy thereafter. Once these values have been learned, the optimal action from any state is the one with the highest Q-value. The main advantage of Q-learning exploited in our solution is that it is able to compare the expected utility of the available actions without requiring a model of the environment. The introduced technique realizes the concept of Adaptive Spectrum Sharing taking into account the effect that Fuzzy Logic Controllers have upon the network as well as the spectrum market. Then the proposed Q-learning scheme tunes the fuzzy logic rules, aiming to find an optimal policy that MNO should follow in order to offer the desirable Quality of Service to its users, while preserving resources (either economical, or radio) when possible. The proposed Adaptive Spectrum Sharing scheme is applicable in sharing scenarios such as Licensed Shared Access, co-primary sharing, etc., that accessing spectrum is granted through spectrum licenses. However, in license-exempt spectrum access scenarios spectrum sharing relies on spectrum sensing and power control mechanisms to avoid harmful interference. The final contribution of this thesis is a mechanism that ensures fair access to spectrum among mobile users in such scenarios. The proposed mechanism caters for underprivileged users by enhancing their transmission power value, generated by the evolutionary execution of Genetic Algorithm. The algorithm's behavior in cases of an incomplete knowledge model (i.e., some of the users may not know all the information) is also assessed as this is particularly important for real systems in which a full knowledge model is typically an unrealistic assumption.

2 Results and Discussion

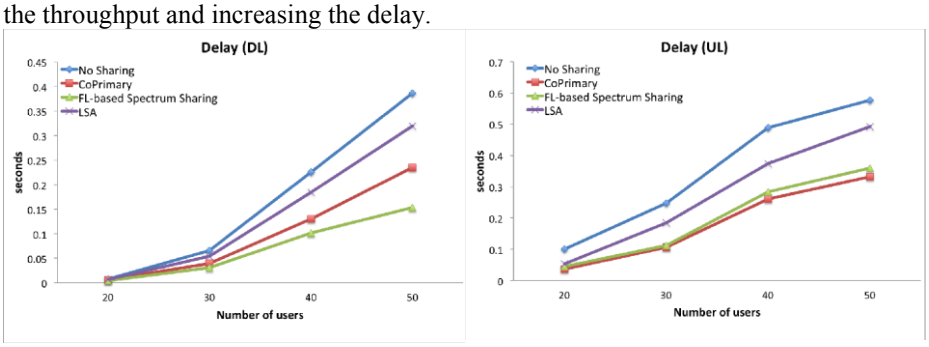
The previously described contributions, for using fuzzy logic controllers and reinforcement learning for Adaptive Spectrum Sharing have been evaluated in the well-known 5G use case, namely shopping mall, firstly introduced by METIS 2020 project in [16]. On the other hand, the introduced Genetic Algorithm for license-exempt spectrum access scenarios is evaluated in small-scale simulation scenario with limited number of users. The aim is to evaluate the mechanisms' efficiency for identifying appropriately events for spectrum sharing, for adapting the model of the environment and for enabling fair resource usage among unlicensed users. In this section, the results of the application of the developed schemes are being provided and analyzed.

2.1 FL-based spectrum sharing

In order to quantify the benefits of the Fuzzy Logic-based spectrum sharing solution we have performed a series of experiments so as to compare its performance against three other schemes namely, no sharing, Co-primary sharing, and LSA sharing based on the available sharing options. All those schemes and our proposed mecha-

nism have been evaluated using the discrete event network simulator NS-3. Our simulation scenario is based on the shopping mall case proposed in METIS project [16]. The considered topology is a 100x50x10 m floor with 10 rooms (that form a 5x2 grid). Three base stations have been deployed in the area; one macro cell located 200 meters away from the building and two femtocells deployed in the considered area.

In the evaluated scenario, UEs that follow a random mobility have been placed in the simulation area and the average delay and throughput both in downlink and uplink communication over a time window of 100 seconds have been measured. The following figures present a comparison between the Fuzzy Logic-based spectrum sharing solution and the other approaches (i.e., no sharing, only co-primary, and only LSA sharing). In all four simulated cases the UEs initiate consuming services. In the three cases where we assume sharing, when the UEs consume a certain portion of the available bandwidth (i.e., 90% - so as to have some resource blocks still available to serve new incoming service requests till the newly acquired spectrum is available, as well as for capturing the nature of the load trend) the spectrum controller is triggered and proceeds in renting spectrum. In the co-primary and LSA cases the controller rents what he is preconfigured to (i.e., co-primary and LSA spectrum respectively), whereas in the FL-based spectrum sharing it rents what the algorithm dictates. When the operator rents spectrum from LSA users, there is the probability that the incumbent user will reclaim his spectrum. In such case the UEs that are being served using LSA spectrum will have to be served by MNO's dedicated resource blocks, thus decreasing the throughput and increasing the delay.



(a) (b)

At this point it should be mentioned that the available economical capacities are the same in all three cases, so the spectrum controller has the same amount of money to consume. Additionally, we assume that the co-primary spectrum has twice the price of the LSA [10][11]. This implies that in the cases of the LSA as well as in the FL-based spectrum sharing the controller may rent extra spectrum, which however, is not guaranteed for the overall time of the sharing. More specifically, in the case of LSA spectrum sharing the operator may acquire twice the co-primary spectrum chunks. Similarly, in the case of the FL-based spectrum sharing if the operator decides to rent only LSA spectrum he may acquire twice the spectrum of the co-primary cases, whereas if he decides to rent only co-primary he may rent exactly as many resource blocks as in the co-primary case. In all the other occasions of the FL-based spectrum sharing scheme the economical capacities are split in the two sharing options.

The results show significant improvement regarding the average delay and throughput when sharing is applied, compared to the no sharing case. Additionally, when comparing the FL-based spectrum sharing to the rest of the sharing schemes we observe that in general the FL-based spectrum sharing and the co-primary sharing perform significantly better than the LSA scheme. This is due to the fact that in the LSA case there is a probability that the incumbent may re-claim his spectrum, thus causing significant delays and throughput reductions. Additionally, for small numbers of UEs the gains from renting spectrum are relative small (since the already available spectrum may cover the user needs), but when the number of UEs increases, the no sharing scheme does not manage to capture the user needs. It is worth mentioning that when the number of UEs increases, the rate of increase in the throughput is reduced even though that the operator rents spectrum, as system's capacity reaches its limitations.
Additionally, it should be highlighted that the developed mechanism outperforms the Co-primary spectrum sharing since it may rent more spectrum when it suits to the users in the vicinity, due to the fact that the controller may split its economic resources to both LSA and co primary spectrum.

2.2 Adaptive Spectrum Sharing through reinforcements

Quantitative analysis is used to evaluate the Reinforcement Learning scheme for adaptive spectrum sharing. In our evaluation in order to be able to compare the scheme against the FL-based scheme we have used the NS-3 simulator to model the behavior of the network. Following similar simulation methodology as in the FL-based scheme evaluation we compare the Adaptive Spectrum Sharing scheme against the four cases (i.e., no-sharing, LSA sharing, Co-primary sharing, Fuzzy-logic based) presented afore in the FL-based scheme evaluation. The Reinforcement Learning scheme applied in the Adaptive Spectrum Sharing mechanism is realized via an offline process that performs training sessions which optimize the behavior of the Fuzzy Reasoners taking into account the reinforcements of their actions (i.e. spectrum cost, monitoring measurements). Then the operator uses the trained Fuzzy Reasoners to make its decisions.

The following methodology has been followed for the Reinforcement learning scheme:

- Initialization: At this point the network topology of the operator is deployed in the simulator. The network topology is based on the afore-mentioned shopping mall. Then, the UEs are placed randomly in the simulation area.
- Monitor-Decision-Execution cycle: During this phase several network parameters are monitored and fed to the Fuzzy Logic Reasoners. Based on the decision making engine the MNO decides whether to obtain additional spectrum resources or not. Finally, the additional spectrum resources are obtained and used in the network, before a new MDE cycle is executed.
- Training sessions: Each experiment is subject to our Q learning mechanism for adapting the decision making process based on reinforcements.

The configuration parameters are the same as in the evaluation process of the Fuzzy Logic based spectrum sharing, so as to have comparable results.

The results reported in thesis show that the Reinforcement Learning improves the average delay and throughput of the FL-based spectrum sharing scheme. The proposed feedback loop optimizes the behavior of the FL-controllers and as the number of UEs increases the benefits of the reinforcement learning are higher. More specifically, the Adaptive Spectrum Sharing scheme demonstrated approximately 20% DL and 30% UL delay reduction compared to FL-based scheme, increasing thus the gains of the proposed spectrum sharing solution against the no-sharing, LSA sharing and the Co-primary sharing schemes. In addition throughput performance of the Adaptive Spectrum Sharing scheme is approximately 5% and 10% increased in DL and UL communication respectively compared to FL-based scheme, and thus is further superior to the no-sharing, LSA sharing and Co-primary sharing schemes.

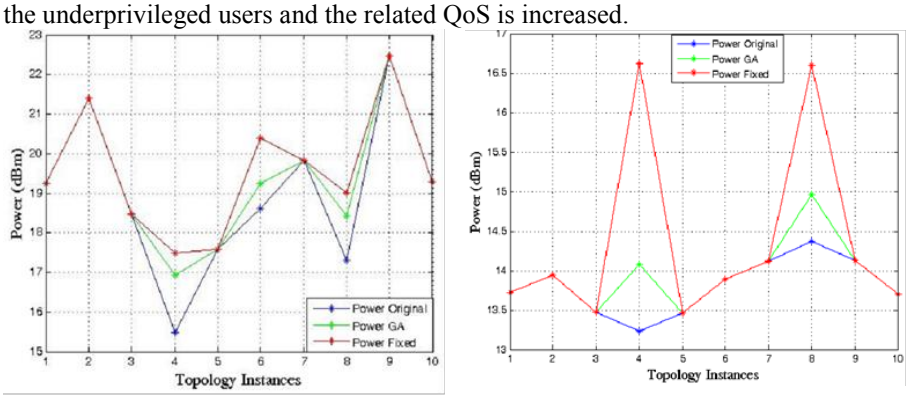
2.3 Fair resource usage through power control in license-exempt scenarios

The proposed Genetic Algorithm for fair resource usage in license-exempt spectrum access scenarios is evaluated in small-scale MATLAB simulation scenario with limited number of users. The proposed algorithm uses a Cooperative Power Control algorithm [17] as a baseline and is compared to a scheme of fixed power value assignment (maximum valid power level). The main objective is to give "fairer" power values to the underprivileged unlicensed users. This concludes to a more "fair" treatment, but incurs loss in system performance, as principles of the baseline algorithm are violated. The major difference between the two proposed techniques is that in case of GA, underprivileged users get better power values, but not the maximum ones due to the negative impact of interference to other users.

The proposed implementation examines a typical network environment with 5 or 10 unlicensed mobile UEs cooperating in order to transmit with an acceptable power value. The Tx power ranges between 10 and 23 dBm and the distances between the unlicensed users is a random number in the [50, 550] meters range. The users set their Tx power levels to maximize the utility function of [17] until the algorithm converges to a steady state for a given topology. The whole procedure lasts for 10 time steps that reflect the mobility of the users in consecutive time frames. For every successive step, our fairness GA-based policy mechanism is triggered, in order to examine whether underprivileged users exist. If so, the GA algorithm is activated, so as to enforce fairness. In order to identify whether an unlicensed user is underprivileged, a time window of previous Tx powers is examined. to detect underprivileged users. The fixed power value schema (FX) lets underprivileged users to transmit with maximum power values usually resulting to a non-cooperative state, where all users are negatively affected. On the other hand, in the proposed fairness scheme the Tx power of the underprivileged users is re-calculated based on the fitness function of the Genetic Algorithm [14].

The results shown in the following figures illustrate the average Tx power values of the 5 and 10 UEs respectively, for each of the 10 time steps and highlight the purpose of the fairness scheme, that is to support the underprivileged users and minimize the negative impact to the network. Indeed, in the proposed scheme the underprivileged users get enhanced power values; however, this is done in a controlled way, so

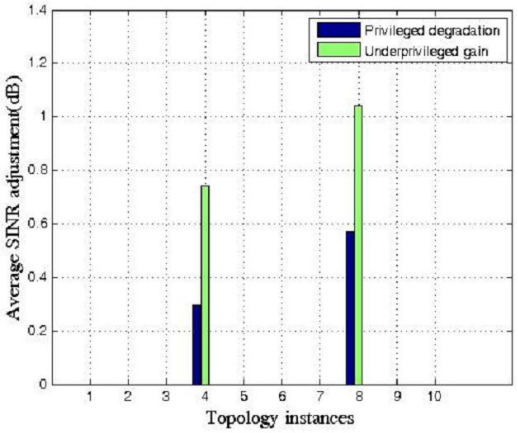
that the impact in the overall performance of the network is limited (marginal reduction of the average network SINR by approximately 0.3 dB). This is a reasonable trade-off for enhancing the overall fairness, especially considering that the SINR of the underprivileged users and the related QoS is increased.



(a) (b)

Furthermore, the proposed solution leads also to enhanced SINR at the receiver for the underprivileged users and the increment of the number of users does not impact the fairness policy. The system remains resilient as more opportunistic users try to transmit and SINR gains remain sufficient [14].

Since the utility function of [17] strikes the optimal balance from a system utilization perspective between the selfish need for transmission at the highest power level and the social conformance of reducing the interference to other neighboring users, altering the Tx Power to the constantly underprivileged users will also have a negative impact to the rest of the users in the environment. Thus, we have measured the consequences of the fairness policy upon the opportunistic users that are not underprivileged compared against the gains of the underprivileged users. The following figure shows a comparative analysis of the average SINR gains of the underprivileged users against the average SINR degradation that the other users will experience. The results show that the gains of the underprivileged users are significantly more compared to the SINR degradation of the rest of the users.



Finally, as mentioned previously many fairness schemes are challenging in their application to real world systems due to the full knowledge requirement and the stringent synchronization constraints among the wireless nodes that this requirement imposes. In our case the genetic algorithm can operate efficiently with a significantly relaxed knowledge model and synchronization scheme. For our evaluation of this highly desirable property we have conducted 1000 experiments assuming the same environment as before; the fundamental difference is that the system suffers a 10-20% message loss, thus leading to undesired effects for the nodes, as they will not have a complete knowledge of the environment. Our results in [14] show that in cases of an incomplete knowledge model the GA (in the scenario of 10 users) is triggered again exactly 2 times (as in the case with full knowledge) with probability equal to 42%. The results also show that cases of not triggering the GA when needed (false negatives) are not possible, but there are some false positive cases where the algorithm is triggered more times than actually needed.

3 Conclusions

Innovative approaches are required for covering the augmented requirements of the future networks to reduce spectrum resources shortage. Considering that other frequency bands remain underutilized due to limited data transmissions of their rightful users the exploitation of such bands becomes very attractive. Up to now, exploitation of TVWS with cognitive radio approaches has been under consideration, though the drawbacks (i.e., complex solutions, interference may be caused to the mobile user, etc.) of such solutions make their realization questionable and discourage their application. On the other hand, the rise of new approaches, such as co-primary spectrum sharing and Licensed Shared Access, which protect both spectrum license holders and spectrum licensees are expected to enable flexible spectrum management.

In this thesis the vision of future mobile networks in which the MNOs share spectrum resources either with other MNOs (co-primary sharing scheme), or with Incumbent Users (LSA sharing scheme) has been thoroughly presented by describing the key characteristics of each approach. The analysis could be summarized in the two main differences between these two spectrum sharing approaches. Both of the differences are related to actors involved in the sharing procedure. The first one is related to the incumbent users that shall not be burdened with complex calculations, which implies that in the LSA case the presence of a translation and coordination entity is required. The second main difference is related to the fact that in the LSA concept potentially several spectrum licensees (i.e., MNOs) may exist, which will not be necessarily coordinated; this may introduce interference among them (the incumbent user is protected from interference), whereas in the co-primary spectrum sharing the spectrum buyer will not experience interference for the time period of the renting. In our work we have presented a common architectural framework for coupling the co-primary and the LSA sharing schemes. For meeting the requirement of reduced complexity in the incumbent users we propose the introduction of a translation engine,

with the prerequisite that the data will be formed in Spectrum Availability structure indicating available spectrum over time, frequency and geographical domains.

The proposed architecture is accompanied by a fuzzy logic based spectrum sharing algorithm for enabling the operators to decide which spectrum authorization option is more suitable given the network conditions. In our analysis, fuzzy reasoners for the LSA and Co-primary sharing schemes have been presented. However, the proposed algorithm could be easily extended to other sharing schemes (e.g., light-licensing). The algorithm has been evaluated against single sharing schemes and also against the typical operation (i.e., without spectrum sharing) of a mobile network in a well-known 5G communication scenario (i.e., shopping mall) related to Ultra Dense Networks. The algorithm influences the decision making process through fuzzy evaluation of the network conditions and the results of the evaluation show significant improvement in achieved throughput and average delay both in uplink and in downlink communication.

However, the way the proposed Fuzzy-Logic based spectrum sharing algorithm evaluated network condition is rather static and thus, it has been further extended with adaptation mechanism (i.e. reinforcement learning technique) to tune the decision making process and realize the concept of Adaptive Spectrum Sharing. Adaptive Spectrum Sharing is the ability of the network to model its environment, assess it and interpret it so as to decide whether spectrum resources (beyond the licensed spectrum of the MNO) will be needed in the near future. Then, evaluate whether the taken decision was beneficiary for the network and tune the decision making process (i.e., adapt the behavior of the Fuzzy Logic Controllers) so as to improve future decisions.

The Adaptive Spectrum Sharing mechanism is then complemented with a fair resource usage mechanism applied in cases of spectrum sharing under general authorization regime. In such scenarios, users are accessing spectrum without having a dedicated license a priori and are obliged to tune their operation in order to avoid harmful interference to the licensed users operating in the same frequency band. This may be done either through spectrum sensing techniques, or via querying a GeoLocation Database before accessing spectrum. However, mutual interference among unlicensed users in such scenarios is not part of any regulatory process and thus, mechanisms that allow fair resource usage are needed. In this thesis we proposed a fair power control mechanism using Genetic Algorithms. The proposed solution has been applied upon a cooperative power control algorithm and the results showed significantly improved SINR for the underprivileged users compared to the original algorithm with minimal impact in the SINR of the privileged users. Furthermore, in comparison to the case of a simplified fairness policy, which assigns underprivileged cognitive users with the maximum valid power level, the proposed scheme offers considerable power gains to the network. Finally, it has been shown that the proposed algorithm can operate efficiently even in cases of partial knowledge models and imperfect message exchange/synchronization between the users, a property that is highly desirable for application in real world system.

4 References

1. K. Chatzikokolakis, P. Spapis, A. Kaloxylos, N. Alonistioti, “Towards spectrum sharing: opportunities and technical enablers”. IEEE Communication Magazine, vol. 53, no. 7,pp.26-33, July 2015.
2. Spapis, P., Chatzikokolakis, K., Alonistioti, N., & Kaloxylos, A. (2014, July). Using sdn as a key enabler for co-primary spectrum sharing. In Information, Intelligence, Systems and Applications, IISA 2014, The 5th International Conference on (pp. 366-371). IEEE.
3. J. Khun-Jush, P. Bender, B. Deschamps, and M. Gundlach, “Licensed shared access as complementary approach to meet spectrum demands: Benefits for next generation cellular systems,” in ETSI Workshop Reconfig. Radio Syst., Cannes, France, Dec. 2012.
4. W.-Y. Lee and I. Akyldiz, “A Spectrum Decision Framework for Cognitive Radio Networks,” IEEE Trans. Mobile Computing, vol. 10, no. 2, pp. 161-174, Feb. 2011.
5. Cao, L., Zheng. H., "Distributed Rule-Regulated Spectrum Sharing," Selected Areas in Communications, IEEE Journal on, vol. 26, no. 1, pp. 130-145, 2008.
6. Y.-T. Lin, H. Tembine, and K.-C. Chen, “Inter-operator spectrum sharing in future cellular systems,” in Proc. IEEE GLOBECOM, Dec. 2012, pp. 2597–2602
7. H. Kamal, M. Coupechoux, P. Godlewski, “Inter-operator spectrum sharing for cellular networks using game theory,” IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2009 , pp.425,429, 13-16 Sept. 2009
8. B. Singh, K. Koufos, O. Tirkkonen, “Co-primary inter-operator spectrum sharing using repeated games,” IEEE International Conference on Communication Systems (ICCS), 2014, pp. 67-71, 19-21 Nov. 2014.
9. Chatzikokolakis, K., Beinas, G., Alonistioti, N., Spapis, P., & Kaloxylos, A. (2015, July). Spectrum sharing: A coordination framework enabled by fuzzy logic. In Computer, Information and Telecommunication Systems (CITS), 2015 International Conference on (pp. 1-5). IEEE.
10. Chatzikokolakis K. ; Spapis P.; Kaloxylos A.; Beinas G.; Alonistioti N.; “Fuzzy-logic enabled spectrum sharing for 5G mobile networks,” to appear in Journal of Networks, 2016
11. K. Chatzikokolakis, P. Spapis, A. Kaloxylos, E. Kiagias, N. Alonistioti, “Adaptive Spectrum Sharing through reinforcements”, submitted at Journal on Selected Areas of Communications (JSAC), 2016
12. R. S. Sutton and A. G. Barto, “Reinforcement Learning: An Introduction”, MIT Press, Cambridge, MA, 1998.
13. E. V. Denardo, ”Dynamic Programming: Models and Applications”, Mineola, NY, 2003.
14. K. Chatzikokolakis, R. Arapoglou, A. Merentitis, N. Alonistioti, “Fair Power Control in Cooperative Systems Based on Evolutionary Techniques”, In the proceedings of Mobile Ubiquitous Computing, Systems, Services and Technologies UBICOMM 23-28 September, Barcelona, Spain, 2012
15. Chatzikokolakis, K., Spapis, P., Stamatelatos, M., Katsikas, G., Arapoglou, R., Kaloxylos, A., & Alonistioti, N. (2013). Spectrum Aggregation in Cognitive Radio Access Networks from Power Control Perspective. Evolution of Cognitive Networks and Self-Adaptive Communication Systems, 105.
16. M. Fallgren and B. Timus (editors), “Future radio access scenarios, requirements and KPIs,” METIS deliverable D1.1, March 2013. Available: https://www.metis2020.com/documents/deliverables/
17. Merentitis, A., & Triantafyllopoulou, D., (2010). Transmission Power Regulation in Cooperative Cognitive Radio Systems Under Uncertainties. IEEE International Symposium on Wireless Pervasive Computing (pp.134-139).

Byzantine fault-tolerant vote collection for D-DEMOS, a distributed e-voting system

Nikos Chondros*

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
n.chondros@di.uoa.gr

Abstract. E-voting systems are a powerful technology for improving democracy by reducing election cost, increasing voter participation, and even allowing voters to directly verify the entire election procedure. Unfortunately, prior internet voting systems have single points of failure, which may result in the compromise of availability, voter secrecy, or integrity of the election results.
In this thesis, we consider increasing the fault-tolerance of voting systems by introducing distributed components. This is non-trivial as, besides integrity and availability, voting requires safeguarding confidentiality as well, against a malicious adversary. We focus on the vote collection phase of the voting system, which is a crucial part of the election process.
We use the DEMOS state-of-the-art but centralized voting system as the basis for our study. This system uses vote codes to represent voters' choices, an Election Authority to setup the election and handle vote collection and result production, and a Bulletin Board for storing the election transcript for the long-term. We extract the vote collection mechanism from the centralized Election Authority component of the original DEMOS system, and replace it with a distributed system that handles vote collection in a Byzantine fault-tolerant manner. In this thesis, we present the design, security analysis, prototype implementation and experimental evaluation of this vote collection component.
We present two versions of this component: one completely asynchronous and one with minimal timing assumptions but better performance. Both versions provide immediate assurance to the voter her vote was recorded as cast, without requiring cryptographic operations on behalf of the voter, while still preserving privacy. This way, a voter may cast her vote using an untrusted computer or network, and still be assured her vote was recorded as cast. For example, she may vote via a public web terminal, or by sending an SMS from a mobile phone.

1 Dissertation Summary

1.1 Problem Description

E-voting systems are a powerful technology to improve the election process. Kiosk-based e-voting systems allow the tally to be produced faster, but require

* Dissertation Advisor: Mema Roussopoulos, Associate Professor

the voter's physical presence at the booth. Internet e-voting systems, however, allow voters to cast their votes remotely. Internet voting systems have the potential to enhance the democratic process by reducing election costs and by increasing voter participation for social groups that face considerable physical barriers and overseas voters. In addition, several internet voting systems allow voters and auditors to directly verify the integrity of the entire election process, providing *end-to-end verifiability*. This is a highly desired property that has emerged in the last decade, where voters can be assured that no entities, even the election authorities, have manipulated the election result. Despite their potential, existing internet voting systems suffer from single points of failure, which may result in the compromise of voter secrecy, service availability, or integrity of the result.

In this thesis, we consider increasing the fault-tolerance of voting systems by introducing distributed components, while still preserving privacy and end-to-end verifiability. We use the DEMOS [13] state-of-the-art but centralized voting system as the basis for our study.

In its current form, the DEMOS voting system is centralized, having an Election Authority (EA) component that handles everything from setup, to vote collection, to result production. This presents a risk to availability, as a failure of this component would prohibit voting. However, it also presents a risk to voters' privacy, as an attacker that takes control of this component can obtain each voter's ballot contents, which directly violates the voter's privacy. Finally, the original centralized DEMOS system had no need to provide feedback to the voter, besides a simple acknowledgment. In a distributed world though, the voter needs to obtain feedback to be assured the vote was actually recorded as cast in enough nodes of the system, something we tackle in this thesis.

One specific attribute of DEMOS is its use of code-voting. In this scheme, there is a setup component which generates vote codes representing the possible voter's choices, and includes them in the voters' ballots. A voter votes by submitting the vote code corresponding to her choice. Because of this technique, the voter does not need to perform cryptographic operations on the device she uses to vote. Expanding on this, we set out to introduce a distributed voting system that uses no client-side cryptography at all. This allows votes to be cast with a greater variety of client devices over public networks, such as feature phones using SMS, or (untrusted) public web terminals, while still preserving voter's privacy.

1.2 Related work

Several end-to-end verifiable e-voting systems have been introduced, e.g. the kiosk-based systems [4, 12, 3, 2, 16] and the internet voting systems [1, 14, 17, 13]. In all these works, the Bulletin Board (BB) is a single point of failure and has to be trusted.

Dini presents a distributed e-voting system, which however is not end-to-end verifiable [11]. In [9], there is a distributed BB implementation, also handling vote collection, according to the design of the vVote end-to-end verifiable e-voting system [8], which in turn is an adaptation of the Prêt à Voter e-voting system [4].

In [9], the proper operation of the *BB* during ballot casting requires a trusted device for signature verification. In contrast, our vote collection subsystem is done so that correct execution of ballot casting can be "human verifiable", i.e., by simply checking the validity of the obtained receipt. Additionally, our vote collection subsystem in D-DEMOS/Async is fully asynchronous, always deciding with exactly $n - f$ inputs, while in [9], the system uses a synchronous approach based on the FloodSet algorithm from [15] to agree on a single version of the state.

1.3 Results

In this thesis, we present the design, security analysis, prototype implementation and experimental evaluation of the vote collection components of the *D-DEMOS* [7] suite of distributed, end-to-end verifiable internet voting systems, with no single point of failure during the election process (that is, besides setup).

We design a distributed *Vote Collection* (*VC*) subsystem that is Byzantine fault-tolerant and able to collect votes from voters and assure them their vote was recorded as cast, without requiring any cryptographic operation from the client device. At election end time, *VC* nodes agree on a single set of votes.

We introduce two versions of the voting components of D-DEMOS that differ in how they achieve agreement on the set of cast votes. The D-DEMOS/Async version is completely asynchronous, while D-DEMOS/IC makes minimal synchrony assumptions but is more efficient. Once agreement has been achieved, *VC* nodes upload the set of cast votes to a second distributed component, the *Bulletin Board* (*BB*). This, in turn, is a replicated service that publishes its data immediately and makes it available to the public forever.

The resulting voting systems are end-to-end verifiable, by the voters themselves and third-party auditors, while preserving voter privacy. To delegate auditing, a voter provides an auditor specific information from her ballot. The auditor, in turn, reads from the distributed *BB* and verifies the complete election process, including the correctness of the election setup by election authorities. Additionally, as the number of auditors increases, the probability of election fraud going undetected diminishes exponentially.

The thesis is structured as follows. Section 1 provides an introduction to the problem, and gives a short description of DEMOS, the system we use as a model and we extend to become fault-tolerant. Section 2 gives background information on voting systems (using information from [5], and tools from distributed systems and cryptography that we employ in our system designs.

Section 3 provides a thorough system description of the two systems we build. It first gives the system model, and the gradually introduces the Vote Collection subsystems we introduce, a mostly-asynchronous one with a single timing assumption (for D-DEMOS/IC) and a completely asynchronous design (for D-DEMOS/Async). It also describes how the systems are initialized by the EA component, and how the voter uses our system to vote. This section also includes the proofs of liveness and safety for both vote collection approaches. In both approaches to vote collection, we design a *voting protocol* that is active

during voting hours and collects votes from the voters, and a *vote set consensus* protocol that ensures agreement between vote collection nodes after voting is finished, and allows the system to progress towards producing the election result.

Section 4 provides answers to questions regarding our system design. First of all, it answers why standard approaches, like a Byzantine Fault Tolerant Replicated State Machine (such as [6]), is not suitable for the problem at hand. It then lists a series of possible attack vectors from different system components, and describes how our system thwarts them.

Section 5 outlines our prototype implementation. It describes our message-passing substrate and its interaction with our Web front-ends, and our implementation of the EA and both versions of the Vote Collection subsystems. It also describes our implementation of Bracha's asynchronous binary consensus, that is used in the asynchronous version of the system.

Section 6 describes the evaluation and presents our experimental results. We perform experiments with a relational database as a data store, and also with an in-memory data storage approach. We perform experiments on a LAN, and we also simulate a WAN. The outcome of the evaluation is that the D-DEMOS/IC vote collection approach is slightly faster during voting (around 15%), and quite faster during vote set consensus (4 times faster). for the disk-based experiments.

2 Vote collection for D-DEMOS

We will now briefly describe the design of our Vote Collection (*VC*) subsystem.

We design the *VC* subsystem as a distributed system of N_v cooperating nodes, tolerating up to f_v Byzantine faults, where $f_v < N_v/3$. Note that, we also tolerate the collusion of an arbitrary number of malicious voters with the malicious *VC* nodes. *VC* nodes have private communication channels to each other, and a public (unsecured) channel for the voters.

We modify the data generation process of DEMOS's *EA*, by adding the following two steps while generating voter's ballots:

1. The (random) vote-code corresponding to each election option is provided in committed form to each *VC* node.
2. A receipt is generated for each vote code, which is itself a random number. The receipt is secret shared across *VC* nodes with a Verifiable Secret Sharing (VSS) scheme. Each *VC* node receives one of these shares.

At step 1, the commitment scheme used hashes the plain text message along with a salt. The salt is provided along with the committed form to each *VC* node, while the opening of the commitment is the vote-code itself.

Before going into detail in the design of the Vote Collection subsystem, we give an overview of its use. *VC* nodes are initialized from the *EA* (as above). The voter receives her ballot also from the *EA*, along with the addresses of the *VC* nodes. During the election hours, *VC* nodes run the *voting protocol*.

For this protocol to start, the voter selects one part of her ballot at random, and posts her selected vote code to one of the *VC* nodes. The *VC* node that

receives her vote validates it, interacts with the other *VC* nodes to reconstruct the receipt from the shares spread across the *VC* nodes, and posts it back to the voter. When she receives a receipt, she compares it with the one on her ballot corresponding to the selected vote code. If it matches, she is assured her vote was correctly recorded and will be included in the election tally. The other part of her ballot, the one not used for voting, will be used for auditing purposes. This design is essential for verifiability, in the sense that the *EA* cannot predict which part a voter may use, and the unused part will betray a malicious *EA* with $\frac{1}{2}$ probability per audited ballot.

At election end time, *VC* nodes run our Vote Set Consensus protocol, which guarantees all *VC* nodes agree on a single set of voted vote codes. After agreement, each *VC* node uploads this set to every *BB* node, which in turn publishes this set once it receives the same copy from enough $(f_v + 1)$ *VC* nodes.

2.1 Synchronous version, for D-DEMOS/IC

The *voting* protocol starts when a voter submits a VOTE⟨serial-no, vote-code⟩ message to a *VC* node. We call this node the *responder*, as it is responsible for delivering the receipt to the voter. The *VC* node confirms the current system time is within the defined election hours, and locates the ballot with the specified serial-no. It also verifies this ballot has not been used for this election, either with the same or a different vote code. Then, it compares the vote-code against every hashed vote code in each ballot line, until it locates the correct entry. Subsequently, it obtains from its local database the receipt-share corresponding to the specific vote-code. Next, it marks the ballot as pending for the specific vote-code. Finally, it multicasts a VOTE_P⟨serial-no, vote-code, receipt-share⟩ message to all *VC* nodes, disclosing its share of the receipt. In case the located ballot is marked as voted for the specific vote-code, the *VC* node sends the stored receipt to the voter without any further interaction with other *VC* nodes.

Each *VC* node that receives a VOTE_P message, first validates the received receipt-share according to the verifiable secret sharing scheme used. Then, it performs the same validations as the responder, and multicasts another VOTE_P message (only once), disclosing its share of the receipt. When a node collects $h_v = N_v - f_v$ valid shares, it uses the verifiable secret sharing reconstruction algorithm to reconstruct the receipt (the secret) and marks the ballot as voted for the specific vote-code. Additionally, the *responder* node sends this receipt back to the voter.

A message flow diagram of our *voting* protocol is depicted in Figure 1. As is evident from the diagram, the time from the multicast of the first VOTE_P message until collecting all receipt shares, is only slightly longer than a single round-trip between two *VC* nodes.

At election end time, each *VC* node stops processing VOTE and VOTE_P messages, and initiates the *vote-set consensus* protocol. It creates a set VS_i of ⟨serial-no, vote-code⟩ tuples, including all *voted* and *pending* ballots. Then, it participates in the Interactive Consistency (IC) protocol of [10], with this set. At the end of IC, each node contains a vector $\langle VS_1, \ldots, VS_n \rangle$ with the Vote

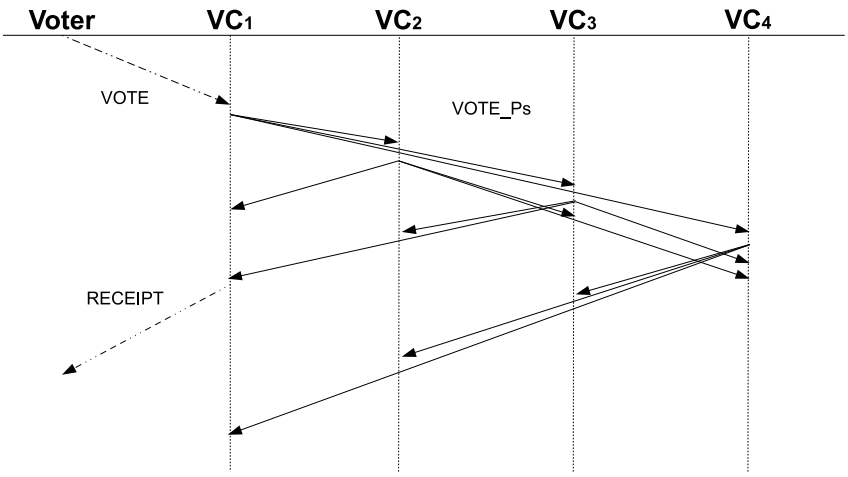


Fig. 1. Diagram of message exchanges for a single vote during the D-DEMOS/IC vote collection phase.

Set of each node, and follows the algorithm of Figure 2. Step 1 makes sure

Cross-tabulate $\langle VS_1, \ldots, VS_n \rangle$ per ballot, creating a list of vote codes for each ballot. Perform the following actions for each ballot:

1. If the list contains two or more distinct vote codes, mark the ballot as NotVoted and exit.
2. If a vote code vc_a appears at least $N_v - 2f_v$ times in the list, mark the ballot as Voted for vc_a and exit.
3. Otherwise, mark the ballot as NotVoted and exit.

Fig. 2. High level description of algorithm after IC.

any ballot with multiple submitted vote codes is discarded. Since vote codes are private, and cannot be guessed by malicious vote collectors, the only way for multiple vote codes to appear is if malicious voters are involved, against whom our system is not obliged to respect our *contract*.

With a single vote code remaining, step 2 considers the threshold above which to consider a ballot as voted for a specific vote code. We select the $N_v - 2f_v$ threshold for which we are certain that even the following extreme scenario is handled. If the *responder* is malicious, submits a receipt to an honest voter, but denies it during *vote-set consensus*, the remaining $N_v - 2f_v$ honest *VC* nodes that revealed their receipt shares for the generation of the receipt, are enough for the system to accept the vote code (receipt generation requires $N_v - f_v$ nodes, of which f_v may be malicious, thus $N_v - 2f_v$ are necessarily honest).

Finally, step 3 makes sure vote codes that occur less than $N_v - 2f_v$ times are discarded. Under this threshold, there is no way a receipt was ever generated.

At the end of this algorithm, each node submits the resulting set of *voted* $\langle$serial-no, vote-code$\rangle$ tuples to each *BB* node, which concludes its operation for the specific election.

2.2 Asynchronous version, for D-DEMOS/Async

We make the following enhancements to the Vote Collection subsystem, to achieve the completely asynchronous version *D-DEMOS/Async*. During voting we introduce another step, which guarantees only a single vote code can be accepted (towards producing a receipt) for a given ballot. We also employ an asynchronous binary consensus primitive to achieve Vote Set Consensus.

More specifically, during voting, the *responder VC* node validates the submitted vote code, but before disclosing its receipt share, it multicasts an ENDORSE$\langle$serial-no, vote-code$\rangle$ message to all *VC* nodes. Each *VC* node, after making sure it has not endorsed another vote code for this ballot, responds with an ENDORSEMENT$\langle$serial-no, vote-code, $\mathsf{sig}_{\mathsf{VC_i}}\rangle$ message, where $\mathsf{sig}_{\mathsf{VC_i}}$ is a digital signature of the specific serial-no and vote-code, with VC_i's private key. The responder collects $N_v - f_v$ valid signatures and forms a uniqueness certificate UCERT for this ballot. It then discloses its receipt share via the VOTE_P message, but also includes the formed UCERT in the message.

Each *VC* node that receives a VOTE_P message, first verifies the validity of UCERT and discards the message on error. On success, it proceeds as per the *D-DEMOS/IC* protocol (validating the receipt share it receives and then disclosing its own receipt share).

The voting process is outlined in the diagram of Figure 3, where we now see two round-trips are needed before the receipt is reconstructed and posted to the voter.

The formation of a valid UCERT gives our algorithms the following guarantees:

a) No matter how many responders and vote codes are active at the same time for the same ballot, if a UCERT is formed for vote code vc_a, no other uniqueness certificate for any vote code different than vc_a can be formed.
b) By verifying the UCERT before disclosing a *VC* node's receipt share, we guarantee the voter's receipt cannot be reconstructed unless a valid UCERT is present.

At election end time, each *VC* node stops processing ENDORSE, ENDORSEMENT, VOTE and VOTE_P messages, and follows the *vote-set consensus* algorithm in Figure 4, for each registered ballot.

Steps 1-2 ensure used vote codes are dispersed across nodes. Recall our receipt generation requires $N_v - f_v$ shares to be revealed by distinct *VC* nodes, of which at least $N_v - 2f_v$ are honest. Note that any two $N_v - f_v$ subsets of N_v contain at least $f_v + 1$ honest nodes (because $f_v > N_v/3$), and at least one of the $f_v + 1$

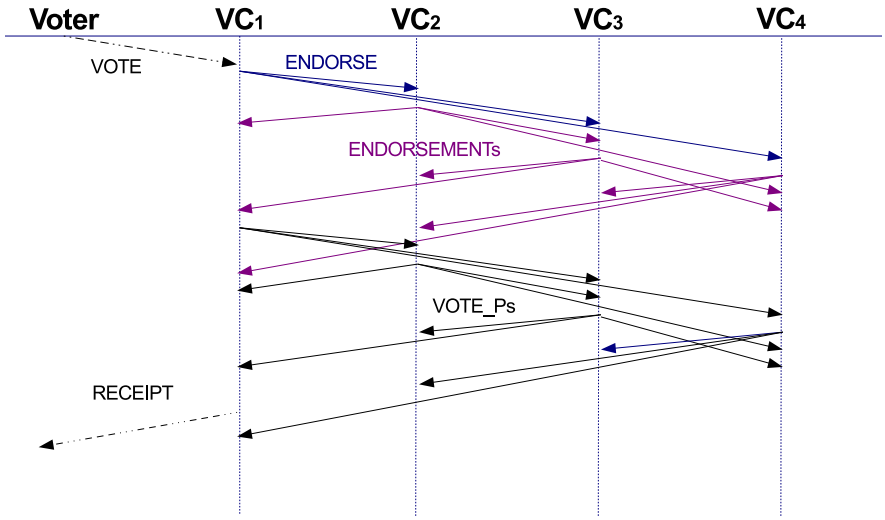


Fig. 3. Diagram of message exchanges for a single vote during the D-DEMOS/Async vote collection phase.

1. Send `ANNOUNCE`⟨serial-no, vote-code, UCERT⟩ to all nodes. The vote-code will be *null* if the node knows of no vote code for this ballot.
2. Wait for $N_v - f_v$ such messages. If any of these messages contains a valid vote code vc_a, accompanied by a valid UCERT, change the local state immediately, by setting vc_a as the vote code used for this ballot.
3. Participate in a Binary Consensus protocol, with the subject "Is there a valid vote code for this ballot?". Enter with an opinion of 1, if a valid vote code is locally known, or a 0 otherwise.
4. If the result of Binary Consensus is 0, consider the ballot not voted.
5. Else, if the result of Binary Consensus is 1, consider the ballot voted. There are two sub-cases here:
   a) If vote code vc_a, accompanied by a valid UCERT is locally known, consider the ballot voted for vc_a.
   b) If, however, vc_a is not known, send a `RECOVER-REQUEST`⟨serial-no⟩ message to all VC nodes, wait for the first valid `RECOVER-RESPONSE`⟨serial-no, vc_a, UCERT⟩ response, and update the local state accordingly.

Fig. 4. High level description of algorithm for asynchronous vote set consensus. This algorithm runs for each registered ballot.

honest nodes has participated in receipt generation. Because of this, if a receipt was generated, at least one honest node's `ANNOUNCE` will be processed by every honest node, and all honest VC nodes will obtain the corresponding vote code in these two steps. Consequently, all honest nodes enter step 3 with an opinion of 1 and binary consensus is guaranteed to deliver 1 as the resulting value, thus

safeguarding our contract against the voters. In any case, step 3 guarantees all *VC* nodes arrive at the same conclusion, on whether this ballot is voted or not.

In the algorithm outlined above, the result from binary consensus is translated from 0/1 to a status of "not-voted" or a unique valid vote code, in steps 4-5. Step 5b requires additional explanation. Assume, for example, that a voter submitted a valid vote code vc_a, but a receipt was not generated before election end time. In this case, an honest vote collector node VC_i may not be aware of vc_a at step 3, as steps 1-2 do not make any guarantees in this case. Thus, VC_i may rightfully enter consensus with a value of 0. However, when honest nodes' opinions are mixed, the consensus algorithm may produce either 0 or 1. In case the result is 1, VC_i will not possess the correct vote code vc_a, and thus will not be able to properly translate the result. Thus we introduce a recovery protocol with which VC_i will issue a `RECOVER-REQUEST` multicast. We claim that another honest node, VC_h, exists that *possesses* vc_a and *replies* with vc_a and the correct UCERT. The reason for the existence of an honest VC_h is straightforward and stems from the properties of the binary consensus problem definition. If all honest nodes enter binary consensus with the same opinion a, the result of any consensus algorithm is guaranteed to be a. Since we have an honest node VC_i, that entered consensus with a value of 0, but a result of 1 was produced, there has to exist another honest node VC_h that entered consensus with an opinion of 1. Since VC_h is honest, it must *possess* vc_a, along with the corresponding UCERT (as no other vote code vc_b can be active at the same time for this ballot). Again, because VC_h is honest, it will follow the protocol and *reply* with a well formed RECOVER-REPLY. Additionally, the existence of UCERT guarantees that any malicious replies can be safely identified and discarded by VC_i.

As per *D-DEMOS/IC*, at the end of this algorithm, each node submits the resulting set of *voted* $\langle \text{serial-no}, \text{vote-code} \rangle$ tuples to each *BB* node, which concludes its operation for the specific election.

2.3 Evaluation

We will now outline some notable results from our evaluation. In Figure 5, we plot the average response time of both our vote collection protocols, versus the number of vote collectors, under different concurrency levels, ranging from 500 to 2000 concurrent clients. The experiment is run with our in-memory data structure, highlighting the performance of our network protocols. Results for both systems illustrate an almost linear increase in the client-perceived latency, for all concurrency scenarios, up to 13 *VC* nodes. D-DEMOS/IC has a slower response time with its single round intra-*VC* node communication, while D-DEMOS/Async is slightly slower due to the extra Uniqueness Certificate round.

Figure 6 shows the throughput of both our vote collection protocols, versus the number of vote collectors, under different concurrency levels, for the same experiment. We observe that, in terms of overall system throughput, the penalty of tolerating extra failures (increasing the number of vote collectors) manifests early on. We notice an almost 50% decline in system throughput from 4 to 7 *VC* nodes

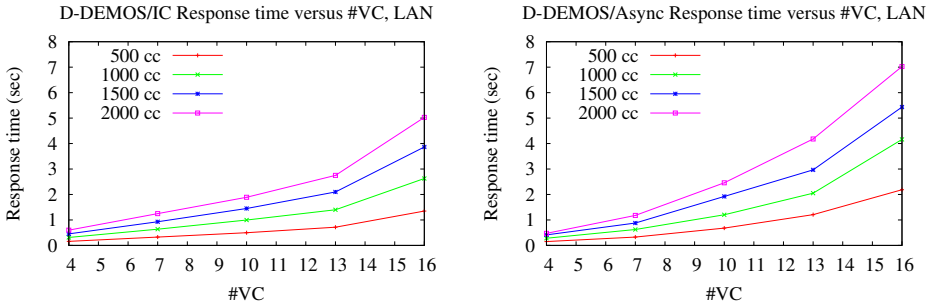


Fig. 5. Vote Collection response time of D-DEMOS/IC (5a) and D-DEMOS/Async (5b), versus the number of *VC* nodes, under a LAN setting. Election parameters are $n = 200{,}000$ and $m = 4$.

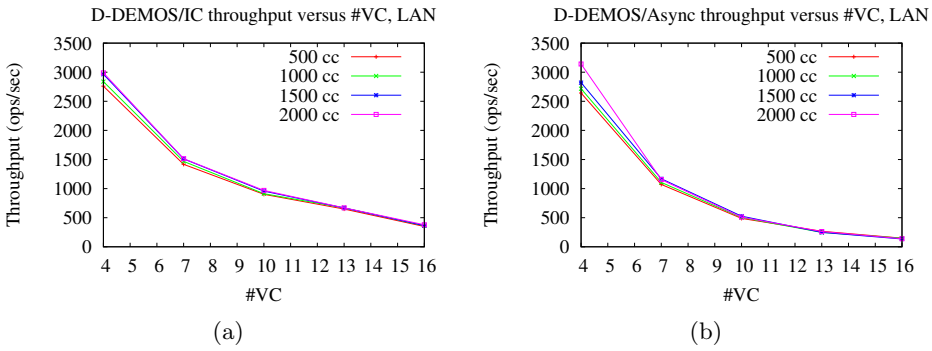


Fig. 6. Vote Collection throughput of D-DEMOS/IC (6a) and D-DEMOS/Async (6b), versus the number of *VC* nodes, under a LAN setting. Election parameters are $n = 200{,}000$ and $m = 4$.

for D-DEMOS/IC, and a bigger one for D-DEMOS/Async. However, further increases in the number of vote collectors lead to a much smoother, linear decrease. Overall, D-DEMOS/IC achieves better throughput than D-DEMOS/Async, due to exchanging fewer messages and lacking signature operations.

3 Conclusion

In this thesis, we presented two different vote collection subsystems for the D-DEMOS suite of distributed vote collection systems. Both resultant voting systems allow voters to verify their vote was tallied-as-intended without the assistance of special software or trusted devices, while maintaining the end-to-end verifiability required for external auditors to verify the correctness of the complete election process. We proved the safety and liveness of both vote collection subsystems, produced prototypes implementing them, measured their performance, and demonstrated their ability to handle large-scale elections.

We believe our vote collection subsystems are applicable to any voting system that uses the code-voting technique. Thus, we believe our work is a required step towards producing higher quality voting systems that can handle large-scale elections efficiently and reliably.

References

1. Adida, B.: Helios: Web-based open-audit voting. In: Proceedings of the 17th USENIX Security Symposium, San Jose, CA, USA. pp. 335–348. USENIX Association (July 2008)
2. Benaloh, J., Byrne, M.D., Eakin, B., Kortum, P.T., McBurnett, N., Pereira, O., Stark, P.B., Wallach, D.S., Fisher, G., Montoya, J., Parker, M., Winn, M.: STAR-vote: A secure, transparent, auditable, and reliable voting system. In: Proceedings of the Electronic Voting Technology Workshop / Workshop on Trustworthy Elections, EVT/WOTE '13, Washington, D.C., USA. USENIX Association (August 2013)
3. Chaum, D., Essex, A., Carback, R., Clark, J., Popoveniuc, S., Sherman, A., Vora, P.: Scantegrity: End-to-end voter-verifiable optical-scan voting. Security & Privacy, IEEE 6(3), 40–46 (2008)
4. Chaum, D., Ryan, P.Y.A., Schneider, S.A.: A practical voter-verifiable election scheme. In: Proceedings of the 10th European Symposium on Research in Computer Security - ESORICS 2005, Milan, Italy. pp. 118–139. Springer (September 2005)
5. Chondros, N., Delis, A., Gavatha, D., Kiayias, A., Koutalakis, C., Nicolacopoulos, I., Paschos, L., Roussopoulos, M., Sotirelis, G., Stathopoulos, P., Vasilopoulos, P., Zacharias, T., Zhang, B., Zygoulis, F.: Electronic voting systems - from theory to implementation. In: E-Democracy, Security, Privacy and Trust in a Digital World. pp. 113–122 (Dec 2013)
6. Chondros, N., Kokordelis, K., Roussopoulos, M.: On the practicality of practical byzantine fault tolerance. In: Proceedings of the ACM/IFIP/USENIX 13th International Middleware Conference (Middleware 2012), Montreal, QC, Canada. pp. 436–455. Springer (December 2012)

7. Chondros, N., Zhang, B., Zacharias, T., Diamantopoulos, P., Maneas, S., Patsonakis, C., Delis, A., Kiayias, A., Roussopoulos, M.: D-demos: A distributed, end-to-end verifiable, internet voting system. In: Distributed Computing Systems (ICDCS), 2016 IEEE 36th International Conference on (Jun 2016)
8. Culnane, C., Ryan, P.Y.A., Schneider, S., Teague, V.: vvote: A verifiable voting system. ACM Transactions on Information and System Security 18(1), 3:1–3:30 (Jun 2015), http://doi.acm.org/10.1145/2746338
9. Culnane, C., Schneider, S.: A peered bulletin board for robust use in verifiable voting systems. In: Proceedings of the IEEE 27th Computer Security Foundations Symposium (CSF 2014), Vienna, Austria. pp. 169–183. IEEE (July 2014)
10. Diamantopoulos, P., Maneas, S., Patsonakis, C., Chondros, N., Roussopoulos, M.: Interactive consistency in practical, mostly-asynchronous systems. In: Proceedings of the IEEE 21st International Conference on Parallel and Distributed Systems (ICPADS 2015). pp. 752–759. IEEE (Dec 2015)
11. Dini, G.: A secure and available electronic voting service for a large-scale distributed system. Future Generation Computer Systems 19(1), 69–85 (2003)
12. Fisher, K., Carback, R., Sherman, A.: Punchscan: introduction and system definition of a high-integrity election system. In: IAVoSS Workshop On Trustworthy Elections (WOTE 2006), Cambridge, United Kingdom (June 2006)
13. Kiayias, A., Zacharias, T., Zhang, B.: End-to-end verifiable elections in the standard model. In: Proceedings of the 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Advances in Cryptology - EUROCRYPT 2015, Sofia, Bulgaria. pp. 468–498. Springer (April 2015)
14. Kutylowski, M., Zagórski, F.: Scratch, click & vote: E2E voting over the internet. In: Towards Trustworthy Elections, New Directions in Electronic Voting, Lecture Notes in Computer Science, vol. 6000, pp. 343–356. Springer (2010)
15. Lynch, N.: Distributed Algorithms. Morgan Kaufmann (1996)
16. Moran, T., Naor, M.: Split-ballot voting: Everlasting privacy with distributed trust. ACM Transactions on Information and System Security 13(2), 16:1–16:43 (Mar 2010), http://doi.acm.org/10.1145/1698750.1698756
17. Zagórski, F., Carback, R.T., Chaum, D., Clark, J., Essex, A., Vora, P.L.: Remotegrity: Design and use of an end-to-end verifiable remote voting system. In: Proceedings of the 11th International Conference on Applied Cryptography and Network Security, ACNS 2013, Banff, AB, Canada. pp. 441–457. Springer (June 2013)

Facial Expression Retrieval Using 3-Dimensional Mesh Sequences

Danelakis E. Antonios*

National and Kapodistrian University of Athens

Department of Informatics and Telecommunications

adanelakis@di.uoa.gr

Abstract. Human emotions are often expressed by facial expressions and are generated by facial muscle movements. In recent years, analysis of facial expressions has emerged as an active research area due to its various applications such as human-computer interaction, human behavior understanding, biometrics, emotion recognition, computer graphics, driver fatigue detection, and psychology. This dissertation introduces a new scheme for dynamic 3D facial expression retrieval. The new scheme employs novel descriptors which exploit facial mesh sequence information of automatically detected facial landmarks. A detailed evaluation of the new retrieval scheme is presented. Experiments have been conducted using the publicly available *BU−4DFE* and *BP4D−Spontaneous* datasets. The obtained results outperform the retrieval results of the state-of-the-art methodologies. Furthermore, the retrieval results are exploited in order to achieve *unsupervised* dynamic *3D* facial expression recognition. The aforementioned *unsupervised* procedure achieves better recognition accuracy compared to *supervised* dynamic *3D* facial expression recognition state-of-the-art techniques. Finally, we present a methodology for detecting primitive facial movements. The obtained results are mostly better than the state-of-the-art and more movements are detected.

1. Introduction

The process of extracting useful content information from large amounts of data, in an automated manner and based on an example or descriptive query, is called *content based information retrieval*. Common types of information that can benefit from such a retrieval process are: textual, visual, audio and video data and most recently, *3D* and *4D* (*3D* over time) data; the latter is also referred to as dynamic *3D* data or *3D* videos.

In recent years, through the creation of inexpensive *3D* scanners and the simplification of *3D* modelling software, a large volume of *3D* and *4D* data has been created. Some of the *4D* datasets that have recently been created involve human facial expressions. These datasets contain *3D* mesh sequences representing people of different ethnicities taking on a number of facial expressions. The creation of the aforementioned datasets gave rise to two new problems for the research community: The problem of *Facial Expression Recognition from 3D mesh sequences* and that of *Facial Expression Retrieval from 3D mesh sequences*. A lot of research has been dedicated to address the problem of *Facial Expression Recognition* in sequences of *3D* facial meshes. On the contrary, to the best of our knowledge, no research on *Facial Expression Retrieval* using *3D* facial mesh sequences appears in the bibliography. The present work thus addresses the latter problem.

Human emotions are often expressed by facial expressions instead of verbal communication. Facial expressions are generated by facial muscle movements, resulting in temporary deformation of the face. Ekman [1] was the first to systematically study human facial expressions. His study categorizes the prototypical facial expressions, apart from neutral expression, into six classes representing anger, disgust, fear, happiness, sadness and surprise. This categorization is consistent across different ethnicities and cultures. Furthermore, each of the six aforementioned expressions is mapped to specific movements of facial muscles, called Action Units (*AU*s). This led to the Facial Action Coding System (*FACS*), where facial changes are described in terms of *AU*s.

In recent years, automatic analysis of facial expressions has emerged as an active research area due to its various applications such as human-computer interaction, engineering, human behavior understanding, biometrics, emotion recognition, computer graphics, driver fatigue detection and psychology.

1.1 Method Overview

This dissertation focuses on the problem of dynamic *3D* facial expression retrieval from large datasets. A lot of research has been dedicated to address the problem of facial expression

---

* Dissertation Advisor: Theoharis Theoharis, Professor

recognition in *4D* data. On the contrary, to the best of our knowledge, no research on facial expression retrieval in *4D* data appears in the bibliography.

In order to address this problem we develop a 3-step retrieval framework: (i) initially, eight *3D* facial landmarks are automatically detected on each *3D* facial mesh of the sequence. (ii) Next, the landmarks are used in order to create a descriptor for the dynamic *3D* facial expression sequence. (iii) Finally, distance functions are used in order for different descriptors (i.e. query descriptor vs dataset descriptor) to be compared and the retrieval list is produced. The pipeline of our scheme is illustrated in Figure 1.

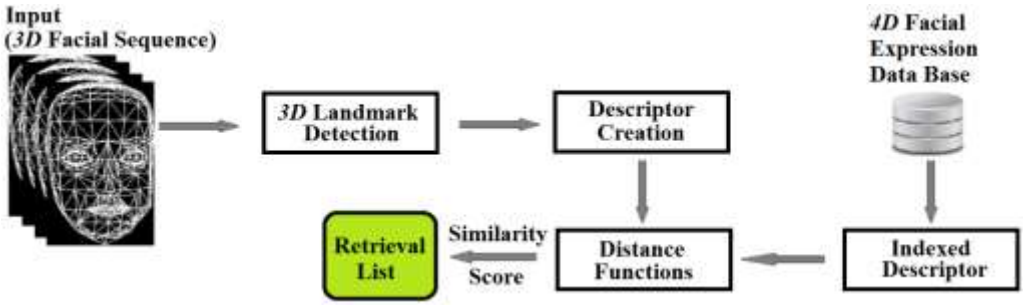


Figure 1: Pipeline of the proposed *4D* facial expression recognition scheme.

At first, eight *3D* facial landmarks are automatically detected on each *3D* facial mesh of the sequence. Each face *3D* mesh is, if not otherwise stated, defined by a set of points in the R^3 space (**vertices**) and a set of triangular **faces** defined in terms of the vertices.

The core of the problem is the computation of a feature set for each dynamic *3D* facial expression sequence. In this step, the structural and/or other special characteristics of the sequence are modelled and a descriptor that faithfully encodes the essence of the mesh sequence, in an efficient manner, is created. Feature selection is tightly connected to the corresponding application and can vary among different *4D* object retrieval systems.

Finally, each *3D* mesh sequence descriptor is used as a signature during the matching procedure. At this step, the signatures of the dynamic *3D* facial expressions, stored in the database, are compared to the corresponding signature of the query dynamic *3D* facial expression, using a specified metric, called *Distance Function*. The selected metric is dependent on both the selected features and the corresponding application. Finally, the response of the dynamic *3D* facial expression retrieval scheme is the set of dynamic *3D* facial expression(s) that correspond to the closest match (es) of the given user query.

1.2 Contributions

This thesis has made the following research contributions in the area of object retrieval: (1) Six new descriptors for the purpose of *Human Facial Expression Retrieval* from *3D* mesh sequences were proposed. For the creation of the descriptors we have used less landmarks than the state-of-the-art methods. (2) A novel mapping from facial features to primitive facial movements is proposed.

The descriptors developed and described in this dissertation are evaluated in terms of retrieval accuracy and demonstrated using both quantitative and qualitative measures via an extensive evaluation against state-of-the-art descriptors on standard datasets. This comparison illustrates the superiority of our descriptors compared to the state-of-the-art.

The overview of this thesis is as follows: In Section 2, the standard *4D* facial expression datasets are reviewed. In Section 3, state-of-the-art methods in the field of Human Facial Expression Recognition from *3D* Mesh Sequences are reviewed. In Section 4, the method for extracting specific *3D* facial landmarks from *3D* facial meshes is presented. In Section 5, the six descriptors, developed during this dissertation for the purpose of Human Facial Expression Retrieval from *3D* Mesh Sequences, are presented. In Section 6, distance functions for descriptor comparison purposes, are illustrated and compared. Section 7 presents the evaluation methodology and illustrates the extensive experimental results of the methods presented in this dissertation, against the state-of-the-art works on standard datasets. In Section 8, a supervised technique for detecting *AUs* is illustrated. Finally, in Section 9, conclusions are drawn.

2. 4D Facial Expression Datasets

The first dataset consisting of faces recorded in 3*D* video is *BU−4DFE*, presented by Yin *et al*. [2]. This dataset was made available in 2008. It involves 101 subjects (58 females and 43 males) of various ethnicities. For each subject the six basic expressions were recorded gradually from neutral face, outset, apex, offset and back to neutral, using the dynamic facial acquisition system *Di3D* (http://www.di3d.com) and producing roughly 60,600 3*D* facial meshes (frames), with corresponding texture images. Finally, each frame is associated with 83 facial landmark points.

Zhang *et al*. [3] presented the *BP4D-Spontaneous* dataset in 2013 to the research community. This dataset contains high-resolution spontaneous 3*D* dynamic facial expressions by encoding 27 *AUs* and their various combinations. It involves 41 subjects (23 females and 18 males) of various ethnicities. The subjects were 18-29 years of age. Each subject was recorded using the dynamic facial acquisition system *Di3D* (http://www.di3d.com). 328 3*D* sequences were created. Finally, each frame is associated with 83 facial landmark points. In Table 1, the basic characteristics of 3*D* video facial expression datasets are shown.

Table 1: Current publicly available datasets of *3D* facial expression mesh sequences.

Dataset	Year	Number of Subjects	Number of Expressions	Number of Landmarks
BU-4DFE	2008	101	6	83
BP4D-Spontaneous	2014	41	8	83

3. Related Work

Due to the lack of works in the area of *3D* video facial expression *retrieval* techniques, in this chapter we will review the *3D* video facial expression *recognition* state-of-the-art methodologies. We will focus on the descriptors which are the common necessities in both areas. The reader is also referred to the surveys presented in [4], [5], [6] and [7].

The typical operational pipeline employed by 3*D* video facial expression recognition methodologies is shown in Figure 2. *3D* video facial expression recognition methodologies take into account 3*D* facial data as 3*D* surfaces. Another common trait of these methodologies is the use of a variety of 3*D* dynamic face analysis techniques to detect and exploit the discrete and well studied facial muscle motions.

3D dynamic face analysis techniques can be divided into two major categories: Tracking–based and 3*D* facial model-based. Tracking-based techniques aim to track specific 3*D* facial model marks using appropriate tracking algorithms. On the other hand, 3*D* facial model-based techniques aim to exploit the facial deformations which take place due to a facial expression. These techniques often use alignment methods to achieve better results.

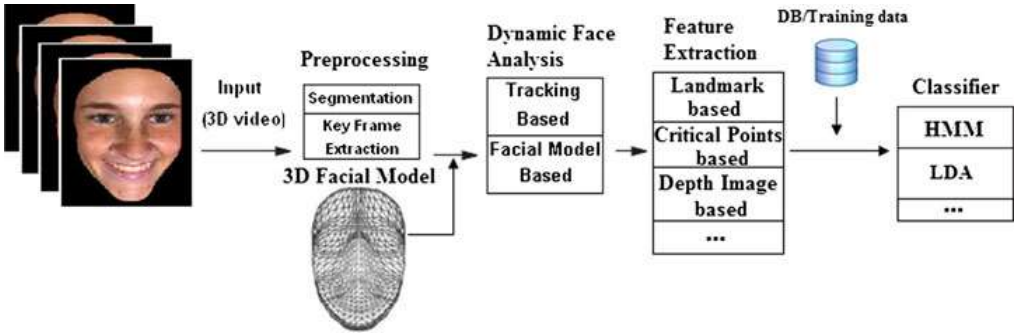


Figure 2: *3D* video facial expression recognition pipeline.

Tracking-based techniques can be further distinguished into two sub- categories: Landmark tracking-based and critical points tracking-based. In the first case, areas are tracked around specific facial landmarks along 3*D* frames and detect temporal changes on their geometry characteristics. In the latter case, techniques aim to track 3*D* model key points along time and detect temporal changes on spatial characteristics that are defined by these points.

Three dimensional facial model-based techniques can also be divided into two subcategories: Facial deformation-based, which aim to detect temporal deformations using a generic face model, and facial surface-based, which create facial surfaces on different face depth levels (i.e., different values on the *z*-axes). Then, estimate the intersection along time between the face and each surface, they extract the final descriptor. A summarization of the state-of-the-art methods is illustrated in Table 2. 'N/A' is used to indicate that the corresponding information is not available.

Table 2: Overview of research work on *3D* video facial expression recognition.

Method	*Dataset*	*Number of Expressions*	*3D Face Analysis*	*Features*	*Classifier*	*Automatic*	*Real-Time Suitability*	*Recognition Accuracy*
Chang *et al.* [8]	Proprietary	6	Landmark tracking	Generalized manifold + Texture	Bayes	NO	NO	N/A
Rosato *et al.* [9]	*BU-3DFE*	6	Landmark tracking	Generalized manifold + Texture	*LDA*	YES	NO	85.90%
Sun *et al.* [10]	*BU-4DFE*	6	Landmark tracking	Gradient + Curvature	*HMM*	YES	NO	94.37%
Tsalakanidou *et al.* [11]	Proprietary	4 (10 *AUs*)	Landmark tracking	Gradient + Curvature	*FACS*	YES	YES	84.00%
Tsalakanidou *et al.* [12]	Proprietary	4	Landmark tracking	Gradient + Curvature	*FACS*	YES	YES	85.00%
Sun *et al.* [13]	*BU-3DFE*	0 (8 *AUs*)	Landmark tracking	Curvature	*HMM*	NO	NO	87.10%
Sun *et al.* [14]	*BU-3DFE*	6	Landmark tracking	Curvature	*HMM*	NO	NO	90.44%
Canavan *et al.* [15]	*BU-4DFE*	6	Landmark tracking	Curvature	*SVM*	YES	NO	84.80%
Berretti *et al.* [16]	*BU-4DFE*	3	Critical point tracking	Average facial distances	*HMM*	YES	YES	76.30%
Jeni *et al.* [17]	*BU-4DFE*	6 (17 *AUs*)	Critical point tracking	Shape index	*SVM*	YES	NO	78.18%
Yin *et al.* [18]	*BU-3DFE*	6	Facial deformation	*FELM* + Motion vectors	*LDA*	NO	NO	80.20%
Sandbach *et al.* [19]	*BU-4DFE*	6	Facial deformation	Vector direction distribution	*GB + HMM*	YES	NO	64.46%
Sandbach *et al.* [20]	*BU-4DFE*	3	Facial deformation	Mean + *STD* of vector direction distribution	*HMM*	YES	NO	81.93%
Fang *et al.* [21]	*BU-4DFE*	6	Facial deformation	*LBP-TOP*	*SVM*	YES	NO	75.82%
Fang *et al.* [22]	*BU-4DFE*	6	Facial deformation	*LBP-TOP*	*SVM*	YES	NO	91.00%
Zhang *et al.* [23]	*BP4D-Spontaneous*	0 (27 *AUs*)	Facial deformation	Curvature + Polar angles	*SVM*	YES	NO	61.33%
Zhang *et al.* [23]	*BU-4DFE*	6	Facial deformation	Curvature + Polar angles	*SVM*	YES	NO	76.12%
Le *et al.* [24]	*BU-4DFE*	3	Facial surface	Chamfer distances	*HMM*	YES	NO	92.22%
Drira *et al.* [25]	*BU-4DFE*	6	Facial surface	*DVF*	*LDA + RMF*	YES	YES	93.21%

4. *3D* Facial Landmarks Detection

The first step of our proposed retrieval scheme is the detection of *3D* facial landmarks on the dynamic *3D* mesh sequence (see Figure 1). Eight facial landmarks, on the *3D* facial scans, are exploited. More specifically, four landmarks for the eyes, two for the mouth, one for the nose and one for the chin are used (see Figure 3). The number of landmarks used here is smaller than the number used by other state-of-the-art techniques.

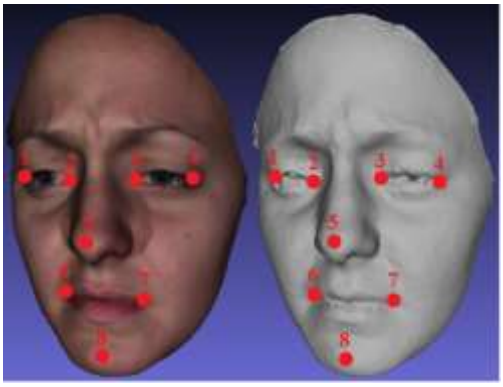


Figure 3: Eight facial landmarks used for the proposed retrieval scheme.

The landmarks are automatically detected using the state-of-the-art methodology previously developed by our team [26], making the proposed retrieval scheme self-contained.

The method presented in [26] performs automatic pose-invariant detection of landmarks on *3D* facial scans under large yaw variations, and is invariant to facial expressions. Three-dimensional information is exploited by fusing *3D* local shape descriptors to extract candidate landmark points. The shape descriptors include the shape index, a continuous map of principal curvature values of a *3D* object's surface and spin images, which are local descriptors of the object's *3D* point distribution.

Landmark detection takes part in two phases. In the training phase, a Facial Landmark Model (*FLM*) representing the landmark positions is created, shape index target values for each landmark are computed and spin image templates for each landmark are generated.

In the detection phase, the algorithm first detects candidate landmarks on the probe facial datasets, by exploiting the *3D* geometry-based information of shape index and spin images. The extracted candidate landmarks are then filtered out and labeled by matching them with the *FLM*. The facial landmark detection method is robust to rotations about the vertical facial axis up to 60 degrees and returns the detected pose; this information is used in our method in order to rotate facial instances that are not frontal.

5. Descriptors for *3D* Facial Mesh Sequences

The first two descriptors proposed in this chapter are spatial, which means that they are based only on spatial changes of the facial expressions across time. The remaining four are spatio-temporal, which means that they are based on both temporal and spatial changes of the facial expressions.

The motivation behind the proposed spatial, hybrid facial expression descriptors *GeoTopo* and *GeoTopo+* is the fact that some facial expressions, like happiness and surprise, are characterized by obvious changes in the mouth topology while others, like anger, fear and sadness, produce geometric but no significant topological changes.

The motivation behind the spatio-temporal descriptors is the expectation that the extra information offered by the temporal dimension can lead to a more accurate descriptor. In addition the descriptor can potentially be made more compact if it stores aggregations of attributes across the time dimension.

GeoTopo Descriptor (GEOmetric & TOPOlogic)

The proposed *GeoTopo* (*Geo*metric and *Topo*logical) descriptor captures geometric, as well as, topological information, which is achieved by the concatenation of two separate sub-descriptors, one expressing the facial geometry and one the facial topology.

The geometric part of the *GeoTopo* descriptor is a simple *2D* function ($G_1(i,j)$), as illustrated in Equation (1). Function G_1 represents the maximum curvature of the *j*-th landmark (L_j) in the *i*-th *3D* mesh ($mesh_i$).

$$G_1\,(i,j) = MaxCurvature\,(mesh_i, L_j) \tag{1}$$

The topological sub-descriptor is also a *2D* function ($T(i, j)$), as illustrated in Equation (2). Function *T* represents the value of the *j*-th feature, related to one or more *AUs*, in the *i*-th *3D* mesh. Ten features are selected in total. One of them is angular, four are areas and five express distances on the face. The calculations of the values of these ten features are performed using exclusively the *3D* coordinates of the eight tracked landmarks (*LM*s) in the *i*-th *3D* time mesh.

$$\boldsymbol{T}(\boldsymbol{i},\boldsymbol{j}) = \begin{cases} \boldsymbol{Angle}_{i,j}(\boldsymbol{LMs}) : j \in \{1\}, \\ \boldsymbol{Area}_{i,j}(\boldsymbol{LMs}) : j \in \{2,\ldots,5\}, \\ \boldsymbol{Distance}_{i,j}(\boldsymbol{LMs}) : j \in \{6,\ldots,10\} \end{cases} \tag{2}$$

Each facial expression can be deconstructed into specific *AUs*. There is a correspondence between each facial muscle and a number of *AUs*. The actual type of the *AU* is determined by the muscle's temporal movement. Figure 4 illustrates the mapping of the ten selected topological features on 3D facial mesh. The concatenation of the aforementioned sub-descriptors, as illustrated in Equations (1) and (2), produces the final *GeoTopo* descriptor: *GeoTopo* = (G_1 ++ *T*).

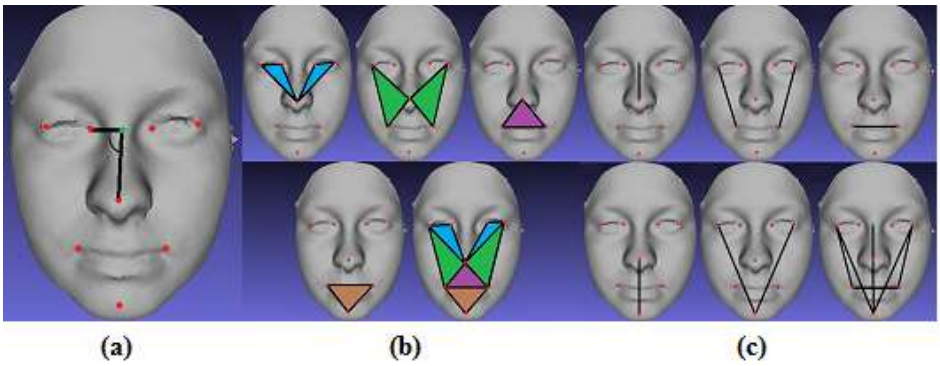


Figure 4: Topological features in use. (a) Angle, (b) Areas, (c) Distances.

GeoTopo+ Descriptor *(GEOmetric & TOPOlogic PLUS)*

GeoTopo+ is a spatial, hybrid descriptor which combines three sub-descriptors capturing topological, as well as, geometric information of the *3D* facial meshes. Two sub-descriptors are used for capturing the facial geometry, based on the heat kernel signature of the *3D* facial surface and the *3D* facial model vertices' normal vectors. The third sub-descriptor is used for capturing facial topology based on *FACS AUs*.

The first geometric sub-descriptor is a simple *2D* function ($G_2(i,j)$), as illustrated in Equation (3). Function G_2 represents the heat kernel signature (*HKS*) [27] of the *j*-th landmark (L_j) in the *i*-th *3D* mesh ($mesh_i$). *HKS* is based on the properties of the heat diffusion process on a shape. It is obtained by restricting the heat kernel in the temporal domain, thus obtaining a local effect.

$$G_2\,(i, j) = HKS\,(mesh_i, L_j) \tag{3}$$

The second geometric sub-descriptor of the *GeoTopo*+ descriptor is the *2D* function ($G_3(i,j)$) presented in Equation (4), which represents the normal vector of the *j*-th landmark (*Lj*) in the *i*-th *3D* mesh ($mesh_i$).

$$G_3\,(i, j) = NormalVector\,(mesh_i, L_j) \tag{4}$$

The topological sub-descriptor is the *2D* function $T(i,j)$ already presented in Equation (2). The concatenation of the aforementioned sub-descriptors, as illustrated in Equations (2), (3) and (4) produces the final *GeoTopo*+ descriptor: *GeoTopo*+ = (G_2 ++ G_3 ++*T*).

DCT-*GeoTopo* Descriptor *(Discrete Cosine Transformation – GeoTopo)*

DCT-GeoTopo uses only a subset of the topological sub-descriptor $T(i,j)$, illustrated in Equation (2). To produce the final descriptor, we apply the *Discrete Cosine Transformation* (*DCT*) on the temporal values of the features producing a transformed sequence for each feature. This transformation maps the features from the temporal to the frequency domain and thus the transformed features represent the spatio-temporal deformation of the initial features. Eight

features of the transformed sequences are selected to construct the final descriptor. Equation (5) represents the final descriptor; this is an *8D* vector irrespective of the number of meshes of the corresponding facial expression *3D* sequence.

$$\text{DCT-GeoTopo}=\begin{bmatrix} 2^{nd}\ DCT \text{ component for area with feature code } \#3, \\ 3^{rd}\ DCT \text{ component for area with feature code } \#3, \\ 3^{rd}\ DCT \text{ component for area with feature code } \#5, \\ 2^{nd}\ DCT \text{ component for distance with feature code } \#6, \\ 4^{\eta}\ DCT \text{ component for distance with feature code } \#6, \\ \text{M}ean\ DCT \text{ components value for distance with feature code } \#8, \\ 2^{nd}\ DCT \text{ component for additional distance of Figure 41}, \\ 2^{nd}\ DCT \text{ component for distance with feature code } \#10 \end{bmatrix} \quad (5)$$

WT-GeoTopo+* Descriptor *(Wavelet Transformation – GeoTopo+)

For the construction of the *WT-GeoTopo+*, the extracted landmarks are used in order to capture the geometric and topological information for each mesh the same way as in *GeoTopo+* descriptor. Then, we perform Wavelet Transformation [28], using *Gaussian Wavelets* at 64 different scales, on each temporal information sequence. We apply the *1D* Wavelet Transformation on all temporal sequences for every feature, as indicated in Equations (6) – (8). The term *WT*() indicates the Wavelet Transformation of the *1D* signal placed within the parenthesis, and the term * indicates the values of the landmarks over the whole set of *3D* meshes of the sequence.

$$W_{G1}(*, j) = WT(G_1(*, j)) \quad (6)$$
$$W_{G2}(*, j) = WT(G_2(*, j)) \quad (7)$$
$$W_T(*, j) = WT(T(*, j)) \quad (8)$$

For each transformed sequence we extract four feature aggregators: the *median*, the *mode* (the most frequently occurring element), the *norm* and the *root mean square* (*RMS*) of the sequence. Thus, we get a *4D* feature vector for each sequence as indicated by Equations (9) – (11). The number of *4D* feature vectors is constant (26 *4D* feature vectors), independent of the number of the frames of the sequence and constitutes the spatio-temporal information of the *3D* facial expression sequence.

$$ST_{G_1} = \{Mean(W_{G_1}(*, j)), Mode(W_{G_1}(*, j)), Norm(W_{G_1}(*, j)), RMS(W_{G_1}(*, j))\} \quad (9)$$
$$ST_{G_2} = \{Mean(W_{G_2}(*, j)), Mode(W_{G_2}(*, j)), Norm(W_{G_2}(*, j)), RMS(W_{G_2}(*, j))\} \quad (10)$$
$$ST_T = \{Mean(W_T(*, j)), Mode(W_T(*, j)), Norm(W_T(*, j)), RMS(W_T(*, j))\} \quad (11)$$

The concatenation of the above sub-descriptors produces the final descriptor: *WT-GeoTopo+* = *(ST_{G1} ++ ST_{G2} ++ ST_T)*.

CVD* Descriptor *(Coordinate Vector Descriptor)

CVD descriptor uses only six of the extracted landmarks (1st, 4th, 5th, 6th, 7th and 8th of Figure 3) and captures their positional information for each frame and, therefore, for the entire dynamic *3D* mesh sequence. The positional information, actually corresponds to the coordinate values of each one of the six facial landmarks.

As a pre-processing step, for each facial mesh, we perform translation so that the nose tip (3rd landmark) coincides with the center of the coordinate system. Thus, even though the initial data are registered, we create even better consistency between the *3D* meshes of the sequence. Next, the Z-coordinates of the 1st, 4th, 6th, 7th and 8th landmark and the Y-coordinates of the 6th, 7th and 8th landmark are kept. The above procedure produces, for each *3D* frame, a feature vector of length eight. Normalization sets the feature values of each vector in the interval [0, 1] and has been performed for each vector item separately. Finally, the set of all the normalized feature vectors of a dynamic *3D* facial expression sequence are averaged, resulting in a single *8D* feature vector.

The function for calculating the *8D* coordinate vector for each *3D* frame *f*, FC_f , is given in Equation (12) and the function for calculating the averaged *8D* coordinate vector, *CVD* , of an entire *3D* sequence is given in Equation (13). *N* is the number of *3D* frames in the sequence.

By capturing specific landmarks' successive coordinate values, even the slightest facial motions are detected. Furthermore, in the proposed descriptors presented earlier in this dissertation, the motion of each landmark is attached to one (in case distance feature is used)

or two (in case angle or area feature is used) other landmarks. On the contrary, the *CVD* coordinate vector captures each landmark's behavior independently of all the other landmarks. The experimental results show that positional features are more descriptive than relational features. However, the results can be further improved using the Wavelet Transformation of the following section.

$$FC_f = \begin{Bmatrix} Z \text{ Coordinate of the } 1^{st} \text{ landmark.} \\ Z \text{ Coordinate of the } 4^{th} \text{ landmark.} \\ Z \text{ Coordinate of the } 6^{th} \text{ landmark.} \\ Z \text{ Coordinate of the } 7^{th} \text{ landmark.} \\ Z \text{ Coordinate of the } 8^{th} \text{ landmark.} \\ Y \text{ Coordinate of the } 6^{th} \text{ landmark.} \\ Y \text{ Coordinate of the } 7^{th} \text{ landmark.} \\ Y \text{ Coordinate of the } 8^{th} \text{ landmark.} \end{Bmatrix} \quad (12)$$

$$CVD = \frac{\sum_{f=1}^{f=N} FC_f}{N} \quad (13)$$

WT-CVD$_b$* Descriptor *(Wavelet Transformation – CVD)

The positional information of the *CVD* descriptor can be used as a basis for the creation of a hybrid spatio-temporal descriptor. To this effect, we perform Wavelet Transformation, using *Gaussian Wavelets* at 64 different scales, on the coordinate information. We apply the *1D* Wavelet Transformation on the *8D* coordinate vector, as indicated in Equation (13). The term *WT*()$_b$ indicates the Wavelet Transformation of the *1D* signal placed within the parenthesis for scale *b*, and the term * indicates that all the components of the vector are used for the calculation.

$$WT - CVD_b = WT(CVD(*))_b, b = 1,...,64 \quad (14)$$

WT-CVD$_b$ is a spatio-temporal descriptor of constant length. Unlike *WT-GeoTopo+* descriptor, we haven't extracted any feature aggregators for the 64 wavelet transformed sequences. If we would have used aggregators, the information for each axis would be confused.

We prefer *Wavelet Transformation* as it has better resolution than *Fourier* and *Cosine Transformation* [29]. This means that each coefficient of the transformation, which expresses both frequency and time domain information, is created in such a way as to capture as much as possible and as precise as possible frequency-time information.

6. Distance Functions

A *Distance Function* is a mathematical expression which is applied on two given time series and produces a scalar metric as output. This output is a non-negative integer number and is called *similarity score*. The similarity score is a mathematical expression of how similar the two input time series are. If the similarity score equals 0, then the two input time series are exactly the same and thus, we have maximum similarity. In general, the bigger the similarity score is, the less similar the two input time series are.

The selection of a distance function is not trivial. It is dependent on both the features that have been selected for the descriptor and the envisaged application. Apart from the initial intuition, the selection of distance function involves extensive experimentation.

The only distance function appropriate for comparing two time series of different length is the Dynamic Time Warping (*DTW*) [30]. That is why, it is suitable for comparing *GeoTopo* and *GeoTopo+* descriptors, since their length differs and is dependent on the number of the *3D* meshes of the mesh sequence corresponding to the descriptor. *DTW* minimizes the effects of shifting and distortion in time by allowing "elastic" transformation of a time series in order to detect similar shapes with different phases. Unfortunately, *DTW* is time consuming compared to other distance functions.

Experimental results show that, in the case of the *WT-GeoTopo+* and *DCT-GeoTopo* spatio-temporal descriptors the most suitable distance functions are *Square of Euclidean Distance* and *Kullback-Leibler Divergence* respectively. Still, there are cases where *DTW* produces significantly better results even for sequences of the same length; such cases are the *CVD* and *WT-CVD$_b$* descriptors.

7. Experimental Results

The experimental evaluation is based on the Precision-Recall curves (or *P-R* Diagram) and five quantitative measures: Nearest Neighbor (*NN*), First Tier (*FT*), Second Tier (*ST*), *E*-measure (*E*-m) and Discounted Cumulative Gain (*DCG*) [31] for the classes of each corresponding dataset. The datasets used for conducting the experiments are the ones presented in Section 2.

Retrieval Evaluation

Several parameters had to be determined in order to conduct the experiments. Initially, descriptor normalization took place. Normalization sets the feature values in the interval [0, 1]. Then a subtraction scheme was implemented; the descriptor values are not used as absolute values corresponding to the current time mesh, but as differences of the current from the initial time mesh.

Table 3 illustrates the retrieval evaluation of the proposed descriptors, compared to the state-of-the-art descriptors for *BU-4DFE* dataset. Table 4 illustrates the corresponding evaluation for *BP4D-Spontaneous* dataset. The proposed descriptors outperform the state-of-the-art in both cases.

Table 3: Comparison of descriptors for *BU-4DFE* dataset.

Descriptor	NN	FT	ST	DCG
WT-CVD$_b$ **[32]**	**0.82**	**0.73**	**0.95**	**0.92**
CVD [32]	0.82	0.72	0.95	0.92
WT-GeoTopo+ [33]	0.81	0.65	0.76	0.92
DCT-GeoTopo [34]	0.75	0.61	0.66	0.86
GeoTopo+ [35]	0.73	0.56	0.77	0.91
GeoTopo [36]	0.71	0.55	0.73	0.89
Berretti *et al.* [16]	0.60	0.50	0.70	0.88
Distribution Vectors	0.52	0.41	0.59	0.82
Curvature	0.47	0.40	0.60	0.82
LBP-TOP	0.43	0.37	0.54	0.80
FELM	0.42	0.38	0.56	0.80
Gradient	0.40	0.36	0.54	0.79
Shape Index	0.35	0.37	0.54	0.79

Table 4: Comparison of descriptors for *BP4D-Spontaneous* dataset.

Descriptor	NN	FT	ST	DCG
WT-CVD$_b$ **[32]**	**0.76**	**0.62**	**0.72**	**0.84**
CVD [32]	0.53	0.60	0.71	0.81
WT-GeoTopo+ [33]	0.75	0.61	0.69	0.83
DCT-GeoTopo [34]	0.70	0.58	0.59	0.78
GeoTopo+ [35]	0.67	0.55	0.72	0.83
GeoTopo [36]	0.61	0.52	0.69	0.82
Berretti [16]	0.59	0.49	0.69	0.81
Distribution Vectors	0.50	0.41	0.57	0.76
Curvature	0.39	0.34	0.47	0.71
LBP-TOP	0.39	0.34	0.47	0.71
FELM	0.63	0.35	0.48	0.77

Gradient	0.38	0.33	0.46	0.71
Shape Index	0.30	0.32	0.47	0.70

Recognition Evaluation

To achieve recognition, a majority voting scheme is implemented among the *k*-top retrieval results. The query expression is classified as belonging to the outvoting class. Experimental results showed that optimal recognition accuracy is achieved for $k = 11$.

Table 5 illustrates the recognition evaluation of the proposed descriptors, compared to the other descriptors, proposed in this thesis, as well as the state-of-the-art descriptors for *BU-4DFE* dataset. Obviously, three descriptors proposed here, outperform the state of the art descriptors by far. Three of our proposed descriptors were tested, for the first time in terms of retrieval, *in BP4D-Spontaneous* dataset and the results are illustrated in Table 6.

Table 5: Comparison of recognition accuracy for *BU-4DFE* dataset.

Method	Nr Facial Expressions	Recognition Accuracy
WT-CVD [32]	**6**	**100.0%**
CVD [32]	6	100.0%
WT-GeoTopo+ [33]	6	96.04%
Sun [10]	6	94.37%
Drira [25]	6	93.21%
Fang [22]	6	91.00%
DCT-GeoTopo [34]	6	90.83%
GeoTopo+ [35]	6	90.00%
Canavan [15]	6	84.80%
GeoTopo [36]	6	84.18%
Berretti [16]	6	79.40%
Jeni [17]	6	78.18%
Zhang [23]	6	76.12%
Fang [6]	6	75.82%
Sandbach [19]	6	64.60%

Table 6: Comparison of recognition accuracy for *BP4D-Spontaneous* dataset.

Method	Nr Facial Expressions	Recognition Accuracy
WT-CVD$_b$ [32]	**8**	**100. 0%**
CVD [32]	8	100.0%
GeoTopo+ [32]	8	88.56%

8. Action Units Detection

The detection of facial Action Units lies in the core of facial analysis and, while related to retrieval, has far broader applications such as biometrics, interaction, behavioral analysis and animation control.

The close connection between the *AU*s and the topological features of the *GeoTopo+* descriptor makes *T* appropriate for the detection of active *AUs* across a dynamic *3D* facial mesh sequence.

We train a classifier on the topological features of Equation (2) for the case of an *AU* activation/no activation. Then, we can decide whether an *AU* is activated/not activated across a facial mesh sequence. The training and the testing sets are defined using the 5-fold cross validation method. The results, compared to the state-of-the-art are illustrated in Table 7. We achieve mostly better detection and we are capable of detecting twelve more *AU*s.

Table 7: Detection of activated *AU*s (%) for *BP4D-Spontaneous* dataset.

*AU*s	Proposed Method	Method [22]	AUs	Proposed Method	Method [22]
1	65.6	58.4	15	61.2	69.0
2	51.9	64.8	16	59.1	-
4	63.3	63.1	17	75.3	65.6
5	57.9	-	18	93.2	-
6	67.4	68.8	19	85.1	-
7	77.4	58.9	20	64.6	-
9	58.1	-	22	90.7	-
10	66.9	66.4	23	61.8	61.4
11	92.0	-	24	59.4	67.6
12	64.8	59.1	27	81.4	-
13	97.8	-	28	59.6	-
14	64.4	59.1	30	96.1	-

9. Conclusions

The present dissertation proposes a robust scheme for facial expression retrieval from sequences of *3D* facial meshes. Our scheme consists of three steps: (i) Detection of landmarks for each *3D* facial mesh of the sequence, (ii) Creation of the descriptor for the sequence and (iii) Comparison of different descriptors (i.e. query descriptor vs dataset descriptors).

The descriptors developed and described in this dissertation are evaluated in terms of retrieval accuracy and demonstrated using both quantitative and qualitative measures via an extensive evaluation against state-of-the-art descriptors on well-known, publicly available datasets. This comparison illustrates the superiority of our descriptors compared to the state-of-the-art.

Furthermore, a technique which exploits the retrieval results, in order to achieve unsupervised dynamic *3D* facial expression recognition is presented. The proposed unsupervised technique exhibits improved performance against supervised state-of-the-art techniques. Finally, the features of the topological part of the *GeoTopo+* descriptor are used for supervised *AU* activation detection, from dynamic *3D* facial mesh sequences. The detection performance of the proposed technique improves on the state-of-the-art for most *AU*s while it can detect twelve more *AU*s than the state-of-the-art.

References

[1] P. Ekman and W. Friesen, *Facial action coding system: a technique for the measurement of facial movement*, Consulting Psychologists Press, Palo Alto, 1978.

[2] L. Yin, X. Wei, Y. Sun, J. Wang and M. J. Rosato, "A 3D facial expression database for facial behavior research," in *IEEE Proceedings on FGR '06*, 2006, pp. 211–216.

[3] X. Zhang, L. Yin, J. Cohn, S. Canavan, M. Reale, A. Horowitz, P. Liu and J. Girard, "BP4D-Spontaneous: A high resolution spontaneous 3D dynamic facial expression database," *Image and Vision Computing*, vol. 32, no. 10, pp. 692-706, 2014.

[4] A. Danelakis, T. Theoharis and I. Pratikakis, "3D mesh video retrieval: A survey," in *Proceedings on 3DTV-CON '12*, 2012, pp. 1–4.

[5] A. Danelakis, T. Theoharis and I. Pratikakis, "A Survey on Facial Expression Recognition in 3D Video Sequences," *Multimedia Tools and Applications*, vol. 74, no. 15, pp. 5577-5615, 2014.

[6] T. Fang, X. Zhao, O. Ocegueda, S. K. Shah and I. A. Kakadiaris, "3D facial expression recognition: A perspective on promises and challenges," in *IEEE Proceedings on FG '11*, 2011, pp. 603–610.

[7] G. Sandbach, S. Zafeiriou, M. Pantic and L. Yin, "Static and dynamic 3D facial expression recognition: A comprehensive survey," *Image and Vision Computing*, vol. 30, no. 10, pp. 683–697, 2012.

[8] Y. Chang, M. B. Vieira, M. Turk and L. Velho, "Automatic 3D facial expression analysis in videos," in *IEEE Workshop AMFG '05*, 2005, pp. 293–307.

[9] M. Rosato, X. Chen and L. Yin, "Automatic registration of vertex correspondences for 3D facial expression analysis," in *IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2008, pp. 1–7.

[10] Y. Sun, X. Chen, M. J. Rosato and L. Yin, "Tracking vertex flow and model adaptation for threedimensional spatiotemporal face analysis," *IEEE Transactions on Systems Man and Cybernetics Part A*, vol. 40, no. 3, pp. 461–474, 2010.
[11] F. Tsalakanidou and S. Malassiotis, "Robust facial action recognition from real-time 3D streams," in *CVPR '09*, 2009, pp. 4–11.
[12] F. Tsalakanidou and S. Malassiotis, "Real-time 2D+3D facial action and expression recognition," *Pattern Recognition*, vol. 43, no. 5, pp.1763–1775, 2010.
[13] Y. Sun, M. Reale and L. Yin, "Recognizing partial facial action units based on 3D dynamic range data for facial expression recognition," in *IEEE FG '08*, 2008, pp 1–8.
[14] Y. Sun and L. Yin, "Facial expression recognition based on 3D dynamic range model sequences," in *Proceedings on ECCV '08*, 2008, pp. 58–71.
[15] S. J. Canavan, Y. Sun, X. Zhang and L. Yin, "A dynamic curvature based approach for facial activity analysis in 3D space," in *CVPR '12 Workshops*, 2012, pp. 14–19.
[16] S. Berretti, A. D. Bimbo and P. Pala, "Automatic facial expression recognition in real-time from dynamic sequences of 3D face scans," *The Visual Computer*, vol. 29, no. 12, pp. 1333-1350, 2013.
[17] L. A. Jeni, A. Lörincz, T. Nagy, Z. Palotai, J. Sebök, Z. Szabó and D. Takács, "3D shape estimation in video sequences provides high precision evaluation of facial expressions," *Image and Vision Computing*, vol. 30, no. 10, pp. 785-795, 2012.
[18] L. Yin, X. Wei, P. Longo and A. Bhuvanesh, "Analyzing facial expressions using intensity-variant 3D data for human computer interaction," in *Proceedings on ICPR '06*, 2006, pp. 1248–1251.
[19] G. Sandbach, S. Zafeiriou, M. Pantic and D. Rueckert, "Recognition of 3D facial expression dynamics," *Image and Vision Computing*, vol. 30, no. 10, pp. 762–773, 2012.
[20] G. Sandbach, S. Zafeiriou, M. Pantic and D. Rueckert, "A dynamic approach to the recognition of 3D facial expressions and their temporal models," in *IEEE FG '11*, 2011, pp. 406–413.
[21] T. Fang, X. Zhao, S. K. Shah and I. A. Kakadiaris, "4D facial expression recognition," in *ICCV '11*, 2011, pp. 1594-1601.
[22] T. Fang, X. Zhao, O. Ocegueda, S. K. Shah and I. A. Kakadiaris, "3D/4D facial expression analysis: An advanced annotated face model approach," *Image and Vision Computing*, vol. 30, no. 10, pp.738–749, 2012.
[23] X. Zhang, M. Reale and L. Yin, "Nebula feature: a space-time feature for posed and spontaneous 4D facial behavior analysis," in *IEEE FG '13*, 2013, pp. 1-8.
[24] V. Le, H. Tang and T. S. Huang, "Expression recognition from 3D dynamic faces using robust spatiotemporal shape features," in *IEEE FG '11*, 2011, pp. 414–421.
[25] H. Drira, B. B. Amor, M. Daoudi, A. Srivastava and S. Berretti, "3D dynamic expression recognition based on a novel deformation vector field and random forest," in *ICPR '12*, 2012, pp. 1104–1107.
[26] P. Perakis, G. Passalis, T. Theoharis and I. A. Kakadiaris, "3D facial landmark detection under large yaw and expression variations," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1552–1564, 2013.
[27] J. Sun, M. Ovsjanikov and L. Guibas, "A concise and provably informative multi-scale signature based on heat diffusion," in *SGP '09 Eurographics Association*, 2009, pp. 1383-1392.
[28] I. Daubechies, *Ten lectures on wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.
[29] R. Q. Quiroga, O. W. Sakowitz, E. Basar and M. Schürmann, "Wavelet transform in the analysis of the frequency composition of evoked potentials," *Brain Research Protocols*, vol. 8, no. 1, pp. 16-24, 2001.
[30] S. Salvador and P. Chan, "Toward accurate dynamic time warping in linear time and space," Intelligent Data Analysis, vol. 11, no. 5, pp. 561–580, 2007.
[31] P. Shilane, P. Min, M. M. Kazhdan, and T. A. Funkhouser, ''The princeton shape benchmark,'' in *IEEE Computer Society*, 2004, pp. 167–178.
[32] A. Danelakis, T. Theoharis and I. Pratikakis, "A Spatio-temporal Wavelet-based Descriptor for Dynamic 3D Facial Expression Retrieval and Recognition," *Pattern Recognition*, submitted, 2015.
[33] A. Danelakis, T. Theoharis and I. Pratikakis, "A Robust Spatio-Temporal Scheme for Dynamic 3D Facial Expression Retrieval," *The Visual Computer*, pp. 1-13, 2015.
[34] A. Danelakis, T. Theoharis and I. Pratikakis, "A Spatio-Temporal Descriptor for Dynamic 3D Facial Expression Retrieval and Recognition," in *3DOR '15 Eurographics Association*, 2015, pp. 63- 70.
[35] A. Danelakis, T. Theoharis, I. Pratikakis and P. Perakis, "An Effective Methodology for Dynamic 3D Facial Expression Retrieval," *Pattern Recognition*, 2015.
[36] A. Danelakis, T. Theoharis and I. Pratikakis, "Geotopo: Dynamic 3D facial expression retrieval using topological and geometric information," in *3DOR '14 Eurographics Association*, 2014, pp. 1-8.

Scaling storage systems for future eXascale environments

Christos Filippidis*

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
cfjs@outlook.com

Abstract. High performance computing (HPC) has crossed the Petaflop mark and is reaching the Exaflop range quickly. The exascale system is projected to have millions of nodes, with thousands of cores for each node. At such an extreme scale, the substantial amount of concurrency can cause a critical contention issue for I/O system. This study proposes a dynamically coordinated I/O architecture for addressing some of the limitations that current parallel file systems and storage architectures are facing with very large-scale systems. The fundamental idea is to coordinate I/O accesses according to the topology/profile of the infrastructure, the load metrics, and the I/O demands of each application. The measurements have shown that by using IKAROS approach we can fully utilize the provided I/O and network resources, minimize disk and network contention, and achieve better performance.

1 Introduction

Large-scale scientific computations tend to stretch the limits of computational power and parallel computing is generally recognized as the only viable solution to high performance computing problems. I/O has become a bottleneck in application performance as processor speed skyrockets, leaving storage hardware and software struggling to keep up. Parallel file systems have be developed in order to allow applications to make optimum use of available processor parallelism. The most important factors affecting performance are the number of parallel processes participating in the transfers, the size of the individual transfers and of course the access patterns. The I/O access patterns are generally divided into the following subgroups [1]:

(1) Compulsory (consist of I/Os that must be made to read a program's initial state from the disk and write the final state back to disk when the program has finished. For example, a program might read a configuration file and perhaps an initial set of data points, and then write out the final set of data points along with graphical and textual representations of the results).
(2) Checkpoint/restart (are used to save the state of a computation in case of a hardware or software error which would require the simulation to be restarted).

* Dissertation Advisor: Yiannis Cotronis, Associate Professor

(3) Regular snapshots of the computation's progress.
(4) Out-of-core read/writes for problems which do not fit to memory.
(5) Continuous output of data for visualization and other post-processing.

Another important factor that may significantly affect performance is the architecture of the storage system, on which we apply the file system. Nowadays, a typical HPC facility uses a small portion of the available nodes for storage purposes (I/O nodes acting as storage servers). Normally each storage server provides a huge number of hard disks through a RAID system. Current globally shared file systems, being deployed at the aforementioned facilities using current storage architectures, have several performance limitations when used with large-scale systems, because [2]:

(1) Bandwidth does not scale economically to large-scale systems.
(2)I/O traffic on the high speed network can be affected by other unrelated jobs.
(3) I/O traffic on each storage server can also be affected by other unrelated jobs.

The three (3) above problems are generally recognized as the most limiting factors for developing future exascale storage infrastructures. Exascale systems will require I/O bandwidth proportional to their computational capacity and it seems that current file systems and storage architectures will not be able to fulfill this requirement. One approach is to configure multiple instances of smaller capacity, higher bandwidth storage closer to the compute nodes (nearby storage) [2]. The multiple instances can provide exascale size bandwidth and capacity in aggregate and can avoid much of the impact on other jobs.

This approach does not provide the same file system semantics and functionality as a globally shared file system. In particular, it does not provide file cache coherency or distributed locking, but there are many use cases where those semantics are not required. Other globally shared file system semantics are required, such as a consistent file name space, and must be provided by a nearby storage infrastructure. In cases where the usage or lifetime of the application data is constrained a globally shared file system provides more functionality than the application’s requirements while at the same time limits the bandwidth which the application can use. Nearby storage provides more bandwidth, but without offering globally shared file system behavior [2].

The factors affecting performance are increasing if we consider the overall data flow (remote-local access) within an international collaborative scientific experiment, like the Large Hadron Collider (LHC) at CERN and KM3NeT. KM3NeT is a future European deep-sea research infrastructure hosting a new generation neutrino detectors that - located at the bottom of the Mediterranean Sea - will open a new window on the universe and answer fundamental questions both in particle physics and astrophysics.

This kind of experiments are generating datasets which are increasing exponentially in both complexity and volume, making their analysis, archival, and sharing one of the grand challenges of the 21st century. These experiments, in their majority, adopt computing models consisting of Tiers (each Tier is made up of several computing

Centers and provides a specific set of services) and for the different steps of data processing (simulation, filtering, calibration, reconstruction and analysis) several software packages are utilized. The computational requirements are extremely demanding and, usually, spans from serial to multi-parallel or GPU-optimized jobs.

In order to confront those challenges we introduced IKAROS as a framework that enables us to create ad-hoc nearby storage formations, able to use a huge number of I/O nodes to increase the available bandwidth (I/O and network) [3,4]. It unifies remote and local access in the overall data flow by permitting direct access to each I/O node, regardless of the tier. In this way we can handle the overall data flow at the network layer, limit the interaction with the operating system, and minimize disk and network contention.

2 Dissertation Summary

IKAROS provides a dynamically coordinated I/O architecture for I/O accesses according to the topology/profile of the infrastructure, the load metrics, and the I/O demands of each application. By referring to the I/O requirements/demands of the application we mean that IKAROS is not using a static/fixed algorithm for data placement. Due to the numerous configuration parameters offered the users and applications are able to choose the preferred strategy for each workload. In figure 1 we show an overview of the IKAROS framework: the input parameters used and the resources managed by IKAROS (I/O nodes and storage media) in all the tiers of the computing model.

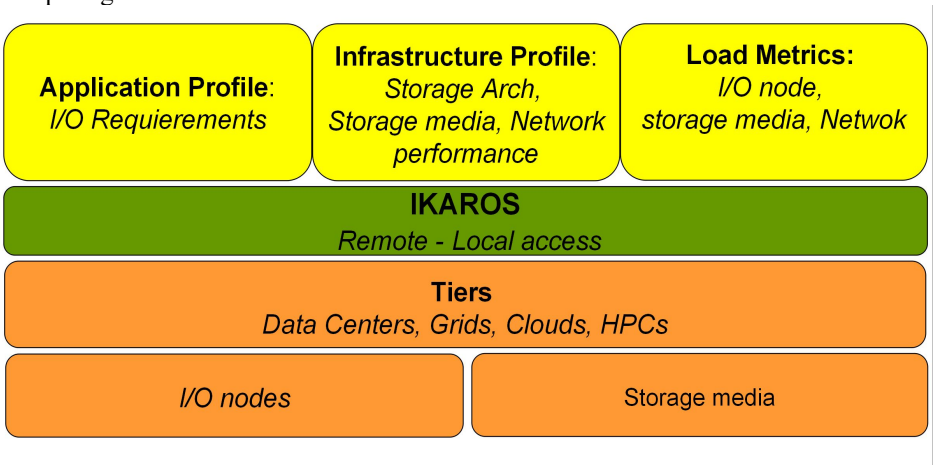


Fig 1. IKAROS Framework

Currently, we feed the appropriate input parameters at the IKAROS framework manually.

IKAROS allows data in a file to be striped across multiple disk volumes on multiple heterogeneous nodes and provides the utility for the storage system to access and

transfer a data file in parts and in parallel mode, without a specific order according to client requests. IKAROS defines three types of nodes: The User Interface (UI)/Client node, the Meta-data node and the I/O node. IKAROS first version was designed as an Apache Dynamic Shared Object (DSO) [3]. The latest versions are written in nodeJS, which provides more flexibility and interoperability with web 2.0 platforms. IKAROS node types are peers with the ability to act in any mode driven by client requests (i.e. any node can act at the same time as a Client/UI, meta-data node or I/O node).

The UI/Client node type is not a typical client but rather is more like a gLite UI [3]. This node type provides services to many users. The users are not forced to use the UI/Client node type. Alternatively, they can access IKAROS by using their own browser or any other HTTP client such as curl and wget. In a typical scenario, a user puts a request to his preferred IKAROS client (e.g: browser, other HTTP client, or UI/Client node) in order to read data from the storage facility. This request triggers the IKAROS module to interact with the meta-data node in order to fetch the necessary information regarding the file partition distribution schema. The client, then, establishes communication with the appropriate I/O nodes.

The IKAROS meta-data service (iMDS) holds a key role in the IKAROS architecture. The iMDS allows us to handle the meta-data sub-systems differently based on the needs. It may respond to a client/application request with three different ways. The client may find the answer: 1) within his own “cache”, 2) locally at a nearby iMDS utility or 3) at an external platform/utility. This approach, focusing on flexibility, can scale both up and down and so can provide more cost effective infrastructures for both large scale and smaller size systems.

Thus, we are able to use existing external infrastructures [5] (as the top level MDS utility, in a tiered system), such as Facebook and Gmail, in order to dynamically manage, share and publish meta-data. In this way we do not have to build our own utilities for searching, sharing and publishing. Additionally, we are enabling users to dynamically use the infrastructure, by creating on demand storage formations and virtual organizations [5].

2.1 IKAROS overall data flow scenario-write request (remote-local access)

In this scenario we analyze a data transfer from a remote storage server to the local parallel file system. In figure 2 we show the implementation of this action by combining GridFTP with the PVFS2. The client initiates a third party data transfer in order to transfer the data file from the the remote GridFTP server to the local parallel file system, in this case we are using PVFS2. We implement the local GridFTP server and the PVFS2 MDS at the frontend machine of the local computing cluster. Due to the client request, the remote GridFTP server starts sending the data file to the local GridFTP server by using N parallel data channels. The local GridFTP server moves the data chunks to the PVFS2 I/O nodes. The combined use of GridFTP with the PVFS2, for implementing the overall data flow, forces us to initiate many independent transfers incurring much overhead to set up and release connections. This

approach can significantly impact performance due to the unnecessary network and disk contention.

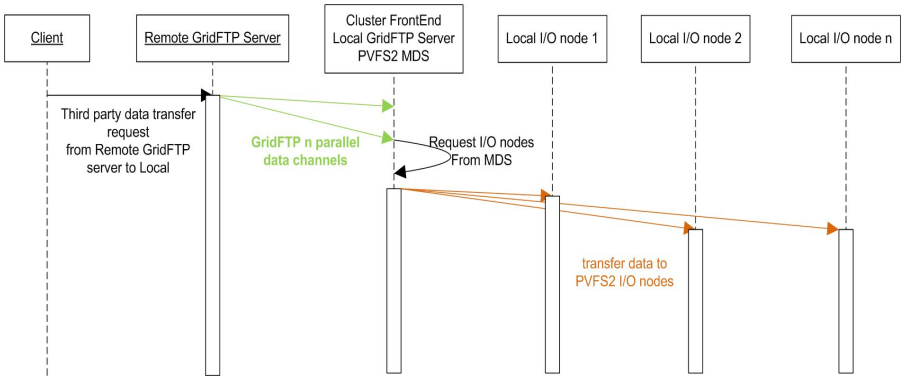


Fig 2. Overall data flow-write request, GridFTP + PVFS2

The network and disk contention mainly appears due to the lack of proper coordination between the two systems, GridFTP/PVFS2. There is no guarantee that the remote access protocol and the local parallel file system, in this case GridFTP and PVFS2, have the same stripe size and the same stripe mapping. At its latest versions PVFS2 provides these kind of information (stripe size and stripe mapping) so we may achieve the required synchronization manually, but this in not also the case for other parallel file systems like GPFS.

A solution could be to use the GridFTP striped server technique, which is not exactly the case we show in figure 2. In figure 2 we are using only one remote GridFTP server and the parallel data channels technique. In the striped server scenario the data file will be striped at several remote GridFTP servers and the local I/O nodes will have to act both as PVFS2 I/O nodes and local GridFTP servers. This means that we must assign public IP addresses to all the local I/O nodes, which in most cases is not desirable. We may bypass that by using a pNFS/PVFS2 combination instead of the GridFTP/PVFS2 combination, but still we will not be able to properly configure the local parallel file system on the fly. Another issue that we will have to consider is that there is no guarantee that the GridFTP servers will stripe the data across the data nodes in the same sequence as PVFS2 does across the I/O servers.

In figure 3 we analyze the same overall data flow scenario, a data transfer from a remote storage server to the local parallel file system, but now we are using IKAROS for the overall data flow in order to avoid the unnecessary network and disk contention. Techniques like the reverse read implementation of the write request, introduced by IKAROS, combined if necessary with reverse HTTP tunneling techniques can help us provide proper coordination between remote and local access and achieve better performance.

This approach allows us to mainly route the data at the network level and minimize the usage of the operating system. In figure 3 the IKAROS client follows the procedure demonstrated in [3] in order to trigger each local I/O node, which

participates in the transfer. Then each I/O node, based on this trigger, makes a request to the remote storage server for the corresponding data chunk. In this way we apply only coordinated parallel data transfers in contrast with the GridFTP/PVFS2 case where we must manually synchronize the stripe size and the stripe mapping between them.

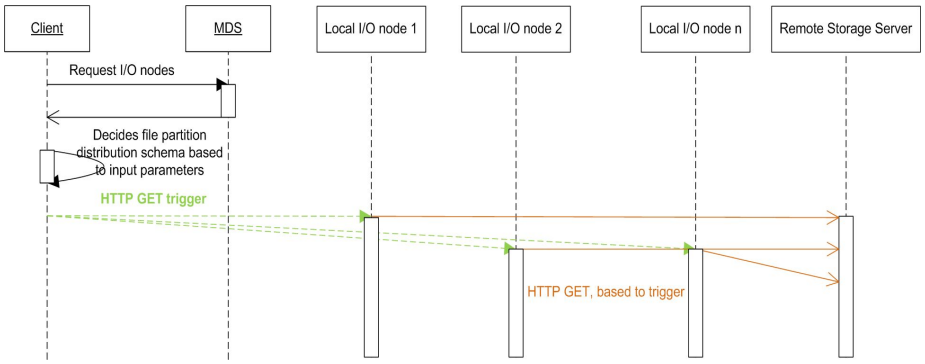


Fig 3. IKAROS overall data flow-write request

3 Results

For the evaluation we are using two different testbeds: the ZEUS computing cluster located at the National Center for Scientific Research "Demokritos" (soho-NAS environment) and the Cy-tera machine which is a Regional European HPC facility located at the Cyprus Institute (HPC environment). In figure 6 we show the two testbeds. The ZEUS computing Cluster provides an infrastructure composed of ten AMD Opteron CPU based systems (each with 8 CPU cores and 16 GB of RAM), four 800Mhz CPU based soho-NAS devices (each with 256 MB of RAM, 1000 Mbps Ethernet controller and 3 TB of storage capacity) and ten 200Mhz CPU based soho-NAS devices (each with 32 MB of RAM, 100 Mbps Ethernet controller and 2 TB of storage capacity). The ZEUS computer Cluster also provides a 1000 Mbps full duplex link between the nodes and an 2.5 Gbps WAN connectivity. In this study we are using the four 800Mhz CPU based soho-NAS devices and the frontend.

The Cy-tera machine consists of 100 nodes (96 Compute nodes) each with 12 Intel Xeon CPU cores, 48 GBs of RAM and one 15K rpms SATA local HDD. The nodes are connected over a QDR (40Gbit/s) infiniband. The GPFS file system is implemented by 4 storage servers. It is supported by 360 TBs raw disk space in 18 Raid-6 arrays each with 10 7200 rpms SATA HDD. The GPFS meta-data is provided by 4 Raid-10 arrays (one associated at each storage server).

As we mentioned before, its extremely important to obtain information like the storage architecture profile, the network topology, the I/O node and storage media performance in order to configure/customize properly each data transfer request. In

figure 4 we show the performance of writing an 80 GBs file from one node to 1,2,4 and 8 storage servers. All nodes are equipped with only one hard disk. In this scenario we are writing/partitioning the data file, which is located at the client side at 1 local HDD, to 1-8 storage servers (1-8 HDDs).

The purpose of this measurement is to help us make an estimation of the optimum load (read/write requests) that a hard disk can handle when we are transferring big data files that do not fit to memory. This estimation applies to both testbeds because the hard disks at the Zeus frontend and the disks at the Cy-tera computing nodes are having similar performance. We can use this estimation in order to decide the optimal number of the I/O nodes that the parallel file system should use in order to stripe/partition the data file.

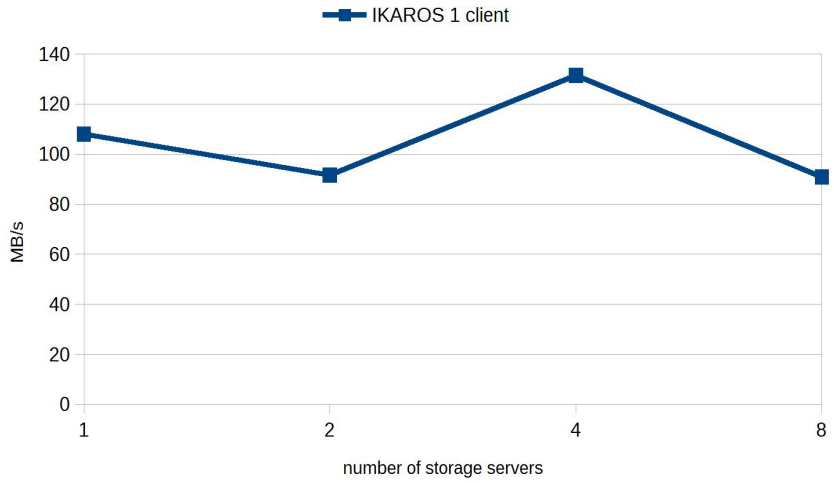


Fig 4. HDD Performance

Figure 4 shows that if the data file is located at a single disk, which is the case most of the times (the compute nodes used by the applications are equipped with a single disk), the best split ratio is 1:4. At the next section we show that if we maintain this ratio (4 dedicated HDDs serving each client for writing/partitioning the file) we can fully utilize the available I/O and network bandwidth.

In figure 4 we observe that by choosing an 1:4 ratio we achieve better performance than by choosing the 1:2 ratio. This is expected because most controllers and data media have queuing mechanisms, which when processing several parallel requests ensure under certain circumstances a higher performance than when processing fewer parallel or even only single requests. This, however, is always done at the price of higher response times. If many parallel requests only entail an increase in response time, and no longer in throughput, the disk subsystem is overloaded [6]. It seems that the 1:4 ratio is the most sufficient ratio for writing a big data file, located in a single disk, to the parallel file system.

Here we compare IKAROS with GPFS in data transfers with 80 GBs file size (in total) and multiple clients, using the Cy-tera machine. In figure 5 we measure the GPFS performance in the Cy-tera machine in order to properly choose the number of concurrent clients at the experiments in figure 6.

For a given GPFS file system, the most important factors affecting performance (aside from the access pattern) are the number of parallel processes participating in the transfers, and the size of the individual transfers. Figure 10 shows that we achieve the highest performance when the ratio of client processes to server nodes is near to 5:1 (though the nodes running our experiments have 12 processors per node, we ran only one client task per node). This ratio is slightly better but close to the 4:1 ratio measured at Lawrence Livermore National Laboratory at 2000 [7] by using 38 servers and 152 clients.

Fig 5. Cy-tera-GPFS Performance

In figure 5 we wrote a 80 GBs file, splitted into separate files, one file for each client process. As mentioned at [7] when the client:server ratio is too low the servers are starved for data; when it is too high the receiving buffers fill up faster than they can be drained, eventually causing packets to be dropped and retries initiated, reducing performance. From figure 10 we conclude that for the Cy-tera machine we must not initiate more than 20 simultaneously data transfers.

As we mentioned previously, the Cy-tera machine uses 180 hard disks distributed at the 4 storage servers in 18 Raid-6 arrays. This Raid-6 configuration in theory could provide a throughput close to 4200 MB/s which is way higher than the measured GPFS peak performance at the this machine (around 1600 MB/s). At the following experiments we show how IKAROS can fully utilize the available resources (I/O nodes and storage media) and achieve better performance.

The following experiments have been chosen because they show the effects of varying the I/O characteristics of application programs. We measured how aggregate

throughput varied depending on the number and configuration of client processes and the size of individual transfers. We also show how GPFS and IKAROS performance scale with system size and throughput of parallel tasks creating and writing a single large file, and of reading an existing file.

To measure the throughput of writes, the benchmark performs a barrier, then each task records a "wall clock" starting time, process 0 creates the file and all other processes wait at a barrier before opening it, then all processes write their data according to the chosen application characteristics (in the tests shown here, always independently of each other, filling the file without gaps and without overlap); finally, all processes close the file and record their ending time.

The throughput is calculated as the total number of bytes written in the total elapsed wall clock time (the latest end time minus the earliest start time). This approach is very conservative, but its advantages are that it includes the overhead of the opening and closing and any required seeks, etc., and measures true aggregate throughput rather than, for example, an average of per process throughput rates. For implementing IKAROS we are using the computing nodes and the local hard disks provided by them. All the nodes being used were on exclusive mode (only used by our process).

The results, for the GPFS, indicate the peak performance the file system is capable of delivering rather than what a user would see in the presence of other jobs competing for the same resources. The I/O performance seen in real applications will depend on complex interactions between the system and the application's run time behavior. Competition with other applications for I/O resources, wild access patterns, and some randomness in the file system can cause performance to be lowered. IKAROS framework enables us to create concurrent client requests through a coordinated I/O architecture in order to prevent I/O system contention [8].

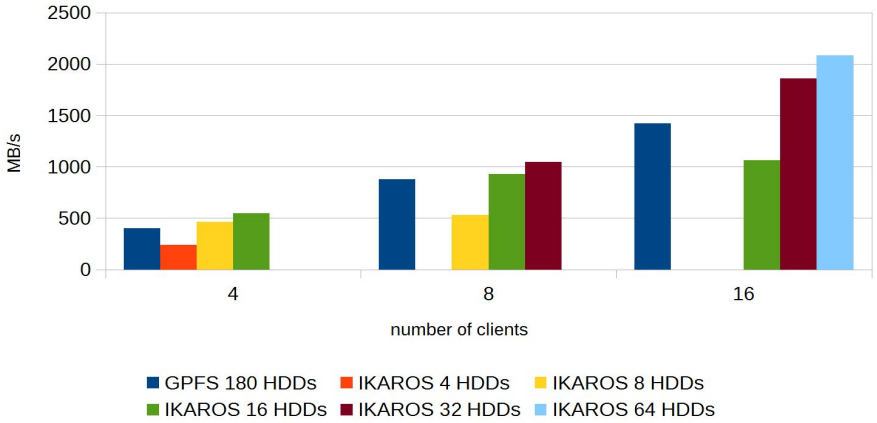


Fig 6. IKAROS vs GPFS by using several numbers of HDDs

In figure 6 we are using up to 16 concurrent clients due to the available nodes at the Cy-tera machine and because the measurements in figure 10 clearly shows that the

GPFS performance, for the current system, is seriously impacted by using more than 20 concurrent clients. Each time we measure the performance of GPFS versus IKAROS on 4:4, 4:2 and 4:1 write ratios (HDDs:clients). In this way we are able to create a virtual cluster of dedicated (4:1) or semi-dedicated (4:2, 4:4) storage facility for each client, where the I/O traffic can not be impacted by other client requests.

In figure 6 we clearly show that IKAROS can outperform GPFS by using 4:2, 4:1 ratios and that this approach can easily scale with only barrier the network bandwidth. IKAROS improves performance by 33% with the 1/3 of the available hard disks [8].

4 Conclusions and future work

This study proposes a dynamically coordinated I/O architecture for addressing some of the limitations that current parallel file systems and storage architectures are facing with very large-scale systems. The fundamental idea is to coordinate I/O accesses according to the topology/profile of the infrastructure, the load metrics, and the I/O demands of each application.

By using the IKAROS reverse read technique, for write operations, we are able to apply only coordinated parallel data transfers for the overall data flow (remote-local access). In contrast with other solutions (e.g: GridFTP/PVFS2 combination) where we are forced to initiate several independent data transfers incurring much overhead to set up and release connections. This can significantly impact performance due to the unnecessary network and disk contention.

The measurements have shown that by using IKAROS approach we can fully utilize the provided I/O and network resources, and minimize disk and network contention. We are able to achieve better performance by creating, on the fly, a virtual cluster of dedicated (4:1) or semi-dedicated (4:2) storage facility for each client, where the I/O traffic can not be affected by other client requests. In this way we managed to improve performance by 33% with the 1/3 of the available hard disks.

As a future work we intend to develop an automated system to feed IKAROS with the appropriate input parameter (figure 1). Probably, such a system could be part of an existing scheduling system like SLURM[46] and HTCondor[47] and or a Message Passing Interface (MPI).

References

1. Schmuck, F., Haskin, R.: GPFS: a shared-disk file system for large computing clusters. In: Proceedings of the FAST 2002 Conference on File and Storage Technologies. IBM Almaden Research Center, San Jose, CA (2002).

2. Dongarra, J., Beckman, P. et al. The International Exascale Software Roadmap," , Volume 25, Number 1, 2011, International Journal of High Performance Computer Applications, ISSN 1094-3420. Exascale Nearby Storage, Cray Position paper.
3. Filippidis C., Cotronis Y.,Markou C.,(2013) IKAROS: an HTTP-based distributed File System, for low consumption and low specification devices, Journal of Grid Computing springer.
4. Filippidis C., Cotronis Y.,Markou C.,(2012) Design and Implementation of the Mobile Grid Resource Management System. Computer Science, 13 (1). pp. 17-24. ISSN 1508-2806.
5. Filippidis, C., Cotronis, Y., Markou, C. (2014). Forming an ad-hoc nearby storage, based on the IKAROS and social networking services, IOP J. Phys.: Conf. Ser. 513 042018.
6. WHITE PAPER, FUJITSU PRIMERGY SERVER BASICS OF DISK I/O PERFORMANCE, http://sp.ts.fujitsu.com/dmsp/Publications/public/wp-basics-of-disk-io-performance-ww-en.pdf.
7. Terry Jones, Alice Koniges, and R. Kim Yates, Performance of the IBM General Parallel File System, Proceedings of the International Parallel and Distributed Processing Symposium, Cancun, Mexico, May 2000.
8. Christos Filippidis, Panayiotis Tsanakas, Yiannis Cotronis, IKAROS: a Scalable I/O Framework for High-Performance Computing Systems,The Journal of Systems & Software (2016),Volume 118, pp. 277-287, DOI:http://dx.doi.org/10.1016/j.jss.2016.05.027

Temporal Search in Document Streams

Dimitrios Kotsakos *

This PhD thesis addresses different challenges in searching temporal document sequences, where documents are created and/or edited over time, and the contents of documents are strongly time-dependent. Examples of temporal document collections are web archives, news archives, blogs, social networking platforms, and personal emails. The main focus of this dissertation is how to exploit temporal information provided in documents and combine it with textual information with the goal of improving the effectiveness of searching temporal document collections.

This summary describes the motivation and research questions addressed in the thesis. In addition, we explain our research context and methods. Our contributions to this thesis are composed of different approaches to solving the addressed research questions. In the end of this chapter, the organization of the rest of the thesis is presented.

1.1 Motivation

The ease of publishing content on social media sites brings to the Web an ever increasing amount of content captured during various types of events and/or before/after these events take place. Event content shared on social media sites such as blogs, Twitter, Facebook, YouTube, and others varies widely, ranging from planned, known occurrences such as a concert or a parade, to unplanned incidents such as an earthquake, floods or death of a celebrity. By exploring and proposing techniques to automatically identify and characterize these events and the relevant user-contributed social media documents (e.g., blog posts, photographs, videos, messages, status updates), we can enable rich search and presentation of all event content. In this dissertation we present approaches for leveraging the wealth of social media documents available on the Web for search purposes and content filtering and characterization.

In this work, we address major challenges in searching temporal document collections. In such collections, documents are created and/or edited over time. Examples of temporal document collections are web archives, news archives, blogs, personal emails and enterprise documents. Unfortunately, traditional IR approaches based on term-matching only can give unsatisfactory results when searching temporal document collections. The reason for this is twofold: the contents of documents and queries are strongly time-dependent, i.e., documents discuss events that took place at particular time periods, and a query representing an information need can be time-dependent as well, i.e., a temporal query.

---

* Dissertation Advisor: Dimitrios Gunopoulos, Professor

One problem faced when searching temporal document collections is the large number of documents possibly accumulated over time, which could result in the large number of irrelevant documents in a set of retrieved documents. Therefore, a user might have to spend more time in exploring retrieved documents in order to find documents satisfying his/her information need. A possible solution for this problem is to take into account the time dimension, i.e. extending keyword search with the creation or published date of documents.

1.2 Document Dating

During the recent years, the amount of user-contributed and digitized content on the Internet has dramatically increased, and makes web search even more challenging. Although well-known search engines (e.g. Google, Bing, etc) deliver very good results for pure keyword searches, they still do not take full advantage of the temporal dimension that characterizes most document collections.

However, in order for temporal text-containment search to give good enough and actually useful results, it is obvious that the timestamps of indexed documents have to be as accurate as possible. In the case of local document archives, trustworthy metadata that includes time of creation and last update is available. However, in the case of web search and web warehousing, having an accurate and trustworthy timestamp is a serious challenge. Another motivational example for research in the area of estimating a document's focus or creation time is that of old digitized documents or of partially failing optical character recognition applications (OCR). Moreover, a web page/document can be relocated and discovery time in this case will be very inaccurate. In some cases metadata about documents on the web can be retrieved but they can also in general not be trusted and often are simply just wrong. Thus, our research challenge is: for a given document with uncertain timestamp, can the contents of the document itself be used to determine the timestamp with a sufficient high confidence? To our knowledge, the only previous work on this topic is the work by de Jong, Rode, and Hiemstra [3], which is based on a statistic language model. In this work, we present approaches that extend the work by de Jong et al. and increases the accuracy of determined timestamps. Our main contributions are 1) a semantic-based preprocessing approach that improves the quality of timestamping, 2) extensions of the language model and incorporating more internal and external knowledge, and 3) an experimental evaluation of our proposed techniques illustrating the improved quality of our extensions.

Several related research efforts have focused on estimating a document's focus or creation time, mostly by the information retrieval community. Purely statistical methods have been proposed [1]. Other approaches have tried to deal with the problem by utilizing information from linguistic constructs with clear references to time periods or moments,

by mentioning, for example, a specific date or year. Another line of work cosiders the entire vocabulary used used in a document in order to reason about when it was created [2]. Kanhabua and Nørvåg in [3, 4] propose a document-dating method that extends the one proposed by De Jong et al. Specifically, the authors propose the application of semantic-based preprocessing of the reference collection, and apply a term-weighting scheme based on their previous work on temporal entropy [3]. The authors further enhance their approach by considering search statistics from Google Zeitgeist.

A serious drawback and disadvantage of most proposed methods, that is being addressed in this dissertation, is that most methods initially pre-segment the timeline of study into intervals of the same fixed length (e.g. a week) and afterwards choose the interval that is most likely to be the temporal origin of the query document, by comparing its vocabulary with the model built for each of the candidate intervals. The drawback of this approach is obvious: it limits the choices of possible time intervals.

1.2.1 Problem Definition

1.2.2 Preliminaries

The problem we address here is defined in the context of a collection of documents $\mathcal{D}$, spanning a timeline of $Y = t_1, t_2, ..., t_n$ of n distinct timestamps. We define a function $t(d)$ to return the timestamp of a given document $d \in \mathcal{D}$. Given a query document $q \notin \mathcal{D}$, for which the timestamp $t(q)$ is unknown, our goal is to find the best possible interval of size ℓ, $I = t_i, ..., t_{\ell+i}, 1 \leq i, j \leq n$ within T, so that $t(q)$ most likley falls within I. We refer to $\mathcal{D}$ as the *reference corpus*

Among other things, our approach considers the *burstiness* of the terms in the query document q. Given a term $x \in q$, we use $\mathcal{B}(x, \mathcal{D})$ to represent the set of non-overlapping bursty intervals for x, as computed over the given corpus $\mathcal{D}$. Each bursty interval is defined within the timeline T spanned by $\mathcal{D}$. In addition, we define $s(b)$ to return the burstiness score of a given bursty interval $b \in \mathcal{B}(x, \mathcal{D})$. Since we are only considering a single corpus, we henceforth refer to $\mathcal{B}(x, \mathcal{D})$ simply as $\mathcal{B}(x)$.

In order to evaluate the effectiveness of our method, we compare its precision against state of the art methods. Given the variety of our datasets as well as their uniform distribution of the time periods examined, we believe that the ability to place a document within a desired timeframe, in other words the precision we achieve, to be the most accurate valuation. At a high-level, the problem can be defined as follows:

Problem 1 [Document Dating]: *Let $\mathcal{D}$ be a collection of documents spanning a timeline of $T = t_1, t_2, ..., t_n$ of n discrete timestamps (e.g. days). Each document $d \in \mathcal{D}$ is associated with exactly one timestamp from T. Let $q \notin \mathcal{D}$ be a query document for which*

the timestamp is unknown. Then, we want to find we want to find the smallest possible timeframe within T during which the document was written.

The definition of the problem assumes that the query document was written within the timeline spanned by the corpus $\mathcal{D}$. No constraints are placed on the size or nature of $\mathcal{D}$. We observe that the presence of such a reference-corpus is necessary, otherwise it would be impossible to arbitrarily assign a timestamp to q.

1.2.3 Our approach

In this section, we introduce our algorithm for the Document Dating problem. As mentioned in the introduction of this chapter, our approach considers (i) the lexical similarity of the query document with the documents in the reference collection $\mathcal{D}$ (ii) the burstiness of the significant terms of the query document q, e.g., top-k terms ranked by *tf-idf*.

The use of lexical similarity captures the intuition that similar documents are more likely to discuss similar topics and events, and are thus more likely to originate in the same timeframe. In practice, however, similar documents may appear on different timestamps across the timeline. We address this, by introducing term burstiness. When an event or topic is recorded in a textual corpus, its characteristic terms exhibit atypically high frequencies. We refer to these timeframes as *bursty intervals*. Our algorithm is orthogonal to the actual mechanism used for computing non-overlapping bursty intervals. By identifying the bursty intervals of different terms, we can identify the timeframe of relevant events, as well as relevant documents that discuss them.

Figure 1.2 illustrates the architecture of our approach. In this section we describe our steps in detail.

We are given a query document q, discussing a disastrous fire in the Jackson theater in Chicago. The figure shows the 7 most lexically similar documents to q: $d_1, d_2, d_3, d_4, d_5, d_6$ and d_7 . Each of these documents has a subset of the following four terms in common with q: *Fire, Chicago, Jackson, Disaster*. The figure shows the bursty intervals for each of these terms. In this example, there are three visible sets of neighboring documents: $\{d_1\}$, $\{d_3, d_4, d_5\}$ and $\{d_5, d_6, d_7\}$. The documents in the second set overlap with multiple bursty intervals from the four characteristic terms, and are thus more likely to discuss the the actual event. Therefore, our approach will report the interval that starts with the first document on the (d_2) and ends with the last one (d_4) as the most likely timeframe for the query document.

We refer to our algorithm as `BurstySimDater`. The pseudocode is given in Algorithm 3.1.

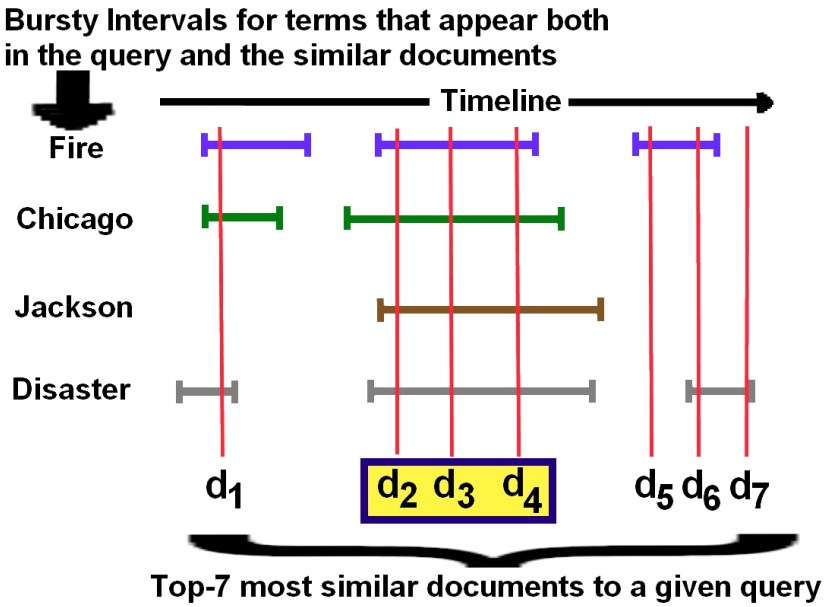


Figure 1.1: An example of how `BurstySimDater` identifies the appropriate timestamp for a given query document. In this case, the three documents d_2, d_3, d_4 will be selected by our algorithm, since they are both close to each other and overlap with multiple bursty intervals of the considered terms.

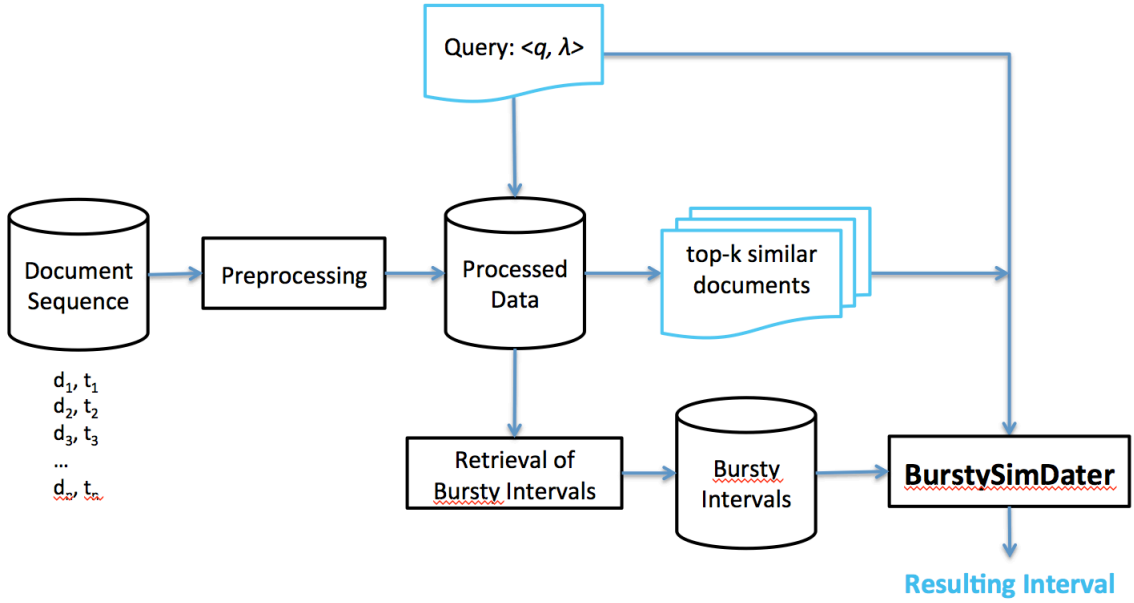


Figure 1.2: The architecture of our approach

The input to the algorithm consists of the query document q, the reference corpus $\mathcal{D}$, the

Algorithm 1 `BurstySimDater`

Input: reference corpus $\mathcal{D}$, bursty intervals $\mathcal{B}$, query document q, max timeframe length ℓ
Output: timeframe of q

1: $\mathcal{S} \leftarrow$ top-k most similar documents to q from $\mathcal{D}$
2: $W_{\mathcal{S}} \leftarrow \emptyset$
3: **for** $d \in \mathcal{S}$ **do**
4: $\quad w_d \leftarrow 0$
5: $\quad \mathcal{Y} \leftarrow d \cap q$
6: $\quad$ **for** $x \in \mathcal{Y}$ **do**
7: $\quad\quad w_d \leftarrow w_d + |\{I \in \mathcal{B}(x) : t(d) \in I\}|$
8: $\quad w_d \leftarrow w_d / |\mathcal{Y}|$
9: $\quad W_{\mathcal{S}} \leftarrow W_{\mathcal{S}} \cup \{w_d\}$
10: $A_{\mathcal{S}} \leftarrow (d \in \mathcal{S}, W_{\mathcal{S}})$
11: $\mathcal{I} \leftarrow \texttt{GetMax}(A_S, \ell)$
12: **Return** $\mathcal{I}$

set of precomputed bursty intervals $\mathcal{B}$ and the upper bound on the reported timeframe ℓ. The output is an interval of length at most ℓ, within the timeline T spanned by $\mathcal{D}$.

First, the algorithm retrieves the top-k most similar documents to q from $\mathcal{D}$. In our own evaluation, we experimented, among others, with the *tf-idf* measure and the Jaccard similarity. We use the latter in our experiments, since it led to the best results. We refer to the retrieved set of the k most similar documents as $\mathcal{S}$.

In steps 2-9, we assign a weight w_d to each document in $d \in \mathcal{S}$, based on its overlap with the burstiness patterns of its terms. Initially, w_d is set to zero. Let $\mathcal{Y}$ be the overlap of d's vocabulary with the vocabulary of the query document q. For each term $x \in \mathcal{Y}$, let $\mathcal{B}(x)$ be the precomputed set of bursty intervals for x. We then increment w_d by the number of the intervals from $\mathcal{B}(x)$ that actually contain $t(d)$. After the iteration over all terms in $\mathcal{Y}$ is complete, we normalize w_d by dividing it by $|\mathcal{Y}|$. Conceptually, the weight w_d of a document d is the average number of bursty intervals that it overlaps with, computed over all the terms that it has in common with the query q. The computed weights are kept in the set $\mathcal{W}_{\mathcal{S}}$.

We want to identify the interval when the most terms from the top-k similar documents are *simultaneously* bursty. This period is the intersection of intervals with the maximum sum of weights. To do this, in steps 10-11, we create an array $A_{\mathcal{S}}$ of size T, where cell i equals to the sum of weights w_d for all documents $d \in \mathcal{S}$ that were written at t_i. Next, we find the interval $\mathcal{I}$ of length ℓ with the maximum sum. By tuning ℓ, we tune the level of desired accuracy. In order to compute the sets of bursty intervals we use the `GetMax` algorithm [6]. Given a discrete time series of frequency measurements, `GetMax` returns a set of non-overlapping bursty intervals with respect to the frequencies.

A *burst* on the timeline is marked whenever the popularity of a specific term dramatically and unexpectedly increases. In order to compute the sets of bursty intervals we use the `GetMax` algorithm, introduced in [5]. Given a discrete time series of frequency measurements for a given term, `GetMax` returns a set of non-overlapping bursty intervals with respect to the number of frequency measurements.

This chapter reviews the literature on the document timestamping problem and proposes a new approach for document dating that overcomes the drawbacks of previous methods: it doesn't depend on temporal linguistic constructs and it can report timeframes of arbitrary length. The proposed method outperforms the previous state-of-the-art in precision and computational efficiency in most of the cases, while being the most versatile of all, since it performs well for many reporting intervals and throughout a variety of datasets. This is achieved by taking into consideration the burstiness of the terms and lexical similarity of testing documents with the timestamped training corpus. An extensive experimental evaluation on real datasets demonstrated the efficacy of the algorithm and its advantage over the state of the art.

1.3 Memes and Events

As a motivational example, consider the part of the homepage of many social media sites that is devoted to *Trending Topics*. Most modern platforms like Twitter, Yahoo!, Facebook, etc. offer such a functionality, where different types of algorithms are used to identify the most popular topics in the platform during a current time window. *Trending Topics* lists may include popular items people search for in an e-commerce site like Amazon.com, trending queries in a web search engine like Google.com, popular topics of interest people write about in a micro-blogging platform like Twitter.com or trending tags people use to annotate their blog posts in a blogging site like Wordpress.com. Most of the items that appear in this lists have been caused by real-life events that triggered the interest of the users and they wrote or posted about them in the social media. They main functionality of the *Trending Topics* lists is that of facilitating search and discovery of new content. However, not all trending items are related to real life events. A significant percentage of popular content has become strategically popular, especially within microblogging environments like Twitter and Tumblr, where fan- and sports-related communities thrive and dominate the usage of the media. Thus, social media platforms and search engines would benefit from better understanding *why* a specific item became popular, and offer a variety of landing pages for different types of content. Specifically, Figure 1.3 illustrates the trending topics on December 28, 2014 in Twitter and Yahoo! respectively. A user clicking on a news-related trending topic would expect to land on a page with a series of news articles, maybe chronologically ordered, describing the timeline and the current state of that specific topic. On the other hand, a user clicking *'iPhone 6 Plus'*, which is a consumer product would expect to find offers, technical specs or reviews of the item. Last, users interested in a celebirty named *Nash*, would expect to find fanpages, photos and videos of the celebrity when clicking on `#HappyBirthdayNash` trending topic.

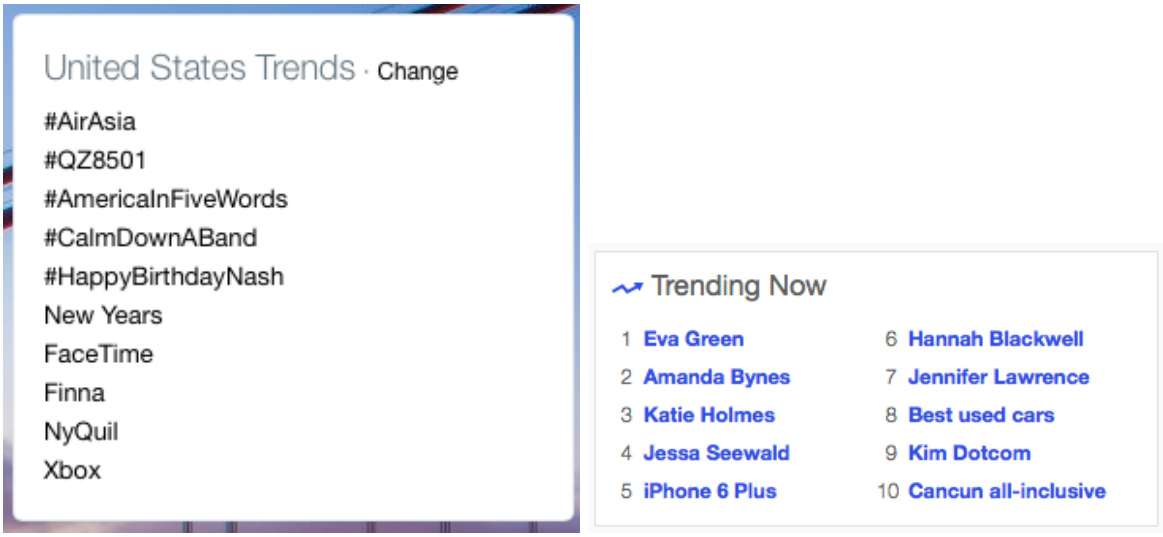


Figure 1.3: Trending Topics in Twitter.com and Yahoo.com on *December 28, 2014*

For our analysis of trending topics, we specifically focus on one social media site, namely, Twitter, due to its transient, large-scale publicly available content. In particular, we collected a vast amount of Twitter messages from various locations posted during various periods of time and focus on the analysis of the most popular hashtags. We show that event-detection methods can not only be based on the detection of terms and phrases that exhibit trending behavior on Twitter, characterized by an unusual increase in message frequency during a particular time period in a Twitter message stream. While some of these trends might refer to actual real-world events, others might include non-event information, triggered by strategically planned advertising campaigns or by user communities trying to promote themselves. Unlike related efforts in this area, which focused on characterizing or analyzing content from individual events on Twitter, or characterizing aggregate trend characteristics for manually identified terms, the features we use in this study in order to discriminate between the various classes of content can be used for more than the two classes that are defined in this dissertation.

Overall, we show that social media sites contain substantial, useful information about different types of popular content that can be exploited and utilized in order to provide more effective and useful services to users of social media sites. With the features we propose in this dissertation, we can effectively identify different types of popular content and their associated social media documents across various social media sites. Regardless of the classifier we use, the type of event/meme, or the social media site, any single popular topic might have hundreds or thousands of associated social media documents. While some of these associated documents might contain interesting and useful information (e.g., event time and location in case of events, participants and opinions in case of memes), others might provide little value (e.g., using heavy slang, incomprehensible language without ac-

companying media) to people interested in learning about an event or meme. Techniques for effective selection of quality event content may then help improve applications such as event browsing and search. Therefore, we propose a noise-filtering mechanism for selecting a subset of the social media documents associated with the significant real-world events.

1.3.1 Hashtag-Based Event Detection: A Proof of Concept Use Case of Meme-Filtering

In order to show the utility of the proposed methodology we applied a hashtag-based event detection approach to our data. As mentioned in the introduction, due to the limited text length of tweets, hashtag analysis is a common approach for micro-blogs mining. For example, most event detection methods in social media rely on time-series analysis of hashtags, inspecting terms that appear bursty for specific time-periods or generally popular terms. The assumption is that the bursty keywords/hashtags will be related to emerging events.

In this section, we argue that these event detection methods can lead to mixed results, since memes are also popular and bursty. In fact, memes appear to have a very well defined popularity period, just like events, so time-series approaches will fail distinguishing one from the other. We applied a burstiness algorithm for event detection in order to study the insufficiency of this type of approach. At a high level, a time-frame is considered bursty if the term exhibits atypically high frequencies for its duration. Bursts in terms of frequency capture the trends in vocabulary usage during each corresponding time-frame and can thus prove useful in event detection. When an event takes place in real life (e.g. an earthquake, sports finals), the event's characteristic terms (e.g. *earthquake, shooting, overtime*) appear more frequently in social media. Unfortunately, memes demonstrate a similar behaviour.

1.3.2 Burstiness Results

In our experiment we split the *Germany* dataset in two sets, one including months *April, June, July* and *August* which served as the training set and one including only *September* which was our testing set. Table 1.1 lists some bursty intervals computation examples along with a short description for the corresponding hashtags. The last column shows the classification result when using meme filtering with the Random Forest classifier, trained over labeled data from the first four months of the dataset. It is apparent that while the bursty intervals computed by `GetMax` algorithm precisely match the actual dates of excessive popularity of the corresponding hashtags, it is not enough to reason about significant real life events that affected the Twitter community. Hence `GetMax` identifies memes and events.

Table 1.1: Bursty Intervals for popular hashtags in *Germany* during September, 2014

hashtag	Bursty Intervals	Description	Meme Filtering
#ff	Sep 5, 12, 19, 25	"Follow Friday" Twitter meme	meme
#eaie2014	Sep 16 - Sep 19	Conference held in Prague during Sep 16 - Sep 19	event
#jaykingslandto60k	Sep 11 - Sep 12	Bot account posting thousands of tweets	meme
#nominateavrillavigne	Sep 11, 15	Celebrity fan campaign	meme
#h96hsv	Sep 14	Soccer match: Hannover 96 vs. Hamburger SV	event
#iphone6	Sep 9, Sep 19	Announcement and release of iPhone 6	event
#iphone6plus	Sep 9-10, 19	Announcement and release of iPhone 6 Plus	event

A closer look at Fig. 1.4 reveals an even further similarity between the different types of popular hashtags in terms of behavior in time. On *Friday, September 19* four hashtags exhibit similarly bursty behavior, being simultaneously and unexpectedly popular. Two of them, namely #iPhone6 and #iphone6Plus, correspond to the event of the release of the new iPhones, #eaia2014 is the hashtag used to annotate discussions and reports from the European Association for International Education held in Prague, while the #ff is a viral Twitter meme with the aim of suggesting people for other users to follow. While the reader would argue that the distinction between an event and a meme in this case is rather trivial, since #ff is a periodically popular hashtag, this is not the case with hashtags like #nominateavrillavigne that have similarly bursty behavior, but only not periodic.

In this chapter we defined the problem of distinguishing a popular topic of interest in a social network between network-generated topics of discussion, denoted as *Memes* and real-life events that triggered the interest of the social network users, denoted as *Events*. We provided a detailed study of the features that affect the classification, applying our experiments on the Twitter network using two different real-life datasets with 27.8 and 6.8 million tweets each and 1.1 million and 491,043 unique hashtags respectively. We evaluated multiple classification methods, among of which the Random Forest classifier performed always best, having been able to reach an accuracy of 89% in its prediction on whether a topic is a *meme* or an *event*. Our study reveals interesting characteristics of the two classes of hashtags, some expected and some not. To demonstrate the utility

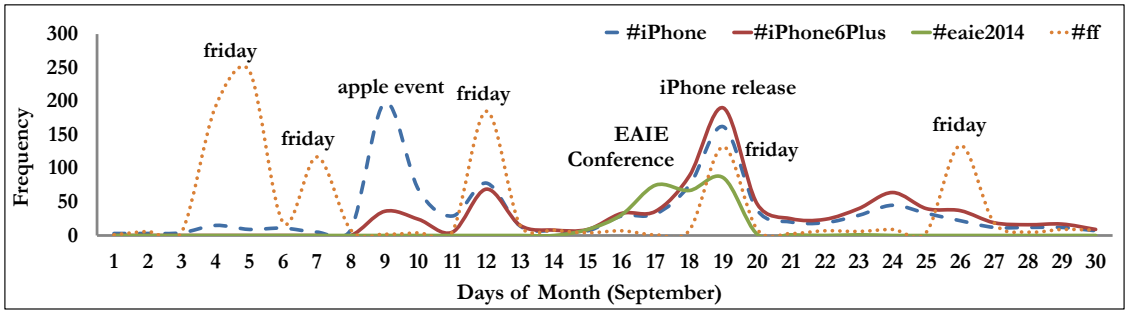


Figure 1.4: Frequency curves of popular hashtags of various kinds as they appeared in *Germany* during *September, 2014*

of our approach we enhance a hash-tag based event detection with meme-filtering and comment on the improved results.

1.4 Contributions and roadmap of this thesis

This dissertation addresses research problems in searching temporal document collections. We have proposed different approaches to solving the addressed research questions. In summary, the contributions of this thesis are:

- We exploited term burstiness in order to detect events in social media document streams.
- We proposed a state of the art technique for for determining the creation time of non-timestamped documents. The proposed approach outperforms the methods of the relevant literature. We improved the quality of document dating by incorporating term burstiness information and textual similarity methods into the algorithm. By conducting extensive experiments, we showed the evaluation of our proposed approach and the improvement over the baseline.
- We formally defined the difference between memes and events in social media and thoroughly examined the differences between the two different types of popular content along various descriptive characteristics, proposing a set of features to aid the classification of various types of content. We showed the usefulness of our method via a burstiness-based event detection approach.

BIBLIOGRAPHY

[1] Nathanael Chambers. Labeling documents with timestamps: Learning from their time expressions. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 98–106, Jeju Island, Korea, July 2012. Association for Computational Linguistics.

[2] F.M.G. de Jong, H. Rode, and D. Hiemstra. Temporal language models for the disclosure of historical text. In *Humanities, computers and cultural heritage: Proceedings of the XVIth International Conference of the Association for History and Computing (AHC 2005)*, pages 161–168, Amsterdam, The Netherlands, September 2005. Royal Netherlands Academy of Arts and Sciences. Imported from EWI/DB PMS [db-utwente:inpr:0000003683].

[3] N. Kanhabua and K. Nørvåg. Improving temporal language models for determining time of non-timestamped documents. *Research and Advanced Technology for Digital Libraries*, pages 358–370, 2008.

[4] Nattiya Kanhabua and Kjetil Nørvåg. Using temporal language models for document dating. In *ECML/PKDD (2)*, pages 738–741, 2009.

[5] T. Lappas, B. Arai, M. Platakis, D. Kotsakos, and D. Gunopulos. On burstiness-aware search for document sequences. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 477–486. ACM, 2009.

[6] Theodoros Lappas, Benjamin Arai, Manolis Platakis, Dimitrios Kotsakos, and Dimitrios Gunopulos. On burstiness-aware search for document sequences. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 477–486. ACM, 2009.

Credible Reputation Systems for P2P e-Communities

Eleni Koutrouli[1]

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
ekou@fi.uoa.gr

Abstract. Reputation mechanisms for distributed e-Communities are vital tools for facilitating trust decisions regarding transactions between entities. Motivated by the current challenges in the area of P2P reputation systems regarding their desgign, credibilty enhancement and objective evaluation, in this thesis we worked towards (1) creating a framework for the development and evaluation of secure reputation systems, and (2) designing and evaluating a credible reputation system for P2P communities with incentives for honest recommendations. We have thus created a conceptual model and a credibility framework for the design of credible reputation systems. We also proposed an evaluation framework for reputation systems for their objective evaluation and comparison. We then developed a credible reputation system (CREPARS) which consists of (a) credibility-enhanced reputation estimation algorithms and processes and (b) a novel recommendation exchange mechanism which is based on recommendation trustworthiness of entities and uses a PKI-based payment scheme. For the evaluation of the proposed reputation system we used credibility analysis and simulation in various attack scenarios and in comparison with other well-known reputation systems. The results have shown that the proposed reputation mechanisms exhibit resilience to various attacks and offer incentives for honest recommendations, leading to increased efficiency.

Keywords: reputation systems, trust, credibility, threat analysis, evaluation of reputation systems, simulation of reputation systems, trust management

1 Dissertation Summary

Contemporary e-Communities have emerged in various application and technological contexts. One of their basic characteristic is the need of their users to be supported in their decision about other users and objects which they have to trust for their transactions. Efficient Reputation Systems (RSs), which integrate the concepts of trust and reputation and support trust decisions in applications for distributed e-Communities, have become vital components for these applications. The systematic

[1] Dissertation advisor: Aphrodite Tsalgatidou, Associate Professor

study of RSs, with a focus on P2P RSs, has revealed a number of issues which impede their efficiency and consequently the efficiency of the application they support. These challenges, specifically the lack of (a) reference reputation models that could facilitate their design, (b) a comprehensive threat analysis and (c) methods and frameworks for the objective evaluation of P2P reputation systems and their comparison, have motivated us towards defining the goals of this thesis, which are the following:

1. Creation of a generic framework for the development and evaluation of secure reputation systems
2. Development of a secure and credible reputation system for P2P e-Communities with incentives for honest recommendations based on the defined generic framework
3. Evaluation of the efficiency and resilience of the proposed reputation system against various attacks and various forms of malicious behavior, in comparison with other RSs, based also on the developed generic framework.

For the satisfaction of the first goal we created a *conceptual model* for the design of reputation systems, a *credibility framework* for the integration of credibility factors in a reputation system, and an *evaluation framework* for the evaluation of reputation systems through suitable methods or through a common evaluation and comparison framework. For the satisfation of the second and third goals we used the proposed conceptual framework for the design, implementation and evaluation of a credible RS for decentralized e-Communities with incentives for honest recommendations. These results are described in Section 2, whereas the main contributions of the thesis are summarized in Section 3.

2 Main Results

2.1 Conceptual Model for P2P Reputation Systems

In order to to facilitate the design of decentralized RSs we present a reference reputation system for P2P RSs, which comprises the concepts, roles, relationships, functionality and design characteristics of such RSs. A detailed description can be found in [12].

Reputation systems use information related with the transactional behavior of entities for the estimation of their reputation and consequently for making trust decisions. They are based either on a centralized structure (e.g. eBay [1]) or on decentralized structures (e.g. [2]-[4]), found mainly in P2P systems, where reputation management is distributed to the participating entities.

In a decentralized RS the participating entities play interchangeably the roles of the *trustor*, the *trustee* and the *recommender*. The trustor is an entity which wants to make a trust decision regarding whether to participate in a transaction with another entity, the trustee. A transaction can involve accessing a resource, an e-Commerce trade, etc. The recommender is the entity that provides the trustor with information regarding the trustworthiness of the trustee (recommendation). In file sharing P2P applications, recommendations may also be given for objects, e.g. files. To make a trust decision the trustor tries to predict the future behavior of the trustee by forming a view of the trustee based on experience about its earlier actions. This subjective view is formed by

estimating an indicator of the quality of the trustee regarding its services and comprises the trustee's *reputation* or *trustworthiness* from the trustor's point of view. To form a reputation view, the trustor needs to gather experience information, either by referring to its own earlier experience with the trustee, or by acquiring it from other entities in the form of recommendations. Recommendations can be based on the recommender's personal experience alone, or on a combination of personal experience and earlier recommendations from others. The various roles of the participating entities in a decentralised RS are illustrated in the UML diagram of Figure 1.

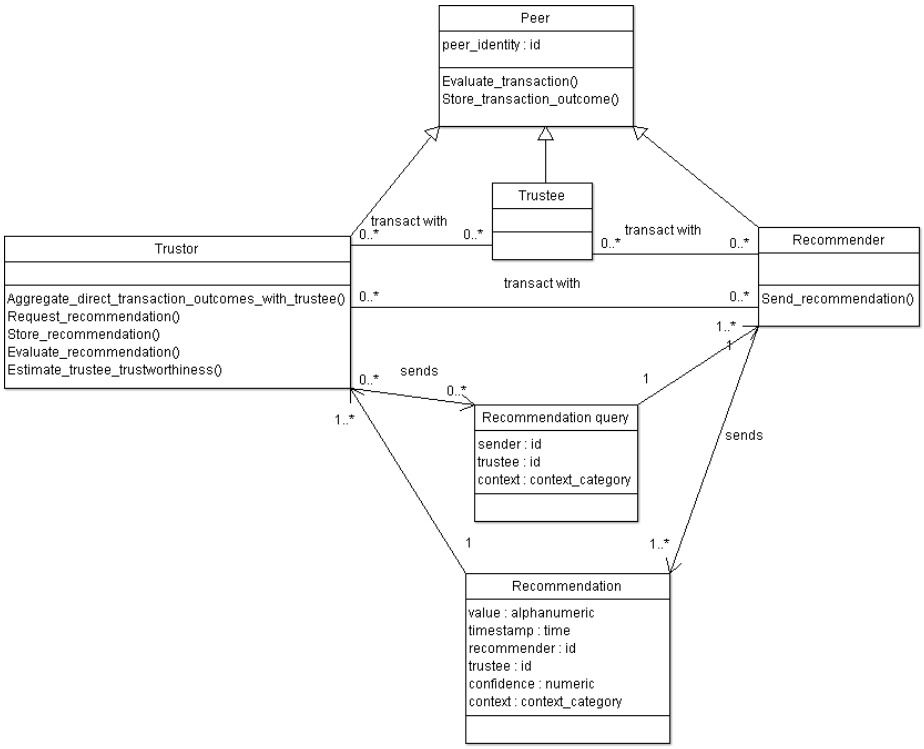


Fig. 1. Conceptual Representation of a Decentralized Reputation System

We outline below the basic characteristics) of a RS (presented also in [5]) for which various design choices can be made in order to cover the related requirements.

1. **Recommendation Content and its Representation**. A recommendation can be an arithmetic value, a combination of a value and associated semantic information, such as confidence or context, etc. Various formats can be used, such as binary, scalar or continuous values in a specific interval.
2. **Recommendation Formation**. It can be done based on the evaluation of a single transaction or on aggregated ratings regarding transactions with the the trustee.
3. **Selection of Recommenders**. This can be done based on recommenders' credibility, on social relationships, on recommendation similarity of the recommender and the trustor regarding commonly evaluated peers, etc.

4. **Reputation Estimation**. As described in [6], reputation estimation approach can be either deterministic, probabilistic or based on fuzzy logic.
5. **Storage and Dissemination of Reputation Information**. Reputation values may be estimated either reactively or proactively. They are stored by the trustor or the trustee or by other special peers. Their communication to the interested parties is done either upon request, or using a disseminating technique.
6. **The Way a Trust Decision is Made**. Trust decisions are threshold-based or rank-based; they are based on the estimated reputation values, therefore, the latter should be translated in a manner that facilitates trust decisions.

2.2 Taxonomy for RSs for Social Network (SN)-based applications

In SN-based applications the concept of reputation is expanded to new meanings, such as "user influence", and RSs use various indirect mechanisms, i.e. mechanisms which are based on social network-related information, rather than ratings. We have thus proposed a taxonomy for such RSs based on their identified dimensions. This taxonomy can be used for the classification of RSs for various types of SN-based applications and for facilitating the design of a RS for a particular SN-based application [7].

2.3 Credibility Framework and Threat analysis of Decentralized RSs

The accuracy of reputation estimation, and thus the credibility of a RS, are affected by a number of factors which we present, grouped in three categories, in Table 1:

Table 1. Credibility Factors of a Reputation System

Factors related to **Recommendation Creation/Content**	Factors related to **Recommendation Selection**	Factors related to **Reputation Reasoning**
Type of recommendation information (value, statement, etc.)	Recommender's credibility	Aggregation method (estimation formula, recency considerations, reputation value translation)
Creation method (transaction rating or opinion)	Uncertainty awareness	History of transactions and recommendation information
Type of experience (negative and/or positive) evaluated in a recommendation	Recommender selection method, considerations about possible bias or pressure	Storage and dissemination methods for reputation values
Recommender's identity	Storage and dissemination methods for recommendations, considerations about possible bias or pressure	Evaluation of estimated reputation
Recommender's confidence on recommendation	Mediator's credibility	Secure storage and retrieval of global reputation values
Binding recommendations with transactions	Who collects recommendations, possible bias	

Entities participating in reputation systems can distort the credibility of the latter in various ways, either as individuals or in cooperation with others, depending on the specific application and social setting of the reputation system. We have classified reputation attacks or misbehavior in the following three main categories:

- **Unfair recommendations**: Entities can spread unfair ratings for other entities in order to lower or increase the reputation of the target entities unfairly. Unfair ratings can be due to lying, misjudging the outcome of a transaction, or making a mistake in the recommending procedure.
- **Inconsistent behavior**: Peers may strategically have an inconsistent behavior that can lead to an incorrect estimation of their reputation allowing them to misbehave and still keep a high reputation. For example, they can misbehave part of the time or towards a subset of peers or change their behavior suddenly or periodically.
- **Identity management related attacks**: A deciding factor for attacks in this category is the identity scheme used in a RS. For example, when the identity scheme permits the use of multiple identities by the same peer, a malicious peer may behave dishonestly and then escape its low reputation by entering the system with a new identity. Furthermore, when an entity A can communicate or store a recommendation produced by an entity B for an entity C without linking its identity and B's identity with the recommendation, then A can easily manipulate the recommendation value. Also, if the system permits it, peers may refuse having sent a recommmendation.

A detailed taxonomy of the attacks against RSs is depicted in Figure 2, and thoroughly described in [10], together with a detailed presentation of the related defense mechanisms. The identified defense mechanisms have been then mapped with the attacks which they confront and with the specific categories of credibility factors to which they belong [10]. This mapping can be used as a guide for the implementation of suitable defense mechanisms in the process of designing a RS.

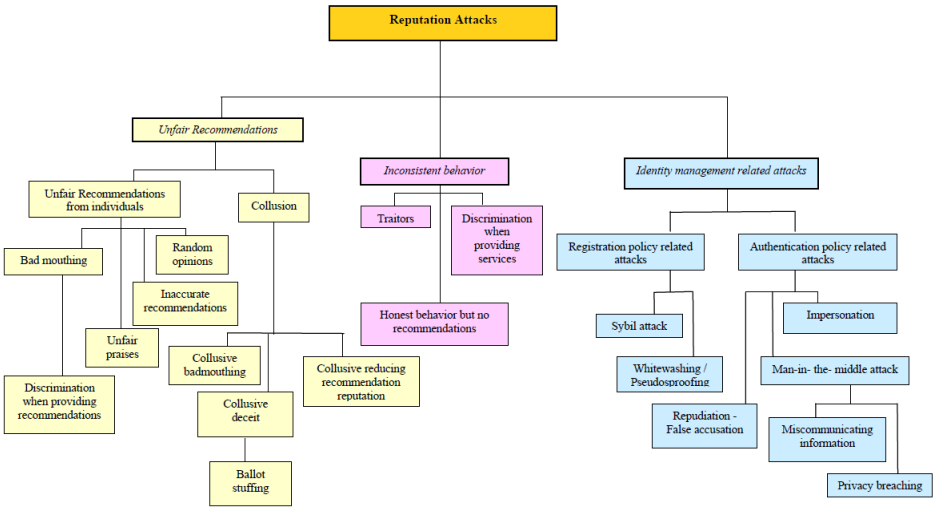


Fig. 2. Taxonomy of Attacks against Reputation Systems

2.4 Evaluation Framework for Reputation Systems

The plethora and heterogeneity of works regarding RSs for various e-Communities creates the need for objective evaluation and comparison between different RSs under the same conditions. Most of the evaluation approaches used in the proposed reputation systems are either proprietary or common experiments under restricted cases. However, the emerged need for generic evaluation approaches led to a number of research works which focus on the development and use of generic frameworks for evaluation and comparison of reputation systems, to which we refer as Common Evaluation Frameworks (CEFs). These works are either theoretic, i.e. they study how a reputation system deals with a number of criteria or attacks, or offer simulation and implementation platforms / tools for the evaluation, comparison and fine-tuning of reputation systems through experimentation. We have classified the various available approaches for RS evaluation according to the taxonomy presented in Figure 3 [11].

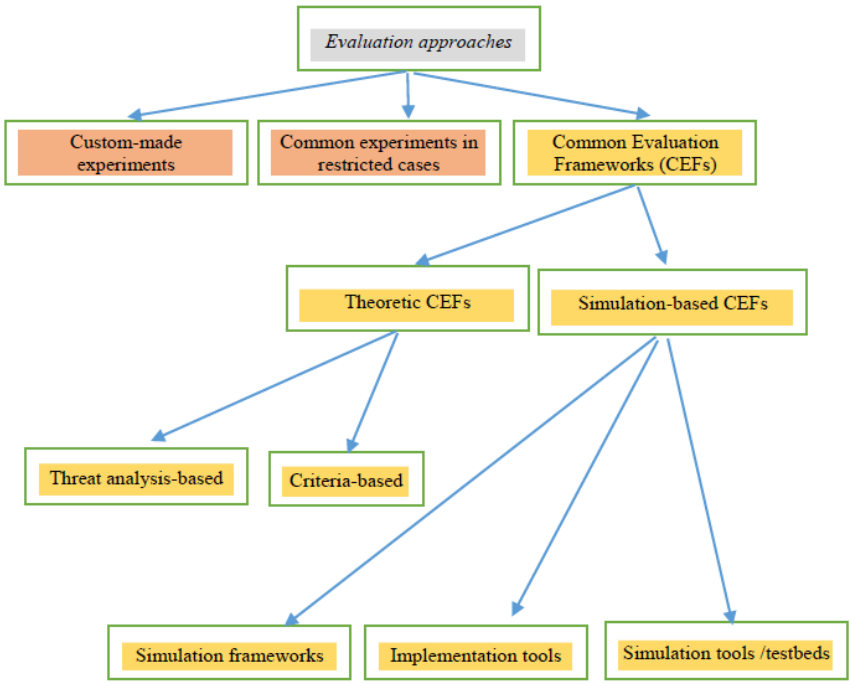


Fig. 3. Taxonomy of Evaluation Approaches

We have focused on works offering CEFs and we have formulated a set of characteristics / properties that are desirable for a generally accepted CEF in order to produce reliable comparisons, as follows: (1) *Standardization*, (2) *Independence of the reputation system characteristics*, (3) *Flexibility*, (4) *Ease of implementation of new reputation systems and new tests*, (5) *Availability of existing implementations of reputation system tests*. In [11] we present the level of conformance of a number of

simulation-based CEF found in the literature to the desirable characteristics, according to the information provided by the authors. We have also defined a number of factors that affect the desirable characteristics. The proposed evaluation framework facilitates (a) finding suitable evaluation methods/CEFs and (b) defining generally accepted CEFs.

2.5 Credibility-enhanced & Payments-based Reputation System for Decentralized Systems (CREPARS)

The proposed reputation system aims at providing credible reputation estimation with incentives for honest recommendations (ratings) and excibiting thus resilience to various attacks. It comprises (a) a reputation model which involves the algorithms for the estimation of the various reputation components and the final reputation value which is based on these components, and (b) a recommendation exchange mechanism which is based on virtual payments and a Public Key Infrastructure (PKI).

Reputation Model. The proposed RS estimates an *overall reputation* value for the trustee which comprises:

a) the *direct reputation* of the trustee from the point of view of the trustor, which is the time weighted average of the transaction evaluation values regarding the direct transactions between the trustor and the trustee.
b) the *indirect repuation* value of the trustee, which is based on third parties' recommendations.

Together with direct reputation, a *confidence factor* is estimated, which takes into considertion the number of direct transactons, the deviaton of the direct transaction evaluation values and the timestamp of the last transaction. A recommendation is the direct reputation estimated by the recommender for the trustee and is provided together with the related confidence. When a transaction takes place, the trustor evaluates the transaction and updates the *recommendation trustworthiness* of the recommending entities, based on the divergence between the transaction evaluation and the provided recommendations. Indirect reputation is estimated as a weighted average of the recommendations, where each recommendation is weighted with the related confidence value and with the recommendation trustworthiness of the recommender. The proposed reputation estimation process, comprising the involved activities and the estimation formulas, are thoroughly presented in [12].

Evaluation. For the evaluation of the proposed reputation model we used the reputation systems simulator TRMSim-WSN [13]. We implemented our model in the simulator and evaluated it using four scenarios (static network, dynamic network, oscillating behavior and collusive bad-mouthing). For each scenario specific network properties and attacks with different percentages of malicious users were simulated. Our model was compared with four other reputation systems (EigenTrust [8], PeerTrust [3], PowerTrust [14] and BTRM-WSN [15]) which are reference reputation systems in the literature. The evaluation metrics that were estimated are the following: (a) *Accuracy* of

the model, i.e. the percentage of the successful selections of honest providers in the all provider selections, and (b) *Average path length*, i.e. the number of the intermediate nodes between the client and the selected service provider, as a performance indicator.

The simulation results show that the proposed system behaves efficiently in all the examined scenarios. It has a scalable performance in static networks, where the number of nodes increases and the number of malicious nodes is 70%, while in dynamic networks, where the topology of the network or the behavior of the nodes changes, the simulation of the proposed system has good results even if the percentage of malicious nodes is quite large. The good performance of the proposed reputation metric is attributed to the integrated credibility factors, namely the *recommendation reputation* of the recommenders, the *time decay function* that is used for weighting recommendations, the estimated *confidence factor* which is attributed to a recommendation and to direct reputation values and which takes into consideration the *number of transactions* and the *deviation of transaction evaluation values*, and to *weighting direct reputation more highly* than indirect reputation in the final reputation estimation. Our various experiments verify the resilience of the proposed reputation model against bad-mouthing, osclillatory behavior and traitor's attack.

CRedibility Enhanced Payments Scheme (CREPS). The proposed reputation system CREPARS involves also a payments scheme for the recommendation exchange, which gives incentives for honest recommendations. The goal of this mechanism is to provide resilience against sybil attack, repudiation, badmouthing and recommendation free-riding. Peers which participate in CREPS have *virtual accounts* and use them to make payments for acquiring recommendations. A peer A (recommendation buyer) which wants a recommendation from a peer B (recommendation seller) pays a value v for it to B. The value depends on both the recommendation reputation values that A and B have estimated for each other (RecRepA(B) and RecRepB(A)) according to the following formula:

$$v = \frac{RecRep_A(B)}{RecRep_B(A)}$$

Each entity has an *Initial Account Balance* for its participation in CREPS. After a recommendation exchange, the account balances of the participating entities are upadated (credited / debited). In order for an entity to participate in CREPS, her recommendation reputation should be higher than a minimum value which is defined according to a threshold value (t_{seller}, t_{buyer} for the recommendation seller and buyer respectively). For the management of virtual accouts we suggest the use of Special Peers (SPs) which are organized in a Distributed Hash Table, so that each SP is responisble for a number of entities. Payment analysis [14] shows that CREPS offers incentives for providing honest recommendations, since (a) the possibility of aquiring honest recommendations is linked with high recommendation reputation, and (b) the access of dishonest reccommenders to the recommendation exchange mechanism is prohibited after a number of recommendation exchanges, depending on the defined threshold values and Initial Account Balance.

CREPS involves also a recommendation exchange protocol based on a Public Key Infrastructure (PKI), which is depicted in Figure 4. According to this protocol signed messages are exchanged between (a) the recommendation buyer and seller, and (b) the participating entities and their SPs, for crediting and debiting the related accounts.

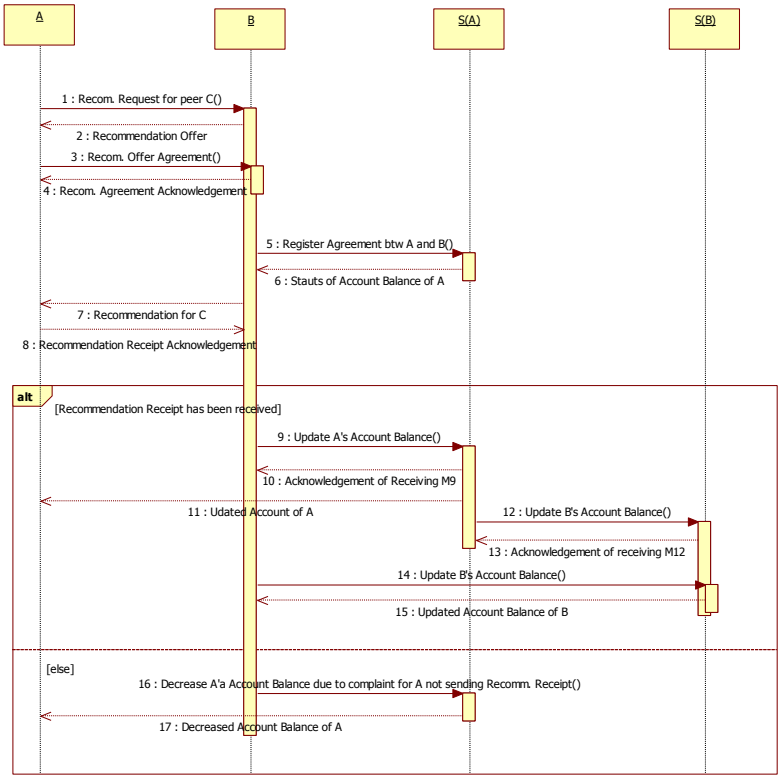


Fig. 4. Exchange of Signed messages during the Recommendation Process

For the evaluation of CREPS we have qualitatively compared it with a number of reputation systems which are based on a PKI. The analysis is based on the level that each reputation system fulfills the following requirements: (i) Privacy / confidentiality, (ii) Non-repudiation, (iii) Traceability, (iv) Ballot-staffing prevention, (v) Sybil Attack prevention, (vi) Whitewashing attack prevention, and (vii) Message integry . Our analysis shows that the various PKI-based reputation mechanisms deal with the aforementioned requirements in various levels depending on the goals and priorities set in each model. The simple PKI-based mechanims of CREPS offer message integrity via encryption, and traceability, non-repudiation and resilience to bad-mouthing via digital signatures. Entities' privacy is covered partially, as the exchanged recommendations and the updated account balances are made aware only to involved entities and to Special Peers which are responsible for them.

3 Contributions

The challenges revealed in the area of reputation systems, regarding their design, threat analysis, credibility enhancement and evaluation, have motivated this thesis, the results of which are composed of the following components:

1. A *generic framework for the development and evaluation of credible reputation systems* for distributed e-Communities, which consists of (a) a conceptual model for reputation systems design, (b) a framework for the integration of various credibility factors in reputation systems, (c) a framework that enables choosing / setting up suitable evaluation methods for specific RSs and also choosing or creating Common Evaluation Frameworks for reputation systems.
2. A *reputation system with integrated credibility factors*, which make it resilient against various attacks. Such factors include recommendation reputation, time decaying, confidence regarding provided recommendaitons and direct reputation values, and adjusting the weights of direct and indirect reputation.
3. A *novel recommendation exchange mechanism based on virtual payments*, which gives incentives for honest recommendations. This mechanism is suitable for reputation systems the efficiency of which depends on honest recommendation provision, as well as on recommendation integrity and confidentiality.

Specifically, the main contributions of the thesis are:

- Two reference models for reputation systems for distributed communities: a conceptual representation of the structure and functionality of a RS which contains the involved entities, attributes, relationships and operations (depicted in Figure 1) and a representation of the workflow of activities of the reputation estimation process in a distributed reputation system which involves a recommendation acquiring activity. The provided formalization of RSs shortens the gap in the research regarding standardization and formalization of reputation systems, which have the following characteristics: (1) reputation estimation is done locally by the trustor, based on direct experience and third-party recommendations, and (2) each peer keeps track of the recommendation reputation of other peers from which it has received recommendations. It also helps researchers in approaching reputation systems in a unified way and thus facilitates their design process.
- Four taxonomies: one for P2P reputation systems [5], one for reputation systems attacks and defence mechanisms [10], one for reputation systems for social network-based applications [7], and one for RS evaluation approaches [11]. The first two taxonomies have been used in a number of research works, appearing in the state-of-the-art of the corresponding fields, or offering a basis for new approaches of RSs and defense mechanisms, e.g. [6-21]. The third taxonomy contributes to the formalization of the more abstract reputation mechanisms which are proposed for the vast and continuously growing area of Social Network-based communities. İt also facilitates the design process of reputation systems for specific types of Social Network-based RSs, as shown in [7]. We note that such RSs are expected to have extensive application in various fields, such as in marketing and social network analysis. The fourth taxonomy enlightens RS designers as to eligible evaluation methods for their RSs.

- A framework for the credibility evaluation of RSs, consisting of a set of credibility criteria which together with the aforementioned taxonomy of attacks and defense mechanisms can be used for assessing the credibility of reputation systems and their resilience to attacks, as presented in [10].
- A thorough survey in the field of reputation systems evaluation, which provides a roadmap for objective evaluation and comparison of reputation systems through a Common Evaluation Framework [11].
- A set of credible reputation metrics for e-Communities and a novel credible recommendation exchange mechanism. The reputation metrics incorporate various credibility factors and provide resilience to attacks against reputation systems. The evaluation results show the efficiency of our reputation metrics in various scenarios and also indicate their usability in real applications. The proposed reputation metrics have been used in a number of research works, such as [22, 23]. The proposed recommendation exchange mechanism uses a credit-based scheme for payments for recommendation exchanges, which offers incentives for honest recommendations, and has been presented in [16]. It has been used as a reference incentive-based mechanism representing state-of-the-art in incentive-based reputation systems, e.g. in [24-26].

We state that our work enhances the area of RSs in various aspects, especially the aspects of design, credibility, evaluation and incentives; this belief has been supported by the adaptation of parts of our work by other research works, as aforementioned. Our future work plans include expanding our work in the fields of credit-based, social network-based and e-Commerce supporting RSs, and further work on benchmarking of reputation systems, i.e. on defining and implementing a CEF which will be grounded on or extend current CEF approaches and will incorporate the desirable characteristics identified in this thesis. Subsequently, we plan to use such a CEF for thoroughly experimenting with the evaluation of the reputation metrics we have developed [16] and other reputation systems in various application environments, contributing thus to the design of optimal RSs for specific e-Community contexts.

References

1. eBay, http://www.ebay.com. Accessed on 1/8/2016
2. Song, S., Hwang, K., Zhou, R.: Trusted P2P Transactions with Fuzzy Reputation Aggregation. IEEE Internet Computing, Special Issue on Security for P2P and Ad Hoc Networks,. 9, 6, 24-34 (2005)
3. Xiong, L., Liu, L.: PeerTrust: Supporting reputation-based trust for peer-to-peer electronic communities. IEEE Transactions on Knowledge and Data Engineering, 16, 7, 843–857 (2004)
4. Dillon, T. S., Chang, E., Hussain, F.K.: Managing the Dynamic Nature of Trust. IEEE Journal Of Intelligent Systems, 19, 5, 79-82 (2004)
5. Koutrouli, E., Tsalgatidou, A.: Reputation-based Trust Systems for P2P Applications: Design Issues and Comparison Framework. In: 3rd Intl. Conf. on Trust, Privacy and Security in Digital Business, pp. 152-161, Springer-Verlag, Berlin, Heidelberg (2006)
6. Hussain, O. K., Chang, E., Hussain, F. K., Dillon, T. S.: A methodology to quantify failure for risk-based decision support system in digital business ecosystems. Data Knowl. Eng. 63, 3, 597-621 (2007) DOI=http://dx.doi.org/10.1016/j.datak.2007.03.014
7. Koutrouli, E., Kanellopoulos, G., Tsalgatidou, A.: Reputation Mechanisms in on-line Social

Networks – The case of an Influence Estimation System in Twitter. Accepted for publication in South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conf. (SEEDA-CECNSM 2016), ACM (2016)
8. Kamvar, S., Schlosser, M., Garcia-Molina, H.: The EigenTrust Algorithm for Reputation Management in P2P Networks. In: World Wide Web Conf. 2003, pp. 640-651. ACM (2003)
9. Stoica, I., Morris, R., Liben-Nowell, D., Karger, D. R., Kaashoek, M. F., Dabek, F., Balakrishnan, H.: Chord: a Scalable Peer-to-Peer Lookup Protocol for Internet Applications, IEEE/ACM Transactions on Networking, 11, 1, 17-32 (2003)
10. Koutrouli E., Tsalgatidou, A.: Taxonomy of Attacks and Defense Mechanisms in P2P Reputation Systems—Lessons for Reputation System Designers. Computer Science Review, 6, 2-3, 47-70 (2012)
11. Koutrouli E., Tsalgatidou, A.: Reputation Systems Evaluation Survey. ACM Computing Surveys, 48, 3, 1-28 (2015)
12. Koutrouli, E., Tsalgatidou, A.: Credibility Enhanced Reputation Mechanism for Distributed e-Communities. In: 19th Euromicro Intl. Conf. on Parallel, Distributed and Network-Based Computing (PDP 2011), pp. 627-634. IEEE Computer Society, Washington, DC, USA (2011)
13. Mármol, F. G., Pérez, G. M.: TRMSim-WSN, Trust and Reputation Models Simulator for Wireless Sensor Networks. In: IEEE Intl. Conf. on Communications (IEEE ICC 2009), pp. 915-919. IEEE Press, Piscataway, NJ, USA (2009)
14. Zhou, R., Hwang, K.: Powertrust: A Robust and Scalable Reputation System for Trusted Peer-to-Peer Computing. IEEE Transactions on Parallel and Distributed Systems, 18, 4, 460-473 (2007)
15. Mármol, F. G. and Pérez, G. M.: Providing Trust in Wireless Sensor Networks Using a Bio-Inspired Technique. Telecommunication Systems, 46, 2, 163-180 (2011)
16. Koutrouli, E., Tsalgatidou, A.: Credible Recommendation Exchange Mechanism for P2P Reputation Systems. In: Trust, Reputation, Evidence and other Collaboration Know-how (TRECK) Track, ACM Symposium on Applied Computing 2013, pp. 1943-1948. ACM (2013)
17. Li, X., Gui, X.: Research on Dynamic Trust Model for Large Scale Distributed Environment. Journal of Software, 18, 4, 460-473 (2007)
18. Clarke, S., Christianson, B., Xiao, H.: Extending Trust in Peer-to-Peer Networks. In: 13th East European conference on Advances in Databases and Information Systems (ADBIS'09), pp. 145-152, Springer-Verlag, Berlin, (2009)
19. Hendrikx, F., Bubendorfer, K., Chard, R.: Reputation Systems: A Survey and Taxonomy. Journal of Parallel and Distributed Computing, 75, 184-197 (2015)
20. Vavilis, S., Petkovic, M., Zannone, N.: A reference model for reputation systems. Decision Support Systems, 61, 147-154 (2014)
21. Sänger, J., Richthammer, C, Rösch A., Pernul, G.: Reusable Defense Components for Online Reputation Systems. In: 9th IFIP WG 11.11 Intl. Conf. (IFIPTM 2015), Hamburg, Germany, pp. 195-202, Springer, Berlin (2015)
22. Vallée, T., Bonnet, G.: Using KL Divergence for Credibility Assessment, In: 2015 Intl. Conf. on Autonomous Agents and Multiagent Systems (AAMAS '15), pp. 1797-1798, Intl. Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2015)
23. Dadhich, P., Dutta, K., Govil, M. C: Detection of Slanders through Euclidean Distance Similarity Assessment for Securing e-Commerce Agents in P2P Decentralised Electronic Communities. Intl. Journal of Security and Networks, 11, 1/2, 48-65 (2016)
24. Lafuente, C. B., Seigneur, J. M.: Extending Trust Management with Cooperation Incentives: A Fully Decentralized Framework for User-Centric Network Environments. Journal of Trust Management, 2, 7, Springer (2015)
25. Seddiki, M., Benchaïba, M.: Gpop: A Global File Popularity Measurement for Unstructured P2P Networks. Int. J. Distrib. Syst. Technol. 6, 3, 51-64 (2015)
26. Haddi F. L., Benchaïba, M.: A survey of Incentive Mechanisms in Static and Mobile P2P Systems. Journal of Network and Computer Applications, 58, C, 108-118 (2015)

Robust Algorithms for Linear and Nonlinear Regression via Spare Modeling Methods: Theory, Algorithms and Applications to Image Denoising

George K. Papageorgiou*

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
geo_papag@hotmail.com

Abstract. In this dissertation, the problem of robust regression is studied, for both the linear and the nonlinear case. For the former case, a novel algorithm, Greedy Algorithm for Robust Denoising (GARD), which is based on sparse optimization techniques, is derived. Moreover, theoretical conditions, which guarantee the identification of the outliers and a bound on the estimation error, are provided. Next, we focus on the nonlinear case, where it is assumed that the unknown nonlinear function belongs to a Reproducing Kernel Hilbert Space (RKHS). A robust scheme, Kernel Greedy Algorithm for Robust Denoising (KGARD), which shares the same concept with GARD, is proposed. The algorithm is compared against other cutting edge methods via extensive simulations, where its enhanced performance is demonstrated. In addition, theoretical results regarding the identification of the outliers are provided. Finally, the proposed robust estimation framework is applied to the task of image denoising, where the advantages of the proposed method are unveiled. The experiments verify that KGARD improves the denoising process significantly, when outliers are present.

Keywords: robust linear regression, robust nonlinear regression in RKHS, greedy algorithm for robust denoising, kernel greedy algorithm for robust denoising, image denoising, outliers

1 Introduction

At the heart of Machine Learning is the task of *regression* or *regression analysis*. In a classic regression task, given a set of training data, the goal is to learn a set of unknown parameters in order to make predictions. In simple words, the task could be seen as a curve fitting problem. Consider a set of training points $(y_i, \boldsymbol{x}_i)$, $y_i \in \mathbb{R}$ and $\boldsymbol{x}_i \in \mathbb{R}^M$ for $i = 1, \ldots, N$. The task is to estimate a function, f, whose graph fits the data. The target function, f, of the independent variables, $\boldsymbol{x}$, is called the *regression function* and can be either linear or nonlinear. The difference

* Dissertation Advisor: Sergios Theodoridis, Professor.

between regression and classification is that in regression the dependent variable belongs to an interval in the real axis (or region in the complex plane), while in classification it is a discrete variable.

Regression analysis is widely used for prediction and forecasting. It is also used as a means to extract information concerning the degree of dependence among the dependent (output) and the independent (input) variables. Thus, useful information and related implications of such dependencies can be revealed.

The earliest form of regression was the method of Least Squares (LS), which was published by Legendre in 1805 and by Gauss in 1809. Legendre and Gauss both applied the method to the problem of determining the orbits of comets, based on astronomical observations. Many techniques that perform regression analysis have been developed, since then. Familiar methods such as linear regression and ordinary Least Squares regression belong to the parametric class of learning techniques; that is, the model function is defined in terms of a finite number of unknown parameters that are estimated from the data. In contrast, nonparametric regression refers to techniques that bypass the need for explicit parameterization of the unknown functional dependence. For example, the regression function can be assumed to lie in a specific set of functions, which may also be infinite-dimensional. A popular example, that will be adopted in the current thesis for the estimation of a nonlinear function, is to assume that the regression function lies in a Reproducing Kernel Hilbert Space (RKHS).

The performance of regression methods, in practice, depends on the form of the data-generating mechanism and how this relates to the regression model being used. Since the true form of the data-generating process is generally unknown, regression analysis often depends, to a large extent, on making assumptions concerning this process. Regression models, that are designed for prediction, are often useful even when the assumptions are moderately violated, although they may not perform optimally. However, if our goal is to make accurate predictions, we should look for a model/method that is *robust* enough, i.e., it can tolerate abnormalities on the data so that the estimation is not significantly affected.

The notion of robustness, i.e., the efficiency of a method to solve a learning task from data under noise uncertainties of various types, has been a major issue in the scientific community for over half a century. The goal is to minimize the effect of the observations that have been corrupted by unexpected high values of noise, known as *outliers*. Outliers are often regarded as erroneous measurements that deviate greatly from the rest of the observations. This is due to the fact: either their values are heavily influenced by another source or they are generated by a different mechanism/distribution.

In such cases, classic estimators, e.g., the Least Squares, are known to fail to perform well. This problem was originally addressed since the 1950s and it was actually solved more than a decade later, by Huber. Eventually, it led to the development of a new field in Statistics, known as *Robust Statistics*. However, the need for development of robust estimators was not only limited within the Statistics scientific community. Similar tasks (involving robust estimators) emerged in

the context of many fields such as Physics, Medicine, Biology, Engineering and Computer Science, to name a few.

The robust tools that have been developed over the years for handling outliers can be classified into two major categories. The first one includes tools that rely on the use of *diagnostics*, whereas the second direction is based on *robust regression* methods. Diagnostics and robust regression have the same goals, only obtained in the opposite order; both approaches have a long history in the field of Robust Statistics. Lately, a different approach has emerged. The recent development of methods in the spirit of robust analysis owes a lot to the emergence of *sparse modeling* methods, during the past decade.

Sparsity-aware learning and related optimization techniques have been at the forefront of the research in signal processing, encompassing a wide range of topics, such as compressed sensing, signal denoising and approximation techniques. Sparsity is closely related to sufficiency or economy of a representation, a mechanism that harmonizes with nature, which tends to be parsimonious. At the heart of this problem lies an underdetermined set of linear equations, which, in general, accepts an infinite number of solutions. Imposing sparsity, is interpreted as seeking for a solution where only a few of the unknown coordinates, which we attempt to estimate, are nonzero. There are two major paths, towards modeling sparse vectors/signals. The first one focuses on minimizing the ℓ_0(pseudo)-norm of a vector, which equals the number of its nonzero coordinates. However, since this is a non-convex optimization task, approximate methods have been established. The family of algorithms that have been developed to address problems involving the ℓ_0 (pseudo)-norm, comprises *greedy* methods, which have been shown to provide the solution of the related minimization task, under certain reasonable assumptions. Even though, in general, this is an NP-Hard task, it has been shown that such methods can efficiently recover a solution in polynomial time. On the other hand, the family of algorithms developed around the methods that employ the ℓ_1-norm, embraces convex optimization, providing a broader set of tools and stronger guarantees for convergence. Both methods have been shown to generate sparse solutions.

A more recent application of sparse modeling and optimization methods, which is also the focus of this work, is that of signal denoising. There, one is interested in recovering the original signal, which apart from the standard inlier noise, e.g., Gaussian, has also been corrupted by outliers. The key to this modeling is to assume that the outliers comprise only a small fraction of the entire data set, thus the outlier vector is modeled as a sparse one.

The goal of this dissertation, is to address the task of robust linear and nonlinear regression via sparse modeling methods, within the context of machine learning. The proposed methods are built on the popular Orthogonal Matching Pursuit algorithm (OMP), by imposing sparsity constraints on the outliers. In particular, two novel robust algorithms are developed. One for the task of linear regression and a second one for the task of nonlinear regression, where it has been also assumed that the function to be estimated lies in a Reproducing Kernel Hilbert Space (RKHS). Various experiments are performed, where both of the

algorithms are compared against state-of-the-art methods. The obtained results demonstrate their performance and highlight their advantages. Moreover, the study of the algorithms has led to the establishment of sound theoretical results. Finally, the focus is turned on the applications of the nonlinear regression scheme to the task of image denoising. As a result, two methods are introduced for the removal of impulsive noise. The most significant results of this novel robust approach are outlined next.

2 Robust Linear Regression

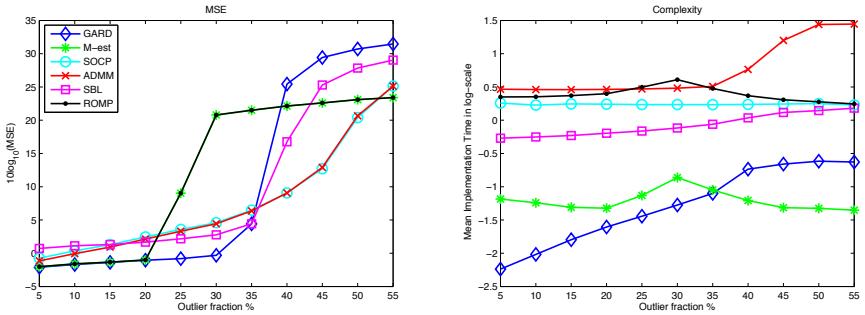
For the linear regression task we have assumed that the output data are corrupted by inlier and outlier noise. Moreover, we have assumed that the outliers are only few compared to the number of the data (thus the outlier vector can be modeled as a sparse one) and that the number, N, of the available data is sufficiently greater than the number, M, of the unknown coefficients. The proposed algorithm is called Greedy Algorithm for Robust Denoising (GARD), and it is based on the classic Orthogonal Matching Pursuit (OMP). The method alternates between a Least Squares (LS) optimization criterion and an OMP-like selection step, that identifies the outliers. The theoretical results that have been established for GARD are:

- The convergence of the scheme in a finite number of steps.
- A bound on the Restricted Isometry Property (RIP) constant, for the case where only outliers are present, which guarantees that GARD successfully identifies the outliers. Moreover, the method recovers both the regression solution and the sparse outlier vector, exactly (with no error), under the existence and uniqueness conditions.
- A second bound on the Restricted Isometry Property (RIP) constant, for the case where the data is corrupted by both inlier and outlier noise, which guarantees that GARD successfully identifies the outliers, assuming that the inlier noise is bounded.
- Performance bounds on the approximation, which guarantee the stability of the algorithm.

It should be noted that, the result concerning the identification of the outliers in the presence of both inlier and outlier noise has been derived for the first time in the robust regression framework.

Next, follows an extended set of experiments that are performed and demonstrate the performance of GARD against other comparative cutting edge methods. For each method, we have computed the Mean-square-error (MSE) and the Mean Implementation Time (MIT), while varying the fraction of the outliers. The most significant results for GARD are:

- It attains the lowest MSE.
- It demonstrates enhanced robustness, compared to all other methods.
- It has very low computational requirements.



(a): $M = 100$ (b): $M = 100$

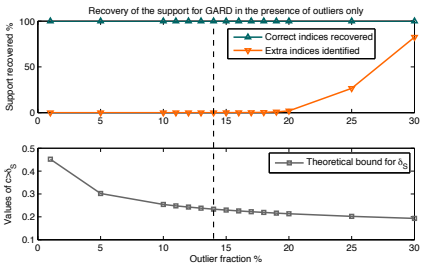
Fig. 1: (a): The attained Mean-square-error (MSE) (in logarithmic scale-dB) versus the fraction of outliers in the output data. (b): Logarithmic scale of the Mean Implementation Time (MIT) versus the outlier fraction. The number of the data is $N = 600$.

In Figure 1 (a), the MSE (in dBs) attained by each method versus the fraction of outliers is depicted, for a fix dimension of the unknown vector at $M = 100$. The Mean Implementation Time (MIT) is also plotted in logarithmic scale in Figure 1 (b). Observe that GARD attains the lowest MSE among its competitors, while in parallel it seems to be the most efficient, operating at the lowest computational cost (the interesting "zone" is for fractions of less than 30%, that is $10\% - 20\%$).

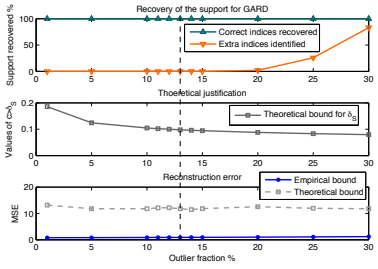
Moreover, in Figure 2 the capability of KGARD to identify the outliers is demonstrated. The green line pointing upwards corresponds to successful outlier identifications, while the orange one pointing downwards corresponds to extra indices that GARD has classified as outliers. In parallel, the relation of the percentage of outliers to the bound of the RIP constant is shown (grey line). Figure 2 (a) corresponds to the noiseless case, while in (b), the data is corrupted by outlier and bounded inlier noise, as the resulting theorem suggests. It is clear, that for small fractions of outliers the support is recovered (one-to-one index), thus we conclude that the condition is valid (the RIP constant cannot be computed).

Figure 3 (a) demonstrates the probability of recovery for each method tested, while varying the fraction of the outliers. In Figure 3 (b), the phase transition curves for each method are given. For each dimension of the unknown vector, we have computed the fraction of outliers for which the method transits from success to failure with probability $p = 0.5$. For example, for $M = 100$ (Figure 3 (a)), the horizontal line at 0.5 corresponds to fractions of outliers (for each method) that are located in the y-axis of Figure 3 (b) for the dimension of $M = 100$. Here, it is clear that up to $M = 200$, GARD succeeds to recover the solution with a higher probability than the rest of the methods.

Finally, in Table 1 we have measured the attained MSE for the case where the noise follows a more general distribution. In columns A, B and C the noise originates from the Lévy alpha-stable distribution, while in column D the noise

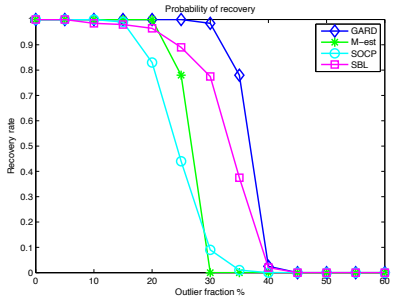


(a): Noise: outliers only.

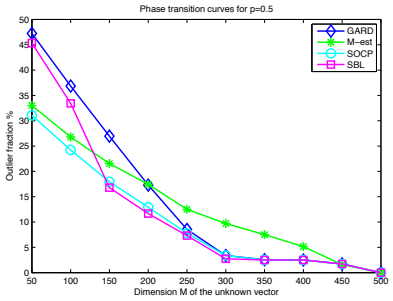


(b): Noise: outliers and bounded inlier.

Fig. 2: The identification of the outliers and the relation to the theoretical bound of the Restricted Isometry Property (RIP), δ_S. (a): The data is corrupted by outliers only. (b): The data is corrupted by outliers and bounded inlier noise. Moreover, the empirical error is computed and the relation to its theoretical upper bound is depicted.



(a): Probability of recovery for $M = 100$



(b): Transition from success to failure.

Fig. 3: (a): The probability of recovery while varying the fraction of outliers, for the the dimension $M = 100$ of the unknown vector, $\boldsymbol{\theta}$, and $N = 600$ observations. As the fraction of the outliers increases, the probability for an accurate estimation drops. (b): Transition from success to failure with probability $p = 0.5$. A vertical line at $M = 100$ indicates the percentage of outliers (for each method respectively) that correspond to the values of the x-axis for probability $p = 0.5$, in (a).

Table 1: Computed MSE, for various experiments. In tests A, B and C, the noise is drawn from the heavy-tailed distribution alpha-stable of Lévy distribution. In test D, noise consists of a sum of two vectors, drawn from 2 independent Gaussian distributions with different variance, plus an outlier noise vector of impulsive noise.

Algorithm	Test A	Test B	Test C	Test D
GARD	0.1772	0.0180	0.0586	0.690
M-est	0.2248	0.2859	1.844e+06	0.704
SOCP	0.4990	0.3502	5.852e+05	1.011
SBL	0.9859	58.3489	2.165e+06	1.292
ROMP	0.2248	0.2859	1.844e+06	0.704

consists of outliers plus inlier noise, with values drawn from two independent Gaussian distributions with different variance.

3 Robust Nonlinear Regression

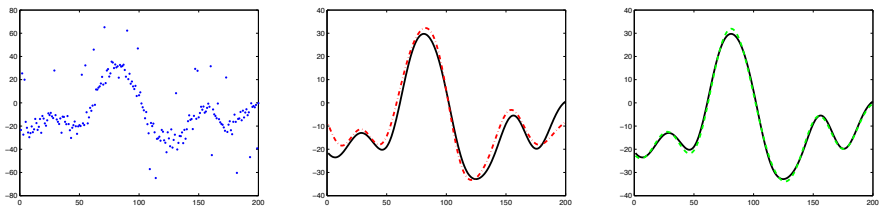
For the study of the nonlinear regression task we have assumed that the original function to be estimated lies in a Reproducing Kernel Hilbert Space (RKHS). Thus, we resort to simple manipulations by replacing the regression matrix with a kernel one. However, since this is a nonparametric estimation task, the proposed robust algorithm had to be modified again (with respect to GARD). The novel scheme, Kernel Greedy Algorithm for Robust Denoising (KGARD), alternates between a Kernel Ridge Regression (KRR) task and an OMP-like selection step. The addition of a regularization term at the estimation steps cannot be avoided and leads to a more complex theoretical analysis for the method. Thus, a different path, than the previously reported one (linear case) is followed. The study of this greedy-based selection scheme led to some interesting results:

- The solution to the regularized Least Squares task, which is performed at each step, is unique.
- The establishment of a bound on the maximum singular value of the kernel matrix, which guarantees that the method identifies the correct locations of all the outliers, first.

However, the method still manages to recover the correct support of the sparse outlier vector in many cases where the theoretical result does not hold. This leads to the conclusion that the provided conditions can be loosen up significantly in the future. The reason that the analysis is carried out for the case where inlier noise is not present is due to the fact that the analysis gets highly involved. The absence of the inlier noise makes the analysis easier and it highlights some theoretical aspects on why the method works. It must be emphasized that, such a

theoretical analysis appears for the first time in the related bibliography. Moreover, in practice, where inlier noise also exists, the method succeeds to correctly identify the majority of the outliers. The significance of the robust nonlinear regression task, is demonstrated in Figure 4, where the estimation with KGARD is compared against the non-robust Kernel Ridge Regression (KRR) method.

On the experimental section, various simulations are performed designating the overall advantages of KGARD against its competitors. In the tests performed, we have measured the MSE, the Mean Implementation Time (MIT) and the number of correct and wrong indices that each method has classified as outliers. In Table 2, the results of the estimation over the nonlinear function $\mathrm{f} = 20sinc(2\pi x)$ are depicted, for various levels of noise (inlier-outlier). It is observed that, KGARD attains the lowest MSE for most of the cases, except for the fraction of outliers at 20%. It should also be noted that, for small fractions of outliers the computational cost of the method is very low, and additionally, it successfully manages to identify the outliers.



(a) (b) (c)

Fig. 4: The significance of robust estimation: (a) Data corrupted by both inlier and 10% of outlier noise. (b) The black and the red dashed lines correspond to the uncorrupted data and the non-robust estimation performed, respectively. The MSE over the training set is 10.79. (c) The black and the green dashed lines correspond to the uncorrupted data and the robust estimation performed with KGARD, respectively. The MSE over the training set is 1.21.

4 Applications to Image Denoising

Finally, we present the applications of the proposed method, i.e., KGARD, in the context of image denoising. In particular, the goal is to approximate the original image that is corrupted by Gaussian (inlier) plus salt and pepper noise (outliers). To this end, the method has been slightly modified and adapted to the task, so that no tuning parameters are involved; instead, the parameters are automatically tuned by the method. As a result, two novel methods are proposed for the task of robust denoising: a) a direct KGARD implementation

Table 2: Computed MSE for f(x) = $20sinc(2\pi x)$ over the training and validation set. Additionally, the percentage of correct and wrong indices that each method has classified as outliers and the Mean Implementation Time (MIT), for various levels of inlier and outlier noise, are evaluated.

Algorithm	MSE_{tr}	MSE_{val}	Cor. ind.	Wr. ind.	MIT (sec)	Inlier - Outlier
RB-RVM	0.0850	0.0851	-	-	0.298	20 dB - 5%
RAM ($\lambda = 0.07, \mu = 2.5$)	0.0344	0.0345	100 %	0.2 %	0.005	20 dB - 5%
KGARD ($\lambda = 0.2, \varepsilon = 10$)	**0.0285**	**0.0285**	100 %	0 %	0.004	20 dB - 5%
RB-RVM	0.0911	0.0912	-	-	0.298	20 dB - 10%
RAM ($\lambda = 0.07, \mu = 2.5$)	0.0371	0.0372	100 %	0.1 %	0.007	20 dB - 10%
KGARD ($\lambda = 0.2, \varepsilon = 10$)	**0.0305**	**0.0305**	100 %	0 %	0.008	20 dB - 10%
RB-RVM	0.0992	0.0994	-	-	0.299	20 dB - 15%
RAM ($\lambda = 0.07, \mu = 2$)	0.0393	0.0393	100 %	0.6 %	0.008	20 dB - 15%
KGARD ($\lambda = 0.3, \varepsilon = 10$)	**0.0330**	**0.0330**	100 %	0 %	0.012	20 dB - 15%
RB-RVM	0.1189	0.1184	-	-	0.305	20 dB - 20%
RAM ($\lambda = 0.07, \mu = 2$)	**0.0421**	**0.0422**	100 %	0.4 %	0.010	20 dB - 20%
KGARD ($\lambda = 1, \varepsilon = 10$)	0.0626	0.0626	100 %	0 %	0.017	20 dB - 20%
RB-RVM	0.3630	0.3631	-	-	0.327	15 dB - 5%
RAM ($\lambda = 0.15, \mu = 5$)	0.1035	0.1036	100%	0.7 %	0.005	15 dB - 5%
KGARD ($\lambda = 0.3, \varepsilon = 15$)	**0.0862**	**0.0862**	100 %	0.1 %	0.005	15 dB - 5%
RB-RVM	0.3828	0.3830	-	-	0.319	15 dB - 10%
RAM ($\lambda = 0.15, \mu = 5$)	0.1117	0.1118	100%	0.4 %	0.006	15 dB - 10%
KGARD ($\lambda = 0.3, \varepsilon = 15$)	**0.0925**	**0.0925**	100 %	0 %	0.008	15 dB - 10%
RB-RVM	0.4165	0.4166	-	-	0.317	15 dB - 15%
RAM ($\lambda = 0.15, \mu = 5$)	0.1186	0.1186	100%	0.3 %	0.007	15 dB - 15%
KGARD ($\lambda = 0.3, \varepsilon = 15$)	**0.1001**	**0.1003**	100 %	0 %	0.012	15 dB - 15%
RB-RVM	0.4793	0.4798	-	-	0.312	15 dB - 20%
RAM ($\lambda = 0.15, \mu = 4$)	**0.1281**	**0.1282**	100%	1.4 %	0.008	15 dB - 20%
KGARD ($\lambda = 0.7, \varepsilon = 15$)	0.1340	0.1349	100 %	0 %	0.016	15 dB - 20%

that can perform the estimation and b) a KGARD scheme combined with a popular wavelet-based method, i.e., Block Matching and 3-D filtering (BM3D). The latter scheme, which first performs the identification and estimation of the outliers via the proposed algorithm (KGARD) and then it removes the remaining of the noise via the BM3D, demonstrated enhanced performance in terms of approximation. The results have been averaged based on the measured Peak signal-to-noise ratio (PSNR).

In Table 3, various results are given for the denoising of the Lena image. In Figure 5, the result of the process is clearly demonstrated. Finally, in Table 4 various results on the denoising of the boat image are depicted, while in Figure 6 the improvement achieved by the combined KGARD-BM3D method is observed.

Table 3: Denoising performed on the *Lena* image corrupted by various types and intensities of noise using the proposed methods, the robust RVM (RB-RVM) approach and the state-of-the-art wavelet method BM3D.

Method	Parameters	Gaussian Noise	Impulses (±100)	PSNR
BM3D	$s = 30$	25 dB	10%	30.84 dB
RB-RVM	$\sigma = 0.3$	25 dB	10%	31.25 dB
KGARD	$\sigma = 0.3$, $\lambda = 1$	25 dB	10%	33.49 dB
KGARD-BM3D	$\sigma = 0.3$, $\lambda = 1$, $s = 10$	25 dB	10%	**35.67 dB**
BM3D	$s = 35$	20 dB	10%	30.66 dB
RB-RVM	$\sigma = 0.4$	20 dB	10%	29.09 dB
KGARD	$\sigma = 0.3$, $\lambda = 1$	20 dB	10%	31.94 dB
KGARD-BM3D	$\sigma = 0.3$, $\lambda = 1$, $s = 15$	20 dB	10%	**33.81 dB**
BM3D	$s = 40$	15 dB	10%	29.94 dB
RB-RVM	$\sigma = 0.4$	15 dB	10%	25.85 dB
KGARD	$\sigma = 0.3$, $\lambda = 2$	15 dB	10%	28.47 dB
KGARD-BM3D	$\sigma = 0.3$, $\lambda = 1$, $s = 25$	15 dB	10%	**30.77 dB**

Table 4: Denoising performed on the *boat* image corrupted by various types and intensities of noise using the state-of-the-art wavelet method BM3D with and without outlier detection.

Method	Parameters	Gaussian Noise	Impulses (±100)	PSNR
BM3D	$s = 25$	25 dB	5%	30.57 dB
KGARD-BM3D	$\sigma = 0.3$, $\lambda = 1$, $s = 10$	25 dB	5%	**34.61 dB**
BM3D	$s = 35$	20 dB	10%	28.97 dB
KGARD-BM3D	$\sigma = 0.3$, $\lambda = 1$, $s = 15$	20 dB	10%	**31.52 dB**
BM3D	$s = 50$	20 dB	20%	27.49 dB
KGARD-BM3D	$\sigma = 0.4$, $\lambda = 1$, $s = 15$	20 dB	20%	**29.7 dB**

(a) (b)

(c) (d)

Fig. 5: (a) The *Lena* image corrupted by 20 dB of Gaussian noise and 10% outliers. (b) Denoising with BM3D (30.66 dB). (c) Denoising with KGARD (31.94 dB). (d) Denoising with joint KGARD-BM3D (33.81 dB).

(a) (b) (c)

Fig. 6: (a) The *boat* image corrupted by 20 dB of Gaussian noise and 10% outliers. (b) Denoising with BM3D (28.97 dB). (c) Denoising with joint KGARD-BM3D (31.52 dB).

5 Conclusions

In this dissertation we studied the problem of robust regression, for both the linear and nonlinear case, under the framework of sparse optimization techniques. Two novel algorithms are derived, for each case, and they are compared against state-of-the-art methods through extensive simulations. The results demonstrated enhanced performance, in terms of estimation and computational cost. Moreover, theoretical results, which guarantee the identification of the outliers, are provided. Finally, the proposed framework is applied to the task of image denoising, where it is shown that the process is significantly improved.

References

1. Papageorgiou, G., Bouboulis, P., Theodoridis, S.: Robust non-linear Regression: A Greedy Approach Employing Kernels. IEEE Transactions on Signal Processing, accepted (2017)
2. Papageorgiou, G., Bouboulis, P., Theodoridis, S.: Robust Linear Regression Analysis - A Greedy Approach. IEEE Transactions on Signal Processing 63(15), 3872–3887 (2015)
3. Papageorgiou, G., Bouboulis, P., Theodoridis, S.: Robust Regression in RKHS - An Overview. In: Proceedings of the European Signal Processing Conference, pp. 2874–2878. IEEE Press, New York (2015)
4. Papageorgiou, G., Bouboulis, P., Theodoridis, S.: Robust Linear Regression Analysis - The Greedy Way. In: Proceedings of the European Signal Processing Conference, pp. 16–20. IEEE Press, New York (2014)
5. Papageorgiou, G., Bouboulis, P., Theodoridis, S.: Robust Image Denoising in RKHS via Orthogonal Matching Pursuit. In: International Workshop on Cognitive Information Processing, pp. 1–6. IEEE Press, New York (2014)
6. Papageorgiou, G., Bouboulis, P., Theodoridis, S.: Robust Kernel-Based Regression Using Orthogonal Matching Pursuit. In: International Workshop on Machine Learning for Signal Processing, pp. 1–6. IEEE Press, New York (2013)

Design and Synthesis of Efficient Circuits for Quantum Computers

Archimedes D. Pavlidis*

Department of Informatics & Telecommunications,
University of Athens, Athens, Greece,
{erevos@di.uoa.gr}

Abstract. The recent advances in the field of experimental construction of quantum computers with increased fidelity components shows that large-scale machines based on the principles of quantum physics are likely to be realized in the near future. As the size of the future quantum computers will be increased, efficient quantum circuits and design methods will gradually gain practical interest. The contribution of this thesis towards the design of efficient quantum circuits is two-fold. The first is the design of novel efficient quantum arithmetic circuits based on the Quantum Fourier Transform (QFT), like multiplier-with-constant-and-accumulator (MAC) and divider by constant, both of linear depth (or speed) with respect with the bits number of the integer operands. These circuits are effectively combined so as they can perform modular multiplication by constant in linear depth and space and consequently modular exponentiation in quadratic time and linear space. Modular exponentiation and modular multiplication operations are integral parts of the important quantum factorization algorithm of Shor and other quantum algorithms of the same family, known as Quantum Phase Estimation algorithms. Important implementation problems like the required high accuracy of the employed rotation quantum gates and the local communications between the gates are effectively addressed. The second contribution of this thesis is a generic hierarchical synthesis methodology for arbitrary complex and large quantum and reversible circuits. The methodology can handle more easily larger circuits relative to the flat synthesis methods. The proposed method offers advantages over the standard hierarchical synthesis which uses Bennett's method of "compute-copy-uncompute".

Keywords: Quantum computer architectures, quantum arithmetic circuits, quantum Fourier transform, quantum circuits synthesis, reversible circuits synthesis

1 Introduction

Quantum Information Theory and Quantum Computing are interdisciplinary research fields that combine different doses of Physics, Informatics and Mathematics depending on which aspect someone focuses. Quantum Computing is

* Dissertation Advisor: D. Gizopoulos, Professor

a relatively recent research field, although Quantum Information Theory has already been developed for the last 40 years, after important results which connect classical Information Theory to Quantum Mechanics (quantum entropies inequalities [2, 18, 19], Holevo bounds for capacities of quantum channels [15, 16], Bekenstein bound [5], etc.)

The theoretical connection of Quantum Mechanics to the Theory of Computation achieved in the 80's [11, 12], while more boost came in the 90's with the invention of efficient quantum algorithms [32, 30, 14], which can be executed on computing machines (quantum computers) exploiting fundamental quantum properties of nature, like superposition and entanglement. Such efficient algorithms can achieve important reduction of time complexity, so that in many instances, problems that cannot be solved in polynomial time on a classical computer with the currently known algorithms, can be solved in polynomial time on a quantum computer. A famous example, with important applications in Cryptography, is the factorization of a composite integer into its prime factors (Shor's algorithm)[30]. Another important example is the efficient simulation of quantum physical systems with many degrees of freedom (like a complex chemical molecule), a computation which is not practically achievable in a classical computer [20].

The physical realization of a quantum computer, while in principle is feasible, requires a complex technological effort to overcome practical problems. An important problem is that the carriers of quantum information, the qubits, are very fragile under the influence of their environment and it is very difficult to maintain them in a constant state for a long enough duration so as they can perform a useful computation. The physical carriers of information can be atoms, ions, nuclei and in general any microscopic system on which quantum mechanical effects can be observed[1]. The disturbance effect on the qubits under the environment influence is known as decoherence and can be thought as an environment noise effect. Decoherence problems increase as the number of qubits increases. Additionally, the basic processing elements of qubits, the quantum gates, introduce another factor of disturbance of quantum information, because usually their operation approximates the ideal theoretical operation with errors which don't allow the construction of useful large quantum computers. These introduced errors can be thought as an additional environment induced noise, converting the ideal gates to noisy or erroneous ones. Thus, although real quantum computers have been already developed using various technologies (photons, ion traps, Josephson junctions), they are limited to about 10 qubits [21, 3, 27, 35].

The decoherence problem has been theoretically addressed in the 90's by exploiting and extending results from classical Error Correcting Codes Theory, leading to the invention of Quantum Error Correcting Codes [31, 7, 33]. Such codes can be applied by combining many noisy quantum physical gates so as to build an ideal quantum logical gate, that is they allow the construction of

[1] Currently, some of the most promising are ion traps [9] and Josepshon junction superconductors [36]

fault tolerant quantum gates. This can be accomplished under some conditions, of which the most important is that the noise percentage introduced by each physical quantum gate is lower than a threshold (Quantum Threshold Theorem) [1]. In such a case, an ideal quantum logical gate can be constructed by using redundancy, that is using many physical gates. During the recent years, the effort to build high reliability quantum gates has been intensified, so as to permit the construction of quantum computers of adequate size in the near future. Results of these efforts are very encouraging.

This thesis contributes two-fold:

1. Design of novel efficient quantum circuits (arrays of interconnected quantum logical gates) for integer arithmetic operations and their combination to a higher hierarchy level to achieve more complex arithmetic operations, like modular exponentiation which is an integral part of Shor's algorithm and important algorithms of the same class [24] . The novelty of the proposed circuits lays in the usage of Quantum Fourier Transform (QFT) on the integers states prior to their processing, resulting in improved efficiency in terms of speed. Problems related to the usage of QFT in arithmetic circuits, such as the requirement for high precision quantum gates and the lack of communications locality between the qubits, are also effectively addressed.
2. A generic hierarchical quantum and reversible circuits synthesis methodology [25, 26]. The majority of existing automatic synthesis methods are flat; they operate on the lowest level of gates and while in many cases they lead to optimal or suboptimal results, they have the disadvantage of not being suitable for large circuits as they have exponential requirements in memory usage and run time. The straightforward incorporation of hierarchical synthesis methods into tools of flat methods uses the methodology of Bennett. In contrast, the proposed hierarchical method offers advantages in terms of derived circuit speed and memory, relative to the few hierarchical ones of the literature.

2 Design of Novel Efficient Quantum Circuits

In the context of this thesis, the used gates are assumed to be reliable (logical level) which have been derived from elementary physical quantum gates incorporating any method of error correction. Thus, the thesis concerns the logical level of quantum gates and not the lower level of physical gates. Therefore, the proposed methods of this doctoral thesis can be applied to any technology of physical realization and fault tolerant implementation of logic gates.

We adopt the computation speed, which is known as circuit depth, as the main criterion of efficiency of the proposed methods in this thesis, and it is the number of required steps to complete the computation. This is an important efficiency criterion when construction of large size, in terms of memory, quantum computers become feasible in the future.

The proposed quantum subsystems concern basic arithmetic operations on integers, like multiplication of a constant with an integer and accumulation

(ΦMAC) and division by constant (GMΦDIV) (quotient and remainder calculation) which are used in important quantum algorithms. The implementations is accomplished by using alternative representation of integers in the Fourier domain (that is we use the Quantum Fourier Transform) instead of the usual representation in the computational basis. Quantum circuits using QFT exist in the literature, but they are limited to various kind of adders only [13], while the straightforward implementation of a MAC with Fourier representation using such adders [4] has quadratic circuit depth relative to the integer size. In contrast, the proposed ΦMAC offers linear depth, a considerably important property for large (and thus practically useful) quantum numbers. Regarding the division circuits, just a few quantum dividers exist in the literature and they are chiefly limited to special purposes (e.g. for Galois fields $GF(2^m)$, that is dividers of polynomials with coefficients 0 and 1). A known general quantum divider based on QFT [17] has a cubic depth, while if the divisor is constant its depth can be reduced to be quadratic. The proposed constant divider in this thesis offers a linear depth.

The above two circuits, effectively combined, can be used to construct other more complex circuits useful in various important quantum algorithms. In this thesis we show how it is possible to construct a constant multiplier modulo N (ΦMULMOD), which is a fundamental element for the operation of modular exponentiation. Modular exponentiation is the most time consuming operation in one of the most important quantum algorithms, the factorization algorithm of Shor, and also in other algorithms of the same family. The proposed design achieves a circuit depth of $O(n^2)$, while the majority of the circuits in the literature ranges between $O(n^2 \log n)$ and $O(n^3)$, and consequently the proposed design offers important speed advantage for large numbers. Some of the circuits in the literature offering quadratic or less depth have the disadvantage of increasing excessively the required space (number of qubits) in order or they have the disadvantage of performing approximate calculation.

In the estimation of the circuit efficiency (being in time or space) we must take into account the physical implementation constraints. Such a constraint is the capability of global interactions between the qubits or the limitation of this interaction to neighborhood qubits only, e.g. in a linear one-dimensional array implementation of qubits, where each one can interact only with its two neighbors (1D-LNN, 1D-Linear Nearest Neighborhood). The proposed architecture for Shor's algorithm, while at first sight seems to require global communications between the qubits, it can be adapted in physical machines requiring local interactions with constant overhead in depth, as we show. That is, we don't have any increase in the quadratic order of depth. In contrast, most of the low $O(n^2 \log n)$ depth architectures when applied in a machine that requires local communications increase the depth (e.g. to $O(n^2\sqrt{n})$ in 2D-LNN or to $O(n^3)$ in 1D-LNN) [8].

The Fourier domain processing of the proposed circuits requires the usage of controlled rotation quantum gates with specific angles. A known drawback of such gates is that they do not belong to the category of gates that may constructed fault tolerantly, unless they are decomposed in a sequence of fault

tolerant capable gates (e.g. H and T gates). But, such a decomposition implies considerable overhead in the depth of the whole modular exponentiation circuit up to an order, that is to $O(n^3)$ from $O(n^2)$. Yet, it is possible, as we show, to have a much lesser overhead of $O(n^2 \log n)$ by permitting approximate computation which allow the Shor's algorithm to operate with minor degradation concerning the probability of success. Therefore, the proposed architecture is one of the most competitive in terms of depth, especially if it is applied to 1D-LNN or 2D-LNN physical machines, which are the most probable to be implemented in the future.

3 Hierarchical Synthesis of Quantum and Reversible Circuits

Design of quantum circuits adopts ideas from classical logical design. Small circuits or circuits with repetitive structure can be designed either ad hoc or with formal synthesis methods based on specifications (e.g. truth tables). In the case of quantum circuits there exist similar synthesis methods based on specifications which in the general case are unitary matrices [10, 29]. In special cases where a quantum circuit is described by a matrix with elements exclusively 0 and 1, then reversible circuits[2] synthesis methods can be exploited [28]. Such quantum circuits cases are met when the circuit computes an arithmetic or logical function in the computational basis (e.g. integer addition).

In such cases, these methodologies are suitable for small circuits only, because the required computation power and memory required for their application increases exponentially with the circuit size. The obvious solution is the hierarchical bottom-up design which is applied in classical circuits. In the hierarchical method, if the desired operation can be described as a splicing of simpler operations, the design starts from the lowest level of simpler operations towards the higher level of the more complex operations. The application of the hierarchical method to quantum circuits is possible but requires special handling of the intermediate computation results that are not useful at the end. The particularity is caused due to the fact that these intermediate results cannot be simply discarded at the end because, in general, they are quantum entangled with the desired results. They must be reset to their initial state by inverse computation. Bennett's method is a well known method that keeps the desired results through copying and resets the intermediate results through uncomputation [6]. Its main characteristic and drawback is that it doubles the computation steps (forward computation and the reverse computation) and it also requires more memory space, equal to the space needed by the desired results due to the copying.

The proposed hierarchical synthesis method transforms the initial specifications of the quantum circuit which are given as arrays and arrays of list representing the classical sequence of operation into a directed acyclic graph called

[2] In a reversible circuit, for every possible output, the respective input can be derived, that is no information erasure happens [34].

forward Quantum Dependence Graph (QDG). The nodes of the forward QDG correspond to the components of a quantum library and they suppose to implement the elementary arithmetic operations. These components could be known constructions from the literature (adders etc), synthesized by other low level synthesis method, or populated by the proposed method applied to a lower level. The arcs connecting the QDG nodes correspond to qubits or quantum registers and they are discriminated in arcs which are affected by their successor node and the ones that control their successor node. The final qubits state of the derived forward QDG describes the desired result along garbage results produced during the computation.

The method adopted to reset the garbage states is to apply uncomputation locally on each node that really needs such an inversion of computation, instead to apply it globally as Bennett's method suggests. Namely, nodes of the forward QDG that are effectively involved in garbage production are marked (these are the nodes which have paths with affected arcs towards final garbage states). These marked nodes of forward QDG are traversed backwards and an inverse of each node is appended to the QDG. The inverse nodes are part of the library as it contains quantum circuits whose inverses are assured to exist.

However, data dependencies between the nodes may not always allow such an inversion, in which case we have a deadlock. Two special procedures are applied to detect and resolve such deadlocks (type I and II deadlocks) before the uncomputation stage. Both procedures have the cost to introduce additional ancilla qubits but they never exceed the additional ancilla qubits that would be needed if Bennett's method would be applied.

The proposed synthesis method requires polynomial execution time and memory space in relation to the number of the functions of the specifications and in any case it produces circuits of equal or better performance in terms of depth and space in compared to the basic Bennett's method.

4 Conclusions

Quantum arithmetic circuits based on the QFT representation of integers, instead of the usual computational basis representation, is an alternative implementation that may offer various advantages if used properly. This is due to the fact that two of the main core blocks are the constant adder, which has a constant depth of 1 when the computation is carried out in a datapath that contains an already QFT transformed integer, and the controlled constant adder which has a linear depth of n. By keeping a sequence of computations in such a datapath without reverting back to the computational basis it is possible to maintain a linear depth which otherwise would be impossible. This can be achieved by exploiting properties of the controlled rotation gates such as commutativity, decomposition and suitable rearrangement so as to pipeline their execution. The initial direct QFT and the final inverse QFT does not alter the linear depth as both transforms can be performed in linear depth. Thus, a computation level

of hierarchy can be climbed onto (e.g. in our case, addition to multiplication), without any respective time complexity increase.

Another advantage of using QFT based arithmetic is the lower space requirements. This is manifested in Beauregard's modular exponentiation [4] , where $2n + 1$ qubits are adequate for the full Shor's algorithm. The reason is that no carry computations are needed in the QFT adder as this is done implicitly with the angle additions. While this advantage is not observed in the proposed modular exponentiation circuit due to the divider complexity, it remains in the multiplier/accumulator ΦMAC where no ancilla qubit is used. Also, robustness of such circuits to gate pruning and rotation angle approximation is observed in various instances.

All these remarks suggest that arithmetic circuits, like the proposed ones, are estimable as building blocks for larger and more complex arithmetic circuits.

The obvious follow up to the QFT arithmetic circuits would be to exploit them to derive more complex arithmetic circuits, useful for various quantum algorithms, like the constant divider was for Shor's algorithm. In the same branch of interest, the subject of approximate computations can be further investigated through simulations. The bounds reported in the thesis may be loose and better results may be obtained with numerical simulations. Numerical simulations of the full Shor's algorithm, like the ones performed in [22, 23], are difficult for the case of the proposed circuit because of the requirement of $8n + 2$ qubits. For example, to factor $N = 15$ we would need to simulate $8 \cdot 4 + 2 = 34$ qubits. The joint state vector of 34 qubits consists of $2^{34} \approx 16 \cdot 10^9$ complex elements leading to about 128Gbytes of memory when using single precision floating point, only for the state vector. Yet, partial simulations can be proven useful. A simulation to derive distances between the ΦMAC and an approximated ΦMAC are feasible ($3n+1$ qubits), or even a similar simulation for the whole divider ($6n+1$ qubits).

The hierarchical design method we propose in the doctoral thesis offers advantages relative to Bennett's method in terms of speed and memory of the target circuit. The specifications of the synthesizable circuit are given as a sequence of arithmetic or logic functions. These functions are supposed to be part of a library of quantum circuits. The library can be constructed by using other lower level synthesis methods, or contain known parametrized circuits of the literature (e.g. adders) or be populated with new circuits of the same hierarchical method. Also, the library contains the inverse circuits due to the necessity described above. The end result of the synthesis in the form of directed acyclic graph (Quantum Dependence Graph - QDG) describes the target circuit, where the nodes of the graph represent the modules of the library and the arcs of the graph represent the interconnections between the modules.

Regarding the hierarchical synthesis method, a next obvious step is to develop a complete software which would include front-end and back-end submodules. The front-end must be a compiler accepting the description of the classical algorithm in a suitable language and transforming it in the internal representation required by the synthesis algorithm. The back-end must combine the final QDG representation with information stored in the library so as to export the syn-

thesized circuit in a low gate-level description such as in a quantum assembly format. Equipped with such an integrated tool, we could do a more systematic comparison with other high level synthesis tools, although the advantages of the proposed synthesis methodology are clear even without the tool.

References

1. Aharonov, D., Ben-Or, M.: Fault-tolerant Quantum Computation with Constant Error. In: Proc. 29th Annual ACM Symposium on Theory of Computing (STOC'97). pp. 176–188 (May 1997)
2. Araki, H., Lieb, E.H.: Entropy inequalities. Communications in Mathematical Physics 18(2), 160–170 (1970)
3. Barends, R., Lamata, L., Kelly, J., García-Álvarez, L., Fowler, A.G., Megrant, A., Jeffrey, E., White, T.C., Sank, D., Mutus, J.Y., Campbell, B., Chen, Y., Chen, Z., Chiaro, B., Dunsworth, A., Hoi, I.C., Neill, C., O'Malley, P.J.J., Quintana, C., Roushan, P., Vainsencher, A., Wenner, J., Solano, E., Martinis, J.M.: Digital quantum simulation of fermionic models with a superconducting circuit. Nature Communications 6, 7654:1–7654:7 (Jul 2015)
4. Beauregard, S.: Circuit for Shor's Algorithm Using $2n+3$ Qubits. Quantum Information & Computation 3(2), 175–185 (Mar 2003)
5. Bekenstein, J.D.: Universal upper bound on the entropy-to-energy ratio for bounded systems. Physical Review D 23, 287–298 (Jan 1981)
6. Bennett, C.H.: Logical Reversibility of Computation. IBM J. Research and Development 17(6), 525–532 (Nov 1973)
7. Calderbank, A.R., Shor, P.W.: Good quantum error-correcting codes exist. Physical Review A 54, 1098–1105 (Aug 1996)
8. Choi, B.S., Van Meter, R.: On the Effect of Quantum Interaction Distance on Quantum Addition Circuits. ACM J. Emerging Technologies in Computing Systems 7(3), 11:1–11:17 (Aug 2011)
9. Cirac, J.I., Zoller, P.: Quantum Computations with Cold Trapped Ions. Physical Review Letters 74, 4091–4094 (May 1995)
10. Cybenko, G.: Reducing Quantum Computations to Elementary Unitary Operations. J. Computing in Science and Engineering 3(2), 27–32 (Mar 1996)
11. Deutsch, D.: Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 400(1818), 97–117 (Jul 1985)
12. Deutsch, D.: Quantum Computational Networks. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 425(1868), 73–90 (Sep 1989)
13. Draper, T.G.: Addition on a Quantum Computer. eprint arXiv:quant-ph/0008033 (Aug 2000)
14. Grover, L.K.: A Fast Quantum Mechanical Algorithm for Database Search. In: Proc. 28th Annual ACM Symposium on Theory of Computing (STOC'96). pp. 212–219 (May 1996)
15. Kholevo, A.S.: Bounds for the quantity of information transmitted by a quantum communication channel. Problemy Peredachi Informatsii (Problems of Information Transmission) 9(3), 3–11 (1973)
16. Kholevo, A.: On the capacity of a quantum communication channel. Problemy Peredachi Informatsii (Problems of Information Transmission) 15(4), 3–11 (1979)

17. Khosropour, A., Aghababa, H., Forouzandeh, B.: Quantum Division Circuit Based on Restoring Division Algorithm. In: Proc. 8th International Conference on Information Technology: New Generations (ITNG '11). pp. 1037–1040 (2011)
18. Lieb, E.H., Ruskai, M.B.: Proof of the strong subadditivity of quantum-mechanical entropy. Journal of Mathematical Physics 14(12), 1938–1941 (1973)
19. Lindblad, G.: Completely positive maps and entropy inequalities. Communications in Mathematical Physics 40(2), 147–151 (1975)
20. Lloyd, S.: Universal Quantum Simulators. Science 273(5278), 1073–1078 (Aug 1996)
21. Monz, T., Nigg, D., Martinez, E.A., Brandl, M.F., Schindler, P., Rines, R., Wang, S.X., Chuang, I.L., Blatt, R.: Realization of a scalable Shor algorithm. Science 351(6277), 1068–1070 (Mar 2016)
22. Nam, Y.S., Blümel, R.: Robustness and performance scaling of a quantum computer with respect to a class of static defects. Physical Review A 88, 062310 (Dec 2013)
23. Nam, Y.S., Blümel, R.: Streamlining Shor's algorithm for potential hardware savings. Physical Review A 87, 060304 (Jun 2013)
24. Pavlidis, A., Gizopoulos, D.: Fast Quantum Modular Exponentiation Architecture for Shor's Factoring Algorithm. Quantum Information & Computation 14(7&8), 649–682 (May 2014)
25. Pavlidis, A., Gizopoulos, D.: Hierarchical synthesis of quantum and reversible architectures. In: Proc. 12th ACM International Conference on Computing Frontiers (CF'15). pp. 13:1–13:8 (2015)
26. Pavlidis, A., Gizopoulos, D.: Hierarchical synthesis of quantum and reversible architectures. International Journal of Parallel Programming 44(5), 1028–1053 (Oct 2016)
27. Politi, A., Matthews, J.C.F., O'Brien, J.L.: Shor's quantum factoring algorithm on a photonic chip. Science 325(5945), 1221–1221 (Sep 2009)
28. Saeedi, M., Markov, I.L.: Synthesis and Optimization of Reversible Circuits - A Survey. ACM Computing Surveys 45(2), 21:1–21:34 (Feb 2013)
29. Shende, V.V., Bullock, S.S., Markov, I.L.: Synthesis of quantum-logic circuits. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 25(6), 1000–1010 (Jun 2006)
30. Shor, P.W.: Algorithms for Quantum Computation: Discrete Logarithms and Factoring. In: Proc. 35th Annual IEEE Symposium on Foundations of Computer Science, (FOCS'94). pp. 124–134 (Nov 1994)
31. Shor, P.W.: Scheme for reducing decoherence in quantum computer memory. Physical Review A 52, R2493–R2496 (Oct 1995)
32. Simon, D.: On the power of quantum computation. In: Proc. 35th Annual IEEE Symposium on Foundations of Computer Science, (FOCS'94). pp. 116–123 (Nov 1994)
33. Steane, A.: Multiple-Particle Interference and Quantum Error Correction. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 452(1954), 2551–2577 (Nov 1996)
34. Toffoli, T.: Reversible computing. Tech. Rep. MIT/LCS/TM-151, Massachusetts Institute of Technology, Laboratory for Computer Science (Feb 1980)
35. Vandersypen, L.M.K., Steffen, M., Breyta, G., Yannoni, C.S., Sherwood, M.H., Chuang, I.L.: Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. Nature 414(6866), 883–887 (Dec 2001)

36. Vion, D., Aassime, A., Cottet, A., Joyez, P., Pothier, H., Urbina, C., Esteve, D., Devoret, M.H.: Manipulating the Quantum State of an Electrical Circuit. Science 296(5569), 886–889 (May 2002)

Tools and Methods for System of Systems Applications in Telecommunication Networks[1]

Kosmas-Christos F. Tsilipanos

Department of Informatics and Telecommunications
National and Kapodistrian University of Athens, 15784 Ilissia Athens, Greece
ktsilipanos@di.uoa.gr

Abstract. This Ph.D. Thesis discusses the System of Systems (SoS) concept adjustment into information and communication technology (ICT) relative fields. The research objective of this Ph.D. thesis is to provide new methods and tools in ICT related enterprises and professionals, aiming into proper analysis and modelling of telecommunications networks. This will enhance the current methodologies and frameworks to battle current challenges such as increased complexity, scaling and uncertainty without of course following a static concept analysis. Pilot studies over a variety of ICT subjects also conducted in order to verify validity and cogency over the developed tools and methods.

Keywords. System of Systems, Reliability, Optimization, Adaptation, Emergence, Reconfiguration

1 Dissertation Summary

1.1 Introduction

Recently a significant interest was observed in the field of System of Systems. Particular attention has been paid on System of Systems Engineering which deals with the development of tools, methods and solutions to the challenges of System of Systems. Nevertheless, for better understanding a SoS, one needs to clarify the characteristics that should be included in it. A System of Systems should consist of a number of operationally and managerially independent systems. The individual systems should be autonomous that is to perform a standalone operation. On the other hand, the independent systems should be integrated in a higher level system (meta-system) to perform

[1] Dissertation Advisor: Dimitris Varoutas, Assistant Professor

a mission / purpose for which each member plays an integral role. Another main feature of System of Systems is that it is a complex system and as such exhibits a dynamic and emergent behavior. Finally, the System of Systems is a dynamic entity as new systems are added and current systems are replaced or removed. In the following sub chapters of the dissertation summary, some pilot networks were SoS concept was exercised will be presented. In detail, the constituent systems of the SoS meta-system will be denoted while outlining the major characteristics of the System of System Engineering theory. For each one of these pilot networks a method/tool was researched and developed in order to enhance knowledge of the SoS while confronting uncertainty. Results on each research study are presented and various details are also disclosed.

1.2 Reliability Study using System of Systems concept.

Service continuity is becoming a critical path for the delivered quality of service (QoS) [1]. Towards this end, telecommunication companies are continuously investing in research and development for reliability analysis. However, the provision of uninterrupted services that require high bandwidth and interactivity often leads to increased complexity of telecommunication systems. This converts the new generation of telecommunications networks into "Systems of Systems"[2], [3], consisting of a mixture of software, hardware and human intervention. The starting point for the development of a new method/tool is to identify its requirements. The System of Systems method/tool should primarily address the high complexity, the large size as well as the dynamic and emergent behavior of telecommunication network. Moreover, it needs to be flexible in order to incorporate and evaluate the impact of unknown events. Following these guidelines, this Ph.D. thesis proposed a new SoS framework[4] for the reliability assessment of telecommunication networks. The proposed framework is a combination of the analytical method of Hazard and Operability Analysis (HAZOP)[5], [6] with the mathematical representation of Fault Tree Analysis (FTA)[7], [8] along with the directed acyclic graphs (DAG) of Bayesian networks (BN)[9]–[11]. In addition, this method encapsulates sensitivity analysis techniques (Monte Carlo Simulations)[8] in order to quantitatively evaluate the impact of unknown

risks and events such as the addition of new systems with unknown characteristics. In order to further reveal real SoS behaviors (e.g. evolution – emergent behavior of SoS), exploratory modeling is implemented. The telecommunication network under investigation is a fiber to the curb (FTTC) access network based on VDSL technology (**Fig. 1**). The network has five independent systems. The Customer Premises Equipment (CPE), the Digital Subscriber Line Access Multiplexer (DSLAM), the Local Exchange (LE), the Central Office (CO) and the Broadband Remote Access Server (BBRAS).

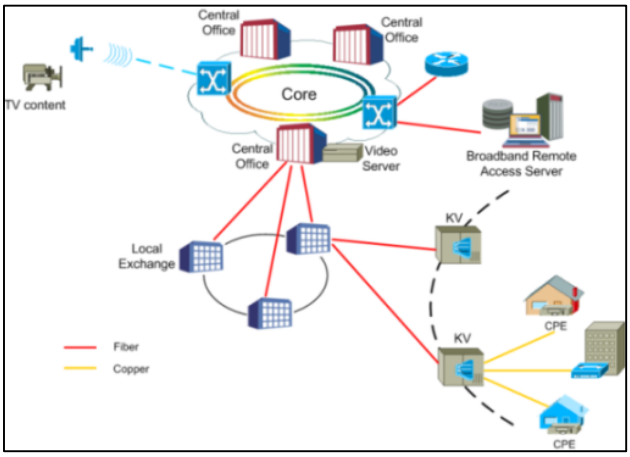


Fig. 1. System under investigation

1.3 Results and Discussion over Reliability Study using SoS.

The input parameters of the obtained model is the exposure time T and the failure rate λ that is the inverse of the mean time between failures (MTBF). In this study, the exposure time T considered arbitrarily as 10 years. However, any other value of the exposure time can be easily introduced in the proposed framework. The mean time between failures (MTBF) that were used in the present study, have been mined from the database of a build in techno-economic tool containing numerous network components [1].Probability of Failure of each system is then estimated (**Fig. 2**). Apart from the Probability of Failure, several importance metrics can also be derived through FTA analysis. Moreover, in order to incorporate complex and unknown events as well as to extract useful guidelines, one should resort to Bayesian network models. Such models will provide us the asset of answering complex probabilistic

queries about the network operation, using the information about nodes and interfaces obtained from the HAZOP and FTA analyses.

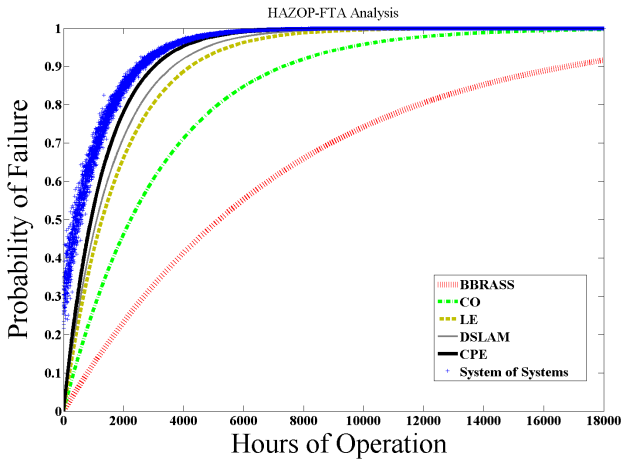


Fig. 2. Probability of failure of the SoS constituent systems for 18000 hours of operation

The inclusion of Bayesian Networks in the Framework of the study makes feasible the generation of quantitative results that deal with complex and unknown events. This way deep uncertainty is also modeled through the proposed Framework. Bayesian Network parameters are actually embedded over the Conditional Probability Tables (CPTs) of each system. Using the BN of each SoS representation, several conditional probabilities can be calculated such as the probability of failure of the SoS given that LE is failing as follows (Eq.1).

$$P(SoS = True \mid CPE, DSLAM, LE = True, CO, BRASS) = \\ = \frac{\sum_{C,D,CO,B \in [True,False]} P(SoS = T, C, D, L = T, CO, B)}{\sum_{SoS,C,D,CO,B \in [True,False]} P(SoS, C, D, L = T, CO, B)} \quad [2] \qquad (1)$$

In order to further study the emergent behavior of the SoS under investigation, one should resort to techniques dealing with large degrees of uncer-

[2] Where CPE, DSLAM, LE, CO, BBRASS, True and False are written as C, D, L, CO, B, T and F for simplicity.

tainty (deep uncertainty). Under conditions of deep uncertainty, it is hard to forecast the realizations and the time-varying relationships of relevant factors in the SoS. Furthermore, these situations of uncertainty can be occurred in a system that has not yet existed. The latter case is examined in this subsection since new versions of the SoS can be obtained through the addition/deletion/replacement of systems or links. Unfortunately, in deep uncertainty the appropriate conceptual models to describe interactions among SoS variables as well as the probability distributions to represent uncertainty about key parameters in the models are unknown. The numerical simulations include evaluations of model outcomes across a large set of possible SoS representations. Each plausible SoS representation can be assumed as one hypothesis about SoS behavior. By investigating a large set of such hypotheses and by evaluating their correctness, the "whole picture" of SoS emergent behavior can be obtained. A simple method to construct the possible SoS models is to include time in the model specification defining the evolution of the SoS model. The evolutionary modeling of SoS is depicted in **Fig. 3**.

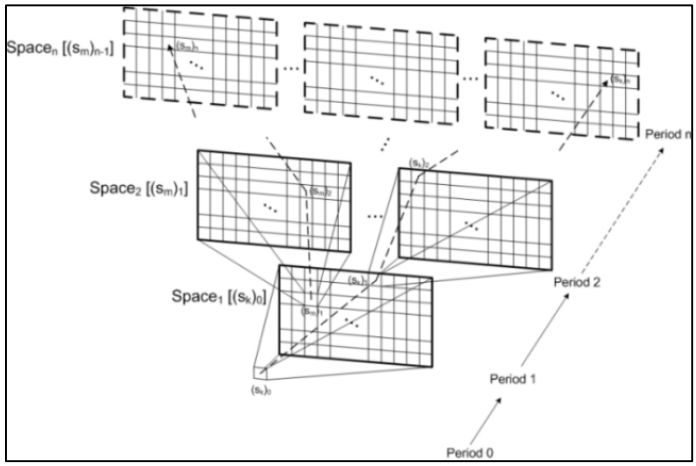


Fig. 3. Network Realization along with Uncertainty spaces

As shown in **Fig. 3**, there is a dependence of uncertainty spaces in a period i on the realizations of all the preceding periods. For example in Period 2, the uncertainty space ($\mathrm{Space}_2\ [(s_m)_1]$) is generated from the realization $(s_m)_1$ of the preceding uncertainty space ($\mathrm{Space}_1\ [(s_k)_0]$) which in turn is originated from the initial condition $(s_k)_0$. Hence, a possible future path (dashed lines) can be represented by the sequence $(s_k)_0 \rightarrow (s_k)_1 \rightarrow (s_k)_2 \ldots \rightarrow (s_k)_n$. In **Fig. 4**,the uncer-

tainty space 2 (after four years) originating from the realization[3] (80% $\lambda_{1,in}$,110 % $\lambda_{2,in}$, 0.3, 0.5) of the previous period (2 years) is illustrated as a color map where different colors correspond to different probability values estimated using (Eq.2).

$$P(SoS = True | CPE = False, DSLAM, LE, CO, BRASS) = \frac{\sum_{D,L,CO,B \in [True,False]} P(SoS = T, C = F, D, L, CO, B)}{\sum_{sos,D,L,CO,B \in [True,False]} P(SoS, C = F, D, L, CO, B)} \quad (2)$$

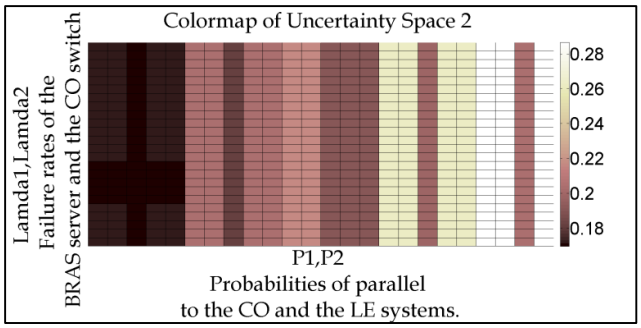


Fig. 4. One of the possible SoS representations after four years (period 2).

In order to simplify **Fig. 4**, the coordinates of each realization are omitted. However, a method to determine the coordinates of a realization *k* (*i*-th line and *j*-th column of **Fig. 4**, is provided by the following equations.

$$\lambda_{1,k} = \lambda_1\left(\left\lfloor \frac{i-1}{5} \right\rfloor + 1\right) \quad (3)$$

$$\lambda_{2,k} = \lambda_2\left(\mathrm{mod}\left((i-1),5\right)+1\right) \quad (4)$$

$$P_{1,k} = P_1\left(\left\lfloor \frac{j-1}{5} \right\rfloor + 1\right) \quad (5)$$

$$P_{2,k} = P_2\left(\mathrm{mod}\left((j-1),5\right)+1\right) \quad (6)$$

[3] The values for the parameters λ1, λ2, P1, P2 are from CPTs, where $\lambda_{i,in}$ are the original values of the component's failure rate used in the current research.

where $\lfloor x \rfloor$ denotes the integer part of x and $\lambda_i(m)$, $P_i(m)$ represent the m-th entries of λ_i and P_i respectively in CPTs. Using the color map **Fig. 4**, one can predict the impact of possible changes (components replacement and/or addition or deletion of links – systems) in the infrastructure by estimating the values of conditional probabilities and determine the emergent behaviors of the SoS by exploring the corresponding paths.

1.4 Optimization and Management of Complex Telecom Investments using System of Systems concept.

The high importance of Information and Communications Technologies for Europe can be viewed by its action to include Digital Agenda in Europe 2020 Strategy. Telecom operators who are the main investors of telecommunications networks have thus to exploit the funding opportunities from the one hand and deal with the high uncertainty influencing such deployments. Hence accurate, quick choices along with budget details and strategy plans need to be constructed on a clear basis. A quick and accurate methodology scheme that will make investors to enhance current methodologies and decide with greater confidence about critical aspects of their investment should consequently be provided. As telecom investments are difficult to be modelled by traditional systems methodologies, given their space and time scale, their multi-dimensional nature, their complexity, the uncertainties arising from demand and price evolution and the emerging needs of users, new approaches incorporating complexity while keeping computational simplicity are needed. In addition, qualitative issues such as the first movement advantage and externalities deriving from the associated networked economies are even more difficult to be incorporated in a simple and accurate manner. Earlier approaches to implement the SoS concept in a techno-economic problem for telecom networks providing a complete and accurate reference to telecom operators and policy makers can be found in [12]. Although this work managed to address the emerging behavior of telecom investments, it could not deal with the complex interdependencies of the constituent systems as well as the externalities arising from the associated networked economies. In this Ph.D. thesis, a methodology based on System of Systems (SoS) framework[13] proposed for modeling telecom investments and defin-

ing strategies leading to profitability under several constraints. Adaptation and reconfiguration concepts on initial decided strategies are also encapsulated in this framework. Having a compact and almost closed-form nature, the proposed framework can be proved an extremely valuable tool for telecom operators. As shown in **Fig. 5**, there are five interdependent systems. The Competitor Analysis (CA), Budget Allocation (BU), Capital and Operational expenses (CAPEX + OPEX), Demand Forecast (DE) and Network Externalities (NE). It should be noted that all the constituent systems are able to interact with each other. In this study, the dependence of customers' number on external factors such as price is incorporated in network externalities system. This, along with the rest systems, will define the expected revenues of the network.

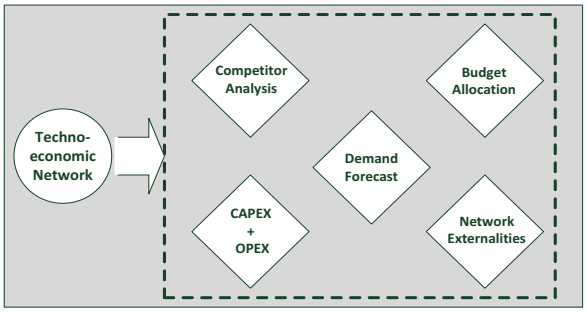


Fig. 5. Techno-economic Network under investigation, typical representation

1.5 Results and Discussion over Techno-economic Study using SoS

FTTH architecture for the last mile is investigated as a case study of the proposed framework. The area is described in terms of subscriber density and geographical characteristics. The area model chosen corresponds to a dense urban area with a surface of 12 km^2 and 5,641 customers per km^2. It is assumed that there is one central office, serving 65,536 customers in total. The total available budget is assumed to be 35M€ for the whole study period of 10 years maximum. The starting year of operator's investment will be decided by the optimization process using genetic algorithms (GAs)[14]. In order to avoid trapping on local extrema, the GA is running 1000 times. In order to maximize provider's profit at the end of the study period, the following non-

linear programming problem is stated following the objective Function as in the indicative equation (Eq.7). Using the estimated parameters along with constituent systems' mathematical models, an optimum strategy for the incremental network deployment (dynamic budget allocation) leading to profit maximization at the end of the study period will be investigated and proposed. The nonlinear programming (NLP) problem is solved using a GA and implemented in Matlab.

Maximize objective Function H(x).

Subject to :

Incentives

Budget

Subscribers

Pricing Complicate Model

Variables: (7)

X0, Starting Year of Investment

X1, X2 competitors entrance case analysis

X3..X12 ,each year of the total 10 of study, for yearly service pricing

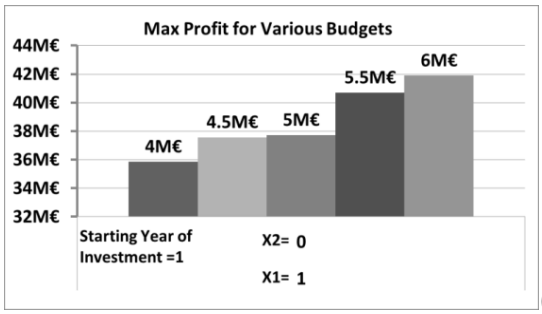
Contraints: $1 \leqslant X0 \leqslant 10$

$0 \leqslant X1, X2 \leqslant 1$

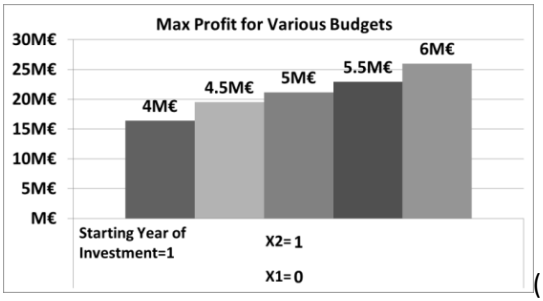
$X3 \leqslant X4 \ldots X11 \leqslant X12$

From the derived results, it is deduced that in almost 90% of the cases the FTTH investment is started in the first two years. This is somehow expected and can be attributed to the longer network's operation period leading to increased revenues and thus profits. A profitability of ~100 M€ is observed in the majority of the simulation runs. Another interesting result is that in more than half of the 1000 cases, the operator decides to invest before his competitor. This can be mainly explained by the extra benefits (incentives) received by the first investor. In these cases, a maximum profit of 103 M€ is observed. Moreover, a delayed decision (invest after the competitor) results in significant profit losses. An interesting case is the simultaneous investment of both operators which is the second best choice as the profitability remains in good levels. One more interesting point is that in the majority of cases where we have a significant profit over the years of study, profitability is observed after the fourth year. The available total budget is a point of great

debated and should be further investigated. Towards this direction, a series of simulation runs were performed assuming -20%, -10%, 10% and 20% budget change respectively. The obtained results are illustrated in **Fig. 6**(a) and **Fig. 6**(b). It is straightforward to understand that a bigger amount of budget is required to further expand the network, needed to meet the high demand due to price reduction policy that was followed. In the analysis of the previous sections, it is assumed that the profits of the investment are shared to the shareholders of the operator's company. However, this strategy is proved to be the worst in terms of available budget. In order to investigate the impact of partial profit re-investment on both the required available budget and the viability of the project, eq. (8) must be transformed to the following form. Assuming an available budget of 25 M€, a series of optimization runs are performed again.



(a)



(b)

Fig. 6. (A) Pivot Chart showing Maximum Profit when Starting at Year 1, BEFORE Competitor for Various Budgets (B) Pivot Chart showing Maximum Profit when Starting at Year 1, AFTER Competitor for Various Budgets

$$\sum_i^{years} Bi \leq \text{BUTotal} + \sum_i^{years} (A\,\%) * Profit \qquad (8)$$

A= [30, 50]

From the obtained results, it is deduced that, the maximum profit (around 93 M€) is achieved by re-investing 50% of the annual profit in network extensions besides the small amount of available budget. The case of starting the investment after competitor is also studied and it produces really low profits. This is a high risk case since if the operator under study misses the investment entrance, a severe amount of money for budgeting will be needed in order to gain market share capable of providing adequate profits.

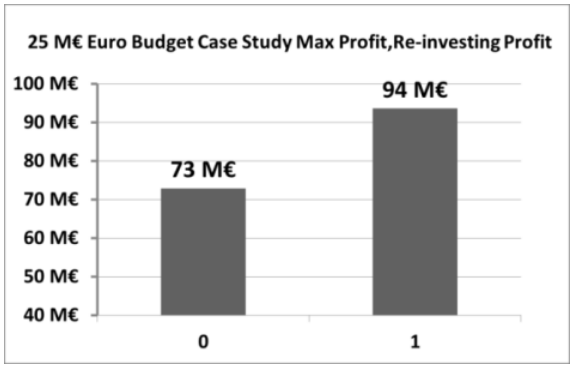


Fig. 7. : Maximum Profit of investment with 25 M€ budget. Starting year of investment is 1. X=0 equals to after competition investing, and X=1 equals to start investing before competition. Assuming aggressive investment in network construction of yearly profits.

In order to compare the results derived with and without re-investment, a series of optimization runs are also performed using eq. (8). The maximum profit was ~70 M€, which is much lower than the case of profit reinvestment. This can be easily explained by the fact that low available budget results in limited network deployment. Thus, the growing demand cannot be met by keeping the number of subscribers in low levels. It should also be highlighted that the same conclusions are derived following the advantageous strategies of pricing policy and investment entrance. This is a clear indication that telecom investments are not static procedures. Contrary, more complicated strategies and adaptive policies should be adopted. This is of high importance, especially in cases where the initial targets are missed over the study period. In order to address such emergent behaviors of the SoS under investigation, it is thus crucial to continuously monitor the defined metrics of effectiveness and performance of a successful investment. Interventions, in terms of pricing policy or percentage of re-investment, deviating from the

initial plan should be performed in order to keep profitability in the desired levels.

2 Conclusions

This thesis and the relative published research material[4] contribute in the reliability study field using the SoS methodology concept with the following:

- All type of risks/hazards covered thru the combination of reliability analyses.
- Enables dynamic characteristics in the proposed reliability analysis Framework, through the Exploratory Modelling as an emergent need of complex systems analysis. Meanwhile with the specific modelling encapsulation the reconfiguration risk/hazard is revealed as part of the emergence of the complex networks analysis.
- Both qualitative and quantitative approach.
- Inclusion of Uncertainty Space in the proposed Reliability Analysis, by using sensitivity analyses. Purpose is to ensure unknown events are taken into account over the cluster of residual mishap risk.

The proposed methodology can be used for evaluation of network performance and monitoring of service quality and service level agreements (SLAs) in a telecommunication network. It can also be exploited in techno-economic studies in order to evaluate the cost of operation, administration and maintenance (OAM). Additionally, in the techno-economics field, this Ph.D. thesis and the relative published material[13] contribute in the specific field the below outlined items:

- Multiple Representation Scenarios of the telecom investment.
- Exhaustive search and strategy alignment based on each case specific action needs and.
- The Optimization Study is also capable for reconfiguration towards decision making enhancement (market or technology related).
- Encapsulation for Dynamic Budget Allocation and Re-investment scenario.

The above mentioned contributions are part of generic methods and tools targeting also over a general Framework. This enables the usage of the developed methods in various fields that are directly related to complexity. Some examples that could follow the SoS engineering theory are the Infor-

mation Systems (IS) and Smart Cities. Both, belong to the author's future research fields in line with the implementation of the SoS methods/tools developed as part of this research effort.

References

[1] W. C. Hardy and L. Preface By-Cardoso, *QoS: measurement and evaluation of telecommunications quality of service*. John Wiley & Sons, Inc., 2001.

[2] W. C. Baldwin, B. J. Sauser, and J. Boardman, "Revisiting 'The Meaning of Of' as a Theory for Collaborative System of Systems," 2015.

[3] W. C. Baldwin and B. Sauser., "Modeling the Characteristics of System of Systems," in *SoSE*, 2009, pp. 1–6.

[4] K. Tsilipanos, I. Neokosmidis, and D. Varoutas, "A system of systems framework for the reliability assessment of telecommunications networks," *Syst. Journal, IEEE*, vol. 7, no. 1, pp. 114–124, 2013.

[5] J.-G. Hwang, H.-J. Jo, and D.-H. Kim, "Hazard analysis of train control system using HAZOP-KR methods," in *Electrical Machines and Systems (ICEMS), 2010 International Conference on*, 2010, pp. 1971–1975.

[6] L. Wei, Z. Laibin, and H. Jinqiu, "Fuzzy information fusion based quantitative HAZOP analysis for gas compressor units," in *Intelligent Systems, 2009. GCIS'09. WRI Global Congress on*, 2009, vol. 2, pp. 423–427.

[7] S. Liu and Z. Han, "Notice of Retraction Reliability Analysis of transformer Based on FTA And Mente Carlo Method," in *Power and Energy Engineering Conference, 2009. APPEEC 2009. Asia-Pacific*, 2009, pp. 1–3.

[8] Z. Liming, Y. Jianwei, C. Guoqiang, and J. Limin, "Monte-Carlo simulation based on FTA in reliability analysis of door system," in *Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on*, 2010, vol. 5, pp. 713–717.

[9] B. Cai, Y. Liu, and Q. Fan, "A multiphase dynamic Bayesian networks methodology for the determination of safety integrity levels," *Reliab. Eng. Syst. Saf.*, vol. 150, pp. 105–115, 2016.

[10] Q. Li, M. Jiang, H. Li, and M. Lu, "Software reliability qualitative evaluation method based on Bayesian networks," in *Education Technology and Computer (ICETC), 2010 2nd International Conference on*, 2010, vol. 4, pp. V4--446.

[11] Z. Yongli, H. Limin, Z. Liguo, and W. Yan, "Bayesian network based time-sequence simulation for power system reliability assessment," in *Artificial Intelligence, 2008. MICAI'08. Seventh Mexican International Conference on*, 2008, pp. 271–277.

[12] T. Rokkas, I. Neokosmidis, D. Katsianis, and D. Varoutas, "Cost analysis of WDM and TDM fiber-to-the-home (FTTH) networks: A system-of-systems approach," *Syst. Man, Cybern. Part C Appl. Rev. IEEE Trans.*, vol. 42, no. 6, pp. 1842–1853, 2012.

[13] K. Tsilipanos, I. Neokosmidis, and D. Varoutas, "Modeling Complex Telecom Investments: A System of Systems Approach," *Eng. Manag. IEEE Trans.*, vol. 62, no. 4, pp. 631–642, 2015.

[14] A. J. Chipperfield, P. Fleming, and H. Pohlheim, *Genetic Algorithm Toolbox: For Use with MATLAB; User's Guide (version 1.2).* University of Sheffield, Department of Automatic Control and Systems Engineering, 1994.

Timing Error Detection and Correction for Reliable Integrated Circuits in Nanometer Technologies

Stefanos Valadimas *

Department of Informatics and Telecommunications
National and Kapodistrian University of Athens
s.valadimas@di.uoa.gr

Abstract. Timing error tolerance turns to be an important design parameter in nanometer technology, high speed and high complexity integrated circuits. This thesis presents three concurrent on-line timing error tolerance techniques which enhance circuit's reliability. The proposed techniques detect and correct timing errors efficiently, in flip-flop based designs, with low power consumption and low silicon area overhead. To validate the three novel techniques, they have been applied in the design of a 32-bit MIPS R2000 pipeline microprocessor.

Keywords: concurrent on-line testing, timing errors, error detection and correction, timing error tolerance, reliability-aware design.

1 Introduction

As technology scales down timing errors are a real concern in high complexity and high frequency integrated circuits. Process, Voltage and Temperature (PVT) variations [1] lead to large spreads in delay, at the system level, which undermine circuit's reliability. Moreover, crosstalk [2], power supply disturbances and resistive IR-drop [3] affect circuit performance increasing the overall impact of timing errors.

In addition, aging mechanisms [4] cause gradual speed degradation of the designs over their service life, mainly due to Bias Temperature Instability (BTI) [5], which is one of the most important phenomena that degrade the performance of nano-scale circuits. BTI primarily accelerates the aging process of MOS transistors by increasing their threshold voltage. BTI-induced delay shifts in logic paths, are related to timing violations during the circuit lifetime.

The increased path delay deviations, due to the above factors, result in timing errors that are not easily detectable in terms of test cost. To mitigate the impact of nanometer scaling, conservative approaches, with wider safety margins, are adopted to guarantee the reliability during system lifetime. In this context, it is evident that timing error tolerance techniques are becoming necessary to provide robustness against timing violations and meet system reliability requirements.

* Dissertation Advisor: Angela Arapoyanni, Professor.

2 Previous Solutions In Error Tolerance

A number of error tolerance techniques have been proposed for flip-flop and latch based designs. Aiming the detection of errors the techniques proposed in [6] and [7] sense the delayed circuit response and provide error tolerance using time redundancy approaches. A well-known and commonly used scheme for flip-flop based designs is the Razor pipeline architecture [8]. The Razor flip-flop consists of the main system flip-flop plus an assistant shadow latch, a multiplexer and a XOR gate (Fig. 1). The shadow latch is clocked by a delayed version of the system clock in order to capture delayed responses of the combinational logic. The XOR gate compares the outputs of the main flip-flop and the shadow latch for error detection. Whenever a timing error occurs the correct data, which are stored in the shadow latch, are injected into the pipeline during the next clock cycle. For every main flip-flop an extra latch, a multiplexer and a XOR gate are required. Hence, this approach suffers from high power consumption and high silicon area cost. Moreover, a metastability detector is required to guarantee high levels of reliability.

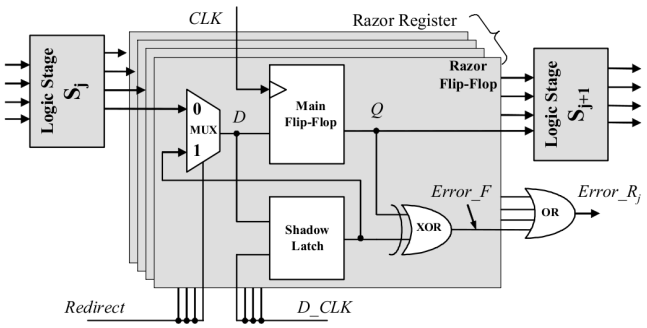


Fig. 1: The Razor flip-flop.

For latch based designs a modified version of the Razor topology (Razor II) is presented in [9]. Its application to a 32-bit ARM microprocessor is discussed in [10]. Also in that case a transition detector is used, at the output of the latch, for error detection while error correction is performed through architectural replay. Another solution to enhance tolerance for latch based designs is the GRAAL architecture [11]. It is based on the XOR comparator for error detection and an additional flip-flop per latch for error correction.

An alternative approach, which masks timing errors by borrowing time from successive pipeline stages, is presented in [12]. An additional latch per main system flip-flop is used to re-sample the input data with a proper delay. Various double-sampling architectures are discussed in [13].

3 Time Dilation Technique

In this section, the first proposed error detection and correction technique is presented. The Time Dilation [14] technique exploits a new scan flip-flop which supports both the standard off-line scan testing capability as well as the on-line (concurrent) error detection and correction capability. According to the proposed technique, after error detection the evaluation time for the logic is automatically extended by a single clock cycle for error correction using correct and valid data stored in each flip-flop. Unlike earlier solutions, no extra memory elements are required in the Time Dilation approach.

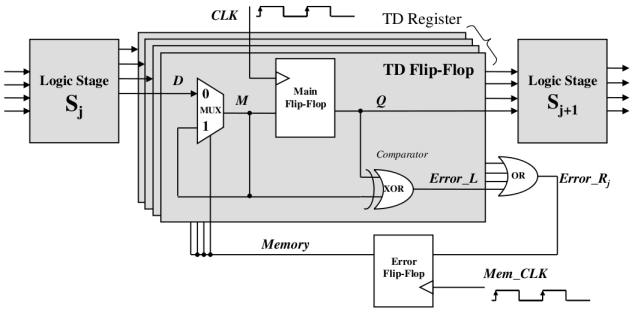


Fig. 2: The Time Dilation flip-flop.

The Time Dilation flip-flop (TD flip-flop) [14] is presented in Fig. 2. This topology utilizes a multiplexer (MUX) and a XOR gate per system flip-flop (Main flip-flop) to provide timing error detection and correction capabilities. The XOR gate compares the input and the output of the Main flip-flop for error detection, while the multiplexer with the feedback configuration forms an extra memory element (a MUX-latch) that captures delayed valid data for error correction. After error detection the logic evaluation time is extended by a clock cycle for error correction, by re-feeding the Main flip-flop with the correct and valid data of the MUX-latch.

The operation of the new flip-flop is quite simple. Initially, the Error flip-flop is reset to low, so that the TD flip-flop is in the normal mode of operation and the D input feeds the Main flip-flop. In the fault free case, the data arrive in time at the D input of the TD flip-flop, they propagate to the M input of the Main flip-flop and they are captured at the Q output by the triggering edge of the clock signal CLK. After the triggering edge the inputs of the XOR gate (signals M and Q) hold the same logic value and the output signal Error_L of the XOR gate is low (no error detection). Consequently, the Error flip-flop retains the low state at its *Memory* output, after the triggering edge of the MEM_CLK clock signal, and the TD flip-flop remains in the normal mode of operation.

The *MEM_CLK* clock signal is a delayed version of the *CLK* clock signal. However, in the presence of a timing failure, which results in a delayed arrival of the data at the signal lines D and M, the logic values on M and Q differ after the triggering edge of the clock signal *CLK*. Thus, the signal Error_L is high indicating an error detection. Consequently, the register error indication signal Error_R will be also high and the same stands for the *Memory* signal after the triggering edge of the *MEM_CLK* clock signal. As a result, the MUX-latch enters the memory state of operation capturing the delayed but correct data at the M input of the Main flip-flop. These correct data feed the Main flip-flop at the next triggering edge of *CLK* for error correction and circuit operation recovery.

The hardware overhead and the power consumption of the TD flip-flop is much lower than this of the Razor topology, since in the latter topology except of the multiplexer and the XOR gate an additional shadow latch is required.

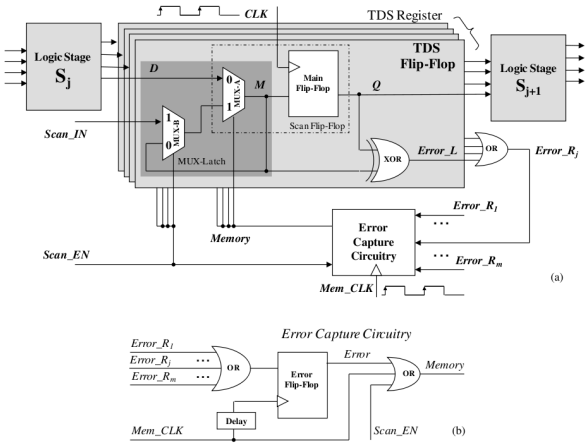


Fig. 3: (a) TDS flip-flop and support circuitry and (b) error capture circuitry.

The scan version of the Time Dilation architecture is presented in Fig. 3. The Time Dilation Scan flip-flop (TDS flip-flop) provides error detection and correction capabilities by appending only a multiplexer (MUX-B) which is utilized for the scan operation, as in a standard scan design. When the scan enable signal (Scan_EN) is "high" the TDS flip-flop operates like a Scan flip-flop to support off-line scan testing procedures. At the same time the *Memory* signal must be also "high". Consequently, the test data are propagated from the Scan_IN port to the input line M of the main flip-flop where they are captured. Then, they are provided through the Q line to the Scan_IN port of the next flip-flop in the chain and so on. In the normal mode of operation (where Scan_EN is "low") the TDS flip-flop behaves like an ordinary flip-flop enhanced with the ability to detect and correct timing errors as it has been analyzed above.

As in the Razor technique, a crucial issue in the proposed technique is the possible existence of short (fast) paths in the combinational logic which may corrupt the data in the MUX-latches. This is the well known hold time problem. As fast paths we define paths with response times inside the monitoring window. To avoid the hold time problem, a minimum path delay constraint is proposed in Razor. This constraint is fulfilled adding delay buffers during logic synthesis to slow down short paths (paths padding).

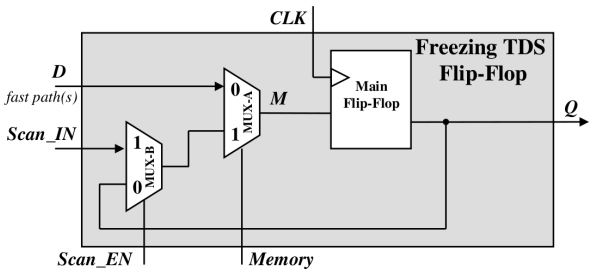


Fig. 4: Freezing TDS flip-flop.

Path padding techniques can be also applied in the proposed TDS technique. However, aiming to reduce the pertinent cost, an alternative design approach can be used instead. The Freezing TDS flip-flop in Fig. 4 is exploited at the end of fast paths that do not intersect with critical paths. Delay buffers are inserted only in the rest fast paths that intersect with time critical paths in order to avoid data corruption in the MUX-latches of the standard TDS flip-flops. In those cases, the minimum path delay constraint is equal to the delay of the *Memory* signal with respect to the system clock *CLK*, plus the hold time of the MUX-latch.

The operation of the Freezing TDS flip-flop in Fig. 4 is based on the fact that the data captured by a flip-flop at the end of a fast path are always correct since they are not affected by timing failures. Consequently, the comparator (XOR gate) is eliminated. The main difference in this new topology is that the Q output of the Main flip-flop drives MUX-B instead of the M line. Thus, in the memory phase of MUX-latch (*Memory*="high") the output data of the Main flip-flop re-feed its input M and latched by the MUX-latch. After a timing error detection at a TDS flip-flop anywhere in the circuit, the correct data of the MUX-latch in a Freezing TDS flip-flop are re-captured at the output Q of the Main flip-flop (data freezing) by the triggering edge of *CLK* in the next correction cycle.

In order to evaluate the proposed timing error detection and correction technique, it has been applied in the design of a 32-bit pipelined MIPS R2000 microprocessor, with scan testing support, in the 90nm CMOS technology of UMC using the standard cells of Faraday Technologies. In parallel, the same microprocessor was designed, in the same technology, using the corresponding

flip-flop oriented Razor technique, with scan support. Comparisons between the two MIPS core designs proved that Time Dilation outperforms over Razor with respect to power consumption and silicon area cost. The Time Dilation based design presents a 12.6% reduction in the power consumption and 1.6% reduction in the silicon area with respect to the Razor based design.

4 Error Detection and Correction Technique

The second proposed technique, the Error Detection and Correction (EDC) technique [15], is based on the bit-flipping flip-flop concept. This is synopsized as follows: in case of error detection at the output of a flip-flop the corresponding logic value is asynchronously complemented for error correction.

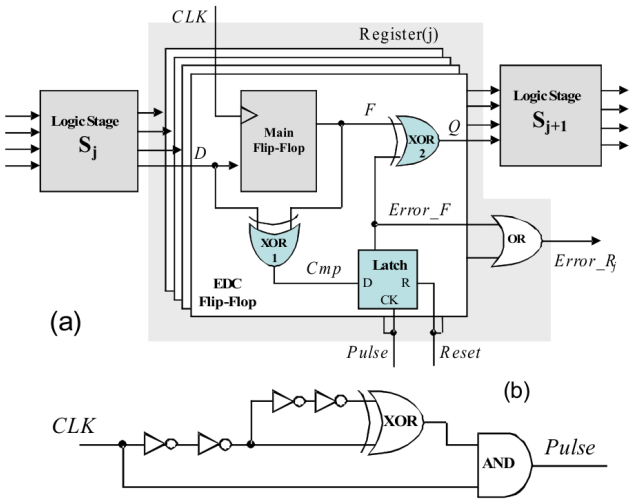


Fig. 5: (a) The EDC flip-flop and (b) The pulse generator.

Fig. 5(a) illustrates the new Error Detection/Correction flip-flop (EDC flip-flop) that is suitable to confront with timing errors. Apart from the original flip-flop (Main flip-flop), it consists of two XOR gates and a latch. The first XOR gate compares the D input and the F output of the Main flip-flop and provides the result to the latch. The latch feeds the second XOR gate at the output of the Main flip-flop. Depending on the comparison result within a specified time interval, either the F signal of the Main flip-flop or its complement is propagated to the output Q of the EDC flip-flop. The Q signal feeds the subsequent logic. Briefly, the proposed timing error detection and correction technique operates as follows. Suppose that a timing error is detected at one or more inputs of the combinational logic stage S_{j+1}, due to a delayed response of the previous stage S_j. Thus, the response of S_{j+1} will be erroneous and must be corrected.

To achieve error correction, the output of each flip-flop, at the register between the two stages, where a timing error has been detected is complemented so that valid values feed the S_{j+1} logic stage. Moreover, in case that this stage is not fast enough (not a shallow stage), the evaluation time of the circuit is extended by one clock cycle to guarantee its correct computation.

Initially, the output Error_F of the latch is reset to zero so that by default the *F* signal of the Main flip-flop propagates to the output *Q* of the XOR gate and feeds the subsequent logic stage. In the error free case the comparison result is a low value at the *Cmp* output of the first XOR gate after the triggering edge of the clock signal *CLK*. This value is captured by the latch. Thus, the *Q* output signal is identical to the *F* signal of the Main flip-flop, which carries the correct value. This signal feeds the subsequent logic stage S_{j+1}.

However, in the presence of a timing fault in logic stage S_j, a delayed signal arrives at the *D* input of the Main flip-flop after the triggering edge of the clock signal *CLK*. In that case, a timing error is present at the *F* output of the Main flip-flop and erroneous data are provided to the subsequent logic stage S_{j+1} through the *Q* output. In addition, the *F* signal value differs from the *D* signal value. The first XOR gate detects this difference and raises its output *Cmp* to high. The latch captures and holds this response. Thus, the second XOR gate provides at its output *Q* the complement of the *F* signal. Now the *Q* output of the EDC flip-flop carries the correct value, which feeds the subsequent logic stage S_{j+1} for its computation. Consequently, the error is locally corrected.

A clock pulse (*Pulse* signal) is used to capture the comparison result of the first XOR gate in the latch (memory state when the *Pulse* is low). This clock pulse can be generated locally from the *CLK* signal using a single Pulse Generator per register like the one illustrated in Fig. 5 (b). Thus, the routing overhead of an extra clock signal is relaxed. The AND gate in Fig. 5 (b) ensures that a single pulse will be generated only during the first phase of every clock cycle. The pulse width is at least equal to the time required by the latch to capture the comparison result. The time interval between the triggering edge of *CLK* and the falling edge of *Pulse* (minus the latch set up time and the XOR propagation delay time) determines the maximum detectable signal delay. Every signal transition at the *D* input of an EDC flip-flop within this time interval is considered as a delayed response. So the circuit design must guarantee that in the fault free case there are no signal transitions at the inputs of EDC flip-flops within this time interval, in order to avoid false alarms.

However, in the general case and in order to ensure the correct operation, extra time is required by the S_{j+1} logic stage to perform its computation after the correction of its input values. For that reason the error indication signal Error_F is used to block the clock signal from feeding the flip-flops during the subsequent clock cycle of the cycle where the error has been detected. Thus, a single clock cycle is dedicated for state recovery.

A core level clock gating technique can be exploited. Note that core level clock gating techniques are in common use for low power operation. To achieve this, the Error_F signals of all EDC flip-flops in a register (j) generate the register's

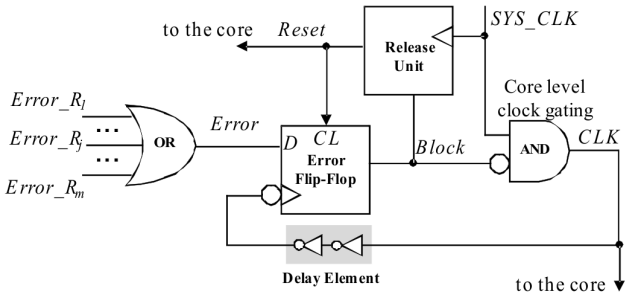


Fig. 6: Clock gating signal generation.

error indication signal Error_R_j through a local OR gate (see Fig. 5(a)). Next, all registers' Error_R_j signals are collected by a second OR gate which generates the core level error indication signal Error, as it is shown in Fig. 6. The Error signal is captured by a single flip-flop, the Error flip-flop. Its output signal Block is used for core level clock gating and to activate the Release unit. The latter releases the clock signal, after the expiration of the next system clock cycle, by the activation of the *Reset* signal which clears the Error flip-flop.

Moreover, the *Reset* signal clears the latches in the EDC flip-flops. Actually, the Release unit is a counter that counts one system clock cycle after its activation. The Error flip-flop is clocked by a delayed copy of the clock signal *CLK*. This delay is equal to the time required for the generation of the Error_F signal and its propagation through the pair of OR gates to the Error flip-flop. Considering small processing cores, the propagation of the Error_F signal will be fast enough to properly block the clock signal.

Comparisons on the MIPS pipelined microprocessor design proved that the EDC technique outperforms over Razor and Time Dilation with respect to power consumption and silicon area cost. The EDC supported design presents 20.8% and 9.2% reduction in the estimated power consumption with respect to the Razor and the Time Dilation supported designs respectively. Considering the silicon area, the EDC supported design presents 11.5% and 10.3% less silicon area with respect to Razor and Time Dilation supported designs respectively.

5 Timing Error Tolerance Technique

The Timing Error Tolerance (TET) technique [16], the third proposed error detection and correction technique, exploits the fact that after the triggering edge of the clock signal in a flip-flop, the data at its output must retain their value until the next triggering edge of the clock. Thus, any signal transition detected at the input of the flip-flop, during this time interval, is related to a timing error that can be corrected by bit-flipping the data stored in the flip-flop. Moreover, according to the adopted scheme, only the flip-flops at the end of

critical paths are replaced by the proposed flip-flop. The timing error tolerant oriented flip-flop structure is presented in Fig. 7(a). It consists of a Transition Detection (TD) unit for error detection and a flip-flop with preset and clear options, which is exploited for error correction.

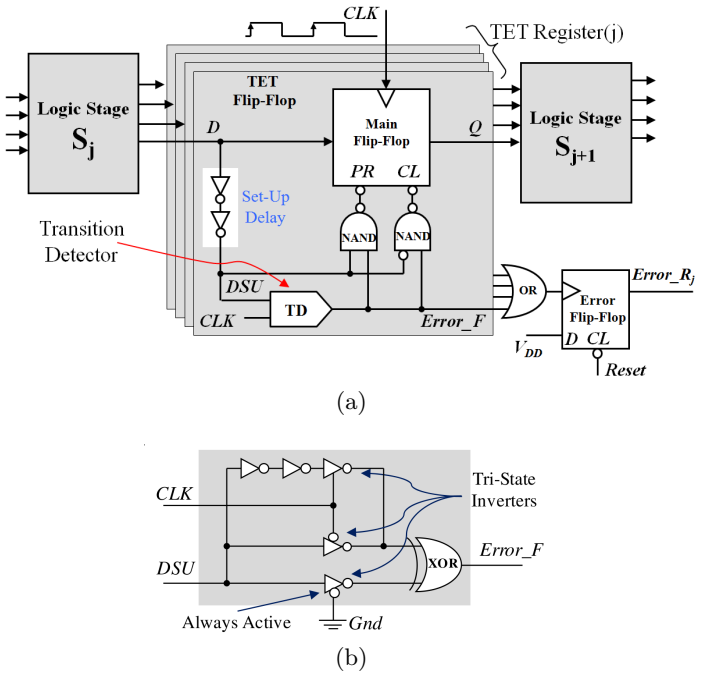


Fig. 7: (a) The proposed Timing Error Tolerant flip-flop and (b) the Transition Detector scheme.

The TD unit monitors the input D of the flip-flop within a time period (monitoring window) after the triggering edge of the clock CLK. During this time interval, no signal transitions are expected at the input of the flip-flop. In case of a signal transition within the monitoring window, the TD unit indicates an error detection by raising its output Error_F to high. A signal transition within the flip-flop's setup time is also considered as a timing violation. In order to be detected as timing error, it must arrive after the triggering edge of the clock. Thus, the TD unit is driven by a delayed version of the flip-flop input signal. This delay is equal to the setup time of the flip-flop. With the signal Error_F at logic "high", the correction operation follows. Two NAND gates are used, which are driven by Error_F and the delayed input signal DSU. If the final input data are at logic "high" then the Error_F signal activates the first NAND gate which presets the flip-flop output to high. If the final input data are at logic "low" then the Error_F signal activates the second NAND gate which clears the

flip-flop. In both cases the output Q of the Main flip-flop turns to the value of the correct but delayed data.

The TD unit design is illustrated in Fig. 7(b). It consists of a two input XOR gate, three tri-state inverters and delay elements. One input of the XOR gate is always driven by the $\overline{\text{DSU}}$ signal, because the bottom tri-state inverter is always active. The other input of the XOR gate is driven either by the $\overline{\text{DSU}}$ signal or by a delayed version of that signal, depending on the value of *CLK*. When the *CLK* signal is at logic "low" the two bottom signal paths are activated. Thus, any transition at the input of the TD unit arrives concurrently at both of the XOR gate inputs and no pulse is generated at its output. When the *CLK* signal is at logic "high" the top and the bottom signal paths are active. In this case, due to the delay elements inserted in the top path, there is a delay between the arrivals of the signals at the two inputs of the XOR gate. Thus, a pulse is generated at the XOR's output. The pulse width is equal to the delay inserted in the top path and adequate to activate the preset or clear operation at the Main flip-flop.

From the above analysis, it is clear that the monitoring window of the TD unit is determined by the "high" pulse width of the *CLK* signal. Any transition at input *D*, in this time interval, is detected as timing error, so that in the normal operation of the circuit, no transitions are permitted at this input. To avoid false alarms, either the duty cycle of *CLK* signal is adjusted or the fast paths are delayed, according to a minimum path delay constraint, or both techniques are applied.

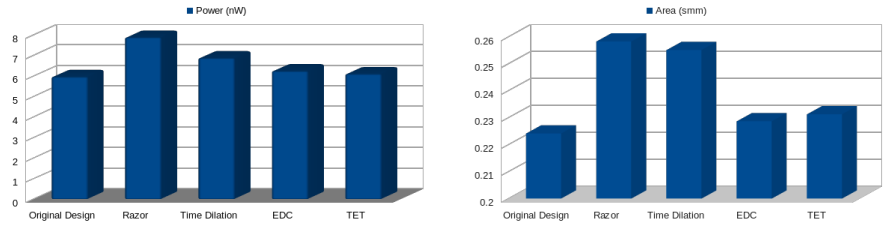


Fig. 8: Comparison graphs for power consumption and silicon area

In the comparisons that follow the standard cells of the 90nm Faraday library are used for the design of all four techniques at the same operating frequency. The TET based design presents 25.59%, 11.21% and 2.24% reduction in power consumption with respect to the Razor [8], the Time Dilation [14] and the EDC [15] based designs respectively. Considering the silicon area, the TET design presents 10.46% and 9.33% less silicon area with respect to Razor and Time Dilation designs respectively and 1.1% increase with respect to the EDC technique. Comparison graphs are presented in Fig. 8.

This scheme was also applied on a 32-bit pipelined MIPS microprocessor, which was fabricated in the 65nm Low Leakage technology of UMC, through the

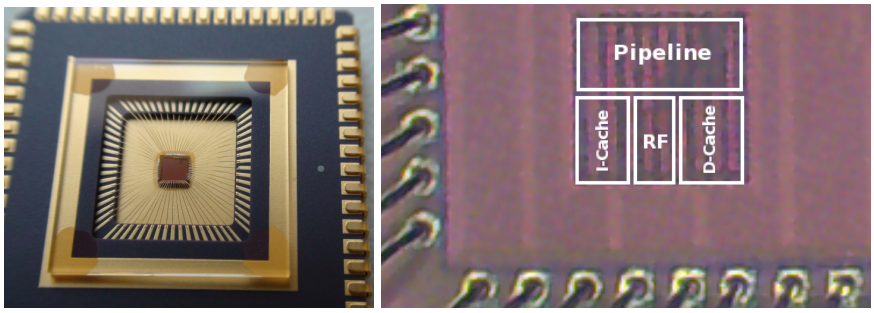


Fig. 9: Fabricated chip and die photo

EUROPRACTICE IC Service, offered by IMEC and Fraunhofer. Fig. 9 shows the fabricated chip and the die photo with the MIPS core.

For the evaluation of the proposed technique on the fabricated chip, timing errors are created by operating the microprocessor at a lower voltage level than the nominal. The design has two outputs: the global error indication signal and the signature output of a Multiple-Input Signature Register (MISR), which is used to compress the response of the design. The error indication signal shows whether timing errors are detected, while the value of the signature shows whether these errors are corrected or not, by comparing this value (i.e. the compacted response of the design) with the expected one. Experimental results show that the proposed technique detects and corrects the generated timing errors efficiently with low power consumption and low silicon area overhead.

6 Conclusions

Timing errors in the memory elements of a design are of increasing importance in nanometer technology microprocessor cores. This thesis presents three low cost timing error detection and correction technique. The first technique provides concurrent error detection and correction in the field of application and also supports off-line manufacturing scan testing. By utilizing a new scan flip-flop, this technique is capable to detect and correct multiple errors at the minimum penalty of one clock cycle delay. The second technique is based on a new bit flipping flip-flop. Whenever a timing error is detected, it is corrected by complementing the output of the corresponding flip-flop. The last technique exploits a transition detector for timing error detection along with asynchronous local error correction schemes to provide timing error tolerance. The proposed approaches are characterized by low cost and reduced design complexity, that also result in reduced power consumption area with respect to earlier design schemes in the open literature.

References

1. J. Semião, J.F. Freijedo, J.J. Rodriguez-Andina, F. Vargas, M.B. Santos, I.C. Teixeira and P.J. Teixeira, "Time Management for Low-Power Design of Digital Systems", ASP Journal of Low Power Electronics (JOLPE), vol. 4, no. 3, pp. 410-419, 2008.
2. M. Cuviello, S. Dey, X. Bai, Y. Zhao., "Fault Modeling and Simulation for Crosstalk in System-on-Chip Interconnects," Int. Conf. on Computer Aided Design, pp. 297-303, 1999.
3. H. Chen, L. Wang., "Design for Signal Integrity: The New Paradigm for Deep-Submicron VLSI Design," Proc. Int. Symp. on VLSI Technology, pp. 329-333, 1997.
4. S.V. Kumar, C.H. Kim, S. Sapatnekar, "Adaptive Techniques for Overcoming Performance Degradation due to Aging in Digital Circuits," Proc. IEEE ASP-DAC, pp. 284-289, 2009.
5. S. Khan, S. Hamdioui, H. Kukner, P. Raghavan and F. Catthoor, "BTI impact on logical gates in nano-scale CMOS technology," Proc. IEEE Int. DDECS, pp. 348-353, 2012.
6. S. Matakias, Y. Tsiatouhas, A. Arapoyanni, and Th. Haniotakis, "A Circuit for Concurrent Detection of Soft and Timing Errors in Digital CMOS ICs," Journal of Electronic Testing: Theory and Applications, vol. 20, no. 5, pp. 523-531, 2004.
7. K. Kang, S.P. Park, K. Kim and K. Roy, "On-Chip Variability Sensor Using Phase-Locked Loop for detecting and Correcting Parametric Timing Failures," IEEE Transactions on VLSI Systems, vol. 18, no. 2, pp. 270-280, 2010.
8. T. Austin, D. Blaauw, T. Mudge and K. Flautner, "Making Typical Silicon Matter with Razor," IEEE Computer, vol. 37, no. 3, pp. 57-65, 2004.
9. S. Das, C. Tokunaga, S. Pant, W-H. Ma, S. Kalaiselvan, K. Lai, D.M. Bull and D.T. Blaauw, "RazorII: In Situ Error Detection and Correction for PVT and SER Tolerance," IEEE Journal of Solid-State Circuits, vol. 44, no. 1, pp. 32-48, 2009.
10. D. Bull, S. Das, K. Shivashankar, G.S. Dasika, K. Flautner and D.T. Blaauw, "A Power-Efficient 32 bit ARM Processor Using Timing-Error Detection and Correction for Transient-Error Tolerance and Adaptation to PVT Variation," IEEE Journal of Solid-State Circuits, vol. 46, no. 1, pp. 18-31, 2011.
11. M. Nicolaidis, "GRAAL: a new fault tolerant design paradigm for mitigating the flaws of deep nanometric technologies," IEEE International Test Conference, 2007.
12. M. Choudhury, V. Chandra, R. Aitken, and K. Mohanram, "Time-borrowing circuit designs and hardware prototyping for timing error resilience," IEEE Transactions on Computers, vol. 63, no. 2, pp. 497-509, 2014.
13. M. Nicolaidis, "Double-Sampling Design Paradigm–A Compendium of Architectures," IEEE Transactions on Device and Materials Reliability, vol. 15, no. 1, pp. 10-23, 2015.
14. S. Valadimas, A. Floros, Y. Tsiatouhas, A. Arapoyanni, X. Kavousianos, "The Time Dilation Technique for Timing Error Tolerance," IEEE Transactions on Computers, vol. 63, no. 5, pp. 1277-1286, 2014.
15. S. Valadimas, Y. Tsiatouhas, A. Arapoyanni, "Timing Error Tolerance in Small Core Designs for SoC Applications," IEEE Transactions on Computers, vol. 65, no. 2, pp. 654-663, 2016.
16. S. Valadimas, Y. Tsiatouhas, A. Arapoyanni, P. Xarchakos, "Effective Timing Error Tolerance in flip-flop Based Core Designs," Springer Journal of Electronic Testing: Theory and Applications, vol. 29, no. 6, pp. 795-804, 2013.

The DEMOS family of e-voting systems: End-to-end verifiable elections in the standard model

Thomas Zacharias**

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications

thzacharias@di.uoa.gr

Abstract. This PhD thesis introduces the DEMOS-A and DEMOS-2 e-voting systems that achieve *end-to-end verifiability in the standard model* for the first time. End-to-end verifiability in the standard model denotes that verification is executed without without putting trust in any administration authority and without assuming any trusted randomness setting. Prior to this thesis, all top-tier e-voting systems (e.g. SureVote, JCJ, Prêt à Voter, Helios, Scantegrity, etc.) assumed honesty of the voting clients, the random oracle model, or the existence a randomness beacon to achieve end-to-end verifiability.
In the core of DEMOS-A and DEMOS-2, is a novel mechanism that extracts the randomness required for verification from the *entropy generated by the voters*, when they engage in the voting phase. This entropy is *internal* with respect to the election environment, therefore the need for trusting an outer source of randomness is removed.
The security analysis is performed under a novel cryptographic framework that constitutes an additional contribution of this thesis. The end-to-end verifiability theorems for DEMOS-A and DEMOS-2 reveal that the security level is in high correlation with the auditing behaviour of the electorate. Motivated by this finding, this thesis extends the framework by modelling e-voting systems as *ceremonies*, inspired by the work of Ellison in 2007. As a case study of an e-voting ceremony, this thesis investigates the security of the well-known Helios e-voting system.

1 Introduction

Political activity in a modern democratic state comprises compositions of individual democratic procedures. At a high level, a democratic procedure consists of three well-defined concepts.

1. An *electorate* formed by the people legitimate to vote,
2. A *voting system*, which serves as means to record and evaluate the electorate's will, and

** *Dissertation Advisor*: Aggelos Kiayias, Associate Professor

3. A *verdict*, which stems from the consensus according to the evaluated electorate's will.

A reliable voting system must incorporate mechanisms for optimising accessibility of the electorate and guaranteeing integrity of the election result while protecting the voters' secrecy. If it does so, then it paves the way for building a healthy democratic society. On the other hand, due to their crucial role in democracy, voting systems have often been top priority targets for attackers that wish to tamper the election result and/or coerce voters to vote against their intention. Voting systems that allow people to sell their votes, or lack verification procedures that convince an auditor of the validity of the election result with minimum doubt, undermine the foundations of any democratic state they are deployed.

e-Voting in democratic procedures

In an e-voting system, election preparation, vote collection and/or tally is executed by electronic devices, partially or fully managed by human authorities. The motivation for introducing e-voting was originally three-fold; (i) facilitating the participation of social groups with considerable physical barriers, (ii) reduction of election cost, and (iii) acceleration of the election preparation, vote casting and tally phase. E-voting emerged in the 60s via punch-card systems, followed by systems based on either optical scan voting, ballot encryption, or vote-code typing. By today, e-voting systems have been used in several countries either in pilot executions (Australia, England, Ireland, Italy, Norway) or binding elections at a municipality or national level (Belgium, Brazil, Canada, Estonia, India, the Netherlands, Switzerland, USA). Nonetheless, they have been subject to often trenchant criticism, mainly due to the disquiet about potential security threats caused by the amount of power now transferred to the machines.

Based on their infrastructure, e-voting systems are classified into (i) *On-site e-voting systems*, where the election is executed in polling stations, and supervision by human authorities is similar to traditional elections, and (ii) *Remote e-voting (i-voting) systems*, where the voters submit their votes using devices (PCs, notebooks, tablets, smartphones) that have internet access.

End-to-end verifiability and e-voting

Besides advancing participation and reduction of election cost and time, several state-of-the-art e-voting systems [10,14,37,13,1,2,44,42] support an attractive and highly non-trivial security feature that traditional voting unavoidably misses by its nature. Namely, the voter can verify that her vote was properly cast, recorded and tallied into the election result without relying to the honesty of any of the election administrators. This strong property is named *end-to-end (E2E) verifiability* and is usually interpreted as the ability of the voter to verify that her vote was (i) cast-as-intended, (ii) recorded-as cast, and (iii) tallied-as-recorded.

Before this PhD thesis, E2E verifiability could not be justified with minimum assumptions. Under a strong cryptographic definition, E2E verifiability could provenly hold only assuming the existence of a *trusted randomness setting*

that could be either a function modelled as a random oracle [1,2,42], or some randomness beacon [10,14,37,13,44].

Objectives and contributions of this thesis

The main objective this thesis investigates, is the feasibility of E2E verifiability *in the standard model*, which denotes that verification is executed with assuming the existence of a trusted randomness setting. As already mentioned, until the writing of this thesis, E2E verifiability in an all-malicious setting could provenly hold only under certain setup assumption for randomness.

In order to illustrate why previous techniques did not work, we elaborate on the previous statement. By its design, Helios [1]-and other client-side encryption E2E verifiable systems as [20,31,2,42]- requires the voter to utilise a voter supporting device to prepare a ciphertext and after an indeterminate number of trials, the voter will cast the produced ciphertext. The submitted ciphertexts are to be homomorphically tallied and thus they should be accompanied by a proof of proper computation. While such proofs are easy to construct based on e.g., [19], they can be argued either (i) interactively or (ii) using a *non-interactive zero-knowledge (NIZK) proof* [6]. Interaction is insufficient in E2E verifiability setting since a corrupt election authority together with a corrupt voter may cook up a malformed proof that is indistinguishable from a proper one. As a result, the non-interactive approach is mandatory. However, NIZK proofs can be sound only under setup assumptions as a *random oracle* or a *common reference string (CRS)* [26]. If the CRS is setup by the election authority, then, in case it is malicious, it will know and exploit the trapdoor; on the other hand, the voters are not interacting with each other and hence cannot setup the CRS by employing a standard multi-party computation protocol [25,12].

On the other hand, in the case of Remotegrity/Scantegrity [13,44] -and other client-side cryptography E2E verifiable systems as [10,14,37]- the random coins need to be obtained from the randomness beacon in order to prove the result correct. It is easy to verify that the system is insecure in terms of E2E verifiability in case the randomness beacon is biased. As before, the only parties active are the election authority and the voters who cannot implement a randomness beacon that is required in the construction.

As a consequence of the aforementioned technical restrictions , the following question remained open until recently:

Q1. *Can the integrity of the election result be proven in the* standard model *i.e. without believing in trusted hardware, random oracles or randomness beacons?*

This PhD thesis answers this question affirmatively by introducing the *DEMOS-A* and *DEMOS-2* e-voting systems that achieve E2E verifiability in the standard model, as long as a *publicly accessible bulletin board* where the election results are posted remains consistent. Furthermore, DEMOS-A and DEMOS-2 preserve privacy given the hardness of a standard cryptographic problem (Decisional

Diffie-Hellman). The core idea for this accomplishment is a novel mechanism for extracting randomness from the entropy injected to the system by the voters' entanglement. This entropy is *internal* with respect to the election environment, a fact that removes the requirement for an external randomness source.

The two systems follow different approaches with respect to their design. In particular, DEMOS-A follows the *code-voting approach*, where the voters obtain ballots that contain independent and random encodings of the election options (typically vote-codes in one-to-one correspondence with the election options). At the voting phase, the voters cast the encodings that correspond to their intended selections in their ballots. Consequently, vote submission becomes a simple procedure which can be run by devices of minimum computational power. However, this flexibility comes with a price of high complexity at the election preparation phase from the election servers side, resulting in important scalability restrictions for DEMOS-A. To resolve this issue, this thesis introduces the DEMOS-2 e-voting system, in the spirit of the *client-side encryption*. Namely, in DEMOS-2, the overhead is distributed to the voting clients, which now must be computationally able to locally encrypt the voters' ballots, hence to perform cryptographic operations. As a result, DEMOS-A and DEMOS-2 have complementary benefits and weaknesses regarding their functionality and security, hence the choice of the most preferable system depends on the given election setting.

The second objective studied in this thesis is the effect of the human factor in the security of an E2E verifiable e-voting system. The security analysis of DEMOS-A provides evidence of a strong correlation between the active participation of honest voters in the auditing procedure and the (parameterised) level of E2E verifiability that can be guaranteed. A natural question follows from this observation:

> **Q2.** *At what extent can human behaviour, even within protocol specification, affect the security of an e-voting system?*

This PhD thesis follows a formal cryptographic direction to deal with this matter. Motivated by the *ceremony framework* introduced by Ellison [22] for the analysis of network protocols, it proposes an extension of standard e-voting security modelling, where human nodes are separated from computer nodes and are formalised as finite state machines (transducers) with limited power, hence incapable of performing cryptographic operations. As a case study of the extended ceremony framework, Helios stands out in terms of the range of possible human behaviour due to (i) the dependence of E2E verifiability on (i.a) the statistics related to the Benaloh audit rate performed by the voters and (i.b) the portion of voters that look up their votes in the bulletin board after election using their ballot trackers and (ii) the dependence of privacy on the trustees auditing the correct uploading of the public key, combined with lack of public key infrastructure (PKI) for support authentication of posted data.

In summary, the contributions of this PhD thesis comprise:

1. The introduction of a robust cryptographic framework for the security analysis of e-voting systems. The said framework captures definitions of E2E verifiability, voter privacy and *passive coercion resistance (PCR)* (often referred as receipt-freeness). The latter property denotes the inability of an e-voting system to allow the voters to prove how they voted or sell their votes, even against an adversary that observes network traffic and requests from the voter the transcript containing their personal view of interaction with the election system. The suggested framework is extended to the ceremony model, suitable for the formal study of human behaviour in an e-voting execution.
2. The presentation of two remote e-voting systems, (i) the vote-coding based DEMOS-A and (ii) the client-side encryption based DEMOS-2 that enrich both major e-voting categories with a member that achieves *E2E verifiability in the standard model* for the first time. The two systems are proven secure under the aforementioned framework and their voter privacy/passive coercion resistance holds assuming the hardness of the extensively studied Decisional Diffie-Hellman problem. These two systems give birth to the *DEMOS family of e-voting systems* sharing the attribute of E2E verifiability in the standard model.
3. A thorough analysis of the Helios e-voting system under the ceremony framework. This analysis is threefold consisting of (i) a rigorous mathematical characterisation of classes of voter behaviours that are assailable or resistant to attacks on verifiability, (ii) an evaluation of the expected E2E verifiability guarantee of Helios based on the previous theoretical context given instantiations of real world Helios applications as well as simulation data, and (iii) a presentation of a standard *man-in-the-middle* attack against Helios's privacy, in cases where election guidelines do not encourage trustees (modelled as human nodes) to verify the correct posting of the election public key in the bulletin board.

Related work

Up to the present moment, numerous noticeable e-voting systems have been introduced [8,18,9,23,20,10,28,30,14,31,40,13,1,17,37,24,2,42,44], adding to cryptographic literature novel directions or ameliorating existing techniques. In the following table, we depict the classification of a list of e-voting systems, according to their infrastructure and vote submission method.

	Client-side encryption	**Code-voting**
On-site	[2]	[14,40,13,37]
Remote	[8,18,9,23,20,28,30,31,1,17,24,42]	[10,44]

End-to-end verifiability in the sense of cast-as-intended, recorded-as-cast, tallied-as-recorded was an outcome of the works in [11] and [36] that introduced the generation of receipts which could be used for simple voter verification while preserving privacy. Prior definitions referring to the weaker notions of *individual* and *universal* verifiability are found in [8,38,29,32,15]. Rigorous end-to-end verifiability definitions have been proposed in [33] and [41]. Definitions of privacy

and receipt-freeness have been introduced in [18,3,16,21,27,35,34,4,5] under the cryptographic, symbolic and universal composability [7] model.

2 Results

In this section, we provide an overview of the components of this PhD thesis that comprise the complete presentation and analysis of DEMOS-A, i.e. the syntax, the end-to-end verifiability and voter privacy/PCR definitions, system's description, and the statement of the security theorems for DEMOS-A. Due to space limitations, we refer the reader interested in the results related to DEMOS-2 and the ceremony framework to [43, Chapter 5] and [43, Chapter 6], respectively.

2.1 Preliminaries

We use λ as the security parameter and consider three additional parameters; the number of voters n, options m, and trustees k, all of which are thought as polynomial in λ.

For an e-voting system $\boldsymbol{\mathcal{VS}}$, we fix the set of options $\mathcal{O} = \{\mathsf{opt}_1, ..., \mathsf{opt}_m\}$. We denote by $\mathcal{U} \subseteq 2^{\mathcal{O}}$ the collection of subsets of options that the voters are allowed to choose to vote for (which may include a "blank" option too). The option selection $\mathcal{U}_\ell$ of voter V_ℓ is an element in $\mathcal{U}$.

Let $\mathcal{U}^*$ be the set of vectors of option selections of arbitrary length. Let f be the *election evaluation function* from $\mathcal{U}^*$ to the set $\mathbb{Z}_+^m$ so that $f(\mathcal{U}_1, \ldots, \mathcal{U}_n)$ is equal to an m-vector whose i-th location is equal to the number of times opt_j was chosen in the option selections $\mathcal{U}_1, \ldots, \mathcal{U}_n$. The entities involved in an e-voting system $\boldsymbol{\mathcal{VS}}$ are the following:

- The *election authority* EA that prepares all the election information.
- The *voters* $\mathcal{V} = \{V_1, \ldots, V_n\}$, possibly equipped with *voting supporting devices (VSDs)*.
- The *vote collector* VC that realises the digital ballot box functionality.
- The set of *trustees* $\mathcal{T} = \{T_1, \ldots T_k\}$ responsible for computing the tally and announcing the election result.
- A publicly accessible and consistent *bulletin board* BB where the election result and all audit information is posted.

2.2 Security framework

Definition of end-to-end verifiability. In order to define E2E verifiability formally, we introduce a suitable notation; given that option selections are elements from a set of m choices, we encode them as m-bit strings, where the bit in the j-th position is 1 if and only if option opt_j is selected. Further, we aggregate the election results as the list with the number of votes each option has received. Thus, the **Result** algorithm outputs a vector in $\mathbb{Z}_+^m$, i.e., the range of the election evaluation function f.

Then, we use the metric d_1 derived by the ℓ_1-norm scaled to half, i.e., $\mathrm{d}_1(R, R') = \frac{1}{2} \cdot \sum_{i=1}^{n} |R_i - R'_i|$, where R_i, R'_i is the i-th coordinate of R, R' respectively, to measure the success probability of the adversary with respect to the amount of tally deviation δ and the number of voters that perform audit θ. In addition, we make use of a *vote extractor* algorithm $\mathcal{E}$ (not necessarily running in polynomial-time) that extracts the non-honestly cast votes.

We define the E2E Verifiability game, $G_{\mathrm{E2E}}^{\mathcal{A},\mathcal{E},\delta,\theta}$, between the adversary $\mathcal{A}$ and a challenger Ch using a vote extractor $\mathcal{E}$. The game takes as input the security parameter, λ, the number of options, m, the number of voters, n, and the number of trustees k. The game is also is parameterised by δ, which is the deviation amount (according to the metric $\mathrm{d}_1(\cdot,\cdot)$) that the adversary wants to achieve and θ, the minimum number of voters that $\mathcal{A}$ must allow to vote honestly and terminate successfully.

The adversary $\mathcal{A}$ starts by selecting the voter, option, and trustee identities for given parameters n, m, k. It also determines the allowed ways to vote as described by the set $\mathcal{U}$. Then, $\mathcal{A}$ fully controls the election by corrupting the EA, the VC, all the trustees $\mathcal{T} = \{T_1, \ldots T_k\}$ and all the VSDs. In addition, it manages the **Cast** protocol executions where it assumes the role of the VC. For each voter, $\mathcal{A}$ may choose to corrupt her or to allow the challenger to play on her behalf. In the second case, $\mathcal{A}$ provides the honest voter with the option selection that will use in the **Cast** protocol. Finally, $\mathcal{A}$ completes the election execution which results to the complete election transcript published in the BB.

The adversary will win the game provided that all θ honest voters that completed the **Cast** protocol successfully will also audit the result successfully, while either (a) the deviation of the tally is at least δ or (b) the extractor fails to produce the option selection of the dishonest voters.

Definition 1. *Let $\epsilon \in [0,1]$ and $m, n, k, \delta, \theta \in \mathbb{N}$ with $\delta > 0$ and $0 < \theta \leqslant n$. Let $\boldsymbol{\mathcal{VS}}$ be an e-voting system with m options, n voters and k trustees w.r.t. the evaluation election unction f. We say that $\boldsymbol{\mathcal{VS}}$ achieves* E2E verifiability *with error ϵ, for a number of at least θ honest successful voters and tally deviation δ if there exists a (not necessarily polynomial-time) vote-extractor $\mathcal{E}$ such that for any adversary $\mathcal{A}$*

$$\Pr\left[G_{\mathrm{E2E}}^{\mathcal{A},\mathcal{E},\delta,\theta}(1^{\lambda}, m, n, k) = 1\right] \leqslant \epsilon \,.$$

Modelling voter privacy/PCR. The definition of privacy concerns the actions that may be taken by the adversary in order to obtain information about the option selections of the honest voters. We specify the goal of the adversary in a very general way; for an attack to succeed, we ask that there is an election result, for which the adversary is capable of distinguishing how the honest voters have voted, while it has access to (i) the individual audit information that the voters obtained after ballot-casting as well as (ii) a set of protocol views that are consistent with all the honest voters' views in the **Cast** protocol instances they participated and the adversary has monitored.

Observe that any system secure against the aforementioned attack scenario would possess also PCR, i.e., voters cannot prove how they voted by showing

the individual audit information they obtain from the **Cast** protocol or even presenting their view in the **Cast** protocol. Given that in the privacy definition we allow the adversary to observe the view of the voter in the **Cast** protocol, we must allow the voter to be able to "lie" about her view, otherwise an attack could be trivially mounted.

In order to capture the PCR property as described above, we utilise an efficient *view simulator* $\mathcal{S}$ that provides a simulated view of the voter in the **Cast** protocol. Intuitively, $\mathcal{S}$ captures the way the voter can lie about her option selection in the **Cast** protocol in case she is coerced to present her view after she completes the ballot-casting procedure. It is imperative that the simulated view is indistinguishable from the actual view the voter obtains.

2.3 The DEMOS-A e-voting system

Description overview

In DEMOS-A, each voter may select 1 out of m options and cast her vote using vote-codes listed in her ballot. Each ballot has two functionally equivalent parts (with a complete list of the m options in each part), instructing the voter to pick one of the two parts at random. The cryptographic payload of DEMOS-A consists of lists of the following primitives:

Additively homomorphic commitments: where a value M is posted in a committed form, denoted by $\mathsf{Com}(M)$, such that (i) when the opening of $\mathsf{Com}(M)$ denoted by $\tilde{M}$ is posted, then no other value than M can be extracted from $\tilde{M}$ (*binding* property), (ii) $\mathsf{Com}(M)$ reveals no information about M to any computationally bounded adversary (*hiding* property) and (iii) for any two values M_1, M_2, it holds that $\mathsf{Com}(M_1) \cdot \mathsf{Com}(M_2) = \mathsf{Com}(M_1 + M_2)$ (*additive homomorphic* property). DEMOS-A utilises ElGamal as a commitment scheme that is (i) perfectly binding and (ii) hiding, assuming that the Decisional Diffie-Hellman problem is hard for the underlying group.

Zero knowledge proofs: these are proofs such that (i) if the honest verifier accepts a proof, then she is assured that the statement is true (*soundness* property) and (ii) the proof reveals no other information than the truth of the statement (*zero-knowledge property*). Specifically, DEMOS-A makes uses of *three-move* zero-knowledge proofs, where the interaction is accomplished by a first move from the prover, a second move that is the verifier's challenge (the source of which is the random choice from the voters regarding the part of the ballot they chose), and a third move where the prover responds to the challenge which completes the proof.

Formally, DEMOS-A consists of five protocols/algorithms: **Setup**, **Cast**, **Tally**, **Result**, and **Verify**. We will briefly present them here, omitting many cryptographic details, for simplicity.

In the **Setup** protocol, the EA generates the initialization data for each election entity. More specifically, each randomly generated vote-code points to a cryptographic payload, consisting of additively homomorphic commitments

of the *option-encoding*, where the i-th option, option_i, is encoded into $(n+1)^{i-1}$. These commitments are associated with necessary zero-knowledge proofs (prover's first move) that allow the EA to show that each commitment is valid (i.e., it commits to an option encoding) later on, without revealing its actual content. The EA then assigns each ballot with two functionally equivalent parts. When the ballot preparation is finalized, the EA distributes the ballots to the voters.

In the **Cast** protocol, the voter randomly chooses one of the two parts of the ballot to vote by submitting the vote-code corresponding to her intended option. The unused part will be kept for auditing after the election ends.

In the **Tally** protocol, the EA fetches the entire election transcript from the BB and posts additional data on the BB. In this step, the tally result is produced using the homomorphic property by "adding" all the option-encoding commitments associated with the vote-codes cast by the voters and are marked as "voted". Note that, the result is in committed form and requires the corresponding opening to be decoded. Furthermore, the commitments that correspond to the unused parts of voter ballots are also revealed for auditing. Finally, the EA derives the challenge (second move) of the zero-knowledge proofs based on the voters' choices of used ballot parts and completes the zero-knowledge proofs that correspond to the option-encoding commitments marked as "voted", by posting all the respective third moves of the prover.

The **Result** algorithm takes as input the entire BB information, and can be executed by anyone. For instance, if $n = 9$, the *option-encodings* of options $1, 2, 3$ are $1, 10, 100$, respectively. Suppose we got 3 votes for option_1, 5 votes for option_3, the sum of the *option-encodings* is $3 * 1 + 5 * 100 = 503$. By the opening of the homomorphic tally, the **Result** algorithm extracts 503 and decodes it as $(3, 0, 5)$, which represents the corresponding votes for each election option.

The **Verify** algorithm can be executed by voters and any third-party auditors. A third-party auditor is able to verify the validity of all the commitments by checking the completed zero-knowledge proofs. Besides, each voter is allowed to perform "print check" by comparing her private ballot with the information on the BB. As the number of auditing voters increases, the probability of election fraud going undetected diminishes exponentially. For example, even if only 10 voters audit, with each one having $\frac{1}{2}$ probability to detect ballot fraud, the probability of ballot fraud going undetected is only $\frac{1}{2}^{10} = 0.00097$.

Security of DEMOS-A

The end-to-end verifiability and voter privacy/PCR that DEMOS-A achieves are formally stated in the following theorems.

Theorem 1. *Assume an election run of DEMOS-A with n voters, m candidates and k trustees. Let q be the size of the group for the of the underlying commitment scheme described. Then, DEMOS-A achieves E2E verifiability information theoretically for at least θ honest successful voters and tally deviation δ with error*

$$2^{-\delta} + 2^{-\theta + \lceil n/\lfloor \log q \rfloor \rceil (\log\log m + 1)} .$$

Theorem 2. *Assume an election run of DEMOS-A with n voters, m candidates and k trustees. Assume there exists a constant $c, 0 < c < 1$ such that for any 2^{λ^c}-time adversary $\mathcal{A}$, the advantage of breaking the hiding property of the commitment scheme is $\mathsf{Adv}_{\mathsf{hide}}(\mathcal{A}) = \mathsf{negl}(\lambda)$. Let $t = \lambda^{c'}$ for any constant $c' < c$. Then, for any constant m and n, k polynomial in the security parameter λ, DEMOS-A achieves voter privacy/PCR against any adversary that corrupts at most t corrupted voters.*

3 Conclusions

The completion of this PhD thesis concludes an extended formal cryptographic argumentation on the boundaries of optimal E2E verifiability and the relation of e-voting security with human auditing behaviour. The introduction of the DEMOS family initialised to the pair of DEMOS-A and DEMOS-2 e-voting systems answers affirmatively to question **Q1** of the introduction, promising election execution where the integrity of the result is proven under the standard model, i.e. without trusting a source of randomness. In addition, the honesty of no election administrator or voting supporting device is required.

As far as studying human behaviour is concerned, this thesis has set the necessary cryptographic background and its mathematically argued results on this matter raise intriguing issues. The security analysis of the widely used Helios e-voting system pointed out its weaknesses, in cases where human verification. Our analysis leads to a debate that, beyond its technical basis, can be viewed from a rather political and philosophical lens; if human behaviour, even within protocol specification, *can affect* the security of an e-voting system, then specifying explicitly the extent of the risks -thus answering question **Q2** of the introduction- becomes a top priority. Can these risks be mitigated by significantly better systems, or do they set a security guarantee upper bound, as price for moving responsibility directly to the voters? In order to ask for end-to-end verifiable security, is people's proper training a prerequisite? Stated abstractly,

Is political maturity an inevitable trade-off for provenly secure direct democratic procedures?

The robust ceremony model of this PhD thesis could be the means for translating these questions into strict mathematical language and thus provide a valuable asset for subsequent research.

References

1. Ben Adida. Helios: Web-based open-audit voting. In *USENIX*, 2008.
2. Josh Benaloh, Michael D. Byrne, Bryce Eakin, Philip T. Kortum, Neal McBurnett, Olivier Pereira, Philip B. Stark, Dan S. Wallach, Gail Fisher, Julian Montoya, Michelle Parker, and Michael Winn. STAR-vote: A secure, transparent, auditable, and reliable voting system. In *EVT/WOTE*, 2013.

3. Josh Cohen Benaloh and Dwight Tuinstra. Receipt-free secret-ballot elections (extended abstract). In *STOC*, 1994.
4. David Bernhard, Véronique Cortier, Olivier Pereira, Ben Smyth, and Bogdan Warinschi. Adapting helios for provable ballot privacy. In *ESORICS*, 2011.
5. David Bernhard, Olivier Pereira, and Bogdan Warinschi. How not to prove yourself: Pitfalls of the fiat-shamir heuristic and applications to helios. In *ASIACRYPT*, 2012.
6. Manuel Blum, Paul Feldman, and Silvio Micali. Non-interactive zero-knowledge and its applications (extended abstract). In Simon [39], pages 103–112.
7. Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, 14-17 October 2001, Las Vegas, Nevada, USA*, pages 136–145. IEEE Computer Society, 2001.
8. David Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Commun. ACM*, 24(2):84–88, 1981.
9. David Chaum. Elections with unconditionally-secret ballots and disruption equivalent to breaking RSA. In *EUROCRYPT*, 1988.
10. David Chaum. Surevote: Technical overview. In *Proceedings of the Workshop on Trustworthy Elections*, WOTE, 2001.
11. David Chaum. Secret-ballot receipts: True voter-verifiable elections. *IEEE Security & Privacy*, 2(1):38–47, 2004.
12. David Chaum, Claude Crépeau, and Ivan Damgård. Multiparty unconditionally secure protocols (extended abstract). In Simon [39], pages 11–19.
13. David Chaum, Aleksander Essex, Richard Carback, Jeremy Clark, Stefan Popoveniuc, Alan T. Sherman, and Poorvi L. Vora. Scantegrity: End-to-end voter-verifiable optical-scan voting. *IEEE Security & Privacy*, 6(3):40–46, 2008.
14. David Chaum, Peter Y. A. Ryan, and Steve A. Schneider. A practical voter-verifiable election scheme. In *ESORICS*, 2005.
15. Benoît Chevallier-Mames, Pierre-Alain Fouque, David Pointcheval, Julien Stern, and Jacques Traoré. On some incompatible properties of voting schemes. In *Towards Trustworthy Elections*, 2010.
16. Benoît Chevallier-Mames, Pierre-Alain Fouque, David Pointcheval, Julien Stern, and Jacques Traoré. On some incompatible properties of voting schemes. In *Towards Trustworthy Elections*, 2010.
17. Michael R. Clarkson, Stephen Chong, and Andrew C. Myers. Civitas: Toward a secure voting system. In *IEEE Symposium on Security and Privacy*, 2008.
18. Josh D. Cohen and Michael J. Fischer. A robust and verifiable cryptographically secure election scheme (extended abstract). In *FOCS*, 1985.
19. Ronald Cramer, Ivan Damgård, and Berry Schoenmakers. Proofs of partial knowledge and simplified design of witness hiding protocols. In *CRYPTO*, 1994.
20. Ronald Cramer, Rosario Gennaro, and Berry Schoenmakers. A secure and optimally efficient multi-authority election scheme. In *EUROCRYPT*, 1997.
21. Stéphanie Delaune, Steve Kremer, and Mark Ryan. Verifying privacy-type properties of electronic voting protocols. *Journal of Computer Security*, 17(4):435–487, 2009.
22. Carl M. Ellison. Ceremony design and analysis. *IACR Cryptology ePrint Archive*, 2007:399, 2007.
23. Atsushi Fujioka, Tatsuaki Okamoto, and Kazuo Ohta. A practical secret voting scheme for large scale elections. In *AUSCRYPT*, 1992.
24. Kristian Gjøsteen. Analysis of an internet voting protocol. *IACR Cryptology ePrint Archive*, 2010:380, 2010.

25. Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play any mental game or A completeness theorem for protocols with honest majority. In *STOC*, 1987.
26. Oded Goldreich and Yair Oren. Definitions and properties of zero-knowledge proof systems. *J. Cryptology*, 7(1):1–32, 1994.
27. Jens Groth. Evaluating security of voting schemes in the universal composability framework. In *ACNS'04*, pages 46–60, 2004.
28. Engelbert Hubbers, Bart Jacobs, and Wolter Pieters. RIES - internet voting in action. In *COMPSAC*, 2005.
29. Ari Juels, Dario Catalano, and Markus Jakobsson. Coercion-resistant electronic elections. *IACR Cryptology ePrint Archive*, 2002:165, 2002.
30. Ari Juels, Dario Catalano, and Markus Jakobsson. Coercion-resistant electronic elections. In *WPES*, 2005.
31. Aggelos Kiayias, Michael Korman, and David Walluck. An internet voting system supporting user privacy. In *ACSAC*, 2006.
32. Steve Kremer, Mark Ryan, and Ben Smyth. Election verifiability in electronic voting protocols. In *ESORICS*, 2010.
33. Ralf Küsters, Tomasz Truderung, and Andreas Vogt. Accountability: definition and relationship to verifiability. In *CCS*, 2010.
34. Ralf Küsters, Tomasz Truderung, and Andreas Vogt. Verifiability, privacy, and coercion-resistance: New insights from a case study. In *32nd IEEE Symposium on Security and Privacy, S&P 2011, 22-25 May 2011, Berkeley, California, USA*, pages 538–553. IEEE Computer Society, 2011.
35. Tal Moran and Moni Naor. Receipt-free universally-verifiable voting with everlasting privacy. In *CRYPTO*.
36. C. Andrew Neff. Practical high certainty intent verification for encrypted votes. Votehere, Inc. whitepaper, 2004.
37. Stefan Popoveniuc and Benjamin Hosp. An introduction to PunchScan. In *WOTE*, 2010.
38. Kazue Sako and Joe Kilian. Receipt-free mix-type voting scheme - A practical solution to the implementation of a voting booth. In *EUROCRYPT*, 1995.
39. Janos Simon, editor. *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*. ACM, 1988.
40. Warren D. Smith. Three voting protocols: Threeballot, vav, and twin. In *USENIX/ACCURATE*, 2007.
41. Ben Smyth, Steven Frink, and Michael R. Clarkson. Computational election verifiability: Definitions and an analysis of helios and JCJ. *IACR Cryptology ePrint Archive*, 2015:233, 2015.
42. Georgios Tsoukalas, Kostas Papadimitriou, Panos Louridas, and Panayiotis Tsanakas. From Helios to Zeus. In *EVT/WOTE*, 2013.
43. Thomas Zacharias. The DEMOS family of e-voting systems: End-to-end vevrifiable elections in the standard model. PhD thesis, National and Kapodistrian University of Athens, July 2016.
44. Filip Zagórski, Richard Carback, David Chaum, Jeremy Clark, Aleksander Essex, and Poorvi L. Vora. Remotegrity: Design and use of an end-to-end verifiable remote voting system. In *ACNS*, 2013.